

Historische Quellenkritik im Digitalen Zeitalter



Dissertation
zur Erlangung der Würde eines Doktors der Philosophie
vorgelegt der Philosophisch-Historischen Fakultät
der Universität Basel

von
Pascal Föhr
aus
Opfikon ZH

Basel 2018
Buchbinderei Bommer

Genehmigt von der Philosophisch-Historischen Fakultät der Universität Basel,
auf Antrag von

Prof. Dr. Martin Lengwiler (Universität Basel),
Prof. Dr. Monika Dommann (Universität Zürich),
Assoz. Prof. Dr. Eva Pfanzelter (Universität Innsbruck).

Basel, den 08.02.2017

Der Dekan Prof. Dr. Thomas Grob

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel
edoc.unibas.ch



Dieses Werk ist lizenziert unter einer
[Creative Commons Namensnennung 4.0 International Lizenz](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Kontakt Verfasser:

Pascal Föhr
pfhist@foehr.ch

Für
Benjamin und Christine

In Erinnerung an Franz[†]

Um die Lesbarkeit zu erleichtern, werden personenbezogene Formulierungen auf die männliche Form beschränkt. Weibliche Personen sind in dieser Formulierung immer mit eingeschlossen.

Coverbild: PANTOGRAPH, Japan¹

¹ URL: <http://www.pangra.net/objet/nichipaso/np103.html>; Urheber: Pangra Studio, 2-10-17 Kanagawa, Kanagawa-ku, Yokohama-shi 221-0045. Bildlizenz gemäss persönlicher Korrespondenz vom 10.07.2015.

Inhaltsverzeichnis

Danksagung.....	1
1. Einleitung.....	2
1.1. Forschungsstand.....	4
1.1.1. Digital Humanities.....	5
1.1.2. Digital History.....	8
1.1.3. Informatik.....	14
1.1.4. Medientheoretische Zugänge.....	18
1.1. Fragestellung, Forschungsfragen und -thesen.....	19
1.2. Aufbau.....	22
1.3. Methodische Herausforderung.....	22
2. 1 und 0.....	25
2.1. Eigenschaften eines Digitalen Objekts.....	29
2.1.1. Digitaldatentypen.....	29
2.1.2. Definition der Eigenschaften.....	31
2.1.3. Abgrenzung.....	39
2.2. Neue Quellentypen.....	42
2.2.1. Integrierender Quellentypus.....	42
2.2.2. Einobjektreduzierender Quellentypus.....	44
2.2.3. Quellentypus der unsichtbar codierten Information.....	44
2.2.4. Keine neuen Quellentypen.....	48
2.3. Zusammenfassung (Fragen und Thesen).....	48
3. Traditionelle Methode und digitale Objekte.....	50
3.1. Traditionelle Prozessschritte und digitale Objekte.....	51
3.1.1. Heuristik.....	52
3.1.2. Quellenkunde.....	53
3.1.3. Quellenkritik.....	56
3.1.4. Interpretation.....	60
3.1.5. Darstellung.....	62
3.2. Eigenschaften digitaler Objekte.....	63
3.3. Objekt, Forschungsressource, Quelle.....	66
3.4. Wissenschaftliches Arbeiten.....	68
3.5. Beurteilung.....	71
3.6. Zusammenfassung (Fragen und Thesen).....	73
4. Digitalmedien.....	75
4.1. Medien – Digitalmedien.....	76
4.2. Neue Medien und Social Media.....	78
4.3. Veränderte Mediennutzung.....	79
4.4. Umbruch in ein Digitales Zeitalter.....	83
4.5. Zusammenfassung.....	86
5. Quellenkritik des Digitalen.....	88
5.1. Heuristik.....	90
5.1.1. Suchstrategie.....	92
5.1.2. Suchräume.....	95
5.1.3. Suchtechniken.....	105
Systematische Suche.....	107
Unsystematische Suche / Suchmaschinen.....	109
5.1.4. Vollständigkeit.....	117
5.1.5. De- / Kontextualisierung.....	119
5.1.6. Recht.....	121
5.1.7. Zusammenfassung.....	131
5.2. Quellenkunde.....	132
5.2.1. Funktion.....	133
5.2.2. Quellengattung.....	134
Lösungsansatz Quellengattungsschema.....	134
5.2.3. Zusammenfassung.....	136

5.3.	Quellensicherung.....	136
5.3.1.	Digitale Langzeitarchivierung (Exkurs).....	139
5.3.2.	Konsistenz.....	141
5.3.3.	Sekundärobjekt und Surrogat.....	147
5.3.4.	In welchem Format – Konversion und Migration.....	150
5.3.5.	Informationssystem – Emulation.....	156
5.3.6.	Sicherungsort.....	158
	Lösungsansatz 'Researcher Driven Archiving'.....	164
	Lösungsansatz 'Trusted Saving Procedure'.....	166
	Lösungsansatz 'Personal Trusted Archive'.....	169
5.3.7.	Metadaten.....	170
	Lösungsansatz 'Metadata Information Section'.....	178
5.3.8.	Zusammenfassung.....	182
5.4.	Quellenkritik.....	184
5.4.1.	Historische versus informationstechnische Authentizität.....	186
5.4.2.	Integrität.....	189
	Lösungsansatz Digitale Signatur und MIS.....	198
5.4.3.	Persistenz.....	198
	Lösungsansatz Persistentes Identifikations- und Adressierungssystem im TDR.....	204
5.4.4.	Objektbeschreibung.....	204
5.4.5.	Echtheit.....	206
5.4.6.	Datierung.....	209
	Lösungsansätze Datierung.....	212
5.4.7.	Entstehungs- / Erscheinungsort.....	212
	Lösungsansatz Erscheinungsort.....	215
5.4.8.	Autorschaft.....	215
5.4.9.	Adressat.....	224
5.4.10.	Darstellung.....	227
5.4.11.	Inhalt.....	230
5.4.12.	Beziehungen zu Objekten.....	238
	Lösungsansatz Link-Check Element.....	240
5.4.13.	Authentizitätsapproximation.....	241
	Lösungsansatz Authentizitätsapproximation.....	242
5.4.14.	Zusammenfassung (Fragen und Thesen).....	251
5.5.	Interpretation.....	255
5.5.1.	Digitaldatentyp.....	258
5.5.2.	Datenform(at).....	259
5.5.3.	Informationssysteme.....	260
5.5.4.	Hypermedialität.....	262
5.5.5.	Zusammenfassung.....	262
5.6.	Darstellung und Publikation.....	263
5.6.1.	Neue Darstellungsformen.....	266
5.6.2.	Schreiben.....	271
5.6.3.	Belege.....	276
	Lösungsansatz Belegkette.....	278
5.6.4.	Visualisieren.....	280
5.6.5.	Open Access und Open Data.....	282
5.6.6.	Zusammenfassung.....	286
5.7.	Digital Humanities Cycle.....	287
5.8.	Zusammenfassung (Fragen und Thesen).....	289
6.	Digitalhistoriker.....	292
6.1.	Ausbildung.....	293
6.2.	Infrastruktur.....	299
6.3.	Zusammenfassung.....	300
7.	Schlussbetrachtungen.....	302
8.	Literaturverzeichnis.....	308
9.	Anhang.....	335
9.1.	Zusammenfassung Lösungsansätze.....	335

Abbildungsverzeichnis

Abbildung 1: Bild einer Schreibmaschine (Pantograph), in eine Webseite integriert (Föhr).....	43
Abbildung 2: Mouseover-Effekt unsichtbar (Quelle: SELFHTML).....	46
Abbildung 3: Mouseover-Effekt sichtbar (Quelle: SELFHTML).....	46
Abbildung 4: Schema 'Dreieck des wissenschaftlichen Belegs' (Föhr).....	70
Abbildung 5: Emoticon 'schnaubender Smiley'; Unicode U+F624 (Unicode Inc.).....	82
Abbildung 6: Historisch-kritischer Prozess mit Quellensicherung nach Pascal Föhr (Föhr).....	89
Abbildung 7: Google Scholar, Metadaten der Lizentiatsarbeit von Pascal Föhr (Google Scholar)	110
Abbildung 8: Webseite von GEO, mit im Browser aktiviertem Werbeblocker (URL: http://www.geo.de).....	118
Abbildung 9: Webseite von GEO, mit im Browser deaktiviertem Werbeblocker (URL: http://www.geo.de).....	119
Abbildung 10: NZZ Online-Artikel (1), Hilfssicherung als Bild. Text des gekürzten Artikels (NZZ Online).....	143
Abbildung 11: NZZ Online-Artikel (2), Hilfssicherung als Bild. Text des gekürzten Artikels mit aktuellen Zusatzinformationen im Webseitenelement 'Lesertrend – Gelesen' (NZZ Online)...	144
Abbildung 12: NZZ Online-Artikel (3), Hilfssicherung als Bild. Text des gekürzten Artikels nach zwei Tagen mit aktuellen Zusatzinformationen im Webseitenelement 'Lesertrend – Empfohlen' (NZZ Online).....	145
Abbildung 13: Webseite der Universität Basel, Standorte (Universität Basel).....	146
Abbildung 14: Sourcecode (Auszug) der Webseite der Universität Basel, Einbettung Google Maps markiert (Universität Basel).....	146
Abbildung 15: Schema Trusted Saving Procedure (Föhr).....	167
Abbildung 16: Metadaten zum Artikel 'Die Debatte über die Schweizer Geschichte wird fahrlässig' auf der Webseite der NZZ (NZZ Online).....	177
Abbildung 17: Metadaten zur Homepage der AB Glanzenburger (AB Glanzenburger).....	177
Abbildung 18: Metadaten zur Webseite des Workshops DH for European Global Studies (Universität Basel).....	177
Abbildung 19: Poster 'Historische Quellenkritik im Digitalen Zeitalter' mit Metadata Information Section (Föhr, 15.01.2015).....	180
Abbildung 20: Vergrößerung der MIS im Poster (Föhr).....	180
Abbildung 21: Ergänzende MIS auf der Webseite (Föhr).....	181
Abbildung 22: Signatur erstellen (Soft Xpansion).....	193
Abbildung 23: Signatur verifizieren (Soft Xpansion).....	194
Abbildung 24: Grenzen Kaschmirs, Ausschnitt Google Maps Schweiz (Google Maps).....	233
Abbildung 25: Grenzen Kaschmirs, Ausschnitt Google Maps Indien (Google Maps).....	234
Abbildung 26: Schema 'Linkhierarchie' (Föhr).....	237
Abbildung 27: Tweet von David M. Berry (Föhr).....	246
Abbildung 28: Tweet von David M. Berry, mit tweetfake erstellt (URL: http://tweetfake.com/).....	248
Abbildung 29: Aerial picture of Princes Street, Edinburgh (Europeana).....	248
Abbildung 30: Digital Humanities Cycle nach Verhoef und Wever (Darstellung: Föhr).....	287
Abbildung 31: Historisch-kritischer Prozess nach Föhr (Kap. 5.).....	336

Tabellenverzeichnis

Tabelle 1: Quellengattungsschema nach Ernst Bernheim (Darstellung: Föhr).....	54
Tabelle 2: 20 Rankingfaktoren von Google Search nach Brian Dean (Darstellung: Föhr).....	115
Tabelle 3: Quellengattungsschema nach Pascal Föhr; Aufzählungen nicht abschliessend.....	135
Tabelle 4: Archivtaugliche Dateiformate nach KOST und BAR (zusammenfassende Darstellung: Föhr).....	155
Tabelle 5: Beispiele von spezifischen Elementen für digitale Objekte zur Ergänzung eines Metadatenschemas (Föhr).....	176
Tabelle 6: Beurteilungsmerkmale Authentizitätsapproximation (Föhr).....	243
Tabelle 7: Bewertungstabelle (Föhr).....	244
Tabelle 8: Authentizitätsapproximation 'Turner et al.: White Paper' (Föhr).....	245
Tabelle 9: Authentizitätsapproximation 'Tweet von David M. Berry' (Föhr).....	247
Tabelle 10: Authentizitätsapproximation 'Europeana: Aerial picture of Princes Street, Edinburgh' (Föhr).....	249
Tabelle 11: Quellengattungsschema nach Föhr (siehe Kap. 5.2.2.).....	335

Abkürzungen

ALLEA	All European Academies
API	Application Programming Interfaces
AVI	Audio Video Interleave
BAR	Schweizerisches Bundesarchiv
CERN	Conseil Européen pour la Recherche Nucléaire
CHNM	Roy Rosenzweig Center for History and New Media
CMS	Content Management System
CSS	Cascading Style Sheets
DARIAH	Digital Research Infrastructure for the Arts and Humanities
DaSCH	Data and Service Center for the Humanities
DBMS	Datenbank Management System
DHC	Digital Humanities Cycle
DMS	Dokumentenmanagementsystem
DOI	Digital Object Identifier
DRM	Digital Rights Management
GIF	Graphics Interchange Format
GIS	Geo Information System
HTML	Hypertext Markup Language
ICT	Information and communication technology
IDC	International Data Corporation
IP	Internetprotokoll oder Intellectual Property
ISO	International Organization for Standardization
IT	Informationstechnik / Informatik
KOST	Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen
MIS	Metadata Information Section
NASA	National Aeronautics and Space Administration
OA	Open Access
OCR	Optical Character Recognition
OD	Open Data
PC	Personal Computer
PDF	Portable Document Format
PDF/A	Portable Document Format für Langzeitarchivierung (A)
PI	Persistenter Identifikator
PTA	Personal Trusted Archive
RDA	Researcher Driven Archiving
RMS	Records Management System
SAGW	Schweizerische Akademie der Geistes- und Sozialwissenschaften
SNF	Schweizerischer Nationalfonds
TDR	Trusted Digital Repository
TEI	Text Encoding Initiative
TIFF	Tagged Image File Format
tl;dr	too long; didn't read
TSP	Trusted Saving Procedure
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
UVC	Universal Virtual Computer
WIPO	World Intellectual Property Organization
Web	World Wide Web
XML	Extensible Markup Language

Länderkürzel nach ISO 3166 Alpha-2.

Danksagung

Die vorliegende Studie ist die gekürzte² und überarbeitete Fassung der Dissertation, die ich 2016 an der Philosophisch-historischen Fakultät der Universität Basel eingereicht und anfangs 2017 erfolgreich verteidigt habe. Dass die Arbeit zu einem guten Ende geführt werden konnte und nun in Buchform vorliegt, ist einer Vielzahl verschiedener Personen und Institutionen geschuldet, denen zu danken mir ein grosses Anliegen ist.

An erster Stelle möchte ich PD Dr. Peter Haber[†] danken, mit dem ich dieses Dissertationsprojekt habe beginnen, aber leider nicht beenden können. Ein grosser Dank geht an Prof. Dr. Martin Lengwiler von der Universität Basel, Prof. Dr. Monika Dommann von der Universität Zürich und vor allem an Assoz. Prof. Dr. Mag. Eva Pfanzelter von der Universität Innsbruck (Österreich), ohne die ich die Arbeit nicht zu Ende gebracht hätte.

Herzlich bedanken möchte ich mich auch bei den Kolleginnen und Kollegen, die mir mit Rat zur Seite gestanden sind und von denen ich Informationen aller Art erhalten habe: Prof. Dr. Raymond George Siemens (r/m/), Prof. Dr. Niels Brügger, Prof. Dr. Christer Clerwall, Prof. Dr. Dominik Schrage, Prof. Dr. Lukas Rosenthaler, Prof. Dr. Arnold F. Rusch, Dr. Mareike König, Dr. Jan Hodel, Dr. Simon Brülisauer, Dr. Catherina Schreiber und PhD cand. Jesper Verhoef. Zudem möchte ich all jenen Personen danken, die mir an verschiedenen Veranstaltungen mit wichtigen Hinweisen, Kommentaren und Fragen weitere Denkanstösse gegeben haben.

Ein grosser Dank an alle, die mich auf diesem Weg begleitet und unterstützt haben: meiner Frau Christine, meiner Familie und meinen Freunden, G!, der Familie Haber-Fäh und deren Freunde, dem Büro 5, Andrea Schüpbach, Stefan Frech, Andreas Affolter und Irina Burgermeister, den Mitarbeitenden des Staatsarchivs Solothurn, Alexander Abramowitsch Woskressenski (Александр Абрамович Воскресенский) und allen weiteren Personen, die ich aus unerfindlichen Gründen nicht namentlich erwähnt habe.

Ich habe versucht, einen meiner Leitsätze im Studium und Beruf umzusetzen, den mir Prof. Dr. em. Urs Bitterli in der ersten Veranstaltung im ersten Semester meines Studiums eingepflichtet hatte: *"Ich bin ja der Meinung, dass man einfache Sachen einfach ausdrücken sollte."*³ Ich hoffe, es ist mir auch in dieser Arbeit gelungen.

² Die Themenbereiche Medientheorie und Traditionelle historisch-kritische Methode wurden stark gekürzt. Ausführungen zur Informationssicherheit und Best Practice-Empfehlungen wurden entfernt.

³ Prof. Dr. em. Urs Bitterli, Proseminar '20. Juli 1944', SS 1997, Universität Zürich.

1. Einleitung

*"The historian who refuses to use a computer as being unnecessary, ignores vast areas of historical research and will not be taken serious anymore."*⁴
(Boonstra, Breure und Doorn, 1990)

Seit dem Aufschwung des Personal Computers in den 1980er-Jahren und dem Durchbruch des Internets Mitte der 1990er-Jahre hat die Zahl digital erstellter Dokumente stetig zugenommen. Digital publizierte Dokumente werden von Wissenschaftlern in der Forschung seit mehr als einem halben Jahrhundert als Quellen genutzt.⁵ Gleichzeitig wird diskutiert, welche dieser Quellen wie zitiert werden dürfen. Ein Aufsatz im Portable Document Format (PDF) auf der Webseite eines Universitätsprofessors wird mittlerweile fast fraglos als zuverlässige und zitierwürdige Quelle akzeptiert, hingegen ist die Verwendung eines Wikipedia-Artikels oft verpönt. Allerdings ist auch beim erstgenannten Aufsatzes ein quellenkritischer Blick nötig, denn digitale Objekte können unmittelbar verschwinden und oder nicht nachvollziehbar mutiert werden. Was jetzt auf der Webseite steht, kann innert Sekundenbruchteilen verändert worden sein.

Innerhalb des Diskurses zur historischen Quellenkritik⁶ sind vor allem drei Diskussionsstränge dominierend. *Erstens* wird seit Längerem die Bedeutung der Archive bei der Aufbewahrung digitaler Quellen diskutiert.⁷ Gesucht werden Lösungen für technische und praktische Probleme wie die Filterung von Informationen, den Wert der Daten für die Gesellschaft, das Urheberrecht, die Macht und Kontrolle der Suchmaschinen oder die Wiederverwendbarkeit von Forschungsdaten. *Zweitens* wird die Wissensgewinnung und -verbreitung thematisiert, wobei das Phantasma der Allwissenheit herumgeistert⁸, die durch Online-Enzyklopädien wie Wikipedia erreicht werden soll. Damit verbunden ist auch die Frage, ob all die

⁴ Boonstra, Onno; Breure, Leen; Doorn, Peter: Past, present and future of historical information science, article, Historical Social Research, 2004, S. 4, URL: <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-30846>, Stand: 06.05.2018.

⁵ Als *Quellen* werden alle Objekte oder Tatsachen bezeichnet, aus denen Kenntnisse der Vergangenheit gewonnen werden können (nach Kirn). Entscheidendes Kriterium für eine Quelle ist "ihr Erkenntniswert für die historische Forschung" (Arnold). Erst durch die Anwendung der Fragestellung auf ein Objekt wird dieses zur Quelle, die zum Erkenntnisgewinn beitragen kann. Siehe: Kirn, Paul: Einführung in die Geschichtswissenschaft, Berlin [-West] 1968 (Sammlung Göschen); Arnold, Klaus: Quellen, in: Jordan, Stefan (Hg.): Lexikon Geschichtswissenschaft: Hundert Grundbegriffe, Stuttgart 2007, S. 251-255.

⁶ *Kritik* ist die "Kunst des Trennens, des Schneidens" (Landwehr), die sowohl das *Aufteilen* in einzelne Elemente, als auch das *Beurteilen* dieser Elemente umfasst. Die Beurteilung erfolgt dabei aus der Sicht des 'Jetzt' mit Alternativen, die sich aus der zeitlichen Diskrepanz zum Untersuchungsobjekt ergeben. 'Kritische Geschichtsschreibung' erfolgt aus einer Selbstkritik der eigenen Wirklichkeitsproduktion, die auf ihrer eigenen Historizität beruht und sich deshalb auch nie allzu sicher sein sollte. Vgl. Landwehr, Achim: Die Kunst, sich nicht allzu sicher zu sein: Möglichkeiten kritischer Geschichtsschreibung, in: geschichte und kritik (61), 2013, S. 7-15, Seiten 7 und 14.

⁷ Siehe: Rosenthaler, Lukas: Archivierung im digitalen Zeitalter. Historische Entwicklung und Wege in eine digitale Zukunft, unveröffentlichte Habilitationsschrift, Universität Basel, Basel 2007, S. 508. Von Lukas Rosenthaler verdankenswerterweise zur Verfügung gestellt.

⁸ Siehe bspw.: Haber, Peter: digital past – Geschichtswissenschaft im digitalen Zeitalter, unveröffentlichte Habilitationsschrift, Universität Basel, Basel 2009, S. 155.

vorhandenen Informationen überhaupt noch verarbeitet werden können.⁹ Schliesslich ist *drittens* etwa 2010 eine leidenschaftliche Diskussion um die Auswirkungen der digitalen Medien auf die Geschichtswissenschaft entstanden:¹⁰ Didaktik, Fachkommunikation¹¹, akademisches Curriculum oder wissenschaftliches Publizieren sowie die Auseinandersetzung zwischen 'close' und 'distant reading'¹² stehen dabei im Fokus. Hingegen wird eine Diskussion über quellenkritische Methoden für die Geschichtswissenschaft im digitalen Zeitalter nur am Rande geführt, was mitunter ein Grund ist, weshalb noch kein konkreter und konsolidierter Vorschlag eines Kanons zur Quellenkritik des Digitalen präsentiert werden konnte.

Da in der Geschichtswissenschaft und in geschichtswissenschaftsnahen Fachgebieten¹³ 'digitale Objekte'¹⁴ als Quellen rege genutzt beziehungsweise von Forschenden erstellt und ausgetauscht werden, untersuche ich in dieser Arbeit die zentrale **Fragestellung, wie sich die historische Quellenkritik durch die Verwendung von digitalen Objekten als Forschungsressource sowie von digitalen Informations- und Kommunikationsmedien verändert.**¹⁵ In der Geschichtswissenschaft ist noch nicht klar, wie wir mit dieser Art von Quelle wissenschaftlich umzugehen haben, denn die bisherigen Methoden und Techniken reichen dafür nicht mehr aus.

⁹ Dazu bspw.: Weinberger, David: *Too Big To Know; Rethinking Knowledge. Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room*, New York 2011.

¹⁰ Bspw.: Tagung 'Rezensionieren – Kommentieren – Bloggen: Wie kommunizieren Geisteswissenschaftler in der digitalen Zukunft?', München 31. Januar / 1. Februar 2013, URL: <http://rkb.hypotheses.org/>; Workshop 'Bloggen in Geschichtswissenschaft und Archivwesen', Wien, 10. November 2014, URL: <http://bioeg.hypotheses.org/202>.

¹¹ Siehe weiterführend bspw.: Prinz, Claudia: *Kommunikation im digitalen Raum*, in: Busse, Laura; Enderle, Wilfried; Hohls, Rüdiger u. a. (Hg.): *Clio Guide. Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften*, Berlin 2016 (Historisches Forum 19). URL: <http://guides.clio-online.de/guides/arbeitsformen-und-techniken/kommunikation-im-digitalen-raum/2016>, Stand: 10.07.2016.

¹² Close reading: eingehendes Befassen mit einem Text; distant reading: computertechnische Auswertung eines Textes. Weiterführend siehe Kap. 4.3 *Veränderte Mediennutzung*.

¹³ Bspw. Kunstgeschichte resp. Digital Art History. Siehe zum Begriff Digital Art History bspw. Klinker, Harald; Surkemper, Liska (Hg.): *What is Digital Art History?*, München 2015 (International Journal for Digital Art History). URL: <http://www.dah-journal.org/>, Stand: 27.02.2016. Diskussionen siehe bspw.: Missomelius, Petra: *Digitale Medienkultur Wahrnehmung, Konfiguration, Transformation*, Bielefeld 2006 (Kultur- und Medientheorie); Kohle, Hubertus: *Digitale Bildwissenschaft*, Glückstadt 2013 (E-Humanities); Pratschke, Margarete: *Die Kunstgeschichte*, SAGW-Tagung 13, Bern 28.11.2013; Baumgärtel, Tilman: *Per Tweet ins Museum: Internetkunst*, in: *Neue Zürcher Zeitung*, 06.09.2013. URL: http://www.nzz.ch/aktuell/feuilleton/kunst_architektur/per-tweet-ins-museum-1.18145506?extcid=Newsletter_06092013_Top-News_am_Morgen, Stand: 06.05.2018. Resultat der internationalen Tagung "Digitale Kunstgeschichte: Herausforderungen und Perspektiven" vom 26. und 27. Juni 2014 in Zürich. URL: <https://www.sik-isea.ch/de-ch/Aktuell/Aktuell/Veranstaltungen/Digitale-Kunstgeschichte>.

¹⁴ Unter einem digitalen Objekt wird die durch ein digitales Informationssystem ('Computer') dargestellte Form einer Ziffernreihe im Binärformat (üblicherweise 0 und 1) verstanden. Vertiefend im Kap. 2. 1 und 0.

¹⁵ Ausführlich in Kap. 1.1. *Fragestellung, Forschungsfragen und -thesen*.

Die vorliegende Arbeit richtet sich an Studenten und Forschende der Geschichtswissenschaft sowie an alle Wissenschaftler und Praktiker in geschichtswissenschaftsnahen Fachgebieten. Die Lösungsansätze sollen Hilfestellungen bieten, wie Digitale Objekte quellenkritisch zu prüfen sind und welche Infrastrukturen die Wissenschaft zur Verfügung stellen sollte, damit die Digitalen Objekte langfristig verwendet werden können.

Die digitalen Geisteswissenschaften stehen vor der Herausforderung, dass neue Begriffe aus der Informatik¹⁶ Einzug ins Fachvokabular hielten. Wo nötig, werden informationstechnische Begriffe im Text oder in einer Fussnote erklärt oder auf Fachliteratur verwiesen.

1.1. Forschungsstand

Die Geistes- und Sozialwissenschaften stehen vor zwei Herausforderungen: *Erstens* haben viele Forschende keine informationstechnischen Grundkenntnisse. Ihnen fehlt das Bewusstsein für den Umgang mit digitalen Objekten, weshalb sie informationstechnische Aspekte oft nicht untersuchen. Die Informationstechnik wird meistens nur als technisches Hilfsmittel betrachtet, beispielsweise als 'Schreibmaschine mit Internetzugang'. *Zweitens* sollte eine Diskussion um die Verwendung von digitalen Quellen im dafür prädestinierten Fachbereich *Digital History*, einer Teildisziplin der *Digital Humanities* geführt werden. Allerdings gibt es keine klare Abgrenzung der Bereiche Digital History und Digital Humanities, im Gegenteil: Besonders in Europa ist wieder ein einseitiger Trend zur corpuslinguistischen Ausrichtung der Digital Humanities zu erkennen, dessen Fokus eben nicht die Spezifika der Digital History sind. Dies erschwert eine vertiefte Auseinandersetzung mit dem hier untersuchten Kernbereich der Historiographie, der *Quellenkritik des Digitalen*.

In den Digital Humanities und der Digital History sind zwei Leitdiskurse feststellbar: *Erstens* wird versucht, die beiden Fachgebiete zu definieren und von anderen abzugrenzen, wobei bisher kein Konsens für eine klare Definition gefunden wurde. *Zweitens* wird die Einführung von Methoden und Techniken aus der Informatik in die Fachgebiete thematisiert. Anwendungen wie Text- oder Netzwerkanalysen werden bereits erfolgreich eingesetzt, jedoch wird über die Verwendung von grossen Datenbeständen, dem sogenannten 'Big Data', noch intensiv diskutiert. Die Quellenkritik des Digitalen wird aber nur am Rande und bruchstückhaft behandelt.

¹⁶ Vertiefende Erklärungen finden sich z.B. bei: Fischer, Peter; Hofer, Peter: Lexikon der Informatik, Berlin Heidelberg 2011, S. 430 f.; Rechenberg, Peter; Pomberger, Gustav (Hg.): Informatik-Handbuch, München 1999.

1.1.1. Digital Humanities

Die nähere Auseinandersetzung mit den sogenannten *Digital Humanities*¹⁷ zeigt schnell, dass keine klare Definition des Begriffs möglich ist. Mit den steigenden Möglichkeiten der computergestützten Verarbeitung von Daten haben sich in den Geisteswissenschaften kontinuierlich neue Forschungsfelder und -themen eröffnet und eine stärkere Interdisziplinarität gefordert. Dies führte zur Vielheit der Definitionsversuche. Die hier verwendete Arbeitsdefinition grenzt deshalb den Begriff nur grob ein und bezieht sich auf die interdisziplinäre Erforschung der Möglichkeiten und Auswirkungen, die sich mit der Verwendung von Informationstechnologie in den Geistes- und Sozialwissenschaften ergeben: *Digital Humanities sind in erkenntnistheoretischem Sinn die Bearbeitung von geistes- oder sozialwissenschaftlichen Fragestellungen anhand von digitalen oder digitalisierten Quellenbeständen, die sinnvollerweise mit interdisziplinären, digitalen Analyse- oder Visualisierungsmitteln durchgeführt werden (müssen).*

Einen noch immer guten Einstieg ins Thema bieten Susan Schreibman, Raymond Siemens und John Unsworth legten mit ihrer Publikation *A Companion to the Digital Humanities* 2004 einen wichtigen Grundstein für das Wissenschaftsteilgebiet, dem sie damit auch den Namen *Digital Humanities* gaben.¹⁸ Veränderungen, die das *Web 2.0* oder *Social Web* in der Informationsgewinnung und -teilung mit Blogs oder Twitter mit sich brachten, wurden von Schreibman noch nicht aufgegriffen. Bereits von diesen Autoren wird der Einsatz der Informatik in den Geisteswissenschaften gefordert, was allerdings weitgehend noch der Umsetzung harret. Weitere Einführungen bietet unter anderem der ALLEA-Report *Going Digital*, in dem die Verfasser auf Chancen und Möglichkeiten mit den neuen Forschungsansätzen, die Notwendigkeit der Standardisierung und auf weitere unterschiedliche Problemfelder hinweisen.¹⁹ Die Autoren in *Debates in the digital humanities*²⁰ des Anglisten Matthew Gold und *Understanding Digital Humanities*²¹ des Medien- und Kommunikationswissenschaftlers David Berry zeigen auf, was alles unter dem Begriff zusammengefasst werden kann. Auch in diesen Publikationen wird allerdings einmal mehr deutlich, dass sich Digital Humanities definitiv wegen ihrer Breite nicht wirklich fassen lassen: Es wird immer noch diskutiert, was

¹⁷ Namensgebend waren: Schreibman, Susan; Siemens, Raymond George; Unsworth, John: *A companion to digital humanities*, Malden 2004 (Blackwell companions to literature and culture 26).

¹⁸ Ebd. Die Publikation nimmt einen weitgehend anglo-amerikanischen Blickwinkel ein. Je nach Definition und Autor werden v.a. im angelsächsischen Wissenschaftsraum auch naturwissenschaftliche Fachrichtungen zu den Digital Humanities gezählt. Eine klare Abgrenzung ist kaum möglich.

¹⁹ Collins, Sandra; Harrower, Natalie; Haug, Dag Trygve Truslew u. a.: *Going Digital: Creating Change in the Humanities*, ALLEA E-Humanities Working Group Report, ALLEA – All European Academies, Berlin 07.05.2015. URL: http://www.allea.org/wp-content/uploads/2015/07/Going-Digital_digital-version.pdf, Stand: 06.05.2018.

²⁰ Gold, Matthew K. (Hg.): *Debates in the digital humanities*, Minneapolis 2012. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.

²¹ Berry, David M.: *Understanding digital humanities*, Houndmills, Basingstoke, Hampshire - New York 2012.

darunter zu verstehen sei, wer *in* und wer *out* sei und ob die Digital Humanities sich um die Praxis oder das Theoretisieren, um Informatik oder Kommunikation, um das Fach oder die Politik kümmern sollen.²²

Die Webseite *whatisdigitalhumanities*²³ zeigt mit jedem Wiederaufbau (refresh) der Seite eine neue Definition an, was eindrücklich das Fehlen einer konzisen Definition von Digital Humanities widerspiegelt. Auch der Blogbeitrag *A Working Definition of Digital Humanities* des Informationswissenschaftlers Scott Weingart zeigt dies deutlich:

"Hah! I tricked you. I don't intend to define digital humanities here – too much blood has already been spilled over that subject."²⁴

Es bleibt festzuhalten: Eine konzise Definition von Digital Humanities gibt es bis dato nicht, denn es sind zu viele Wissenschaftszweige mit ihren eigenen Ansprüchen und Ansichten in diesem Diskurs involviert, wie Raymond Siemens an der Digital Humanities Summer School of Switzerland im Sommer 2013 erklärte.²⁵

Aufgrund dieser Undifferenziertheit oder dem Unwillen, klare Grenzen zu ziehen, muss sich die Digital Humanities Community den Vorwurf gefallen lassen, sie demonstriere damit selbstgefällig ihre Neuartigkeit und Integrationsfähigkeit für alle möglichen Wissenschaftskulturen.²⁶ Zudem ist zu hinterfragen, ob Digital Humanities als eigenes Forschungsfeld überhaupt gerechtfertigt sind, falls das Attribut 'digital' nur als Verteidigung und Erneuerung der Geisteswissenschaft in Zeiten rückläufiger Forschungsgelder dient, aber keine neuen Erkenntnisse liefert.²⁷ Der Medienwissenschaftler Theo Röhle fragt danach, inwiefern den

²² Vgl. Spiro, Lisa: «This Is Why We Fight»: Defining the Values of the Digital Humanities, in: Gold, Matthew K. (Hg.): *Debates in the digital humanities*, Minneapolis 2012, S. 16. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.

²³ Heppler, Jason: URL: <http://whatisdigitalhumanities.com/>.

²⁴ Weingart, Scott: *A Working Definition of Digital Humanities*, the scottbot irregular, 08.12.2013, URL: <http://www.scottbot.net/HIAL/?p=39749>, Stand: 25.02.2014.

²⁵ Siemens, Raymond George: *Understanding and Responding to (Inter)Disciplinary Change*, unpubl. Vortrag, DH Summer School Switzerland, University of Berne 26.03.2013. Zu weiteren Definitionsversuchen siehe bspw. auch Beiträge in: Gold, Matthew K. (Hg.): *Debates in the digital humanities*, Minneapolis 2012 (URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>); Alvarado, Rafael C.: *The Digital Humanities Situation*, S. 50; Gold, Matthew K.: *Day of DH: Defining the Digital Humanities – Blog Posts*, S. 67-71; Fitzpatrick, Kathleen: *The Humanities, Done Digitally*, S. 12-15; Svensson, Patrik: *Beyond the Big Tent*, S. 36-49. Weiter: Terras, Melissa; Nyhan, Julianne; Vanhoutte, Edward (Hg.): *Defining digital humanities: a reader*, Farnham 2013; Schmale, Wolfgang: *Digital Humanities: Praktiken der Digitalisierung, der Dissemination und der Selbstreflexivität*, Stuttgart 2015 (Geschichte).

²⁶ Siehe dazu bspw.: Hermann, Felix: Rezension zu: Gold, Matthew K. (Hrsg.): *Debates in the Digital Humanities*, Minneapolis 2012 / Berry, David M. (Hrsg.): *Understanding Digital Humanities*, New York 2012, Buchrezensionen, H-Soz-Kult, 09.01.2015, URL: <http://www.hsozkult.de/publicationreview/id/rezbuecher-21673>, Stand: 09.01.2015.

²⁷ Vgl.: Hall, Gary: *There Are No Digital Humanities*, in: Gold, Matthew K. (Hg.): *Debates in the digital humanities*, Minneapolis 2012, S. 133-138. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>. Hall argumentiert zudem, dass das 'Digitale' den Humanities nicht zugefügt werden kann, weil die Geisteswissenschaften schon vor der Schaffung dieser Begriffskombination ein Verständnis von Computern und dem Digitalen hatten und auch mit

Digital Humanities eine strategische Rolle zukomme, was die Vermarktung von Daten durch grosse Datensammler wie Google oder Facebook angehe.²⁸ Es wäre demnach auch zu untersuchen, ob von diesen Unternehmen mit der Forcierung von Trends wie *Big Data* oder *Social Media* geisteswissenschaftliche Kompetenzen²⁹ eingekauft werden, die dann in der wissenschaftlichen Forschung fehlen würden.³⁰

Zu Big Data und Social Media erscheinen zurzeit fast zahllose Studien, Artikel und Bücher, eine Übersicht und aktuelle Darstellung ist nur cursorisch möglich. Eine Einführung zu Big Data bietet Ramón Reicherts gleichnamiges Werk³¹, in dem verschiedene Autoren auf Technik, Wissen, Theorien und Methoden eingehen und das Thema in Bezug zu weiteren, mit den Digital Humanities konfrontierten Wissenschaftsgebieten wie der Informatik oder Medienwissenschaft setzen.

In den meisten Definitionsversuchen über Digital Humanities werden zwei gemeinsame Punkte genannt. Es geht erstens um die Nutzung von digitalen Informationssystemen und zweitens um den breiten Themenbereich der Humanities, wobei dieser wiederum nicht scharf abgegrenzt wird. Die von der Schweizerischen Akademie der Geistes- und Sozialwissenschaften (SAGW) formulierte Definition nimmt diese Punkte auf und legt zudem den Fokus auf einen *epistemologischen* Ansatz:

diesen Techniken arbeiteten. Das ist korrekt, allerdings vernachlässigt Hall das Verhältnis der Durchdringung der Gesellschaft durch die Informationstechnologie, die erst in den 1980er-Jahren begann und rapide zunahm, was neue Forschungsfelder eröffnete und die Schaffung der Begriffskombination rechtfertigt.

²⁸ Siehe: Röhle, Theo: Big Data – Big Humanities? Eine historische Perspektive, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 169.

²⁹ Kompetenz: Vermögen, Fähigkeit. Für Dieter Baacke ist der Begriff schwerwiegend, weil er sowohl eine anthropologische Voraussetzung, als auch einen Zielwert formuliert. Die Voraussetzung liegt in der Annahme begründet, dass alle Menschen kompetente Lebewesen sind und deren Kompetenz gefördert werden müsse; der Zielwert besteht in der Förderung dieser Annahme. Nach: Baacke, Dieter: Medienpädagogik, Tübingen 1997 (Grundlagen der Medienkommunikation), S. 96.

³⁰ Aktuelle Diskussionen zum Thema Digital Humanities siehe bspw.: Warwick, Claire; Terras, Melissa; Nyhan, Julianne (Hg.): Digital humanities in practice, London 2012; Rogers, Richard: Digital methods, Cambridge (USA) 2013; Kelly, T. Mills: Teaching history in the digital age, Ann Arbor 2013 (Digital humanities). Des Weiteren das Forum: The Status Quo of Digital Humanities in Europe, (URL: <http://www.hsozkult.de/hfn/text/id/texte-2535>) und die Webseiten und Newsreader: Digitale Geschichtswissenschaft (URL: <http://digigw.hypotheses.org/feed>); G&I (URL: <http://blog.ahc-ch.ch/feed/>); H-Soz-u-Kult (URL: <http://hsozkult.geschichte.hu-berlin.de/rss.xml>); infoclio (URL: <https://www.infoclio.ch/de/news/rss.xml>); Digital Humanities Now Editors Choice (URL: <http://feeds.feedburner.com/DHNowEditorsChoiceAndNews?format=xml>); edwired (URL: <http://feeds.feedburner.com/Edwired?format=xml>); heise online (URL: <http://www.heise.de/newsticker/heise-atom.xml>); Hypotheses URL: <http://hypotheses.org/>; Alliance of Digital Humanities Organizations (URL: <http://adho.org/>); Humanist Discussion Group (URL: <http://dhumanist.org/>).

³¹ Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014. Die Beiträge der Autoren wurden teilweise ins Deutsche übersetzt.

"Von eigentlichen 'e-humanities' oder 'digital humanities' in einem erkenntnistheoretischen Sinn ist dann die Rede, wenn bestimmte Fragestellungen an Quellenbestände herangetragen werden, die sinnvollerweise nur aufgrund ihrer digitalen Verfügbarkeit bearbeitet werden können oder deren Untersuchung digitale Analyse- oder Visualisierungsmittel erfordern, wie es etwa bei der computergestützten Linguistik, der datenbankbasierten quantitativen Geschichtsforschung, der Archäometrie oder jüngst der Neuroästhetik der Fall ist."³²

Diese Definition liegt der eingangs des Kapitels erwähnten Arbeitsdefinition zugrunde, obwohl auch hier das Forschungsgebiet nicht klar eingegrenzt wird. Die beispielhafte Nennung von Forschungsgebieten ist wenig zielführend und wurde in der Arbeitsdefinition auf das Fachgebiet der Geistes- und Sozialwissenschaften eingegrenzt.

1.1.2. Digital History

Nach den problematischen Definitionsversuchen in den Digital Humanities ist es nicht verwunderlich, dass auch die Definition von Digital History nicht klar ist. Zwar versteht sie sich als eine Art *geschichtswissenschaftliches* Fachgebiet der Digital Humanities, aber was sie ist oder was sie nicht ist, welche wissenschaftlichen Fachrichtungen und Personen dazu gehören und wie sie sich selber definiert, ist unklar. Die in dieser Arbeit verwendete Arbeitsdefinition grenzt den Begriff wie bei den Digital Humanities nur grob ein. Digital History bezieht sich auf die interdisziplinäre Erforschung der Möglichkeiten und Auswirkungen, die sich mit der Verwendung von Informationstechnologie in den Geistes- und Sozialwissenschaften ergeben: *Digital History ist in erkenntnistheoretischem Sinn die Bearbeitung von geschichtswissenschaftlichen Fragestellungen an digitale oder digitalisierte Quellenbestände, die sinnvollerweise mit interdisziplinären, digitalen Analyse- oder Visualisierungsmitteln durchgeführt werden (müssen).*

Erstaunlich ist jedoch, dass sowohl in der 'traditionellen' Geschichtswissenschaft, als auch in der Digital History eine intensive Auseinandersetzung mit der Quellenkritik des Digitalen bisher weitgehend ausgeblieben ist. Aufgrund des hohen Stellenwertes, der die historisch-kritische Methode in der Geschichtswissenschaft inne hält, wären Lösungsvorschläge zur Frage, wie die historische Quellenkritik auf digitale Objekte als Forschungsressource anzuwenden ist, zu erwarten gewesen. In den beigezogenen physischen und digitalen Publikationen und Beiträgen sind die für die Fragestellung interessanten Methoden und Techniken zur Quellenkritik des Digitalen nur bruchstückhaft untersucht worden.

Digital History als Forschungsfeld haben die Historiker Daniel Cohen und Roy Rosenzweig (erst) 2006 mit der Publikation *Digital History* aufgearbeitet, vor allem um den lehrenden

³² Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Digital Humanities und Web 2.0, Bern 2012 (SAGW Bulletin), S. 29. URL: http://www.sagw.ch/dms/sagw/bulletins_sagw/bulletins_2012/SAGWbulletin1-12.

Historikern das World Wide Web näher zu bringen.³³ Die Anerkennung von Digital History als wissenschaftliches Fachgebiet ist jedoch nicht unbestritten. Beispielsweise bezeichnete der Bildungswissenschaftler und Historiker Thomas Nygren an einer Veranstaltung Digital History als Bastard – "it is not true history" – denn es handle sich bloss um die Aneignung erlernbarer Techniken.³⁴ Die Erklärung dafür, was der Unterschied zum Erlernen des traditionellen Historikerhandwerks ist, blieb er jedoch schuldig.

Weitere Autoren beleuchten weitere, teils andere Aspekte von Digital History, ohne den Begriff in seiner Gesamtheit zu definieren. Beispielsweise versucht der Mediävist Valentin Groebner in *Wissenschaftssprache digital*, mit verschiedenen Essays zu beschreiben, wie Digital History betrieben werden soll.³⁵ Er gibt darin zwar nützliche Hinweise, allerdings beschränken sich seine Überlegungen auf das Thema wie *digital immigrants* und *digital natives* mit dem Informationsüberfluss umgehen und selbst digital publizieren sollen.³⁶ Vor allem Letztere hätten weniger Ahnung vom Funktionieren des Internets und von Suchmaschinen als sie glauben.³⁷ Groebner gibt aber weder eine konkrete Anleitung noch definiert er Digital History.

Wolfgang Schmale geht in *Digitale Geschichtswissenschaft*³⁸ unter anderem der Frage nach, was digitale Geschichtswissenschaft ist. Er definiert sechs Kriterien: die individuelle Historical Workstation, die Geschichtswissenschaft im Netz, das Management historischen Wissens unter Nutzung semantischer Methoden, das E-Learning, das Forschen unter Zuhilfenahme digitaler Arbeitstechniken und das digitale Erzählen, womit die digitale Geschichtswissenschaft per definitionem trans- und interdisziplinär und multimedial ist.³⁹ Für Schmale erfordert die digitale Geschichtswissenschaft, "die Ästhetik des traditionellen Systems der Geschichtswissenschaften mit den Vorteilen des digitalen Mediums Hypertext zu verbinden."⁴⁰ Weiter führt er aus, dass dieses Unterfangen "[...] die für eine Wissenschaft typischen Qualitätskontrollstufen, [...] das Schreiben von Artikeln nach den wissenschaftlichen Regeln und Methoden, [...] eine hochgradig komplexe Kombination aus Software-

³³ Cohen, Daniel J.; Rosenzweig, Roy: *Digital history: A Guide to Gathering, Preserving, and Presenting the Past on the Web*, Philadelphia 2006.

³⁴ Schriftliche Notiz von Pascal Föhr: Nygren, Thomas; Vikström, Lotta: *A Bastard in the History Classroom: Benefits and Obstacles in Using Digital Databases in History Teaching*, dhlu-Panel: *Challenges and potentialities of online archives*, Luxembourg 20.03.2012.

³⁵ Groebner, Valentin: *Wissenschaftssprache digital. Die Zukunft von gestern*, Paderborn 2014.

³⁶ *Digital Immigrants*: Personen, die sich den Umgang mit Digitaltechnik aneigneten. *Digital Natives*: Personen, die von klein auf mit den Digitaltechniken vertraut sind. Zudem 'Google-Generation': *Digital Natives*, die von klein auf mit dem Internet und WWW vertraut sind (ab ca. Jahrgang 1993). Diese Begriffe sind zwar pauschalisierend, bringen jedoch die unterschiedliche Nutzung von und Lebensweise mit Digitalmedien zur Geltung, weshalb sie im Folgenden verwendet werden.

³⁷ Zu den ersten (analogen) Suchmaschinen siehe auch: Tantner, Anton: *Die ersten Suchmaschinen: Adressbüros, Fragämter, Intelligenz-Comptoirs*, Berlin 2015.

³⁸ Schmale, Wolfgang: *Digitale Geschichtswissenschaft*, Wien, Köln 2010.

³⁹ Ebd., S. 61.

⁴⁰ Ebd., S. 72.

Anwendungen, [...] eine Vielzahl von Mitspielern, nicht nur WissenschaftlerInnen, sondern auch TechnikerInnen, Informatikdiensten, MediendesignerInnen, Institutionen, Geldgebern [umfasst].⁴¹ Schmale bezieht sich zwar nur auf im Internet verfügbare Objekte und erwähnt andere digitale Quellen nicht, aber seine Kriterien hätten auch für eine konkrete, aber leider nicht erfolgte Definition von Digital History verwendet werden können. Zudem wäre ein vertiefender Blick auf die Qualitätskontrollstufen der Quellenkritik wünschenswert gewesen.

In *Geschichte digital*⁴² beschreibt der Historiker Guido Koller den Stand der Forschung im Bereich der Digital History. Koller weist in seinen kurz gehaltenen Kapiteln auf wichtige Probleme und Lösungsansätze hin, die in dieser Arbeit vertiefend dargestellt werden. Er streicht heraus, dass sich die Bedeutung der Quellenkritik sich mit dem Einbezug von digital produzierten Unterlagen ändere, ohne aber eine neue Methodik vorzustellen oder anzustossen. Jedoch versucht er die Eigenschaften von Digital History zu beschreiben, indem er sie als Disziplin versteht, die "Entwicklungen in temporären virtuellen Räumen und an spezifischen Orten diachron und synchron zu beschreiben, analysieren und zu verstehen versucht."⁴³ Damit erweitert er Schmales Konzept, das sich auf das hypertextuelle Darstellen von Geschichten im Netz' beschränkt.

Einige Werke beschäftigen sich mit dem Einsatz von konkreten elektronischen Hilfsmitteln in der Geschichtswissenschaft, wie Klaus van Eickels' (Mediävist), Ruth Weichselbaumers (Mediävistin) und Ingrid Bennewitz' (Philologin) *Mediaevistik und Neue Medien*.⁴⁴ Die Arbeit fasst Beiträge einer Veranstaltung zum Thema zusammen, bleibt aber in der Gesamtheit eine Darstellung von Projekten und Möglichkeiten, wie Informationstechnik in der Mediävistik eingesetzt werden kann und wie der Zugang zu digitalen Inhalten standardisiert werden soll. Eine methodologische Auseinandersetzung mit der digitalen Quelle an sich findet nur in Ansätzen statt, aber die Beiträge geben einen guten Einblick in die Möglichkeiten, wie digitale Geschichte betrieben werden kann. Andere Werke sind aus archivischer Sicht verfasst und sind ausschliesslich als Hilfsmittel für das Auffinden von digitalen Quellen gedacht und bieten keine oder nur beschränkte methodologischen Diskussionen. Als Beispiel dient *Elektronische Informationsressourcen für Historiker*⁴⁵ des Bibliothekaren Klaus Gantert oder die *online-Guides*⁴⁶ von clio-online, die online publiziert sind und bald auch gedruckt erscheinen werden. Solche Arbeiten sind unter dem Blickwinkel der Verbesserung der digitalen Informati-

⁴¹ Ebd.

⁴² Koller, Guido: *Geschichte digital: Historische Welten neu vermessen*, Stuttgart 2016.

⁴³ Ebd., S. 39.

⁴⁴ Eickels, Klaus van; Weichselbaumer, Ruth; Bennewitz, Ingrid (Hg.): *Mediaevistik und Neue Medien*, Ostfildern 2004.

⁴⁵ Gantert, Klaus: *Elektronische Informationsressourcen für Historiker*, Berlin 2011 (Bibliotheks- und Informationspraxis 43).

⁴⁶ Siehe URL: <http://guides.clio-online.de/>. Busse, Laura; Enderle, Wilfried; Hohls, Rüdiger u. a. (Hg.): *Clio Guide. Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften*, Berlin 2016 (Historisches Forum 19).

onskompetenz und der Einarbeitung in die Besonderheiten der jeweiligen Ressource dennoch sehr hilfreich.

Hauptsächlich mit den Arbeitstechniken der Digital History befassen sich beispielsweise der Historiker Martin Gasteiner und Peter Haber in *Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften*.⁴⁷ Haber und Gasteiner zeigen in diesem Sammelband Arbeitstechniken für den einzelnen Forschenden auf. Die Beiträge sind zwar hilfreich, stellen aber nur eine Momentaufnahme dar, denn die schnell fortschreitende Entwicklung lässt die Techniken rasch veralten. In diesem Band findet sich zudem der bisher einzig auf Deutsch publizierte, sich explizit auf historische Quellenkritik beziehende Beitrag *Von der Quellenkritik zum kritischen Umgang mit digitalen Ressourcen* der Historikerin Eva Pfanzelter.⁴⁸ Auf wenigen Seiten legt sie die grundsätzlichen Probleme dar und bietet pragmatische Lösungsansätze. Sie beschäftigt sich wohl aus Platzgründen nicht eingehend mit methodischer Quellenkritik, weist aber darauf hin, "dass es *die* methodisch einwandfreie und generell anwendbare Methode dafür nicht gibt."⁴⁹ Dies trifft besonders auf digitale, aber auch auf physische Quellen zu.

In der Einführungsliteratur in Werken zur Einführung in das Studium der Geschichtswissenschaft wird nur bruchstückhaft auf die Quellenkritik von digitalen Daten eingegangen wie beispielsweise in *Werkstatt des Historikers*⁵⁰ der Historiker Martha Howell und Walter Prevenier, *Geschichte Online*⁵¹ des Wirtschafts- und Sozialhistorikers Franz Eder oder im Kapitel über E-History in *Praxisbuch Geschichte*⁵² des Historikers Martin Lengwiler. Der Geisteswissenschaftler Richard Rogers bietet in *Digital Methods*⁵³ einen guten methodologischen Einstieg für Forschungen im und über das Internet, eine vertiefte Auseinandersetzung mit den Nutzungsmöglichkeiten und Funktionalitäten von Suchmaschinen sowie einen Ausblick auf Big und Small Data. Rogers beschränkt sich aber auf diese Themen und führt keine explizit quellenkritische Diskussion. Dies würde eher *Doing History: Research and Writing in the Digital Age*⁵⁴ der Historiker Michael Galgano, Chris Arndt und Raymond Hyser versprechen, das als Einführung für Studenten und Interessierte konzipiert ist. Leider wird die digitale Quellenkritik nur am Rande thematisiert, womit ein wichtiger Aspekt von 'Doing'

⁴⁷ Gasteiner, Martin; Haber, Peter (Hg.): *Digitale Arbeitstechniken: für die Geistes- und Kulturwissenschaften*, Wien, Köln, Weimar 2010.

⁴⁸ Pfanzelter, Eva: *Von der Quellenkritik zum kritischen Umgang mit digitalen Ressourcen*, in: *Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften*, 2010, S. 39-49.

⁴⁹ Ebd., S. 42 f. Kursiv im Original.

⁵⁰ Howell, Martha; Prevenier, Walter: *Werkstatt des Historikers: Eine Einführung in die historischen Methoden*, Köln 2004.

⁵¹ Zur Quellenkritik in: Eder, Franz X.; Berger, Heinrich; Casutt-Schneeberger, Julia u. a.: *Geschichte Online: Einführung in das wissenschaftliche Arbeiten*, Wien, Köln, Weimar 2006, S. 266-270.

⁵² Lengwiler, Martin: *Praxisbuch Geschichte: Einführung in die historischen Methoden*, Zürich 2011. Zur E-History S. 254-272.

⁵³ Rogers: *Digital methods*, 2013.

⁵⁴ Galgano, Michael J.; Arndt, J. Chris; Hyser, Raymond M.: *Doing history research and writing in the digital age*, Boston, Mass 2008.

vernachlässigt wird. Auch bei *History in the Digital Age*⁵⁵ der Geisteswissenschaftlerin Toni Weller wird die Quellenkritik des Digitalen leider nur beiläufig erwähnt. Allerdings wird von den Autoren sehr deutlich hervorgehoben, dass sich die Arbeitsweise der Historiker unter dem Einfluss des digitalen Zeitalters ändert.⁵⁶ Weller und Galgano bieten einen kompakten und strukturierten Einstieg ins Thema der Digital History.

Verschiedene Aspekte der neuen, digitalen Arbeitstechniken haben Gerben Zaagsma in Artikel *On Digital History*⁵⁷, John Fritch und Robert Cromwell in *Evaluating Internet Resources*⁵⁸, Bernhard Rieder und Theo Röhle in *Digital Methods*⁵⁹ und Jakob Krameritsch in den Beiträgen *Geschichte(n) im Netzwerk* und *Hypertext schreiben*⁶⁰ diskutiert. Der Historiker Zaagsma bietet eine kompakte und gute Einführung in das gesamte Thema und beleuchtet mit einer seiner Kernaussagen, dass sich Geschichte im digitalen Zeitalter ihrer Hybridität bewusst sein muss, einen wichtigen Aspekt der digitalen Geschichtswissenschaft. Der Archiv- und Kommunikationswissenschaftler Fritch und der Elektroingenieur Cromwell konzentrieren sich darauf, wie digitale Objekte aus dem Internet kritisch untersucht werden können. Die Geistes- und Kommunikationswissenschaftler Rieder und Röhle sprechen vor allem die Herausforderungen an, welchen sich die Geschichtswissenschaft im digitalen Zeitalter zu stellen hat, geben aber nur einen Anstoss zur Auseinandersetzung mit neuen digitalen Methoden. Der Historiker Jakob Krameritsch beschäftigt sich ebenfalls mit digitalen Arbeitstechniken, insbesondere mit der Darstellung von Geschichte in Digitalmedien, in der er den hypertextuellen einem chronologischen Ablauf vorzieht. Techniken für die Quellenkritik stellt er aber nur am Rande vor.

Wie Eva Pfanzelter beschäftigt sich Peter Haber in seiner Habilitationsschrift *digital past*⁶¹ explizit mit historischer Quellenkritik des Digitalen. Einerseits zeigt er auf, wie der Computer in die Geschichtswissenschaft eingeführt wurde, andererseits setzt er sich mit der Problematik der Digitalisierung der Wissenschaft auseinander. Zur historischen Methode im

⁵⁵ Weller, Toni (Hg.): *History in the digital age*, London 2013.

⁵⁶ Siehe Zusammenfassung: Weller, Toni: Conclusion. A changing field, in: Ebd., S. 195-205.

⁵⁷ Zaagsma, Gerben: On Digital History, in: *BMGN – Low Countries Historical Review* 128 (4), 16.12.2013, S. 3-29.

⁵⁸ Fritch, John W.; Cromwell, Robert L.: *Evaluating Internet Resources: Identity, Affiliation, and Cognitive Authority in a Networked World*, in: *Journal of the American Society for Information Science and Technology* 52 (6), 2001, S. 499-507. DOI: 10.1002/asi.1081.

⁵⁹ Rieder, Bernhard; Röhle, Theo: *Digital Methods: Five Challenges*, in: *Understanding Digital Humanities*, Houndmills, Basingstoke, Hampshire; New York 2012, S. 67-84.

⁶⁰ Krameritsch, Jakob: *Geschichte(n) im Netzwerk. Hypertext und dessen Potenziale für die Produktion, Repräsentation und Rezeption der historischen Erzählung*, Münster 2007 (Medien in der Wissenschaft); Krameritsch, Jakob: *Hypertext schreiben*, in: *Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften*, 08.04.2010, S. 83-95.

⁶¹ Haber, Peter: *Digital Past: Geschichtswissenschaft im digitalen Zeitalter*, München 2011. Zusätzlich zu dieser Publikation wurde die unveröffentlichte Habilitationsschrift verwendet, die in verdankenswerter Weise von der Familie Haber zur Verfügung gestellt wurde. Siehe: Haber: *digital past* (unv. Habil.), 2009.

21. Jahrhundert stellt er die Frage, ob eine "den digitalen Medialitäten angepasste Hilfswissenschaft gefordert ist, in deren Zentrum die Quellenkritik des Digitalen stehen müsste"⁶² oder ob der digitale Wandel aufgrund der methodischen Veränderungen als den Kernbestandteil des Faches zu konzeptualisieren ist. Haber gibt nach einer kurzen Einführung in die digitale Quellenkritik keine explizite Antwort, sondern verweist nur auf die pragmatische Verwendung von Checklisten zur Beurteilung von Quellen. Haber, der leider 2013 überraschend verstarb, konnte diese Fragen nicht mehr selber weiterverfolgen. Dennoch: da sich Haber mit den Anfängen des Einsatzes der elektronischen Datenverarbeitung in den Geschichtswissenschaften und Arbeitstechniken rund um Google, Data Driven History und Geschichtswissenschaft 2.0⁶³ beschäftigte, gehört das Werk tatsächlich zur Pflichtlektüre für Interessierte der digitalen Geschichtswissenschaft, wie der Historiker Mills Kelly es ausdrückte.⁶⁴

Simon Margulies' *Digitale Daten als Quelle der Geschichtswissenschaft*⁶⁵ ist die bisher einzige Arbeit, die sich im Kontext der Geschichtswissenschaft mit dem digitalen Objekt an sich beschäftigt. Die beiden Hauptkapitel zur Quellengenese und Quellensuche sind allerdings ohne informationstechnische Kenntnisse schwer nachvollziehbar. Im letzten Hauptkapitel stellt der Historiker Margulies die Quellenkritik digitaler Daten detaillierter dar, vor allem im Hinblick auf deren Überlieferungsgeschichte und die nötigen archivarisches Massnahmen zur Erhaltung. Margulies bietet einen fundierten Einstieg in das Thema, weil er auch auf die Eigenschaften digitaler Daten eingeht, verbleibt aber hauptsächlich im informationstechnischen Bereich. Er weist darauf hin, dass "diese Konzepte bedingen, wie Gegenstände in verschiedenen Attributen und Beziehungen abgebildet werden und definieren somit die Herstellung digitaler Daten. Im übertragenen Sinne entsteht dabei das, was Droysen als die Art der Materialien bezeichnet hat."⁶⁶ Margulies beschränkt sich allerdings auf Konzepte von Informationssystemen und ordnet die Thematik nicht in die Methodenproblematik der Geschichtswissenschaft im Umgang mit digitalen Objekten der gesamten digitalen Welt ein.

Mit einer Anpassung der traditionellen historisch-kritischen Methode auf die Anforderungen für die Untersuchung von digitalen Objekten haben sich die Geisteswissenschaftler Jesper Verhoef und Melvin Wever der Universität Utrecht (NL) befasst. Sie schlagen mit dem ,Digital

⁶² Haber: Digital Past, 2011, S. 104-112, hier S. 106.

⁶³ Geschichtswissenschaft 2.0 ist ein Synthesebegriff zwischen Geschichtswissenschaft und Web 2.0. Siehe zu Web 2.0 Kap. 4.1. Medien.

⁶⁴ Siehe Buchrückseite von Haber: Digital Past, 2011.

⁶⁵ Zur Quellenkritik in: Margulies, Simon B.: Digitale Daten als Quelle der Geschichtswissenschaft: Eine Einführung, Hamburg 2009 (Kölner Beiträge zu einer geisteswissenschaftlichen Fachinformatik 2), S. 319-408.

⁶⁶ Ebd., S. 196.

Humanities Cycle⁶⁷ einen strukturierten Prozess vor, wie digitale Quellen unter Berücksichtigung ihrer Eigenschaften vor. In ihrem Lösungsansatz legen sie den Fokus allerdings mehr auf digitale Tools, als auf die quellenkritische Analyse von digitalen Quellen.

In der aktuellen Diskussion wird die Quellenkritik des Digitalen weiterhin nur bruchstückhaft behandelt. So findet sich beispielsweise in der Publikation *WerkstattGeschichte* (hrsg. vom Verein für kritische Geschichtsschreibung) nur der Artikel *Wikipedia – Zum Verhältnis von Struktur und Wirkungsmacht eines heimlichen Leitmediums* der Historikerin Maren Lorenz, in dem sie nur kurz auf quellenkritische Fragen eingeht.⁶⁸ Auch die *Digitale Geschichtswissenschaft* (Arbeitsgruppe des Verbands der Historiker und Historikerinnen Deutschlands e.V., DE) mit wissenschaftlichen Beiträgen und Veranstaltungshinweisen zur Digital History im deutschsprachigen Raum, *Geschichte & Informatik G&I* (Verein, CH) mit Beiträgen und Eigenpublikationen zur Schweizer Digital History oder *infoclio* (Fachportal zur Schweizer [digitalen] Geschichtswissenschaft, CH) mit Beiträgen zu Veranstaltungen, Projekten in der Schweizer Digital History sowie hilfreichen Tools für die Digitalgeschichte, veröffentlichen selten Beiträge zum vorliegenden Thema.⁶⁹

1.1.3. Informatik

Die Grundlagen der Informationstechnik werden in den Geistes- und Sozialwissenschaften nur rudimentär oder gar nicht vermittelt. Nur in wenigen Studiencurricula sind entsprechende Veranstaltungen verpflichtend, wenn es sich nicht um einen der wenigen spezifischen Lehrgänge zu Digital Humanities wie an den Universitäten Würzburg, London oder Chicago handelt oder als Zusatzkurs wie in Dublin belegt wird.⁷⁰

Die Informatik⁷¹ oder Informationstechnik (IT) als Wissenschaft der automatisierten Informationsverarbeitung⁷² hat die Grundlagen der elektronischen Datenverarbeitung entwickelt. Ohne informationstechnische Infrastruktur⁷³ ist die Erstellung, Bearbeitung oder Weiterverarbeitung eines digitalen Objekts nicht möglich. Die Kenntnis der elektronischen Prozesse und

⁶⁷ Teilweise von ihnen auch als 'Digital History Cycle' bezeichnet. Siehe: Verhoef, Jesper; Wevers, Melvin: *The Digital Humanities Cycle: Hermeneutics, Heuristics, and Source Criticism in a Digital Age*, in: *Digital Humanities Benelux Conference*, Antwerpen 2015. URL: <http://www.dhbenelux.org/wp-content/uploads/2015/04/20.pdf>, Stand: 06.05.2018. Eine kurze Beurteilung erfolgt in Kap. 5.7.

⁶⁸ Siehe: Lorenz, Maren: *Wikipedia – Zum Verhältnis von Struktur und Wirkungsmacht eines heimlichen Leitmediums*, in: *Empire is coming home* (43), 2006, S. 84-95.

⁶⁹ *Digitale Geschichtswissenschaft* (URL: <http://digigw.hypotheses.org/feed>); *G&I* (URL: <http://blog.ahc-ch.ch/feed/>); *infoclio* (URL: <https://www.infoclio.ch/de/news/rss.xml>).

⁷⁰ Für Würzburg siehe URL: <http://www.uni-wuerzburg.de/fuer/studierende/angebot/faecher/digihum>; London siehe URL: <http://www.ucl.ac.uk/dh/courses/mamsc>; Chicago siehe URL: <http://www.luc.edu/ctsdh/academics/maindigitalhumanities/>; Dublin siehe URL: <http://dh.tcd.ie/dhc/>.

⁷¹ Im Englischen wird dafür der Begriff 'computer science' (Computerwissenschaft) verwendet.

⁷² Herold, Helmut; Lurz, Bruno; Wohlrab, Jürgen: *Grundlagen der Informatik: praktisch – technisch – theoretisch*, München, Boston 2007, S. 24.

⁷³ Bspw. Personal Computer, Netzwerke, Computerprogramme, Speichermedien etc.

Eigenschaften von digitalen Objekten ist deshalb für die Erarbeitung einer Quellenkritik des Digitalen unerlässlich. Fundierte Einführungen in die IT bieten Standardwerke wie *Grundlagen der Informatik*⁷⁴ der Informationswissenschaftler Helmut Herold, Bruno Lurz und Jürgen Wohlrab.

Die (Weiter-)Entwicklung von Hard- und Software und auch die Entwicklung des Internets schreiten rasch voran, grundsätzlich unabhängig von den Geisteswissenschaften. Forschungen in den Digital Humanities haben sich folglich stets auf Neuentwicklungen aus der Informatik einzustellen und können darauf nur marginal Einfluss nehmen, indem sie beispielsweise Ansprüche an die Hardware in Bezug auf Archivierungssysteme definieren. Die Softwareentwicklung für Anwendungen in den Digital Humanities wie beispielsweise in der Text Encoding Initiative⁷⁵ (TEI) basiert auf bestehenden Informationssystemstrukturen, die Entwicklung gänzlich neuer wird weder als zielführend noch als wirtschaftlich erachtet.⁷⁶

Was ein digitales Objekt ist, welche Eigenschaften diesem zuzuschreiben sind und wie es als historische Quelle verwendet werden kann, wurde bisher nicht umfassend untersucht. Der Physiker Lukas Rosenthaler hat in seiner unveröffentlichten Habilitationsschrift *Archivierung im digitalen Zeitalter*⁷⁷ verschiedene Eigenschaften aus der Sicht einer technischen Archivierungsproblematik vorgestellt. Die Historikerin Catherina Schreiber hat in ihrer Untersuchung *Genuine Internetdaten als historische Quellen*⁷⁸ einen multidimensionalen und auf eine korrealistische Quellentheorie bezogenen Quellenbegriff vorgestellt. Beide führen zwar wichtige Merkmale für eine allgemeine Definition auf, sind aber noch zu spezifisch auf ihre Blickwinkel fokussiert. Simon Margulies hat sich, wie bereits erwähnt, vertiefend mit digitalen Objekten auseinandergesetzt. Allerdings fokussiert sich Margulies in seinem Definitionsansatz auf die quellengenetische Modellierung von Datenbanken und schliesst damit weitere potenzielle digitale Quellen aus.

Die Problematik der Langzeitarchivierung digitaler Objekte hat Rosenthaler ebenfalls in seiner Arbeit aufgearbeitet. Des Weiteren wurde für diese Arbeit *Preserving Digital Information*⁷⁹ des Computerspezialisten und Herausgebers von *Digital Document Quarterly DDQ* Henry Gladney und *Web Archiving*⁸⁰ des heutigen Direktors der *Internet Memory Foundation*⁸¹ Julien Masanès verwendet. Zwar hat sich die Speicher- und Archivtechnik in

⁷⁴ Herold; Lurz; Wohlrab: Grundlagen der Informatik, 2007.

⁷⁵ URL: <http://www.tei-c.org/>.

⁷⁶ Aus Gesprächen mit diversen Personen an Veranstaltungen.

⁷⁷ Rosenthaler: Archivierung im digitalen Zeitalter, 2007.

⁷⁸ Schreiber, Catherina: Genuine Internetdaten als historische Quellen – Entwurf einer korrealistischen Quellentheorie, in: Zeitschrift für digitale Geschichtswissenschaften 1, 16.08.2012. URL: <http://universaar.uni-saarland.de/journals/index.php/zdg/article/view/292>, Stand: 11.12.2012.

⁷⁹ Gladney, Henry M.: Preserving Digital Information, Berlin 2007.

⁸⁰ Masanès, Julien: Web Archiving, Berlin 2006.

⁸¹ URL: <http://internetmemory.org/>.

den letzten Jahren stark weiterentwickelt, aber die Autoren stellen die Grundlagen der Bestandserhaltung digitaler Objekte dar, welche sich bisher wenig veränderten. Ergänzend zu Masanès wurden verschiedene Beiträge des Geisteswissenschaftlers Niels Brügger (Universität Århus, DK) verwendet, der intensiv zur Archivierung von Webseiten forscht.⁸² Seine Ausführungen bilden die Basis der Diskussion zur quellenkritischen Untersuchung von Webseiten.

Für die Wissenschaft von grosser Wichtigkeit sind digitale Forschungsinfrastrukturen, weil sie grosse Mengen an Forschungsdaten mit entsprechenden Vorkehrungen dauerhaft, verlässlich und authentisch zur Verfügung stellen. Unter digitaler Forschungsinfrastruktur werden digitale Informationssysteme verstanden, die von einer öffentlichen oder privaten Institution betrieben werden und die auf die Ansprüche von wissenschaftlich Forschenden ausgerichtet sind. Dies bedingt eine enge Zusammenarbeit der Geisteswissenschaften mit dem Fachbereich Informatik, der auf ihre Bedürfnisse abgestimmte Hard- und Software zur Verfügung stellt respektive mit ihr solche entwickelt.

Die Langzeitarchivierung von digitalen Daten, die deren dauerhafte Speicherung und Wiederverwendung erst ermöglicht, erfordert spezielle Strategien. Nationale und internationale Kompetenznetzwerke wie nestor⁸³, KOST⁸⁴, DARIAH⁸⁵ tragen zur Standardisierung und Umsetzung dieser spezifischen Ansprüche bei. Projekte wie das *Data and Service Center for the Humanities (DaSCH)* der Schweizerischen Akademie für Geistes- und Sozialwissenschaften (SAGW) stecken noch im Aufbau.⁸⁶

Die Archivwissenschaften befassten sich schon früh mit dem Problem der Langzeitarchivierung digitaler Objekte. Bereits 1975 erschien mit *EDV und Archive – Ein Ratgeber*⁸⁷ des Archivaren Horst Romeyk das erste Handbuch zum Thema, in dem er mit der Anwendung der IT in der öffentlichen Verwaltung die Entstehung eines neuartigen Registraturgutes⁸⁸ fest-

⁸² Brügger, Niels: *Archiving Websites. General Considerations and Strategies*, Aarhus 2005; Brügger, Niels: *Web history*, New York 2010; Brügger, Niels: *Web History and the Web as a Historical Source*, in: *Zeithistorische Forschungen/Studies in Contemporary History*, Online-Ausgabe 9 (2), 2012. URL: <http://www.zeithistorische-forschungen.de/16126041-Bruegger-2-2012>, Stand: 06.05.2018; Brügger, Niels: *When the Present Web is Later the Past: Web Historiography, Digital History, and Internet Studies*, in: *Historical Social Research* 37 (4), 2012, S. 102-117.

⁸³ URL: <http://www.nestor.sub.uni-goettingen.de/>.

⁸⁴ KOST: Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen. URL: <http://kost-ceco.ch/>.

⁸⁵ DARIAH: Digital Research Infrastructure for the Arts and Humanities. Siehe: Romary, Laurent; Neuroth, Heike; Fritze, Christiane u. a.: *DARIAH-EU – Digital Research Infrastructure for the Arts and Humanities*, in: *Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Digital Humanities und Web 2.0*, 2012, S. 36-39.

⁸⁶ Siehe URL: <http://www.sagw.ch/de/sagw/die-akademie/unternehmen/DaSCH.html>.

⁸⁷ Romeyk, Horst: *EDV [Elektronische Datenverarbeitung] und Archive*, Siegburg 1981², S. 11. 1. Auflage maschinenschriftlich 1975.

⁸⁸ Registraturgut: alle Objekte, die aus der Tätigkeit von Behörden entstehen, aber noch nicht dem Archiv angeboten wurden.

stellte, das wie bestehende Objekte bewertet, erschlossen und gelagert werden müsse.⁸⁹ Diese wichtige Erkenntnis, dass es sich um ein *neuartiges* Gut handelt, setzt sich erst langsam durch, denn sowohl in der Praxis, als auch in der Wissenschaft wird auch heute noch oft ein physisches Objekt einem digitalen gleichgesetzt.⁹⁰ Romeyk irrt allerdings in der Behauptung, dass es wie bestehende Objekte gelagert werden kann, denn digitale Objekte haben keine Materialität, sind also nicht physisch vorhanden.

Die Historiker Janusch Carl und Andreas Rutz stellten in einem kurzen Aufsatz über Archive und IT fest, dass in den Archivwissenschaften der Umgang mit digitalen Daten selbst 2012 noch immer marginal behandelt wird.⁹¹ Einzig der Aufsatz *Neue Informationstechnologien und Archive*⁹² des Archivaren Peter Worm wird als Ausnahme genannt. Allerdings haben sie wahrscheinlich die Publikationen des Schweizer Vereins *Geschichte und Informatik* übersehen, der seit 1990 einige Beiträge und 1999 auch ein Band zum Thema *Informatik im Archiv = L'Informatique des archives* veröffentlichte.⁹³ Mittlerweile sind weitere Publikationen zum Thema erschienen wie *Advanced Digital Preservation* des Physikers David Giarretta oder *Practical Digital Preservation* von Adrian Brown (Direktor von *The Parliamentary Archives*, UK).⁹⁴

Viele Gedächtnisinstitutionen⁹⁵ setzen mittlerweile standardisierte Methoden zur digitalen Langzeitarchivierung produktiv ein, nachdem sie den Experimentalbetrieb erfolgreich abgeschlossen haben.⁹⁶ Es besteht allerdings in der Archivwissenschaft kein Konsens, was unter

⁸⁹ Weitere Ausführungen in: Carl, Janusch; Rutz, Andreas: Bits and Bytes statt Pergament und Papier? Das digitale historische Archiv Köln und die Zukunft des Kölner Stadtarchivs im Web 2.0, in: *Archivar*, Zeitschrift für Archivwesen 2, 2012, S. 143-153.

⁹⁰ Eigene Erfahrungen als Forscher und als wissenschaftlicher Assistent in einem Staatsarchiv.

⁹¹ Carl; Rutz: Bits and Bytes statt Pergament und Papier?, 2012, S. 145.

⁹² Worm, Hans-Jürgen: *Neue Informationstechnologien und Archive*, in: Reimann, Norbert (Hg.): *Praktische Archivkunde: ein Leitfaden für Fachangestellte für Medien- und Informationsdienste: Fachrichtung Archiv*, Münster 2008, S. 219-237.

⁹³ Verein 'Geschichte & Informatik' (URL: <http://blog.ahc-ch.ch/>). Zeitschrift von 1990-1995 'Bulletin Geschichte und Informatik / Histoire et Informatique', ab 1996/97 ohne 'Bulletin'; ISSN 1420-5955.

⁹⁴ Giarretta, David: *Advanced Digital Preservation*, Berlin 2011; Brown, Adrian: *Practical Digital Preservation: a How-to Guide for Organizations of any Size*, London 2013. Weiterführende archivwissenschaftliche Literatur: Nicholas, David; Herman, Eti (Hg.): *Information Science*, London 2014 (*Critical concepts in media and cultural studies*); Gladney: *Preserving Digital Information*, 2007; Masanès: *Web Archiving*, 2006; Brügger: *Archiving Websites*, 2005; Brügger, Niels: *Website Analysis. Elements of a conceptual architecture*, Aarhus 2010 (CFI monograph series 12). URL: http://cfi.au.dk/fileadmin/www.cfi.au.dk/publikationer/cfis_skriftserie/012_brugger.pdf, Stand: 24.09.2015; Rosenthaler: *Archivierung im digitalen Zeitalter*, 2007.

⁹⁵ Als Gedächtnisinstitutionen werden gesellschaftliche, staatliche und kirchliche Einrichtungen bezeichnet, die (oft gesetzlich beauftragt und staatlich unterstützt) Überlieferungsgut und Kulturerbe sammeln und langfristig/dauerhaft aufbewahren. Dazu gehören insbesondere Gallerien, Bibliotheken, Archive und Museen (sog. BAM- (dt.) oder GLAM (eng.) -Institutionen). Nicht dazu gezählt werden reine Leihbibliotheken und Institutionen ohne eigene Sammlung.

⁹⁶ Nach Adam Farquhar von der British Library. Siehe: Lindstädt, Birte; Schmidt, Christoph: «Digital Preservation Summit 2011», Globales Gipfeltreffen zur digitalen Langzeitarchivierung; Tagungsbericht, in: *Archivar*, Zeitschrift für Archivwesen 2, 2012, S. 176.

digitaler Langzeitarchivierung genau zu verstehen ist.⁹⁷ Der Digitalwissenschaftler Adam Farquhar weist zumindest darauf hin, was digitale Langzeitarchivierung nicht sei: nämlich die Auswahl oder Beschaffung digitaler Inhalte, die Dokumentation des Web oder die Digitalisierung analoger Inhalte.⁹⁸

1.1.4. Medientheoretische Zugänge

Der Diskurs über die Medialität von Computern und Daten wird hauptsächlich in den Medienwissenschaften geführt und hier nur ansatzweise und beispielhaft behandelt. Die sich daraus ergebenden Fragen, Probleme und Herausforderungen für die Geistes- und Sozialwissenschaften, und damit auch für die Geschichtswissenschaft, werden hauptsächlich im Rahmen der Diskussionen über den 'Computational Turn' respektive das digitale Zeitalter geführt.⁹⁹ Für die Quellenkritik des Digitalen sind medientheoretische Grundlagen von Interesse, weil das Medium mitbestimmt, wie eine Information weitergegeben wird. Eine umfassendere Untersuchung hat der Jurist und Technikforscher Michael Nentwich in *Cyberscience 2.0* präsentiert. Darin untersucht er die Rolle der digitalen sozialen Kulturen, die durch das Web 2.0 eingeleitet wurden, in der akademischen Welt.¹⁰⁰ Laut Nentwich ist Cyberscience nun Realität und wird sich immer weiter in den *Cyberspace*¹⁰¹ verlagern. Unter Cyberscience versteht er die Wissenschaft, die sich in der virtuellen Welt bewegt und sich mit ihr beschäftigt. Wie die Cyberscience mit einem digitalen Objekt als Quelle wissenschaftlicher Forschung konkret umzugehen hat, diskutiert er nicht – das ist aber auch nicht Gegenstand seiner Untersuchung. Eine aktuelle medientheoretische Auseinandersetzung mit dem Digitalen leistet der Medienwissenschaftler Marcus Burkhardt in *Digitale Datenbanken*¹⁰², indem er die Medialität digitaler Datenbanken zu bestimmen und zu beschreiben versucht und eine kritische Datenbankkultur fordert. Zu den Themen Medienwandel und Copyright

⁹⁷ Farquhar, in: Ebd.

⁹⁸ Farquhar, in: Ebd.

⁹⁹ Vertiefend in Kap. 0

Umbruch in ein Digitales Zeitalter.

¹⁰⁰ Nentwich, Michael: *Cyberscience 2.0 research in the age of digital social networks*, Frankfurt a.M. 2012 (Interaktiva).

¹⁰¹ Cyberspace: Cybernetic Space; von digitalen Informationssystemen generierte virtuelle Welt, die der Nutzer mit Hilfe eines solchen Systems 'besuchen' kann. Nach: Prevezanos, Christoph: *Computer Lexikon 2013*, München 2012 (Always learning). Der Begriff cyberspace suggeriert zwar die Vorstellung von Raum und Ort, aber Distanz und Zeit spielen keine Rolle unter den interagierenden Objekten. Der Begriff steht mehr als Metapher für ein 'grenzenloses Medium für Kommunikation'. Vgl.: Tripathi, Arun Kumar: *The Significance of Digital Hermeneutics for the Philosophy of Technology*, in: Kelly, Matthew; Bielby, Jared (Hg.): *Information Cultures in the Digital Age: A Festschrift in Honor of Rafael Capurro*, Wiesbaden 2016, S. 152.

¹⁰² Burkhardt, Marcus: *Digitale Datenbanken: eine Medientheorie im Zeitalter von Big Data*, Bielefeld 2015 (Digitale Gesellschaft).

habe ich vor allem Monika Dommanns *Autoren und Apparate* verwendet.¹⁰³ Die Historikerin und Wirtschaftswissenschaftlerin Dommann hat in ihrer Habilitationsschrift den Konflikt zwischen Autoren(rechten) und Apparaten aufgearbeitet und wie dieser mit gesetzlichen Bestimmungen (Urheberrecht) zu entschärfen versucht wird.

1.1. Fragestellung, Forschungsfragen und -thesen

Obwohl die Geschichtswissenschaft seit den 1960er Jahren Erfahrungen im Umgang mit digitalen Objekten sammeln konnte, wurden die verschiedenen Aspekte der quellenkritischen Analyse separat behandelt. In der Digital History hat sich die Forschung zwar an das Thema der Quellenkritik des Digitalen angenähert (Pfanzer, Haber, Margulies) und einige Digitalhistoriker¹⁰⁴ haben sich mit Aspekten der Quellenkritik beschäftigt (Schmale, Groebner, Eickels, Lengwiler), aber eine systematische Untersuchung zu 'traditionellen' quellenkritischen Methoden und den 'neuartigen' Quellen ist jedoch bisher nicht erfolgt.

Informatik, Medienwissenschaft sowie Archiv- und Bibliothekswissenschaft sind Fachbereiche mit grossem Einfluss auf die Quellenkritik des Digitalen und für die Entwicklung einer allfälligen neuen quellenkritischen Methode miteinzubeziehen. Die Informatik bildet einerseits die Grundlagen des Digitalen, andererseits entwickelt sie sich so rasch weiter, dass andere Wissenschaften nur darauf reagieren können. Konzepte der Medienwissenschaften können zur Gewinnung von Erkenntnissen über die Nutzung von digitalen Medien in den Geschichtswissenschaften beitragen. Eine Anwendung solcher Erkenntnisse auf den quellenkritischen Prozess wurde jedoch noch nicht vorgenommen. Hingegen sind Methoden der Langzeitarchivierung bereits im Einsatz, Langzeiterfahrungen fehlen allerdings. Viele Wissenschaftler vertrauen auf die wissenschaftliche Verwendbarkeit von Digitalarchiven, aber wie sich darin verfügbare digitale Objekte quellenkritisch überprüfen lassen, ist nicht geklärt. Aus diesen Ausführungen geht hervor, dass eine Quellenkritik des Digitalen ein Forschungsdesiderat ist.

Die vorliegende Arbeit ist deshalb von der zentralen Fragestellung geleitet, **wie sich die historische Quellenkritik durch die Verwendung von digitalen Objekten als Forschungsressource sowie digitalen Informations- und Kommunikationsmedien verändert.**

¹⁰³ Dommann, Monika: *Autoren und Apparate*. Die Geschichte des Copyrights im Medienwandel, Frankfurt a.M. 2014 (Fischer Wissenschaft).

¹⁰⁴ Zum Begriff siehe: Föhr, Pascal: 'just a historian', *Historical Source Criticism*, 09.08.2013, URL: <http://hsc.hypotheses.org/241>, Stand: 06.11.2014.

Spezifisch werden sieben Fragen untersucht:

1. Welche Eigenschaften hat ein digitales Objekt und wie grenzt es sich von den bisher bekannten Objekten ab?

Bis zur 'digitalen Wende' wurden von den Historikern hauptsächlich physische, also haptisch erfahrbare Quellen – sei dies eine Urkunde oder ein Buch. Bei der Verwendung von digitalen Objekten gilt es also zunächst deren Eigenschaften zu definieren, um sie überhaupt adäquat quellenkritisch überprüfen zu können. Thematisiert werden muss in diesem Zusammenhang, ob bei einer starken Unterscheidung digitaler Objekte von bisher bekannten Objekten eventuell von einer neuartigen Quellengattung zu sprechen ist.

2. Gibt es neue Quellen, die es analog nicht gibt, und wenn ja, welche?

Da davon ausgegangen werden kann, dass es Quellen gibt, die in der analogen Welt nicht zu finden sind, ist zu untersuchen, welche Quellentypen digitaler Objekte existieren.

3. Welche Probleme entstehen durch die Eigenschaften eines digitalen Objekts bei der historisch-quellenkritischen Überprüfung (im Gegensatz zu physischen Objekten)?

Es ist zu erwarten, dass bei der quellenkritischen Überprüfung von digitalen Objekten aufgrund ihrer besonderen Eigenschaften neue Herausforderungen entstehen, da die Methode hauptsächlich auf physische Objekte ausgelegt ist.

4. Können digitale Objekte mit der historisch-kritischen Quellenkritik überprüft werden?

5. Muss für digitale Objekte eine neue historisch-kritische Methode erarbeitet werden?

Trotz der oben genannten Herausforderungen soll geklärt werden, ob während des Überprüfungsprozesses die Authentizität des Objekts festgestellt werden kann. Sollte dies nicht der Fall sein, ist der historisch-kritische Prozess neu zu definieren beziehungsweise es gilt zu fragen:

6. Welche Methoden können für die Quellenkritik digitaler Objekte in der Geschichtswissenschaft angewendet werden?

Entsprechende Methoden sollen vorgestellt werden. Woraus die letzte Frage resultiert:

7. Verändert sich die Arbeitsweise der historischen Zunft durch die veränderten quellenkritischen Anforderungen?

Die Arbeit mit den digitalen Quellen wird durch die festgestellten Änderungen im historisch-kritischen Überprüfungsprozess beeinflusst. Es ist deshalb zu untersuchen, inwiefern die Arbeitsweise anzupassen ist.

Ein digitales Objekt ist physisch nicht greifbar. Deshalb nehme ich an, dass es eine neue Art von Quelle darstellt, die sich aufgrund deren Eigenschaften von bisher bekannten abgrenzt (These 1). Auch gibt es neue Quellentypen, für die es bisher analog keine Analogien gibt (These 2). Aufgrund der Besonderheit des digitalen Objekts ist zudem davon auszugehen, dass die traditionelle historische Quellenkritik keine oder ungenügende Methoden zur Überprüfung der Echtheit zur Verfügung stellt (These 3). Digitale Objekte benötigen zur quellenkritischen Überprüfung daher (auch) Methoden aus der Informatik (These 4). Die quellenkritische Überprüfung eines digitalen Objekts bedeutet, dass der Forschende einen Näherungswert zu erarbeiten hat, der die subjektive Wahrscheinlichkeit ausdrückt, für wie authentisch er das Objekt hält (These 5). Aufgrund dieser Annahmen lässt sich schliessen, dass sich die Arbeitsweise der historischen Zunft durch Digitalität verändert (These 6).

Die Aufarbeitung des Themas mit allen damit zusammenhängenden Aspekten ist aus Platzgründen nicht möglich und wäre nur durch ein interdisziplinäres Forschungsprojekt mit mehreren Beteiligten zu erreichen. Die für die Quellenkritik des Digitalen wesentlichen Aspekte der verschiedenen Fachbereiche werden berücksichtigt, aber Grundlagenkenntnisse werden nur wo nötig vermittelt. Wichtig sind vor allem Begriffe, Methoden und Konzepte aus der Informationstechnik, die nur soweit erklärt werden, wie sie zum Verständnis der Argumentation in dieser Arbeit beitragen. Die Speicherung und Archivierung digitaler Daten werden zwar behandelt, die Methoden und Praxis der Archiv- und Bibliothekswissenschaften aber nur am Rande erwähnt. Vertiefende Einblicke in die Methoden der Langzeitarchivierung von digitalen Daten oder des Records Managements bleiben deshalb ebenso aussen vor wie die Organisation von Bibliotheken.¹⁰⁵ Die vorliegende Arbeit geht zudem auch nur kurz auf die didaktische Verwendung digitaler Medien ein. Fragen des Copyrights, des Leistungsschutzrechts, der Anti-Piracy-Gesetze, des Lizenzrechts etc. können nicht eingehend behandelt werden, da diese Themen eigener Darstellungen in Zusammenarbeit mit Juristen bedürfen. Der Leser soll aber durch die vorliegenden Ausführungen für die Problematik sensibilisiert werden. Medientheoretische Grundlagen und die Auswirkungen von Digitalmedien auf die Geschichtswissenschaft werden nur soweit dargelegt, als dass sie für das Verständnis des ‚neuen‘ Mediums nötig sind.

Manche der vorgeschlagenen Lösungsansätze verfolgen einen hohen, idealistischen Anspruch, der in der Praxis kaum um- oder durchsetzbar scheint. Forschung und Wissenschaft schreiten aber nur dadurch voran, dass hoch gesteckte Ziele angestrebt und hoffentlich irgendwann erreicht werden. Dafür sind Anstrengungen in der Ausbildung der Historiker, finanzielle Mittel und weiterer Forschung nötig. Die vorliegende Arbeit soll zu einer fruchtbaren Diskussion beitragen.

¹⁰⁵ Dazu weiterführend: Masanès: Web Archiving, 2006; Gladney: Preserving Digital Information, 2007; Toebak, Peter M.: Records Management ein Handbuch, Baden 2007.

Die historisch-kritische Untersuchung der verwendeten digitalen Quellen erfolgte nach derzeit gültigen Standards, da die digitalen Mechanismen zu deren Authentizitätsprüfung grösstenteils erst noch zu implementieren sind.

1.2. Aufbau

Zuerst stelle ich die Grundlagen für die Erarbeitung der Quellenkritik des Digitalen kurz dar. Im folgenden, zweiten Kapitel *1 und 0* untersuche ich die Eigenschaften eines digitalen Objekts und diskutiere Antworten auf die ersten zwei Fragen und Thesen. Im dritten Kapitel wird geklärt, ob digitale Objekte respektive Forschungsressourcen durch die traditionelle Quellenkritik überprüft werden können, was die Fragen drei bis fünf beantwortet. Ich erwarte, dass bei der quellenkritischen Überprüfung von digitalen Objekten aufgrund ihrer Eigenschaften Probleme entstehen, da die Methode hauptsächlich auf physische Objekte ausgelegt ist. Im Kapitel 4 gehe ich der Frage nach, wie Digitalmedien die Wissenschaft beeinflusst und ob von einem 'Digitalen Zeitalter' gesprochen werden kann.

Im zentralen fünften Kapitel Quellenkritik des Digitalen zeige ich anhand des überarbeiteten Prozesses die konkreten Probleme bei der quellenkritischen Überprüfung von digitalen Objekten auf und stelle einige Lösungsansätze vor. Im anschliessenden Kapitel Digitalhistoriker zeige ich auf, welche Massnahmen ergriffen werden sollten, um Digitalhistoriker auszubilden und ihnen adäquate Infrastrukturen zur Verfügung zu stellen. In den Schlussbetrachtungen werden die Erkenntnisse zusammengefasst.

Im ersten Unterkapitel des Anhangs sind die von mir vorgeschlagenen Lösungsansätze mit einer kurzen Erklärung aufgelistet, im zweiten das Literaturverzeichnis.

1.3. Methodische Herausforderung

Die Entwicklung einer Quellenkritik des Digitalen basiert auf der 'traditionellen' historisch-kritischen Methode (im Folgenden als traditionelle Methode¹⁰⁶ bezeichnet). Diese wird seit ihrer eigenen Entwicklung im 19. Jahrhundert¹⁰⁷ für die Untersuchung von hauptsächlich analogen Quellen erfolgreich angewendet. Digitale Objekte unterscheiden sich aber von analogen durch mehrere Eigenschaften, insbesondere durch fehlende Materialität.¹⁰⁸ Zudem beginnen wir auch erst langsam zu verstehen, wie die neue Technologie der Digitalmedien den Menschen in seiner Denkweise, seinem Handeln und seinen Beziehungen zur Umwelt

¹⁰⁶ Methode: Auf einem Regelsystem aufbauendes Verfahren, das zur Erlangung von wissenschaftlichen Erkenntnissen oder praktischen Ergebnissen dient. Wermke, Matthias; Klosa, Annette; Kunkel-Razum, Kathrin u. a. (Hg.): Duden - das Fremdwörterbuch, Bd. 5, Mannheim, Leipzig, Wien, Zürich 2001 (Der Duden in zwölf Bänden).

¹⁰⁷ Siehe ausführlicher in Kap. 3. Traditionelle Methode und digitale Objekte.

¹⁰⁸ Siehe ausführlicher in Kap. 2. 1 und 0.

und zu dieser Technologie verändert¹⁰⁹, was besonders in den Medienwissenschaften intensiv diskutiert wird.¹¹⁰ In den Geschichtswissenschaften dringt erst langsam die Erkenntnis durch, dass digitale Objekte eine andere Quellenart darstellen:

*"The use of digital sources, in other words, completely changes the landscape of information and transaction costs that historians have traditionally faced."*¹¹¹

Mangels Alternative wurde die traditionelle historisch-kritische Methode als Basis für die Untersuchung von neuartigen Quellen verwendet. Dies führte dazu, dass 'Methoden von gestern für Artefakte von heute'¹¹² verwendet wurden. Diese methodische Herausforderung war aber wenig gravierend, denn bei der traditionellen Methode ist zwischen dem Prozess an sich und den anzuwendenden Untersuchungstechniken zu unterscheiden. Letztere konnten nicht oder nur sehr eingeschränkt auf digitale Objekte angewendet werden, insbesondere ergeben Techniken zur Untersuchung von Materialität wenig sinnvolle Resultate.¹¹³ Der historisch-kritische Prozess blieb jedoch valide, wie weiter unten gezeigt wird. Es geht nach einem der Begründer der Methode, dem Historiker Johann Gustav Droysen, um die Vorgehensweise, wie "forschend zu verstehen"¹¹⁴ ist. Die vom Historiker Ernst Bernheim definierten Hauptaufgaben dieses Prozesses bestehen aus dem Sammeln der Überlieferungen und der Erkenntnis des Zusammenhangs und sind von der Materie des Objekts unabhängig.¹¹⁵ Sammeln und Verstehen bleiben aus offensichtlichen Gründen auch bei der Verwendung von digitalen Objekten wichtige Schritte im Prozess, denn ohne Quellen und Verstehen sind keine Erkenntnisse möglich. Das kritische Hinterfragen der vorliegenden, potenziellen Quellen bleibt eine grundsätzliche Aufgabe der historischen Arbeit und auch die Grundsätze des wissenschaftlichen Arbeitens sind weiterhin zu beachten.¹¹⁶

Für eine Untersuchung eines digitalen Objekts sind darauf angepasste Methoden und Hilfsmittel, insbesondere aus der IT, anzuwenden. Ich präsentiere verschiedene Lösungsvorschläge für eine Quellenkritik des Digitalen, wobei die meisten aufgrund fehlender informationstechnischer Infrastruktur (Hard- und Software) nur als Idee und beispielhaft formuliert

¹⁰⁹ Nach Sherry Turkle: Jenkins, Henry: «Does this Technology serve Human Purposes?» A «Necessary Conversation» with Sherry Turkle, in: Svensson, Patrik; Goldberg, David Theo (Hg.): *Between humanities and the digital*, Cambridge 2015, S. 72.

¹¹⁰ Siehe ausführlicher in Kap. 3. Traditionelle Methode und digitale Objekte.

¹¹¹ William J. Turkel in: Cohen, Daniel J.; Frisch, Michael; Gallagher, Patrick u. a.: *Interchange: The Promise of Digital History*, in: *The Journal of American History* 95 (2), 01.09.2008, S. 455. DOI: 10.2307/25095630.

¹¹² Nach Eva Pfanzelter. Persönliche Notiz vom 04.10.2015.

¹¹³ Das wird in Kap. 3. Traditionelle Methode und digitale Objekte untersucht.

¹¹⁴ Droysen, Johann Gustav: *Grundriss der Historik*, Leipzig 1875², §8.

¹¹⁵ Nach: Bernheim, Ernst: *Lehrbuch der historischen Methode und der Geschichtsphilosophie mit Nachweis der wichtigsten Quellen und Hilfsmittel zum Studium der Geschichte*, Leipzig 1903, S. 163-167.

¹¹⁶ Siehe ausführlicher in Kap. 3.4. Wissenschaftliches Arbeiten.

werden können.¹¹⁷ Diese Ideen werden hier erstmals der Wissenschaftsgemeinschaft vorgestellt und stehen deshalb erst jetzt für eine kritische Auseinandersetzung zur Verfügung.¹¹⁸

In dieser Arbeit werden bestehende Grundsätze und Konventionen des wissenschaftlichen Arbeitens eingehalten, auch wenn ich deren Anpassung propagiere.¹¹⁹

¹¹⁷ Die erwähnten Hard- und Softwares dienen nur als Beispiele und dürfen nicht als Empfehlung aufgefasst werden.

¹¹⁸ Weil in einer Dissertation neue Erkenntnisse, Theorien und Methoden gefordert werden, wurde von einer breiten Präsentation und Diskussion der vorliegenden Ideen auf dem Forschungsblog hsc.hypotheses.org abgesehen. Diese wurden nur in kleinen Kreisen vorgestellt und diskutiert.

¹¹⁹ Insbesondere die Anpassung der Belegkette in Kap. 5.6.3. Belege, Lösungsansatz.

2. 1 und 0

Die allermeisten Informationssysteme speichern ein digitales Objekt in einem binären Code: 0 oder 1.¹²⁰ Daraus leiten sich spezifische Eigenschaften ab, die sich teilweise von denen eines analogen Objekts unterscheiden. Für eine quellenkritische Untersuchung müssen diese Eigenschaften bekannt sein, was bisher kaum und wenn, dann mit Bezug zum dominierenden Thema Internet untersucht wurde.¹²¹ Auch wurde bisher ein digitales nicht von einem analogen abgegrenzt, was jedoch fundamental für das Verständnis der neuen Quellenart ist. Des Weiteren werden neue Quellentypen definiert, die zuvor nicht bekannt waren.

In diesem ersten Schritt werden die ersten zwei Forschungsfragen beantwortet und die zwei jeweils zugehörigen Thesen verifiziert. Zuerst gilt es folgende, zentrale Begriffe zu klären:

Daten, der Plural von Datum, sind Zeichen, die kleinste und unteilbare Elemente eines Wertebereichs darstellen.¹²² Daten werden immer generiert, und sind nie bloss gegeben.¹²³ In einen sinnvollen Zusammenhang gesetzt, können daraus durch Verarbeitung und Interpretation *Information(en)* entstehen, sofern durch entsprechendes, immanentes oder durch Metadaten erhaltenes Vorwissen ein Sinn erkennbar ist.¹²⁴ Für die Geschichtswissenschaft steht "die inhaltliche Interpretation und damit die semantische Bedeutung der Information im Zentrum."¹²⁵

*"Während Daten für sich genommen lediglich bedeutungslose Punkte in Raum und Zeit sind, entsteht Information dadurch, dass Daten zueinander in Beziehung gesetzt werden. Wissen wiederum entsteht aus dem Verständnis von Informationsmustern, also aus der Kontextualisierung von Informationen."*¹²⁶

¹²⁰ Andere Systeme wie das ternäre werden heute nicht mehr (weiter-)entwickelt. Siehe dazu bspw.: Hunger, Francis: SETUN eine Recherche über den sowjetischen Ternärcomputer, Leipzig 2007.

¹²¹ Ausnahmen sind bspw. Lukas Rosenthaler, Catherina Schreiber oder John Haugeland. Siehe Kap. 2.1.2. Definition der Eigenschaften.

¹²² «Datum, Data»: Fischer; Hofer: Lexikon der Informatik, 2011, S. 217.

¹²³ Vgl. Giessmann, Sebastian; Burkhardt, Marcus: Was ist Datenkritik? Zur Einführung, S. 3. URL: <http://www.medialekontrolle.de/wp-content/uploads/2014/09/Giessmann-Sebastian-Burkhardt-Marcus-2014-03-01.pdf>, Stand: 28.04.2016.

¹²⁴ Bspw. ergeben die Daten C, H, 1, 2, 3, 4, 5, 6 erst dann einen Sinn, wenn auch bekannt ist, dass es sich um eine IBAN-Nummer handelt. Zur Konzeption des Begriffs siehe bspw.: Case, Donald O.: The Concept of Information (2012), in: Information Science, Bd. I / IV, London 2014 (Critical concepts in media and cultural studies), S. 15-42.

¹²⁵ Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 16.

¹²⁶ Haber, Peter: «Google-Syndrom». Phantasmagorien des historischen Allwissens im World Wide Web, in: Haber, Peter; Epple, Angelika (Hg.): Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0, Zürich 2004 (Geschichte und Informatik 15), S. 75 f.

Verwandt, aber nicht gleichbedeutend sind die Begriffe *Fakt* und *Aussage*, die bereits mit einer Bedeutung aufgeladen sind, wohingegen ein Datum als 'neutral' angesehen werden kann – "Data has no truth."¹²⁷

Digitale Daten können im weitesten Sinne als mit symbolischen Zeichen codierte Informationen verstanden werden.¹²⁸ Als Zeichen kann jedes Element aus einem definierten und endlichen Zeichenvorrat aus unterscheidbaren physischen Markierungen verwendet werden, beispielsweise Ziffern, Buchstaben oder akustische Signale. Wenn sich der Zeichenvorrat aus genau zwei Zeichen wie 0 und 1 oder hell und dunkel zusammensetzt, wird von binären Daten gesprochen. Digitale Daten sind für den Menschen nicht lesbar und benötigen zu deren Darstellung ein digitales Informationssystem (Computer), bestehend aus Speicher, Steuer- und Recheneinheit sowie Peripherie mit Ein- und Ausgabeeinheiten und Software.

Den digitalen Daten entgegengesetzt werden *analoge Daten*. Diese sind "[...] durch einen kontinuierlichen Wert einer physikalischen Größe definiert."¹²⁹ Die Quecksilbersäule eines Thermometers oder der Uhrzeiger auf dem Zifferblatt einer Uhr sind Beispiele dafür. Sowohl analoge wie auch digitale Daten können in elektronischer und nicht-elektronischer Form festgehalten werden. Ein auf Papier ausgedruckter Binärcode ist zwar physisch greifbar, aber nur digital verwendbar. Umgekehrt ist ein analoges Videoband zwar ein elektronisches Speichermedium, aber von der Natur her analog.

Als *Metadaten* werden datenbeschreibende Informationen (Daten über Daten¹³⁰) bezeichnet.

*"Metadata is data or information that enables people to perform certain functions in relation to the information resources that the metadata is about. Metadata is information that is distinct from the resource which it is about, [...]"*¹³¹

*"[...] metadata is structured information that describes the attributes of information resources for the purposes of identification, discovery, selection, use, access, and management."*¹³²

¹²⁷ Für die Verwendung der Begriffe 'data, fact, evidence' in der englischen Sprache siehe: Rosenberg, Daniel: Data before the Fact, in: Gitelman, Lisa (Hg.): «Raw data» is an oxymoron, Cambridge (USA) 2013 (Infrastructures series), S. 15-40, hier S. 37.

¹²⁸ Folgend nach: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 27 f.

¹²⁹ Ebd., S. 28. Zur Unterscheidung zwischen analog und digital siehe auch: Böhnke, Alexander; Schröter, Jens (Hg.): Analog / Digital – Opposition oder Kontinuum? Zur Theorie und Geschichte einer Unterscheidung, Bielefeld 2004 (Medienumbrüche). URL: <http://dokumentix.uni-siegen.de/opus/volltexte/2009/381/pdf/AnalogDigital.pdf>, Stand: 19.04.2013.

¹³⁰ Nach Rosenthaler: "Unter Metadaten werden strukturierte Daten verstanden, welche eine Informationsquelle oder Ressource beschreiben und ihr Eigenschaften zuordnen, die aus der Informationsquelle selbst nicht direkt ersichtlich sind." Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 400.

¹³¹ Nach: Miller, Stephen J.: Metadata for Digital Collections: A How-to-do-it Manual, London 2011, S. 1..

¹³² Taylor, Arlene G.; Joudrey, Daniel N.: The organization of information, Westport (USA) 2009 (Library and information science text series), S. 89.

Digitale Dokumente, Bilder, Videos, Graphiken, Texte etc. in digitaler Form (Binärformat) werden in ihrer durch den Computer dargestellten Form als *digitale Objekte* bezeichnet. Digitale Objekte bestehen aus digitalen Daten und können nur mittels eines Ausgabegerätes wahrgenommen werden.

Der *Algorithmus* ist ein Problemlösungsverfahren, das in eine "endliche Folge von eindeutig bestimmten und tatsächlich durchführbaren Teilhandlungen"¹³³ unterteilt ist, ähnlich einem Kochrezept. Algorithmus steht als Begriff beispielhaft für alle Faktoren, die für einen Nutzer nicht ohne weiteres sichtbar sind und nur mit hochspezialisiertem Wissen nachvollzogen werden können.¹³⁴

Ein *Programm* ist ein Algorithmus, der in einer für eine Maschine verständlichen Folge von Anweisungen in einer spezifischen Syntax codiert ist. Algorithmen werden sequenziell abgearbeitet und können Selektionen oder Iterationen enthalten.¹³⁵ Zur Gliederung können Algorithmen in Methoden, darunter Module, Prozeduren, Routinen und Funktionen, zusammengefasst werden. Als Parameter wird ein Eingabewert bezeichnet, der an den Algorithmus für die Abarbeitung übergeben wird.

Ein *Informationssystem* besteht aus Menschen und/oder Maschinen, die Informationen erzeugen und/oder nutzen und die durch eine bestimmte Organisation vernetzt sind. Informationssysteme setzen die Bedingung, "was als Information zur Erscheinung kommt, als solche adressiert, gesucht, gefunden und verarbeitet werden kann."¹³⁶ Unter einem *digitalen (rechnergestützten, elektronischen) Informationssystem* wird die "Gesamtheit der Betriebssysteme, Applikationen, Datenbanken und Kommunikation ermöglichenden Ressourcen sowie der sie operativ haltenden technischen Einrichtungen, organisatorischen Regeln und betrieblichen Einheiten"¹³⁷ verstanden. Es bietet Zugriff auf digitale Datenbestände, die in digitaler Form vorgehalten werden und auf die mittels bestimmter Programme (Software) zugegriffen werden kann. In einem solchen System werden Daten eingegeben, gespeichert, gegliedert, (weiter-)verarbeitet, analysiert und dargestellt. Im Folgenden wird unter einem Informationssystem stets ein rechnergestütztes Informationssystem verstanden. Auf analoge resp. analog-digitale Informationssysteme wird explizit hingewiesen.

¹³³ Fischer; Hofer: Lexikon der Informatik, 2011, S. 35.

¹³⁴ Vgl.: Schweizerische Gesellschaft für Geschichte: Le pouvoir des algorithmes. Technologies numériques et rapports de force au XXe-XXIe siècles, Panel Detail, 4. Schweizerische Geschichtstage, 09.06.2016, URL: <https://www.geschichtstage.ch/panel/52/le-pouvoir-des-algorithmes-technologies-numeriques-et-rapports-de-force-au-xxe-xxie-siecles>, Stand: 10.04.2016.

¹³⁵ Selektion: 'Wenn Bedingung erfüllt, dann...'; Iteration: 'Wiederhole bis Bedingung erfüllt, dann...'

¹³⁶ Burkhardt: Digitale Datenbanken, 2015, S. 159.

¹³⁷ Fischer; Hofer: Lexikon der Informatik, 2011, S. 432.

Das *Internet* ist ein weltweites Netzwerk. Es verbindet alle angeschlossenen Computer, die das TCP/IP-Protokoll¹³⁸ verwenden. Das *World Wide Web* (kurz: *Web* oder *WWW*) ist ein Dienst des Internets für textbasierte und mit Hyperlinks verknüpfte Informationen, basierend auf der Nutzung von Web-Protokollen. Für die strukturierte Identifikation einer Ressource im Internet wird der *Unified Resource Identifier URI* verwendet, der das Auffinden dieser Ressource ermöglicht. Der *Unified Resource Locator URL* ist eine Untermenge des URI und beschreibt mittels einer festgelegten Syntax und Semantik die Adresse des Speicherorts einer Ressource im Internet.¹³⁹ Umgangssprachlich wird die URL als Web- oder Internetadresse bezeichnet.

Ein *Hypertext* ist eine aus mehreren modularen Objekten wie Text, Bild oder Multimediainhalten bestehende und über Elemente wechselseitig verknüpfte Informationseinheit, die von einem Hypertextsystem, bestehend aus einem Redaktions- oder Datenbanksystem, zur Verfügung gestellt und verwaltet wird. Die Literaturwissenschaftlerin Katherine Hayles schälte die grundlegenden Eigenschaften eines Hypertextobjekts präzise heraus:¹⁴⁰

*"[...] they are dynamic images; they include both analogue resemblance and digital coding; they are generated through fragmentation and recombination; they have depth and operate in three dimensions; they are written in code as well as natural language; they are mutable and transformable; they are spaces to navigate; they are written and read in distributed cognitive environments; and they initiate and demand cyborg reading practices."*¹⁴¹

¹³⁸ Siehe dazu bspw.: Herold; Lurz; Wohlrab: Grundlagen der Informatik, 2007, S. 402-411.

¹³⁹ Syntax URL: zugriffsdienst://[bereich.]domäne[:port]/verzeichnis(se)/dateiname.

¹⁴⁰ Hayles verwendet für die Textanalyse eine spezifische Methode: "Media-specific analysis (MSA) attends both to the specificity of the form [...] and to citations and imitations of one medium in another. Attuned not so much to similarity and difference as to simulation and instantiation, MSA moves from the language of "text" to a more precise vocabulary of screen and page, digital program and analogue interface, code and ink, mutable image and durably inscribed mark, texton and scripton, computer and book." Siehe: Hayles, Katherine N.: Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis, in: Poetics Today 25 (1), 2004, S. 69. DOI: 10.1215/03335372-25-1-67.

¹⁴¹ Ebd., S. 68. Siehe auch: Krameritsch: Geschichte(n) im Netzwerk. Hypertext und dessen Potenziale für die Produktion, Repräsentation und Rezeption der historischen Erzählung, 2007, S. 138-147. In weiten Teilen noch immer grundlegend: Kuhlen, Rainer: Hypertext: Ein nicht-lineares Medium zwischen Buch und Wissensbank, Berlin 1991 (Edition SEL-Stiftung). Der konzeptionelle 'Vater' des Hypertextes, allerdings ohne diesen Begriff zu verwenden, war Vannevar Bush mit dem hypothetischen Gedächtnissystem 'memex', das er im Juli 1945 im Artikel 'As We May Think' in 'The Atlantic Monthly' vorstellte (URL: <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>). Siehe auch: Bush, Vannevar: Wie wir denken sollten, in: Wirth, Uwe (Hg.): Kulturwissenschaft, Frankfurt a.M. 2008, S. 392-416. Den Begriff 'Hypertext' bestimmte Ted Nelson im 1981 publizierten Konzept zu 'Xanadu', einem global vernetzten System: Nelson, Theodor Holm: Literary machines 93.1: the report on, and of, project Xanadu concerning word processing, electronic publishing, hypertext, Sausalito 1993. Siehe auch: Koskimaa, Raine: From the Gutenberg Galaxy to the Internet Galaxy. Digital Textuality and the Change of the Cultural Landscape, in: Runnel, Pille; u.a. (Hg.): The digital turn – user's practices and cultural transformations, Frankfurt a.M. 2013, S. 271-283.

Allerdings ist der Hypertext weder als Buch, noch als Wissensbank, sondern als etwas dazwischen aufzufassen.¹⁴² Das Konzept des Hypertexts verortet nicht nur das Objekt unter vielen Objekten, sondern auch die Idee des Vielen auch im einzelnen Objekt.¹⁴³ Ein Hypertext schwankt zwischen Objekt und Objektsammlung.

Hypertexte sind unsequenziert, das heisst, sie haben keinen Anfang und kein Ende und können in beliebiger Reihenfolge gelesen werden. Zur Verknüpfung von Hypertexten resp. Objekten werden *Hyperlinks* verwendet.¹⁴⁴ Hypertext darf nicht mit E-Text gleichgesetzt werden, denn Letzterer ist ein digital vorliegender Text ohne hypertextuelle Bezüge zu anderen Objekten. Als Beispiel ist das E-Book zu nennen. Der Hyperlink ist für die Kohärenz einer Sammlung digitaler Objekte, insbesondere Webseiten, grundlegend ist, weil sonst keine direkten Bezüge unter den Objekten geschaffen werden können. Bei einem Buch sind die einzelnen Seiten physisch aneinandergebunden. Webseiten werden hingegen mittels Hyperlinks virtuell verknüpft werden. Nils Brügger definiert drei Arten von Hyperlinks, die gemeinsam, aber unterschiedlich verwendet werden:

*"The hyperlink is in other words both a semantic, formal and physically performative entity, all three elements of which define each other's existence and are thereby equally necessary, if in separate ways."*¹⁴⁵

2.1. Eigenschaften eines Digitalen Objekts

Digitale Objekte entstehen auf drei Arten, die als Digitaldatentypen bezeichnet werden. Diese Typen haben jedoch dieselben Eigenschaften, die es zu definieren gilt. Anschliessen ist zu untersuchen, ob und wie sie sich von bisher bekannten (analogen) Objekten abgrenzen.

2.1.1. Digitaldatentypen

Digitale Daten, die in ihrer dargestellten Form als digitales Objekt bezeichnet werden, können aufgrund ihrer Entstehungsart in drei Typen, den *Digitaldatentypen* unterschieden werden. Der erste Typ sind digitale Daten, die aus einem Prozess der Digitalisierung entstehen, in dem von einem analogen Objekt eine digitale Form, das *Digitalisat* erzeugt wird,

¹⁴² Dies drückt Rainer Kuhlen im Untertitel seiner Monographie aus: Ein nicht-lineares Medium zwischen Buch und Wissensbank. Siehe: Kuhlen: Hypertext, 1991.

¹⁴³ Vgl.: Burkhardt: Digitale Datenbanken, 2015, S. 105.

¹⁴⁴ Zu Hypertext und Hyperlink siehe bspw. auch: Schwabe, Astrid: Historisches Lernen im World Wide Web: Suchen, flanieren oder forschen? Fachdidaktisch-mediale Konzeption, praktische Umsetzung und empirische Evaluation der regionalhistorischen Website Vimu.info, Bd. 4, Göttingen 2012 (Beihefte zur Zeitschrift für Geschichtsdidaktik), S. 92-97. Schwabe unterscheidet inter- (innerhalb derselben Webseite), intra- (innerhalb desselben Webauftritts) und extrahypertextuelle Links (externe Webseite).

¹⁴⁵ Brügger: Website Analysis, 2010, S. 28.

beispielsweise das Digitalbild einer Urkunde. Während des Prozesses gehen verschiedene Informationen wie Haptik oder Geruch verloren und nur ein Teil davon kann mit Metadaten, die den digitalisierten Daten hinzugefügt werden, übertragen werden. Der zweite Typ sind digitale Daten, die in einem digitalen Arbeitsprozess, meist mit Interaktion des Menschen im Zusammenspiel von Hard- und Software *generiert* werden, beispielsweise eine im Personal Computer erstellte Datei. Solche Daten existieren ausschliesslich in digitaler Form und haben kein analoges Pendant, wie auch der dritte Typ. Dieser entsteht durch eine elektronische *Verarbeitung*, in der digital vorhandene Daten verändert werden und nicht mehr unbedingt identisch mit den unweigerlich verlorenen Ausgangsdaten sind. Diese 'neuen' Daten sind meist unvollständig und kontextlos – vor allem die dynamischen und verlinkten Elemente fehlen oder können nicht mehr aufgerufen werden. Wird beispielsweise eine Webseite als Bild auf dem PC gespeichert¹⁴⁶, so können keine Hyperlinks verwendet, automatische Bildfolgen angezeigt oder Texte markiert werden.

Diese drei Digitaldatentypen werden nach Niels Brügger als *digitized data*, *born-digital data* und *reborn-digital data* bezeichnet.¹⁴⁷ Allerdings meint Brügger in seinen Ausführungen, dass reborn-digital data nur während eines digitalen Archivierungsprozesses entstehen würden. Brügger forscht zur Langzeitarchivierung von Webseiten und nimmt aus diesem Blickwinkel diese unzutreffende Eingrenzung vor, denn reborn-digitale Daten entstehen bei jeder Migration¹⁴⁸ und bei jeder elektronischen Verarbeitung. Bei einer Migration unterstützen die neuen Datenträger- und/oder Dateiformate andere Eigenschaften, verarbeiten Daten anders und benötigen andere Software zur Bearbeitung oder Darstellung. Nur wenn ein Dateiformat gewählt wird, das den Funktionsumfang des bestehenden Objekts erweitert und keine bestehenden Eigenschaften verloren gehen lässt, wird von einer verlustlosen Migration gesprochen. Andererseits erzeugen andere Verarbeitungsprozesse aus digitalen Daten neue, aber nicht zwingend andersartige. Eine Datentabelle und die daraus computergenerierte Darstellung als Kuchendiagramm besagen inhaltlich zwar dasselbe. Das Diagramm ist aber als neues, reborn-digitales Objekt aufzufassen.

Die Digitaldatentypen sind nicht zwingend eindeutig, sondern können auch kombiniert vorkommen. Beispielsweise kann ein Digitalisat mit entsprechender Software beschriftet oder durch Zeichnungselemente wie einem Pfeil ergänzt werden. Unter Umständen können die Ergänzungen nicht wieder vom Digitalisat getrennt werden.

¹⁴⁶ Möglich durch Browser-Ad-Ons wie 'ScreenGrab!' unter Firefox (URL: <https://addons.mozilla.org/de/firefox/addon/screengrab-fix-version/>).

¹⁴⁷ Vgl.: Brügger: *When the Present Web is Later the Past*, 2012, S. 104.

¹⁴⁸ Migration (IT): Prozess der Umwandlung von Daten aus einem Format in ein anderes oder von einer Applikation aus einer Systemumgebung in eine andere. Nach: Fischer; Hofer: *Lexikon der Informatik*, 2011.

Alle drei Digitaldatentypen sind digitale Objekte mit denselben Eigenschaften und können als Quellen verwendet werden. Ein Umstand, dem oft – besonders von Geisteswissenschaftlern – wenig Beachtung geschenkt wird.¹⁴⁹

2.1.2. Definition der Eigenschaften

Ein digitales Objekt ist eine Entität (blob oder Datei)¹⁵⁰, bestehend aus einer geregelten Abfolge der Zeichen '0' und '1'¹⁵¹, unabhängig von seiner Darstellung. Die physische Repräsentation dagegen ist analoger Natur, die beispielsweise als elektrische Spannung, Stromfluss oder Magnetfeld auftritt. Eine Manipulation an einem digitalen Objekt erfolgt durch eine analoge Zustandsänderung einer Zeichenfolge, die durch ein Über- oder Unterschreiten eines Schwellenwertes ausgelöst wird.¹⁵² Bestehende Konventionen bestimmen, wie der Binärcode von der verarbeitenden Maschine gelesen werden muss. Jeder in einer höheren Programmiersprache geschriebener Quellcode¹⁵³ wird von einem Informationssystem in einen nur für Maschinen lesbaren Code übersetzt, den sogenannten Binär- oder Maschinencode. Dieser Code wird anschliessend verarbeitet und das Resultat wieder in einem für die verwendete Programmiersprache sinnvollen Format ausgegeben.

Digitale Objekte lassen sich nicht materialisieren. Beispielsweise stellt der Ausdruck eines E-Mails ein neues Objekt dar, mit ganz anderen Eigenschaften und kann nicht mit dem, meist immer noch vorhandenen, digitalen Objekt gleichgesetzt werden. Auch das Ausdrucken des Codes ist keine Fixierung, denn der Code kann nur im Zusammenspiel mit entsprechender Hard- und Software die vom Autor intendierte Darstellung oder Form annehmen oder Prozesse auslösen. Auf diesen Umstand weist das Titelbild dieser Arbeit schnippisch hin.

Der Digitalarchivar Jeff Rothenberg hält fest, dass viele digitale Artefakte inhärent digital sind und nicht vom Digitalen gelöst werden können:

¹⁴⁹ Diverse Beiträge von Geisteswissenschaftlern beschränken sich auf digitalisierte Primär- oder Sekundärquellen, ohne zu beachten, dass ein genuines oder reborn-digitales Objekt ebenfalls eine Quelle sein kann. Siehe bspw. Schmale, Wolfgang: Digital Humanities – Einleitung: Begriff, Definition, Probleme, in: Historische Mitteilungen. HMRG 26, 2013, S. 87.

¹⁵⁰ Blob: binary large object. Technische Bezeichnung für Datei (im Folgenden wird Datei verwendet). Siehe Gladney: Preserving Digital Information, 2007, S. 109.

¹⁵¹ Grundsätzlich wären auch andere Zeichen möglich, die Verwendung von 0 und 1 als verschriftlichte Beschreibung der zwei Zustände ist Standard.

¹⁵² Vertiefend: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 333 f. sowie Literatur zur Einführung in die Informatik.

¹⁵³ Der Quellcode ist in einer höheren, d.h. auf die Bedürfnisse des Menschen ausgerichteten Programmiersprache geschrieben (Java, C++, HTML, ...). Unter Verwendung von Compilern oder Interpretern wird aus dem Quellcode der Maschinen- oder Bytecode zur Verarbeitung oder Speicherung generiert (Binärcode). Code bezeichnet einen Quell- oder Binärcode.

*"They cannot be meaningfully represented as page images – Doing so loses essential aspects of their contents and/or behaviour."*¹⁵⁴

Dynamische, aktive oder interaktive Artefakte sind auch nicht in einer anderen Form darstellbar. Rothenberg führt dazu als Beispiele Multimedia, dynamisch generierte Inhalte wie JavaScript oder CGI¹⁵⁵, Präsentationen mit Animationen oder Simulationen, Interaktivität oder Digitale Kunst auf.¹⁵⁶ Deshalb ist ein digitales Objekt *ausschliesslich digital*.

Ein digitales Objekt liegt in verschiedenen Formen gleichzeitig vor, beispielsweise in der Darstellung als Bild und als Code. Dabei kann unter den einzelnen Informationsformen beliebig gewechselt werden, ohne dass sie sich gegenseitig beeinträchtigen. Jedoch stellt jede Form andere Informationen zur Verfügung: Präsentiert eine Webseite eine blaue Fläche, wirkt die Farbe je nach Betrachter heller oder dunkler, intensiver oder matter. Im Sourcecode der Webseite hingegen ist der genaue Wert des Blaus festgehalten, beispielsweise #0000b5 nach HTML Farbcode.¹⁵⁷ Ob das Blau exakt so dargestellt wird, hängt allerdings von der verwendeten Hardware ab, beispielsweise verfälscht ein Bildschirm mit eingeschränkter Farbwiedergabe oder mit einem Rot-Stich die Darstellung entsprechend.¹⁵⁸ Zusätzlich könnte im Sourcecode ein Kommentar des Programmierers vermerkt sein, der auf der Webseite nicht sichtbar ist, wie beispielsweise diese Codezeile:

```
<!-- Das Blau der Augen meiner Frau -->.
```

Diese Aussage ist als Metainformation des Objekts vorhanden, die aber nicht dargestellt wird. Trotzdem beschreiben beide Formen dasselbe Objekt. Des Weiteren kann ein digitales Objekt mehrere Objekte beinhalten, im Falle der genannten Webseite könnte die Fläche von einer anderen Webseite stammen, die in die aufgerufene integriert worden ist. Jedes Mal, wenn auf der externen Webseite die Farbe ändert, geschieht dies auch auf der vorliegenden. Im Sourcecode sind dafür ein Platzhalter und die entsprechende Quelle definiert. Zusätzlich könnte die vorliegende Webseite mit einem Hyperlink auf ein anderes Objekt verweisen und

¹⁵⁴ Rothenberg, Jeff: Digital Preservation in Perspective: How far have we come, and what's next?, Technologie, Future Perfect 2012: Digital Preservation by Design, Wellington 26.03.2012, Fol. 11. URL: http://de.slideshare.net/FuturePerfect_/jeff-rothenberg-digital-preservation-perspective#, Stand: 22.07.2014.

¹⁵⁵ CGI: Computer Generated Imagery (Computeranimation), bspw. 'Ice Age' (2002, Twentieth Century Fox Film Corporation (presents), Blue Sky Studios, Twentieth Century Fox Animation).

¹⁵⁶ Siehe: Rothenberg: Digital Preservation in Perspective, 2012, Fol. 11.

¹⁵⁷ Der HTML-Farbcode wird mit einem # eingeleitet, gefolgt von drei Paar Hexadezimalziffern von 00 bis FF für die Intensität der Farben Rot, Grün, Blau – die Mischung dieser Farben ergibt die dargestellte Farbe. Siehe zu HTML weiterführend: Freeman, Elisabeth: HTML mit CSS und XHTML von Kopf bis Fuss, Beijing 2006 (Ein Buch zum Mitmachen und Verstehen), hier S. 363-374.

¹⁵⁸ Zu eingeschränkter Farbwiedergabe siehe bspw.: Fischer, Martin: Bitte 10 Bit! Eingeschränkte Farbwiedergabe beim Mac Pro, in: c't: Magazin für Computertechnik (10), 2015, S. 36.

damit den Informationskontext erweitern. Dem Betrachter ist freigestellt, ob er die Informationen aufrufen will, ganz im Gegensatz zur Integration externer Quellen.¹⁵⁹

Eine weitere Dimension erhält ein digitales Objekt durch den Betrachter und seine individuelle Art, das Objekt zu lesen und zu verarbeiten. Er kann das Lesen unterbrechen und einem Link folgen, wieder zurückkehren oder nicht, das Objekt in eine andere Form oder in ein anderes Format bringen, indem er es beispielsweise auf Papier oder in verschiedenen Grössen dreidimensional druckt oder mit einem Skript automatisch anders darstellen lässt¹⁶⁰, oder beispielsweise als Wiki-Artikel weiterentwickelt. Er kann das Objekt oder Teile des Objekts extrahieren und lokal speichern, verändern und in neuem Kontext verwenden – ohne dass dies vom Autor intendiert worden wäre. Ein digitales Objekt ist durch die unterschiedlichen Formen, die es annehmen kann, die Verknüpfungsmöglichkeiten mit anderen Objekten und den zum Teil sehr unterschiedlichen Rezeptionen durch die Nutzer höchst flexibel und interaktiv. Es ist losgelöst von zeitlichen und räumlichen Dimensionen, es ist nicht auf den Gegenstand selbst beschränkt und dadurch als *multidimensional* zu bezeichnen.¹⁶¹

Eine weitere Eigenschaft digitaler Objekte hängt mit der Art und Weise der Informationsübermittlung zusammen, das heisst, wie ein digitales Objekt kommuniziert respektive wie es mit einem Nutzer durch die sogenannte *Human Computer Interaction* interagiert.¹⁶² Die Kommunikation geschieht dabei auf mehreren Sinnesebenen: Sehen, Fühlen, Hören und mittlerweile auch Riechen mittels Versprühen von Duftstoffen. Dabei ist allerdings das digitale Objekt auf das Zeichensystem seiner Programmiersprache beschränkt und interagiert daraus mit Befehlen an die angeschlossene Peripherie. Dabei kann eine mit entsprechender Hard- und Software auch vom Autor nicht intendierte Kommunikation stattfinden, indem beispielsweise ein geschriebener Text durch ein zusätzliches Programm über einen angeschlossenen Lautsprecher in Form des Vorlesens ausgegeben wird. Digitale Objekte sind folglich *multimodal*. Obwohl zu den Begriffen 'Mode' und 'Modalität' keine einheitliche Terminologie existiert¹⁶³, wird *Multimodalität* im Sinne einer *Interaktion mit mehreren unterschiedlichen Sinnen* verstanden.

Digitale Objekte können nicht fixiert werden. Unter Fixieren wird ein Vorgang verstanden, der ein Objekt in einen unveränderbaren, materiellen Zustand überführt, dessen Veränderung

¹⁵⁹ Dazu weiterführend in Kap. 5.3.2. Konsistenz.

¹⁶⁰ Skript: Anweisung für Computer (kurzes Programm). Bspw. kann mit entsprechenden Skripten eine Webseite auf die Bedürfnisse des Nutzers angepasst werden, ohne dass der Autor darauf Einfluss hat (bspw. mit Greasemonkey; URL: <https://addons.mozilla.org/de/firefox/addon/greasemonkey/>).

¹⁶¹ Siehe auch: Schreiber: *Genuine Internetdaten als historische Quelle*, 2012, S. 8-12.

¹⁶² Vertiefend bspw.: Reich, Siegfried (Hg.): *Human Computer Interaction*, Heidelberg 2013 (HMD - Praxis der Wirtschaftsinformatik).

¹⁶³ Siehe bspw.: Schneider, Jan Georg; Stöckl, Hartmut: *Medientheorien und Multimodalität: Zur Einführung*, in: Schneider, Jan Georg; Stöckl, Hartmut (Hg.): *Medientheorien und Multimodalität: ein TV-Werbespot – sieben methodische Beschreibungsansätze*, Köln 2011, S. 24-28.

stets nachvollzogen werden kann. Da Manipulationen eines digitalen Objekts durch elektronische Impulse ausgelöst werden, kann jedes Objekt mit geeigneten Mittel nicht nachvollziehbar verändert werden. Grundsätzlich besteht stets die Möglichkeit, eine Datei zu ändern oder zu löschen, sofern darauf zugegriffen werden kann. Zwar sind heutzutage revisionssichere Informationssysteme im Einsatz, die Zugriffe verhindern oder zumindest protokollieren, aber es gibt keines, das auf Dauer alle Änderungsversuche zu unterbinden vermag.¹⁶⁴ Mit fortschreitender Entwicklung der Systeme werden auch Wege entdeckt, wie in ein bestehendes System eingebrochen werden kann. Wenn eine Manipulation durch nicht nachvollziehbare, elektronische Zustandsveränderungen der Datenträger erfolgt, können die Veränderungen eben auch nicht nachvollzogen werden. Elektronisch protokollierte Zustandsänderungen können selbst wieder elektronisch manipuliert werden – die Nachvollziehbarkeit ist nicht gewährleistet. Mit Methoden der Digitalarchäologie oder Computer Forensik¹⁶⁵ können in vielen Fällen zwar Manipulationen festgestellt und/oder Daten nach einem Löschvorgang wiederhergestellt werden, aber ein Erfolg ist keineswegs garantiert. Ein Recovery¹⁶⁶ ist besonders dann aussichtslos, wenn eine darauf spezialisierte Person die Vorgänge durchgeführt hat. Diese Volatilität und Manipulierbarkeit eines digitalen Objekts werden von der Archiv- und Bibliothekswissenschaftlerin Abby Smith als sehr wichtige Merkmale bezeichnet:

"One of the most important qualities of information in digital form is that by its very nature it is not fixed in the way that texts printed on a paper are. Digital texts are neither final nor finite, and are fixed neither in essence nor in form except when a hard copy is printed out, for they can be changed easily and without trace of erasures or emendations."¹⁶⁷

Abby Smiths Aussage, dass durch einen Ausdruck das digitale Objekt fixiert werden kann, ist jedoch nicht korrekt, wie ich oben bereits ausgeführt habe. Selbst der Aussagegehalt eines Textes kann durch die Materialisierung verändert werden, indem beispielsweise mit Hyperlinks verknüpfte Elemente nun nicht mehr in die Aussage integriert und damit 'stumm' sind. Es kann festgehalten werden, dass digitale Objekte *volatil* und nicht nachvollziehbar *manipulierbar* sind.

Die physische Repräsentation eines Codes ist grundsätzlich unabhängig vom Datenträger, denn wenn ein Medium die eindeutige Unterscheidung von 0 und 1 ermöglicht, kann

¹⁶⁴ Zu revisionssicheren Informationssystemen siehe weiterführend S. 41.

¹⁶⁵ Siehe weiterführend bspw.: Kuhlee, Lorenz; Völzow, Victor: Computer-Forensik Hacks, Köln 2012 (Hacks series); Geschonneck, Alexander: Computer-Forensik: Computerstraftaten erkennen, ermitteln, aufklären, Heidelberg 2011; Graves, Michael W.: Digital archaeology: the art and science of digital forensics, London 2014.

¹⁶⁶ Recovery: Wiederherstellung oder Rekonstruktion eines vorherigen, konsistenten Zustands bei Daten und Parametern. Nach: Fischer; Hofer: Lexikon der Informatik, 2011.

¹⁶⁷ Smith, Abby: Why digitize?, Washington (D.C.) 1999, S. 3.

jeglicher Binärcode darauf gespeichert werden.¹⁶⁸ Ein Binärcode kann auch versendet werden, ohne dass ein physischer Datenträger bewegt werden muss, beispielsweise durch ein elektromagnetisches Signal, von einer Sendeantenne auf der Erde mehrere Stunden zur Raumsonde Voyager 1 der NASA unterwegs ist oder durch die kabellose Verbindung zwischen zwei Computern, bei der abhängig von der Dateigrösse nur wenige Mikrosekunden verstreichen.¹⁶⁹ Nur aus praktischen Gründen werden physische Datenträger verwendet, die möglichst viel Binärcode speichern können – Lukas Rosenthaler spricht von "transitorischer Materialität."¹⁷⁰ Der Begriff *immateriell* sollte nach dem Historiker William Turkel nicht verwendet werden. Er bevorzugt den Terminus *Transduction* im Sinne einer energetischen Konversion von der einen in eine andere Form.¹⁷¹ Zusammenfassend kann definiert werden, dass das Medium für den Binärcode irrelevant ist und jederzeit geändert werden kann, ohne dass das digitale Objekt selbst beeinträchtigt wird. Diese Eigenschaft digitaler Objekte wird (hier) als *Datenträgerunabhängigkeit* bezeichnet.

Informationssysteme können digitale Objekte theoretisch verlustfrei vervielfältigen. Tatsächlich können während des 'Kopierens' Fehler auftreten, die aber von den meisten Systemen durch geeignete Prüf- und Korrekturverfahren automatisch berichtigt werden.¹⁷² Ist die Überprüfung erfolgreich und konnte festgestellt werden, dass jedes einzelne Bit identisch ist, so "wurde eine hundertprozentige 1:1-Kopie erstellt"¹⁷³, wie Rosenthaler festhält und von *Klonen* spricht.¹⁷⁴ Da digitale Objekte auch keine Materialität besitzen, kann von *echten Klonen* gesprochen werden, im Gegensatz zu 1:1-Kopien, die sich wenigstens in ihrer Materialität unterscheiden. Zu beachten ist, dass dies nur dann gilt, wenn die Klone isoliert vorliegen und keine Informationen darüber vorhanden sind, in welchem Informationssystem sie sich befinden respektive befanden oder in welches System das Objekt geklont wurde. Anhand solcher Metadaten können die Klone eventuell unterschieden werden. Verlustfrei bedeutet, dass bei der Anfertigung einer Kopie die Vorlage in keiner Weise beeinträchtigt wird und dass im Idealfall das neu entstandene Objekt der Vorlage exakt gleicht. Selbst bei Vorgängen wie der Kompression eines digitalen Objekts kann eine verlustfrei Kopie der Vorlage erstellt werden, denn bei diesem Prozess werden nur Redundanzen eliminiert, der ursprüngliche Bitcode kann aber jederzeit vollständig und ununterscheidbar wiederhergestellt

¹⁶⁸ Siehe: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 380.

¹⁶⁹ Siehe: Ebd., S. 381. Jim Linder hat ein medienloses Archiv vorgeschlagen, in dem Daten ständig als elektromagnetische Signale unterwegs sind. Siehe dazu: Lindner, Jim: Toward a MediaLESS Archive, in: Aubert, Michelle; Billeaud, Richard (Hg.): Archiver et communiquer l'image et le son: les enjeux du 3ème millénaire: actes du Symposium technique mixte – JTS Paris 2000, Paris 2000.

¹⁷⁰ Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 40.

¹⁷¹ Turkel, William J.: Intervention. Hacking history, from analogue to digital and back again, in: De Groot, Jerome (Hg.): Public and popular history, London 2012, S. 142 f.

¹⁷² Siehe: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 333 f.

¹⁷³ Ebd., S. 376.

¹⁷⁴ Siehe: Ebd.

werden.¹⁷⁵ Zudem kann der Kopierprozess immer wieder vorgenommen werden und es wird immer wieder eine 1:1-Kopie erstellt. Kopierschutzverfahren wie Normverletzungen, Dongle, Verschlüsselung, versteckte Programme, Lizenzschlüssel oder Digital Rights Management (DRM)¹⁷⁶ können das Kopieren über eine gewisse Zeit verhindern. Allerdings besteht ein sportlicher Wettlauf zwischen Produzenten von Kopierschutzmechanismen und Kopierschutzknackern, der auch auf rechtlichem Wege ausgefochten wird.¹⁷⁷ Bis jetzt wurde kein Kopierschutzsystem entwickelt, das nicht irgendwann umgangen werden konnte.¹⁷⁸ Trotz Kopierschutzmechanismen kann festgehalten werden, dass digitale Objekte aus einer Vorlage heraus *verlustfrei vervielfältigt* werden können.

Die fehlende Materialität ermöglicht theoretisch, dass digitale Objekte unendlich oft verlustfrei wiederverwendet werden können. Grenzen setzt nur das Informationssystem, das aufgrund seiner Materialität nach einiger Zeit nicht mehr lauffähig ist, weil beispielsweise Komponenten physisch verschleissen. Dabei ist es unerheblich, ob es sich um ein darzustellendes, ein abzuspielendes oder eine Software als prozessuales Objekt handelt. Zudem ist es möglich, dass ein digitales Objekt gleichzeitig verwendet wird, beispielsweise ein Bild, das in verschiedenen Websites¹⁷⁹ erscheint. Digitale Objekte sind stets *verlustfrei wiederverwendbar*.

Ein Code wird zu einer Datei als Einheit zusammengefasst, umgekehrt entspricht aber nicht jede Datei einer Einheit – sie kann wiederum aus mehreren Dateien bestehen. Zudem ist der Dateiinhalt nicht auf eine physische Repräsentation beschränkt, sondern kann sich über mehrere Dateien erstrecken, die in verschiedenen Systemen liegen.¹⁸⁰ Die physische Repräsentation einer Datei wird auf einem Datenträger von einem Dateisystem organisiert. Ohne

¹⁷⁵ Zu verlustfreier und verlustbehafteter Kompression sowie Redundanz siehe: Ebd., S. 342-359.

¹⁷⁶ Siehe dazu: Ebd., S. 377-379.

¹⁷⁷ Bspw. das Verbot der Umgehung von Kopierschutzmechanismen, ungeachtet dessen, ob der Mechanismus umgangen werden kann. Für die Schweiz: Bundesversammlung der Schweizerischen Eidgenossenschaft: Bundesgesetz über das Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz), 01.01.2011. URL: <https://www.admin.ch/opc/de/classified-compilation/19920251/index.html>, Stand: 09.10.2015, Art. 39a. Für Deutschland: Bundesministerium der Justiz und für Verbraucherschutz: Gesetz über Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz), 31.08.2015. URL: <http://www.gesetze-im-internet.de/urhg/>, Stand: 09.10.2015, §95a. Für Österreich: Bundesministerium für Justiz: Bundesgesetz über das Urheberrecht an Werken der Literatur und der Kunst und über verwandte Schutzrechte (Urheberrechtsgesetz), 13.08.2015. URL: <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10001848>, Stand: 09.10.2015, §90c.

¹⁷⁸ Siehe dazu einschlägige Foren, teilweise im Darknet.

¹⁷⁹ Es wird zwischen Webauftritten (engl. website) und Webseiten (engl. webpage) unterschieden. Ein Webauftritt kann eine gesamte Domain und damit mehrere Webseiten umfassen. Die Homepage ist die Haupt- oder Einstiegsseite eines Webauftritts. Umgangssprachlich wird für Webauftritt meist 'Webseite' verwendet. Eine andere Begriffsunterscheidung verwendet bspw. Peter Haber: Web-Auftritt resp. synonym Website (website) und Web-Seite (webpage). Siehe: Haber: digital past (unv. Habil.), 2009, S. 71, Anm. S. 73.

¹⁸⁰ Siehe: Margulies, Simon B.: Digitale Daten als Quelle der Geschichtswissenschaft: Eine Einführung, Hamburg 2009 (Kölner Beiträge zu einer geisteswissenschaftlichen Fachinformatik 2), S. 18.

die vom Dateisystem in Metadaten festgehaltenen Informationen, wie und wo die einzelnen Teile einer Datei abgespeichert sind, ist eine Zusammenführung und Darstellung kaum mehr möglich.¹⁸¹ Die Inhalte der meisten Datenträger für digitale Objekte sind zudem für die menschlichen Sinne nicht zugänglich, denn es ist unmöglich, zu sehen oder zu fühlen, ob und wenn ja, welche Daten sich auf einer Festplatte befinden.¹⁸² Das Manipulieren oder Löschen einer Datei geschieht wiederum durch einen elektronischen Prozess, in dem bestimmte Zustandsänderungen vorgenommen oder der Umfang der zur Datei gehörenden Daten verkleinert oder vergrößert werden. Jedes Erstellen oder Ändern einer Datei erfolgt durch elektronische Impulse, die auf der Ebene der einzelnen Bits nicht nachvollziehbare Änderungen vornehmen.

Für die Verarbeitung und Speicherung von digitalen Objekten muss das Informationssystem die Bedeutung eines Codes kennen. Dazu werden Informationen über die physikalischen Eigenschaften des Datenträgers an sich (das physische Format), über die verwendeten Datenträgerformate (die Anordnung der Bits und Bytes auf dem Datenträger) und das Dateiformat (die interne Struktur einer Datei) benötigt und es müssen entsprechende Treiber vorhanden sein, um diese Informationen korrekt zu interpretieren.¹⁸³ Ohne ein solches, spezifisches Informationssystem ist es deshalb nicht möglich, eine Datei darzustellen oder zu manipulieren. Für die Darstellung wird die Datei anhand vorhandener Metainformationen wie der Dateiendung durch das Informationssystem verarbeitet respektive interpretiert und am Ausgabegerät dargestellt. Besteht die Datei aus einer Prozessanweisung für andere Maschinen oder Programme, so ist eine Darstellung am Bildschirm nicht zwingend nötig. Um ein digitales Objekt darstellen, ausführen oder manipulieren zu können, wird ein *spezifisches Informationssystem* benötigt, das aus Hard- und Software sowie peripheren Geräten besteht und zum Funktionieren Strom benötigt. Informationssysteme dürfen allerdings nicht als fehlerfreie Rechenmaschine angesehen werden. Sowohl Software als auch Hardware kann fehlerhaft sein und bei Berechnungen oder Prozessabläufen falsche Ergebnisse liefern.¹⁸⁴

Lukas Rosenthaler führt als zusätzliche Eigenschaft auf, dass digitale Objekte "ohne Verlust, mit Lichtgeschwindigkeit quasi instantan, ohne den physischen Transport von Materie,

¹⁸¹ Mithilfe der Datenforensik können unter Umständen die verstreuten Dateibestandteile wieder zusammengeführt werden.

¹⁸² Nach: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 417.

¹⁸³ Treiber: Steuerungssoftware zur Ansteuerung von Komponenten eines elektronischen Informationssystems. Siehe dazu: Ebd., S. 381-399; Prevezanos: Computer Lexikon 2013, 2012, S. 277.

¹⁸⁴ Beispiel für Fehler in der Hardware: Als FDIV-Bug wird der vom Mathematiker Thomas Nicely 1994 nachgewiesene Fehler bezeichnet, der dem Pentium Prozessor von Intel beim Dividieren unterlief. Siehe: Stiller, Andreas: 20 Jahre FDIV-Bug: Ein Prozessor-Rechenfehler macht Geschichte, Newsticker, heise online, 30.10.2014, URL: <http://www.heise.de/newsticker/meldung/20-Jahre-FDIV-Bug-Ein-Prozessor-Rechenfehler-macht-Geschichte-2438283.html>, Stand: 18.11.2014.

beliebig zwischen zwei Punkten im Raum transportiert werden"¹⁸⁵ können und sich dadurch Raum und Zeit aufheben würden. Zu entgegnen ist, dass jegliche Datenübertragung Zeit benötigt und dieses physikalische Gesetz nicht aufgehoben wird, auch wenn der Mensch dies nicht nachvollziehen kann. Der Raum kann hingegen als aufgehoben erklärt werden, weil ein digitales Objekt keinen Datenträger und damit keine Materie und keinen Platz beansprucht.

Nach Catherina Schreiber ist eine digitale Quelle (born digital source) genuin digital, multi-modal und -dimensional, veränderbar, prozessorientiert sowie zeitlich und räumlich unabhängiger als ein physisches Objekt. Zudem kann sie für einen grösseren Umfang an quantitativen und qualitativen Analysen verwendet werden.¹⁸⁶ Dieser Definitionsansatz bestimmt wichtige Merkmale eines digitalen Objekts: das genuin Digitale, das Multimodale und -dimensionale und die Manipulierbarkeit. Hingegen ist jegliches Objekt das Ergebnis eines Prozesses, wie lange er auch dauerte oder zurücklag. Die grössere Unabhängigkeit von Raum, im Sinne von Nähe des Autors zum Geschehen, ist zwar in der Quellenkritik zu berücksichtigen, ist selbst aber keine Eigenschaft eines digitalen Objekts. Auch die zeitliche Unabhängigkeit ist kein spezifisches Merkmal eines digitalen Objekts. Das von Schreiber genannte Beispiel eines Chatbeitrages ist vergleichbar mit einem physischen Brief in einem regen Briefwechsel über Kontinente hinweg.¹⁸⁷ Chats können zwischen zwei Parteien (privat), halböffentlichen oder halbprivaten Gruppen oder öffentlich entstehen und die Einträge können sehr schnell aufeinander folgen. Auch ein Briefwechsel kann privat, halböffentlich oder halbprivat oder öffentlich erfolgen. Die Partizipierenden beider Kommunikationsformen können ebenso nah oder fern vom Geschehen sein. Eine digitale Kommunikation kann schneller sein – muss aber nicht, wie Schreiber selbst feststellte.

Der Philosoph John Haugeland weist digitalen Objekten drei Eigenschaften zu: fehler- respektive verlustfreie Kopierbarkeit, Komplexität, wobei noch so komplizierte Netzwerke aus einfachsten Standardkomponenten wie 0 und 1 abgebildet werden können, und Medienunabhängigkeit, wobei in unterschiedlichen Medien genau gleichwertige Strukturen vorkommen können.¹⁸⁸ Diese Eigenschaften wurden in der obigen Definition bereits aufgegriffen, Haugeland lässt aber einige Eigenschaften ausser Acht oder impliziert sie in seiner Aufzählung, ohne es darzulegen. Für Peter Haber umfasst Digitalität "ein Bündel von Merkmalen, die in ihrer Summe dazu führen, dass der Umgang mit Information flexibel, punktgenau und immer wieder neu adaptierbar, gleichzeitig aber auch abstrakt und nur durch die Vermittlung von

¹⁸⁵ Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 381.

¹⁸⁶ Nach: Schreiber: Genuine Internetdaten als historische Quelle, 2012.

¹⁸⁷ Siehe: Ebd., S. 11.

¹⁸⁸ Nach: Haugeland, John: analog und analog, in: Böhnke, Alexander; Schröter, Jens (Hg.): Analog/Digital – Opposition oder Kontinuum? Zur Theorie und Geschichte einer Unterscheidung, Bielefeld 2004, S. 34.

Maschinen rezipierbar wird."¹⁸⁹ Die Merkmale wurden oben als Eigenschaften aufgeführt, darunter auch die Abhängigkeit von einem digitalen Informationssystem. Mit flexibel, punktgenau und neu adaptierbar spricht Haber jedoch nicht Merkmale eines digitalen Objektes an, sondern diejenigen von Digitalmedien, die von Objekten klar zu trennen sind.¹⁹⁰

Mit diesen Ausführungen können die Eigenschaften eines digitalen Objekts definiert und der erste Teil der ersten Frage beantwortet werden:

Ein digitales Objekt ist genuin digital, multimodal und -dimensional, nicht nachvollziehbar manipulierbar, volatil, datenträgerunabhängig, verlustfrei vielfältig- und (gleichzeitig) wiederverwendbar und von einem spezifischen, digitalen Informationssystem abhängig.

2.1.3. Abgrenzung

Ich gehe davon aus, dass sich unter den oben genannten Eigenschaften eines digitalen Objekts mindestens eines befindet, das die Abgrenzung des digitalen Objekts von anderen, bisher bekannten Objekten rechtfertigt. Damit wären digitale Objekte als neue Quellenart zu betrachten. Zur Verifikation dieser Behauptung wird überprüft, ob es sich bei den definierten Eigenschaften eines digitalen Objekts um Alleinstellungsmerkmale handelt oder nicht.

Ein digitales Objekt entsteht und besteht ausschliesslich in einem Informationssystem. Bei einer Transformation von analog zu digital oder digital zu analog wird jeweils ein neues Objekt erzeugt, das andere Eigenschaften aufweist und deshalb nicht gleichwertig behandelt werden kann. Zum Beispiel kann in einem Ausdruck einer Webseite kein Hyperlink angeklickt werden. Die Unterscheidung zwischen analog und digital ist noch eindeutig.¹⁹¹

Sowohl digitale, als auch physische Objekte sind multimodal. Ein Buch ist nicht nur durch Sehen und Lesen erfahrbar, sondern auch durch Fühlen und Riechen. Dadurch, dass die sprachliche Kommunikation nur selten isoliert vorkommt, kann auch hier von Multimodalität gesprochen werden. Deshalb ist diese Eigenschaft nicht exklusiv für digitale Objekte zu beanspruchen.

Multidimensionale Objekte finden sich auch in bisher Bekanntem, beispielsweise in wissenschaftlichen Texten mit Verweisen auf weiterführende Literatur. Auch handwerklich gefertigte

¹⁸⁹ Haber: digital past (unv. Habil.), 2009, S. 81.

¹⁹⁰ Weitere Ausführungen zu Digitalmedien in Kap. 4. Digitalmedien.

¹⁹¹ William G. Thomas III hatte in einem Blogbeitrag interessante Ausblicke für Digital History beschrieben, u.a. die zukünftig vollständige Verschmelzung von digitaler und physischer Welt in einem Menschen als eine Art Cyborg. Die Verflechtung wird sicherlich weiter zunehmen, aber ob sich ein solches Stadium erreichen lässt, ist spekulativ. Hingegen ist klar, dass die Unterscheidung von digital und analog weiterhin eminent wichtig ist. Siehe: Thomas, William G., III: The Future of Digital History, #rrchnm20, William G. Thomas III, 16.11.2014, URL: <http://railroads.unl.edu/blog/?p=1146>, Stand: 19.11.2014.

Utensilien können in verschiedenen Dimensionen untersucht werden. Beispielsweise können liturgische Geräte von ihrer Beschaffenheit und vom Material von Interesse sein, zum Beispiel wie weit die Handwerkskunst entwickelt ist, und/oder ob und wie sie als Bestandteil eines liturgischen Prozesses eingesetzt wurden (Funktion). Diese Eigenschaft ist demnach nicht ausschliesslich einem digitalen Objekt zuzuschreiben.

Volatilität besteht sowohl bei digitalen als auch physischen Objekten. Zwar mag das Löschen eines digitalen Objekts viel schneller und einfacher erfolgen als das Verbrennen eines Buchs, aber beide Objekte können nicht mehr in der intendierten Art genutzt werden. Bei Manipulationen an physischen Objekten können diese meist durch forensische Verfahren nachgewiesen werden. Beispielsweise hinterlässt das Überschreiben eines Worts in einem Dokument zwangsläufig Spuren, denn das Material wird beschädigt. Bei digitalen Objekten hingegen ist es meist nicht möglich, die Manipulation nachzuweisen, weil sie nicht auf Materialität angewiesen sind. Das Überschreiben eines Worts bedeutet, dass der Binärcode der Datei verändert wird, was aber nur im Vergleich mit dem ursprünglichen Code ersichtlich ist. Physisch rekonstruierbare Spuren sind nicht vorhanden und auch die Digitalforensik kann wenige Anhaltspunkte liefern, wenn nur diese eine Datei untersucht werden kann.¹⁹² Die Entwicklungen bei forensischen Untersuchungen von Bildmanipulationen ist mittlerweile schon weit fortgeschritten, aber dennoch können solche nicht in jedem Fall festgestellt werden.¹⁹³

Eine bessere Sicherheit vor Löschung oder Manipulation von Daten bieten *revisionssichere* Informationssysteme, die beispielsweise von der Steuerverwaltung für die elektronische Geschäftsführung einer Unternehmung verlangt werden. Solche Systeme basieren auf einem ausgeklügelten System von Zugangsrechten und Protokollmassnahmen. Zum einen wird genau geregelt, welcher Nutzer auf welche Objekte zugreifen und welche Eingaben respektive Manipulationen vornehmen darf. Es ist auch festgehalten, welcher Systemadministrator welche Rechte vergeben darf und wie im System ein Systemadministrator eröffnet wird. Zum anderen werden alle Vorgänge versioniert und manipulationssicher protokolliert, sodass diese stets nachvollziehbar sind. Aber auch solche Systeme sind angreifbar, wenn es

¹⁹² Eine andere Ausgangslage bietet sich für die Digitalforensik, wenn auf das Informationssystem zugegriffen werden kann, das die Datei erstellt/manipuliert hat, denn evtl. können in Protokolldateien Hinweise gefunden werden. Weiterführend bspw.: Graves: Digital archaeology, 2014, Kap. 9 (Document Analysis). Graves zeigt unter Verwendung des Programms Directory Snoop (URL: <https://www.briggsoft.com/dsnoop.htm>) und dem Dateieigenschaftendialog im MS Windows Datei Explorer, welche Informationen relativ einfach ausgelesen werden können.

¹⁹³ Zur Bildforensik siehe bspw.: Krawetz, Neal: A picture's worth: Digital image analysis, in: Black Hat DC, Washington (USA) 2008. URL: <https://www.blackhat.com/html/bh-dc-08/bh-dc-08-archives.html>, Stand: 17.06.2015. Siehe auch Diskussionen um Bildmanipulationen im Zusammenhang mit dem Flugzeugabsturz von MH17 am 21.07.2014 in der Ukraine, die v.a. in den Publikumsmedien reise-risch geführt wurden: Bidder, Benjamin: Interview zu angeblichen MH17-Manipulationen: «Bellingcat betreibt Kaffeesatzleserei», in: Spiegel Online, 03.06.2015. URL: <http://www.spiegel.de/politik/ausland/mh17-satellitenbilder-bellingcat-betreibt-kaffeesatzleserei-a-1036874.html>, Stand: 01.07.2015.

beispielsweise gelingt, auf den Quellcode der Software zuzugreifen. Mit entsprechend abgeändertem Code könnte der Angreifer das oben genannte System deaktivieren und nicht nachvollziehbare Manipulationen vornehmen. Dennoch akzeptiert der Gesetzgeber Daten aus solchen Systemen als Beweismittel, weil durch verschiedene Massnahmen wie Prüfung durch externe Experten, die Wahrscheinlichkeit einer Manipulation sehr klein ist.

Datenträgerunabhängige Objekte wie die menschliche Sprache sind ebenfalls bereits bekannt. In der Oral History wurden Erfahrungen gesammelt, wie mit sprachlichen Quellen umzugehen ist. Das Ergänzen mündlicher Aussagen mit schriftlichen Kommentaren zu Phonetik, Umfeld oder Verfassung des Sprechenden hilft, die Aussagen besser zu verorten. Mit dem Einsatz von Audio- und (audio-)visuellen Aufzeichnungen können noch bessere Ergebnisse erzielt werden, weil sich damit Informationen festhalten lassen, die kaum protokolliert werden können. Für die Oral History sind auch die nichtsprachlichen Kommunikationsebenen wie Mimik oder Gestik von Interesse.

Von verlustfreier und fehlerfreier Kopierbarkeit und Verwendbarkeit kann nur bei einem digitalen Objekt gesprochen werden. Aufgrund der Datenträgerunabhängigkeit spielt die physische Materialität bei digitalen Objekten keine Rolle. Physische Objekte unterscheiden sich hingegen nur schon im Material von einer Kopie, unerheblich wie exakt die Reproduktion vonstattenging. Beispielsweise können zwei Schrauben zwar genau gleich gefertigt sein, sie unterscheiden sich jedoch immer im Material, denn jede Schraube benötigt ihr eigenes Material. Zudem kann eine Schraube nicht gleichzeitig an zwei verschiedenen Stellen eingesetzt werden. Sprachliche Kommunikation bedarf zwar auch keiner Materialität, kann aber, im Gegensatz zu digitalen Objekten, weder gleichzeitig an mehreren Orten, noch genau gleich reproduziert werden. Ein Beispiel: Jedes von einer Person gesprochene oder gerufene 'Au!' unterscheidet sich vom vorherigen. Wird das 'Au!' hingegen als digitales Audio-Objekt abgespielt, so bleibt es sich immer gleich.

Nur digitale Objekte benötigen zwingend ein digitales Informationssystem zu ihrer Darstellung oder Verarbeitung. Sowohl physische Objekte als auch verbale/nonverbale Kommunikation respektive Darstellungen sind nicht auf ein Informationssystem angewiesen. Zwitter-systeme wie das Verschlüsselungssystem Enigma funktionieren in einer Kombination von mechanischen und elektronischen Verfahren. Zur Dechiffrierung eines verschlüsselten Texts ist dieses System zwar nötig, aber sowohl das Ursprungsobjekt wie das Endprodukt als Buchstaben-/Zahlenkombination sind physisch und verbleiben nicht als eigenständiges Objekt im System.

Mit diesen Ausführungen wurde aufgezeigt, dass sich digitale Objekte von bisher bekannten Objekten unterscheiden, womit der zweite Teil der ersten Frage beantwortet ist:

Digitale Objekte unterscheiden sich von bisher bekannten Objekten durch die ausschliessliche Digitalität, die verlustfreie und fehlerfreie Vervielfältig- und Wiederverwendbarkeit, die nicht nachvollziehbare Manipulation sowie dadurch, dass sie zwingend auf ein spezifisches, digitales Informationssystem angewiesen sind.

2.2. Neue Quellentypen

In den obigen Ausführungen ist festgestellt worden, dass sich ein digitales Objekt von bisher bekannten Objekten abgrenzt und dass es eine neue Quellenart darstellt. Aus dieser Erkenntnis folgere ich, dass digitale Objekte, wenn sie als Quelle verwendet werden, neue Quellentypen darstellen. Der erste, neue Quellentypus basiert auf der bisher unbekanntem Möglichkeit externe Inhalte in ein Objekt nicht erkennbar einzubinden. Die bisher nicht mögliche, verlustfreie Vervielfältigbarkeit stellt einen zweiten und die nicht sichtbaren Ebenen eines digitalen Objekts einen dritten Quellentypus dar.

2.2.1. Integrierender Quellentypus

Eine Webseite wirkt mit ausschliesslich statischen Elementen oberflächlich gesehen als ein geschlossenes Ganzes, vergleichbar mit einer Zeitungsseite. Statische wie dynamisch strukturierte Webseiten¹⁹⁴ werden jedoch durch ein Zusammenspiel von mehreren Dateien dargestellt: ein Bild liegt als Bilddatei vor und wird durch einen Befehl im Quellcode der Webseite an der entsprechenden Position in definierter Grösse dargestellt. Die Bilddatei kann im Dateisystem der Webseite gespeichert sein oder auch mittels eines Hyperlinks von einem anderen Dateisystem stammen, das nicht unter der Kontrolle des Webseitenadministrators ist. Auch ein Brief, geschrieben in einem Textverarbeitungsprogramm¹⁹⁵, kann externe Elemente enthalten, die beispielsweise aus Gründen der schnelleren Darstellung nicht in das Dokument selbst integriert wurden. Dabei bleibt das integrierte Objekt an seinem bestehenden Speicherort intakt und ist gleichzeitig wieder- und mehrfachverwendbar. Der Betrachter erhält weder über den externen Speicherort noch über den Urheber Informationen, solche sind nur im Quellcode ersichtlich.

Ein in genannter Konzeption gleichwertiger Quellentypus ist bisher nicht bekannt. Ein physisches Objekt kann sich kein weiteres Objekt so zu Eigen machen, dass es einerseits an seinem bisherigen Standort verbleiben kann und gleichzeitig auch in ein anderes Objekt integriert werden kann. Zum Beispiel ist zwar ein Puzzleteil in ein Gesamtbild integriert, es muss

¹⁹⁴ Statisch: Auf jeder Seite eines Webauftritts werden die einzelnen Elemente (Text, Bild,...) fest verankert (analog einer Zeitungsseite). Dynamisch: Der gesamte Inhalt eines Webauftritts wird aus Datenbanken generiert.

¹⁹⁵ Bspw. Apache OpenOffice Writer oder Microsoft Word.

aber physisch eingefügt werden und steht dann nicht gleichzeitig anderen Bildern zur Verfügung. Das im untenstehenden Blog-Beitrag¹⁹⁶ dargestellte Bild der Schreibmaschine kann direkt von der Webseite von Pantograph in den Beitrag integriert werden, ohne dass es vom Autor in der Arbeitsumgebung verwaltet, das heisst gespeichert und eingebunden werden muss.

Ein Bild sagt mehr...

© 13. Juli 2015 Off Topic, Problemfelder, Publikation digitale Quelle, Hypertext, physisch, Wissenschaftliches Arbeiten

Kürzlich bin ich auf ein Bild gestossen, das sehr gut visualisiert, in welchem Dilemma die (Geschichts-)Wissenschaft an vielen Universitäten sich befindet:



Zwar wollen wir die digitalen (0 und 1) Möglichkeiten nutzen, sind aber oft an bestehende Restriktionen und Vorgaben gebunden, welche auf analogen Publikationsformen (Papierblatt) beruhen.

Diese Problematik wurde auf verschiedenen Veranstaltungen und in diversen Beiträgen diskutiert, weshalb auf weitere Ausführungen verzichtet wird. Aber das Bild musste hier publiziert werden, weil es das Thema pointiert darstellt.

Bild: PANTOGRAPH

Abbildung 1: Bild einer Schreibmaschine (Pantograph), in eine Webseite integriert (Föhr)

Wenn ein Objekt andere Objekte zu seiner Darstellung und Informationsvermittlung internalisieren kann, ohne dass dabei die externen Objekte lokal vorhanden sein müssen und damit auch nicht deren Weiterverwendung beeinträchtigt wird, kann von einem *Externalressource internalisierenden Quellentypus* oder vereinfacht von einem *integrierenden Quellentypus* gesprochen werden.

¹⁹⁶ URL: <http://hsc.hypotheses.org/359>. Bildlizenz durch PANTOGRAPH (wie Coverbild) gemäss persönlicher Korrespondenz vom 10.07.2015.

2.2.2. Einobjektreduzierender Quellentypus

Die oben erwähnte gleichzeitige Wieder- und Mehrfachverwendbarkeit eines digitalen Objekts bringt unter dem Aspekt, dass dieses verlustlos und fehlerfrei kopierbar ist, einen weiteren Quellentypus hervor. Da ein echter Klon, eine 1:1-Kopie, eines digitalen Objekts sich von diesem nicht unterscheidet, können alle vorhandenen Klone unter Verwendung von Hyperlinks auch wieder ohne Informationsverlust auf ein einzelnes Objekt reduziert werden. Dies ist bei physischen Objekten, die sich durch die verwendete Materialität grundsätzlich unterscheiden, nur unter der Annahme möglich, dass Kopien mit vernachlässigbaren Unterschieden geschaffen wurden. Bei digitalen Objekten hingegen ist ein Klon stets perfekt.¹⁹⁷ Weil Klone ohne Informationsverlust auf ein einzelnes Objekt reduziert respektive alle Klon-Redundanzen eliminiert werden können, kann von einem *einobjektreduzierenden Quellentypus* gesprochen werden.

2.2.3. Quellentypus der unsichtbar codierten Information

Verschiedene Techniken erlauben es, einem Objekt zusätzliche Informationen anzufügen. Beispielsweise kann ein Buch mit handschriftlichen Annotationen auf der Seite oder einer Haftnotiz versehen werden. Die Annotation wird zum sichtbaren Bestandteil des Buchs und ist mit diesem verbunden. Ähnliche Techniken können auch bei digitalen Objekten angewendet werden, beispielsweise mit der Kommentarfunktion eines PDF-Readers. Zusätzlich aber können einem digitalen Objekt Informationen angefügt werden, die nicht sofort ersichtlich sind, sondern erst nach einer bestimmten Aktion durch den Nutzer. Die Aktionen sind zwar im Quellcode ersichtlich, in der dargestellten Form werden sie aber erst dann aufgerufen, wenn die definierte Handlung durch den Nutzer erfolgt ist. Wird beispielsweise der Mauszeiger über das im Code definierte Objekt geführt (`mouseover`¹⁹⁸), können zusätzliche Informationen, Bilder etc. als Elemente in einer das ursprüngliche Objekt überlagernden Ebene erscheinen. Alternativ kann die Nutzeraktion auch direkt eine weitere Aktion auslösen, beispielsweise eine andere Webseite aufrufen, ohne dass der Nutzer aktiv hätte durch Klicken bestätigen müssen. Die Informationen sind dadurch, dass sie im Code enthalten sind, fester Bestandteil des Objekts, werden aber nicht oder nur nach bestimmten Aktionen dargestellt.

Folgende Codes in HTML, CSS und JavaScript lösen (gemeinsam) beim Überfahren einer Fläche mit dem Mauszeiger eine sichtbare Änderung aus:¹⁹⁹

¹⁹⁷ Siehe: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 376.

¹⁹⁸ Häufig auf Webseiten genutzt. Erklärung siehe bspw. unter URL: <http://wiki.selfhtml.org/wiki/JavaScript/Event-Handler/mouseover/> oder URL: http://www.w3schools.com/cssref/sel_hover.asp.

¹⁹⁹ Beispiel von SELFHTML, siehe URL: <https://wiki.selfhtml.org/wiki/JavaScript/Event/mouseover>, Stand: 19.10.2016. Verwendung gemäss Lizenz Creative Commons Namensnennung – Weitergabe

HTML

```
<!DOCTYPE html>
<html>
<head>
  <meta charset="utf-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0;" />
  <link rel="stylesheet" type="text/css" media="screen" href="/Beispiel:SELFHTML-Beispiel-
Grundlayout.css" />
  <style>
h1 span {
  margin-left: 4.5em;
}
</style>
  <title>Beispiel: mouseover Event-Handler</title>
</head>

<body>
  <h1>Beispiel: mouseover Event-Handler</h1>

  <main>
    <p>Bewegen Sie die Maus auf den Button und wieder weg,
      um den Event-Handler in Aktion zu sehen:</p>
    <button id="Eingabe" type="button">Klick mich!</button>
    <p id="Ausgabe"></p>
  </main>

  <script>
var elem = document.getElementById('Eingabe');
var ausgabe = document.getElementById('Ausgabe');
elem.addEventListener('mouseover', mouseOver);
elem.addEventListener('mouseout', mouseOut);

function mouseOver() {
  ausgabe.innerHTML = 'Ich bin dynamisch!';
  elem.innerHTML = 'Drüber!';
}

function mouseOut() {
  ausgabe.innerHTML = ' ';
  elem.innerHTML = 'Wieder weg!';
}
</script>
</body>
</html>
```

CSS²⁰⁰

```
h1 span {
  margin-left: 4.5em;
}
```

unter gleichen Bedingungen 3.0 Deutschland (CC BY-SA 3.0 DE), einsehbar unter <https://creativecommons.org/licenses/by-sa/3.0/de/>.

²⁰⁰ CSS: Cascading Style Sheets legen die Gestaltung eines HTML-Tags fest. Diese kann innerhalb einer HTML-Datei festgelegt werden oder in einer separaten Datei, die für alle HTML-Tags eine zentrale Verwaltung bieten kann.

JavaScript

```
var elem = document.getElementById('Eingabe');
var ausgabe = document.getElementById('Ausgabe');
elem.addEventListener('mouseover', mouseOver);
elem.addEventListener('mouseout', mouseOut);

function mouseOver() {
    ausgabe.innerHTML = 'Ich bin dynamisch!';
    elem.innerHTML = 'Drüber!';
}

function mouseOut() {
    ausgabe.innerHTML = ' ';
    elem.innerHTML = 'Wieder weg!';
}
```



Abbildung 2: Mouseover-Effekt unsichtbar (Quelle: SELFHTML)



Abbildung 3: Mouseover-Effekt sichtbar (Quelle: SELFHTML)

Es sind nicht immer alle Elemente eines Objekts sichtbar, unmittelbar zugänglich oder bekannt.²⁰¹ Manche Elemente können als unabhängige Paratexte²⁰² verstanden werden, andere sind vom Autor als fester, vom Objekt nicht zu trennender Objektbestandteil definiert. Niels Brügger konstatiert auch, dass eine Unterscheidung zwischen eigentlichen und

²⁰¹ Vgl.: Brügger: Website Analysis, 2010, S. 26. Brügger bezieht sich zwar hauptsächlich auf Textelemente einer Webseite, aber die Aussage gilt für digitale Objekte im Allgemeinen.

²⁰² Paratext: Ein textbegleitender, d.h. einen den Haupttext ergänzenden oder kommentierenden Text oder Textteil (Vor- oder Nachwort, Widmung, Motto, Titel, Anmerkung, Illustration etc.). Nach: Wrana, Daniel: DiskursNetz: Wörterbuch der interdisziplinären Diskursforschung, Berlin 2014 (Suhrkamp Taschenbuch Wissenschaft), S. 290. Siehe auch: Brügger: Website Analysis, 2010, S. 38.

paratextuellen Inhalten schwierig sein kann, weil diese ineinander greifen oder sogar beide Funktionen übernehmen können.²⁰³

Eine besondere und sehr verbreitete Form von Objekten mit unsichtbar codierten Informationen sind digitale *Datenbanken*. Eine Datenbank ist eine auf einem Modell basierende Datenorganisation, bestehend aus Metadaten, der Datenbasis und einem Datenbankmanagementsystem (DBMS).²⁰⁴ Die Datenbasis ist eine Kollektion verschiedener einzelner Elemente, den Entitäten, die alle die genau gleiche Signifikanz aufweisen und nicht in einer Reihenfolge stehen.²⁰⁵ Das letzte Merkmal unterscheidet digitale von physischen Datenbanken wie Register oder Karteikartensammlungen, die immer in einer bestimmten Reihenfolge geordnet sind, ansonsten die Informationen nicht auffindbar wären. Das DBMS ermöglicht neuartigen Umgang mit Daten wie beispielsweise den schnellen Zugriff auf Daten und Datensammlungen oder das Sortieren, Umsortieren und Verknüpfen von Daten innerhalb der Möglichkeiten des Datenbankmodells. Es ist für die Speicherung, Auswahl und Sichtbarmachung der einzelnen Elemente zuständig, die sonst verborgen bleiben. Mittels einer Suchanfrage, dem Query, werden Suchbedingungen formuliert, die nach bestimmten Regeln automatisch mit der Datenbasis abgeglichen werden und die Ergebnisse werden vom System präsentiert. Das Ergebnis hängt von den Kompetenzen des Nutzers im Umgang mit der vom Datenbankmodell zur Verfügung gestellten Suchtechnologie ab. Ob das Ergebnis korrekt ist, kann selbst ein erfahrener und kompetenter Nutzer meistens nicht beurteilen, weil er weder die Algorithmen des DBMS, noch den Datenbestand und damit die potenziell auffindbaren Informationen kennt. Damit kann die Datenbank als eine Art von Black Box bezeichnet werden, die dem Quellentypus der unsichtbar codierten Information zugeordnet werden kann.

Im Analogen bestehen keine bisher bekannten Objekte, die solche Eigenschaften aufweisen, denn Annotationen sind immer mit dem Objekt in irgendeiner Art und Weise sichtbar fest verbunden.²⁰⁶ Deshalb kann der *Quellentypus der unsichtbar codierten Information*²⁰⁷ als neuer Typus bezeichnet werden.²⁰⁸

²⁰³ Siehe: Brügger: Website Analysis, 2010, S. 45.

²⁰⁴ Nach: Fischer; Hofer: Lexikon der Informatik, 2011.

²⁰⁵ Nach: Manovich, Lev: The language of new media, Cambridge (USA) 2001 (Leonardo), S. 218. URL: http://dss-edit.com/plu/Manovich-Lev_The_Language_of_the_New_Media.pdf, Stand: 29.04.2016.

²⁰⁶ Auch innovative Projekte nutzen weiterhin physische Verbindungen. Siehe bspw.: Thirlwell, Adam: Kapow!, London 2012.

²⁰⁷ Brügger nennt diese Form "not-immediately-visible, non-indicated post festum paratexts". Siehe: Brügger: When the Present Web is Later the Past, 2012, S. 41 f.

²⁰⁸ Dass sich die *unsichtbar codierten Informationen* von bisher bekannten Annotationsformen unterscheiden, ist auch rechtlich von Belang. Das Landesgericht München I hat mit Urteil vom 17.12.2014 (Az. 37 O 8778/14) entschieden, dass die Urheberbenennung eines verwendeten Bildes auf einer Webseite nur mit mouseover der Anforderung der Creative Commons-Lizenz nicht genügt.

2.2.4. Keine neuen Quellentypen

Viele digitalen Quellentypen sind mit bereits bekannten analogen Quellentypen vergleichbar. Dabei ist es unerheblich, ob sie tatsächlich existieren oder sinnvoll sind. Entscheidend ist nur, ob ein digitaler einem analogen Quellentypus entsprechen könnte, auch wenn es umständlich, kostenintensiv oder nicht zielführend wäre. Als Beispiele können genannt werden:

- Verbindungsdaten, die Protokolle der verwendeten IP-Adressen eines Kommunikationsvorgangs, sind vergleichbar mit dem Protokoll einer Korrespondenz (wer hat wann wem einen Brief geschrieben)
- Der Hyperlink stellt eine Beziehung zwischen Objekten dar, die eine Herkunft nachweisen und auf weiterführende Inhalte hinweisen kann.²⁰⁹ Ein Hyperlink kann beide Funktionen auch gleichzeitig übernehmen.
- Der Programmcode ist vergleichbar mit Text, der eine Art Code darstellt, der zu interpretieren ist
- Metadaten sind vergleichbar mit Zusatzinformationen wie Betriebsanleitungen
- Ein (unbekannter) Algorithmus ist vergleichbar mit einer Anfrage an eine Bibliothek, die einen Stapel Bücher bereitstellt. Suchraum, Suchalgorithmus und Vollständigkeit sind hierbei unklar.

Solche digitalen Quellentypen können nicht als neu bezeichnet werden, weil sie ein analoges Pendant haben.

2.3. Zusammenfassung (Fragen und Thesen)

Digitale Objekte können in die drei Digitaldatentypen *digitized data*, *born-digital data* und *reborn-digital data* unterschieden werden. Sie sind nicht zwingend eindeutig, sondern können auch kombiniert vorkommen. Alle drei Digitaldatentypen sind digitale Objekte mit denselben Eigenschaften und können als Quellen verwendet werden.

Die Eigenschaften eines digitalen Objekts wurden definiert. Diese bildeten die Grundlage für die Untersuchung der Unterschiede zwischen einem digitalen und bisher bekannten Objekten. Mit diesen Ergebnissen ist die erste Frage ‚Welche Eigenschaften hat ein digitales Objekt und wie grenzt es sich von den bisher bekannten Objekten ab?‘ beantwortet und die erste These verifiziert:

F1.1: Ein digitales Objekt ist genuin digital, multimodal und -dimensional, nicht nachvollziehbar manipulierbar, volatil, datenträgerunabhängig, verlustfrei vielfältig-

²⁰⁹ Siehe dazu auch: Hodel, Jan: Verkürzen und Verknüpfen: Geschichte als Netz narrativer Fragmente. Wie Jugendliche digitale Netzmedien für die Erstellung von Referaten im Geschichtsunterricht verwenden, Bern 2013 (Geschichtsdidaktik heute), S. 280.

und (gleichzeitig) wiederverwendbar und von einem spezifischen, digitalen Informationssystem abhängig.

F1.2: Ein digitales Objekt unterscheidet sich durch seine ausschliessliche Digitalität, seine verlustfreie und fehlerfreie Vervielfältig- und Wiederverwendbarkeit, die Möglichkeit nicht nachvollziehbarer Manipulation und seiner Abhängigkeit von einem spezifischen, digitalen Informationssystem von bisher bekannten Objekten ab.

T1: Ein digitales Objekt ist eine neue Quellenart, die sich durch bestimmte Eigenschaften von den bisher bekannten Objekten abgrenzt.

Aufgrund dieser Erkenntnisse wurde untersucht, ob und welche Quellentypen digitale Objekte darstellen. Es wurden drei neue Quellentypen definiert, womit die zweite Frage ‚Gibt es neue Quellen, die es analog nicht gibt, und wenn ja, welche?‘ beantwortet und die zweite These verifiziert ist:

F2: Drei Quellentypen wurden definiert: der integrierende, der einobjektreduzierende und jener der unsichtbar codierten Information.

T2: Mit digitalen Objekten sind auch neue Quellentypen entstanden, die es analog nicht gibt.

Diese Erkenntnisse bilden die Basis für die Beurteilung, ob die traditionelle historisch-kritische Methode auch für digitale Objekte verwendet werden kann, ob sie anzupassen ist oder gänzlich verworfen werden muss.

3. Traditionelle Methode und digitale Objekte

Die Geschichtswissenschaft nutzt mit der *historisch-kritischen Methode* einen Arbeitsprozess, um potenzielle Quellen²¹⁰ zu finden und einzuordnen, die vorhandenen und zur Verwendung vorgesehenen Objekte und Aussagen kritisch zu hinterfragen, diese anschließend zu interpretieren und, falls sie Verwendung finden, in einer geeigneten Form darzustellen. Oft wird die historisch-kritische Methode verkürzt als *Quellenkritik* bezeichnet, was aber nur der Kernbestandteil der Methode ist. Mit *traditionell* wird hier die historisch-kritische Methode bezeichnet, die sich nicht mit digitalen Forschungsressourcen auseinandersetzt.

Die traditionellen Methode wird auf ihre Verwendbarkeit für eine Quellenkritik des Digitalen überprüft. Es wird hinterfragt, ob ein hauptsächlich auf physische Objekte ausgelegter Prozess auf alle digitalen und damit datenträgerlosen Objekte angewendet werden kann.²¹¹ Dabei wird herausgearbeitet, ob die Prozessschritte Heuristik – Quellenkunde – Quellenkritik – Interpretation – Darstellung weiterhin verwendet werden können, zu ergänzen sind oder ob eine neue Methode zu erarbeiten ist. In einem weiteren Schritt gilt es die sich aus den Eigenschaften des Objekts aufdrängenden, spezifischen Fragestellungen der Quellenkritik zu untersuchen. In einem kurzen Kapitel werden die Gebote des wissenschaftlichen Arbeitens erläutert, die mit der historisch-kritischen Methode verknüpft sind. Anschliessend erfolgt eine Beurteilung, inwiefern die traditionelle Methode (weiter-)verwendet werden kann. Die vierte und fünfte Frage wird beantwortet und die vierte These verifiziert.

Die Untersuchung bezieht sich allgemein auf digitale Objekte und nicht nur auf digitale Quellen aus dem Internet oder dem Web. Solche werden zwar gerne als Forschungsobjekte verwendet und erwähnt, sie stellen aber nur einen Teil der Gesamtheit an digitalen Daten dar, die alle privaten, geschäftlichen oder wissenschaftlichen Daten, Sensordaten oder Datenflüsse zwischen Maschinen umfassen, unabhängig von ihrem Speicherort oder Verarbeitung.²¹² Leider werden diese Datenarten in der Forschung noch immer ignoriert, denn es werden fast ausschliesslich Problem mit digitalen Objekten aus dem Web diskutiert. Fragen nach Authentizität, Nachvollziehbarkeit oder Vertrauenswürdigkeit sind aber nicht nur an die oft als Beispiel genannte Wikipedia²¹³ zu stellen, sondern beispielsweise auch an die selbst erstellte Datei auf dem eigenen Computer.

²¹⁰ Hier kumulativ verstanden als Text, Gegenstand oder Tatsache.

²¹¹ Siehe: Pfanzer, Eva: Von der Quellenkritik zum kritischen Umgang mit digitalen Ressourcen, in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken: für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 40.

²¹² Nicht wenige Autoren beziehen sich bei digitalen Quellen ausschliesslich auf solche im Internet oder WWW. Siehe bspw.: Lengwiler: Praxisbuch Geschichte, 2011, S. 265-267.

²¹³ URL: <https://www.wikipedia.org/>.

3.1. Traditionelle Prozessschritte und digitale Objekte

"[...] wie es eigentlich gewesen."²¹⁴ (Ranke, 1885)

Eine methodischen Vorgehensweise zur kritischen Überprüfung von Objekten, die als Quellen in einer geschichtswissenschaftlichen Untersuchung verwendet werden sollten, wurde Anfang des 19. Jahrhunderts durch den Historiker Barthold Georg Niebuhr²¹⁵ initiiert. Der 'Vater der Geschichtswissenschaft'²¹⁶, der Historiker Leopold von Ranke, verfeinerte die Ansätze Niebuhrs und forderte unter anderem, dass von der Geschichtsschreibung nur festgehalten werden dürfe, "wie es eigentlich gewesen"²¹⁷ sei. Weiterentwickelt und systematisiert wurde die Methode vor allem vom Historiker und Geschichtstheoretiker Johann Gustav Droysen in seinem Grundlagenwerk *Grundriss der Historik*.²¹⁸ Darin kritisierte er zwar Ranke, attestierte ihm aber "die grösste historische Gelehrsamkeit."²¹⁹ Droysens besonderer Verdienst ist die Ausformulierung der Grundlagen der historischen Quellenkritik.²²⁰ Droysens Methodik wurde von Ernst Bernheim im *Lehrbuch der historischen Methode und der Geschichtsphilosophie*²²¹ ausführlich dargelegt und erweitert.²²² Die Grundlagen der bis heute wichtigsten traditionellen Arbeitsmethode waren damit erarbeitet und besteht aus den fünf Prozessschritten Heuristik, Quellenkunde, Quellenkritik, Hermeneutik und Darstellung.²²³ Weiterentwicklungen erfolgten beispielsweise durch die Erkenntnis, dass eine Quelle ein

²¹⁴ Ranke, Leopold von: Sämtliche Werke, Bd. 33/34, Leipzig 1885, S. 7.

²¹⁵ Vgl. Maltzahn, Christoph Frhr von: Niebuhr, Barthold Georg, in: Vom Bruch, Rüdiger; Müller, Rainer A. (Hg.): Historikerlexikon : von der Antike bis zur Gegenwart, München 2002, S. 236. Auf eine ausführliche Geschichte der Geschichtswissenschaft wird verzichtet, viele verdiente Historiker können nicht erwähnt werden. Weiterführend: Cornelissen, Christoph: Geschichtswissenschaften: Eine Einführung, Frankfurt a.M. 2004.

²¹⁶ Nach: Maltzahn, Christoph Frhr von: Droysen, Johann Gustav, in: Vom Bruch, Rüdiger; Müller, Rainer A. (Hg.): Historikerlexikon : von der Antike bis zur Gegenwart, München 2002, S. 267.

²¹⁷ Nach: von Ranke: Sämtliche Werke, 1885, S. 7. Rankes Aussage ist nach Konrad Reppen ein Zitat von Thukydides 2.48.3: Reppen, Konrad: Über Rankes Diktum von 1824: «Bloss sagen, wie es eigentlich gewesen», in: Historisches Jahrbuch 102, 1982, S. 439-449. Dem widerspricht bspw.: Stroud, Ronald S.: «Wie es eigentlich gewesen» and Thucydides 2.48.3, in: Hermes 115 (H. 3), 1987, S. 379-382. Vielleicht handelt es sich eher um eine Anlehnung als um ein Zitat, denn Thukydides bezieht sich auf einen bestimmten historischen Kontext, Ranke hingegen entwickelte daraus eine Geschichtstheorie.

²¹⁸ Droysen: Grundriss der Historik, 1875.

²¹⁹ Droysen, Johann Gustav: Briefwechsel, Bd. 2, S. 434, in: Muhlack, Ulrich (Hg.): Leopold von Ranke und die Begründung der quellenkritischen Geschichtsforschung, in: Historische Debatten und Kontroversen im 19. und 20. Jahrhundert: Jubiläumstagung der Ranke-Gesellschaft in Essen 2001, Wiesbaden 2003 (Historische Mitteilungen. Beiheft), S. 25.

²²⁰ Droysen: Grundriss der Historik, 1875, S. 16 f., insb. §§ 29-33.

²²¹ Bernheim: Lehrbuch, 1903, B. 1889-1908 (6 Auflagen).

²²² In der Literatur über Geschichtsphilosophie, Geschichtsschreibung und Methoden der Geschichte wird weniger auf Bernheim, als hauptsächlich auf Droysen verwiesen, der mit seiner 'Historik' eine erste Theorie erarbeitet hat. Siehe bspw.: Lengwiler: Praxisbuch Geschichte, 2011, S. 61; Metzger, Franziska: Geschichtsschreibung und Geschichtsdenken im 19. und 20. Jahrhundert, Bern 2011.

²²³ Je nach Autor werden Schritte zusammengenommen (bspw. Quellenkunde als Teil der Heuristik oder der Quellenkritik), ausgegliedert oder hinzugefügt, wie bspw. die 'Relevanzfrage' von: Jordan, Stefan: Einführung in das Geschichtsstudium, Stuttgart 2013 (Reclams Universal-Bibliothek), S. 108 f.

Informationsträger für verschiedene Fragestellungen ist.²²⁴ Weitere Geistes- und Geschichtswissenschaftler wie Marc Bloch, Pierre Bourdieu, Robin Collingwood, Lucien Febvre, Michel Foucault, Johann Huizinga, Karl Marx, Friedrich Meinecke, Theodor Mommsen, Jörn Rüsen, Paul Veyne oder Max Weber beeinflussten die traditionelle Methode massgeblich.²²⁵ Die unterschiedlichen Auffassungen über die Konstruktion, das Erklären und das Verstehen der Vergangenheit, die geschichtsphilosophischen Theorien sowie die Arbeitsweise des Historikers werden weiterhin heftig diskutiert.²²⁶ Der geschichtstheoretische Diskurs wird hier allerdings ausgeklammert, denn die traditionelle Methode wurde nie grundsätzlich infrage gestellt. Über die grosse Anzahl der Gegner der kritischen Methode meinte Marc Bloch pointiert, es sei leichter sie zu tadeln oder zu verspotten, als sich an ihre strengen Gebote zu halten.²²⁷

3.1.1. Heuristik

Die *Heuristik* setzt sich mit der systematischen Suche nach Quellen und Darstellungen in teils hybriden²²⁸ Such- und Informationsräumen²²⁹ auseinander, deren Ergebnis eine sorgfältig erarbeitete und möglichst vollständige Materialbasis sein sollte.²³⁰ Ausgehend von einer Fragestellung wird das verwertbare Material erarbeitet, ergänzt und erweitert.²³¹ Dazu werden je nach Fragestellung verschiedene Findmittel und gegebenenfalls Hilfsmittel aus

²²⁴ Vgl. Lorenz, Chris: Heuristik, in: Jordan, Stefan (Hg.): Lexikon Geschichtswissenschaft: Hundert Grundbegriffe, Stuttgart 2007, S. 140.

²²⁵ Siehe dazu z.B.: Stern, Fritz; Osterhammel, Jürgen (Hg.): Moderne Historiker: Klassische Texte von Voltaire bis zur Gegenwart, München 2011.

²²⁶ Vgl. dazu Beiträge in der Einführungsliteratur in das Studium der Geschichte, bspw.: Cornelissen: Geschichtswissenschaften, 2004. Goertz, Hans-Jürgen: Geschichte. Ein Grundkurs, Reinbek bei Hamburg 1998; Lorenz: Heuristik, 2007, S. 141; Borowsky, Peter; Vogel, Barbara; Wunder, Heide: Einführung in die Geschichtswissenschaft I: Grundprobleme, Arbeitsorganisation, Hilfsmittel, Opladen 1989 (Studienbücher moderne Geschichte 1), Kap. VI. Zur Methode bspw.: Landwehr, Achim: Geschichte des Sagbaren. Einführung in die historische Diskursanalyse, Tübingen 2001 (Historische Einführungen), S. 103-134.

²²⁷ Bloch, Marc: Aus der Werkstatt des Historikers. Zur Theorie und Praxis der Geschichtswissenschaft, Frankfurt a.M. 2000, S. 27.

²²⁸ Unter Hybridität wird im Folgenden das Vermischen des Analogenen und Digitalen verstanden.

²²⁹ Suchraum: Ort, an dem Daten aller Art zur Verfügung gestellt werden und Wissen entstehen kann. Ein Suchraum kann einen oder mehrere 'Informationsräume' enthalten, die aus Fachinformationen bestehen, die durch Auswahlverfahren und Sortierung vorstrukturiert und aufbereitet wurden (bspw. Zentralbibliothek einer Universität, Internet, Institut etc.). Nach: Krüger, Stefanie: Die Erschliessung digitaler und analoger Suchräume. Anforderungen an heuristische Verfahren, in: Haber, Peter; Epple, Angelika (Hg.): Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0, Zürich 2004 (Geschichte und Informatik 15), S. 95.

²³⁰ Zur Hermeneutik siehe bspw.: Jordan: Einführung in das Geschichtsstudium, 2013, S. 109-111; Droysen: Grundriss der Historik, 1875, §§ 19-27; Bernheim: Lehrbuch, 1903, S. 227-293; Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 203.

²³¹ Jordan: Einführung in das Geschichtsstudium, 2013, S. 109-111; Droysen: Grundriss der Historik, 1875, §§19-27; Bernheim: Lehrbuch, 1903, S. 227-293.

anderen Wissenschaftszweigen verwendet. Dieser Prozess wird massgeblich davon beeinflusst, wer die entsprechenden Such- und Informationsräume kontrolliert.²³²

Das Suchen und Finden eines digitalen Objektes ist zwar nur mit der Nutzung eines digitalen Informationssystems möglich, aber als Prozessschritt unerlässlich. Aus analog vorhandenen Objekten wie beispielsweise einer gedruckten Liste, können zwar Hinweise auf digitale Objekte entnommen werden, das digitale Objekt selbst kann aber nur mit und in einem spezifischen digitalen Informationssystem mit den dazu nötigen Komponenten aufgerufen werden.²³³ Steht kein solches zur Verfügung, kann der Suchprozess nicht durchgeführt und allenfalls vorhandene digitale Objekte können nicht gefunden und als digitale Quellen verwendet werden.

Dieser Umstand verlangt jedoch nicht nach einer Anpassung der traditionellen Heuristik. Der fehlende Zugang zu einem Informationssystem ist vergleichbar mit einem nicht zugänglichen Archiv, bei dem nicht sicher ist, ob sich darin relevante Informationen befinden.

Der Prozessschritt der Heuristik ist weiterhin durchzuführen, um die für die Fragestellung relevanten Objekte zu finden. Der Suchprozess verlangt aber durch die Verwendung eines digitalen Informationssystems neue Suchmethoden und -strategien innerhalb neuer Suchräume sowie eine gewisse Erfahrung im Umgang mit solchen Informationssystemen. Unabhängig von den Eigenschaften eines Objekts muss dieses zuerst gesucht werden. Somit lässt sich festhalten, dass die Heuristik als Prozessschritt der traditionellen Methode bei der Anwendung auf digitale Objekte nicht anzupassen ist.

3.1.2. Quellenkunde

In der *Quellenkunde* werden die Zeugnisse der Vergangenheit nach Quellengattungen unterteilt. Die Gruppierung der Quellen erfolgt nach äusserer Beschaffenheit und inhaltlicher Aussagekraft. Quellen werden zunächst nach ihrer *Funktion* in Primär-, Sekundär- und Tertiärquellen unterschieden.²³⁴ Primärquellen werden in *Gattungen* eingeteilt. Trotz des

²³² Siehe: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 203.

²³³ Nötige Komponenten sind bspw.: für das Aufrufen einer URL u.a. eine Internetverbindung (Modem, WLAN-Modul, Bildschirm ...); für die Darstellung eines Objekts mit proprietärem Format die entsprechende Software.

²³⁴ Siehe dazu bspw. Burchert, Heiko; Sohr, Sven: Praxis des wissenschaftlichen Arbeitens: eine anwendungsorientierte Einführung, München 2008 (Studien- und Übungsbücher der Wirtschafts- und Sozialwissenschaften), S. 42-44.

Strebens nach Vereinheitlichung werden verschiedene Schemata verwendet, denen eigene theoretische Ansätze zu Grunde liegen.²³⁵

Als Beispiel das von Ernst Bernheim entwickelte und oft gelehrte und verwendete Schema.²³⁶

Überlieferungen				
Überreste (unwillentlich, funktional)		Tradition (willentlich, intentional)		
Überreste im engeren Sinne	Denkmäler	bildlich	mündlich	schriftlich (veröffentlicht/ unveröffentlicht)
<ul style="list-style-type: none"> - körperliche Reste - Sprache - Zustände und Institutionen - Produkte - geschäftliche Akten 	<ul style="list-style-type: none"> - Inschriften - Monumente - Urkunden 	<ul style="list-style-type: none"> - historische Gemälde - topographische Darstellungen - historische Skulpturen 	<ul style="list-style-type: none"> - Erzählung - Sage - Anekdoten - Sprichwörter - Historische Lieder 	<ul style="list-style-type: none"> - historische Inschriften - Genealogien - Kalender - Annalen - Chroniken - Biographien - Memoiren

Tabelle 1: Quellengattungsschema nach Ernst Bernheim (Darstellung: Föhr)

Eine Unterteilung in Tradition und Überrest sowie in willkürliche und unwillkürliche Überlieferungen birgt die Gefahr, dass damit eine weitere Nutzung der Quelle präjudiziert werden könnte.²³⁷ Eine Quellenkritik hat jedoch unabhängig und "unbeeinflusst von a priori-Kategorisierungen"²³⁸ zu erfolgen. Zudem ist eine strenge Anwendung der Schemata kaum möglich, weil beispielsweise eine Unterscheidung nicht immer zweifelsfrei möglich ist. Ob die Überlieferung absichtlich oder unabsichtlich respektive willkürlich oder unwillkürlich erfolgte, ist meist erst nach der quellenkritischen Überprüfung ersichtlich und die Quelle kann erst dann der Gattung Überrest oder Tradition zugeordnet werden. Eine Tradition ist zudem auch immer ein Überrest, weil sie auch unwillentlich vermittelte Informationen enthält, wohingegen bei einem Überrest nicht immer von Tradition gesprochen werden kann. Die Unterscheidung zwischen Tradition und Überrest darf folglich nur als grober Raster angesehen werden.²³⁹

²³⁵ Andere Schemata siehe bspw.: Borowsky; Vogel; Wunder: Einführung in die Geschichtswissenschaft I: Grundprobleme, Arbeitsorganisation, Hilfsmittel, 1989, S. 124-126; Institut für Geschichtswissenschaft der Universität Bremen: Studienleitfaden Geschichtswissenschaft. Einführung in die Techniken und Methoden des wissenschaftlichen Arbeitens, 2012, S. 14.

²³⁶ Nach: Bernheim: Lehrbuch, 1903, S. 233.

²³⁷ Nach: Schulze, Winfried: Einführung in die neuere Geschichte, Stuttgart 2010, S. 46.

²³⁸ Ebd.

²³⁹ Vgl. auch: Borowsky; Vogel; Wunder: Einführung in die Geschichtswissenschaft I: Grundprobleme, Arbeitsorganisation, Hilfsmittel, 1989, S. 125 f.

Auch digitale Objekte werden einer Quellengattung zugeordnet, weil dies vor allem für die Auswahl der einzusetzenden Hilfsmittel für die quellenkritische Untersuchung hilfreich ist. Je nach Gattung werden andere Komponenten eines digitalen Informationssystems benötigt.

Auch eine digitale Quelle kann nach seiner Funktion unterschieden werden. Beispielsweise kann ein Digitalbild als Primärquelle bezeichnet werden, ein Digitalisat einer Urkunde als Sekundärquelle und ein Artikel aus der Wikipedia als Tertiärquelle. Problematisch wird die Bestimmung der Funktion einer Resultatliste einer Suchmaschinenanfrage. Diese kann nicht als Tertiärquelle aufgefasst werden, weil sie nicht als von einer Person kuratierte, thematische Überblicksdarstellung über Primär- und Sekundärquellen bezeichnet werden kann. Sie ist eine algorithmisch generierte Ansammlung von Hinweisen in nicht nachvollziehbarer Reihenfolge. Des Weiteren stellt sich die Frage, wie die Funktion von Quellen bezeichnet werden soll, die durch einen algorithmisch generierten Algorithmus, die sogenannte künstliche Intelligenz (KI), erstellt wurden. Schliesslich ist zu klären, ob das ganze digitale Informationssystem oder einzelne seiner Komponenten eine Funktion erhalten, weil sie für die Verwendung des digitalen Objekts notwendig sind.

Die in der traditionellen Methode geforderte Unterscheidung zwischen Original und Kopie kann bei digitalen Objekten nicht durchgeführt werden, weil bei digitalen Objekten das Original der Kopie gleichgesetzt ist und jederzeit durch diese ersetzt werden kann. Die Beziehung zwischen Original und Kopie kann als 'metaleptische Umkehrung der Ursache-Wirkung-Relation' bezeichnet werden.²⁴⁰ Einzig Metadaten liefern Hinweise darauf, welches von zwei vorhandenen Objekten zuerst entstanden sein *könnte*.

Im Zusammenhang mit digitalen Objekten sollte allerdings *weder von Original noch von Kopie* gesprochen werden, weil die Verwendung des einen ein Vorhandensein des anderen impliziert. Die Sprachwissenschaftlerin Gisela Fehrmann verspricht Abhilfe durch den Begriff *Originalkopie*²⁴¹, der die Vorstellung der Singularität eines Objekts und der Vorher-nachher-Beziehungen aufheben soll. Zusätzlich würde zum Ausdruck gebracht werden, dass der Status von Original und Kopie nicht einfach als gegeben, sondern als Produkt transkriptiver Prozesse anzusehen ist.²⁴² Von der Verwendung von Originalkopie wird hier aber abgesehen, denn der Rückgriff auf die traditionellen Begriffe Original und Kopie und deren Kombination kann nicht nur Verwirrung stiften, sondern auch dazu verleiten, die traditionelle Methode anzuwenden.

²⁴⁰ Metalepse: rhetorische Figur, bei der das Nachfolgende mit dem Vorhergehenden vertauscht wird. Nach: Fehrmann, Gisela; Linz, Erika; Schumacher, Eckhard u. a.: Originalkopie: Praktiken des Sekundären – Einleitung, in: Originalkopie: Praktiken des Sekundären, Bd. 11, Köln 2004 (Mediologie), S. 9.

²⁴¹ Vgl. diverse Beiträge in: Fehrmann, Gisela; Linz, Erika; Schumacher, Eckhard u. a.: Originalkopie: Praktiken des Sekundären, Bd. 11, Köln 2004 (Mediologie).

²⁴² Nach: Fehrmann u. a.: Originalkopie: Praktiken des Sekundären – Einleitung, 2004, S. 9.

Wird ein Original kopiert, entsteht nicht eine Kopie, sondern es entstehen zwei *Klone*.²⁴³ Diese sind nicht unterscheidbar, abgesehen vom Entstehungsdatum, das aber mit entsprechenden Mitteln ebenfalls geklont werden kann. Welcher der beiden Klone das Original darstellt, kann nicht mehr aufgrund des Früheren oder Späteren nachvollzogen werden, sondern kann nur durch Zuschreibung erfolgen.²⁴⁴ Zudem ist das digitale Objekt medienunabhängig, weshalb keine physische Komponente als Original bezeichnet werden kann.

Wenn keine chronologische Unterscheidung möglich ist, wird Droysens Echtheitsprüfung, die bei physischen Objekten zur Ermittlung der Originalität im Gegensatz zur Kopie oder Fälschung dient, obsolet.²⁴⁵ Das wiederum bedeutet, dass andere Kriterien für die Beurteilung von digitalen Objekten entwickelt werden müssen.

Welche Funktion algorithmisch generierte Objekte und durch einen Algorithmus algorithmisch generierte Objekte übernehmen können, welche Komponenten eines digitalen Informationssystems Teil der Funktion sind und wie das Verhältnis zwischen Original und Kopie geklärt werden kann, sind weiterhin ungelöste Problemkreise. Trotzdem erachte ich den Prozessschritt der Quellenkunde für die Untersuchung digitaler Objekte weiterhin als notwendig, weil darin die für eine Untersuchung geeigneten Hilfsmittel bereitgestellt werden können. Der Prozessschritt der Quellenkunde ist weiterhin durchzuführen.

3.1.3. Quellenkritik

Die *Quellenkritik* ist der Hauptbestandteil der traditionellen Methode und der "Inbegriff einer methodisch geleiteten Auswertung der Materialien der historischen Erkenntnis."²⁴⁶ Grundsätzlich ist kein Objekt wertlos, es ist die Fragestellung, die den Quellenwert bestimmt. Der Aussagewert einer Quelle ist zu eruieren, indem die besonderen räumlichen, zeitlichen, subjektiven und formalen Bedingungen einer Quelle erfasst und geklärt sowie der Wahrheitsgehalt und die Glaubwürdigkeit der Quelle geprüft werden.²⁴⁷ In der Quellenbeschreibung werden die Art der Quelle respektive die Quellengattung, die Überlieferungsgeschichte und der äussere Erhaltungszustand des Objekts untersucht.

Bei gewissen Quellen wie mittelalterlichen Urkunden erfolgt eine *Textsicherung* im Sinne der Feststellung des authentischen Wortbestandes. Dies impliziert das ‚Lesen-Können‘ eines

²⁴³ Siehe Kap. 2.1. Eigenschaften eines Digitalen Objekts.

²⁴⁴ Siehe: Haber: digital past (unv. Habil.), 2009, S. 195; Fehrmann u. a.: Originalkopie: Praktiken des Sekundären – Einleitung, 2004, S. 9.

²⁴⁵ Nach: Droysen: Grundriss der Historik, 1875, S. 16, § 30.

²⁴⁶ Jaeger, Friedrich: Geschichtstheorie, in: Goertz, Hans-Jürgen (Hg.): Geschichte. Ein Grundkurs, Reinbek bei Hamburg 1998, S. 736.

²⁴⁷ Nach: Fuchs, Konrad; Raab, Heribert: Quellenkritik, in: Wörterbuch zur Geschichte, München 2002, S. 653.

Textes, womit sowohl die Entzifferung der Zeichen (verblasst, überschrieben etc.) als auch die Sprache (Fremdsprache, Begriffe etc.) gemeint ist.²⁴⁸

Oft wird für diese *Prüfung der Authentizität* einer Quelle auf die Bernheim'sche Unterteilung in eine *innere (höhere) und äussere (niedere) Kritik* zurückgegriffen.²⁴⁹ Letztere beurteilt die "Echtheit und Vollständigkeit der Quelle, klärt Indizien für inhaltliche Fälschungen, falsche Autorangaben und Überlieferungsvarianten ab und erhellt den Entstehungskontext der Quelle und deren Überlieferungsgeschichte."²⁵⁰ Die innere Quellenkritik beurteilt den Inhalt und fragt nach dem Aussagewert der Quelle. Die Unterteilung in eine innere und eine äussere Quellenkritik ist allerdings nicht immer zweifelsfrei gegeben, weshalb von einigen Autoren zu Gunsten eines einfacheren oder strukturierteren Prozesses darauf verzichtet wird. Beispielsweise verwendet der Historiker Stefan Jordan die Schritte *Prüfung der Echtheit, der Chronologie und der Richtigkeit*, in denen die bereits erwähnten Fragen im Rahmen der Quellenkritik beantwortet werden, um die Authentizität der Quelle und damit deren Aussagewert zu festzustellen.²⁵¹

Die *Feststellung der Authentizität* ist das wichtigste und umfassendste Konzept der traditionellen Methode, denn eine Quelle sollte *echt* und *glaubwürdig* sein.²⁵² Nicht was eine Quelle als wahr ausgibt, sondern die tatsächliche Wahrheit soll festgestellt werden.²⁵³

Die kritische Überprüfung eines jeden Objekts, das als Quelle infrage kommt, ist das Herzstück der traditionellen Methode und ist auch für ein digitales Objekt weiterhin durchzuführen. Diese Erkenntnis ist offensichtlich und würde grundsätzlich keiner weiteren Ausführungen bedürfen. Jedoch sind die in diesem Schritt zu stellenden Fragen auf die Eigenschaften eines digitalen Objekts anzupassen.

Auch bei digitalen Objekten ist eine Textsicherung durchzuführen. Allerdings ist diese allgemeiner zu fassen, denn es ist schon fraglich, ob das Objekt mit den vorhandenen Möglichkeiten technisch ‚gelesen‘ werden kann. Das digitale Objekt muss in einem Dateiformat vorliegen, das vom zur Verfügung stehenden Informationssystem korrekt dargestellt werden kann. Dazu muss die digitale Forschungsressource vorgängig in einem langzeitarchivfähigen

²⁴⁸ Siehe zur Textsicherung bspw.: Borowsky; Vogel; Wunder: Einführung in die Geschichtswissenschaft I: Grundprobleme, Arbeitsorganisation, Hilfsmittel, 1989, S. 157-160.

²⁴⁹ Nach: Bernheim: Lehrbuch, 1903, S. 300 und 429.

²⁵⁰ Lengwiler: Praxisbuch Geschichte, 2011, S. 98f.

²⁵¹ Vgl.: Jordan: Einführung in das Geschichtsstudium, 2013, S. 111-115.

²⁵² Siehe: Net Art as a Methodological Case Study, in: Daniels, Dieter; Reisinger, Gunther (Hg.): Net pioneers 1.0 contextualizing early net-based art, Berlin 2009, S. 123-142. URL: <http://www.netzpioniere.at/node/52>, Stand: 06.05.2018.

²⁵³ Weiter zur Quellenkritik: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 322 f; Büttner, Sabine: Einführung: Arbeiten mit Quellen, [historicum-estudies.net](http://www.historicum-estudies.net), 12.02.2007, URL: <http://www.historicum-estudies.net/etutorials/tutorium-quellenarbeit/?L=962>, Stand: 06.05.2018; Buchmann, Bertrand Michael: Einführung in die Geschichte, Wien 2002 (Facultas-Manuals).

Format und in einem vertrauenswürdigen Archiv gespeichert worden sein. Dieser Schritt fehlt im traditionellen Prozess!

Zur Feststellung der Authentizität werden insbesondere die Echtheit, Vollständigkeit und Provenienz des Objekts untersucht. Besonders problematisch ist das Kriterium der Echtheit, denn wenn Echtheit wie bei Droysen als *Originalität des Objekts* aufgefasst wird, wird deren Prüfung obsolet, weil ein Original nicht von einer Kopie unterschieden werden kann.²⁵⁴ Eine Unterscheidung zwischen älterer und neuer Version mittels diakritischem Verfahren wird bei digitalen Klonen nicht erfolgreich sein. Falls es sich tatsächlich um eindeutige Klone handelt, können diese jedoch auf ein einziges Objekt reduziert werden, denn die Unterscheidung ist nicht nötig.

Wird jedoch unter einem echten Objekt ein Objekt verstanden, das sich von einer Fälschung unterscheidet, kann mithilfe geeigneter Methoden die Wahrscheinlichkeit der Echtheit eines Objekts eruiert werden. Als geeignet werden informationstechnische und traditionelle Methoden betrachtet, die in einem meist hybriden Vorgehen die quellenkritischen Fragen beantworten helfen. Beispielsweise kann ein Digitalbild nur mittels eines digitalen Informationssystems überprüft werden. Liegt das Bild im RAW-Format²⁵⁵ vor und ist die zugehörige EXIF-Datei²⁵⁶ vorhanden, können mit einer entsprechenden Software die Daten des Bilds und dessen Metadaten ausgelesen werden. Erst mit diesen Informationen können inhaltliche Aussagen über das Bild gemacht werden. Wenn beispielsweise der Dateiname 'Rhodos-2015-01' lautet, aber in den zugehörigen Metadaten der GPS-Ort Kreta und das Datum 2011 gefunden werden, dann können diese auf eine Fälschung hinweisen.²⁵⁷

Damit ist bereits die Überlieferungsgeschichte angesprochen, die mit der Entstehungszeit weitere Hinweise auf die Echtheit einer Quelle liefert. Jedoch ist bei digitalen Objekten auch diese Angabe nicht eindeutig. Es ist beispielsweise bei einem wissenschaftlichen Paper nicht klar, welches Datum ausschlaggebend sein soll: Das im Dokument erwähnte, das aber vor- oder rückdatiert sein kann, das letzte Speicherdatum des Objekts als Metadatum, das Speicherdatum der Speichersystems einer Webseite, womit das Objekt veröffentlicht wurde, oder das Speicherdatum beim Download durch den Nutzer. Ebenso ist bei digitalen Objekten die Bestimmung des Entstehungsorts problematisch, denn dieser ist beispielsweise bei einem kollaborativ erarbeiteten Dokument nicht eindeutig. Zu dessen Erstellung wird ein Online-

²⁵⁴ Vgl.: Haber: digital past (unv. Habil.), 2009, S. 195 f.

²⁵⁵ RAW-Format: Rohdatenformat für Digitalbilder. In anderen Formaten (bspw. GIF – Graphics Interchange Format) werden aus den Rohdaten durch die Kamerasoftware viele Informationen nicht mit abgespeichert, obwohl diese vorhanden wären. Das Format enthält nur Farb- und Helligkeitswerte der einzelnen Pixel.

²⁵⁶ EXIF-Daten: Exchange Imaging Format for Still Cameras. Metadaten zu Kamerahersteller, Modell, Blendenwert, Verschlusszeit, ISO-Einstellung, Objektiv, Weissabgleich, Bildauflösung, Einsatz des Blitzes, Datum, Uhrzeit etc. Bei GPS-fähigen Kameras auch diese Angaben.

²⁵⁷ GPS: Global Positioning System. Siehe URL: <http://www.gps.gov/>.

Programm genutzt, das eine Institution mit bekanntem Hauptsitz aber unbekanntem Speicherort des Objekts zur Verfügung stellt. Die Autoren können zudem aus verschiedenen Ländern stammen. Wird das Objekt online publiziert, werden wiederum zwei mögliche Orte erzeugt, den Sitz der Institution und den Standort des Speichersystems. Weiter kann die Authentifizierung schwierig sein, wenn beispielsweise ein Artikel auf einer Online-Plattform wie Wikipedia anonym oder unter einem Pseudonym platziert wird.

Ein weiteres Problem stellt sich, wenn die Vollständigkeit einer digitalen Quelle festgestellt werden soll. Hyperlinks können mehrere Funktionen übernehmen. Beispielsweise kann auf externe Inhalte verwiesen werden, die damit viel enger als bei physischen Objekten eingebunden werden können. Durch einen Hyperlink kann der Autor den externen Inhalt als Voraussetzung oder als Teil seines Beitrags einfügen, das heisst, der Leser muss ihn lesen, wenn er nicht zufälligerweise Vorkenntnisse zum Thema hat. Zudem ist es bei einem digitalen Objekt viel einfacher, Teile daraus zu entfernen. Beispielsweise bewirkt die Löschung einer Graphik in einem Textverarbeitungsprogramm meistens, dass der gesamte restliche Text vorgeschoben wird. Die Löschung ist für einen Leser, der die vorherige Version nicht kennt, nicht nachvollziehbar. Derselbe Vorgang ist hingegen bei einem Buch mit physischem Trennen verbunden, was sehr wahrscheinlich entdeckt wird.²⁵⁸

Schliesslich sind die Fragen nach Inhalt, Intention und Tendenz, Bezügen zu anderen Quellen und dem/den Adressat(en) zu klären. Insbesondere die letzte Frage kann schwierig zu beantworten sein, wenn beispielsweise ein auf einer offenen Online-Plattform publizierter Fachartikel nicht mehr nur für ein bestimmtes Publikum zugänglich ist, sondern von jeder Person gelesen und kritisiert werden kann. Damit ändert sich das Verhältnis zwischen Autor und Adressat, denn beide haben die Möglichkeit, aktiv miteinander in Kontakt zu treten und mit Kommentieren und Editieren den Artikel unmittelbar zu verändern. Sie kommunizieren sichtbar miteinander.

Für die Untersuchung des Inhalts und des Aussagewerts einer digitalen Quelle sind trotz der erwähnten Schwierigkeiten die gleichen Fragen wie bei der Anwendung der traditionellen Methode zu stellen. Sowohl die Fragen nach dem Verfasser, dessen Intention oder Tendenz, der Überlieferung selbst, dem Adressaten, dem Inhalt und den Bezüge zu anderen Quellen. Aber auch die Klärung der Überlieferungsgeschichte, Vollständigkeit und Provenienz bilden weiterhin den Inhalt einer quellenkritischen Untersuchung. Ob zwischen einer inneren und äusseren Quellenkritik unterschieden wird oder die Fragen in einer anderen Art an das zu untersuchende Objekt gestellt werden, ist unerheblich. Festgestellt wird, dass der bisherige Fragenkatalog auch bei digitalen Objekten anzuwenden ist.

²⁵⁸ Das Herausschneiden bezieht sich nur auf die Graphik, d.h., es wird in der Buchseite ein Loch entstehen.

Allerdings zeichnet sich ein digitales Objekt nebst anderen Eigenschaften dadurch aus, dass es ein digitales Informationssystem mit entsprechenden Komponenten für dessen Darstellung benötigt. Ohne dieses System kann weder das Objekt noch dessen Metadaten gelesen werden. Deshalb ist zusätzlich zum digitalen Objekt auch das er- und das darstellende Informationssystem zu untersuchen. Dabei werden nicht die formalen Aspekte des Objekts, sondern diejenigen des Informationssystems untersucht, ein Punkt, in dem sich die Quellenkritik des Digitalen von der traditionellen Methode unterscheidet.

Der Fragenkatalog der traditionellen Methode ist folglich um informationstechnische Fragen zu erweitern.²⁵⁹

- In welchem Format liegt das Objekt vor?
- Welche Hard- und Software wurde bei der Erstellung verwendet?
- Welche Hard- und Software ist zur gleichwertigen Darstellung nötig? Eventuell mit einer Emulation?
- Welche Migrationsschritte hat das Objekt durchlaufen, unter Angabe der verwendeten Hard- und Software?
- Wie wird das Objekt dauerhaft archiviert?
- Wie werden die Angaben zu obigen Fragen erstellt, gespeichert und dargestellt respektive wie werden die Metadaten erstellt?
- Welche Sicherheitsmerkmale sind vorhanden?

Antworten auf diese Fragen sollten in den zugehörigen Metadaten zu finden sein. Diese sind aber in vielen Fällen nicht Teil des Objekts selbst, sondern als zusätzliche Datei(en) oder Einträgen in Informationssystemen abgelegt. Digital vorhandene Metadaten sind wiederum als eigenständige Objekte zu behandeln und entsprechend zu untersuchen. Wie bei bekannten analogen Objekten gilt der Grundsatz, dass ein Objekt umso echter ist, je mehr überprüfbare Informationen im Objekt oder in den zugehörigen Metadaten zur Verfügung stehen.

Abhängig von der Art des Objekts sind für die genannten Überprüfungen verschiedene Methoden und Techniken einzusetzen, die teilweise aus anderen Wissenschaftsbereichen wie der Informatik stammen und für die historisch-kritische Methode angepasst werden. Welche Methode wie angewendet wird, ist hier unerheblich. Es ist hier nur festzuhalten, dass in der Geschichtswissenschaft jedes Objekt mit geeigneten Methoden kritisch zu hinterfragen ist.

3.1.4. Interpretation

In der *Quelleninterpretation* (Quellendeutung, -auswertung oder Auffassung; Hermeneutik)²⁶⁰ wird das vorliegende historische Material objektiv gewürdigt und in einen weiteren histori-

²⁵⁹ Auf die genannten Fragen wird in Kap. 5. Quellenkritik des Digitalen eingegangen.

schen Kontext eingliedert. Beides erfolgte teilweise bereits während der Quellenkritik respektive im dialogischen Prozess zwischen kritischem Hinterfragen und Interpretation.

Der Begriff der (historischen) Hermeneutik ist nicht eindeutig definiert.²⁶¹ Für die Historiker Konrad Fuchs und Heribert Raab ist Hermeneutik in der Geschichtswissenschaft schlicht "die Lehre von der Auslege- oder Interpretationskunst". Die Interpretation definieren sie als "die Auslegung und Erklärung von Quellen".²⁶² Die Historiker Ulrich Muhlack und Thomas Welskopp ordnen der Interpretation *Verstehen* und *Erklären* als die beiden Grundformen oder -konzepte des historischen Begreifens zu.²⁶³ Verstehen und Erklären beeinflussen sich gegenseitig und stehen deshalb in einem engen Verhältnis zu- aber nicht gegeneinander.

Für Droysen ist das Wesen der historischen Methode *forschend zu verstehen*²⁶⁴ Der Historiker bemüht sich dabei um eine ganzheitliche Deutung mit dem Ziel, "geschichtliche Wirklichkeiten im Vergleich mit der Gegenwart deutlich werden zu lassen".²⁶⁵ Mit Erklären ist die Analyse von Faktoren mittels 'Warum-Fragen' gemeint, um allfällige und möglichst plausible Zusammenhänge unter diesen feststellen zu können. Droysen argumentiert zwar, dass die historische Forschung nicht erklären will, beschränkt seine Ablehnung aber auf die Darlegung von vermeintlich notwendigen oder gesetzesmässigen, zeitlichen Abfolgen.²⁶⁶ Dem entgegen hat für Thomas Welskopp ein Historiker historische Phänomene zu erklären, indem diese explizit "als Wirkungen von Ursachen beschrieben werden".²⁶⁷

Mareike König fasst in einem Blog-Beitrag sehr gut zusammen, wie in den Geschichtswissenschaften Erkenntnisse gewonnen werden:

"[...] durch die hermeneutische Interpretation²⁶⁸ von Texten, Bildern und Objekten, durch das narrative Füllen von Leerstellen, die uns diese aufgeben, durch spezifische Verknüpfungen, die in der Geschichtswissenschaft zumeist eine temporale Struktur aufweisen. Idealtypisch folgen wir dabei einer Fragestellung, verwenden standardi-

²⁶⁰ Droysen verwendet den Begriff 'Quelleninterpretation', Bernheim 'Auffassung'. Für Geschichte und aktuelle Diskussion um den Begriff Hermeneutik siehe bspw.: Jung, Matthias: Hermeneutik zur Einführung, Hamburg 2012⁴.

²⁶¹ Ausführlich bspw.: Ebd.

²⁶² Fuchs; Raab: Quellenkritik, 2002, S. 334 und 337.

²⁶³ Vgl.: Muhlack, Ulrich: Verstehen, in: Goertz, Hans-Jürgen (Hg.): Geschichte. Ein Grundkurs, Reinbek bei Hamburg 1998, S. 99; Lorenz, Chris: Konstruktion der Vergangenheit, Köln 1997, S. 91. Auf die Diskussionen über die beiden Begriffe und die Geschichte der Hermeneutik wird hier aus Platzgründen nicht weiter eingegangen.

²⁶⁴ Droysen: Grundriss der Historik, 1875, §8.

²⁶⁵ Jordan, Stefan: Theorien und Methoden der Geschichtswissenschaft. Orientierung Geschichte, Paderborn 2008, S. 47.

²⁶⁶ Droysen: Grundriss der Historik, 1875, §37.

²⁶⁷ Welskopp, Thomas: Erklären, in: Goertz, Hans-Jürgen (Hg.): Geschichte. Ein Grundkurs, Reinbek bei Hamburg 1998², S. 132.

²⁶⁸ Wenn, wie oben ausgeführt, Hermeneutik mit Interpretation(skunst) gleichgesetzt wird, ist Königs Formulierung einer 'hermeneutischen Interpretation' als Pleonasmus zu bezeichnen.

*sierte Methoden, bemühen uns um Berücksichtigung verschiedener Standpunkte und sind uns der eigenen Subjektivität beim Abfassen einer Darstellung bewusst.*²⁶⁹

Welche Ansätze und innovative Methoden²⁷⁰ für die Interpretation verwendet werden können, erschliesst sich aus dem Quellenmaterial und werden vom Historiker aufgrund der Forschungsfrage, seinen Erfahrungen, seiner Ausbildung, seinem Werdegang, seinem Wissensstand und seiner Vorlieben ausgewählt.²⁷¹ Die Interpretation ist deshalb immer subjektiv.

Die Interpretation der verwendeten Quelle ist trotz unklarer Definition auch bei einem digitalen Objekt der auf die Quellenkritik folgende Schritt. Ohne die Interpretation kann weder die Quelle verwendet und in die Darstellung eingebettet noch können neue Erkenntnisse gewonnen werden.

Bei digitalen Objekten ist allerdings zusätzlich zu deren Interpretation auch zu beurteilen, ob und inwiefern die Erkenntnisse der Untersuchung der verwendeten digitalen Informationssysteme mit einfließen. Digitale Objekte sind mit digitalen Informationssystemen untrennbar verknüpft, weshalb Letztere in die Interpretation miteinzubeziehen sind. Deshalb ist die Interpretation einer Quelle um die Beurteilung der verwendeten digitalen Informationssysteme zu ergänzen.

3.1.5. Darstellung

Schliesslich mündet die Arbeit des Historikers in eine *Darstellung* der Forschungsergebnisse. Sie ist ein unverzichtbarer Teil der Forschung, ohne die sonst das Erforschte "unreif und ungeschlossen"²⁷² bliebe. Dazu werden die Tatsachen unter Einhaltung bestimmter formaler Kriterien wie Quellenangaben, Verweise, Definitionen oder präzise Formulierung²⁷³ objektiv

²⁶⁹ König, Mareike: Jenseits der Metaphorik: Experimente in den Digital Humanities #dhiha6, Billet, digitale:geschichte, 07.06.2015, URL: <https://dguw.hypotheses.org/257>, Stand: 19.02.2018.

²⁷⁰ Bspw. wurden bis in die 1980er-Jahre Fallstudien oder Stichproben in der Geschichtswissenschaft höchst selten verwendet. Danach fanden sie als 'innovative Methode' in spezifischen Forschungsfeldern Verwendung und sind heute weitgehend akzeptiert, obwohl noch immer Diskussionen um deren Verwendungszweck geführt werden. Siehe dazu bspw.: Lengwiler: Praxisbuch Geschichte, 2011, S. 55-73.

²⁷¹ Achim Landwehr argumentiert dagegen, dass "die Frage nach dem, was hinter dem Papier liegt, nach den 'eigentlichen Intentionen', nach der Mentalität, nach den 'leitenden Interessen'" nur spekulative Ergebnisse liefern könne. Vgl.: Landwehr: Geschichte des Sagbaren. Einführung in die historische Diskursanalyse, 2001, S. 103.

²⁷² Droysen: Grundriss der Historik, 1875, § 44.

²⁷³ Quellenangaben, Verweise etc. sind Kriterien des wissenschaftlichen Arbeitens. Siehe Kap. 3.4. Wissenschaftliches Arbeiten.

und angemessen dargestellt, in den historischen Kontext eingebettet und die Ergebnisse und deren Relevanz in Bezug auf die Fragestellung verständlich erläutert.²⁷⁴

Die gewonnenen Erkenntnisse sind auch weiterhin in einer adäquaten Form darzustellen. Erst mit der Darstellung tritt der Autor (der Forscher) in die Diskussion ein und ermöglicht eine Auseinandersetzung mit der Forschungsgemeinschaft. Aus und mit den verwendeten Quellen wird etwas Neues dargestellt.

Wenn digitale Objekte als Quellen verwendet werden, bietet sich eine digitale Darstellungsform an. Mit dieser lassen sich die Möglichkeiten der Vernetzung nutzen, indem auf digitale Quellen verlinkt werden oder diese direkt integriert werden. Damit kann auch der Medienbruch verhindert werden, der durch eine analoge Darstellung entstehen würde, denn das Objekt kann nicht wie ursprünglich vorhanden dargestellt werden. Diese Problematik bestand auch bisher schon, denn beispielsweise ist ein auf einer Fotografie abgebildetes Haus durch den Betrachter nicht begehbar. Im Unterschied zum Haus wäre jedoch ein digitales Objekt in seiner intendierten Form weiterhin erfahrbar, weil es in seinem Kontext verbleiben könnte. Wird beispielsweise der Zugriff auf eine Webseite dauerhaft gewährleistet, kann der Nutzer diese durch Aufrufen des entsprechenden Links in der Darstellung und unabhängig von Zeit und Ort 'erfahren'.

Nicht nur die Masse an digitalen Objekten, sondern auch deren Vielfältigkeit (Bild, Datenbank, Netzwerkvisualisierung, Twitter-Archiv, virtuelle Welt, Protokolldatei etc.) verlangt nach adäquaten wissenschaftlichen Darstellungsformen. Bisher verwendete, analoge Formen und Formate verursachen durch die Transformation des Digitalen ins Analoge einen Medienbruch, der die meisten Eigenschaften eines Objekts zerstört. Damit ist auch keine adäquate Darstellung möglich. Bei digitalen Objekten ist eine digitale Form vorzuziehen.

3.2. Eigenschaften digitaler Objekte

Es konnte bestätigt werden, dass der Prozess der traditionellen historischen Methode grundsätzlich beibehalten werden kann. Jedoch erfordert die Verwendung von digitalen Objekten mit ihren Eigenschaften Digitalität und Volatilität den zusätzlichen Schritt der Quellensicherung. Die bisher vorhandenen datenträgerlosen Quellen – etwa die mündlichen Aussagen in der Oral History – sind in ihrer Quantität sehr begrenzt und können als Spezialfälle angesehen werden. Da die digitalen Quellen die Anzahl der anderen Quellengattungen längst über-

²⁷⁴ Stefan Jordan fordert explizit so wenige Fremdwörter wie möglich zu verwenden, denn ein wissenschaftlicher Text soll sich durch seinen Inhalt auszeichnen, komplexitätsmindernd wirken und einen verständlichen Satzbau aufweisen. Siehe: Jordan: Einführung in das Geschichtsstudium, 2013, S. 130 f. Weiter: Lengwiler: Praxisbuch Geschichte, 2011, S. 240-244.

holt haben und weiter zunehmen, ist die Methode auf deren Bedürfnisse anzupassen. Nicht mehr physische Objekte sind als *Leitmedium* zu betrachten, sondern die digitalen.

Für die Untersuchung, ob die einzelnen Eigenschaften eines digitalen Objekts innerhalb des Prozesses der traditionellen Methode überprüft werden können, wird auf die obige Definition eines digitalen Objekts zurückgegriffen:

Ein digitales Objekt ist genuin digital, multimodal und -dimensional, nicht nachvollziehbar manipulierbar, volatil, datenträgerunabhängig, verlustfrei vielfältig- und (gleichzeitig) wiederverwendbar und von einem spezifischen, digitalen Informationssystem abhängig.

Das genuin Digitale bezieht sich auf die Auffindbarkeit und Zugänglichkeit sowie Lesbarkeit des Objekts. In der traditionellen Methode wird davon ausgegangen, dass Gedächtnisinstitutionen bestehen, welche die Auffindbarkeit, Zugänglichkeit und Lesbarkeit langfristig und vertrauenswürdig gewährleisten. Für digitale Objekte sind dementsprechend digitale Langzeitarchive nötig, die digitale Objekte sicher archivieren und für die Wissenschaft dauerhaft zur Verfügung stellen zu können. Das *genuin Digitale* stellt für die traditionelle Methode ein Problem dar, denn digitale Objekte können nicht materialisiert werden und stehen somit auch nicht in traditionellen Institutionen dauerhaft zur Verfügung.

Physische Objekte sind multimodal, denn sie können durch verschiedene Sinne wahrgenommen werden. Digitale Objekte können aber mit keinem Sinn wahrgenommen werden, sondern nur deren Interpretation. Allerdings liegt ein digitales Objekt in verschiedenen Formaten gleichzeitig vor, denn es kann sowohl als Ganzes betrachtet werden oder als Algorithmus oder Code. Die kritische Überprüfung dieser verschiedenen Formate kann für jedes Format einzeln erfolgen. Die Eigenschaft der *Multimodalität* benötigt keine Erweiterung der traditionellen Methode.

Die Frage nach der *Multidimensionalität*, den verschiedenen Informationsebenen und Autoren sowie der Verbreitung und Originalität, kann ebenfalls mit der traditionellen Methode behandelt werden. Traditionelle multidimensionale Quellen sind zum Beispiel Lexika.

Als volatil werden hauptsächlich Zustände wie Wertänderungen von Aktien in der Finanzbranche oder die Flüchtigkeit von chemischen Stoffen beschrieben. Physische Informationsträger wie Papier werden nicht als volatil bezeichnet, weil sie grundsätzlich für die Aufbewahrung über einen längeren Zeitraum verwendet, archiviert und auch wiederverwendet werden können. Dies ist bei digitalen Objekten wegen ihrer nicht fixierbaren Volatilität nicht möglich. Die *Volatilität* digitaler Objekte stellt für die traditionelle Methode ein Problem dar.

Jedes Objekt ist manipulierbar, weshalb bei jeder quellenkritischen Überprüfung nach Indizien auf eine Manipulation gesucht wird. Bei digitalen Objekten kann diese, unter Verwendung entsprechender Mittel, jedoch nicht nachgewiesen werden. Es ist sogar eher davon

auszugehen, dass das Objekt manipuliert wurde, weshalb Indizien zu seiner Authentizität gesucht werden sollten. Die *nicht nachvollziehbare Manipulierbarkeit* digitaler Objekte stellt für die traditionelle Methode ein Problem dar.

Bei datenträgerunabhängigen Quellen entfällt die quellenkritische Untersuchung des Datenträgers, der bei physischen Objekten meist weitere, wichtige Informationen enthält. Bei physischen Objekten ist mit dem Datenträger meist auch ein Aufbewahrungsort verbunden, der selbst weitere Informationen überliefert, wie auch die zusätzlich an diesem Ort aufbewahrten Objekte im Bestand oder der Serie. Für eine quellenkritische Untersuchung und vor allem für die Nachvollziehbarkeit der Forschungstätigkeit ist es notwendig, dass digitale Quellen an *vertrauenswürdigen und dauerhaften Speicherorten* gesichert werden können. Dies ist in der traditionellen Methode nicht vorgesehen.

Mit der Erfindung des Buchdrucks wurde eine Art von verlustfreiem Kopieren möglich. Verschiedene Institutionen können ein Exemplar respektive eine Kopie desselben Buchs vorhalten, was bei einem digitalen Objekt mit der Verteilung auf verschiedene Speicherorte oder Systemkomponenten zu vergleichen ist. Digitale Objekte können anders als ein Buch unmitelbar und tatsächlich verlustfrei kopiert und damit geklont werden. Die *Unmittelbarkeit* kann aber als eine Weiterentwicklung des Kopierprozesses angesehen werden, was im Rahmen der traditionellen Methode untersucht werden kann.

Die *verlustfreie Wiederverwendbarkeit* ist bei einem physischen Objekt nicht möglich²⁷⁵, selbst wenn es aus besonders resistentem Materialien wie Stein besteht. Digitale Objekte sind jedoch tatsächlich verlustfrei wiederverwendbar, was zwar als eine Weiterentwicklung von sehr resistenten Materialien angesehen werden kann, aber keinen zusätzlichen Untersuchungsschritt nötig macht. Problematisch ist hingegen, dass ein digitales Objekt *gleichzeitig wiederverwendbar* ist, das heißt zeitlich und räumlich von verschiedenen digitalen Informationssystemen unabhängig. Damit entstehen neue Beziehungsformen unter Objekten, deren Untersuchung in der Quellenkritik bereits vorgesehen ist und keine Anpassung erfordert.

Eine von einem *spezifischen digitalen Informationssystem* abhängige Quelle ist nur dann zugänglich, wenn ein entsprechendes System mit allen nötigen Komponenten zur Verfügung steht. Ein nicht zugängliches digitales Objekt gleicht einem Dokument, das in einem verschlossenen physischen Archiv liegt und auf das nicht zugegriffen werden kann. Obwohl man vielleicht weiß, dass es dort vorhanden ist, kann es nicht verwendet werden. Auch diese Eigenschaft ist also nicht neu.

²⁷⁵ Erfahrungen aus der Archivwissenschaft resp. eines jeden Archivars.

Nach dieser Untersuchung kann die dritte Frage 'Welche Probleme entstehen durch die Eigenschaften eines digitalen Objekts bei der historisch-quellenkritischen Überprüfung (im Gegensatz zu physischen Objekten)?' beantwortet werden:

F3: Die genuine Digitalität kann mit bisherigen Methoden nicht untersucht werden, sondern bedarf Methoden aus der Informationstechnik, weil unter anderem für die Entstehung, Verarbeitung und Darstellung eines digitalen Objekts immer ein digitales Informationssystem benötigt wird. Ebenso stellt die mit der Digitalität verbundene Volatilität ein Problem dar, da das digitale Objekt nicht fixiert werden kann. Problematisch ist zudem, dass ein digitales Objekt nicht nachvollziehbar manipuliert werden kann.

Ob ein Objekt echt ist (ist es das, wofür es gehalten wird?), oder ob es als Fälschung oder Irrtum zu bezeichnen ist, wird aufgrund der Untersuchung seiner Überlieferungsgeschichte festgestellt. Wenn die Digitalität eines digitalen Objekts nicht mit der traditionellen Methode untersucht und eine Unterscheidung zwischen Echtheit und Fälschung/Irrtum aufgrund nicht nachvollziehbarer Manipulation nur schwerlich festgestellt werden kann, sind die bestehenden Methoden als ungenügend zu bezeichnen. Damit wird die dritte These insofern teilweise bestätigt, dass eine Untersuchung digitaler Objekte mit der traditionellen Methode zwar möglich ist, aber nur ungenügende Resultate liefern kann. Die These 'Aufgrund der Besonderheit des digitalen Objekts ist davon auszugehen, dass die traditionelle historische Quellenkritik keine oder ungenügende Methoden zur Überprüfung der Echtheit zur Verfügung stellt' wird daher geändert:

T3: Aufgrund der Besonderheit des digitalen Objekts stellt die traditionelle historische Quellenkritik nur ungenügende Methoden zur Überprüfung der Echtheit zur Verfügung.

3.3. Objekt, Forschungsressource, Quelle

Mit der Einführung von Begriffen aus der Informationstechnik ergeben sich zusätzliche Definitionsprobleme. Quelle wird von Informationstechnikern und Geisteswissenschaftlern unterschiedlich verwendet, weshalb in der Kommunikation untereinander der Begriff geklärt oder ein anderer verwendet werden sollte. Die Geschichtswissenschaft verwendet hauptsächlich Texte als Quelle ihrer Forschung. Die elektronische Datenverarbeitung hat mit digitalen Objekten eine neue Quellenart mit neuen Quellentypen hervorgebracht, die seit der Jahrtausendwende mit der Ausbreitung des Internets und der damit erfolgten Verbreitung von

Audio-, Video- und Bilddaten enorme quantitative Ausmasse erreichte.²⁷⁶ Digitale Objekte als grafische oder dynamisch-grafische, mehrdimensionale Darstellungen sind in der Wissenschaft mittlerweile etabliert. Wird die Hauptaussage aber nun vermehrt visualisiert dargestellt und knapp textlich erklärt, nimmt die Bedeutung des Textes als Leitdarstellungsform ab. Auch stellt sich bei digitalen Objekten die Frage, was denn überhaupt Text ist:

*"The Code is not the Text (unless it is the Text)."*²⁷⁷

Der Literaturwissenschaftler John Cayley spricht hier das Problem an, dass es sich bei digitalen Objekten stets um Text in der Form von Programmcode oder Bitcode handelt.

Eva Pfanzelter betonte an der Tagung *.hist2011 – Geschichte im digitalen Wandel*²⁷⁸, dass sich Historiker nicht mehr lange hauptsächlich mit standardisierten Quellen in Bibliotheken oder Datenbanken beschäftigen werden.²⁷⁹ Die Verwendung von digitalen Daten mit all ihren verschiedenen und kombiniert auftretenden Formaten wird immer häufiger nötig. Es sollte für diese Quellenart besser der Begriff *digitale Forschungsressource* verwendet werden, um ihrer Vielfalt Rechnung zu tragen. Digitale Quellen werden oft nicht in den klassischen, von der Forschung konsultierten Archivinstitutionen gesammelt oder vorgehalten.

Die von Pfanzelter angeregte Unterscheidung zwischen analoger Quelle und digitaler Ressource basiert auf einer notwendigen informationstechnischen Präzisierung. Unter *Source* (Quelle) werden in der Informatik zwei verschiedene Dinge verstanden: einerseits die Datenherkunft, die beispielsweise bei der Verschiebung einer Datei von einem Ort (source) an einen anderen (target) verwendet wird. Andererseits wird bei der Programmierung vom Source Code, dem Quellcode gesprochen, der im Klartext als lesbares Programm vor der Compilierung vorliegt. Die Ressource ist dagegen informationstechnisch eindeutig ein Betriebsmittel des Datenverarbeitungssystems, also alle binär verwalteten Objekte, wie zum Beispiel Dateien, Daten, Betriebssystemkommandos, Threads oder Warteschlangen.²⁸⁰ Informationstechnisch ist deshalb bei digitalen Objekten von Ressourcen und im Zusammenhang mit wissenschaftlicher Forschung von Forschungsressourcen zu sprechen.

²⁷⁶ Die Datenmengen werden mit dem *Internet der Dinge* exponentiell steigen. Siehe dazu Studie von EMC²: Turner, Vernon; Gantz, John F.; Reinsel, David u. a.: White Paper: The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things, 04.2014. White Paper unter URL: <http://idccdocserv.com/1678>, Stand: 02.10.2014; Präsentation unter URL: <https://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>, Stand: 06.05.2018.

²⁷⁷ Cayley, John: The Code is not the Text (unless it is the Text), electronic book review, 10.09.2002, URL: <http://www.electronicbookreview.com/thread/electropoetics/literal>, Stand: 25.08.2014.

²⁷⁸ Tagungsorganisation durch clio-online.de; Humboldt-Universität zu Berlin, 14.-15. September 2011.

²⁷⁹ Pfanzelter, Eva: Quellenkritik vor der Zerreißprobe? Vom kritischen Umgang mit digitalen Ressourcen. Video/Abstract, Dokumentation, *.hist2011*, 15.09.2011, URL: [http://www2.hu-berlin.de/historisches-forschungsnetz/tagung/index.php?conference=hist2011&schedConf=index&page=pages&op=view&path\[\]=documentation%2Fpfanzelter](http://www2.hu-berlin.de/historisches-forschungsnetz/tagung/index.php?conference=hist2011&schedConf=index&page=pages&op=view&path[]=documentation%2Fpfanzelter), Stand: 19.12.2014.

²⁸⁰ Zur Definition von Quelle und Ressource (Betriebsmittel) siehe: Fischer; Hofer: Lexikon der Informatik, 2011.

Ressource kann aber den Begriff *Quelle* nicht ersetzen, weil im Deutschen unter *Ressource* einerseits ein natürliches Produktionsmittel für die Wirtschaft und andererseits ein Hilfsmittel, eine Hilfsquelle, eine Reserve oder ein Geldmittel verstanden wird. In der Wissenschaft wird unter *Ressource* ein Hilfsmittel verstanden, um zur *Quelle* zu gelangen oder um diese, beispielsweise in der Geschichtswissenschaft, überhaupt lesen zu können. Die wissenschaftliche *Ressource* entspricht nicht der wissenschaftlichen *Quelle*.

Obwohl der Begriff der digitalen *Ressource* für digitale Objekte präziser wäre, ist die Einführung des Begriffs problematisch. In der Wissenschaft ist der Begriff 'Quelle' zu stark verankert, als dass diese sprachliche Unterscheidung durchgesetzt werden könnte. Forschende können auch weiterhin den Begriff *Quelle* für alle Arten von Objekten verwenden, auch für digitale. Die Unterscheidung zwischen *Quelle* und *Ressource* kann als sprachliche Finesse angesehen werden, die allerdings dann nötig ist, wenn Geisteswissenschaftler mit Informatikern kommunizieren.

3.4. Wissenschaftliches Arbeiten

Parallel zur traditionellen Methode in der Geschichtswissenschaft entwickelten sich moralische und methodische Ansprüche für das wissenschaftliche Arbeiten. Wenn Wissenschaftler diese Ansprüche nicht erfüllen und gegen die teils ungeschriebenen Regeln verstossen, kann nicht nur ihre Reputation leiden, sondern sie können sogar ihre Anstellung oder ihren Auftrag verlieren.²⁸¹ Mit den heute zur Verfügung stehenden Mitteln ist die Entdeckung von Verfehlungen gegen die Regeln einfacher geworden, was unter Umständen ein politisches, wirtschaftliches oder rechtliches Nachspiel haben kann. Rechtliche Konsequenzen können Abmahnungen oder Unterlassungsklagen sein, wirtschaftliche Konsequenzen können die Abführung von Tantiemen oder die Vernichtung aller Exemplare des Plagiats sein. Mittlerweile werden Masterarbeiten und Dissertationen systematisch auf Plagiate untersucht.²⁸² Die Texte werden mit der Datenbank einer Software wie turnitin oder Docoloc²⁸³ abgeglichen und die verdächtigen Passagen werden anschliessend angezeigt.²⁸⁴ Ein Verdacht muss in jedem Fall noch durch eine Fachperson überprüft werden, da die Softwares nur identische

²⁸¹ Laut *The Korea Herald* sind in Südkorea etwa 200 Professoren von 50 Universitäten befragt worden, die des Plagiarismus verdächtigt werden. Schuldige verlieren sehr sicher ihre Stelle und können zu bis zu fünf Jahre Haft sowie bis zu 50 Millionen Won (ca. CHF 40'000.-) verurteilt werden. Hyun-ju, Ock: 200 professors to be indicted over brazen copyright breach, Social Affairs, The Korea Herald, 24.11.2015, URL: <http://www.koreaherald.com/view.php?ud=20151124001042>, Stand: 26.06.2016.

²⁸² Debora Weber-Wulff ist im deutschsprachigen Raum führend in der Forschung zu Plagiarismus und betreibt ein eigenes Themenportal. Siehe URL: <http://plagiat.htw-berlin.de/>.

²⁸³ turnitin (URL: <http://turnitin.com/>), Docoloc (URL: <http://www.docoloc.de/>).

²⁸⁴ Eine kollaborative Plagiatsdokumentationsplattform ist bspw. VroniPlag (URL: <http://de.vroniplag.wikia.com/>). Suchmaschinen für die Suche nach plagiierten Bildern sind bspw. Plaghunter (URL: <http://www.plaghunter.com/>) oder TinEye (URL: <http://www.tineye.com/>).

Textstellen finden können.²⁸⁵ Zu spät erkannte Plagiate werfen nicht nur ein schlechtes Licht auf den Verfehlenden, sondern auch auf die Institution, die den Titel verliehen hat. Beispielsweise hat die Universität Bayreuth dem damals amtierenden deutschen Bundesminister der Verteidigung Karl-Theodor von Guttenberg aufgrund bewiesener Plagiatsvorwürfe den Dokortitel aberkannt, der darauf zurücktrat.²⁸⁶ Auch die Universität musste sich Vorwürfe gefallen lassen, wieso das Plagiat nicht erkannt wurde.

Wissenschaftliches Arbeiten lernen Studierende in den ersten Semestern des Studiums, damit ihre Abschlussarbeiten alle formalen und inhaltlichen Kriterien ihres Wissenschaftszweigs erfüllen. Werden zunächst bei Studienarbeiten nur die größten Unterlassungen geahndet, steigen die Ansprüche bei Abschluss- oder Diplomarbeiten und weiter bei Dissertationen oder Fachpublikationen aller Art. In Handbüchern, Einführungen zum Studium, rechtsverbindlichen Erlassen der Universitäten und auf diversen Webseiten werden verschiedene, aber wenig einheitliche Kriterien des wissenschaftlichen Arbeitens beziehungsweise der guten wissenschaftlichen Praxis aufgeführt. Leider werden in vielen solchen Studienhilfen wenig aussagekräftige Kriterien wie die Forderung nach einer wissenschaftlichen Terminologie aufgeführt.²⁸⁷ Detaillierte Erklärungen oder eine entsprechende Veranstaltung sind unerlässlich, damit die Studenten profitieren und die Ansprüche durchgesetzt werden können.

Zu den wichtigsten und oft genannten Kriterien des wissenschaftlichen Arbeitens gehören eine klare Fragestellung, kritische Würdigung der Quellen, eigenständiges Arbeiten, systematisch und methodisch korrekte Bearbeitung des Themas, logischer Aufbau und Gliederung, prägnanter und korrekter Ausdrucksstil, klare Argumentation, Verwendung präziser Definitionen sowie die mit einem hohen Stellenwert versehene, ausführliche Dokumentation. Darunter wird hauptsächlich die Gewährleistung der Nachvollziehbarkeit verstanden, insbesondere die Angaben, welche Primär-, Sekundär- oder Tertiärquellen in welcher Art und Weise verwendet wurden und vollständiger, korrekter Zitierung aus der Quelle.

Unter Nachvollziehbarkeit werden wiederum zwei unterschiedliche Kriterien verstanden: Einerseits ist eine verwendete Quelle stets zu zitieren oder nachzuweisen, denn eine Unterlassung wird als Plagiat gewertet. Andererseits sollte eine verwendete Quelle durch eine Drittperson selbständig (wieder) aufgefunden, begutachtet und kritisiert werden können. Ist eine

²⁸⁵ Im letzten von Debora Weber-Wulff durchgeführten Softwaretest (2013) werden nur drei Produkte als 'teilweise nützlich' bezeichnet. Siehe URL: <http://plagiat.htw-berlin.de/software/2013-2/>.

²⁸⁶ Die 'Guttenberg-Affäre' beschäftigte die Deutsche (Bildungs-)Politik 2011 monatelang. Siehe dazu den die Affäre einleitenden Artikel: Preuss, Roland: Guttenbergs Doktorarbeit: Summa cum laude? – «Mehr als schmeichelhaft», in: [sueddeutsche.de](http://www.sueddeutsche.de/politik/guttenbergs-doktorarbeit-summa-cum-laude-mehr-als-schmeichelhaft-1.1060779), 16.02.2011. URL: <http://www.sueddeutsche.de/politik/guttenbergs-doktorarbeit-summa-cum-laude-mehr-als-schmeichelhaft-1.1060779>, Stand: 03.10.2014.

²⁸⁷ Siehe bspw.: Diederich, Jörg; Müller-Rowold, Boris: Kriterien für wissenschaftliche Arbeiten, Text, Stiftung Universität Hildesheim, 16.09.2010, URL: <http://www.uni-hildesheim.de/index.php?id=2239>, Stand: 30.11.2012.

unabhängige quellenkritische Überprüfung nicht möglich, so mindert sich die Aussagekraft der Quelle. Sie kann oder sollte deshalb unter Umständen besser nicht verwendet werden, da sich der Autor sonst dem Vorwurf der Behauptung oder Fälschung aussetzt. Schematisch kann die von mir als *Dreieck des wissenschaftlichen Belegs* bezeichnete Nachvollziehbarkeit von Quellen wie folgt dargestellt werden:

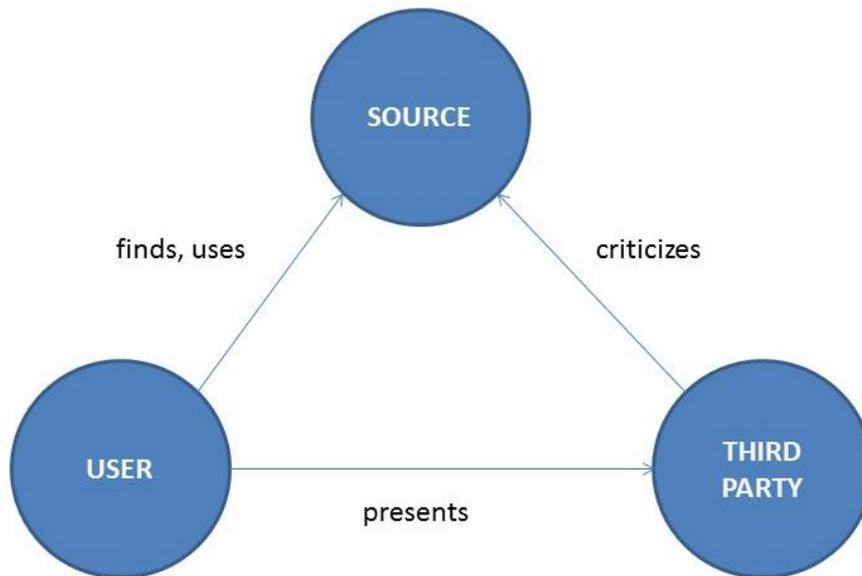


Abbildung 4: Schema 'Dreieck des wissenschaftlichen Belegs' (Föhr)

Das wissenschaftliche Arbeiten gleicht der juristischen Beweisführung, die der Autor zu erbringen hat. Eine mittels Anmerkungsapparat und Literaturverzeichnis überprüfbare Argumentation wird als Kriterium wissenschaftlicher Arbeit explizit gefordert, denn sie legt offen, "auf welchen Quellen die Überlegungen des Autors oder der Autorin beruhen, und machen die Texten nicht nur nachvollziehbar, sondern ermöglichen auch eine Kontrolle der Forschungsergebnisse."²⁸⁸ Die Nachvollziehbarkeit ist ein eminent wichtiges Kriterium, denn "ein Text ist dann ein guter wissenschaftlicher Text, wenn er erstens genau ist; wenn er zweitens seine Grundlagen und Ergebnisse für seine Leserinnen nachprüfbar macht (denn *dafür* sind Anmerkungen da), und drittens: Wenn er neue Informationen zur Verfügung stellt."²⁸⁹ Wird demnach eine Argumentation nicht belegt, steht sie auf wackligen Füßen und dem Verfasser kann der Vorwurf des unwissenschaftlichen Arbeitens oder unter Umständen sogar des Plagiats gemacht werden.

Wissenschaftliche Forschung ist kollaborativ und verlangt nach Einhaltung der Formalitäten. Das Verschweigen von Quellen wird als Verstoss gegen wissenschaftliche Integrität und

²⁸⁸ Lengwiler: Praxisbuch Geschichte, 2011, S. 230-233, hier S. 231.

²⁸⁹ Groebner, Valentin: Welches Thema? Was für eine Art Text? – Vorschläge zum wissenschaftlichen Schreiben 2009 ff., in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 16. Kursiv im Original.

damit der Interessen der Wissenschaft angesehen.²⁹⁰ Dies nimmt nicht nur den Verfasser in die Pflicht, sondern auch die Trägerinstitution, bei der eine Arbeit eingereicht wurde. Sie hat bei einem begründeten Verdacht umgehend ein entsprechendes Verfahren einzuleiten.²⁹¹

3.5. Beurteilung

*"Um es klarzustellen: Die Quellenkritik muss nicht neu erfunden werden. Doch die fehlenden Bezugspunkte und Orientierungshilfen erschweren sie."*²⁹²
(Hodel, 2005)

Zwei Erkenntnisse können aufgrund obiger Ausführungen gewonnen werden:

Erstens können digitale Objekte zwar mit der traditionellen Methode überprüft werden, aber der Prozess ist aufgrund der Eigenschaften *genuin digital*, *Volatilität* und *nicht nachvollziehbare Manipulierbarkeit* digitaler Objekte anzupassen. Wenn die Zahl dieser Art von Quellen sehr gering wäre, könnten auch die damit zusammenhängenden Probleme marginalisiert oder deren Lösung durch eine Hilfswissenschaft erarbeitet werden. Da aber digitale Objekte eine riesige und ständig zunehmende Masse an forschungsrelevanten Daten und Informationen darstellen, ist der Prozess der traditionellen Methode anzupassen.

Die Grundsätze des wissenschaftlichen Arbeitens verlangen nach eindeutigen Quellennachweisen, die unabhängig vom Autor überprüft werden können müssen. Dies ist bei digitalen Objekten aufgrund ihrer Volatilität und Manipulierbarkeit nicht ohne zusätzlichen Aufwand möglich. Das *genuin Digitale* bedingt digitale Infrastrukturen, in denen digitale Objekte dauerhaft, authentisch und vor allem technisch lesbar gesichert werden können. Dieser Vorgang ist bei digitalen Objekten unumgänglich und wird als *Objekt- oder Quellensicherung*²⁹³ bezeichnet. Da die Quellensicherung bei jeder potenziellen digitalen Quelle vorzunehmen ist und dieser Prozess neue, insbesondere Medien- und Informatikkompetenzen voraussetzt, ist die Quellensicherung als neuer Schritt in der historisch-kritischen Methode zu definieren. Professionelle und gesetzlich verankerte Infrastrukturen, die eine Sicherung eines beliebigen

²⁹⁰ Siehe: Akademien der Wissenschaften Schweiz (Hg.): Wissenschaftliche Integrität. Grundsätze und Verfahrensregeln, 2008, S. 20 f. URL: http://www.snf.ch/de/derSnf/forschungspolitische_positionen/wissenschaftliche_integritaet/Seiten/default.aspx.

²⁹¹ Bspw. sind die an der Universität Zürich Lehrenden verpflichtet, erhebliche Plagiate dem Universitätsanwalt zu melden, der bei erhärtetem Verdacht ein Disziplinarverfahren einleitet (URL: <http://www.lehre.uzh.ch/plagiate.html>). An anderen Universitäten wird ähnlich vorgegangen, bspw. an der University of Oxford (GB) (URL: <http://www.ox.ac.uk/students/academic/guidance/skills/plagiarism>).

²⁹² Hodel, Jan: Historische Online-Kompetenz. Überlegungen zu einem hybriden Kompetenzmodell, in: Haber, Peter; Epple, Angelika (Hg.): Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0, Zürich 2004 (Geschichte und Informatik 15), S. 154.

²⁹³ Analog zu den nicht ganz korrekten, aber etablierten Begriffen *Quellenkunde* und *Quellenkritik* wird im Folgenden der Begriff *Quellensicherung* verwendet.

digitalen Objekts erlauben, sollten zur Verfügung gestellt werden. Nur damit kann sichergestellt werden, dass die erwähnten Probleme auf ein Mindestmass reduziert werden können.

Somit sind in der historisch-kritischen Methode neu folgende *sechs Schritte* iterativ durchzuführen:

1. Heuristik
2. Quellenkunde
3. *Quellensicherung*
4. Quellenkritik
5. Interpretation
6. Darstellung

Zweitens kann zwar auf den Grundsätzen und den Fragen der traditionellen Methode aufgebaut werden, aber digitale Objekte erfordern zusätzliche, neue Untersuchungsmethoden. Da digitale Objekte nur durch die Verwendung von Informationstechnik entstehen und genutzt werden können, sind entsprechende Verfahren aus diesem Wissenschaftszweig zu übernehmen und anzuwenden.

Durch die massive Verwendung und Verbreitung von digitalen Objekten hat sich der Geschichtswissenschaftler zusätzliche Kompetenzen in verschiedensten Wissenschaftszweigen anzueignen, insbesondere der Informationswissenschaft, der Archivwissenschaft, der Bibliothekswissenschaft und der Informatik.²⁹⁴ Diese Fachgebiete werden aufgewertet, indem sie nicht mehr nur als Hilfswissenschaften für die Geschichtswissenschaft angesehen werden, sondern ihre Grundlagen fest in die historisch-kritische Methode verankert und ständig aktualisiert werden, damit zukünftig eine Quellenkritik des Digitalen erfolgreich durchgeführt werden kann. Wie bisher werden bei der Untersuchung einer Quelle nicht alle Methoden, Verfahren oder Prozessschritte zur Anwendung kommen, jedoch benötigen die Geschichtswissenschaftler das Rüstzeug, digitale Objekte kritisch beurteilen zu können.

Die Fragen vier und fünf, 'Können digitale Objekte mit der historisch-kritischen Quellenkritik überprüft werden?' und 'Muss für digitale Objekte eine neue historisch-kritische Methode erarbeitet werden oder können digitale Objekte mit der bisherigen überprüft werden?' können aufgrund obiger Ausführungen beantwortet werden:

Digitale Objekte können mit der historisch-kritischen Quellenkritik teilweise, aber nicht vollständig überprüft werden. Eine neue Methode ist zwar nicht nötig, aber es sind der Prozessschritt der Quellensicherung und neue Fragen und Arbeitsmethoden, insbesondere aus der Informatik einzuführen.

²⁹⁴ Dazu auch: Haber: digital past (unv. Habil.), 2009, S. 194.

Damit ist auch die vierte These bestätigt:

Ein digitales Objekt benötigt zur quellenkritischen Überprüfung (auch) Methoden aus der Informatik.

3.6. Zusammenfassung (Fragen und Thesen)

Die traditionellen Prozessschritte Heuristik – Quellenkunde – Quellenkritik – Interpretation – Darstellung sind weiterhin zu verwenden, es muss keine neue Methode entwickelt werden. Allerdings sind digitale Objekte in geeigneten digitalen Infrastrukturen zu sichern, weil sie aufgrund ihrer Volatilität und Manipulierbarkeit nicht dauerhaft und authentisch verwendet werden können. Dazu sind digitale Infrastrukturen nötig, in denen die digitalen Quellen im neuen Prozessschritt ‚Quellensicherung‘ gespeichert und Dritten zur kritischen Beurteilung zur Verfügung gestellt werden. Nur damit können die Gebote des wissenschaftlichen Arbeitens, im Speziellen die Nachweispflicht von verwendeten Quellen, erfüllt werden.

Die dritte Frage ‚Welche Probleme entstehen durch die Eigenschaften eines digitalen Objekts bei der historisch-quellenkritischen Überprüfung (im Gegensatz zu physischen Objekten)?‘ ist beantwortet:

F3: Die genuine Digitalität kann mit bisherigen Methoden nicht untersucht werden, sondern bedarf Methoden aus der Informationstechnik, weil unter anderem für die Entstehung, Verarbeitung und Darstellung eines digitalen Objekts immer ein digitales Informationssystem benötigt wird. Ebenso stellt die mit der Digitalität verbundene Volatilität ein Problem dar, da das digitale Objekt nicht fixiert werden kann. Problematisch ist zudem, dass ein digitales Objekt nicht nachvollziehbar manipuliert werden kann.

Die dritte These ‚Aufgrund der Besonderheit des digitalen Objekts ist davon auszugehen, dass die traditionelle historische Quellenkritik keine oder ungenügende Methoden zur Überprüfung der Echtheit zur Verfügung stellt‘ wurde abgeändert:

T3: Aufgrund der Besonderheit des digitalen Objekts stellt die traditionelle historische Quellenkritik nur ungenügende Methoden zur Überprüfung der Echtheit zur Verfügung.

Die Fragen vier und fünf, ‚Können digitale Objekte mit der historisch-kritischen Quellenkritik überprüft werden?‘ und ‚Muss für digitale Objekte eine neue historisch-kritische Methode erarbeitet werden oder können digitale Objekte mit der bisherigen überprüft werden?‘ wurden ebenfalls beantwortet:

F4/5: Digitale Objekte können mit der historisch-kritischen Quellenkritik teilweise, aber nicht vollständig überprüft werden. Eine neue Methode ist zwar nicht nötig, aber es sind der Prozessschritt der Quellensicherung und neue Fragen und Arbeitsmethoden, insbesondere aus der Informatik einzuführen.

Damit ist auch die vierte These bestätigt:

T4: Ein digitales Objekt benötigt zur quellenkritischen Überprüfung (auch) Methoden aus der Informatik.

Des Weiteren wurde festgestellt, dass ein digitales Objekt besser als *Ressource* und nicht als *Quelle* zu bezeichnen ist. In wissenschaftlichen Forschungsprojekten sollte der Begriff *digitale Forschungsressource* verwendet werden. Weil aber der Begriff 'Quelle' in der Wissenschaft zu stark verankert ist, wird anstelle von *Ressource* der Begriff *digitale Quelle* verwendet. In der Kommunikation mit Informationstechnikern sollte jedoch zwischen der informationstechnischen *Quelle* (*Source*) und der geisteswissenschaftlichen *Quelle* (*Ressource*) unterschieden werden.

4. Digitalmedien

*"L'archiduc François Ferdinand assassiné!"*²⁹⁵
(Léon Vivien, 28.06.1914 auf Facebook)

Die Aussage des Zitates, dass der Kronprinz Franz Ferdinand ermordet wurde, erstaunt, denn sie wurde am 28.06.1914 auf der erst 2004 gegründeten *Social Media*-Plattform Facebook publiziert.²⁹⁶ Der Grund für diese Unstimmigkeit ist schnell gefunden: Das Musée de la Grande Guerre du Pays de Meaux hat in Zusammenarbeit mit der Werbeagentur DDB (beide Paris) ein Facebook-Profil des fiktiven Franzosen und späteren Soldaten Léon Vivien erstellt, um dessen Kommunikation so darzustellen, wie wenn Vivien während des Ersten Weltkriegs nicht Briefpapier, sondern Facebook zur Verfügung gehabt hätte.²⁹⁷ Die vom Musée de la Grande Guerre initiierte Übertragung von einem physischen Medium in moderne Facebook-Posts war sehr erfolgreich, denn das Profil von Vivien erhielt über 66'000 Likes und das Museum konnte eine markante Zunahme seiner Likes verzeichnen.²⁹⁸

Das Beispiel wirft mehrere Fragen auf: Wie ist der Medienbegriff und wie sind die damit zusammenhängenden Begriffe wie *Neue Medien* und *Social Media* zu definieren? Hat ein Medienwandel, -bruch oder -umbruch stattgefunden (oder findet er statt)?

In diesem exkursartigen Kapitel werden zuerst Arbeitsdefinitionen für die Begriffe Medien und Digitalmedien festgelegt und die Begriffe Neue Medien und Social Media werden diskutiert. Anschliessend werden die veränderte Mediennutzung thematisiert und im letzten Kapitel wird der Frage nachgegangen, ob es ein Digitales Zeitalter gibt. Es würde an dieser Stelle zu weit führen, den aktuellen medientheoretischen Diskurs, umfangreiche medienkritische Analysen oder die Diskussion um den noch nicht geklärten Medienbegriff umfassend darzustellen, da diese weder im Fokus der Fragestellung noch der Geschichtswissenschaft liegen.²⁹⁹

²⁹⁵ URL: <https://www.facebook.com/leon1914>.

²⁹⁶ Zu Social Media siehe Kap. 4.2. Neue Medien und Social Media. Das Unternehmen Facebook wurde 2004 gegründet (URL: <https://www.facebook.com/>) und stellt die zurzeit erfolgreichste Social-Media Plattform im Internet zur Verfügung.

²⁹⁷ Das Musée de la Grande Guerre du Pays de Meaux hat das Projekt in Zusammenarbeit mit dem Historiker Jean-Pierre Verney durchgeführt.

²⁹⁸ Post: Beitrag. Likes: In Facebook können Mitglieder Personen, Institutionen oder Beiträge (Posts) mit einem 'Like' markieren und damit ihre Zustimmung ausdrücken. Anzahl Likes Stand 20.11.2014.

²⁹⁹ Siehe zur Kritik an Medientheorien und -theoretikern bspw.: Lovink, Geert: Zero comments Elemente einer kritischen Internetkultur, Bielefeld 2008 (Kultur- und Medientheorie), S. 129-145. Lovink plädiert für einen Wechsel von der Medientheorie zur Medienphilosophie, die "Möglichkeitenräume" (S. 138 und 145) eröffnen würde.

4.1. Medien – Digitalmedien

Oft werden die Begriffe *Digitalmedien* oder *Neue Medien* verwendet, wobei nicht klar ist, was unter Medium³⁰⁰ zu verstehen ist. In älteren Theorien zu Kommunikation und Medien wird vor allem auf die Materialität von Medien Bezug genommen. Vor allem Claude Shannon³⁰¹, Harold Innis³⁰² Herbert Marshall McLuhan³⁰³, Vilém Flusser³⁰⁴ und Neil Postman³⁰⁵ haben die Grundlagen der Medientheorie geschaffen, definieren den Begriff technisch, aber jeweils anders. Für jüngere Medientheoretiker greift der technische Medienbegriff zu kurz, denn ein Medium umfasst weit mehr als Kommunikation, Mittel oder Information. Medien müssen "zugleich als technisch basierte Zeichensysteme, arbeitsteilig verfahrenende Organisationen und als Institutionen (Normen- und Regelsysteme) mit jeweils spezifischem Leistungsvermögen für andere soziale Systeme und Funktionen für die Gesellschaft begriffen werden."³⁰⁶

Was ein Medium ist und welche Eigenschaften es hat, sind die zwei zentralen Fragen in Definitionsversuchen. Allerdings ist die Unterscheidung von Medien und Nicht-Medien äusserst schwierig und wird kaum je zu einem Konsens führen, wie Marcus Burkhardt feststellte.³⁰⁷ Selbst die Unterscheidung von einzelnen Medien ist kaum möglich. Eine Zeitung ist zwar als Medium anerkannt, stellt aber gleichzeitig ein Metamedium dar, das die Medien Papier, Schrift, Sprache, Bild, Druck und Druckerschwärze kombiniert.³⁰⁸

Der Medienbegriff wird sehr unterschiedlich aufgefasst und sehr heterogen verwendet³⁰⁹, aber "fatalerweise meinen die meisten, sie meinten das Gleiche, wenn sie den Begriff Medium verwenden."³¹⁰

³⁰⁰ Zum Medienbegriff siehe bspw.: Schwabe: Historisches Lernen im World Wide Web, 2012, S. 78-80.

³⁰¹ Shannon, Claude Elwood: A Mathematical Theory of Communication, in: The Bell System Technical Journal (27), 1948, S. 379-423. Siehe auch: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 338-342.

³⁰² Innis, Harold A: Empire and Communications, London 1950; Innis, Harold A: Bias of Communication, Toronto 1951. Zu Innis (kurz): Spahr, Angela: Magische Kanäle. Marshall McLuhan, in: Kloock, Daniela; Spahr, Angela (Hg.): Medientheorien: eine Einführung, Paderborn 2012 (UTB. Medienwissenschaft, Kommunikationswissenschaft 1986), S. 47.

³⁰³ McLuhan, Marshall: The Gutenberg Galaxy: The Making of Typographic Man, London 1962; McLuhan, Marshall (Hg.): Understanding Media: The Extensions of Man, New York 1964; McLuhan, Marshall; Fiore, Quentin; Agel, Jerome: The medium is the message: An inventory of effects, New York 1967 (Bantam Books).

³⁰⁴ Zu Vilém Flusser siehe: Rosner, Bernd: Telematik. Vilém Flusser, in: Kloock, Daniela; Spahr, Angela (Hg.): Medientheorien: eine Einführung, Paderborn 2012 (UTB. Medienwissenschaft, Kommunikationswissenschaft 1986), S. 77-98.

³⁰⁵ Postman, Neil: Technopoly: The surrender of culture to technology, New York 1992.

³⁰⁶ Beck, Klaus: Medien, in: Lexikon Kommunikations- und Medienwissenschaft, Wiesbaden 2013 (Studienbücher zur Kommunikations- und Medienwissenschaft), S. 201-202, in Anlehnung an Ulrich Saxer (1931-2012).

³⁰⁷ Siehe dazu: Burkhardt: Digitale Datenbanken, 2015, S. 15 und 28.

³⁰⁸ Nach: Ebd., S. 32.

³⁰⁹ Siehe auch die Bestimmungs- und Definitionsversuche in: Münker, Stefan; Roesler, Alexander (Hg.): Was ist ein Medium?, Frankfurt a.M. 2008 (Suhrkamp Taschenbuch Wissenschaft).

³¹⁰ Ebd., S. 11.

Aufgrund dieser Definitionsproblematik wird im Folgenden eine vereinfachte Arbeitsdefinition des Medienbegriffs verwendet:³¹¹

Ein Medium ist ein Vermittlungssystem, das von Menschen für Menschen zur beabsichtigten oder unbeabsichtigten Übermittlung von Informationen verwendet wird, wobei sich Medium und Kommunikation gegenseitig bedingen.

Aus dieser allgemeinen Arbeitsdefinition für Medium kann eine Arbeitsdefinition für Digitalmedien abgeleitet werden:

Ein Digitalmedium ist ein Vermittlungssystem, das von Menschen für Menschen zur beabsichtigten oder unbeabsichtigten digitaler Übermittlung von Informationen verwendet wird, wobei sich Medium und Kommunikation gegenseitig bedingen.

Als Vermittlungssystem werden hauptsächlich Computer eingesetzt. Aber Computer "entziehen sich nicht nur einer eindeutigen Funktionszuschreibung, vielmehr kann nahezu alles, was mit, durch und in Computern getan wird, auf unterschiedliche Weise getan werden."³¹² Computer sind programmierte Maschinen, die einem Nutzer mit Befehlen, Algorithmen und Programmen sowie Datenformaten, -strukturen und -modellen vielfältige Möglichkeiten bieten, eine Tätigkeit auszuführen. Für Marcus Burkhardt ist es die wahrscheinlich grösste Herausforderung der Medientheorie des Digitalen, dass der Gebrauch des Computers sehr flexible Handlungsoptionen bietet, wie mediale Konstellationen artikuliert, bedient, verarbeitet und distribuiert werden können. Das elektronische Informationssystem ist ein Vermittler zwischen einer Nutzeroberfläche³¹³ für menschliches Handeln und der technischen Ausführung im unsichtbaren Innern, die auf eine ganz andere Weise erfolgt. Ein digitales Objekt ist demnach nicht nur auf der Ebene der Oberfläche, sondern gleichzeitig auch in der Tiefe des elektronischen Informationssystems zu verorten und medial zu untersuchen.³¹⁴

Dies verlangt vom Nutzer neue Medien- und Informationstechnikkompetenzen. Das reine Betrachten oder Lesen reicht im Umgang mit digitalen Objekten und elektronischen Informationssystemen nicht mehr aus. Der Nutzer hat beispielsweise im Umgang mit Datenbanken die nötigen formalen Sprachen zu beherrschen und die ihm auf der Oberfläche präsentierten Informationen unter Berücksichtigung der verwendeten Technik zu interpretieren. Damit kann festgestellt werden, dass sich die Medialität von digitalen Objekten von bisher bekannten unterscheidet.

³¹¹ Siehe dazu bspw. Thilloßen, Anne: Schreiben im Netz: Neue literale Praktiken im Kontext Hochschule, Münster 2008 (Medien in der Wissenschaft), S. 49-51; Bentele, Günter; Brosius, Hans-Bernd; Jarren, Otfried (Hg.): Lexikon Kommunikations- und Medienwissenschaft, Wiesbaden 2013 (Studienbücher zur Kommunikations- und Medienwissenschaft), S. 199-227.

³¹² Burkhardt: Digitale Datenbanken, 2015, S. 73.

³¹³ Engl.: Graphical User Interface (GUI).

³¹⁴ Siehe vertiefend: Luhmann, Niklas: Die Gesellschaft der Gesellschaft, Frankfurt a.M. 1998 (Suhrkamp Taschenbuch Wissenschaft), S. 304.

4.2. Neue Medien und Social Media

Der Begriff *Neue Medien* wird zwar in Publikumszeitschriften, Blogs und anderen Beitragsformen gerne und oft verwendet, aber er ist missverständlich und sollte nicht verwendet werden. Ein neues Medium verändert die Strukturen von Wahrnehmung, Intersubjektivität und Kultur.³¹⁵ Es ist ein neues Vermittlungssystem, das bis anhin nicht zur Verfügung stand. Mit der Einführung eines neuen Mediums wird ein bereits bestehendes, neues zu einem alten und darf nicht mehr mit dem Begriff bezeichnet werden. Darauf weist auch der Kommunikationswissenschaftler Klaus Beck hin, denn es wurden "in den 1980er-Jahren Kabelrundfunk, Videotext und Bildschirmtext u.a. [...], in den 1990er-Jahren hingegen digitale Medien"³¹⁶ als neue Medien bezeichnet. Es stellt sich jedoch die Frage, wie gross die Veränderungen vom alten zum neuen Medium respektive wie stark die Anpassung an die neuen Strukturen sein müssen, um ein Medium als *neu* zu bezeichnen. Bei gänzlich neuartigen Medien wie beim Übergang vom Bild zur Schrift ist dies eventuell noch möglich, aber bei einer graduellen Anpassung wie vom handgeschriebenen zum gedruckten Buch ist dies kaum klar festzulegen. Zudem ist bei jedem neu auftretenden Medium die Bedeutung des Begriffs anzupassen, denn das Neue wird nun alt. Diese Problematik unterstreicht der Medientheoretiker Geert Lovink, der meint, dass alles anders wird, nur nicht die neuen Medien.³¹⁷ Weil dem Begriff eine ständige Erneuerung inhärent ist und deshalb keine Abgrenzung möglich ist, sollte grundsätzlich nicht von neuen Medien, sondern von Digitalmedien gesprochen werden.

Unter *Social Media* oder *Social Web* werden partizipative Formate des vom Softwareentwickler Tim O'Reilly 2005 ausgerufenen Web 2.0 verstanden.³¹⁸ Darunter ist die Entwicklungsstufe des World Wide Web zu verstehen, die nach der reinen Informationspräsentation des Web 1.0, das allerdings nie so bezeichnet wurde, nun die aktive Beteiligung der Nutzer ermöglicht: Das Web fungiert als Service-Plattform, die kollektive Intelligenz der Nutzer wird genutzt, Daten und deren Management stehen im Mittelpunkt, neue Formen der Software-Entwicklung wie Crowd-Testing³¹⁹ oder Interoperabilität werden eingeführt etc.³²⁰ *Kollaboration*, *Information* und *Beziehungspflege* stehen folglich im Mittelpunkt und sind wesentlicher Teil von Social Media. Der Konsument ist gleichzeitig Produzent von Inhalten und

³¹⁵ Nach McLuhan. Siehe: Mersch, Dieter: Medientheorien zur Einführung, Hamburg 2013 (Zur Einführung 318, Ed. 3), S. 113.

³¹⁶ Beck: Medien, 2013, S. 202.

³¹⁷ Lovink: Zero comments, 2008, S. 123, in Anlehnung an Paul Valéry.

³¹⁸ Siehe: O'Reilly, Tim: What is Web 2.0?, 30.09.2005, URL: <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>, Stand: 06.05.2018.

³¹⁹ Crowd-Testing: Das Testen eines noch nicht ausgereiften (Software)Produkts durch eine grosse Anzahl von Nutzern. Weiterführend: Kurzidim, Michael: Crowd-Testing: pro Bug 25 Franken, Businesspraxis, Computerworld.ch, 01.12.2014, URL: <http://www.computerworld.ch/businesspraxis/artikel/weg-mit-den-software-bugs-66855/>, Stand: 02.12.2014.

³²⁰ Siehe zur Charakteristik des Web 2.0 bspw.: Nentwich: Cyberscience 2.0, 2012, S. 5-7; Haber: Digital Past, 2011, S. 126 f.

wird somit zum "Prosumer."³²¹ Die Social Media heben sich durch ihre Unmittelbarkeit, die Möglichkeit, Inhalte zu verknüpfen und zu ergänzen, und durch die Interaktion von miteinander in Beziehung stehenden Menschen von den traditionellen Massenmedien ab:

Unter Social Media werden Digitalmedien für die kommunikative Interaktion von miteinander in Beziehung stehenden Menschen verstanden, die Informationen als Prosumer unmittelbar erstellen, teilen und verknüpfen können.

4.3. Veränderte Mediennutzung

"I now have almost totally lost the ability to read and absorb a longish article on the web or in print."³²² (Friedman, 2008)

Die Nutzungsgewohnheiten haben sich in den letzten 20 Jahren massiv verändert. War damals der Besuch einer Bibliothek für eine Recherche unumgänglich, sind Informationen heute mit Smartphones wortwörtlich sofort 'zur Hand'. Nutzer verlangen nach digitalen Inhalten und wollen überall mit ihrem präferierten digitalen Informationssystem darauf zugreifen können.

Das Verarbeiten der Informationen hat sich ebenfalls verändert, insbesondere in der kritischen Beurteilung der aufgefundenen Informationen. David Nicholas hat unter Rückgriff auf mehrere Studien Folgendes aufgezeigt:

"[...] the speed of young people's web searching meant that little time was spent in evaluation, either for relevance, accuracy or authority – they skitter along the digital surface never staying for long in one place or going very deep."³²³

Dies hat der Psychologe Barrie Gunter schon etwas früher festgestellt:

"[...] although young people demonstrate an apparent ease and familiarity with computers, they rely heavily on search engines, view rather than read and do not possess the critical and analytical skills to assess the information that they find on the web."³²⁴

³²¹ Krameritsch: Hypertext schreiben, 2010, S. 84.

³²² Friedman, Bruce: How Google is Changing Our Information-Seeking Behavior, Lab Soft News, 06.02.2008, URL: http://labsoftnews.typepad.com/lab_soft_news/2008/02/how-google-is-c.html, Stand: 20.03.2015.

³²³ Nicholas, David; Rowlands, Ian; Clark, David u. a.: Google Generation II: web behaviour experiments with the BBC, in: Aslib Proceedings 63 (1), 18.01.2011, S. 30. DOI: 10.1108/00012531111103768.

³²⁴ Gunter, Barrie; Rowlands, Ian; Nicholas, David: The Google Generation. Are ICT innovations changing information-seeking behaviour?, Oxford 2009 (Chandos information professional series), S. 153.

Das Problem der oben pauschalisiert angesprochenen *tl;dr-generation*³²⁵ ist, dass sie vom Lesen längerer Textpassagen arbeitstechnisch überfordert sind, weil sie nur Kurztexte gewohnt sind.³²⁶ Selbst bekannte Wissenschaftsblogger und Digital Immigrants wie Bruce Friedman oder Nicholas Carr geben zu, dass auch sie nicht mehr gewohnt sind, längere Passagen zu lesen. Carr meint: "The deep reading that used to come naturally has become a struggle".³²⁷ Er fragt sogar, ob uns Google verblödet, bleibt aber eine klare Antwort schuldig.

*Deep Reading*³²⁸ oder *Close Reading* meinen ein aufmerksames und reflektierendes Lesen, um möglichst alle Facetten eines Objekts zu erfassen. Darunter befinden sich Aspekte wie die Struktur und der Aufbau der Quelle, die genannten Personen, Institutionen oder Organisationen, das Fehlen von Inhalten, Verweise oder thematische Bezüge.³²⁹ Ziel ist, ein grösstmögliches Verständnis davon zu erhalten, um das daraus gewonnene Wissen bei nachfolgenden anwenden zu können. Dem entgegengesetzt hat sich in den letzten Jahren das *Powerbrowsen* oder *Skimming* – eine Kurzform von Querlesen – sowohl bei Digital Natives, als auch bei Digital Immigrants durchgesetzt. Beim Powerbrowsen werden Titel, Inhalte, Zusammenfassungen etc. horizontal gelesen, nur auf der Suche nach "quick wins."³³⁰ Dabei ist die Geduld und Konzentration nach relativ kurzer Zeit erschöpft, denn was nicht auf den ersten oder zweiten Blick ersichtlich ist, wird verworfen und das nächste Objekt wird aufgerufen.³³¹ Die Aufmerksamkeit wird sehr oft auf andere Objekte wie Werbung oder Hyperlinks gelenkt, was einen reflektierten, langgesponnenen Gedanken verhindert.³³² Die digitalen Publikationsformen mit ihren Informationsschnipseln scheinen tatsächlich eine Veränderung der Aufnahmefähigkeit, der Hirntätigkeit und der Verbindungsmuster zu bewirken.³³³ Da aber wissenschaftliches Arbeiten Deep Reading erfordert, ist diese Veränderung negativ zu werten. Wenn Geschichtswissenschaftler nicht mehr fähig sind, ihre

³²⁵ tl;dr: too long; didn't read. In Abwandlung der ursprünglichen Bedeutung von 'Kritik am Text, der unnötig überlang formuliert ist' zur Kritik am Leser, der nicht mehr fähig ist, längere Texte zu lesen. Zudem unterscheidet sich das Lesen von webbasierten Artikeln von jenem physischer Artikel. Siehe dazu bspw.: Baron, Naomi S.: The Plague of tl;dr, in: The Chronicle of Higher Education, 09.02.2015. URL: <http://chronicle.com/article/The-Plague-of-tl-dr/151635/>, Stand: 07.04.2015; Nielsen, Jakob: F-Shaped Pattern For Reading Web Content, Articles, NN/g Nielsen Norman Group, 17.04.2006, URL: <http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/>, Stand: 07.04.2015.

³²⁶ 20 Minuten und Blick am Abend sind zwei Schweizer Gratiszeitungen. Siehe URL: <http://www.20min.ch/> und URL: <http://www.blickamabend.ch/>.

³²⁷ Carr, Nicholas: Is Google Making Us Stupid?, in: The Atlantic, 08.2008. URL: <http://www.theatlantic.com/magazine/archive/2008/07/is-google-making-us-stupid/306868/>, Stand: 06.06.2014.

³²⁸ Der Begriff wurde geprägt von: Birkerts, Sven: The Gutenberg elegies the fate of reading in an electronic culture, Boston (Mass.) [etc.] 1994.

³²⁹ Vgl.: Koller: Geschichte digital, 2016, S. 80.

³³⁰ Carr: Is Google Making Us Stupid?, 2008.

³³¹ Beim bekannten 'Querlesen' wird zwar dieselbe Technik angewendet, allerdings wird dabei das gesamte Objekt durchgesehen.

³³² Nach: Weinberger: Too Big To Know, 2011, S. XII.

³³³ Siehe: Nicholas u. a.: Google Generation II, 2011, S. 44.

Quellen intensiv zu studieren sowie hermeneutisch und epistemologisch zu erfassen, wird die Qualität wissenschaftlicher Arbeiten massiv leiden. Veränderte technische Möglichkeiten beeinflussen die Art und Weise, wie Forschende Informationen und Daten sammeln, verwalten, analysieren und visualisieren.³³⁴ Mit der massenhaften Verwendung von digitalen Objekten haben sich neue Arbeitsweisen etabliert, unter anderem das durch den Literaturwissenschaftler Franco Moretti propagierte *Distant Reading*³³⁵ oder das algorithmische Durchsuchen und Vernetzen von grossen Textkorpora. Durch Letzteres können beispielsweise neue Erkenntnisse über Unterschiede der im Korpus enthaltenen Schriften gewonnen werden. Algorithmische Auswertungen liefern aber bloss oberflächliche Informationen, solange ihnen keine intensive Beschäftigung mit der Entwicklung und dem Inhalt des Korpus vorausgeht. Intensive Beschäftigung heisst wiederum Deep Reading. Distant Reading kann aber die Suche nach Passagen unterstützen, die es anschliessend genauer zu lesen gilt.³³⁶ Distant und close reading müssen nicht als Gegensatz verstanden werden, denn das Springen zwischen beiden Techniken kann sinnvoll sein und zu den gewünschten Ergebnissen führen.³³⁷

Mit der Entwicklung von Programmiersprachen geht die Entwicklung der menschlichen Ausdrucksweise in mündlicher oder schriftlicher Form einher. Seit der Einführung des SMS-Standards³³⁸ bei Mobiltelefonen werden Kurzmitteilungen ausgetauscht, die teilweise nur innerhalb einer Gruppe und im Kontext verständlich sind, wie 'thx 4 pic – rofl – xxx' (thanks for the picture – I am rolling on the floor laughing – kisses) – eine Entwicklung, die auch weitere Kommunikationssysteme wie Twitter oder Facebook betrifft. Solche sprachliche Kommunikation, die teilweise auch mit Visualisierungen wie Smileys oder Emoticons/Emoji kombiniert wird³³⁹, ist Gegenstand verschiedener Forschungsprojekte.³⁴⁰ Vor allem Symbolsysteme wie Emoticons sind bei der Kommunikation wichtig, denn sie können Aussagen

³³⁴ Vgl. bspw.: Collins, Ellen; Jubb, Michael: How do Researchers in the Humanities Use Information Resources?, in: *Liber Quarterly* 21 (2), 2012, S. 177; Grey, Anthony J. G.: Jim Gray on eScience: A Transformed Scientific Method, in: Hey, Anthony J. G.; Tansley, Stewart; Tolle, Christin (Hg.): *The fourth paradigm data-intensive scientific discovery*, Redmond, Wash 2009, S. XVII-XXXI. URL: http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf.

³³⁵ Siehe: Moretti, Franco: *Distant reading*, London 2013. Nach Moretti ist distant reading nötig, weil mit deep reading in derselben Zeitspanne viel weniger Text verarbeitet werden kann. Die stetig wachsende Menge an Publikationen macht aber ein Schnelllesen nötig (S. 48 – kursiv im Original): "...we really need is a little pact with the devil: we know how to read texts, now let's learn how *not* to read them."

³³⁶ Vgl.: Koller: *Geschichte digital*, 2016, S. 62.

³³⁷ Vgl.: Gibbs, Fred; Owens, Trevor: *The Hermeneutics of Data and Historical Writing*, in: Dougherty, Jack; Nawrotzki, Kristen (Hg.): *Writing history in the digital age*, Ann Arbor 2013 (*Digital humanities*), S. 162. URL: <https://muse.jhu.edu/book/27633>, Stand: 05.02.2018.

³³⁸ SMS: Short Message System, beschränkt auf 160 Zeichen.

³³⁹ Smiley: standardisierte, grafische Darstellung eines Gesichtsausdrucks, bspw. ☺, einer Emotion oder eines sonstigen Objekts, bspw. eines Hauses; Emoticon: abstrakte Darstellung eines Smileys aus ASCII-Zeichen, bspw. :-).

³⁴⁰ Publikationen zur (kurz-)sprachlichen Kommunikation: http://www.sms4science.uzh.ch/research/publikationen_en.html.

reduziert wiedergeben (bspw. ein Daumen-hoch-Smiley für 'Ich bin einverstanden'), aber auch zu Missverständnissen führen. Beispielsweise wird das Emoticon mit schnaubender Nase in Europa als Ausdruck von Ärger, in Japan aber von Triumph verstanden.³⁴¹



Abbildung 5: Emoticon 'schnaubender Smiley'; Unicode U+F624 (Unicode Inc.)³⁴²

Die sprachliche Ausdrucksweise in der Realität unterscheidet sich teilweise von jener anonym oder unter einem Pseudonym auftretender Autoren. Auf Online-Plattformen können Meinungen anonym oder unter einem Pseudonym hinterlassen werden, die in einer realen Diskussion in dieser Form eventuell nicht geäußert würden. Aussagen, die früher im kleinen gesellschaftlichen Kreis hinter vorgehaltener Hand gemacht wurden, werden nun (frei) zugänglich und schriftlich veröffentlicht.

Von diesem Informationsaustausch können zwar viele profitieren, er hat aber für die Wissenschaft negative Auswirkungen. Mit dem einfachen Zugang zu vielen Informationen aus nahezu allen Themengebieten wird das professionelle oder wissenschaftliche Wissen durch *Crowd Knowledge*³⁴³ angegriffen und die Urteile ausgewiesener Experten werden infrage gestellt. Eine Person kann sich beispielsweise aufgrund von Symptomen bereits vor der Konsultation eines Arztes in Foren oder auf Webseiten über mögliche Krankheiten informieren.³⁴⁴ Der Patient ist zwar besser informiert, er stellt aber das Fachurteil immer öfter infrage und sein Vertrauen in die Fachperson schwindet. Diese Haltung hat sich auf alle Wissenschaftszweige ausgedehnt. Professionelles Wissen wird hinterfragt, aber unprofessionell beurteilt und mit Gerüchten, Halbwahrheiten oder Verschwörungstheorien vermischt. Die Wissenschaft kommt dadurch in Zugzwang, weil sie sich öffentlich gegen Laienwissen zu

³⁴¹ Siehe auch: Miller, Hannah: Investigating the Potential for Miscommunication Using Emoji, GroupLens, 04.05.2016, URL: <http://grouplens.org/blog/investigating-the-potential-for-miscommunication-using-emoji/>, Stand: 14.04.2016; Miller, Hannah; Thebault-Spieker, Jacob; Chang, Shuo u. a.: "Blissfully happy" or "ready to fight": Varying Interpretations of Emoji, in, Köln 2016. URL: http://grouplens.org/site-content/uploads/Emoji_Interpretation.pdf, Stand: 14.04.2016.

³⁴² Unicode U+F624, URL: <http://unicode.org/emoji/charts/full-emoji-list.html>. Verwendung gemäss Lizenzvereinbarung Unicode Inc., einsehbar unter URL: <http://www.unicode.org/copyright.html#License>.

³⁴³ Crowd knowledge: Das Wissen der Vielen – ähnlich zu crowd wisdom (Weisheit der Vielen); Theorie, die besagt, dass mit steigender Anzahl beteiligter Personen auch das Wissen besser, genauer etc. ist. Theorie aufgestellt von: Galton, Francis: Vox Populi, in: Nature (75), 07.03.1907, S. 450-451. DOI: 10.1038/075450a0. Siehe auch: Surowiecki, James: The wisdom of crowds : why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations, London 2004.

³⁴⁴ Aussage von verschiedenen Ärzten im persönlichen Gespräch.

behaupten hat. Mills Kelly zeigte anhand des Wikipedia-Artikels zu *Donner Party*³⁴⁵ auf, dass neue wissenschaftliche Erkenntnisse wieder gelöscht werden können – Crowd Knowledge übertrumpft wissenschaftliches Wissen.³⁴⁶ Der Soziologe Dominik Schrage nennt dies eine "Entmystifizierung der wissenschaftlich-technischen Expertise"³⁴⁷ und meint, dass sich die Experten und Laien durch die extensive Veröffentlichung und Verbreitung von Wissen einander annähern, wobei Expertenwissen vor allem zur Bestätigung des Laienwissens gefragt ist.³⁴⁸ Die Gefahr besteht allerdings, dass damit nur das Expertenwissen Verbreitung findet, das den Laien argumentativ dient, weil diese keine Auseinandersetzung wünschen. Das Nachvollziehen der Entstehung solcher Texte ist hingegen für Fragen der Dekonstruktion eines Textes interessant. Texte werden hinterfragt und angepasst, damit sie wieder einen Sinn ergaben, nachdem Widersprüche aufgedeckt worden sind.

4.4. Umbruch in ein Digitales Zeitalter

*"The way we understand cultural dynamics is now identical with computational logic. It is not anymore a method, methodology, or tool but, more fundamentally, the way we see and act in the world."*³⁴⁹ (Yuk Hui)

Es hat sich in den letzten Jahren gezeigt, dass wir uns im Umbruch in ein *Digitales Zeitalter*³⁵⁰ respektive im *Computational Turn* befinden. Dieser Umbruch wird von Wissenschaftlern unterschiedlich bezeichnet und charakterisiert, aber nur wenige³⁵¹ verneinen ihn. Beispielsweise spricht Dieter Mersch von einem *Medial Turn*³⁵², Wolfgang Schmale deutet ihn

³⁴⁵ Donner-Party: Gruppe von Amerika-Siedlern, die 1846 gegen Westen zogen und auf der Reise eingeschneit wurden. Ihnen wird Kannibalismus nachgesagt (sic!). Siehe URL: http://en.wikipedia.org/wiki/Donner_party.

³⁴⁶ Der Eintrag in der englischsprachigen Wikipedia zu 'Donner Party' wurde laut der History am 24.04.2002 erstellt. Kelly änderte den ersten Absatz aufgrund inzwischen erfolgter neuer Erkenntnisse am 28.04.2006 (Tkelly7), aufgrund derer der Kannibalismus unter den Mitgliedern dieser Gemeinschaft in Zweifel gezogen wird. Bereits am 1.05.2006 wurde der Absatz wieder geändert, mit leicht abweichendem Sinn. Bis zum 11.03.2014 erfolgten am gesamten Artikel insgesamt 3597 Änderungen von 1339 Usern (viele davon nur mit IP-Adresse, also anonym). Der damals aktuelle Eintrag enthielt die von Kelly eingebrachte Änderung nicht mehr, sondern wies sogar explizit auf den Kannibalismus hin. Siehe: Kelly: Teaching history in the digital age, 2013, S. 99 f.

³⁴⁷ Schrage, Dominik: Flüssige Technokratie, in: Merkur deutsche Zeitschrift für europäisches Denken (Jg. 66, H. 9/10 (H. 760/761)), 2012, S. 820.

³⁴⁸ Ebd., S. 821.

³⁴⁹ Hui, Yuk: The computational turn, or, a new Weltbild, in: Junctures – The Journal for Thematic Dialogue (13), 2010, S. 41.

³⁵⁰ Zeitalter: ein längerer Zeitraum der Geschichte verstanden, "der sich durch die Aus- und Nachwirkung eines bestimmten Ereignisses, einer Persönlichkeit oder einer Idee auszeichnet" (Fuchs, Raab) Synonyme: Abschnitt, Epoche, Periode, Phase. Spanne.Vgl.: Fuchs, Konrad; Raab, Heribert: Zeitalter, Weltalter, Äon, in: Wörterbuch zur Geschichte, München 2002.

³⁵¹ Bspw. spricht nicht von einem Turn oder Umbruch sprechen, sondern meint, es sei schlicht eine neue Perspektive oder ein Additionsvorgang, der in die Geisteswissenschaften Einzug halte. Vgl.: Hagner, Michael: Vor und nach den Digital Humanities. Eine Übersicht, SAGW-Tagung 13, Bern 28.11.2013.

³⁵² Mersch: Medientheorien zur Einführung, 2013, S. 131.

unter Vorbehalt als Medienrevolution³⁵³ und auch für Helmut Schanze ist der Umbruch erkennbar.³⁵⁴ Michael Nentwich hat sich in *Cyberscience 2.0* mit der Rolle der digitalen sozialen Kulturen, die durch das Web 2.0 in der akademischen Welt eingeführt wurden, untersucht.³⁵⁵ Die Cyberscience³⁵⁶ sei Realität und werde sich immer weiter in den Cyberspace verlagern, lautet Nentwichs Fazit. Die Anglistin Cathy Davison ist ähnlicher Ansicht, denn Computational Humanities veränderten die Art und Weise des Forschens, vor allem "the kinds of questions we can ask; and the depth, breadth and detail of the answers we can provide to those questions."³⁵⁷ Der Computer Humanist Willard McCarty meint ergänzend: "Everything changed for computing in the humanities"³⁵⁸. Auch David Berry unterstützt diese Ansichten: "computational techniques are not merely an instrument wielded by traditional methods; rather they have profound effects on all aspects of the disciplines. Not only do they introduce new methods, which tend to focus on the identification of novel patterns in the data as against the principle of narrative and understanding, they also allow the modularisation and recombination of disciplines within the university itself."³⁵⁹ Und ergänzend: "...it is becoming more and more evident that research is increasingly being mediated through digital technology."³⁶⁰ Peter Haber und Martin Gasteiner, bezeichnen den Umbruch (allerdings einschränkend auf wissenschaftliche Arbeitstechniken) als "Paradigmenwechsel von analog zu digital."³⁶¹

Der Paradigmenwechsel oder die Verlagerung (shift) bedeutet demnach für die Wissenschaft neue Diskussionen über Sprachverständlichkeit und damit die Frage, ob Code eine Sprache ist, über Ontologie und damit die Frage, ob die computergenerierte Darstellung das *Seiende* ist oder nur der Code, und über die Art des Einsatzes von Computertechnologie in der Wissenschaft und damit Fragen über das nötige technische Vorwissen sowie über die Gültigkeit von computergenerierten Ergebnissen. Des Weiteren sind Fragen nach Abhängigkeiten,

³⁵³ Siehe: Schmale: Digitale Geschichtswissenschaft, 2010, S. 23.

³⁵⁴ Vgl.: Schanze, Helmut: Gibt es ein digitales Apriori?, in: Böhnke, Alexander; Schröter, Jens (Hg.): Analog/Digital – Opposition oder Kontinuum? Zur Theorie und Geschichte einer Unterscheidung, Bielefeld 2004, S. 67-79.

³⁵⁵ Nentwich: *Cyberscience 2.0*, 2012.

³⁵⁶ *Cyberscience*: Wissenschaft, die sich in der virtuellen Welt bewegt und sich mit ihr beschäftigt.

³⁵⁷ Davidson, Cathy N.: *Humanities 2.0: Promise, Perils, Predictions*, in: Gold, Matthew K. (Hg.): *Debates in the digital humanities*, Minneapolis 2012, S. 479. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.

³⁵⁸ McCarty, Willard: *A Telescope for the Mind?*, in: Ebd. S. 116.

³⁵⁹ Berry, David M.: *The Computational Turn: Thinking About the Digital Humanities*, in: *Culture Machine* 12 (0), 18.02.2011, S. 13. URL: <http://www.culturemachine.net/index.php/cm/article/view/440>, Stand: 14.01.2014.

³⁶⁰ Ebd., S. 1.

³⁶¹ Haber, Peter; Gasteiner, Martin: *Digitale Arbeitstechniken. Eine Einführung*, in: Gasteiner, Martin; Haber, Peter (Hg.): *Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften*, Wien, Köln, Weimar 2010, S. 14. Weitere Ausführungen zu analog/digital siehe: Haber: *digital past* (unv. Habil.), 2009, S. 76-82. Zu Paradigmenwechsel sei auf das Konzept von Kuhn verwiesen: Kuhn, Thomas S.: *The structure of scientific revolutions*, Bd. 2, Chicago 1962 (*International encyclopedia of unified science* 2).

sozialer Interaktion, Privatsphäre oder Viralität zu stellen. Mit der Miniaturisierung und dem Einsatz digitaler Technologien hat sich beispielsweise auch die Schreib- und Schriftkultur verändert. Funktion des Erinnerns, die nun auf einen technischen Träger übertragen werden kann.³⁶²

Der Computational Turn ist nicht nur einfach ein weiterer Turn, sondern eine einschneidende Wende, weil er "die Wertehierarchie der kulturwissenschaftlichen Disziplinen betrifft."³⁶³ Er wirkt sich auch auf das wissenschaftliche Methodeninventar aus. Die computertechnische Durchdringung hat vor der Wissenschaft nicht Halt gemacht, denn der Computer wird in allen Forschungsfeldern als Universalmaschine eingesetzt.³⁶⁴ Neue Arbeitsmethoden haben sich etablieren können, sei dies die einfache Nutzung des Computers als elektronische Schreibmaschine oder die Entwicklung von komplexer Software für die Auswertung von Experimenten, wie beispielsweise vom Conseil Européen pour la Recherche Nucléaire (CERN) in Genf.³⁶⁵ Der Computational Turn verlangt nach einer Anpassung der quellenkritischen Methode im Umgang mit den neuartigen Quellen.

Fragen zur *objektiven* Darstellung der Erkenntnisse mittels Sprache oder Symbolen, beispielsweise aus dem linguistic, interpretive oder poststrukturalistischen Turn, bleiben auch im Computational Turn dieselben und sind wieder zu stellen. Aufgrund der universellen Einsatzmöglichkeiten eines Computers und der erfolgten computertechnischen Durchdringung der Gesellschaft und Wissenschaft kann der Computational Turn als herausragendster Wechsel der letzten drei Jahrzehnte bezeichnet werden, was die dritte Hypothese bestätigt.

Es findet ein Umbruch in ein Digitales Zeitalter statt. Dessen Merkmale sind: Konvergenz der Medien (Bitstrom als universales Medium), Abhängigkeit von Informationssystemen, computerunterstützte Arbeitsprozesse, Produktion von überwiegend digitalen Daten, Verlust von Materialität, ständige Auseinandersetzungen zwischen Virtualität und Realität, instantaner Informationsaustausch (Vernetzung) sowie (abnehmende) Notwendigkeit computertechnischen Vorwissens.

³⁶² Nach: Stiegler, Bernard: Licht und Schatten im digitalen Zeitalter. Programmatische Vorlesung auf dem Digital Inquiry Symposium am Berkeley Center for New Media, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 35-46, hier S. 44 f.

³⁶³ Lauer, Gerhard: Die digitale Vermessung der Kultur. Geisteswissenschaften als Digital Humanities, in: Geiselberger, Heinrich; Moorstedt, Tobias (Hg.): Big Data: das neue Versprechen der Allwissenheit, Berlin 2013 (edition unseld), S. 100.

³⁶⁴ Vgl.: Ebd., S. 101; Foster, Ian: How Computation Changes Research, in: Bartscherer, Thomas; Coover, Roderick (Hg.): Switching codes: Thinking through digital technology in the humanities and the arts, Chicago 2012, S. 15-37.

³⁶⁵ Das CERN stellt Daten öffentlich zur Verfügung. Siehe URL: <http://opendata.cern.ch/>.

Der Umbruch kann als *Computational Turn* bezeichnet werden, den ich als 'Turn, der durch den gesellschaftsdurchdringenden Einsatz von elektronischen Informationssystemen und digitalen Objekten sowie Virtualität und instantane Hypertextualität (Vernetzung) charakterisiert ist' bezeichne.

4.5. Zusammenfassung

Die Bedeutung des Medienbegriffs konnte nicht abschliessend geklärt werden, weil es zu viele verschiedene Definitionsansätze gibt. Es wird deshalb eine Arbeitsdefinition verwendet, aus der wiederum eine Arbeitsdefinition für Digitalmedien abgeleitet und verwendet wird:

Ein Medium ist ein Vermittlungssystem, das von Menschen für Menschen zur beabsichtigten oder unbeabsichtigten Übermittlung von Informationen verwendet wird, wobei sich Medium und Kommunikation gegenseitig bedingen.

Ein Digitalmedium ist ein Vermittlungssystem, das von Menschen für Menschen zur beabsichtigten oder unbeabsichtigten digitaler Übermittlung von Informationen verwendet wird, wobei sich Medium und Kommunikation gegenseitig bedingen.

Der Begriff Neue Medien ist abzulehnen, da jedes neue Medium mit fortschreitender Zeit und technischer Entwicklung veraltet. Hingegen kann der Begriff Social Media weiterhin verwendet werden:

Unter Social Media werden Digitalmedien für die kommunikative Interaktion von miteinander in Beziehung stehenden Menschen verstanden, die Informationen als Prosumer unmittelbar erstellen, teilen und verknüpfen können.

Die *tl;dr-generation*³⁶⁶ ist vom Lesen längerer Textpassagen arbeitstechnisch überfordert und kann die Konzeption des Deep Reading nicht vollständig umsetzen. Wenn Geschichtswissenschaftler nicht mehr fähig sind, ihre Quellen intensiv zu studieren sowie hermeneutisch und epistemologisch zu erfassen, wird die Qualität wissenschaftlicher Arbeiten massiv leiden. Zudem wird der Wissenschaftler nicht mehr in der Lage sein, sich gegen Laienwissen durchzusetzen.

Die Verbreitung des Computers hat grosse soziologische und kulturelle Auswirkungen auf die Gesellschaft und damit auch auf die Wissenschaft, weshalb von einem Umbruch zu einem Digitalen Zeitalter gesprochen werden kann.

³⁶⁶ tl;dr: too long; didn't read. In Abwandlung der ursprünglichen Bedeutung von 'Kritik am Text, der unnötig überlang formuliert ist' zur Kritik am Leser, der nicht mehr fähig ist, längere Texte zu lesen. Zudem unterscheidet sich das Lesen von webbasierten Artikeln von jenem physischer Artikel. Siehe dazu bspw.: Baron: The Plague of tl;dr, 2015; Nielsen: F-Shaped Pattern For Reading Web Content, 2006.

Konvergenz der Medien (Bitstrom als universales Medium), Abhängigkeit von Informationssystemen, computerunterstützte Arbeitsprozesse, Produktion von überwiegend digitalen Daten, Verlust von Materialität, ständige Auseinandersetzungen zwischen Virtualität und Realität, instantaner Informationsaustausch (Vernetzung) sowie (abnehmende) Notwendigkeit computertechnischen Vorwissens sind Merkmale des Digitalen Zeitalters.

Der Computational Turn ist als eine weiterhin andauernde, relativ schnelle und grundsätzliche Verlagerung und Perspektivenerweiterung von physischen auf digitale Objekte aufzufassen. Insbesondere die grundsätzlichen Veränderungen auf thematischer, linguistischer und methodischer Ebene machen den Computational Turn zu einem herausragenden Umbruch.

Der Computational Turn ist durch den gesellschaftsdurchdringenden Einsatz von elektronischen Informationssystemen und digitalen Objekten sowie Virtualität und instantane Hypertextualität (Vernetzung) charakterisiert.

5. Quellenkritik des Digitalen

Die Aufgabe der Historiker besteht darin "verlässliche *Quellen zu wählen, sie in verlässlicher Weise zu lesen und sie so zusammenzustellen, dass sie verlässliche Darstellungen der Vergangenheit ermöglichen.*"³⁶⁷ (Howell und Prevenier, 2004)

Die Quellenkritik des Digitalen findet innerhalb des neu etablierten historisch-kritischen Prozesses statt. Die traditionelle Methode wird dabei um den Schritt der Objektsicherung ergänzt, der bei digitalen Objekten aufgrund deren Volatilität und Manipulierbarkeit zwingend durchzuführen ist. Die Eigenschaften eines digitalen Objekts und die damit verbundenen Probleme für die Geschichtswissenschaft erfordern zudem neue Methoden aus der Informationstechnik, um eine quellenkritische Analyse adäquat durchführen zu können. Neue, auf der Informationstechnik basierende Verfahren für die Geschichtswissenschaft fordert auch Monika Dommann:

*"Die Historiker sind ihrerseits nun vor die höchst produktive Aufgabe gestellt, neue Analysetechniken zu entwickeln, um elektronische Daten zu suchen, analytisch zu entschlüsseln und die veränderten medialen und technischen Bedingungen der Datengenerierung auch epistemologisch im Rahmen einer erweiterten Quellenkritik zu erfassen."*³⁶⁸

Damit positioniert sich die historisch-kritische Methode mit dem Kernelement der Quellenkritik neu, denn die interdisziplinäre Erweiterung setzt neue Kompetenzen voraus, die jeder Historiker und Geschichtsforschende erlernen muss. Wissenschaftler haben demnach, wie vielfach gefordert, ihre Informations- und Medienkompetenzen zu erweitern, ohne allerdings die bisher erforderlichen Kompetenzen zu vernachlässigen.³⁶⁹ Erst die Möglichkeit, sowohl physische als auch digitale Medien nutzen zu können, erlaubt es, Qualitätsunterschiede und medienspezifische Besonderheiten zu erkennen.³⁷⁰ Peter Haber fordert denn auch eine allgemeine *Fähigkeit zur Medien- und Informationskritik*³⁷¹ und für den Historiker und Wissensforscher Philipp Sarasin hat der kompetente Mediennutzer "zumindest eine Ahnung von medialen Produktionsbedingungen, kann politische Standorte erkennen und nutzt verschiedene Medien, um nicht von einem Kanal abhängig zu sein."³⁷² Geisteswissenschaftler haben sich demnach informationstechnische Kenntnisse anzueignen, wenn sie Projekte durchführen wollen, die die Verwendung digitaler Informationssysteme beinhalten. Nur schon

³⁶⁷ Howell; Prevenier: Werkstatt des Historikers, 2004, S. 6. Schriftschnitt im Original invers.

³⁶⁸ Dommann, Monika: Schlüssel für Clio: Digitale Daten und die Suche nach einer neuen Quellenkritik, in: Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Digital Humanities und Web 2.0, 2012, S. 57.

³⁶⁹ Siehe dazu auch: Haber: digital past (unv. Habil.), 2009, S. 188-191 & 201.

³⁷⁰ Nach: Sarasin, Philipp: Schlaue Maschinen. Peter Habers kritische Medienwissenschaft und unsere Lage im Netz heute, in: Gugerli et al. (Hg.): Digital Humanities, Zürich 2013 (Nach Feierabend 9), S. 196.

³⁷¹ Siehe: Haber: Digital Past, 2011, S. 104.

³⁷² Sarasin: Schlaue Maschinen, 2013, S. 197.

die Kenntnis der Verständigungs-codes in den verschiedenen Disziplinen, insbesondere der Informatik, ist notwendig, sollen Missverständnisse und fehlerhafte Umsetzung vermieden werden.³⁷³ Datenintegrität, Authentizität und Authentifizierung sind dabei nur drei Stichworte.³⁷⁴

In diesem Kapitel werden zuerst aktuelle Probleme in der quellenkritischen Untersuchung von digitalen Objekten in den sechs einzelnen Prozessschritten diskutiert und Lösungsansätze präsentiert, wie der wissenschaftliche Umgang mit digitalen Objekten verbessert werden kann.

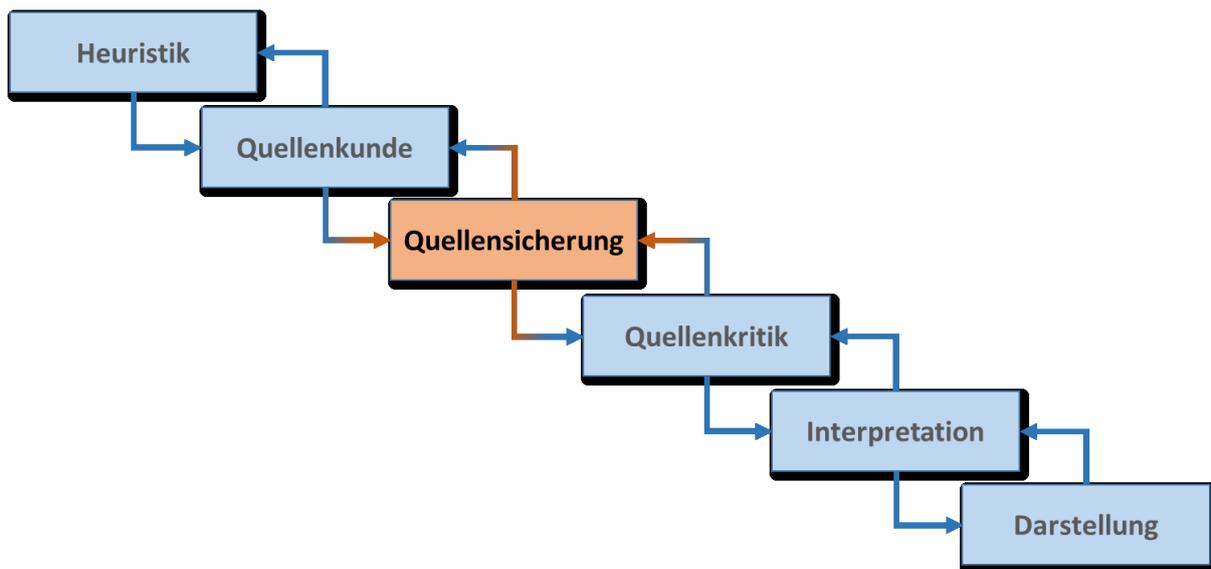


Abbildung 6: Historisch-kritischer Prozess mit Quellensicherung nach Pascal Föhr (Föhr)

In der Anwendung der angepassten historisch-kritischen Methode zeigt sich, wie sich die historische Quellenkritik durch die Verwendung von digitalen Objekten als Forschungsressource sowie digitalen Informations- und Kommunikationsmedien verändert. Anschliessend wird der von Jesper Verhoef und Melvin Wever für die quellenkritische Überprüfung digitaler Objekte entwickelte Digital Humanities Cycle vorgestellt und dargelegt, warum dieser (noch) nicht verwendet werden sollte.

Es ist nicht möglich, alle Methoden und Techniken der Informatik zur Untersuchung von digitalen Objekten vorzustellen, da eine nicht überschaubare Anzahl an Objektarten existiert, die ständig weiterentwickelt werden, und zudem laufend neue entstehen. Allgemein entwickelt sich die Informationstechnik rasant weiter, weshalb hier grundlegende Überlegungen zu technischen Problemen formuliert und konkrete Probleme nur exemplarisch behandelt werden. Es werden *Problembereiche* definiert, die grössere Auswirkungen auf die Quellenkritik der Geschichtswissenschaft und deren Arbeitsweise haben. Der Leser soll für

³⁷³ Dazu: Pawliczek, Aleksandra: Geschichtsforschung und Geschichtsschreibung zwischen Digital Humanities und Archivquellen, in: geschichte und kritik (61), 2013, S. 39 und 42.

³⁷⁴ Vgl.: Haber: Digital Past, 2011, S. 152.

die Probleme im Umgang mit digitalen Objekten sensibilisiert werden und Methoden kennenlernen, die einem Objekt Authentizität, Integrität und Nachvollziehbarkeit attestieren können.

5.1. Heuristik

Über Digital Natives: *"Den meisten von ihnen muss man die Techniken der elektronischen Recherche sehr gründlich erklären; von nichtkommerziellen Plattformen und vom deep web haben sie gewöhnlich kaum eine Ahnung."*³⁷⁵ (Groebner, 2014)

In der Heuristik, verstanden als "Kunst, Wege zur Gewinnung wissenschaftlicher Erkenntnisse zu finden"³⁷⁶, sollte strukturiert vorgegangen werden – was oft nicht der Fall ist. David Nicholas stellte unter Berücksichtigung mehrerer Studien wie folgt fest: "young people have a poor understanding of their information needs and thus find it difficult to develop effective search strategies as a result; [...]"³⁷⁷. Der Recherchierende sollte für die im Voraus formulierte Forschungsfrage³⁷⁸ eine Strategie definieren und mit dafür geeigneten Techniken die Suchräume auswählen und durchsuchen. Für eine erfolgreiche Recherche³⁷⁹ wird ein hybrides Vorgehen verlangt, um alle potenziellen Quellen zu finden. Die Bestände an Archiv- und Sammlungsgut haben sich seit der Durchsetzung des Computers als Arbeitsinstrument mehr und mehr zum Digitalen verschoben und bereits sind Bestände nur noch digital verfügbar. Mittlerweile ist es auch unwahrscheinlich geworden, dass zu einem Forschungsthema keine digitalen oder digitalisierten Publikationen existieren oder neuere physische Publikationen mit einer digitalen Suche nicht gefunden werden können. Aber es kann auch nicht davon ausgegangen werden, dass alle relevanten Publikationen ausschliesslich digital vorliegen. Diese stehen beispielsweise im Web nicht zur Verfügung, weil eine digitale Version für die Gedächtnisinstitution(en) zu teuer ist, Inhalte aufgrund von Urheberrechtsbestimmungen wie bisher nur analog eingesehen werden können, ein Archiv immer noch keine digitalen Findmittel im Internet zur Verfügung stellt oder weil das gesuchte Werk eventuell noch nicht digitalisiert wurde.³⁸⁰ Die im Rechercheprozess vielfältig nutzbaren Möglichkeiten mit jeweils

³⁷⁵ Groebner: Wissenschaftssprache digital. Die Zukunft von gestern, 2014, S. 13. Schriftschnitt im Original invers.

³⁷⁶ Fuchs, Konrad; Raab, Heribert: Wörterbuch zur Geschichte, München 2002.

³⁷⁷ Nicholas u. a.: Google Generation II, 2011, S. 30.

³⁷⁸ Die Erarbeitung der Forschungsfrage erfolgt unabhängig von zu verwendenden Quellenarten und wird deshalb hier nicht weiter ausgeführt.

³⁷⁹ Es wird grundsätzlich von einer *Recherche* ausgegangen, die als Zusammentragen von komplexen und verknüpften Objekten und Informationen zum Zweck der Klärung von Hintergründen und Umständen aufgefasst wird. Unter *Retrieval* wird eine Suche nach einem bestimmten Objekt oder einer konkreten Information verstanden.

³⁸⁰ Allgemein zu Recherche (v.a. für Deutschland) siehe bspw.: Kühmstedt, Estella: Klug recherchiert: für Historiker, Göttingen 2013 (UTB 3940). Kommentiertes Verzeichnis elektronischer Informationsres-

spezifischen Eigenschaften stellen hohe Anforderungen an die Informationskompetenz des Recherchierenden. Damit verbunden ist einerseits die Frage, welche 'Geschichte' geschrieben werden kann, wenn ausschliesslich digitale Quellen verwendet werden, und andererseits die Fragen, welches Material, welche Sichtweisen, Ereignisse, historischen Akteure und möglichen Narrative dadurch ausgeschlossen werden.³⁸¹ Weil sich viele Geisteswissenschaftler dieses Problems nicht bewusst sind, wird in den ersten drei Unterkapiteln auf Veränderungen in der Suchstrategie und den -techniken sowie auf Suchräume eingegangen.

Das Kapitel 5.1.5 widmet sich kurz der Problematik der De- und Kontextualisierung sowie der damit verbundenen Rekontextualisierung. Bei einer Suche werden die aufgefundenen Objekte unter Umständen aus ihrem bestehenden Kontext herausgerissen. Um beurteilen zu können, ob es sich um ein für die Fragestellung relevantes Objekt handelt, ist es wieder in seinen ursprünglichen Kontext zu setzen und dann zu untersuchen.

Verbunden mit der Heuristik ist auch die Frage, wem das publizierte oder aufgefundene Wissen gehört sowie ob und wie es genutzt werden darf. Diverse rechtliche Probleme entstehen beim Austausch und bei der Nutzung von digitalen Objekten, insbesondere weil die bestehenden Rechtssysteme nur innerhalb geographischer Grenzen gelten, das Internet aber weltweit zur Verfügung steht. Von verschiedenen Institutionen wird regelmässig betont, dass *das Internet*, besonders was die Verwendung von Unterhaltungsmedien anbelangt, keinen rechtsfreien Raum darstelle.³⁸² Der Suchende sollte sich der rechtlichen Problematiken bewusst sein, die in Kapitel 5.1.6 Recht angesprochen werden. Er sollte beispielsweise nachvollziehen können, warum auf einen Suchraum oder ein Objekt nicht zugegriffen werden kann. Forschende sollten auch gewisse juristische Kompetenzen erwerben, um potenzielle Quellen rechtskonform verwenden zu können.

Im heuristischen Prozess ist ein besonderes Augenmerk auf die anzuwendenden Suchtechniken zu legen. Der Einsatz eines Informationssystems als Hilfsmittel bedingt eine vertiefte Auseinandersetzung mit den Eigenschaften digitaler Suchtechniken, insbesondere weil heutzutage einer der ersten Schritte einer Recherche die Eingabe einer Zeichenfolge in Google Search ist.³⁸³ Die Beurteilung der Relevanz von Suchergebnissen bleibt ein wichtiger Teil der Heuristik.

sourcen: Gantert: El. Informationsressourcen, 2011, mit teilaktualisiertem Ressourcenverzeichnis unter URL: <http://www.degruyter.com/view/product/41979>.

³⁸¹ Nach: Zaagsma: On Digital History, 2013, S. 21.

³⁸² Siehe bspw.: Team #hwzmobile: Das Internet; kein rechtsfreier Raum, CAS Mobile Busines, Hochschule für Wirtschaft Zürich – Center for Digital Busines, 23.01.2015, URL: <http://hwzdigital.ch/das-internet-kein-rechtsfreier-raum/>, Stand: 03.03.2015; Schüler, Hans-Peter: Urheberrecht: Spanien besteuert Web-Links, Newsticker, heise online, 26.07.2014, URL: <http://www.heise.de/newsticker/meldung/Urheberrecht-Spanien-besteuert-Web-Links-2268747.html>, Stand: 19.09.2014.

³⁸³ Google Search als Beispiel für eine Online-Suchmaschine (URL: <https://www.google.ch/>).

5.1.1. Suchstrategie

Der Historiker Jan Hodel stellte fest, dass digitale Arbeitstechniken eigentlich hybride Arbeitstechniken sind, die in vielfacher Hinsicht von Praktiken der *analogen* Zeit geprägt sind.³⁸⁴ Dies bezieht er insbesondere auf das Vorwissen des Recherchierenden, die Gestaltung des Suchvorgangs und die Art und Absicht der Recherche. Hodel verwendet zwar das Schlüsselwort *Technik*, er bezieht es aber auf Elemente der Suchstrategie. In einer kurzen Recherche ist aufgefallen, dass in vielen Wegleitungen, Studienleitfäden oder anderer Einführungsliteratur zum Geschichtsstudium oder zum wissenschaftlichen Arbeiten nicht klar zwischen Strategie und Technik unterschieden wird, was zu Missverständnissen führen kann.³⁸⁵ Wie der Begriff *Suchstrategie* in der Geschichtswissenschaft zu verwenden ist, scheint bisher nicht klar erläutert worden zu sein.³⁸⁶ Der Begriff findet sich nur in Zusammenhang mit spezifischen Regeln, Methoden, Techniken, Arbeitsanweisungen oder als einfache Tipps: Eine erste Anlaufstelle ist die Bibliothek; Suchen Sie nach Quelleneditionen; Fragen Sie sich 'Was suche ich überhaupt?'). Ausgehend vom allgemeineren Begriff *Strategie*, als ein "genauer Plan des eigenen Vorgehens, der dazu dient, ein militärisches, politisches, psychologisches oder ähnliches Ziel zu erreichen, und in dem man diejenigen Faktoren, die in die eigene Aktion hineinspielen könnten, von vornherein einzukalkulieren versucht"³⁸⁷, wird dieser für die Verwendung im heuristischen Prozess Geschichtswissenschaft in drei Punkten angepasst. Erstens ist das Ziel das Finden von Objekten zur Beantwortung einer zuvor formulierten, spezifischen Forschungsfrage. Zweitens sind Faktoren wie Zeit, Raum, finanzielle Möglichkeiten und zur Verfügung stehende Suchräume einzukalkulieren. Diese Faktoren können sich während des Suchvorgangs verändern. Drittens sollte das Vorgehen nach den Anforderungen an das wissenschaftliche Arbeiten nachvollziehbar und systematisch sein. Aufgrund dieser Ausführungen wird der Begriff Suchstrategie wie folgt definiert:

³⁸⁴ Nach: Hodel, Jan: Recherche: Google – and Far Beyond, in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 26.

³⁸⁵ Auf Hinweise zu Studienleitfäden wird verzichtet. Literatur bspw.: Freytag, Nils: Kursbuch Geschichte. Tipps und Regeln für wissenschaftliches Arbeiten, Paderborn 2011 (UTB 2569, Ed. 5), insb. Kap. 2 «Literatur und Quellen finden»; Lengwiler: Praxisbuch Geschichte, 2011; Eder u. a.: Geschichte Online, 2006; Rogers: Digital methods, 2013; Franck, Norbert; Stary, Joachim (Hg.): Die Technik wissenschaftlichen Arbeitens: Eine praktische Anleitung, Paderborn 2013 (UTB 724); Kühmstedt: Klug recherchiert (Hist), 2013, insb. S. 147-150. Beispiel für Online-Einführung: Baumann, Jan; Hügi, Jasmin: compas: Strukturiertes Forschen im Web, infoclio, 03.2015, URL: <http://www.compas.infoclio.ch/de>, Stand: 24.03.2015.

³⁸⁶ Hypothese anhand einer Durchsicht geschichtswissenschaftlicher Wörterbücher, bspw.: Fuchs; Raab: Wörterbuch zur Geschichte, 2002; Jordan, Stefan: Lexikon Geschichtswissenschaft: Hundert Grundbegriffe, Stuttgart 2007; Eder u. a.: Geschichte Online, 2006; Gasteiner; Haber (Hg.): Digitale Arbeitstechniken, 2010; Eh, Doris; Schütte, Simone: Literatur finden, in: Franck, Norbert; Stary, Joachim (Hg.): Die Technik wissenschaftlichen Arbeitens eine praktische Anleitung, Paderborn 2013 (UTB 724), S. 33-64.

³⁸⁷ Wermke u. a. (Hg.): Duden - das Fremdwörterbuch, 2001.

In der Geschichtswissenschaft wird mit Suchstrategie ein nachvollziehbarer und systematischer Vorgehensplan bezeichnet, wie unter Berücksichtigung bestimmender Faktoren nach Objekten recherchiert wird, die für die Beantwortung einer bestimmten Fragestellung massgeblich sind.

Zur Abgrenzung der Suchstrategie von Suchtechnik: Eine *Technik* ist eine Fähigkeit oder Kunstfertigkeit, die zur richtigen Ausübung einer Sache notwendig ist, oder die Gesamtheit der Kunstgriffe und Verfahren, die auf einem bestimmten Gebiet üblich sind.³⁸⁸ In der Strategie wird festgelegt, *was* getan wird, die Technik hingegen besagt, *wie* etwas getan wird.

Beim Festlegen der Suchstrategie sind vier Bereiche zu berücksichtigen: *Erstens* wird für die Erarbeitung der Suchstrategie die Kenntnis der Fragestellung vorausgesetzt – 'Was will ich?'. Es sollte bekannt sein, welche Ziele mit der Beantwortung der Frage verfolgt werden und wie das Thema kontextualisiert ist. *Zweitens* sind die theoretischen Grundlagen zu erarbeiten. Aufgrund der Fragestellung kann antizipiert werden, ob zusätzliche Wissenschafts- und Fachgebiete wie die Sozialwissenschaft oder die Sphragistik mit einbezogen werden sollten oder müssen und ob zusätzliche theoretische Grundlagen zu erarbeiten sind. Die theoretischen Grundlagen beeinflussen die Auswahl der Suchräume, insbesondere dann, wenn die Fragestellung den Einbezug weiterer Fachgebiete erfordert. *Drittens* sind einschränkende Faktoren zu eruieren. Ideal wäre, wenn dem wissenschaftlichen Forschungsprozess unbegrenzte Ressourcen zur Verfügung stehen. Die Praxis hingegen beweist das Gegenteil, weil Faktoren wie Raum, Zeit oder Geld das Ergebnis einer Recherche und des Projekts beeinflussen. Einschränkende Faktoren haben Auswirkungen auf das Resultat einer Recherche. Dies ist bereits bei der Erarbeitung der Suchstrategie einzukalkulieren, um ergebnislose oder nicht verwertbare Recherchen eventuell zu verhindern und allenfalls negative Auswirkungen auf andere Faktoren zu minimieren. Beispielsweise spart ein nicht zu bearbeitender Suchraum Zeit. Zu den einschränkenden Faktoren gehören auch die für eine Recherche notwendigen technischen Ressourcen. Beispielsweise wäre eventuell eine Online-Recherche im heutigen Web mit der Hardware von 1995 möglich, sie würde aber enorm oder sogar zu lange dauern. Demgegenüber ist eine Software von 1995 auf einer heutigen Hardware-Umgebung wahrscheinlich nicht ohne zusätzlichen Komponenten wie Treiber oder Emulatoren einsatzfähig. *Viertens* sind die Suchräume festzulegen, denn je spezifischer diese definiert werden, desto schneller wird ein Recherchevorgang abgeschlossen. Einfallsreichtum und Kreativität sind hier gefragt, um auch auf den ersten Blick nicht evidente Suchräume einzubeziehen. Interessante Objekte finden sich unter Umständen auch in

³⁸⁸ Nach: Ebd. Weiteres in Kap. 5.1.3. Suchtechniken.

Suchräumen thematisch entfernter Gebiete. Je nach Resultat der Suche kann oder sollte die Fragestellung angepasst und geschärft werden.

Es ist nicht möglich eine allgemein gültige Suchstrategie festzulegen, denn diese beruht auf individueller Erfahrung, Einfallsreichtum, Kreativität und auf persönlichen Präferenzen eines Recherchierenden. Je mehr dieser über den Untersuchungsgegenstand weiss, desto schneller steht die persönliche Strategie fest. Sie wird aber vermutlich während der Recherchetätigkeit verändert werden, denn es bestehen wechselseitige Beziehungen zwischen den einzelnen Teilbereichen, was ein iteratives Vorgehen verlangt. Beispielsweise beeinflussen Rechercheergebnisse die vorher festgelegte Fragestellung nachträglich. Je weniger Vorwissen über das Forschungsthema bestand, desto wahrscheinlicher ist eine Anpassung der Fragestellung oder sogar deren Reformulierung. Eine Anpassung der Fragestellung kann wiederum einschränkende Faktoren verändern.

Für den Erfolg eines Rechercheprojekts ist eine gute Suchstrategie elementar, weil damit die zu durchsuchenden Suchräume und die zu erwartenden, verwendbaren Objekte bestimmt werden.³⁸⁹ Recherchierende haben mit den in den jeweiligen Suchräumen verfügbaren Medien vertraut zu sein, um die aufgefundenen Objekte quellenkritisch überprüfen zu können. Dazu ist eine *hohe Digitalmedienkompetenz* nötig, die der Wissenschaftler durch Ausbildung und Erfahrung im entsprechenden Gebiet oder durch die Beratung von einem Spezialisten erhält. Unter *Medienkompetenz* wird in konzeptioneller Ausweitung des älteren Begriffs *Lesekompetenz* ein 'Bündel verschiedener Teilkompetenzen'³⁹⁰ verstanden, die nicht nur rein technische und intellektuelle Kompetenzen wie die Auswahl problemrelevanter Medien umfasst, sondern auch die Fähigkeit zur Medien- und Informationskritik³⁹¹, die auch eine Medienreflexionskompetenz einschliesst – 'the medium is the message', wie Marshal McLuhan sagte.³⁹² Jedem Nutzer sollten folglich die Funktionalitäten, Eigenschaften und Auswirkungen digitaler Medien und Objekte bekannt sein, bevor er sie als digitale Quellen verwendet. Erst dann ist eine kritische Untersuchung möglich.

Erst eine fundierte Medienkompetenz ermöglicht es, sinnvolle und neue Fragestellungen für Themenbereiche der Digital History zu entwickeln. Informationstechnische Methoden können einerseits (endlich) Antworten auf bestehende Forschungsfragen liefern oder gänzlich neue generieren, die wiederum nur mit Hilfe der Informationstechnik beantwortet werden können. Bei der Erarbeitung neuer Fragestellungen steht die Geschichtswissenschaft noch am

³⁸⁹ Hilfreich für die Erstellung einer Suchstrategie sind vor allem Visualisierungstechniken wie Flussdiagramme, Suchbäume oder Mindmaps, wie bspw.: TheBrain (URL: <http://thebrain.com/>), Dia (URL: <http://dia-installer.de/>).

³⁹⁰ Nach: Haber: Digital Past, 2011, S. 104. Siehe weiterführend im Kap. 6.1. Ausbildung.

³⁹¹ Nach: Strohm, Klauspeter: Kulturwissenschaften und Medienkompetenz. Einige Überlegungen, in: Historische Anthropologie Kultur, Gesellschaft, Alltag (1), 2001, S. 128.

³⁹² McLuhan; Fiore; Agel: The medium is the message, 1967.

Anfang, denn bisher stehen hauptsächlich quantitative Analysen wie bei der TEI oder die Geschichte des Digitalen im Fokus der Forschung. Zusätzlich ist bei aller Euphorie über die neuen Möglichkeiten auch danach zu fragen, welche Antworten auf welche Fragen (weiterhin) nicht berechnet werden können und eine qualitative Analyse erfordern. Die Erarbeitung einer Suchstrategie, die die Kenntnis der konkreten Fragestellung, die Erarbeitung theoretischer Grundlagen, das Feststellen einschränkender Faktoren und das Festlegen der Suchräume umfasst, unterscheidet sich auf der prozessualen Ebene nicht von der bisherigen Methode. Allerdings wird in der Quellenkritik des Digitalen vom Recherchierenden zusätzlich eine hohe Digitalmedienkompetenz erwartet, die informationstechnische und intellektuelle Grundlagen über digitale Medien und Objekte umfasst. Die Kompetenzen über physische Medien dürfen jedoch nicht vernachlässigt werden – ein hybrides Vorgehen ist gefordert.

5.1.2. Suchräume

Das heute allgegenwärtige Internet erweckt den Eindruck, dass digitale Suchräume einfach zugänglich, alle nötigen Informationen dort vorhanden und verlinkt sind und dass alles problemlos verwendet werden könne.³⁹³ Dieser Eindruck täuscht, denn es ist aufgrund von (Selbst-)Zensur, Schutzfristen, versteckten Inhalten oder technisch-infrastruktureller Abschottung nur ein kleiner Teil der im Internet vorhandenen Daten zugänglich. Es ist vielen Wissenschaftlern nicht klar, dass und welche Einschränkungen in digitalen Suchräumen bestehen. Die Herkunft der potenziellen Quellen und welche Quellen aus welchen Gründen nicht verwendet werden können, weil sie beispielsweise nicht auffindbar oder nicht vorhanden sind, ist jedoch ein wichtiger Bestandteil der Quellenkritik. Entsprechend sind Lösungsansätze zu entwickeln und umzusetzen, um bekannte Hürden abzubauen und den Zugang zu Informationen zu gewährleisten.

Die Recherche in rein analogen Suchräumen hängt von räumlichen und zeitlichen Einschränkungen sowie dem Vorwissen des Recherchierenden ab. Je nach zur Verfügung stehender Zeit und finanziellen Ressourcen, kann eine mehr oder weniger grosse Anzahl an Institutionen persönlich besucht und in deren Beständen nach verwertbaren Objekten gesucht werden. In den vorhandenen Findmitteln oder in der Sekundärliteratur zum Thema finden sich eventuell Hinweise auf weitere Bestände, die ausgewertet werden könnten. Je

³⁹³ Einführung zum Thema Suchräume siehe bspw.: Krüger: Die Erschliessung digitaler und analoger Suchräume. Anforderungen an heuristische Verfahren, 2005.

nach Forschungsthema sind Reisen an verschiedene Orte unumgänglich, um die Bestände einsehen zu können.³⁹⁴

Suchräume können anhand der *Zugriffs- und Zugangsmöglichkeiten* unterschieden werden. Der Zugang zu privaten, unternehmerischen oder institutionellen Datenbeständen ist in der Regel stark beschränkt, mit Ausnahme einer bestimmten Nutzergruppe. Informationen über den Status, den Umfang, die Art oder sogar die Existenz des Datenbestands können fehlen. Ein Zugriff ist erst möglich, wenn das Einverständnis des Dateneigners vorliegt und entweder die gewünschten Informationen dem Nutzer direkt zugestellt werden oder eine Zugriffsmöglichkeit explizit eingerichtet wird. Die Nutzung beschränkt sich hierbei auf einen klar definierten Bestand, weitere Bestände können nicht durchsucht werden. Solche Suchräume werden als *geschlossen* bezeichnet. Ein geschlossener Suchraum ist beispielsweise das Firmenarchiv der *Richtblick GmbH* in Frauenfeld, auf das nur Mitarbeitende Zugriff haben. Externe können Informationen persönlich oder schriftlich anfordern, worauf diese ihnen nach Zustimmung durch die Dateneigner entweder physisch oder digital zugesendet oder am Standort der Unternehmung eingesehen werden können. Auch kann ein direkter Zugriff, beispielsweise via VPN-Verbindung³⁹⁵ eingerichtet werden.

Davon wird ein Suchraum unterschieden, der einem Nutzer zwar Informationen über dessen Datenbestand zur Verfügung stellt, aber den direkten Zugriff auf den Bestand selbst jedoch nicht erlaubt. Ein Nutzer kann erst nach einer Anmeldung und/oder einer Form von Gegenleistung wie einem Abonnement, einer Zugangsgebühr oder der Zugehörigkeit zu einer Institution auf den Datenbestand zugreifen. Die Nutzung beschränkt sich auf einen klar definierten Bestand, weitere Bestände können nicht durchsucht werden. Die Nutzung ist vom Dateneigner grundsätzlich vorgesehen und erwünscht.³⁹⁶ Solche Suchräume werden als *halboffen* bezeichnet. Dies ist beispielsweise bei einem Zeitungsarchiv der Fall, das zwar Online-Findmittel zur Verfügung stellt, aber keinen grundsätzlichen Zugriff auf die Artikel gewährt. Es sind zwar Informationen über den Bestand vorhanden, aber nur autorisierte Personen oder Systeme können auf diesen zugreifen. Bei digitalen, halboffenen Suchräumen entscheidet oft die IP-Adresse des Recherchierenden über die Zutrittsrechte zu elektronischen Daten. So erhält beispielsweise der Nutzer mit einer IP-Adresse der Universität Zürich Zugriff

³⁹⁴ Wenn eine Zusendung von physischen Objekten nach Hause oder in eine nahegelegene Institution möglich und finanziell tragbar ist, sind Reisen unter Umständen nicht nötig. Ist jedoch zusätzlich der Archivierungskontext des Objektes von Interesse, kann auf eine Reise an den Aufbewahrungsort nicht verzichtet werden.

³⁹⁵ VPN: Virtual Private Network. Durch strenge Authentisierung, Autorisierung und Verschlüsselung gesicherter Zugang zu einem Netzwerk. Siehe: Fischer; Hofer: Lexikon der Informatik, 2011.

³⁹⁶ Bspw.: swisslex, kostenpflichtige Plattform für Schweizer Rechtsinformationen (Gesetze, Dissertationen etc.). URL: <https://swisslex.ch>.

auf bestimmte Datenbestände bei JSTOR³⁹⁷, einer der Universität Basel auf andere Bestände und einer der Universität Kampala (UG). Gute IP-Adressen funktionieren wie Passepartouts und "gewähren Zugang zu den aktuellsten Zeitschriften, sie verschaffen Einblick in die elaborierten Datenbanken von kommerziell operierenden Informations Providern und erschliessen den Weg zu den Digitalisaten von vergriffenen Drucksachen."³⁹⁸ Allen anderen bleiben der Zugang und der Zugriff verwehrt.

Eine dritte Form von Suchraum sind *offene* Datenbestände, die ohne Zugriffsbeschränkungen durchsucht und unabhängig von einschränkenden Copyrightbestimmungen genutzt werden können. Der Dateneigner hat mit der Publikation eine Nutzung explizit vorgesehen und gewährt einen unbeschränkten Zugriff auf den gesamten Datenbestand.

Geschlossene, halboffene und offene Suchräume existieren sowohl in analogen wie auch in digitalen Suchräumen. Unterschieden werden kann die Nutzung der Suchräume durch Digital Natives und Digital Immigrants. Zwar wird von allen Forschenden ein hybrides Vorgehen gefordert, aber Digital Immigrants befassen sich weiterhin lieber mit analogen Beständen wohingegen Digital Natives diese immer weniger beachten. Beispielsweise meinte eine Studentin der Bildungswissenschaftlerin Tara Brabazon, dass sie grosse Probleme habe, die in den Kursunterlagen erwähnte weiterführende Literatur von zu Hause aus zu finden:

"When, with raised eyebrow, I stated that the further readings were books and articles sitting on the shelves of the library, she became exasperated: 'You mean I have to go into the library and get them?'"³⁹⁹

Das ist weitaus kein Einzelfall, denn die Bildungsspezialistin Julie Carpenter zeigte in ihrer Studie *Researchers of Tomorrow* auf, dass relativ wenige Doktoranden der Geistes- und Sozialwissenschaften 'Primär-'Quellen wie Zeitungen, Archivmaterial und Sozialdaten nutzen würden.⁴⁰⁰ Aber gleichzeitig wurde festgestellt, dass für Doktoranden der Geisteswissenschaften gedruckte Publikationen weiterhin einen hohen Stellenwert haben, insbesondere weil ihnen im Verhältnis zu den Naturwissenschaften weniger digitale Journals zur Verfügung stehen. Bei den Doktoranden handelte es sich übrigens um Digital Immigrants, was die Verschiebung der Wahl von bearbeiteten Suchräume zu Gunsten der digitalen noch bemerkenswerter macht.

³⁹⁷ JSTOR: Digitale Bibliothek für wissenschaftliche Artikel, Bücher und Primärquellen. Siehe URL: <http://jstor.org/>.

³⁹⁸ Dommann: Schlüssel für Clio, 2012, S. 56.

³⁹⁹ Brabazon, Tara: The university of google. Education in the (post) information age, Aldershot 2008, S. 15.

⁴⁰⁰ Untersuchung der Recherchetätigkeiten bei über 17'000 Doktoranden aus 70 Bildungsinstitutionen Grossbritanniens. Siehe: Carpenter, Julie; Wetheridge, Louise; Tanner, Sophie: Researchers of Tomorrow: The research behaviour of Generation Y doctoral students, JISC, Bristol, London 28.06.2012, S. 5 und 24. URL: <http://www.jisc.ac.uk/publications/reports/2012/researchers-of-tomorrow>, Stand: 04.03.2014.

Die Vernachlässigung analoger Suchräume hat zur Folge, dass unter Umständen wichtige Quellen ungenutzt bleiben. Nicht oder noch nicht digital zugängliche Forschungsergebnisse in aktuellen Publikationen werden ungenügend rezipiert und diskutiert.⁴⁰¹ Ältere und noch nicht digitalisierte Quellenbestände werden in aktuellen Diskussionen ebenfalls nicht mehr erwähnt, obwohl sie wichtige und grundlegende Erkenntnisse enthalten könnten. Diese Problematik bestand zwar schon vor dem Digitalen Zeitalter, aber mit dem bewussten Vernachlässigen oder Ignorieren analoger Bestände verschlimmert sich der Zustand zunehmend, insbesondere weil immer weniger Wissenschaftler das Wissen über bereits vorhandene Erkenntnisse aktiv besitzen.

Digitale oder virtuelle Suchräume, auf die im Folgenden Bezug genommen wird, erlauben Recherchen, die nahezu unabhängig von Raum, Zeit und finanziellen Möglichkeiten sind. Ein solcher Suchraum kann als *entgrenzter Raum*⁴⁰² bezeichnet werden, unabhängig von den tatsächlichen Zugriffs- und Zugangsmöglichkeiten.⁴⁰³ Beispielsweise kann die offene Online-Datenbank eines Archivs grundsätzlich⁴⁰⁴ jederzeit und von jedem Ort mit Internetzugang aufgerufen und durchsucht werden, weshalb sie als offen entgrenzter Raum bezeichnet werden kann. Das Ergebnis der Suche ist abhängig von den zur Verfügung gestellten Daten respektive davon, welche Berechtigungen der Dateneigner dem Nutzer zur Durch- und Ansicht überlässt. Bei restriktiver Berechtigungsvergabe kann beispielsweise zwar der Zugriff auf die Katalogeinträge erlaubt sein, aber das Abrufen des Archivguts selbst ist nur an speziellen Computern vor Ort möglich. Extensiver ausgelegte Berechtigungen können den Zugriff auf das digitale Objekt selbst sowie zusätzlich unter anderem auf Metadaten bieten und Bearbeitungs- und Ergänzungsmöglichkeiten wie Tagging⁴⁰⁵ oder Verlinkungen erlauben.

Je nachdem, ob nach digitalisierten (digitized), digital erstellten (born-digital) oder bereits digital vorhandenen Daten (reborn-digital) gesucht wird, sind andere Suchräume zu definieren. Suchräume können sich auch überschneiden. So digitalisieren beispielsweise verschiedene Schweizer Staatsarchive bestimmte Objekte oder Teilbestände und stellen diese online

⁴⁰¹ Siehe dazu: Wampfler, Philippe: «online fist». Geisteswissenschaften als Social Media, in: Gugerli et al. (Hg.): Digital Humanities, Zürich 2013 (Nach Feierabend 9), S. 89.

⁴⁰² Nach: Schwabe: Historisches Lernen im World Wide Web, 2012, S. 103. Schwabe bezeichnet das World Wide Web als 'entgrenzten' Raum.

⁴⁰³ Zur Orientierung in digitalen Suchräumen siehe bspw.: Krüger: Die Erschließung digitaler und analoger Suchräume. Anforderungen an heuristische Verfahren, 2005, S. 100 f.

⁴⁰⁴ Abgesehen von Wartungszeiträumen, Hard- und Softwareproblemen sowie mutwilliger Manipulation der Bestände und/oder Findmittel (Webauftritt).

⁴⁰⁵ Tagging: Freies, 'unprofessionelles' Verschlagworten durch Nutzer (collaborative/social/crowd tagging oder folksonomy). Je häufiger einem Objekt dasselbe 'Tag' vergeben wird, desto prominenter erscheint es in der Resultatliste einer Suche. Vgl.: Nentwich, Michael: Cyberscience 2.0 oder 1.2? Das Web 2.0 und die Wissenschaft, in: ITA manu:script (ITA-ms-09-02), 11.2009, S. 5. DOI: 10.1553 Anm. 4.

zur Verfügung.⁴⁰⁶ In der gleichen Tektonik befinden sich auch genuin digitale und reborn-digitale Objekte, die direkt in ein digitales Langzeitarchiv übernommen worden sind.

Unterschiedliche Suchräume können unter Verwendung von Hyperlinks auch direkt verbunden werden. Das eröffnet zwar neue Möglichkeiten, aber dadurch vermischen sich gleichzeitig auch der kontrollierte Raum für geschichtswissenschaftliche Faktenrecherchen und der amorphe Raum des Internets.⁴⁰⁷ Die an den Recherchierenden gestellte Forderung, alle möglichen Suchräume zu durchforsten, kann einen grossen zeitlichen Aufwand bedeuten, weil sich insbesondere der Zeitpunkt des Rechercheabschlusses nur schwer festlegen lässt. Ständig entstehen durch das Publizieren von neuen digitalen Objekten auch neue Suchräume und Recherchemöglichkeiten. Diese Schwierigkeit bestand schon in der analogen Zeit, nun verschärft sie sich jedoch durch relativ kurze Publikationszyklen. Beispielsweise kann eine neue Datenbank publiziert, kurze Zeit später schon mit einer anderen verknüpft und wieder wenige Zeit später in eine andere Plattform integriert werden. Solche Veränderungen gehen bei analogen Datenbeständen sehr viel langsamer vonstatten. Zusätzlich verstärkt wird das Problem durch den stetigen und zunehmenden *Informationszufluss* in die Suchräume selbst. Bei Recherchen zu aktuellen Themen – wie für die vorliegende Arbeit – bedarf es zu dessen Bewältigung einer gut ausgearbeiteten Suchstrategie, die nur mit digitalen Hilfsmitteln umsetzbar ist.

Wird ein analoger Suchraum digitalisiert, darf keinesfalls von einem äquivalenten Informationsangebot ausgegangen werden. Der digitale Suchraum kann einerseits mit Tags, Verlinkungen, Metadaten oder hybriden Beständen zusätzliche Informationen bieten als sein analoges Pendant. Andererseits kann er auch weniger Informationen enthalten, wenn beispielsweise nur Verzeichnisse, aber keine Datenbestände enthalten sind. Vor allem bei der Digitalisierung von öffentlich verfügbaren Datenbeständen von staatlichen Verwaltungsstellen ist feststellbar, dass die digitalisierten Suchräume eher halboffen konzipiert werden. Peter Haber meinte diesbezüglich, dass die Diskrepanz zwischen einfach zugänglichen Online-Quellen und fast gar nicht zugänglichen Offline-Quellen zunehmen werde.⁴⁰⁸ Der Historiker Sacha Zala⁴⁰⁹ kommt in Bezug auf das Schweizerische Bundesarchiv (BAR) zu einem anderen Schluss. Laut Zala sind beispielsweise die Entscheide des Bundesrats nach Ablauf der dreissigjährigen Schutzfrist zwar physisch im BAR einsehbar, jedoch nicht über dessen

⁴⁰⁶ Bspw. das Staatsarchiv Zürich, URL: <http://www.staatsarchiv.zh.ch/>. Siehe Archivkatalog StAZH.

⁴⁰⁷ Vgl.: Enderle, Wilfried: Fakten und Informationen im digitalen Raum, in: Busse, Laura; Enderle, Wilfried; Hohls, Rüdiger u. a. (Hg.): Clio Guide. URL: <http://guides.clio-online.de/guides/arbeitsformen-und-techniken/fakten-und-informationen/2016>, Stand: 10.07.2016.

⁴⁰⁸ Haber, Peter: Zeitgeschichte und Digital Humanities, Version: 1.0, Archive, Docupedia, 24.09.2012, URL: http://docupedia.de/zg/Digital_Humanities?oldid=84594, Stand: 17.09.2014.

⁴⁰⁹ Direktor der Forschungsgruppe 'Diplomatische Dokumente der Schweiz' (URL: <http://dodis.ch/>) und Präsident der Schweizerischen Gesellschaft für Geschichte (URL: <http://www.sgg-ssh.ch/>).

Online-Plattform abrufbar.⁴¹⁰ Das BAR wendet anscheinend die Regeln zur Öffnung und Sperrung von Beständen strenger an, denn der vollständige Online-Zugang und -Zugriff auf öffentliche Bestände ist auch nach Ablauf der Schutzfristen nicht verlässlich möglich.⁴¹¹ Dazu ist anzumerken, dass Online-Quellen speziell wegen des Daten- und Persönlichkeitsschutzes restriktiven Zugangs- und Zugriffsbeschränkungen unterliegen müssen, weil bereits mit einer einfachen, unsystematischen Suche unverhältnismässig viele Daten aufgefunden und (miss-)interpretiert werden können.⁴¹² Der Dateneigner kann dies nur verhindern, indem er den Online-Zugang beschränkt. Als Wissenschaftler will Zala auf alle Bestände online zugreifen, aber er muss auch den Schutz vor unverhältnismässiger Bearbeitung von Datenbeständen aus wirtschaftlichen oder persönlichen Interessen akzeptieren.

Der Einsatz von elektronischen Dokumenten- oder Records Management Systemen (DMS/RMS)⁴¹³ und digitaler Kommunikation hat zusätzliche Veränderungen in den Beständen zur Folge. In DMS werden alle digitalen Daten einer Unternehmung oder einer Institution nach klar definierten Regeln abgelegt. Eine Änderung der darin eingebundenen Daten ist kaum mehr möglich oder führt zu einem neuen Eintrag. RMS sind darauf ausgelegt, Records⁴¹⁴ von der Entstehung bis zur Archivierung oder Kassation in einem klar definierten Prozess (lifecycle) zu begleiten und zu steuern, um einen definierten Qualitätsanspruch zu erfüllen sowie um Datenverlust vorzubeugen und Geschäftsbeziehungen zur Beweisführung rechtlich abzusichern. Deshalb wird in vielen Unternehmen und Institutionen die gesamte digitale Kommunikation der Mitarbeitenden langfristig gespeichert. Für Unternehmen und Institutionen sind solche elektronischen Systeme im Vergleich zu physischen kostengünstiger, sicherer und schneller. Für die Forschung hingegen sind sie eine der Ursachen für den Verlust von Informationsgut, auch *schwarze Löcher* genannt.⁴¹⁵ Bei elektronisch gesteuerten Prozessen werden meist alle anfallenden Daten aufbewahrt, damit sie schnell wieder aufgerufen werden können. Da dies den Nutzern bewusst ist, werden Informationen aus Angst vor Regressansprüchen des Arbeitgebers oder als 'Revolte' gegen die Überwachung

⁴¹⁰ Aussage von Sacha Zala an der infoclio-Tagung 'Datenschutz und Geschichtswissenschaften', Bern, 28.11.2014. Tagungsbericht von Peppina Beeli: URL: <https://www.infoclio.ch/de/node/137237>.

⁴¹¹ Amrein, Marcel: Archivzugang mit Hindernissen, in: Neue Zürcher Zeitung, Zürich 28.07.2014, S. 7.

⁴¹² Bspw. wenn dadurch über eine Person ein verjährtes Delikt publik wird. Siehe Kap. 5.1.6 Recht.

⁴¹³ Zu DMS siehe bspw.: Götzer, Klaus (Hg.): Dokumenten-Management Informationen im Unternehmen effizient nutzen, Heidelberg 2014. Zu RMS siehe bspw.: Toebak: Records Management ein Handbuch, 2007.

⁴¹⁴ Record: strukturierte Dateneinheit (Daten-Record) oder wenig oder unstrukturierte Dateneinheit (Unterlagen-Record), die in einem betrieblichen Kontext mit Struktur- und Prozesswissen eingebunden ist. Vgl. Toebak: Records Management ein Handbuch, 2007, S. 19 f.

⁴¹⁵ Zur Verhinderung von Informations- und Datenverlust in der Forschung siehe bspw.: Hochschulrektorenkonferenz Deutschland: Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien, 10.11.2015. URL: <https://www.hrk.de/positionen/beschluss/detail/wie-hochschulleitungen-die-entwicklung-des-forschungsdatenmanagements-steuern-koennen-orientierungsp/>, Stand: 06.05.2018.

teilweise nicht mehr verschriftlicht und im System eingetragen. Der Mitarbeitende betreibt damit eine Art von (*Selbst-*)*Zensur*. Dass er damit unter Umständen wissenschaftliche Forschung verhindert, ist ihm nicht bewusst, egal oder wird von ihm in Kauf genommen. Zudem ist zu hinterfragen, wie solche Eingriffe in die Überlieferungsbildung überhaupt bewertet werden sollen, wenn sie denn überhaupt entdeckt werden.⁴¹⁶ Deshalb sind Bestände eines Suchraums, die aus solchen Systemen stammen, sehr kritisch zu hinterfragen, wofür Kenntnisse über die Funktionalität der verwendeten Systeme wichtig sind.

Der Suchraum mit dem grössten Datenvolumen ist zurzeit das *Internet*, in dem sich wiederum der grösste Dienst *World Wide Web* befindet. Über die gesamthaft im Internet vorhandene Datenmenge liegen zwar nur Schätzungen vor⁴¹⁷, sie wird aber einen sehr grossen Teil der im Jahr 2014 weltweit vorhandenen digitalen Daten von etwa 4,4 Trillionen Gigabyte für sich beanspruchen.⁴¹⁸ Allerdings kann nur auf einen kleinen Teil der im Internet enthaltenen und indexierbaren Informationen mit einfachen Mitteln zugegriffen werden. Der Datenbestand des *Surface Web* wird mit 10-20%, manchmal nur mit 5% des gesamten Datenvolumens beziffert.⁴¹⁹ Der Rest befindet sich im *Deep Web*⁴²⁰, das drei Gruppen von digitalen Datenbeständen beinhaltet: erstens Objekte, die zwar öffentlich zugänglich sind, aber von den Suchmaschinen nicht indexiert wurden oder werden können, weil sie beispielsweise zu verschachtelt oder in komplexen Datenstrukturen wie Datenbanken⁴²¹ gespeichert sind. Zweitens durch eine Paywall, ein Login oder ähnliches geschützte Objekte und drittens Objekte, die sich in einem separaten Netz wie dem Darknet⁴²² befinden. Diese *Unter- oder*

⁴¹⁶ Vgl.: Bischoff, Frank M.: Bewertung elektronischer Unterlagen und die Auswirkungen archivarischer Eingriffe auf die Typologie zukünftiger Quellen, in: *Archivar, Zeitschrift für Archivwesen* 1, 2014, S. 40-52.

⁴¹⁷ Siehe: Bergman, Michael K.: White Paper: The Deep Web: Surfacing Hidden Value, in: *The Journal of Electronic Publishing* 7 (1), 01.08.2001. DOI: 10.3998/3336451.0007.104; Lewandowski, Dirk; Mayr, Philipp: Exploring the Academic Invisible Web, in: *Library Hi Tech* 24 (4), 2006, S. 529-539; Kunder, Maurice de: Geschatte grootte van het geïndexeerde World Wide Web, 29.03.2007. URL: <http://www.dekunder.nl/Media/Scriptie%20Maurice%20de%20Kunder%20-%20Grootte%20geïndexeerde%20web.pdf>, Stand: 12.03.2015.

⁴¹⁸ Datenvolumen vgl.: Turner u. a.: *The Digital Universe of Opportunities*, 2014, S. 2.

⁴¹⁹ Es existieren keine verlässlichen Informationen über die Volumengrösse und Grössenverhältnisse. Siehe bspw.: Barrenechea, Mark: *Enterprise Information Management (EIM): The Next Generation of Enterprise Software*, OpenText White Paper Series, 12.2012, S. 4. URL: <http://www.legaltechnology.com/wp-content/uploads/2013/07/OpenText-EIM-Summary.pdf>, Stand: 02.11.2015.

⁴²⁰ Deep Web wird auch als Hidden oder Invisible Web bezeichnet. Da sich darin Datenbestände befinden, die weder versteckt noch unsichtbar sein müssen, wird der Begriff 'Deep Web' bevorzugt. Analog wird Surface Web statt Visible Web verwendet.

⁴²¹ Zu den Eigenschaften von Datenbanken siehe bspw.: Margulies: *Digitale Daten als Quelle der Geschichtswissenschaft*, 2009, S. 196; Burkhardt: *Digitale Datenbanken*, 2015, S. 139; Jonas, Jeff; Sokol, Lisa: *Data Finds Data*, in: Segaran, Toby; Hammerbacher, Jeff (Hg.): *Beautiful data: [the stories behind elegant data solutions]*, Sebastopol, Calif 2009 (*Data modeling & design*), S. 105-118.

⁴²² Das Darknet (oder Dark Web) ist ein Teil des Deep Web, das nur mit spezieller Software zugänglich ist und dessen Daten oft nicht visualisiert (graphische Webseite) vorliegen. Hier wird der Begriff

*Teilwebs*⁴²³ sind zwar mit dem gesamten Web verlinkt und damit theoretisch erreichbar, werden aber durch software- und hardwaretechnische Barrieren vom Web abgeschottet.

Zu den Teilweb gehören die in den letzten Jahren entstandenen *Mesh-Netze*. Das sind autonome Internetstrukturen, die ähnliche Dienste wie das 'grosse Internet' zur Verfügung stellen. Beispielsweise versuchen sich Staaten wie China oder der Iran mit solchen Infrastrukturen vom restlichen Web zu separieren. Aber auch Bürgergemeinschaften haben erfolgreich Mesh-Netze aufgebaut. Diese sind vollständig unabhängig und existieren parallel zur globalen Internetinfrastruktur. Apps wie FireChat⁴²⁴ erlauben sogar die direkte Kommunikation zwischen Mobilgeräten, ohne dass ein Provider⁴²⁵ oder Netzknotenpunkt zwischengeschaltet ist.⁴²⁶ Ein Zugriff auf Datenbestände eines Mesh-Netzes wird durch die errichteten Barrieren fast verunmöglicht. Zudem ist fraglich, wer überhaupt Zugang zu einem solchen Netz erhält. Wird es durch einen Staat oder eine Unternehmung zur Verfügung gestellt, so können diese auch Inhalte, Dienste und Nutzer selektieren. Unabhängigkeit und Netzneutralität können kaum gewährleistet werden.⁴²⁷ Ein aufgeteiltes Web ist unberechenbar, unzuverlässig und damit praktisch nutzlos, meint der Informatiker Vinton Cerf, Mitentwickler des Arpanets und damit einer der Väter des Internets.⁴²⁸ Einer der Gründe für diese fortschreitende Zersplitterung des Internets ist die 2013 von Edward Snowden aufgedeckte Datensammelwut der amerikanischen National Security Agency (NSA).⁴²⁹ Den Internet-Usern ist damals bewusst geworden, dass eine tatsächliche Gefahr für den Schutz ihrer Daten besteht – allerdings ignorieren sie diese bewusst, oder der Aufwand für Gegenmassnahmen ist ihnen zu

Darknet verwendet, um begrifflich klarer abzugrenzen, dass das Darknet eher als Teil des Internets und weniger des World Wide Web angesehen werden sollte.

⁴²³ Rogers: Digital methods, 2013, S. 33-35.

⁴²⁴ FireChat: Kostenloser Sofortnachrichtendienst. Für Android siehe Google App Store, URL: <https://play.google.com/store/apps/>.

⁴²⁵ Provider: Eine Unternehmung oder Behörde, die ans Internet angeschlossen ist und Kommunikationsdienstleistungen anbietet.

⁴²⁶ Bspw. wurde von den Teilnehmern der Demonstrationen in Hongkong im Oktober 2014 solche Mesh-Netze erfolgreich eingesetzt. Siehe: Talbot, David; Schwan, Ben: Hongkong: Mesh-Netzwerk ermöglicht Demonstranten die Kommunikation, Infotech, Technology Review, 06.10.2014, URL: <http://www.heise.de/tr/artikel/Hongkong-Mesh-Netzwerk-ermoeglicht-Demonstranten-die-Kommunikation-2411653.html>, Stand: 07.10.2014.

⁴²⁷ Dazu weiterführend bspw.: Bundesamt für Kommunikation: Netzneutralität, Bericht zur Arbeitsgruppe, Schweizerische Eidgenossenschaft, Bern 23.10.2014. URL: <http://www.bakom.admin.ch/themen/internet/04810/index.html?lang=de>, Stand: 30.06.2015; Lanier, Jaron: Wem gehört die Zukunft? Du bist nicht der Kunde der Internet-Konzerne, du bist ihr Produkt, Hamburg 2014. Eine tiefgehende Auseinandersetzung für Deutschland bietet bspw.: Bortnikov, Vyacheslav: Netzneutralität und Bedingungen kommunikativer Selbstbestimmung: Pflichten des freiheitlichen Verfassungsstaates zur Gewährleistung der Neutralität des Internets im Lichte der grundrechtlichen Schutzpflichtenlehre, München 2013 (Schriftenreihe des Instituts für Rundfunkrecht an der Universität zu Köln).

⁴²⁸ Nach: Grassegger, Hannes: Ein zerteiltes Internet wäre nutzlos – Interview mit Vinton Cerf, in: NZZ am Sonntag, Zürich 09.02.2014, S. 26.

⁴²⁹ NSA: National Security Agency, einer der Geheimdienste der USA; Edward Snowden: Whistleblower, der mit den durch ihn veröffentlichten Geheimdienstdaten das Ausmass der weltweiten Überwachung durch die NSA und 'befreundete' internationale Geheimdienste aufzeigte. Eine Einführung widerspricht bspw.: Lyon, David: Surveillance after Snowden, Cambridge 2015.

gross.⁴³⁰ Die Folge davon ist ein Wettkampf zwischen Datensammlern und einem kleinen Teil der überwachten Nutzer, die ihre Daten möglichst durch Verstecken, Abschotten und Verschlüsseln vor Einsicht, Manipulation oder Diebstahl schützen wollen.⁴³¹

Zugriffshürden werden nicht nur durch technische Abschottung geschaffen, sondern auch durch softwaregestützte Dienste wie *Geoblocking* oder inhaltlicher *Zensur*. Geoblocking erlaubt den Zugriff auf digitale Objekte nur in bestimmten geographischen Regionen, was besonders für privatwirtschaftliche Unternehmen interessant ist, die dadurch ihr Angebot gezielter vermarkten oder eine Nutzung verhindern können.⁴³² Auch (eher autoritär regierte) Länder filtern die allgemein erreichbaren Webauftritte. In Russland beispielsweise werden Inhalte zensiert, womit gleichzeitig auch der Zugriff auf diverse Informationen verhindert wird.⁴³³ Bei den Filtern handelt es sich um zwischengeschaltete Programme, die den Zugang zu Daten aufgrund bestimmter Kriterien unterbinden. Meistens wird ein ganzer Webauftritt gesperrt, wenn sich darin eines dieser Kriterien findet. In China, der Türkei und im Iran ist die

⁴³⁰ Zu persönlicher Informationssicherheit siehe bspw. Bundesamt für Sicherheit in der Informationstechnik: Leitfaden Informationssicherheit. IT-Grundschutz kompakt, 02.2012. URL: <https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzUeberblick/LeitfadenInformationssicherheit/leitfaden.html>; Poguntke, Werner: Basiswissen IT-Sicherheit: Das Wichtigste für den Schutz von Systemen und Daten, Dortmund 2013 (Informatik); Raley, Rita: *Dataveillance and Countervailance*, in: Gitelman, Lisa (Hg.): «Raw data» is an oxymoron, Cambridge (USA) 2013 (Infrastructures series), S. 121-145; Rogers, Russ: *The Basics of Information Security: Understanding the Fundamentals of Infosec in Theory and Practice*, 2014; Diverse Beiträge oder Sonderhefte von Fachzeitschriften, bspw.: c't wissen Überwachung abwehren, Heise Medien Verlag, Hannover 2015. Internetressourcen: Digitale Gesellschaft: URL: <https://www.digitale-gesellschaft.ch/>; Electronic Frontier Foundation EFF: <https://www.eff.org/>; Information Security Society Switzerland (ISSS). URL: <http://www.issss.ch/>; «Swiss Internet Security Alliance» (SISA). URL: <https://www.swiss-isa.ch/>.

⁴³¹ Siehe dazu bspw.: Grassegger, Hannes: Das Ende des Internets, in: NZZ am Sonntag, Zürich 09.02.2014, S. 24-26.

⁴³² Beispiel aus dem Analogen: DVDs wurden mit einem 'Regionencode' (eine Art von Verschlüsselung) versehen, was ein Abspielen in einem dafür nicht vorgesehenen Gerät verunmöglichte. In einer Region waren nur Geräte mit entsprechendem Code erhältlich, ein Parallelimport von Geräten und DVDs war sogar gesetzlich verboten. In der Schweiz währte dies allerdings nur kurze Zeit (Artikel 12 Abs. 1bis des Bundesgesetzes über das Urheberrecht und die verwandten Rechte (URG) vom 01.08.2002, revidiert auf 01.04.2004; Bericht des Bundesrates über Motionen und Postulate der gesetzgebenden Räte im Jahre 2003 vom 23.06.2004, 2002 P 02.3389).

Im Internet wird das Geoblocking auf der Basis der IP-Adressen umgesetzt. Ländern oder Regionen wurden bestimmte IP-Adressblöcke zugeordnet, die ganz oder teilweise blockiert werden können. Eine IP-Adresse ist eine rein numerische, in vier Gruppen zu 8 Bits (dezimal 0 bis 255) unterteilte Kennung für den Anschluss von Gerätekomponenten aller Art im Internet (Bspw.: 192.168.0.1.). Siehe: Fischer; Hofer: *Lexikon der Informatik*, 2011. Der Dienst 'utrace' (URL: <http://www.utrace.de/>) zeigt kartographisch die Lokalisierung von IP-Adressen an.

⁴³³ Zu den russischen Überwachungs- und Zensurmechanismen siehe die Zusammenfassung der Videokonferenz von Andrei Soldatov vor dem EP Committee on Civil Liberties, Justice and Home Affairs. Auszüge unter: European Commission: *Inquiry on Electronic Mass Surveillance of EU Citizens – Exchange of views on the Russian communications interception practices (SORM)*, Audiovisual Services, 22.01.2014, URL: <http://ec.europa.eu/avservices/video/shotlist.cfm?ref=95402>, Stand: 13.10.2015. Ausführlicher Report des Europäischen Parlaments zur Überwachung durch Geheimdienste siehe: Moraes, Claude: *Report on the US NSA surveillance programme, surveillance bodies in various Member States and their impact on EU citizens' fundamental rights and on transatlantic cooperation in Justice and Home Affairs*, Report A7-0139/2014, Committee on Civil Liberties, Justice and Home Affairs, Bruxelles 21.02.2014. URL: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+REPORT+A7-2014-0139+0+DOC+PDF+V0//EN>, Stand: 04.05.2018.

Zensur hauptsächlich politischer Natur,⁴³⁴ in Grossbritannien werden allgemein Inhalte, die für Kinder unter 18 Jahren nicht geeignet scheinen, blockiert.⁴³⁵ So hehr das Ziel in Grossbritannien sein mag, so problematisch ist die Umsetzung, weil die Filter unter anderem nicht nur pornographische Webinhalte, sondern auch Webauftritte zur sexuellen Aufklärung automatisch sperren. Anbieter müssen sich in *Whitelists*⁴³⁶ eintragen lassen und Bürger müssen sich aktiv melden und registrieren, wenn sie auf einen gesperrten Inhalt zugreifen möchten. Die britische Bürgerrechtsbewegung *Open Rights Group* hat nach einem halben Betriebsjahr der von der britischen Regierung eingeführten Filter festgestellt, dass rund 20% der beliebtesten Webseiten blockiert waren (Stand Juli 2014).⁴³⁷ Damit stellt dieses System einen kaum zumutbaren Eingriff in die Privatsphäre der Bürger dar und wird deshalb auch von der Bürgerrechtsbewegungen und der Presse stark kritisiert.

Der Trend zu Netzensur oder Netzsperrern scheint zuzunehmen. In der Schweiz wird zurzeit die Revision des Geldspielgesetzes, das die Sperrung von URLs ausländischer Spielbanken ermöglicht, intensiv diskutiert.⁴³⁸ Es wird befürchtet, dass mit diesem ersten Schritt Netzsperrern salonfähig und auf weitere Inhalte ausgeweitet würden. Dass Netzsperrern gemäss einer Studie von Florent Thouvenin und Burkhard Stiller (Universität Zürich) "faktisch unwirksam"⁴³⁹ sind, wird anscheinend zu wenig Beachtung geschenkt.

⁴³⁴ In China zensieren selbst grosse Unternehmen wie Springer ihr Angebot, Google hat sich mittlerweile aus dem chinesischen Markt wieder zurückgezogen. Siehe: Editorial Board: A prestigious research publisher gives in to China's censorship, *The Washington Post*, 04.11.2017, URL: https://www.washingtonpost.com/opinions/a-prestigious-research-publisher-gives-into-chinas-censorship/2017/11/04/4e4ff9de-c0c6-11e7-959c-fe2b598d8c00_story.html, Stand: 07.02.2018. In der Türkei sollen Betreiber einer Webseiten bald eine Lizenz benötigen. Siehe: Wittenhorst, Tilman: Zensur befürchtet: Türkei will Websites erst nach Lizenzvergabe zulassen, *Newsticker, heise online*, 06.02.2018, URL: <https://www.heise.de/newsticker/meldung/Zensur-befuerchtet-Tuerkei-will-Websites-erst-nach-Lizenzvergabe-zulassen-3961490.html>, Stand: 07.02.2018.

⁴³⁵ Siehe Berichte diverser Bürgerrechtsorganisationen und Forschungsinstitutionen, bspw.: Reporters Without Borders (URL: https://www.reporter-ohne-grenzen.de/feinde_internet/), OpenNet Initiative (URL: <https://opennet.net/>), Internet Monitor des Berkman Center for Internet & Society der Universität Harvard (USA) (URL: <https://thenetmonitor.org/>).

⁴³⁶ Whitelist: Positivliste mit vertrauenswürdigen Einträgen für privilegierten Zugang o.ä. Gegenteil: Blacklist (Liste gesperrter Inhalte/Zugänge).

⁴³⁷ Siehe: Cowburn, Pam: ORG's Blocked project finds almost 1 in 5 sites are blocked by filters, *Blog, Open Rights Group*, 02.07.2014, URL: <https://www.openrightsgroup.org/blog/2014/blockedproject>, Stand: 08.05.2015; Kannenberg, Axel: Bürgerrechtler: Britischer Pornofilter blockt fast jede fünfte Website, *Newsticker, heise online*, 03.07.2014, URL: <http://www.heise.de/newsticker/meldung/Buergerrechtler-Britischer-Pornofilter-blockt-fast-jede-fuenfte-Website-2248673.html>, Stand: 08.05.2015.

⁴³⁸ Siehe zum laufenden Rechtssetzungsprojekt des Schweizerischen Bundesamtes für Justiz unter URL: <https://www.bj.admin.ch/bj/de/home/wirtschaft/gesetzgebung/geldspielgesetz.html>.

⁴³⁹ Stiller, Burkhard; Thouvenin, Florent: Gutachten: Netzsperrern, 16.09.2016, S. 18. URL: <http://www.itsl.uzh.ch/dam/jcr:9607879d-3e04-4de2-b25a-bf46fa58d5ac/Gutachten%20Netzsperrern-final.pdf>, Stand: 26.04.2018. Der Informatiker und Netzaktivist Markus Ritzmann überprüft regelmässig die Erreichbarkeit von über einer Mio. URLs und trägt die Ergebnisse in eine 'Zensur-Datenbank' ein. Siehe URL: <https://dnszensur.ch/>.

5.1.3. Suchtechniken

*"We are in a crisis of knowledge"*⁴⁴⁰ (Weinberger, 2011)

Welche Techniken sollen nun verwendet werden, um sich in den riesigen digitalen Datenbeständen zurechtzufinden? Ist das verfügbare Wissen tatsächlich 'too big to know' wie der Philosoph und Internetforscher David Weinberger behauptet? Haben wir einen Wissensüberfluss oder sogar zu viel Wissenschaft? Weinberger warf diese Fragen auf und meinte, dass wir uns in einer *Wissenskrise* befänden.⁴⁴¹ Ich stimme Weinberger zu, dass die Gesellschaft und auch die Wissenschaft an *Informationsüberfluss* leidet, denn die Fülle an bestehendem und stets zunehmendem Wissen ist für eine Person schon lange nicht mehr zu bewältigen, wenn es sich nicht um ein sehr spezifisches Wissensgebiet handelt. Zudem werden durch die Wissenszunahme die Fachkanons infrage gestellt, und es werden immer wieder neue Quellenarten und Suchräume eingeführt, was immer komplexere Suchtechniken erfordert.

Dem Problem des *Information Overload* kann aber mit geeigneten Suchtechniken begegnet werden. Diese Techniken *filtern* Informationen nach bestimmten Kriterien und reduzieren es damit auf einen für den Nutzer bearbeitbaren Umfang. Dazu sind allerdings genaue, computertechnisch auswertbare und damit strukturierte Metadaten über die vorhandenen Informationen notwendig, was die Informationsmenge noch einmal vergrößert. Auch für den Kommunikationswissenschaftler Klaus Merten ist dies die Lösung für das Problem des Informationsüberflusses:

*"Die Überforderung durch ein Zuviel an Information lässt sich nur durch ein[e] weitere Überforderung, durch ein Nachmehr an Information – freilich besonderer Art – reduzieren."*⁴⁴²

Damit sind die eine Informationsgesellschaft charakterisierenden Meta-Medien gemeint, die durch Selektion und Zusammenfassung von Informationseinheiten einen Überblick zu verschaffen versprechen. Dazu gehören beispielsweise Manuals, Abstracts oder Verzeichnisse von Suchmaschinen, die schon zu den Metametamedien gezählt werden können. Spezifische Suchtechniken für die Recherche in digitalen Datenbeständen versprechen den Informationsüberfluss bewältigen zu können.

Unter *Suchtechnik* versteht man die Fähigkeit, die zur richtigen Ausübung einer Recherche notwendig ist. Suchtechniken mit und in digitalen Informationssystemen unterscheiden sich von analogen. Sie erfordern deshalb neue Kompetenzen, vor allem in der Beurteilung der

⁴⁴⁰ Weinberger: Too Big To Know, 2011, S. 173.

⁴⁴¹ Ebd.

⁴⁴² Merten, Klaus: Zur Ausdifferenzierung der Mediengesellschaft. Wirklichkeitsmanagement als Suche nach Wahrheit, in: Arnold, Klaus; Tonnemacher, Jan (Hg.): Alte Medien – neue Medien: Theorienperspektiven, Medienprofile, Einsatzfelder. Festschrift für Jan Tonnemacher, Wiesbaden 2005 (Public relations), S. 23.

Suchresultate. Insbesondere wird nachfolgend aufgezeigt, dass bei der Verwendung von digitalen Suchsystemen bisher nicht bekannte Probleme auftreten, weshalb diesen Systemen nicht vertraut werden sollte.

Digitale Suchtechniken hängen vor allem von den vorhandenen Metadaten ab, die entweder durch eine Institution oder Firma oder durch Private angelegt wurden. Tendenziell ist die Qualität der Metadaten bei Gedächtnisinstitutionen bedingt durch Standardisierungen und wissenschaftliche Richtlinien eher höher als bei Privaten. Beispielsweise werden Schlüsselbegriffe in Bibliotheken jeweils nach einem bestimmten und einheitlichen Schema vergeben, was bei Privaten eher selten der Fall ist. Diese hohe Qualität der Metadaten ist jedoch gefährdet, wenn die Institutionen die Indexierung und den Suchdienst als Teile ihrer Kernkompetenzen nicht mehr weiterführen und dafür externe Anbieter wie den Indexdienst von Google Search einsetzen. Automatisierte Indexdienste kommerzieller Anbieter arbeiten keineswegs besser oder genauer, denn "Metadaten (die nicht nur zur korrekten Erfassung von Publikationen, sondern auch für ihr Ranking entscheidend sind) erweisen sich häufig als fehlerhaft."⁴⁴³ Folglich sind Suchresultate aus solchen Diensten kritisch zu hinterfragen, wenn sie denn als solche überhaupt zu erkennen sind.

Drei Hauptarten von aktiven Suchtechniken können unterschieden werden, deren Einsatz von der Auswahl des Suchraums und allenfalls vorhandener Systematik abhängig ist:

Bei einer *unsystematischen / ungerichteten Suche* werden mittels Suchmaschinen grosse / riesige Datenbestände grosser Suchräume durchsucht (Indexdatenbanken, Big Data etc.).

Bei einer *systematischen / gerichteten Suche* werden kleinere Bestände bestimmter Institutionen in spezifischen Informations- oder Fachsuchräumen mittels Datenbanken, Archiv- und Bibliothekskatalogen⁴⁴⁴ und ähnlichem durchsucht.⁴⁴⁵

Es kann auch eine unsystematische Suche in einem eingegrenzten Suchraum erfolgen. Falls dieser an sich schon eine bestimmte Systematik mit einem zu erwartenden Suchresultat aufweist, wie beispielsweise eine Fachdatenbank, kann in Kombination der vorgenannten Suchtechniken von einer *halbsystematischen Suche* gesprochen

⁴⁴³ König, René; Nentwich, Michael: Cyberscience 2.0: Wissenschaftskommunikation in der Beta-Gesellschaft, in: kommunikation @ gesellschaft 15 (Sonderausgabe), 2014, S. 26. Unter Verweis auf: Jacso, Peter: Google Scholar's Ghost Authors, Library Journal, URL: <http://lj.libraryjournal.com/2009/11/industry-news/google-scholars-ghost-authors/>, Stand: 30.12.2013; Nunberg, Geoffrey: Google's Book Search: A Disaster for Scholars, in: The Chronicle of Higher Education, 31.08.2009. The Chronicle of Higher Education, URL: <http://chronicle.com/article/Googles-Book-Search-A/48245/>, Stand: 01.10.2015.

⁴⁴⁴ Ebenfalls als Datenbank aufgebaut.

⁴⁴⁵ Zu Vor- und Nachteilen siehe Einführungsliteratur Geschichtsstudium, bspw.: Kühmstedt: Klug recherchiert (Hist), 2013.

werden. Da die halbsystematische Suche eine Kombination oben genannter Techniken ist, wird auf weitere Ausführungen verzichtet.

Systematische Suche

Sowohl bei physischen wie auch bei digitalen *Such- und Findmitteln* ist das Vertrauen in die Vollständigkeit und Integrität der zur Verfügung gestellten Mittel von grosser Wichtigkeit. Gedächtnisinstitutionen wie Bibliotheken oder Archiven wurde bis anhin per Konvention vertraut – man konnte davon ausgehen, dass alle verfügbaren Objekte nach objektiven Kriterien ausgewählt und in den Bestand übernommen wurden und dass die Objekte vollständig und korrekt in den Findmitteln verzeichnet sind. Beispielsweise legt der seit 1996 bestehende Schweizer Kodex für Archivare fest, dass die Integrität von Archivgut zu schützen ist, damit es ein zuverlässiger Beweis der Vergangenheit bleibt.⁴⁴⁶ Nach dem Kodex besteht die wichtigste Aufgabe der Archivarinnen und Archivare darin, die Unversehrtheit der von ihnen verwalteten und verwahrten Unterlagen zu erhalten. Dies trifft auch für die zur Verfügung gestellten Findmittel zu, was auch digitale Informationssysteme einschliesst. Der Nutzer muss und kann darauf vertrauen, dass die Datenbank(en) aktuell ist / sind, der Suchprozess dem vorgeschriebenen Prozess entspricht und alle vorhandenen Daten integer und vollständig sind. Dies ist durch die Zusammenarbeit zwischen Informatikspezialisten und Archivaren, Bibliothekaren und Historikern zu gewährleisten, was aber noch nicht vorausgesetzt werden darf.⁴⁴⁷

Die in den meisten Bibliotheks- und Archivinstitutionen eingesetzten digitalen Findmittel sind Datenbanksysteme, die via Terminal und/oder Webseite abgefragt werden können. Damit ein solches Informationssystem dafür genutzt werden kann, ist zunächst der Zugriff darauf zu gewährleisten. Im Unterschied zu physischen Findmitteln reicht das blosse Vorhandensein des Systems nicht, es benötigt entsprechende Hard- und Software sowie Strom. Ist der Zugriff gewährleistet und kann das System genutzt werden, hat sich der Suchende darüber im Klaren zu sein, wie das System funktioniert, um das Zustandekommen des Resultats zu rekonstruieren.⁴⁴⁸ Die Suchmöglichkeiten, beispielsweise mit Booleschen Operatoren, Synonymschreibweisen, Selektionsfiltern oder *Crowd Tagging*, sind bei digitalen Systemen mannigfaltiger als bei physischen und das Nachvollziehen einer Resultatliste ist dementsprechend schwieriger. In einem physischen Karteikastensystem ist der Bestand nach bestimmten

⁴⁴⁶ Siehe: Verein Schweizerischer Archivarinnen und Archivare (VSA): Kodex ethischer Grundsätze, VSA – Verein Schweizerischer Archivarinnen und Archivare, 1997, S. §1, URL: <http://vsa-aas.ch/beruf/ethikkodex/>, Stand: 06.05.2018.

⁴⁴⁷ Hier spielt die Kommunikationsfähigkeit der beteiligten Personen eine grosse Rolle, denn sie müssen eine gemeinsame Sprache finden, was grundlegende Kenntnisse der anderen Fachbereiche voraussetzt. Personen, die in beiden Fachgebieten zu Hause sind, sind (leider) rar.

⁴⁴⁸ Siehe dazu: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 204.

Kriterien wie Autorname oder Stichwort sortiert, die nur einzeln bearbeitet werden können. In digitalen Systemen hingegen werden die Bestandeseinheiten als Entitäten in einer Datenbank mit definierten Kriterien erfasst und indexiert. Die Kriterien können miteinander verknüpft werden. Eine Entität wird gelistet, wenn das Kriterium oder die Kriterien bei einem Abfrageprozess zutrifft oder zutreffen. Zwar kann so die Resultatliste möglicherweise präziser und schneller erstellt werden, aber das Vorgehen birgt die Gefahr, dass durch Kombination oder Ausschluss von Stichworten relevante Einträge nicht gelistet werden. Die Resultatliste hängt nicht mehr wie bisher hauptsächlich von der Kriterienvergabe oder Verschlagwortung durch Mitarbeiter der Institution ab, sondern im Wesentlichen von den Kenntnissen des Recherchierenden über die Funktionalitäten des Informationssystems – seinen informationstechnischen Kompetenzen.

Besondere Aufmerksamkeit erfordern *gewichtete Resultatlisten*, die Einträge nach ihrer algorithmisch berechneten Relevanz nacheinander aufführen. Eine gewichtete Liste führt die Einträge nach Abnahme der Wahrscheinlichkeit eines relevanten Verweises auf der Basis eines stochastischen Datenmodells auf. Wie die Gewichtung zu Stande kommt, ist selten nachvollziehbar, weshalb die Liste als *Relevanzkonstruktion* zu betrachten ist. Auskunft geben könnten nur die Entwickler oder Programmierer des eingesetzten Informationssystems. Allerdings unterliegt diese Information in sehr vielen Fällen dem Geschäftsgeheimnis. Zwar kann eine gewichtete Liste vorteilhaft sein, weil nicht jeder Verweis auf seine tatsächliche Relevanz zu überprüfen ist.⁴⁴⁹ Der Vorteil schwindet aber, wenn das Zustandekommen der Gewichtung nicht bekannt ist und die Relevanzprüfung deshalb trotzdem durchgeführt werden muss. Zudem kann das Datenmodell einer Gewichtung relativ schnell geändert werden, was sowohl in Fachkatalogen als auch in Webkatalogen grössere Auswirkungen auf die Rangierung in der Ergebnisliste haben kann. Änderungen können beispielsweise die Entfernung eines Eintrags bewirken, wie bei der Suchmaschine Google Search, die seit dem 21. April 2015 eine Webseite stärker gewichtet, wenn sie für Mobilgeräte optimiert ist.⁴⁵⁰ Noch nicht optimierte Webauftritte sollten schnellstmöglich umgestaltet werden, was meist einen nicht unerheblichen Aufwand bedeutet. Wenn sie nicht angepasst werden, kann der Traffic⁴⁵¹ um ca. 10% sinken, was als 'Mobilegeddon' bekannt ist.⁴⁵² Wissenschaftliche Fachportale oder Institutionen sind davon ebenfalls betroffen und rangieren nun in den

⁴⁴⁹ Vgl.: Ebd., S. 219.

⁴⁵⁰ Makino, Takaki; Phan, Doantam: Rolling out the mobile-friendly update, Official Google Webmaster Central Blog, 21.04.2015, URL: <http://googlewebmastercentral.blogspot.com/2015/04/rolling-out-mobile-friendly-update.html>, Stand: 13.10.2015.

⁴⁵¹ Traffic: Anzahl Webseitenbesucher. Siehe: Grabs, Anne; Bannour, Karim-Patrick: Follow me!: Social Media Marketing mit Facebook, Twitter, XING, YouTube und Co. Inkl. Empfehlungsmarketing, Crowdsourcing und Social Commerce, 2011, S. 31.

⁴⁵² Adobe Systems Incorporated: Digital Advertising Report. Adobe Digital Index Q2 2015, S. 4. URL: http://www.cmo.com/content/dam/CMO_Other/ADI/ADI_Mobilegeddon/Q2-ADI-Digital-Ad-Report.pdf, Stand: 24.07.2015.

Resultatlisten weiter hinten. Damit wird es schwieriger, relevante und verlässliche Informationen zu finden, weil die Nutzer sich nur mit den ersten Ergebniseinträgen intensiv beschäftigen und schon die zweite Seite einer Resultatliste kaum mehr beachten.⁴⁵³

Unsystematische Suche / Suchmaschinen

Eine unsystematische Suche beginnt in den allermeisten Fällen im Web mit einer *webbasierten Suchmaschine* oder, anders gesagt, wenn eine Suche mit einer Suchmaschine erfolgt, ist sie unsystematisch. Bei allen Online-Suchmaschinen basiert das Resultat einer Anfrage auf einer Suche innerhalb des Indexes des Suchmaschinenanbieters und nicht auf der Suche im gesamten Web.⁴⁵⁴ Es wird deshalb nur das gefunden, was die Datenbank zu diesem Zeitpunkt beinhaltet und der Algorithmus anhand der Suchbegriffe als relevant erachtet. Es kann demnach sein, dass signifikante Informationen nicht gefunden werden, denn wenn sie im Deep Web oder zwischen zwei Durchläufen eines Crawler⁴⁵⁵ des Suchmaschinenanbieters veröffentlicht wurden, ist im Index der Suchmaschine nichts verzeichnet. Die meisten Crawler durchsuchen das Web nur bis zu einer bestimmten Strukturtiefe, meistens der dritten Ebene eines Webauftritts. Eine Suchmaschine ist daher nicht vollständig und auch nicht objektiv.⁴⁵⁶

Der *Index* einer Suchmaschine ist eine Datenbank, die Textinformationen aus Textelementen oder Metadaten enthält. Auch eine Suche nach digitalen Bildern, Film- oder Tonobjekten erfolgt hauptsächlich noch durch das automatische Auslesen der Metadaten des Objekts und eventuell den auf derselben Seite vorhandenen Text. Das Erkennen eines solchen Objekts durch eine Suchmaschine steht noch am Anfang der Entwicklung.⁴⁵⁷ Spezielle Bildersuchmaschinen⁴⁵⁸ die ähnliche Bilder wie ein Vorlagenbild sucht, basieren nicht auf einer inhaltlichen, sondern auf einer Farbanalyse der Vorlage. Die Pixelanzahl einer Farbe

⁴⁵³ Siehe: Petrescu, Philip; Ghita, Mircea; Loiz, Dana: Google Organic CTR Study, Studie, Caphyon / Advanced Web Ranking, 2014, S. 24. URL: <https://www.advancedwebranking.com/google-ctr-study-2014.html>, Stand: 28.03.2016.

⁴⁵⁴ Zur Indexierung siehe: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 218.

⁴⁵⁵ Crawler oder Searchbot: selbständiges Programm, das das Internet nach Informationen durchsucht. Gefundene Inhalte werden an den Betreiber gesendet, der anschliessend Auswertungen vornimmt. Siehe: Prevezanos: Computer Lexikon 2013, 2012.

⁴⁵⁶ Vgl.: Burkhardt: Digitale Datenbanken, 2015, S. 13.

⁴⁵⁷ Erste Erfolge durch: Karpathy, Andrej; Fei-Fei, Li: Deep Visual-Semantic Alignments for Generating Image Descriptions, Stanford University, Stanford 11.2014, S. 10. URL: <http://cs.stanford.edu/people/karpathy/deepimagesent/devisagen.pdf>, Stand: 19.11.2014; Vinyals, Oriol; Toshev, Alexander; Bengio, Samy u. a.: Show and Tell: A Neural Image Caption Generator, arXiv – Cornell University Library, arXiv:1411.4555 [cs.CV], 17.11.2014. URL: <http://arxiv.org/abs/1411.4555>, Stand: 19.11.2014. Im 'The Wolfgang Language Image Identification Project' wird ebenfalls eine Bildinhaltserkennung durchgeführt. URL: <https://www.imageidentify.com/>.

⁴⁵⁸ Bspw. durch tineye (URL: <http://www.tineye.com/>), Google (URL: <https://images.google.com/>).

und deren Verteilung werden bestimmt, worauf anhand dieser Werte weitere Bilder gesucht werden.

Was auch immer gesucht wird, die im Index enthaltenen Daten und der eingesetzte *Algorithmus* des Suchmaschinenbetreibers bestimmen, was an welcher Stelle der Resultatliste erscheint. Dieser Algorithmus *muss* ständig optimiert werden, einerseits, um Manipulationsversuche zur Beeinflussung der Gewichtung abzuwehren⁴⁵⁹, andererseits zur Fehlerbereinigung und Ergänzung. Beispielsweise wurden in Google Scholar 2009 die Begriffe *Password*, *P Login*, *Subscribe* und *Ltd* als sehr produktive und oft zitierte Autoren aufgeführt.⁴⁶⁰ Mittlerweile ist das Problem allerdings fast behoben.

Fehler können auch aufgrund falscher oder falsch interpretierter Metadaten entstehen, wie bei der Lizentiats- respektive Masterarbeit von Pascal Föhr über *Jakobinische Tendenzen in Meinungsbildenden Sozietäten der Deutschschweiz zwischen 1789 und 1798*, die in Google Scholar am 01.08.2016 als Dissertation geführt wird.

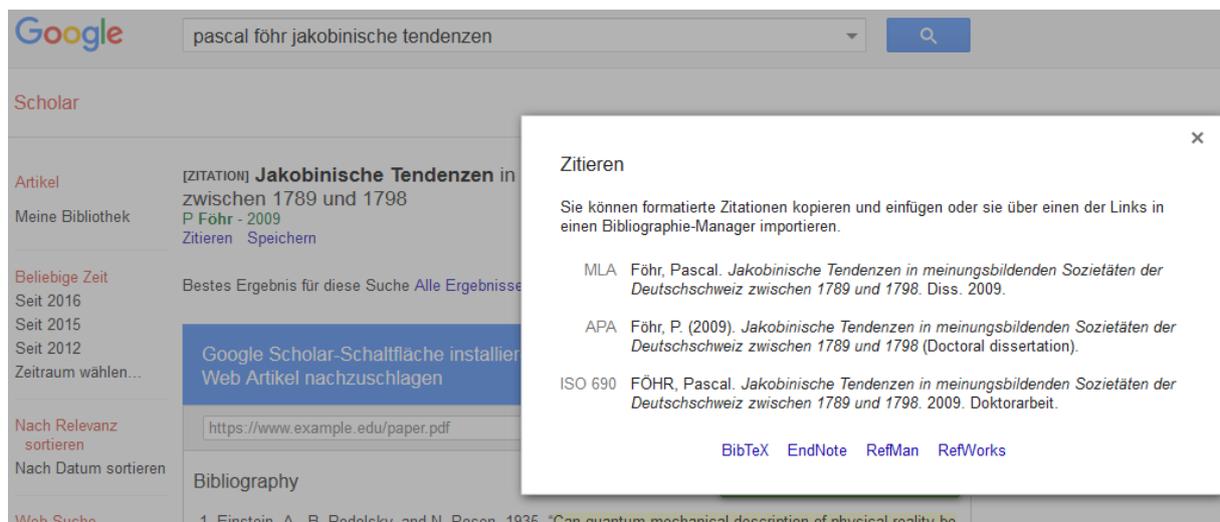


Abbildung 7: Google Scholar, Metadaten der Lizentiatsarbeit von Pascal Föhr (Google Scholar)⁴⁶¹

Bei einer Suche in einer Datenbank ist die Semantik des Suchbegriffs relevant. Je spezifischer der Suchbegriff, der sogenannte search string, unter Verwendung von Booleschen Operatoren formuliert wird, desto kürzer ist die Resultatliste. Dies bedeutet aber nicht zwangsläufig eine höhere Qualität, denn die Suche nach 'Big Data' ergibt je nach Schreibweise und verwendeten Operatoren wie *Big Data*, *"Big Data"* oder *Big NEAR Data* lediglich

⁴⁵⁹ Unter Suchmaschinenoptimierung SEO werden Massnahmen von Webaufttrittsbetreibern zur Verbesserung der Rangierungsposition in Suchmaschinen verstanden. Gewisse Anpassungen sind von den Suchmaschinenbetreibern sogar erwünscht.

⁴⁶⁰ Jacso: Google Scholar's Ghost Authors. Nur noch Ltd wird am 01.10.2015 mit rund 590'000 Ergebnissen signifikant als Fehler geführt.

⁴⁶¹ Bildzitat Google Scholar (URL: <https://scholar.google.ch/>); URL: https://scholar.google.ch/scholar?lookup=0&q=pascal+f%C3%B6hr+jakobinische+tendenzen&hl=de&as_sdt=0,5#d=gs_cit&p=&u=%2Fscholar%3Fq%3Dinfo%3AyY091e5_uuwJ%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D0%26hl%3Dde. Stand: 01.08.2016.

unterschiedliche Resultate.⁴⁶² Die technische Handhabung einer Suchmaschine ist für den Recherchierenden zentral, denn jede Zeichenfolge hat Auswirkungen auf die Resultatliste.⁴⁶³ Gefunden werden kann nur das, was sich in der Datenbank befindet und das, was die vorhandene Suchtechnologie mittels ihrer Abfragesprache (Suchsyntax) ermöglicht.⁴⁶⁴ Je grösser die Kenntnisse eines Nutzers über digitale Suchtechnologien sind, desto nützlicher sind Suchresultate. Aber auch ein kompetenter Nutzer kann nicht sicher sein, dass alle relevanten Objekte gelistet sind. Da weder die potenziell auffindbaren Informationen, noch die Regeln für die Zuordnung der Relevanz und die Selektion bekannt sind, ist eine Suchmaschine als *Black Box* zu betrachten.⁴⁶⁵

Die meistbenutzte aller Suchmaschinen ist unbestritten Google Search.⁴⁶⁶ Anlässlich der Summer School of Switzerland, Berne 2013, stellte ich den Teilnehmenden meines Workshops eine Rechercheaufgabe zum Thema *History of the internet*.⁴⁶⁷ Auf Nachfrage hin habe ich festgestellt, dass von den etwa 25 Teilnehmenden rund 80% dafür als erstes Google Search verwendet und ausschliesslich vielversprechende Einträge von der ersten Seite der Resultatliste aufgerufen hatten.⁴⁶⁸ Diese Zahl spiegelt in etwa den statistischen Marktanteil von Google Search wieder, der im deutschsprachigen Raum bei ca. 95% (2013/2014), global bei ca. 77% (2013) / 68% (2014) liegt.⁴⁶⁹

Google hat das ehrgeizige Ziel "die Informationen der Welt zu organisieren und für alle zu jeder Zeit zugänglich und nützlich zu machen."⁴⁷⁰ Die Dienste⁴⁷¹ eines Privatunternehmens mit der Absicht, *die Informationen der Welt zu organisieren*, sollten von der Wissenschaft stets mit kritischem Blick genutzt werden. Googles Dienstleistungen stehen zwar grundsätzlich kostenlos zur Verfügung, dennoch erwirtschaftet das Unternehmen Milliardengewinne. Die Informationen der Welt kostenlos zu organisieren, scheint folglich lukrativ zu sein, was die Skepsis gegenüber der wissenschaftlichen Verwendbarkeit der durch Google zur Verfügung

⁴⁶² *Suchbegriff mit Sonderzeichen*.

⁴⁶³ Nach: Burkhardt: Digitale Datenbanken, 2015, S. 14.

⁴⁶⁴ Siehe: Ebd., S. 179.

⁴⁶⁵ Vgl.: Ebd., S. 260.

⁴⁶⁶ Aufgrund der marktbeherrschenden Stellung von Google wird die Problematik anhand dieser Suchmaschine aufgezeigt. Die in der Beliebtheit folgenden Suchmaschinen Bing, Yahoo und Baidu sind vergleichbar problematisch.

⁴⁶⁷ Summer School of Switzerland, Berne 2013. Workshop: Föhr, Pascal: Historical Sources Criticism in the Digital Age, 28.06.2013.

⁴⁶⁸ Siehe auch: Nentwich: Cyberscience 2.0, 2012, S. 128.

⁴⁶⁹ Die Zahlen divergieren je nach Studie und Verfasser. Die zwei wichtigsten Anbieter für Webstatistiken: Netmarketshare (URL: <http://netmarketshare.com/>) und StatCounter (URL: <http://statcounter.com/>). Genannte Zahlen nach Netmarketshare, Jahr 2014.

⁴⁷⁰ Google, Unternehmen, URL: <http://www.google.com/intl/de/about/company/>, Stand: 24.12.2013.

⁴⁷¹ Bspw.: Search (Suchmaschine); Youtube (Video); Google Scholar und Google Books (akademische Publikationen und allg. Bücher); Google Maps, Earth und Streetview (Geographie); Google Docs ([kollaborative] Wissensproduktion); Gmail und Google+ (Kommunikation und Social Media Plattform); [mobile] Applikationen und Android-OS; Google Phone Nexus (Smartphone); AdWords (Werbung) etc.

gestellten Informationen erhöhen sollte.⁴⁷² Die angebrachte Skepsis scheint aber unter Akademikern genauso wenig verbreitet zu sein wie unter anderen Nutzern.⁴⁷³ Eine Untersuchung unter US-amerikanischen Lehrern aus Middle und High Schools hat ergeben, dass sowohl Lehrer als auch Studenten der Ansicht sind, "'research' means 'Googling'".⁴⁷⁴ Zum selben Ergebnis kommen auch die Kommunikationswissenschaftlerin Birgit Stark in ihrer Untersuchung zur Suchmaschinennutzungskompetenz sowie der Historiker Max Kemman in seinem Werk *Digital Research Practices of Humanities Scholars*.⁴⁷⁵ Laut Stark ist für die meisten Studienteilnehmenden Google Search eine "automatisierte (Such)Maschine, wobei sie implizit Neutralität und Objektivität voraussetzen."⁴⁷⁶ Kemman konstatiert:

*"Google is the key player among available search engines. This dominant use illustrates the paradoxical attitude of scholars toward Google: while provenance and context are deemed key academic requirements, the workings of the Google algorithm remain unclear."*⁴⁷⁷

Des Weiteren stellt er fest, dass durch die Verwendung von Google Search mit seinem unbekanntem Algorithmus zwar eine Black Box im Suchprozess vorhanden sei, diese aber das Vertrauen in die Suchresultate bei den meisten Studenten nicht negativ beeinflusse.⁴⁷⁸

Die Problematik der oben genannten Gewichtung wird verschärft, wenn der Nutzer sich im Google-System anmeldet, denn die Suchanfragen werden nutzerbezogen protokolliert und bei der nächsten Abfrage mit einbezogen. In Google Search werden deshalb zwei Personen am selben Standort trotz der Verwendung derselben Suchbegriffe unterschiedliche Resultatlisten erhalten. Zudem werden dem Nutzer mit der Autocomplete-Funktion bekannte Begriffe aus seinem jeweiligen Abfrageverlauf vorgeschlagen, was zwar eine Hilfe sein kann, aber auch die Kreativität einschränkt. Die Geisteswissenschaftlerin Anna Jobin bezeichnet die Autocomplete-Funktion sogar als Ort der Produktion von Geschichtsbewusstsein, weil die

⁴⁷² Die Bemerkungen sind keine Kritik an der Unternehmung Google Inc. oder der Suchmaschine. Aber jeder Forschende hat kritisch zu hinterfragen, was für Konsequenzen der monopolartige Einfluss von Google auf seine Arbeit hat. Kritik an Google siehe bspw.: Vaidhyanathan, Siva: *The Googlization of everything (and why we should worry)*, Berkeley 2011; Perri, Pascal (Hg.): *Google un ami qui ne vous veut pas que du bien*, Paris 2013.

⁴⁷³ Nach: Nentwich: *Cyberscience 2.0*, 2012, S. 128.

⁴⁷⁴ Purcell, Kristen; Rainie, Lee; Heaps, Alan u. a.: *How Teens Do Research in the Digital World. A survey of Advanced Placement and National Writing Project teachers finds that teens' research habits are changing in the digital age*, Pew Internet & American Life Project, 01.11.2012, S. 3. URL: <http://www.pewinternet.org/Reports/2012/Student-Research.aspx>, Stand: 30.12.2013.

⁴⁷⁵ Stark, Birgit; Dörr, Dieter; Aufenanger, Stefan: *Die Googleisierung der Informationssuche – Suchmaschinen im Spannungsfeld zwischen Nutzung und Regulierung*, Universität Mainz, Mainz 16.12.2014; Kemman, Max; Kleppe, Martijn; Scagliola, Stef: *Just Google It – Digital Research Practices of Humanities Scholars*, in: Mills, Clare; Pidd, Michael; Ward, Esther (Hg.): *Proceedings of the Digital Humanities Congress 2012. Studies in the Digital Humanities – HRI Online Publications*, arXiv:1309.2434 [cs.DL], Sheffield 2014. URL: <http://arxiv.org/abs/1309.2434>, Stand: 06.08.2015.

⁴⁷⁶ Stark; Dörr; Aufenanger: *Die Googleisierung der Informationssuche*, 2014, S. 3.

⁴⁷⁷ Kemman; Kleppe; Scagliola: *Just Google It*, 2014, S. 1.

⁴⁷⁸ Siehe: Ebd., S. 16.

Vorschläge neue und teilweise falsche Zusammenhänge herleiten.⁴⁷⁹ Jobins Erkenntnis entspricht den oben erwähnten Untersuchungen, weshalb ihr zuzustimmen ist. Der Nutzer befindet sich schon bei der Eingabe in einer *Filterblase*.⁴⁸⁰

Weitere Unterstützung erhält der Nutzer unter anderem durch die *IP-to-geo-Technologie* oder *address location technology*. Zum Beispiel wird er in einem digitalen Informationssystem mit Schweizer IP-Adresse nach der Eingabe der URL *google.com* in das Adressfeld des Browsers automatisch auf *google.ch* umgeleitet. Damit werden die Einträge der Resultatliste aus der Region des Nutzers viel stärker gewichtet als andere, und unter Umständen werden andere Einträge wiederum nicht angezeigt, obwohl sie grundsätzlich zugänglich wären. Dies ist nur einer von über 200 Faktoren, die Google Search für die Ermittlung einer Suchanfrage verwendet.⁴⁸¹

Für die meisten Teilnehmer der Stark'schen Studie scheinen diese 'Hilfen' klare Vorteile zu bringen, die allfällige Bedenken bezüglich Datensammlung und Personalisierung aufwiegen. Für eine wissenschaftliche Recherche ist eine solche Haltung jedoch nicht angemessen, denn die Resultate einer personalisierten Suche verstossen gegen das Objektivitätsprinzip. Die Auswahl der potenziellen Quellen erfolgt nicht mehr durch den Recherchierenden allein, sondern wird durch die Auswertung der Präferenzen eines Nutzers algorithmisch präjudiziert. Je mehr Informationen über den Nutzer zur Verfügung stehen, desto kleiner wird seine Filterblase und damit der extern bestimmte Suchraum. Google Search wählt die anzuzeigenden Resultate seit dem 4. Dezember 2009 anhand von 57 Merkmalen wie Standort des Nutzers oder verwendeter Browser aus, bevor die Liste präsentiert wird.⁴⁸² Durch diese Vorfilterung der Suchresultate erhält der Nutzer in seiner Resultatliste nur das aufgeführt, was er vermeintlich zu sehen wünscht. Auch Social Media Plattformen wie Twitter filtern Inhalte, wie die Untersuchungen der Politikwissenschaftler Mathias und Wolfgang König ergaben.⁴⁸³ Dem Nutzer werden bei einer Suche nicht alle Tweets angezeigt, nur diejenigen,

⁴⁷⁹ Jobin, Anna: La mémoire kaléidoscopique : l'histoire au prisme des algorithmes d'autocomplétion, Panel: Le pouvoir des algorithmes. Technologies numériques et rapports de force au XXe-XXIe siècles, Lausanne 09.06.2016. URL: <https://geschichte.ch/referat/77/la-memoire-kaleidoscopique-l-histoire-au-prisme-des-algorithmes-d-autocompletion>, Stand: 26.06.2016.

⁴⁸⁰ Zur Filterblase oder 'filter bubble' siehe: Pariser, Eli: Filter Bubble: Wie wir im Internet entmündigt werden, München 2012.

⁴⁸¹ Nach: Google: Algorithmen – Alles über die Suche, Google, URL: <http://www.google.com/intl/de/insidesearch/howsearchworks/algorithms.html>, Stand: 30.12.2013; Burkhardt: Digitale Datenbanken, 2015, S. 264 f. Bsp. für Änderung der Faktoren: Makino, Takaki; Jung, Chaesang; Phan, Doantam: Finding more mobile-friendly search results, Official news on crawling and indexing sites for the Google index, Official Google Webmaster Central Blog, 26.02.2015, URL: <http://googlewebmastercentral.blogspot.com/2015/02/finding-more-mobile-friendly-search.html>, Stand: 17.06.2015.

⁴⁸² Die Anzahl und Merkmale variieren seither. Vgl. Pariser: Filter Bubble, 2012, S. 10.

⁴⁸³ König, Mathias; König, Wolfgang: Twitter kontrolliert Auffindbarkeit von Tweets, Praxis, politik & kommunikation, 10.03.2016, URL: <http://www.politik-kommunikation.de/ressorts/artikel/twitter-kontrolliert-auffindbarkeit-von-tweets-1486134101>, Stand: 26.06.2016.

die gemäss Twitter relevant sind. Der Algorithmus, der die Relevanz ermittelt, ist jedoch nicht nachvollziehbar.

Der von einer Suchmaschine verwendete Algorithmus wird aus Gründen des Geschäftsgeheimnisses vom Suchmaschinenanbieter nicht offengelegt. Selbst Suchmaschinen, die keine Gewichtung oder eine besondere Neutralität versprechen⁴⁸⁴, geben diesen aus wirtschaftlichen Gründen nicht bekannt. Findige Nutzer würden ansonsten eventuell vorhandene Lücken ausnutzen, um mit entsprechenden optimierenden Massnahmen wie sogenannte Linkfarmen ihr Webangebot möglichst weit oben in der Liste platzieren zu können. Zusätzlich wird das Problem durch Meta-Suchmaschinen⁴⁸⁵ verstärkt, die keinen eigenen Index erstellen, sondern die Indizes anderer Suchmaschinenanbieter verwenden. Der Suchalgorithmus greift auf verschiedenartig erstellte Datenbanken mit eventuell bereits gewichteten Daten zurück, die nochmals neu gewichtet werden.

Der Aphorismus 'quod non est in internet, non est in mundo'⁴⁸⁶ oder etwas zugespitzt 'quod non est in *Google*, non est in mundo' bringt es auf den Punkt: Wenn als einziger Suchraum das Web verwendet wird und der Zugang dazu nur über die Dienstleistungen von Google führt, dann sind sehr viele potenzielle Quellen verloren, weil Google kontrolliert, was genutzt wird. Der Medienwissenschaftler Siva Vaidhyanathan erklärt die Problematik präzise mit dem Hinweis, dass Google zunehmend die Linse ist, wie die Gesellschaft die Welt sieht:

*"Its process of collecting, ranking, linking, and displaying knowledge determines what we consider to be good, true, valuable, and relevant. The stakes could not be higher."*⁴⁸⁷

Weil Suchmaschinen einen so wichtigen Stellenwert in der täglichen Arbeit einnehmen, ist das kritische Hinterfragen der Resultate eminent wichtig, aber kaum möglich. Die Funktionsweisen und die verwendeten Algorithmen respektive die Rankingfaktoren mit deren Gewichtung sowie der verwendete Index einer Suchmaschine müssten durch die Betreiber offengelegt werden, was aus den oben teilweise genannten Gründen kaum möglich ist.

Es existiert eine ungesicherte Liste der Rankingfaktoren, die von Google Search verwendet werden könnten. Angaben zur Gewichtung der Faktoren sind nicht vorhanden oder nur ungenau. Im Folgenden die ersten zwanzig Faktoren.⁴⁸⁸

⁴⁸⁴ Bspw. DuckDuckGo (URL: <https://duckduckgo.com/>), unbubble (URL: <https://unbubble.eu/>), WolframAlpha (URL: <http://www.wolframalpha.com/>), qwant (URL: <https://prev.qwant.com/>) oder MetaGer (URL: <https://metager.de/>).

⁴⁸⁵ Bspw. ZOO (URL: <http://www.zoo.com/>), der Nachfolger von MetaCrawler.

⁴⁸⁶ Aphorismen von verschiedenen Personen verwendet. Zweiter Aphorismus erstmals gehört von Sacha Zala an der infoclio-Tagung vom 02.11.2012 in Bern zum Thema 'Zugangs- und Nutzungsrechte für historische Quellen im Web' (persönliche Notiz).

⁴⁸⁷ Hier und folgend nach: Vaidhyanathan: *The Googlization of everything*, 2011, S. 7.

Faktor	Bedeutung
Domain Age	Je älter eine Webseite, desto besser das Ranking
Keyword Appears in Top Level Domain	Besseres Ranking, wenn das Suchwort im Domainnamen enthalten ist
Keyword As First Word in Domain	Besseres Ranking, wenn das Suchwort als Erstes im Domainnamen enthalten ist
Domain registration length	Je länger die Domain im Voraus reserviert / bezahlt ist, desto besser das Ranking
Keyword in Subdomain Name	Besseres Ranking, wenn das Suchwort im Subdomainnamen enthalten ist
Domain History	Je beständiger die Besitzverhältnisse (Whols), desto besser das Ranking
Exact Match Domain	Stimmt das Suchwort exakt mit einer Domain überein, ist das Ranking höher. Allerdings nur bei qualitativ hoch eingestuften Domains
Public vs. Private Whols	Werden die Besitzinformationen einer Domain nicht offengelegt, sinkt das Ranking (etwas zu verstecken?)
Penalized Whols Owner	Gehört eine Domain einem bekannten Spammer, sinkt das Ranking
Country Top Level Domain extension	Länderspezifische Top Level Domains werden dort höher gewichtet
Keyword in Title Tag	Höhere Gewichtung, wenn Suchwort im Titel der Webseite vorkommt
Title Tag Starts with Keyword	Höhere Gewichtung, wenn das Suchwort am Anfang des Titels der Webseite steht statt am Ende
Keyword in Description Tag	Höhere Gewichtung, wenn das Suchwort in der Seitenbeschreibung enthalten ist
Keyword Appears in H1 Tag	Höhere Gewichtung, wenn das Suchwort im Untertitel enthalten ist
Keyword is Most Frequently Used Phrase in Document	Höhere Gewichtung, wenn das Suchwort öfters auf der Webseite enthalten ist
Content Length	Je länger der Seiteninhalt, desto eher eine positive Gewichtung
Keyword Density	Je grösser das Verhältnis des Suchwortes zu anderen Wörtern, desto besser das Ranking – allerdings nur bis zu einer (un-)bestimmten Dichte
Latent Semantic Indexing Keywords in Content (LSI)	Implementierte LSI-Wörter helfen Suchmaschinen, Wörter mit mehreren Bedeutungen zu unterscheiden. Fehlen diese, ist das Ranking schlechter
LSI Keywords in Title and Description Tags	Sind LSI-Wörter im Titel und der Beschreibung vorhanden, steigt das Ranking
Page Loading Speed via HTML	Je schneller eine Webseite geladen wird, desto besser das Ranking

Tabelle 2: 20 Rankingfaktoren von Google Search nach Brian Dean (Darstellung: Föhr)

Die Aussagekraft dieser Faktoren ist sehr beschränkt, denn die Faktoren sind nicht offiziell von Google bestätigt und es fehlen konkrete Gewichtungsinformationen, insbesondere was 'besser' und 'schlechter' bedeuten.

Bei vielen Faktoren ist zwar nachvollziehbar, warum sie in den Katalog aufgenommen wurden. Aber es ist nicht klar, ob ein Faktor in einem konkreten Fall angewendet und falls, wie dieser gewichtet wird. Was beispielsweise unter einer qualitativ hoch eingestuften Domain zu verstehen ist, kann nicht eruiert werden.

⁴⁸⁸ Eine Liste mit über 200 Rankingfaktoren von Google Search. Nach: Dean, Brian: Google Ranking Factors: The Complete List, Backlinko, 21.03.2016, URL: <http://backlinko.com/google-ranking-factors>, Stand: 28.03.2016.

Diese Probleme sind bisher nicht bekannt gewesen, denn im Analogen fehlt eine vergleichbare Technik. Da Suchmaschinen wie Google Search Ergebnislisten nach nicht nachvollziehbaren Kriterien aus ihren Indizes erstellen, und selbst die quellenkritische Überprüfung wahrscheinlicher Faktoren hypothetisch ist, sind Suchmaschinen als unzuverlässige Recherchemittel zu betrachten.

Mittlerweile sind weitere Suchdienste für verschiedene und spezielle Inhalte entstanden, die eine Suche oder den Suchraum selbst personalisieren und damit algorithmisch vorfiltrieren. Beispielsweise kann durch die News-App *Newsron*⁴⁸⁹ festgelegt werden, welcher Prozentsatz an Artikeln aus welcher Rubrik welcher Zeitung in welchem Land gelistet werden soll. Der Algorithmus der App blendet vermeintlich Uninteressantes aus und präsentiert dem Nutzer nur das, was für diesen interessant sein könnte. Damit entstehen zwei Probleme: erstens ist dem Nutzer die Datenbasis nicht bekannt, also welche Artikel zur Verfügung stehen. Zweitens ist dem Nutzer auch das Auswahlverfahren nicht bekannt. Vielleicht hätte der Nutzer, wenn er die Datenbasis gekannt hätte, ganz andere Artikel ausgewählt. Von diesem Automatismus unterscheidet sich die Auswahl eines Objektes durch eine Redaktion. Zwar ist diese Vorselektion oft ebenfalls kaum nachvollziehbar, aber sie ist mit einer intellektuellen Arbeit verbunden. Für die Digitalhistorikerin Mareike König, eine der Verantwortlichen sowie Redakteurin der wissenschaftlichen Blogplattform *de.hypotheses*⁴⁹⁰, dient die Redaktion der Qualitätssicherung in der Informationsflut.⁴⁹¹ Damit fungiert die Redaktion aber ähnlich einer Suchmaschine als Informationsregulator oder Gatekeeper, indem sie Informationen nach subjektiven Kriterien wie Interesse oder Qualität filtert und Relevanz in ihrem Sinne definiert. Für einen Recherchierenden sollte aber *jedes verfügbare Objekt* auf seinen Nutzen hin kritisch überprüft und ausgewählt oder verworfen werden.

Bereits im Analogen wurden Informationen reguliert. Beispielsweise werden bei Quelleneditionen nur bestimmte Beiträge aus einem überblickbaren und meist abgeschlossenen Bestand aufgenommen, bei wissenschaftlichen Projekten nach klaren und publizierten Richtlinien. Die in Zeitungen veröffentlichten Leserbriefe werden hingegen von der Redaktion ausgewählt, mit dem Anspruch, die Meinung der eingegangenen Beiträge abzubilden. Eine Überprüfung ist hier kaum möglich. *Newsron* unterscheidet sich davon durch die automatische Selektion aus einem unbekanntem und nicht überblickbarem Bestand, der erst mit Hilfe

⁴⁸⁹ URL: <http://www.newsron.com/>.

⁴⁹⁰ URL: <http://de.hypotheses.org/>.

⁴⁹¹ Nach: König, Mareike: Blogs als Wissensorte der Forschung, Präsentation, Die Zukunft der Wissensspeicher: Forschen, Sammeln und Vermitteln im 21. Jahrhundert, Konstanz 06.03.2015, Fol. 17. URL: <http://cms.uni-konstanz.de/wissenschaftsforum/veranstaltungen/veranstaltungsarchiv/veranstaltungen-2015/die-zukunft-der-wissensspeicher/>, Stand: 17.06.2015.

der Informationstechnik erschlossen werden kann. Der Algorithmus entscheidet darüber, was dem Nutzer zur Verfügung gestellt wird.

Suchtechniken für digitale Suchräume mit ausschliesslich digitalen Objekten sind zwar von "Praktiken der 'analogen Zeit' geprägt"⁴⁹², unterscheiden sich von diesen aber vor allem durch die Nutzung eines elektronischen Informationssystems sowie der Präsentation der Suchresultate.

5.1.4. Vollständigkeit

Die Suche nach und das Auffinden von digitalen Objekten wird durch die eingesetzten Algorithmen nicht nur unterstützt, sondern auch aktiv verhindert. Algorithmen, welche Objekte oder Inhalte nach bestimmten Kriterien automatisiert filtern, können für Nutzer sehr praktisch sein. Er hat aber darauf zu vertrauen, dass der Algorithmus einwandfrei funktioniert. Dieser darf nur irrelevante Objekte oder Informationen entfernen oder unterdrücken und muss die relevanten vollständig anzeigen. Anhand der zwei Beispiele *Spam*⁴⁹³ und *Werbeblocker*⁴⁹⁴ wird das Problem verdeutlicht.

Die meisten E-Mail-Verwaltungssoftwares enthalten ein Modul, das unerwünschte E-Mails automatisch herausfiltert und in einem separaten Ordner ablegt oder direkt löscht. Wird der Spam-Filter aktiviert, erlernt die Software anhand der vom Nutzer eingegebenen Kriterien, was dieser als Spam bewertet und herausgefiltert werden muss. In dieser Phase filtert die Software meist noch offensichtlich fehlerhaft und es gelangen für den Nutzer interessante E-Mails in den Spam-Ordner und uninteressante in den Posteingangsortner. Der Nutzer überprüft aber den Spam-Ordner noch regelmässig. Diese regelmässigen Überprüfungen nehmen ab, je länger der Filter eingesetzt wird und je weniger Fehler er verursacht, bis der Nutzer nur noch in begründeten Fällen wie nach einer telefonischen Nachfrage eines Kontaktes im Spam-Ordner nachsieht. Das Problem verschärft sich zusätzlich, wenn der Nutzer eine automatische Löschung der im Spam-Ordner enthaltenen Objekte einrichtet, oder wenn in einer Firmenumgebung ein Spam-Filter eingesetzt wird, auf den der einzelne Nutzer keinen Zugriff mehr hat.

⁴⁹² Hodel: Recherche, 2010, S. 26.

⁴⁹³ Spam (Spam-Mail oder Junk-Mail): Unerwünschte E-Mails mit unterschiedlichen Inhalten (Werbung, Forenbeiträge, persönliche Konversation), die in hoher Zahl versendet werden und damit die elektronische Kommunikation überlasten (sowohl Provider, als auch Nutzer). Begriff nach dem gleichnamigen Büchsenfleisch in einem Sketch von Monty Python. Siehe: Prevezanos: Computer Lexikon 2013, 2012; Fischer; Hofer: Lexikon der Informatik, 2011.

⁴⁹⁴ Werbeblocker: Programme oder Programmodule, die die Anzeige unerwünschter Inhalte einer Webseite unterdrücken. Nach: Prevezanos: Computer Lexikon 2013, 2012.

Werbeblocker wie Adblock Plus⁴⁹⁵ sind für Nutzer ebenfalls sehr praktisch, weil nicht nur unerwünschte Inhalte auf einer Webseite ausgeblendet werden, sondern weil damit auch die Seite schneller geladen wird. Der Nutzer kann aber nur selten nachvollziehen, welche Inhalte der Werbeblocker tatsächlich unterdrückt, wenn dessen Source Code⁴⁹⁶ sowie White- und Blacklists nicht veröffentlicht sind. Der Nutzer muss darauf vertrauen, dass der Algorithmus nur Inhalte ausblendet, die für ihn tatsächlich nicht relevant sind. Selbst wenn der Werbeblocker einwandfrei funktioniert, können eventuell relevante Informationen verborgen bleiben. Dies kann dann eintreffen, wenn von einem Website-Betreiber sogenannte Werbeblocker-Blocker⁴⁹⁷ eingesetzt werden, die den Werbeblocker des Nutzers erkennen und dann den Zugriff auf Informationen verhindern.

Im Webauftritt von GEO Deutschland ist ein Werbeblocker-Blocker implementiert. Wenn ein Adblocker erkannt wird, werden keine Inhalte angezeigt. Dies stellt für einen Nutzer den günstigsten Fall dar, denn es können durch den Webseitenbetreiber auch nur selektiv Inhalte ausgeblendet werden, was für den Nutzer nicht erkennbar ist. In diesem Fall sind die vorhandenen Objekte oder Informationen unvollständig:

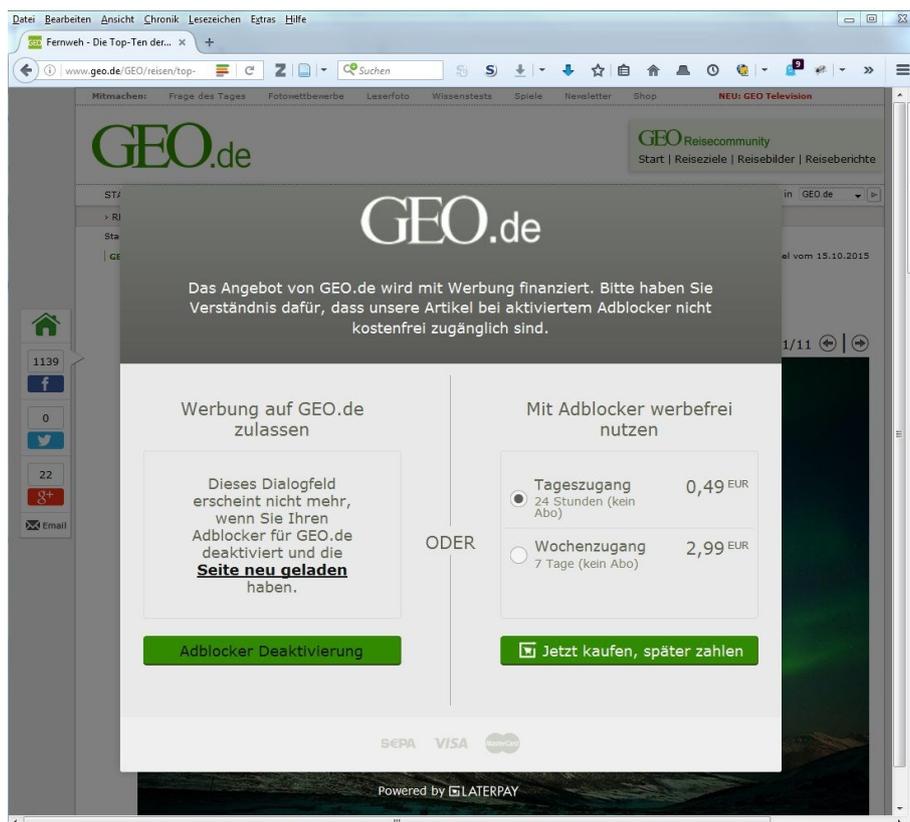


Abbildung 8: Webseite von GEO, mit im Browser aktiviertem Werbeblocker (URL: <http://www.geo.de>)⁴⁹⁸

⁴⁹⁵ URL: <https://adblockplus.org/>.

⁴⁹⁶ Bspw. uBlock, URL: <https://github.com/chrisaljoudi/uBlock/>.

⁴⁹⁷ Bspw. BlockAdblock, URL: <http://blockadblock.com/>.

⁴⁹⁸ Bildzitat GEO.de (URL: <http://www.geo.de>); URL: <http://www.geo.de/GEO/reisen/top-ten/fernweh-die-top-ten-der-beliebtesten-naturphaenomene-81702.html>, Stand: 09.04.2016.

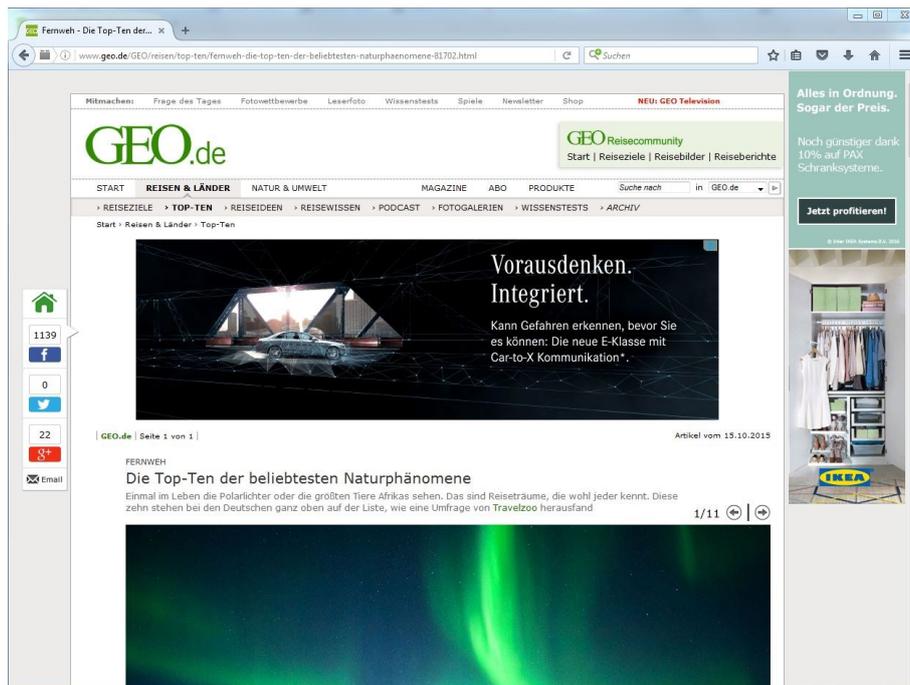


Abbildung 9: Webseite von GEO, mit im Browser deaktiviertem Werbeblocker (URL: <http://www.geo.de>)⁴⁹⁹

Der Nutzer hat sich bewusst zu sein, dass der Einsatz von Computerprogrammen zur Filterung oder Selektierung von Objekten und Inhalten zu Informationsverlust führen kann. Ein Spam-Filter kann mit einem Assistenten verglichen werden, der Briefe sortiert, allfällige Werbung entfernt und anschliessend an den oder die Empfänger weiterverteilt. Mit einem E-Mailprogramm hat nun fast jeder Nutzer einen Assistenten, der die Post filtert, selektiert und neu ordnet. Ein Werbeblocker ist hingegen in der analogen Welt nicht umsetzbar, denn dies wäre ein Assistent, der beispielsweise beim Spaziergang durch eine Stadt jede Werbeanzeige entfernt, ohne dass es dem Nutzer auffallen würde.

5.1.5. De- / Kontextualisierung

Die während einer Suche aufgefundenen potenziellen Quellen stehen meistens mit anderen Objekten in Beziehung. Zum Beispiel wird der Kontext eines physischen Objekts in einem Archiv durch die Abteilung, der Bestand und/oder das Dossier und die darin enthaltenen weiteren Objekte hergestellt. Dieser Kontext bietet für eine potenzielle Quelle eine Orientierungshilfe und Möglichkeiten, systematisch nach weiteren potenziellen Quellen zu suchen, sowie wichtige Hintergrundinformationen zur Beurteilung der Authentizität des Objekts. Bei einer Suche mittels Suchmaschine geht in den meisten Fällen der Kontext eines einzelnen Suchresultats verloren.

⁴⁹⁹ Bildzitat GEO.de (URL: <http://www.geo.de>); URL: <http://www.geo.de/GEO/reisen/top-ten/fernweh-die-top-ten-der-beliebtesten-naturphaenomene-81702.html>, Stand: 09.04.2016. Die beiden Screenshots zeigen zudem die Problematik auf, dass der Nutzer mit der Verwendung eines Werbeblockers auf Werbung verzichten will, Webseitenbetreiber aber auf Werbeeinnahmen angewiesen sind, um die Inhalte und Dienste zur Verfügung stellen zu können.

Das algorithmisch generierte Resultat einer unsystematischen Suche wird entweder als Ergebnisliste präsentiert, oder es werden bestimmte Elemente eines digitalen Objekts hervorgehoben, wie beispielsweise bei einer Volltextsuche die gesuchte Zeichenfolge. In beiden Fällen werden die aufgefundenen Elemente meistens aus dem Zusammenhang gerissen und damit dekontextualisiert, oder es wird zumindest deren Kontext durch eine sogenannte Kontextmodifikation aktiv geändert, indem sie als separates Element hervorgehoben werden und nicht mehr mit anderen Objekten oder Elementen in Beziehung stehen.⁵⁰⁰ Die Ergebnisliste enthält keine systematische Reihenfolge, die als Orientierungshilfe dienen und für eine systematische Suche verwendet werden könnte.⁵⁰¹ Sie bildet auch keinen neuen Kontext, denn die Einträge stehen untereinander nicht in einer Beziehung, sondern haben nur den gesuchten Ausdruck gemeinsam. Durch diese Dekontextualisierung können auch Paratexte wie ein Vorwort oder das Impressum verloren gehen, die für den Nutzer als Zusatzinformation hilfreich sind. Zudem ist für den Nutzer unter Umständen auch nicht mehr ersichtlich, in welchem Bestand welcher Institution sich das digitale Objekt tatsächlich befindet, denn mit Hyperlinks kann ein Objekt in verschiedene Bestände integriert werden. Jedes Element der Liste ist deshalb an seinem Speicher- oder Fundort zu rekontextualisieren, falls dies überhaupt möglich ist.⁵⁰² Das (Re-)Kontextualisieren kann durch einzelne Elemente des Objekts verhindert werden, beispielsweise wenn der Autor ein Pseudonym verwendet. Informationen über diesen stehen dann in den meisten Fällen nicht zur Verfügung, was auch der Sinn eines Pseudonyms und der Anonymität ist.⁵⁰³

Mit digitalen Elementen entstanden auch neue Paratexte, die zusätzliche Informationen zu den Objekten liefern können. Für Webseiten beispielsweise hat Niels Brügger vier zusätzliche paratextuelle Elemente identifiziert.⁵⁰⁴ Das erste Elementen nennt er *double-paratext* und meint damit einen Paratext, der verschiedene Inhalte gleichzeitig begleitet. Beispielsweise kann ein Titel sowohl über einem Artikelanrisstext als auch über dem ganzen Artikel stehen. Ein zweites Element ist das *drop-down menu*, mit dem aus Platzgründen hypertextuelle Elemente graphisch verborgen werden, wie beispielsweise die Navigation eines Webauftritts, die Hinweise auf zu erwartende Inhalte gibt. Mit *not-immediately-visible, non-*

⁵⁰⁰ Vgl.: Reichert, Ramón: Im Kino der Humanwissenschaften: Studien zur Medialisierung wissenschaftlichen Wissens, Bielefeld 2007 (Kultur- und Medientheorie), S. 35. Auch: Kemman; Kleppe; Scagliola: Just Google It, 2014, S. 4.

⁵⁰¹ Siehe: Lingelbach, Gabriele: Ein Motor der Geschichtswissenschaft? Zusammenhänge zwischen technologischer Entwicklung, Veränderungen des Arbeitsalltags von Historikern und fachlichem Wandel, zeitenblicke, 09.08.2011, Abs. 21, URL: http://www.zeitenblicke.de/2011/1/Lingelbach/index_html, Stand: 30.07.2015.

⁵⁰² Bspw. bietet der Webauftritt des Staatsarchivs Zürich eine Archivplansuche, in der Fundstücke verortet werden können. Siehe URL: <http://suche.staatsarchiv.djiktzh.ch/archivplansuche.aspx>.

⁵⁰³ Siehe zu Pseudonymen und Anonymität Kap. 5.4.8. Autorschaft.

⁵⁰⁴ Weiterführend: Brügger: Website Analysis, 2010, S. 40-43. Paratexte für Webseiten sind: Überschriftensystem, referenzierende Textelemente (Siehe...), organisierende Listen (Menüsystem), zugehörige Elemente (Textbox), Kopf- und Füsselemente, Titel, URL etc.

indicated post festum paratexts bezeichnet Brügger die dritte Art von Paratexten, die nicht unmittelbar sichtbar sind, deren Existenz nicht erkennbar ist oder die erst durch eine Aktion wie beispielsweise durch Mouse-Over sichtbar werden. Die Informationen sind im Code enthalten und nur nach bestimmten Aktionen oder Prozessen sichtbar. Dieses Element wurde schon als neuer *Quellentypus der unsichtbar codierten Information* identifiziert.⁵⁰⁵ Das vierte paratextuelle Element sind *search terms*. Darunter werden Suchbegriffe verstanden, die in ein auf der Webseite implementiertes, ebenfalls paratextuelles Suchfeld eingegeben werden. Durch das Listen der aufgefundenen, effektiven Inhalte werden Suchbegriffe zu einem Paratext, weil damit neue Beziehungen geschaffen werden, die vorher nicht intendiert waren. Dieser Vorgang ist durch hohe Zufälligkeit, Flexibilität und Dynamik charakterisiert.

Neue Formen von Paratexten sind auch für andere digitale Objekte vorstellbar, dazu liegen aber bisher aufgrund der Vielfalt der Objekte noch keine Untersuchungen vor. Jedes Objekt ist einzeln auf seine individuellen Paratexte, mit denen es kontextualisiert werden kann, zu untersuchen.

Werden die Objekte losgelöst vom Kontext verwendet, besteht einerseits die Gefahr der inhaltlichen Fehlinterpretation, andererseits können dadurch für eine kritische Beurteilung notwendige Informationen und Metadaten wie Autor oder Zeit verloren gehen. Damit mindert sich der Aussagewert eines Objekts, mit entsprechenden Auswirkungen auf die direkte Verwendbarkeit für die Forschungsfrage. Folglich sind potenzielle Quellen für eine adäquate Beurteilung zuerst zu rekontextualisieren. Das bedeutet ein *Herauszoomen* aus der Mikro-Ebene in eine den grösseren Zusammenhang darstellende Makro-Ebene.⁵⁰⁶ Erst auf dieser Ebene sind Beziehungen und Systematiken erkennbar.

5.1.6. Recht

*"Die Geltendmachung von Urheberrechten an wissenschaftlichen Daten, Informationen und weiteren Inhalten entbehrt [...] oft jeglicher rechtlichen Grundlage."*⁵⁰⁷ (Egloff, 2014)

Die geltende nationale Gesetzgebung und internationale Abkommen zum Urheberrecht und verwandten Rechten erschweren die Nutzung von vorhandenem Wissen. Diese rechtlichen Rahmenbedingungen schränken Suchräume ein und können die Verwendung von digitalen Objekten verhindern. Die folgenden Ausführungen sollen einerseits den Leser für die Prob-

⁵⁰⁵ Siehe Kap. 2.2.3. Quellentypus der unsichtbar codierten Information.

⁵⁰⁶ Dazu: Zaagsma: On Digital History, 2013, S. 26.

⁵⁰⁷ Egloff, Wili: Der Zugang zu Fachinformationen – zwischen gesellschaftlicher Notwendigkeit und privatem Verwertungsinteresse, in: Gartska, Hansjürgen; Coy, Wolfgang (Hg.): Wovon – für wen – wozu. Systemdenken wider die Diktatur der Daten. Wilhelm Steinmüller zum Gedächtnis, Berlin 2014, S. 355. URL: <http://edoc.hu-berlin.de/docviews/abstract.php?lang=ger&id=40657>, Stand: 19.03.2015.

ematik sensibilisieren und andererseits bilden sie eine der argumentatorischen Voraussetzungen für ein 'Trusted Digital Repository (TDR)'.⁵⁰⁸

Urheberrecht⁵⁰⁹ (engl. Copyright⁵¹⁰) und Leistungsschutzrecht⁵¹¹ sind gesetzliche Grundlagen für die Durchsetzung der Ansprüche dieser Rechteinhaber, die massgeblichen Einfluss auf die Verbreitung und Verwendung von Wissen haben und als bedeutender einschränkender Faktor im gesamten historisch-kritischen Prozess zu betrachten sind. Urheberrecht / Copyright gehören zum Sammelbegriff *geistiges Eigentum*, unter den auch verwandte Schutzrechte, Patente, Gebrauchsmusterrechte, Geschmacksmusterrechte, Marken und geschäftliche Bezeichnungen als subjektive Rechte fallen.⁵¹² Sie sind national unterschiedlich geregelt, teilweise von supranationalen Organisationen wie der EU und internationalen Abkommen wie im WIPO-Abkommen (World Intellectual Property Organization) beeinflusst. Zwar wird vermehrt und mit wachsendem Druck die freie Verwendung von Wissen gefordert⁵¹³, was unter den Bezeichnungen Access to Knowledge (A2K), Open Source⁵¹⁴, Creative Commons⁵¹⁵, Public Domain⁵¹⁶, Sharing⁵¹⁷, Copyleft⁵¹⁸ oder Open Access (OA)⁵¹⁹ bekannt ist. Die Forderung stösst aber auf teils massiven

⁵⁰⁸ Siehe dazu Kap. 5.3.6 Sicherungsort, Lösungsansatz ‚Personal Trusted Archive‘.

⁵⁰⁹ Urheberrecht: Schutz von Werken, d.h. geistige Schöpfungen der Literatur und Kunst, die individuellen Charakter haben. Für Schweiz siehe bspw.: Eidgenössisches Institut für Geistiges Eigentum: Grundlegendes zum Urheberrecht URL: <https://www.ige.ch/urheberrecht/urheberrecht.html>.

⁵¹⁰ Zur Entwicklung des Copyrights im internationalen Kontext siehe: Dommann: Autoren und Apparate, 2014. Zur Zukunft siehe: Saunders, Joss: The future of copyright: what are the pressures on the present system?, in: Cope, Bill; Phillips, Angus (Hg.): The future of the academic journal, Oxford 2014 (Chandos information professional series), S. 249-257. Saunders meint, dass bisherige Protagonisten wie Publikationsgesellschaften den Einfluss auf das Copyright immer stärker an die politische Legislative und Exekutive sowie die Gerichte abgeben werden.

⁵¹¹ Leistungsschutzrecht: Immaterialgüterrechtsform, zum Schutz von Leistungen von Publikationsgesellschaften gegenüber Suchmaschinenbetreibern.

⁵¹² Nach: WIPO: WIPO-Administered Treaties: Convention Establishing the World Intellectual Property Organization (28.09.1979), WIPO, URL: http://www.wipo.int/wipolex/en/treaties/text.jsp?file_id=283854#P50_1504, Stand: 31.08.2015, Art. 2 viii.

⁵¹³ Weiterführend bspw.: Krikorian, Gaëlle; Kapczynski, Amy (Hg.): Access To Knowledge In The Age Of Intellectual Property, New York 2010. Ein (teilweise berechtigter) Rundumschlag gegen die wirtschaftliche Nutzbarmachung von Wissen; vernachlässigt werden von Krikorian allerdings u.a. die Kosten der Wissensgenerierung, dafür schreibt sie mit viel Idealismus für eine offene Wissenskultur.

⁵¹⁴ Open Source: Quelldaten werden veröffentlicht und können genutzt werden ausser für die profitorientierte Verwendung. Es bestehen unter Open Source viele Lizenzmodelle (General Public Licence GPL). Fischer; Hofer: Lexikon der Informatik, 2011.

⁵¹⁵ Creative Commons CC: Lizenzmodelle (Public Domain, General Public Licence etc.) zur freien Verwendung und Verbreitung von Inhalten mit feiner Abstufung der gewährten Rechte. Der Urheber muss bei einer CC-Lizenz stets genannt werden. Prevezanos: Computer Lexikon 2013, 2012. Siehe auch URL: <http://creativecommons.org/>.

⁵¹⁶ Public Domain: Urheber verzichtet vollständig auf sein geistiges Eigentum. Fischer; Hofer: Lexikon der Informatik, 2011.

⁵¹⁷ Sharing: gemeinsame Nutzung.

⁵¹⁸ Copyleft: erzwingt die Freiheit zum Verändern und Verteilen des Urwerks und steht dem Copyright entgegen. Prevezanos: Computer Lexikon 2013, 2012.

⁵¹⁹ Open Access: freier Zugang (zu Wissen). In der Wissenschaft sollen Forschungsergebnisse physisch und digital frei zugänglich sein.

Widerstand der Rechteinhaber, die über die Verwendung ihres geistigen Eigentums⁵²⁰ bestimmen wollen und dafür eine Gegenleistung verlangen.

Diskussionen um rechtliche Einschränkungen wurden bereits mit der Einführung und Verwendung von *Reproduktionsmethoden* wie Photostat, Offset und Mikrofilm in der ersten Hälfte des 20. Jahrhunderts geführt.⁵²¹ Für Verleger waren *originalgetreue* Kopien eine Gefährdung ihrer Absatzinteressen für die *Originale*, einer Ansicht, der sich die Wissenschaft aktiv und auch entgegen gerichtlicher Entscheide widersetzte.⁵²² Einer der grössten Verfechter einer Vision von frei verfügbaren Ideen und von freiem Austausch innerhalb der Weltgesellschaft war der Historiker Robert Binkley.⁵²³ Er begriff das exklusive Copyright der Verleger als ein den Interessen der Wissenschaftler auf möglichst breite Diskussion ihrer Forschungsergebnisse entgegengerichteter Standpunkt und forderte vergeblich eine "neue Ethik der Kopie."⁵²⁴ Die Vereinbarungen zum Schutz des literarischen und künstlerischen Eigentums wie die Berner Konvention von 1886 und deren Revision von 1908 in Berlin, sowie die Entwicklungen zu einer immer restriktiveren Auslegung des Urheberrechts wie die Trans-Pacific Partnership (TPP) und die Transatlantic Trade and Investment Partnership (TTIP)⁵²⁵, berücksichtigen hauptsächlich die Interessen der Urheber und Verwertungsgesellschaften und entsprechen nicht einer offenen kulturellen Entwicklung der Gesellschaft.⁵²⁶ Insbesondere Historiker sollten sich für die Behandlung von historischem Wissen als öffentlichem Gut einsetzen, denn es gibt keine durchsetzbaren geistigen Eigentumsrechte und niemand muss mit dem Ausschluss aus der Forschungsgemeinschaft rechnen, solange das von Reinhard Koselleck formulierte Diktum des Vetorechts der Quellen beachtet wird.⁵²⁷

Seit den 1990er-Jahren wird nun darüber diskutiert und gestritten, wie das Urheberrecht von einer physischen auf die virtuelle Welt ausgedehnt werden kann. Das sei allerdings ein Ansinnen, das zu einem *Zusammenstoss des Urheberrechts mit dem Internet* führe, so der

⁵²⁰ Geistiges Eigentum: Rechte an Werken der Literatur, Kunst und Wissenschaft sowie Schöpfungen technischer und ästhetischer Natur sowie Kennzeichen / Marken. Siehe auch: 'Intellectual Property' (IP) nach WIPO (URL: <http://www.wipo.int/about-ip/en/>). Weniger verwendet wird 'Immaterialgut'; Begriff nach: Kohler, Josef: Das Autorenrecht – Eine zivilistische Abhandlung zugleich ein Beitrag zur Lehre vom Eigentum, vom Miteigentum, vom Rechtsgeschäft und vom Individualrecht, Jena 1880.

⁵²¹ Siehe dazu eingehend: Dommann: Autoren und Apparate, 2014, insb. S. 135-169.

⁵²² Vgl.: Ebd., S. 160 f.

⁵²³ Vgl.: Ebd., S. 162 f.

⁵²⁴ Ebd., S. 165.

⁵²⁵ In den Geheimverhandlungen zu den Handelsabkommen TPP (USA und div. Pazifikstaaten) hat anscheinend die USA die Schutzfrist auf 70 Jahre nach dem Tod des Urhebers durchgesetzt. Siehe Vertragstext: New Zealand Foreign Affairs & Trade: Text of the Trans-Pacific Partnership, Trans-Pacific Partnership, 2016, Kap. 18 (Intellectual Property), URL: <https://www.tpp.mfat.govt.nz/text>, Stand: 03.04.2016. In den geheimen Verhandlungen um TTIP (USA und EU) wird analog der TPP ebenfalls ein strenger Schutz geistigen Eigentums angestrebt.

⁵²⁶ Für alternative Urheberrechtsmodelle setzten sich nach 1945 u.a. Arno Winter und die Library of Congress (USA) ein. Siehe: Dommann: Autoren und Apparate, 2014, S. 239 f.

⁵²⁷ Vgl.: Dommann, Monika; Gugerli, David: Geschichtswissenschaft in Begutachtung: acht Kommentare zur historischen Methode der Gegenwart, in: Traverse, 2011, S. 162.

Softwareentwickler Hartwig Thomas.⁵²⁸ Monika Dommann hat die aktuelle Debatte mit den rechtlichen Konflikten bei der Regulierung von Bibliothekskopien und Musikaufnahmen in Beziehung gesetzt und viele Parallelen gefunden.⁵²⁹ Doch handelte es sich bei den von ihr untersuchten Medien nicht um digitale Objekte wie in den aktuellen urheberrechtlichen Auseinandersetzungen, sondern um materielle Objekte. Die Kritik, dass sich eine physische Publikationen regelnde Gesetzgebung nicht einfach auf virtuelle Medien übertragen lasse, weil sich Letztere in der Nutzung stark von Ersteren unterscheiden, wurde bisher von den Gesetzgebern grossmehrheitlich ignoriert.⁵³⁰ Die Auswirkungen auf die Forschung sind nicht unerheblich, weil aufgrund dieser rechtlichen Unsicherheit Daten und Informationen aus Eigenschutz vor juristischen Auseinandersetzungen eher zurückgehalten oder nur schwer zugänglich publiziert werden. Forschende üben damit eine Selbstzensur aus.

Welche Probleme mit der Anwendung des Urheberrechts für Forschende entstehen können, wird anhand eines aktuellen Beispiels dargelegt:

Eine Bibliothek hat eine Zeitschrift abonniert, auf deren Artikel, sofern die betreffende Ausgabe nicht ausgeliehen ist, unmittelbar zugegriffen werden kann. Die digitale Version des Artikels, heute fast immer born-digital, kann zwar auf der Webseite des Verlags verfügbar sein, aber der Zugriff darauf ist eingeschränkt. Aus marktwirtschaftlichen oder urheberrechtlichen Gründen wird nur jenen der Zugriff ermöglicht, die über einen gültigen Zugang zu den durch einen Anmeldeprozess geschützten Inhalten verfügen (Abonnement). Die Publikationsgesellschaft bestimmt damit, wann, für wen und in welcher Form der Artikel digital zur Verfügung steht. Der Artikel wird beispielsweise nie digital öffentlich zugänglich sein oder erst nach einer bestimmten Zeit. Oder er ist zu Beginn öffentlich verfügbar, wird aber nach einer bestimmten Zeit hinter eine *Moving Wall* oder *Paywall* ins kostenpflichtige Archiv verschoben. Neben verschie-

⁵²⁸ Nach: Thomas, Hartwig: Erster Entwurf eines Versuchs über den Zusammenstoss des Urheberrechts mit dem Internet, Zürich 2014. URL: <https://buchundnetz.com/online-buch/erster-entwurf-eines-versuchs-ueber-den-zusammenstoss-des-urheberrechts-mit-dem-internet/inhaltsverzeichnis/>. Thomas' Buch wird nach und nach online publiziert: <https://buchundnetz.com/online-buch/erster-entwurf-eines-versuchs-ueber-den-zusammenstoss-des-urheberrechts-mit-dem-internet/inhaltsverzeichnis/>. Siehe weitere urheberrechtskritische Literatur, bspw.: Falkvinge, Rick: The Entire Copyright Monopoly Idea is Based on a Colossal Lie, Opinion, TorrentFreak, 21.06.2015, URL: <https://torrentfreak.com/the-entire-copyright-monopoly-idea-is-based-on-a-colossal-lie-150621/>, Stand: 22.06.2015.

⁵²⁹ Siehe: Dommann: Autoren und Apparate, 2014.

⁵³⁰ Siehe dazu bspw. die kontroversen Debatten nach der Veröffentlichung des Schlussberichts der AGUR12 (vom Bundesrat eingesetzte Arbeitsgruppe zum Urheberrecht): Arbeitsgruppe zum Urheberrecht: Schlussbericht AGUR12, Eidgenössisches Institut für Geistiges Eigentum, Bern 28.11.2013. URL: https://www.ige.ch/fileadmin/user_upload/recht/national/d/urheberrecht/Schlussbericht_der_AGUR12_vom_28_11_2013_DE.pdf, Stand: 06.05.2018; Nicolussi, Ronny: Arbeitsgruppe schlägt Zensur und Überwachung vor: Urheberrecht im digitalen Zeitalter, Schweiz, Neue Zürcher Zeitung, 06.12.2013, URL: <http://www.nzz.ch/aktuell/schweiz/arbeitsgruppe-schlaegt-zensur-und-ueberwachung-vor-1.18199613>, Stand: 04.05.2018; Universität Zürich: Chaosseminar: AGUR12 und Urheberrechtsentwicklungen, Agenda, Universität Zürich, 11.12.2014, URL: <http://www.agenda.uzh.ch/record.php?id=25153>, Stand: 04.04.2015.

denen Formaten wie PDF oder EPUB kann der Artikel auch mit gekürztem Inhalt, dem Abstract, online zur Verfügung stehen, während der vollständige Artikel kostenpflichtig ist. Zusätzlich kann ein digitales Objekt mit einem DRM-Verfahren geschützt werden, das die Darstellung des Inhalts nur auf einem bestimmten Gerät erlaubt und Kopien verhindert. Es ist sogar möglich, dass auf Endgeräten digitale Objekte gelöscht werden, ohne dass der Nutzer eingreifen kann.⁵³¹

Vor allem *technische Schutzmassnahmen* wie Digital Rights Management (DRM)⁵³² können die Nutzung und Archivierung eines Werks einschränken, indem dadurch auch der vom Urheberrecht ausgenommene *Werkgenuss* verhindert werden kann. Solche Massnahmen erlauben es, nicht nur berechnete Schutzansprüche durchzusetzen, sondern auch irrelevante und vom Urheberrecht nicht betroffene digitale Objekte zu schützen respektive nur einem bestimmten Kreis zugänglich zu machen. Zudem bieten einige Publikationsgesellschaften den Bibliotheken ihre Fachpublikationen nur im Paket mit anderen Zeitschriften und mit entsprechend hohem Preis an. Aus Spargründen müssen Bibliotheken eventuell auf diese Publikationen verzichten, wodurch Wissenschaftlern der Zugang zu wichtigen und aktuellen Fachinformationen verwehrt bleibt.⁵³³

Dagegen versuchen sich namhafte Wissenschaftler mit der *Budapest Open Access Initiative*⁵³⁴ von 2002 oder der *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*⁵³⁵ von 2003 zu wehren. Zwar hat die Wissenschaftsgemeinschaft schon Erfolge gegenüber einzelnen Verlagen erzielt, aber bisher haben die Bemühungen für eine grundsätzliche Änderung der Situation in der Politik und Gesetzgebung noch wenig Widerhall gefunden.

Wenig beachtet wird die gesetzlich festgehaltene Richtlinie, dass es einer geistige Schöpfung mit individuellem Charakter bedarf, die sogenannte *Schöpfungshöhe*, um ein Werk schützen lassen zu können.⁵³⁶ Beispielsweise erreicht ein Digitalisat eines gemeinfreien

⁵³¹ Bspw. hat Amazon 2009 diverse von Nutzern gekaufte Bücher auf den durch Amazon vertriebenen und im Besitz der Nutzer befindlichen Kindle-Geräten aus Copyright-Gründen gelöscht. Siehe dazu: Stone, Brad: Amazon Erases Orwell Books From Kindle, in: The New York Times, 18.07.2009. URL: <http://www.nytimes.com/2009/07/18/technology/companies/18amazon.html>, Stand: 19.03.2015. Mittlerweile wenden sich verschiedene Verlage von harten, d.h. einschränkenden DRM-Verfahren ab, siehe bspw.: Schumacher-Gebler, Christian: PIPER ebooks mit digitalem Wasserzeichen, News, Piper Verlag, 23.06.2015, URL: <http://www.piper.de/aktuelles/news/piper-e-books-ohne-kopierschutz>, Stand: 24.07.2015.

⁵³² Siehe dazu bspw.: Gladney: Preserving Digital Information, 2007172 f.

⁵³³ Siehe: Egloff: Wovon – für wen – wozu, 2014, S. 354 f.

⁵³⁴ URL: <http://www.budapestopenaccessinitiative.org/>. Dazu auch: Wilbanks, John: Licence restrictions: A fool's errand, in: Nature 495 (7442), 28.03.2013, S. 440-441. DOI: 10.1038/495440a.

⁵³⁵ Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, Berliner Erklärung, Open Access – Max-Planck-Gesellschaft, 22.10.2003, URL: <http://openaccess.mpg.de/Berliner-Erklärung>, Stand: 18.11.2014.

⁵³⁶ Siehe dazu bspw.: Weber, Rolf H.: E-Commerce und Recht: rechtliche Rahmenbedingungen elektronischer Geschäftsformen, Zürich 2010, S. 254 f.

Werks die geforderte Schöpfungshöhe nicht und kann auch nicht geschützt werden, weil der Digitalisierungsprozess nicht genügend schöpferisch wirkt. Wird das Digitalisat mit zusätzlichen Informationen versehen, können diese, und nur diese born-digitalen Informationen, allenfalls urheberrechtlich geschützt werden. Folglich sind beispielsweise die Copyright-Bestimmungen der Herzog August Bibliothek Wolfenbüttel zum Digitalisat des Sachsenspiegels aus dem 14. Jahrhundert kaum haltbar, weil das ursprüngliche Werk gemeinfrei ist.⁵³⁷ Da die Institution keine Rechte daran besitzt, können keine Rechte daraus erwachsen, die sie weitergeben könnte. Trotzdem wird von der Institution darauf hingewiesen, dass die Bibliothek mit diesem Lizenzierungssystem einen "Akzent für eine offene und freie Wissenschaftskultur"⁵³⁸ setzen will. Die Diskussion um die Schöpfungshöhe hat zu Abmahnungen und Gerichtsfällen geführt und ist aufgrund unklarer gesetzlicher Definition und international verschiedenen Gesetzgebungen noch nicht beendet.⁵³⁹

Das Copyright soll die *Ausgewogenheit* zwischen den Rechten der Urheber eines geistigen Eigentums und den sozialen und kulturellen Ansprüchen der Gesellschaft sicherstellen.⁵⁴⁰ Die Schranken des Urheberrechts werden aber auch mittels Verlängerung von Schutzfristen umgangen. Beispielsweise wird im auch unter *Mickey Mouse Protection Act* bekannten *Copyright Term Extension Act*⁵⁴¹ der USA die Schutzfrist für viele Werke auf 70 Jahre erhöht oder in der geplanten Revision des Urheberrechtsgesetzes von Südafrika⁵⁴² sollen ewiger Schutz auf viele Immaterialgüter garantiert werden.

Mit der Schaffung von digitalen Objekten in all ihren Formaten und Verwendungsmöglichkeiten sind neue Ansprüche auf beiden Seiten entstanden, die verhandelt und durch die

⁵³⁷ Unter CC-Lizenz. Hingegen könnte eine gedruckte Neuauflage des Werks in ihrer Darstellungsform geschützt werden. Nur der ursprüngliche Inhalt und dessen Darstellung bleiben gemeinfrei.

⁵³⁸ Dauer, Antje: Pressemitteilung der Herzog August Bibliothek Wolfenbüttel: Freie Nutzung von Digitalisaten dank neuer Lizenzen, 23.05.2013. URL: <http://www.hab.de/files/2013-05-22-creativecommons-lizenz.pdf>, Stand: 20.03.2015.

⁵³⁹ Bspw. London National Portrait Gallery vs. Wikipedia 2009; Mannheimer Reiss-Engelhorn-Museen vs. Website-Betreiber bzgl. Richard-Wagner-Porträt. Siehe dazu bspw.: Pachali, David: Foto-Reproduktionen: Wikimedia verliert Bilderstreit mit Mannheimer Museum (Update), Museen, iRights – Kreativität und Urheberrecht in der digitalen Welt, 23.06.2016, URL: <https://irights.info/artikel/foto-reproduktionen-wikimedia-verliert-bilderstreit-mit-mannheimer-museum/27582>, Stand: 16.07.2016.

Forderung zur Klarstellung siehe bspw. Reda-Bericht: Reda, Julia: Implementation of Directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society, 2014/2256(INI)-P8_TA-PROV(2015)0273, European Parliament, 09.07.2015, Abschn. 31. URL: <http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2015-0273&language=EN>, Stand: 18.09.2015. Forderung nach klaren Rechtsnormen für Digitalisate von Kulturgut siehe bspw. 'Hamburger Note' vom 13.10.2015, URL: <http://hamburger-note.de/>.

⁵⁴⁰ Siehe: Weber: E-Commerce und Recht: rechtliche Rahmenbedingungen elektronischer Geschäftsformen, 2010, S. 273-277.

⁵⁴¹ Senate and House of Representatives of the United States of America: Copyright Term Extension Act, 27.10.1998. URL: <http://www.copyright.gov/legislation/pl105-298.pdf>, Stand: 05.04.2015.

⁵⁴² Siehe: Padayachy, Meshendri: Copyright Amendment Bill 2015 – Draft, 27.07.2015, S. 8. URL: <http://www.gov.za/documents/copyright-amendment-bill-comments-invited-27-jul-2015-0000>, Stand: 16.08.2015.

Gesetzgebung festgelegt werden müssen.⁵⁴³ Die Weiterentwicklung der Informationstechnik verlangt nach einer steten Anpassung der Gesetze, was zu rechtlichen Unsicherheiten führt. Beispielsweise stellen nicht wenige Lehrbeauftragte ihre Unterrichtsmaterialien aus Furcht vor rechtlichen Konsequenzen wieder nur physisch im Semesterapparat zur Verfügung, obwohl keine gesetzlichen Eigentumsrechte an Informationen bestehen.⁵⁴⁴ Auch die Gedächtnisinstitutionen stehen vor grossen Herausforderungen, denn die Gesetzgebung regelt viele Aspekte des Urheberrechts wie die Ausleihe von E-Books, Pflichtexemplar- oder Zweitveröffentlichungsrecht, Verzeichnisprivileg, verwaiste Werke oder Entschädigung von Autoren für Ausleihe ihrer Werke noch immer nicht abschliessend.⁵⁴⁵ Urheberrechtsreformen werden intensiv diskutiert, sind aber äusserst schwierig um- und durchzusetzen, weil unter anderem die Berner Konvention von den Vertragspartnern nur einvernehmlich geändert werden kann.⁵⁴⁶

Doch auch die erfolgreiche *Durchsetzung* kann bei Rechteinhabern zu unliebsamen Überraschungen führen, wie zwei Beispiele zeigen. In Deutschland hat der Axel Springer Verlag auf der Basis des dort geltenden Leistungsschutzrechts gefordert, dass Google Search jene Titel zu vergüten hat, die in der Resultatliste einer Suche mit Textaufriss und/oder Bild aufgeführt werden. Google nahm darauf diese Titel aus dem Index, was laut dem Axel Springer Verlag zu einem Traffic-Einbruch auf seiner Webseite von fast 40% führte.⁵⁴⁷ Der Verlag richtete daraufhin relativ zügig einen unentgeltlichen Zugriff für Google Search ein. In Spanien wurden durch eine Änderung des Gesetzes über das geistige Eigentum, das auch unter 'Google-Steuer' bekannte *Ley de Propiedad Intelectual* von 2014, selbst geringfügige Textzitate bis

⁵⁴³ Siehe bspw. für die Schweiz: Eidgenössisches Institut für Geistiges Eigentum, URL: <https://www.ige.ch/>.

⁵⁴⁴ Vgl.: Egloff: Wovon – für wen – wozu, 2014, S. 355. Egloff bezieht sich hier vor allem auf die Rechtslage in Deutschland, weist aber darauf hin, dass in allen europäischen Rechtssystemen ähnliche Gesetze gelten.

⁵⁴⁵ Zur Problematik des Copyrights in Archiv- und Bibliotheksinstitutionen siehe bspw. für die Schweiz: arbedo, Herausforderung Urheberrecht, Ausgabe 2, 2015. In der aktuellen Diskussion um die Modernisierung des Urheberrechts (Schweiz) werden diverse genannte Aspekte angegangen. Siehe: Eidgenössisches Institut für Geistiges Eigentum: Bundesgesetz über das Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz, URG) – Entwurf, 11.12.2015. URL: <https://www.ige.ch/urheberrecht/modernisierung-des-urheberrechts-2015.html>, Stand: 03.04.2016.

⁵⁴⁶ Siehe bspw. 'Julia Reda Report': Reda: Harmonisation of certain aspects of copyright and related rights in the information society, 2015. Urheberrechtsreformen werden u.a. unter den Begriffen 'Copyleft', 'Copyfight', 'Copy me' oder 'Copywrong' gefordert. Siehe weiterführend bspw.: McGuinness, Phillipa (Hg.): Copyfight, Sydney 2015. Die Australische Regierung äussert sich in ihrem Untersuchungsberichtes sehr kritisch zu den Bestimmungen des geistigen Eigentums und schlägt die Einführung von stark verbesserten 'Fair Use-Bestimmungen' für die Nutzer vor. Siehe: Australian Government Productivity Commission: Intellectual Property Arrangements, Report 78, Canberra 23.09.2016. URL: <http://www.pc.gov.au/inquiries/completed/intellectual-property/report>, Stand: 26.06.2016.

⁵⁴⁷ Siehe URL: http://www.axelspringer.de/presse/Axel-Springer-schliesst-Datendokumentation-ab-Gravierender-Schaden-durch-verschlechterte-Suchanzeigen-bei-Google_22070688.html.

zu Hyperlinks unter Urheberrechtsschutz gestellt und für abgabepflichtig erklärt.⁵⁴⁸ Google hat daraufhin per Mitte Dezember 2014 seinen Newsdienst in Spanien komplett eingestellt. Spanische Nachrichtenportale werden seither nicht mehr automatisch in der Suchmaschine aufgeführt, mit entsprechend negativen Auswirkungen auf die Nutzerzahlen. Akademische Institutionen sind von diesem Gesetz übrigens nicht ausgenommen und ebenfalls zu Urheberrechtsabgaben verpflichtet, was in Zeiten knapper Budgets zu Einsparungen in Lehre und Forschung führen wird.

Allerdings wird weiterhin über die Verlinkung von urheberrechtlichen Inhalten verhandelt. Der EU-Gutachter Melchior Wathelet (Sen.) ist zum Schluss gekommen, dass ein Link, der auf Inhalte verweist, die das Urheberrecht verletzen, selbst nicht rechtswidrig sei.⁵⁴⁹ Es ist zu erwarten, dass die Rechteinhaber die Gesetzgeber dahingehend zu beeinflussen versuchen, dass solche Verlinkungen schnellstmöglich unter Strafe gestellt werden, was wiederum Bürgerrechtsgruppen und Internetaktivisten wie die Electronic Frontier Foundation (EFF) oder die Piratenpartei zu verhindern versuchen. Für den Recherchierenden ist dieses Gutachten jedoch vorteilhaft, da die Verknüpfung von Inhalten (wieder) verbessert werden kann.

Auf EU-Ebene wird auch über die Auslegung der informationellen Selbstbestimmung diskutiert. Einerseits wird ein *Recht auf Vergessen* gefordert⁵⁵⁰, andererseits werden Datenschutzbestimmungen so stark gelockert, dass Unternehmen unter Angabe von *legitimen Interessen* persönliche Daten der Nutzer sammeln, speichern, verarbeiten und sogar Drittparteien zur Verfügung stellen dürfen.⁵⁵¹ Dagegen wird von Seiten der Verbraucher- und Datenschützer heftig opponiert, was weitere Verhandlungen über die Grenzen des erlaubten Datensammelns nötig macht.⁵⁵² Gestärkt wird die Position der Verbraucher- und Datenschützer durch mehrere Gerichtsentscheide. Google wurde in Spanien dazu verpflichtet, Indexteinträge in Google Search auf berechtigten Antrag hin zu löschen und darf in Suchresul-

⁵⁴⁸ Siehe: Brey, Mariano Rajoy: Modifica el texto refundido de la Ley de Propiedad Intelectual (11404), in: Boletín Oficial Del Estado (268), 05.11.2014, S. 90404-90439.

⁵⁴⁹ Wathelet, Melchior: Opinion of Advocate General Wathelet. Case C-160/15 (GS Media BV v Sanoma Media Netherlands BV, Playboy Enterprises International Inc., Britt Geertuida Dekker), ECLI:EU:C:2016:221, Gerichtshof der Europäischen Union, Luxembourg 07.04.2016. URL: <https://curia.europa.eu/>.

⁵⁵⁰ Vgl.: EU-Ratspräsidentschaft (Lettland): Proposal from Presidency to Permanent Representatives Committee: General Data Protection Regulation – Chapter III, 27.04.2015. URL: <http://data.consilium.europa.eu/doc/document/ST-6834-2015-INIT/en/pdf>.

⁵⁵¹ Siehe: Data; Protection Working Party (DAPIX); Data: Proposal from Presidency to Council: General Data Protection Regulation, 09.03.2015. URL: <http://data.consilium.europa.eu/doc/document/ST-6834-2015-INIT/en/pdf>, insb. Art. 38. Kommentar: Briegleb, Volker: EU-Länder einigen sich auf Eckpunkte der Datenschutzreform, Newsticker, heise online, 13.03.2015, URL: <http://www.heise.de/newsticker/meldung/EU-Laender-einigen-sich-auf-Eckpunkte-der-Datenschutzreform-2574894.html>, Stand: 01.05.2015.

⁵⁵² 66 Nichtregierungsorganisationen haben in einem offenen Brief den Schutz der Privatsphäre gefordert: McNamee, Joe: Open Letter to the President of the European Commission Jean-Claude Juncker concerning Data Protection, E-Mail, 21.04.2015. URL: https://edri.org/files/DP_letter_Juncker_20150421.pdf, Stand: 01.05.2015.

taten nicht mehr auf diese Inhalte verweisen.⁵⁵³ In England geht der Datenschutz so weit, dass Google sogar Indexeinträge löschen muss, die über die Löschung von Indexeinträgen berichten.⁵⁵⁴ Die Informationen verbleiben zwar auf den ursprünglichen Websites, aber die Suche danach wird erschwert – sie werden also aus dem Surface ins Deep Web verschoben. Das Recht auf informationelle Selbstbestimmung wird zwar wieder einmal bestätigt, aber damit wird auch die Verfälschung eines Suchergebnisses durch willentliche Zensur von eigentlich auffindbaren Objekten legitimiert.

Selbst für Behörden können die geltenden Urheberrechtsbestimmungen hinderlich sein. Beispielsweise ist die Rechtslage von in staatlichem Auftrag archivierten Webauftritten meist ausserhalb lizenzrechtlicher Vereinbarungen und deshalb problematisch. Eine Ausnahme ist das dänische Webarchiv *netarkivet*, das per Gesetz⁵⁵⁵ ermächtigt ist, jeden unter der *Top Level Domain*⁵⁵⁶ *.dk* publizierten Webauftritt zu archivieren. Das Netarkivet ist zusätzlich berechtigt, Webauftritte unter anderen Domains zu archivieren, die sich an ein dänisches Publikum richten, das heisst, wenn der Inhalt in dänischer Sprache geschrieben ist oder er sich auf Dänemark, seine Einwohner oder dänische Sachen bezieht. Damit hat Dänemark klare rechtliche Rahmenbedingungen geschaffen, wie sie in vielen anderen Ländern noch nicht bestehen. Ohne solche Gesetze könnten allenfalls auch staatliche Archive zur Löschung von Inhalten gezwungen werden.

Ein weiterer Problemkreis ist die Verwendung von Hyperlinks. Während diese in der Schweiz grundsätzlich nicht als Aneignung des darunter enthaltenen Gedankenguts gilt⁵⁵⁷, wird sie in anderen Ländern als technische Verbreitung gewertet und ist einklagbar, wenn sich der Linksetzer nicht mit einem Haftungsausschluss, dem sogenannten Disclaimer, deutlich

⁵⁵³ Siehe: EuGH: Urteil des Europäischen Gerichtshofes; Rechtssache C-131/12, InfoCuria – Rechtsprechung des Gerichtshofs, 13.05.2014, URL: <http://curia.europa.eu/>.

⁵⁵⁴ Siehe: Information Commissioner's Office: Data Protection Act 1998 – Superprovisory Powers of the Information Commissioner – Enforcement Notice, 18.08.2015. URL: <https://ico.org.uk/media/action-weve-taken/enforcement-notice/1432380/google-inc-enforcement-notice-18082015.pdf>, Stand: 21.08.2015.

⁵⁵⁵ URL: <http://netarkivet.dk/>. Gesetz: Margrethe II., Königin von Dänemark: Lov om pligtaflevering af offentliggjort materiale (Lov nr. 1439 af 22. december 2004), Pligtaflevering, 22.12.2004, URL: <http://www.pligtaflevering.dk/loven/index.htm>, Stand: 31.03.2015. Ebenso gesetzlich ermächtigt ist bspw. die Österreichische Nationalbibliothek: Nationalrat der Republik Österreich: Bundesgesetz über die Presse und andere publizistische Medien (Mediengesetz – MedienG), 19.04.2015. URL: <http://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10000719>, Stand: 09.09.2015, § 43b. Abs. 1; Recht, Christian: Webarchivierung und Webharvesting der ÖNB, in: *ipCompetence* (2), 2009, S. 42-53.

⁵⁵⁶ Domain: Klartextnamen von Webseiten im Internet, hierarchisch gegliedert. Top Level Domain: oberste Ebene der Hierarchie mit von der Verwaltungsorganisation ICANN (URL: <https://www.icann.org/>) vergebenen Standard- und Länderdomains wie *.com*, *.ch* oder *.dk*. Nach: Prevezanos: *Computer Lexikon* 2013, 2012.

⁵⁵⁷ Siehe: Volz, Stephanie: Trennungsgebot und Internet: Ein medienrechtliches Prinzip in Zeiten der Medienkonvergenz, Bd. 58, Zürich 2014 (Publikationen aus dem Zentrum für Informations- und Kommunikationsrecht der Universität Zürich), S. 168.

distanziert, wobei er selbst damit nicht gänzlich geschützt ist.⁵⁵⁸ Die Verwendung eines Hyperlinks ist folglich auch in der wissenschaftlichen Forschung und Lehre nicht unproblematisch, weil die Rechtslage nicht gänzlich geklärt ist.

Ein Ansatz zur Minderung der Schutzfristen von urheberrechtlich geschützten analogen und digitalen Werken orientiert sich am Patentwesen.⁵⁵⁹ Die Schutzrechte werden ans Werk und nicht mehr an den Urheber gebunden und die Schutzfrist eines Werks wird auf zwanzig Jahre ab Publikationsdatum reduziert.⁵⁶⁰ Sie erwächst weiterhin mit der Schaffung des Werks, im Gegensatz zum anzumeldenden Patent, und der Urheber hat in dieser Zeit die Möglichkeit zu dessen Verwertung. Dieser Ansatz würde klare, begrenzte und sinnvolle Rahmenbedingungen schaffen, die sowohl dem Urheber die finanzielle Verwertung seines Werks ermöglichen, aber auch der kulturellen Weiterentwicklung der Gesellschaft und der wissenschaftlichen Forschung zugute kämen. Damit würden absurde Debatten wie diejenige um das Urheberrecht der Tagebücher von Anne Frank gar nicht erst entstehen.⁵⁶¹ Eine Schutzfrist sollte die kulturelle Weiterentwicklung der gesamten Gesellschaft nicht behindern. Die rege geführten Diskussionen zeigen auf, dass ein grosser gesetzlicher Handlungs- und Klärungsbedarf besteht.⁵⁶² Die rechtlichen Unsicherheiten hemmen die wissenschaftliche

⁵⁵⁸ Siehe bspw. Urteil des Bundesgerichtshofes der Bundesrepublik Deutschland Az.: I ZR 74/14: Ein Anbieter haftet dann für einen von ihm gesetzten Link, wenn der rechtswidrige Inhalt auf der verlinkten Seite deutlich erkennbar ist oder er auf die Rechtswidrigkeit hingewiesen wurde. Siehe URL: <http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=en&nr=73259&pos=0&anz=1>.

⁵⁵⁹ Idee an verschiedenen Veranstaltungen diskutiert. Siehe: Föhr, Pascal: Anpassung Urheberrechtsgesetz, Historical Source Criticism, 01.03.2016, URL: <http://hsc.hypotheses.org/382>, Stand: 26.03.2016; Die Grünen - Europäische Freie Allianz: Kreatives Schaffen und Urheberrecht im digitalen Zeitalter – Positionspapier, 21.11.2011. URL: <http://www.greens-efa.eu/de/kreatives-schaffen-und-urheberrecht-im-digitalen-zeitalter-4539.html>, Stand: 31.08.2015, §26. Zu gesetzlichen Bestimmungen zum Patentwesen (Schweiz) siehe: Bundesversammlung der Schweizerischen Eidgenossenschaft: Bundesgesetz über die Erfindungspatente (Patentgesetz), 01.01.2012. URL: <https://www.admin.ch/opc/de/classified-compilation/19540108/index.html>, Stand: 06.10.2015; Schweizerischer Bundesrat: Verordnung über die Erfindungspatente (Patentverordnung), 01.09.2014. URL: <https://www.admin.ch/opc/de/classified-compilation/19770250/index.html>, Stand: 06.10.2015.

⁵⁶⁰ Es ist zu unterscheiden zwischen datenschutzrechtlichen und urheberrechtlichen Schutzfristen. Verwaltungsunterlagen unterliegen meist datenschutzrechtlichen Fristen, sind aber urheberrechtlich nicht geschützt.

⁵⁶¹ Der Anne Frank Fonds in Basel bezeichnet den 1980 verstorbenen Otto Frank seit kurzem als Mitautor, was den Schutz bis 2050 ausdehnen würde. Siehe: Carvajal, Doreen: Anne Frank's Diary Gains 'Co-Author' in Copyright Move, in: The New York Times, New York (USA) 13.11.2015. URL: <http://www.nytimes.com/2015/11/14/books/anne-frank-has-a-co-as-diary-gains-co-author-in-legal-move.html>, Stand: 04.12.2015.

⁵⁶² Diverse Vorschläge werden von verschiedenen Seiten zur Diskussion vorgelegt. In der Schweiz ist im Sommer 2015 eine Expertengruppe zur Überarbeitung des Urheberrechtsgesetzes eingesetzt worden, die den genannten Entwurf vorlegten: Finanzdepartement der Schweizerischen Eidgenossenschaft: Expertengruppe "Zukunft der Datenbearbeitung und Datensicherheit" eingesetzt, Medienmitteilungen, 27.08.2015, URL: <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-58475.html>, Stand: 28.08.2015; Eidgenössisches Institut für Geistiges Eigentum: Entwurf URG, 2015. Für Deutschland siehe bspw.: Pfeifer, Karl-Nikolaus; Frey, Dieter; Rudolph, Matthias: Urhebervertragsrecht in der Reform: Der "Kölner Entwurf", in: Gewerblicher Rechtsschutz und Urheberrecht, Praxis im Immaterialgüter und Wettbewerbsrecht (1), 2015, S. 1-25;

Arbeit und ermöglichen die Ausdehnung wirtschaftlicher Interessen von Privaten, Unternehmen und Institutionen auf Informationen, die eigentlich gemeinfrei und frei zugänglich sind oder sein sollten. Die Durchsetzung von Open Access in der Wissenschaft sowie klare Richtlinien für die Zweitverwertung, die die Rechte des originären Autors und die des dessen Werk verwendenden Autors regelt, wären wichtige Grundlagen für eine freie Entfaltung der wissenschaftlichen Forschung. Es sollten deshalb relativ zügig klare gesetzliche Grundlagen für die Verwendung von urheberrechtlich geschützten Werken in der Wissenschaft geschaffen werden. Die Geschichtswissenschaft soll dabei möglichst frei über Werke verfügen können, sofern Herkunft und Urheber klar nachvollziehbar sind. Den dazu per Konvention vertrauenswürdigen Institutionen wie Archive oder Bibliotheken für die Aufbewahrung von möglichen Quellen ist der entsprechende Handlungsspielraum zu gewähren, damit beispielsweise der Versand von Kopien rechtskonform erfolgen kann. Es sind Vorstösse zu begrüßen, die die Offenlegung von Forschungsdaten erwirken⁵⁶³, den Umgang mit geschützten Werken für die Wissenschaft erleichtern oder eine grundsätzliche Reform des Urheberrechts zu Gunsten einer für die Allgemeinheit vorteilhaften Nutzungsregelung umsetzen wollen.⁵⁶⁴ Hingegen ist eine Ausdehnung von Schutzbestimmungen wie im Copyright Term Extension Act oder der TPP und TTIP strikt abzulehnen.

5.1.7. Zusammenfassung

Bei der Suche nach potenziellen Quellen ist unter Verwendung von elektronischen Informationssystemen, des Internets mit seinen Diensten sowie von digitalen Objekten noch grössere Vorsicht geboten als bei einer traditionellen Suche. Für die Festlegung der Suchstrategie hat der Recherchierende eine hohe (Digital-)Medienkompetenz über digitale Objekte sowie digitale Informations- und Kommunikationsmedien vorzuweisen.

In der Geschichtswissenschaft wird mit *Suchstrategie* ein nachvollziehbarer und systematischer Vorgehensplan bezeichnet, wie unter Berücksichtigung bestimmender Faktoren nach Objekten recherchiert wird, die für die Beantwortung einer bestimmten Fragestellung massgeblich sind.

Zudem sollte eine konkrete Fragestellung vorliegen, die theoretischen Grundlagen erarbeitet, die einschränkenden Faktoren zu eruiert und die Suchräume festgelegt werden.

Christiansen, Per; Gehlen, Dirk von; Hofmann, Jeanette u. a.: Ein Gedankenexperiment zur Neuordnung des Urheberrechts, irights.info, 27.08.2015. URL: <http://irights.info/2015/08/29/ein-gedankenexperiment-zur-neuordnung-des-urheberrechts/25980>, Stand: 01.09.2015.

⁵⁶³ Bspw.: Holdren, John P.: Increasing Access to the Results of Federally Funded Scientific Research. Memorandum For The Heads Of Executive Departments And Agencies, Executive Office of the President. Office of Science and Technology Policy, 22.02.2013.

⁵⁶⁴ Siehe erwähnte Forderung: Die Grünen - Europäische Freie Allianz: Kreatives Schaffen und Urheberrecht im digitalen Zeitalter – Positionspapier, 2011, S. § 26.

Dem Nutzer sollte bekannt sein, wie eine Suchmaschine funktioniert, damit er die Resultatliste einer Suche kritisch beurteilen kann. Insbesondere die algorithmisch berechnete Gewichtung ist problematisch, weil sie nicht nur die Zahl der aufgefundenen Objekte beschränkt, sondern diese in eine vermeintliche Ordnung bringt und damit die Relevanzbeurteilung durch den Nutzer teilweise vorwegnimmt. Wissenschaftler dürften sich nicht auf eine algorithmische Gewichtung verlassen, sondern haben alle potenziellen Quellen zu erfassen, zu untersuchen und deren Relevanz für die Forschungsfrage zu eruieren. Suchmaschinen sind als unzuverlässige Recherchemittel zu betrachten.

Die aufgefundenen Objekte und Informationen sind auf ihre Vollständigkeit hin zu überprüfen. Beim Einsatz von algorithmisch gestützten Selektions- und Filterverfahren ist mit Informationsverlust zu rechnen, weil einerseits die Algorithmen nicht immer zuverlässig funktionieren und andererseits eventuell relevante Informationen unterdrückt werden.

Mit der Hervorhebung des gesuchten Elements wird dieses dekontextualisiert. Für eine adäquate Beurteilung ist es zu rekontextualisieren. Das Objekt oder das Element ist mit anderen Objekten oder Elementen in Beziehung zu setzen, damit eine fundierte Auswahl an Objekten für die weitere Bearbeitung getroffen werden kann.

Rechtliche Rahmenbedingungen können die Suchräume und die Verwendung der potenziellen Quellen einschränken, insbesondere bei digitalen Objekten. Schutzfristen oder spezielle Lizenzen können einen freien Zugriff auf und die freie Verwendung von Informationen und Wissen verhindern respektive an bestimmte Bedingungen knüpfen, was sich hemmend auf die wissenschaftliche Entwicklung auswirken kann.

Jan Hodels Aussage, dass eine Recherche mit einer hybriden Suchtechnik am erfolgreichsten sei, sollte ergänzt werden: es ist nicht nur die Suchtechnik relevant, sondern es ist auch das Wissen darüber relevant, was mit einer Suche nicht gefunden werden kann und wie mit den gefundenen Objekten umgegangen werden kann oder darf. Der heuristische Prozess ist in der digitalen Welt komplexer als in der analogen und fordert hohe Medienkompetenzen.

5.2. Quellenkunde

Digitale Objekte sind in ein Quellengattungsschema einzuordnen, wobei sie als neue Quellenart nicht in die bekannten Schemata eingefügt werden können. Digitale Quellen sind als virtuelle, elektronische Überreste in einem neuen Schema zu verorten. Aus digitalen Objekten algorithmisch generierte neue digitale Objekte (reborn-digital objects) können die Funktion von Primär-, Sekundär- oder Tertiärquellen übernehmen und zusätzlich, und dies im Gegensatz zu analogen Objekten, diejenige von Quartär- und Quintärquellen. Jedes digitale Objekt wurde durch einen Algorithmus in einem elektronischen Informationssystem und nicht

von einer natürlichen Person erzeugt. Reborn-digitale Objekte werden algorithmisch aus bereits digitalen Objekten generiert und stellen eine neue Art von Funktion dar.

5.2.1. Funktion

Die zwei *Funktionen* der Quartär- und Quintärquellen ergänzen die bisher in der traditionellen Methode⁵⁶⁵ bekannten Unterscheidungen. Dass bei digitalen Objekten eine solche aufgrund ihrer fehlenden Materialität nur intellektuell erfolgen kann, spielt eine untergeordnete Rolle. Auch bei physischen Objekten ist die Materialität nur ein Kriterium, der Überlieferungszusammenhang, der Inhalt oder die Angaben über den oder die Autor(en) liefern weitere Hinweise.

Bei einer *Quartärquelle* erstellt das elektronische Informationssystem das Ergebnis mittels eines Algorithmus quasi autonom. Allerdings ist das Ergebnis aufgrund von Verarbeitungsproblemen des Systems oder Fehlern im Code des Algorithmus nicht zwingend fehlerfrei.⁵⁶⁶ Zudem ist es möglich, dass ein Algorithmus kein Ergebnis liefert, weil er entweder eine unendliche Schleife enthält oder ein Problem aufgreift, das mathematisch nicht gelöst werden kann.⁵⁶⁷ Auch sollte sichergestellt werden, dass das Ergebnis nicht durch Optimierungen durch Dritte beeinflusst wurde.⁵⁶⁸

Die Aussagekraft der Quartärquelle wird durch die genannten Probleme der algorithmischen Verarbeitung gemindert und verringert sich weiter, je weniger über den Algorithmus und dessen Faktoren bekannt ist.

Die fünfte Funktion betrifft vor allem den Bereich der künstlichen Intelligenz (KI). Wird ein Algorithmus durch ein Computersystem geschrieben, das heisst ein Algorithmus generiert selbständig einen Algorithmus, wird dessen Ergebnis als *Quintärquelle* zu bezeichnen sein.⁵⁶⁹

Reborn-digitale Objekte entstehen per Definition aus digitalen Daten und sind deshalb weder Primär-, noch Sekundär- oder Tertiärquellen. Sie können nur die Funktionen von Quartär- und Quintärquellen übernehmen. Quartär- und Quintärquellen sind nur algorithmisch gene-

⁵⁶⁵ Siehe Kap. 3.1.2. Quellenkunde.

⁵⁶⁶ Die Firma Intel, die einen Grossteil der in heutigen Computern verwendeten Prozessoren produziert, musste eingestehen, dass eine Produktionsreihe fehlerhaft rechnet. Vgl.: Stiller: 20 Jahre FDIV-Bug, 2014.

⁵⁶⁷ Bspw. das 'P=?NP' - Problem: Meyer, Steven: The Effect of Computers on Understanding Truth, in: Hagengruber, Ruth; Ess, Charles (Hg.): The Computational Turn: Past, Presents, Futures? Proceedings IACAP 2011, Münster 2011, S. 49-52. URL: <http://homepages.uc.edu/~freemago/papers/IACAP.pdf>, Stand: 06.05.2018.

⁵⁶⁸ Bspw. durch sog. Search Engine Optimization SEO. Damit wird eine Webseite so angepasst, dass ein externer Suchalgorithmus die Webseite möglichst weit oben im Ranking aufführt.

⁵⁶⁹ Dies ist in Ansätzen bereits möglich. Siehe: Murali, Vijayaraghavan; Qi, Letao; Chaudhuri, Swarat u. a.: Neural Sketch Learning for Conditional Program Generation, arXiv:1703.05698 [cs.PL], 16.03.2017. URL: <http://arxiv.org/abs/1703.05698>, Stand: 30.04.2018.

rierte Objekte. Daraus ist zu folgern, dass Quartär- und Quintärquellen ausschliesslich aus reborn-digitalen Objekten bestehen und reborn-digitale Objekte ausschliesslich die Funktion von Quartär- und Quintärquellen übernehmen.

Somit wird definiert:

Reborn-digitale Objekte übernehmen die Funktion von *Quartär-* (algorithmisch generiertes Objekt) und *Quintärquellen* (Objekt, das durch einen von einem Algorithmus geschaffenen Algorithmus entstanden ist), die wiederum nur aus reborn-digitalen Objekten bestehen.

5.2.2. Quellengattung

In den bisher verwendeten Quellengattungsschemata, die vor allem physische Objekte berücksichtigen, lässt sich die neue Quellenart der digitalen Objekte nicht einordnen. Zwar konnten auch bisher nicht alle Quellenarten in solchen Schemata eindeutig verortet werden, da diese nur einen groben Raster als Hilfeleistung für die Auswahl von Untersuchungsmethoden zur Verfügung stellen. Da digitale Objekte immer mehr an Bedeutung gewinnen, mittlerweile einen Grossteil der verfügbaren Quellen darstellen und sie sich durch eindeutige Merkmale von physischen unterscheiden, ist ein Schema unter Berücksichtigung dieser Quellenart zu definieren. Aufgrund der obigen Darlegungen⁵⁷⁰ empfiehlt sich zudem eine Loslösung von der Einteilung in Überrest und Tradition.

Lösungsansatz Quellengattungsschema

Es wird unter Berücksichtigung bereits bekannter Schemata ein Quellengattungsschema vorgeschlagen, das die Überreste in die zwei eindeutigen Klassen *materiell / physisch* und *immateriell / virtuell* unterteilt. Materiell ist alles physisch Vorhandene wie ein Haus, ein Dokument oder ein Siegel und alles nicht Materielle ist immateriell oder virtuell wie die Sprache, eine Institution oder ein digitales Objekt. Innerhalb der materiellen Klasse können vier Kategorien unterschieden werden: *Überreste* im engeren Sinn, also alles, was mit dem täglichen Leben eines Lebewesens verbunden ist, *Denkmäler* sowie *bildliche* und *schriftliche Überreste*. In der immateriellen Klasse wird nach den drei Kategorien *Audio / Audiovisuell*, *Zustände* und *Elektronisch* unterschieden. Obwohl Musik sowohl in der Kategorie Audio als auch in Elektronisch aufgeführt wird, ist die Unterscheidung eindeutig: ein Musikstück, das als .mp3-Datei vorliegt, ist als digitales Objekt der elektronischen Kategorie zuzuordnen. Wird es abgespielt, ist das, was gehört wird, der Kategorie Audio zuzuordnen.

⁵⁷⁰ Siehe insb. Kap. 3.1.2. Quellenkunde.

Innerhalb der Kategorie Elektronisch ist zwischen zur Darstellung vorgesehenen und nicht zur Darstellung vorgesehenen Objekten zu unterscheiden. Erstere sind beispielsweise Bild-, Musik- oder Textobjekte. In die zweite Gruppe gehören teilweise versteckte Objekte wie CSS-Dateien, die die Darstellung eines Objekts der ersten Gruppe bestimmen, oder Hilfsdateien wie Programmbibliotheken, die das Betriebssystem, ein Programm oder andere Elemente des Informationssystems unterstützen. Ein Code kann in den meisten Fällen ebenfalls der zweiten Gruppe zugeordnet werden, da dieser an sich nicht dargestellt wird, sondern als zusätzliche Objektebene zum dargestellten Objekt besteht und durch das Informationssystem interpretiert wird (the code is not the text, unless it is the text). Ebenso sind Datenbankelemente nicht sichtbar und werden nur mit Hilfe des Datenbanksystems dargestellt. Eine Unterscheidung zwischen digitalisierten, genuin digitalen und reborn-digitalen Objekten ist im Schema jedoch nicht nötig, da alle drei die gleichen Eigenschaften aufweisen.

Daraus ergibt sich das folgende Schema, in das sowohl traditionell bekannte als auch neue Quellenarten eingeordnet werden können:

Überreste						
materiell / physisch				immateriell / virtuell		
Überreste im engeren Sinne	Denkmäler	Bildlich	Schriftlich	Audio/ audiovisuell	Zustände	Elektronisch
<ul style="list-style-type: none"> - körperliche Reste - Werkzeuge - Wohnen - Produkte 	<ul style="list-style-type: none"> - Inschriften - Monumente - Urkunden 	<ul style="list-style-type: none"> - historische Gemälde - topographische Darstellungen - historische Skulpturen 	<ul style="list-style-type: none"> - historische Inschriften - Genealogien - Kalender - Annalen - Chroniken - Biographien - Memoiren - Urkunden und Akten (Vorgänge rechtlicher Natur) - Literatur - Geschäftsakten 	<ul style="list-style-type: none"> - Erzählung - Sage - Anekdoten - Sprichwörter - historische Lieder - Sprache - Ton, Musik 	<ul style="list-style-type: none"> - Sitten - Prozesse - Institutionen 	<ul style="list-style-type: none"> Darstellung - Graphiken (dynamisch / statisch) - Texte - Webseiten - Musik - Film Versteckt - Programm bibliotheken - Code - Datenbank-elemente

Tabelle 3: Quellengattungsschema nach Pascal Föhr; Aufzählungen nicht abschliessend

Die Bestimmung der Quellenart eines digitalen Objekts ist einerseits für die Quellensicherung, andererseits für die Quellenkritik nützlich respektive notwendig. Die Art des Objekts bestimmt in der Quellensicherung, welche Teile des Objekts in welcher Form mit welchen Hilfsmitteln gespeichert und vorgehalten werden.⁵⁷¹ In der Quellenkritik dient die Verortung dazu, die Komponenten des zu verwendenden Informationssystems zu bestimmen, mit dem das Objekt wieder dargestellt werden soll.

⁵⁷¹ Teile: bspw. einzelne Datei, mit oder ohne Metadaten, einzelne Webseite eines Webauftritts; Form: bspw. Dateiformat, Binärcode, Bild, Text; Hilfsmittel: bspw. Datenbank, bestimmtes Programm.

Die oben erwähnte quellenkundliche Medienkompetenz wird hier benötigt, um aus den Suchresultaten die für die Fragestellung sinnvollen digitalen Objekte auswählen zu können und Klone zu erkennen.

5.2.3. Zusammenfassung

In der Quellenkunde werden digitale Objekte nach Funktion und Gattung verortet. Das Vorgehen entspricht zwar dem bereits bekannter Objekte, aber digitale können zwei zusätzliche Funktionen übernehmen und sind separat zu klassifizieren.

Reborn-digitale Objekte übernehmen die Funktion von Quartär- (algorithmisch generiertes Objekt) und Quintärquellen (Objekt, das durch einen von einem Algorithmus geschaffenen Algorithmus entstanden ist), die wiederum nur aus reborn-digitalen Objekten bestehen.

Es wird ein Quellengattungsschema vorgeschlagen, das die Überreste in die zwei eindeutigen Klassen materiell / physisch und immateriell / virtuell unterteilt. Innerhalb der materiellen Klasse werden die vier Kategorien Überreste (im engeren Sinn), Denkmäler, bildliche und schriftliche Überreste unterschieden. Die immateriellen Klasse enthält die drei Kategorien Audio / Audiovisuell, Zustände und Elektronisch. In letztgenannter Kategorie befinden sich zur Darstellung vorgesehene und versteckte, digitale Objekte.

5.3. Quellensicherung

"[...] die Sicherung digitaler Quellen und der Umgang mit ihnen [stellen] eine besondere Herausforderung dar, [...]."⁵⁷² (Koller, 2016)

Die Quellensicherung wird als eigenständiger Prozessschritt der historischen Methode definiert, weil die Wahrscheinlichkeit des Verlustes bei digitalen Objekten besonders aufgrund ihrer Volatilität, Manipulierbarkeit und der Abhängigkeit von einem spezifischen, digitalen Informationssystem viel höher ist als bei materiellen. Diesen Umständen wurde bisher zu wenig Beachtung geschenkt oder sie wurden marginalisiert. Wenn digitale Objekte als potenzielle Quellen verwendet und untersucht werden sollen, sind sie in einer geeigneten Infrastruktur zu sichern. Dies bedingt neue methodische Arbeitsweisen, die sich Geschichtswissenschaftler bisher nicht gewohnt waren. Fragen zur Auswahl des Objekts, zu Form und Formaten, dem Sicherungsort und den zugehörigen Metadaten sind bei jeder Sicherung für jedes Objekt individuell zu beantworten und erfordern neue Lösungsansätze.

Dieser Prozessschritt war bisher nicht vorgesehen, da grundsätzlich davon ausgegangen wurde und wird, dass sich eine Quelle an einem per Konvention *sicheren Ort* wie einem

⁵⁷² Koller: Geschichte digital, 2016, S. 22.

Archiv, einer Bibliothek oder einem Museum befindet.⁵⁷³ Werden durch Ausgrabungen, Denkmalsicherungen oder andere Tätigkeiten neue zu sichernde Objekte gefunden, werden sie oder zumindest eine schriftliche und photographische Dokumentation davon oft in eine solche sichere Institution überführt. Dadurch können die meisten physischen Quellen bei korrekter Lagerung zu einem späteren Zeitpunkt ohne zusätzliche Verfahren wiederverwendet und gelesen werden.⁵⁷⁴

Bei digitalen Objekten ist dies nicht grundsätzlich möglich, weil sie zur Darstellung ein digitales Informationssystem benötigen. Dieses besteht aus Hard- und Softwarekomponenten, die einem schnellen Alterungsprozess unterworfen sind und unter Umständen binnen relativ kurzer Zeit nicht mehr funktionieren. Deshalb genügt es nicht, ausschliesslich den binären Code zu erhalten, um ein künftiges Auffinden, Lesen und Interpretieren gewährleisten zu können. Zudem bedingen die Eigenschaften der Volatilität und nicht nachvollziehbaren Manipulierbarkeit eines digitalen Objekts grundsätzlich eine *Objektsicherung*, um die Verfügbarkeit, Integrität und Authentizität des digitalen Objekts möglichst lange gewährleisten zu können. Vor allem Forschungsressourcen im Web können innerhalb sehr kurzer Zeit verändert oder gelöscht werden und bestehen deshalb später allenfalls nicht mehr oder nicht mehr so, wie zum Zeitpunkt der Verwendung durch den Nutzer. Die Quellensicherung ist deshalb nicht nur für die eigene quellenkritische Überprüfung, sondern auch für eine spätere, von einer Drittperson durchgeführte Quellenkritik von grosser Wichtigkeit.⁵⁷⁵ Der Nutzer ist verpflichtet zu beweisen, dass die von ihm verwendete Forschungsressource so oder in sehr ähnlicher Form existierte, weshalb das Objekt manipulationssicher zur Verfügung zu stellen ist. Wenn digitale Objekte für eine langfristig nachvollziehbare Argumentation verwendet werden sollen, ist eine zusätzliche Sicherung in einem vertrauenswürdigen Archiv zwingend notwendig.

Die für physische Objekte verwendete *Textsicherung*⁵⁷⁶ reicht hierbei nicht aus. Eine Quellensicherung ist unabhängig von der Quellenart bei allen zur Verwendung vorgesehenen digitalen Objekten durchzuführen. In Anbetracht des immer grösser werdenden digitalen Datenvolumens ist es sinnvoll, die Quellensicherung als eigenständiger Prozessschritt in die historische Methode einzuführen.

Damit verbunden sind teilweise bereits angesprochene Probleme, die bei der Archivierung von materiellen Objekten selten oder nicht auftreten. Zu erwähnen sind hierbei die sich

⁵⁷³ Zu Vertrauenswürdigkeit und Nachvollziehbarkeit von Archiven siehe: Gänser, Georg: Archive, Nachvollziehbarkeit und Vertrauen, in: *scrinium* 71-2017, 2017, S. 102-113.

⁵⁷⁴ Unter korrekter Lagerung werden die für Quellen idealen Bedingungen in einer archivischen Institution verstanden.

⁵⁷⁵ Siehe Kap. 3.4. Wissenschaftliches Arbeiten.

⁵⁷⁶ Bezogen auf die Sicherung der Lesbarkeit von Textelementen in analogen Objekten. Siehe auch Kap. 3.1.3. Quellenkritik.

grundsätzlich zu stellenden Fragen, ob das elektronische Informationssystem mit zu archivieren ist, welche Code-Ebenen zu sichern sind, wie die Langzeitarchivierung gewährleistet wird oder welche Metadaten zusätzlich gespeichert werden sollen oder müssen.⁵⁷⁷ Solche Fragen können nur beantwortet werden, wenn Experten aus dem Archivwesen, der Informatik und den jeweiligen Fachgebieten eng zusammenarbeiten. Die zunehmende Komplexität der Informationstechnik erfordert nach wissenschaftlichen Standards aufgebaute, interdisziplinäre Lösungen für die Quellensicherung. Nur gemeinschaftlich erarbeitete Lösungen stellen die Nutzung von digitalen Forschungsressourcen für alle Wissenschaftszweige auch in Zukunft sicher. Dafür sind neue Methoden zu erarbeiten und anzuwenden sowie ständig zu überprüfen und weiterzuentwickeln. Die Ansprüche der einzelnen wissenschaftlichen Fachgebiete und Forschungsgruppen sind sehr individuell, weshalb hier nur Lösungsansätze präsentiert werden können.

Aufgrund obiger Ausführungen wird vorgeschlagen, die traditionelle historische Methode um den Schritt der Objektsicherung zu ergänzen:

Der Prozess der traditionellen historischen Methode wird um den Schritt *Quellensicherung* erweitert, um den Zugriff auf eine digitale Forschungsressource sowie deren Verfügbarkeit dauerhaft zu gewährleisten.

Ziel der Quellensicherung ist es, ein digitales Objekt unter grösstmöglicher Wahrung von dessen Integrität und Authentizität⁵⁷⁸ so zu speichern, dass es möglichst dauerhaft für die wissenschaftliche Forschung nachvollziehbar (wieder-)verwendet werden kann. Da damit einige informationstechnische Probleme verbunden sind, die bei einer physischen Objektsicherung nicht auftreten, werden Fragen nach der Objektauswahl und -form, dem Format, den Komponenten des Informationssystems, dem Sicherungsort und den Metadaten ausführlich behandelt. Der folgende Exkurs zur Langzeitarchivierung zeigt die Probleme der dauerhaften Sicherung von digitalen Objekten auf, was diesen Schritt im quellenkritischen Prozess massiv beeinflusst. Des Weiteren werden in diesem Kapitel Konzepte und Techniken der Langzeitarchivierung diskutiert und dargestellt, weil sie eng mit dem Schritt der Quellensicherung und späteren Wiedernutzung digitaler Objekte von grundlegender Bedeutung sind.

⁵⁷⁷ Zur Archivierungsproblematik siehe bspw.: Rosenthaler: Archivierung im digitalen Zeitalter, 2007; Gladney: Preserving Digital Information, 2007; Masanès: Web Archiving, 2006.

⁵⁷⁸ Zur Präzisierung der hier verwendeten Begriffe: *Authentizität* wird als Aussage über die Echtheit, Zuverlässigkeit oder Glaubwürdigkeit einer Person oder Sache verstanden. Die *Authentifikation* ist der Prozess der Identitätsüberprüfung mittels *Authentifizierung* (Überprüfung der Person oder Sache durch Externe) und/oder *Authentisierung* (Nachweis der Identität durch die Person oder Sache selbst). Im Englischen wird allerdings der Begriff 'authentication' sowohl für Authentifizierung als auch für Authentisierung verwendet. Eine Identitätsprüfung ist im Kontext der historisch-kritischen Methode die Frage an das Objekt, ob es das ist, was es zu sein scheint. Die Autorisierung hingegen ist die Berechtigung oder Ermächtigung einer Person oder Sache, Zugang oder Zugriff auf eine Person oder Sache zu haben.

5.3.1. Digitale Langzeitarchivierung (Exkurs)

Die Archivwissenschaften beschäftigen sich seit längerer Zeit mit den Problemen der Langzeitarchivierung von digitalen Objekten. Verschiedene grössere und kleinere Archive haben bereits entsprechende Infrastrukturen geschaffen. Die in solchen Projekten gewonnenen Erkenntnisse bilden die Grundlage für die Quellensicherung in der historisch-kritischen Methode. Einzelne Aspekte werden später nochmals ausführlicher diskutiert.

Bei der Archivierung von digitalen Objekten sind vor allem fünf Problembereiche zu beachten:⁵⁷⁹

- *Alterung der Medien ('Verdampfung' der Bits):* Durch verschiedene Einwirkungen altern Datenträger und werden mit der Zeit unlesbar. Der Verlust eines einzigen Bits kann zum Totalverlust des Datenträgers führen. Ein physisches Objekt bleibt hingegen meistens auch dann weiterhin nutzbar, wenn Teile davon fehlen. Wird beispielsweise das Inhaltsverzeichnis aus einem Buch herausgetrennt, ist der Inhalt immer noch zugänglich und lesbar.
- *Obsoleszenz der Datenträger:* Alte Datenträgertypen werden durch neue abgelöst oder nicht mehr hergestellt. Die für die Darstellung nötigen Verarbeitungsgeräte (lesen, schreiben, speichern) werden ebenfalls nicht mehr hergestellt, wenn der Datenträger nicht mehr produziert wird.
- *Abhängigkeit von obsoleter Software:* Bestimmte Dateiformate bedingen spezifische Software zur Verarbeitung. Wenn die Software nicht mehr lauffähig oder erhältlich ist, können die Objekte auch nicht mehr gelesen und auf andere Datenträger geschrieben oder gespeichert werden.
- *Abhängigkeit von obsoleten Software-Umgebungen:* Ein Programm ist oft nur in einem spezifischen Softwaresystem lauffähig, dessen Komponenten wie das Betriebssystem oder die Treiber in bestimmten Versionen vorhanden sein müssen.
- *Abhängigkeit von obsoleter Hardware:* Software benötigt ein spezifisches Hardwaresystem, dessen Komponenten wie Graphikkarte oder Laufwerke in einer bestimmten Version vorhanden sein müssen.

Jeff Rothenbergs Ausspruch "Digital objects last forever – or five years, whichever comes first"⁵⁸⁰ zeigt die ungefähre zeitliche Dimension auf, in der die Integrität eines digitalen Objekts und dessen *archetypischen Systems*⁵⁸¹ gewährleistet werden kann. Nach etwa fünf Jahren ist mit einem Totalverlust eines digitalen Objekts zu rechnen, wenn im Archivierungsprozess keine erhaltenden Massnahmen durchgeführt wurden.

⁵⁷⁹ Vgl.: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 450 f.

⁵⁸⁰ Rothenberg: Digital Preservation in Perspective, 2012, Fol. 16.

⁵⁸¹ Archetypisches System: Ursprüngliches System, das alle nötigen Komponenten in der benötigten Konfiguration enthält.

Um diesen drohenden Verlust zu verhindern oder zu minimieren, wurden verschiedene *Methoden* entwickelt.⁵⁸²

- *Datenträger und Datenformate verbessern*: Einerseits sollen Datenträger länger haltbar sein, damit der darauf gespeicherte Bitcode gelesen werden kann und andererseits sollen Datenformate entwickelt werden, die über lange Zeit gelesen werden können und von spezifischer Soft- und Hardware unabhängig sind.
- *Sicherung auf neuer Hardware*: Daten sollen sowohl ständig gesichert (Backup), als auch immer wieder auf neue Hardware kopiert werden.
- *Migration*: Datenträger und Daten sollen in regelmässigen Abständen migriert werden.
- *Emulation*: Das archetypische System kann virtuell nachgebaut werden (Emulation). Aber: "Keep in mind that in 2050 we will need to emulate not just the word processing program but elements of the operating system of yesteryear as well. And if emulation fails to work, we are in trouble."⁵⁸³
- *Digital archaeology oder digital forensics*: Mit physischen Geräten und spezieller Software werden von Datenträgern noch lesbare Daten im Source- oder Maschinencode ausgelesen.

Diese Methoden beeinträchtigen teilweise die Integrität eines digitalen Objekts. Unproblematisch ist die Verbesserung von Datenträgern oder Datenformaten. Eine verlustfreie Datenträgermigration beeinflusst die Integrität des Objekts nicht, sie kann aber dann zu Problemen führen, wenn der neue Datenträger nicht mehr ins System eingebunden werden kann. Bei einer Emulation wird zwar die Integrität des Informationssystems kompromittiert, das Objekt selbst aber bleibt integer. Es ist jedoch zu testen, ob die Emulation die Darstellung des Objekts verändert. Bei der Digitalarchäologie stellt sich die Frage, ob die noch lesbaren Daten des Objekts vollständig sind (unverletzte Integrität), oder ob Teile davon fehlen (verletzte Integrität). Da das Speichersystem des Datenträgers die Speicherorte zuweist, kann die Frage je nach Zustand des Datenträgers nicht beantwortet werden.

Viele Archiver versuchen zurzeit mit einer Kombination von Datenformat- und Datenträgermigration die dauerhafte Verfügbarkeit von digitalen Objekten sicherzustellen. Einerseits wird ein Objekt sowohl im Originalformat, als auch in einem langzeitarchivfähigen Dateiformat gespeichert. Andererseits werden die digitalen Objekte regelmässig auf einen neuen physischen Datenträger kopiert. Auf diese Prozesse wird unten vertiefend eingegangen.

⁵⁸² Siehe bspw.: Caroline R. Arms von 'The Library of Congress's Office of Strategic Initiatives', in: Cohen; Rosenzweig: Digital history, 2006, S. 244f.

⁵⁸³ Ebd., S. 245.

5.3.2. Konsistenz

Für die Darstellung eines digitalen Objekts in der vom Autor intendierten Weise müssen alle dafür nötigen Informationen vorhanden sein. In der Quellensicherung ist dafür zu sorgen, dass das Objekt konsistent gesichert wird, um eine quellenkritische Überprüfung überhaupt erst zu ermöglichen. Unter der *Konsistenz* eines digitalen Objekts wird verstanden, dass alle Informationen vorhanden sind, um es in der intendierten Weise und mit vollständigem Ausagewert darstellen zu können, was unter Umständen eine Rekontextualisierung voraussetzt.

Hierbei sind vor allem die *Abhängigkeiten* zu anderen Objekten zu überprüfen, was von vielen Nutzern oft unterschätzt wird. Bei einem Objekt, das als einzelne Datei vorliegt, wäre eine Sicherung als Einzelobjekt naheliegend. Jedoch ist vorab zu klären, ob in den Metadaten zusätzliche zu sichernde Informationen über die Datei vorliegen und ob es tatsächlich alle zur Darstellung nötigen Informationen enthält. Bei einem dynamischen Objekt wie einer GIF-Animation⁵⁸⁴ ist dies der Fall, hingegen liegen bei Film-DVDs⁵⁸⁵ die Audio- und Video-Spuren in getrennten und mehreren Dateien vor, die alle zu sichern sind. Je nach Fragestellung können die zusätzlich vorliegenden Informationen aber auch irrelevant sein und die Abhängigkeiten können aufgelöst werden, wenn beispielsweise nur die Audiodateien von einer DVD benötigt werden.

Wenn ein digitales Objekt einen Hyperlink zu einem anderen digitalen Objekt enthält, knüpfen sie eine "kulturelle Assoziation."⁵⁸⁶ Deshalb ist abzuklären und abzuwägen, ob die verlinkten Objekte als (inhaltliche) Abhängigkeit aufzufassen und ebenfalls zu sichern sind. Abhängigkeiten können auch in Bezug auf die Darstellung eines digitalen Objekts bestehen, wie beispielsweise die Darstellung einer HTML-Webseite durch *Cascading Style Sheets* (CSS). Innerhalb einer HTML-Datei wird für deren Darstellung auf eine CSS-Datei verwiesen, die entsprechende Parameter enthält. Der Webbrowser nutzt die Informationen aus beiden Dateien für die Darstellung, weshalb bei der Sicherung einer so strukturierten Webseite auch zwingend beide Dateien zu speichern sind. Der Quellcode der Webseite gibt Aufschluss darüber, wie sie konzipiert ist und welche Dateien für die intendierte Darstellung zu sichern sind. Besonders bei reborn-digitalen Objekten ist zu überprüfen, ob diese ohne die Basisdaten verwendet werden können. Dies kann beispielsweise der Fall sein, wenn ein Objekt Daten aus einer Datenbank nur aufbereitet und die daraus gewonnenen

⁵⁸⁴ GIF: Graphics Interchange Format; stellt ein singuläres Bild oder eine animierte Bildfolge dar.

⁵⁸⁵ DVD: Digital Versatile Disc, optischer Datenträger.

⁵⁸⁶ Bassler, Moritz: Die kulturpoetische Funktion und das Archiv: eine literaturwissenschaftliche Text-Kontext-Theorie, Tübingen 2005 (Studien und Texte zur Kulturgeschichte der deutschsprachigen Literatur), S. 315. Bassler bezieht sich zwar auf Texte, aber die Aussage kann auf Objekte verallgemeinert werden.

Informationen nicht in sich selbst speichert. Ohne die Datenbank wird das Objekt keine Informationen darstellen und ist somit (inhaltlich) wertlos.

Abhängigkeiten können auch durch eine automatisierte Verknüpfung von Inhalten entstehen. Beispielsweise versucht die Firma IBM die E-Mail mit ihrem Produkt *Verse*, das mit dem Slogan 'Mail that understands you' beworben wird, zu modernisieren.⁵⁸⁷ In *Verse* werden zu einer E-Mail auch die mit ihr zusammenhängenden Aufgaben, allfällig angeforderte Unterlagen oder ähnliche Elemente sowie die involvierten Personen angezeigt. Wird die E-Mail als potenzielle Quelle gesichert, ist zu entscheiden, welche der von der Software zusätzlich verknüpften Informationen und Objekte zu sichern sind. Der Nutzer hat diese algorithmisch generierten Informationen zur Verfügung gehabt und hat seine Tätigkeiten auf diese abgestützt. Aus einem relativ einfachen digitalen Objekt wird durch die automatisierte Vernetzung von Informationen ein komplexes Konstrukt, dessen Sicherung diverse Entscheide bezüglich Einbindung und Auslassung von Informationen erfordert. Zudem ist zu überprüfen, ob *Verse* die zusätzlichen Informationen vollständig darstellt.⁵⁸⁸

Oft befinden sich Dateien in einem Datencontainer, der als einzelne Datei in der Datenstruktur aufgeführt wird.⁵⁸⁹ Wird nur ein Teil der im Container befindlichen Dateien benötigt, sind deren Abhängigkeiten zu prüfen, worauf sie entweder herausgelöst werden können oder gemeinsam zu sichern sind. Es ist dabei zu entscheiden, ob das Containerformat weiterverwendet werden soll. Einige Formate werden zu einem späteren Zeitpunkt eventuell obsolet, weshalb die Daten extrahiert werden sollten. Containerformate wie das Audio Video Interleave (AVI)⁵⁹⁰ hingegen beinhalten verschiedene Datenströme für Video und Audio, die nicht separiert werden können beziehungsweise sollten. Solche standardisierten Formate wurden als Container konzipiert, um die Einheit der darin enthaltenen Daten zu garantieren, beispielsweise für die Synchronisierung von Video- und Audiospuren.⁵⁹¹ Ein unsachgemässes Separieren kann zu Daten- und/oder Informationsverlust führen. Hier ist eventuell die Migration in ein aktuelleres Format sinnvoller.⁵⁹²

Wenn sich digitale Objekte oder Informationen in einem bestimmten Kontext wie einem Online-Zeitungsartikel befinden, ist zu entscheiden, welche Elemente als *Zusatzinformationen*

⁵⁸⁷ Siehe Artikel: Weber, Volker: *Verse: Die Neuerfindung der E-Mail*, according to IBM, News, heise online, 18.11.2014, URL: <http://www.heise.de/newsticker/meldung/Verse-Die-Neuerfindung-der-E-Mail-according-to-IBM-2459147.html>, Stand: 19.11.2014. Produktwebseite unter URL: <https://www.ibm.com/ch-en/marketplace/business-email-platform>.

⁵⁸⁸ Siehe Kap. 5.1.4. Vollständigkeit.

⁵⁸⁹ Bspw. ZIP-Datei; Kompressions- und Archivierungsformat.

⁵⁹⁰ AVI: Dateiformat für audiovisuelle Inhalte.

⁵⁹¹ Bspw. MPEG-4 nach ISO/IEC 14496-10:2014; ISO Moving Picture Experts Group. URL: http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=66069.

⁵⁹² Zu Migration siehe Kap. 5.3.4. In welchem Format – Konversion und Migration.

mitzuspeichern sind. Beispielsweise ist vordergründig bei einem NZZ Online-Artikel nur der darin enthaltene Text von Interesse.



Abbildung 10: NZZ Online-Artikel (1), Hilfssicherung als Bild. Text des gekürzten Artikels (NZZ Online)⁵⁹³

Die auf der Website zusätzlich vorhandenen Informationen wie die von anderen Nutzern gelesene Artikel, weisen gleichzeitig auf aktuelle oder interessante Themen hin und betten so den Artikel wie bei physischen Zeitungen in einen gewissen Kontext ein.

Der Nutzen der Zusatzinformationen nimmt allerdings ab, je älter der Artikel ist. Zum gleichen, aber nun zwei Tage alten Artikel werden jetzt aktuelle Artikel empfohlen, womit der Artikel in einen anderen Kontext gestellt wird. Darauf sollte bei einer Sicherung mit zusätzlichen Informationen hingewiesen werden. Diese sich automatisch aktualisierende Verknüpfung ist gänzlich neu und bei physischen Objekten nicht möglich.

⁵⁹³ Für alle drei Bildzitate wurde die Website als Bild gesichert mit Screengrab! als Add-on für den Browser Firefox 40.0.3; URL: <https://addons.mozilla.org/de/firefox/addon/screengrab-fix-version/>). Bildzitate NZZ Online (Abb. 10-12) (URL: <http://www.nzz.ch/>): URL: <http://www.nzz.ch/international/asien-und-pazifik/die-hydra-des-separatismus-1.18508554#tab-Empfohlen>. Abbildungen 10 und 11: 24.03.2015, 15.45 Uhr; Abbildung 12: 26.03.2015, 15.34 Uhr.

Jobs | Immo | Trauer | eBalance | Services | Shops Abo | Angebote | Hilfe | [Login](#) | [Suchen](#)

Neue Zürcher Zeitung

Zürich 14° Dienstag 24. März 2015 – E-Paper / Webpaper

International | Wirtschaft | Finanzen | Schweiz | Zürich | Meinung | Sport | Feuilleton | Wissenschaft | Panorama | Lebensart | mehr ▾

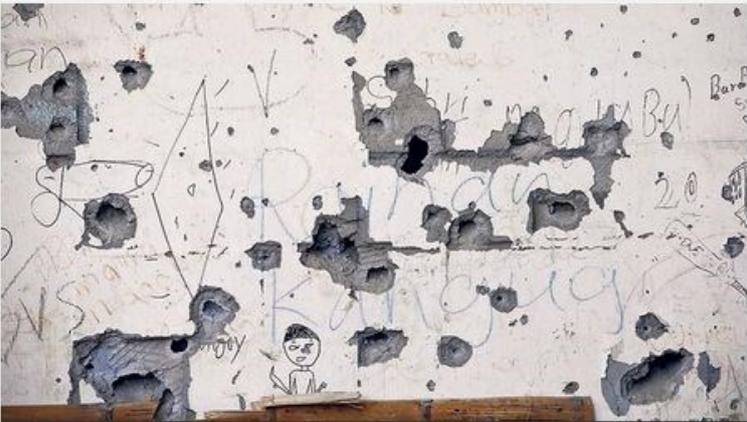
ASIEN & PAZIFIK

Unruhiges Mindanao

Die Hydra des Separatismus

Manfred Rist, Lupon 24.3.2015, 05:30 Uhr

Tweet ✉ 📧 📧 📧



Der Kampf gegen Rebellenruppen auf den Philippinen hinterlässt Spuren. Einschusslöcher in der Wand einer öffentlichen Schule in Zamboanga. (Bild: Dennis M. Sabangan / EPA)

Seit Jahrzehnten kämpfen auf Mindanao Rebellengruppen für mehr Autonomie. Nun scheint ein Autonomieabkommen mit einer Gruppe in Griffnähe. Längst nicht alle Separatisten sind damit einverstanden – sie formieren sich bereits neu.

Wenn Hussein Muñoz auf das Massaker von Mamasapono zu sprechen kommt, gerät sein Bleistift in wilde Fahrt. Mit hektischen Strichen zeichnet der stellvertretende Stabschef der Moro Islamic Liberation Front (MILF) den Hinterhalt nach, der Ende Januar 44 Special Forces der philippinischen Polizei

LESERTREND

GELESEN EMPFOHLEN KOMMENTIERT

ist das Leben nüchtern überhaupt auszuhalten?

Zerstörerische Missverständnisse

Einen Neuanfang wagen

Offene Machfrage bei Silka trotz Gerichtsentscheid

Mit Waffen gegen Reformen

Günther Jauch sollte entlassen werden

Eine Weltordnung ohne Europa

Teil Aviv wird zur bedrohten Enklave

CS hätte Konto nicht einfrieren dürfen

Euro taucht im Währungskrieg zum Dollar ab

[Finanzen](#)

zurück [jobs.NZZ.ch](#) weiter

ZÜRICH INSURANCE COMPANY LTD
Senior Actuary Reserving GI 

PERSONAL SIGMA LUZERN AG
ICT SYSTEM SENIOR ENGINEER LINUX-
Luzerner Kantonalbank

ENGINEERING MANAGEMENT SELECTION ...
IT Operations Manager

ADLATUS
Unternehmerpersönlichkeit als Mitglied
der Geschäftsleitung

ELIGENDO AG
CFO / MITGLIED DER GL (M/ VV)

Abbildung 11: NZZ Online-Artikel (2), Hilfssicherung als Bild. Text des gekürzten Artikels mit aktuellen Zusatzinformationen im Webseitenelement 'Lesertrend – Gelesen' (NZZ Online)

Mit extern eingebundenen Inhalten (integrierender Quellentypus) können sich die Beziehungen unter den einzelnen Elementen ebenfalls verändern. Zwar werden solche Elemente vor allem für Werbung eingesetzt, aber es existieren auch Webseiten, die selbst keinen Inhalt anbieten und ausschliesslich externe Daten darstellen.

Jobs | Immo | Trauer | eBalance | Services | Shops Abo | Angebote | Hilfe | [Login](#) | [Suchen](#)



Neue Zürcher Zeitung

Zürich 8° Donnerstag 26. März 2015 – E-Paper / Webpaper

International | Wirtschaft | Finanzen | Schweiz | Zürich | Meinung | Sport | Feuilleton | Wissenschaft | Panorama | Lebensart | [mehr](#) ▾

ASIEN & PAZIFIK

Unruhiges Mindanao

Die Hydra des Separatismus

Manfred Rist, Lupon 24.3.2015, 05:30 Uhr

[f](#) [t](#) [g+](#)
[✉](#) [📄](#) [📱](#)



Der Kampf gegen Rebellen Gruppen auf den Philippinen hinterlässt Spuren. Einschusslöcher in der Wand einer öffentlichen Schule in Zamboanga. (Bild: Dennis M. Sabangan / EPA)

Seit Jahrzehnten kämpfen auf Mindanao Rebellen Gruppen für mehr Autonomie. Nun scheint ein Autonomieabkommen mit einer Gruppe in Griffnähe. Längst nicht alle Separatisten sind damit einverstanden – sie formieren sich bereits neu.

Wenn Hussein Muñoz auf das Massaker von Mamasapono zu sprechen kommt, gerät sein Bleistift in wilde Fahrt. Mit hektischen Strichen zeichnet der stellvertretende Stabschef der Moro Islamic Liberation Front (MILF) den Hinterhalt nach, der Ende Januar 44 Special Forces der philippinischen Polizei und 18 ihrer Kämpfer das Leben kostete. Die MILF habe die ortsfremden Einheiten, die einem international gesuchten Terroristen auf der Spur waren, bis zum Morgengrauen in einem Kornfeld bekämpft und von ihrem nahe gelegenen Lager ferngehalten. Auf deren Rückzug seien sie dann von Einheiten der

LESERTREND

GELESEN
EMPFOHLEN

- Grandiose Flut am Mont-Saint-Michel
[Alltagsgeschichten](#)
- Ist das Leben nüchtern überhaupt auszuhalten?
- Günther Jauch sollte entlassen werden
- «Zunehmendes Demokratiedefizit und sinkendes Engagement»
- Ein Naturtalent
- Wachsende Kluft zwischen den Kontinenten
- Eine Weltordnung ohne Europa
- Euro taucht im Währungskrieg zum Dollar ab
- A World Order without Europe?
- Deutscher Airbus in den französischen Alpen abgestürzt

[← zurück](#)
[jobs.NZZ.ch](#)
[weiter →](#)

- [KANTONSSPITAL BASELSTADT](#)
Oberärztin/-arzt Frauenklinik


- [ZURICH INSURANCE COMPANY LTD](#)
Senior Actuary Reserving GI


- [PERSONAL SIGMA LUZERN AG](#)
ICT SYSTEM SENIOR ENGINEER LINUX-
Luzerner Kantonalbank
- [ENGINEERING MANAGEMENT SELECTION...](#)
IT Operations Manager
- [ADLATUS](#)
Unternehmerpersönlichkeit als Mitglied
der Geschäftsleitung
- [ELIGENDO AG](#)
CFO / MITGLIED DER GL (M/ W)

Abbildung 12: NZZ Online-Artikel (3), Hilfssicherung als Bild. Text des gekürzten Artikels nach zwei Tagen mit aktuellen Zusatzinformationen im Webseitenelement 'Lesertrend – Empfohlen' (NZZ Online)

Ein zur Visualisierung von Standortinformationen gerne verwendetes externes Element ist Google Maps⁵⁹⁴, das auch die Universität Basel unter Kontakt – Standorte anbietet:

⁵⁹⁴ URL: <https://www.google.ch/maps>. Anleitung zur Einbettung siehe URL: <https://support.google.com/maps/answer/3544418>.

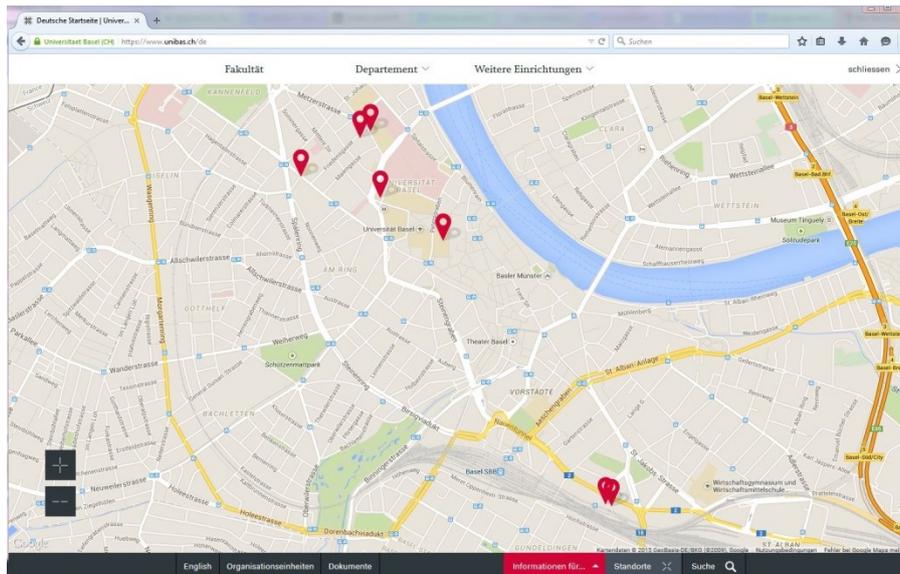


Abbildung 13: Webseite der Universität Basel, Standorte (Universität Basel)⁵⁹⁵

Die Karte ist vollständig in der Website der Universität integriert, was für den Nutzer nicht sichtbar ist, denn die im Browser angezeigte URL lautet weiterhin <https://www.unibas.ch/de>. Welche Elemente originär auf dieser Webseite publiziert und welche nur eingebunden sind, kann nur im Quelltext der Webseite nachgewiesen werden.⁵⁹⁶ Im in diesem Beispiel verwendeten Quelltext (Auszug) wird mit der Zeile

URL: `script src="https://maps.googleapis.com/maps/api/js?key=AIzaSyCWrNkhNj4V6r9hYrUr6lxG8MgzgNx7xQc&sensor=false" type="text/javascript"`

auf die externe digitale Quelle verwiesen.⁵⁹⁷

```
<!DOCTYPE html>
<html class="js canvas no-touch geolocation multiplebgs backgroundsize boxshadow opacity cssanimations cssecolumns csstransforms csstransforms3d csstransitions fontface video audio localstorage sessionstorage inlinesvg boxsizing texttrackapi track display-table cubicbezierrange mediaqueries no-details no-mobile advancedcssanimation greater-than-1280 js-on polyfill-important swf em-increased-4 em-increased-3 em-increased-2 em-increased-1" lang="de" xml:lang="de" style="">
<head>
  <link type="text/css" rel="stylesheet" href="https://fonts.googleapis.com/css?family=Roboto:300,400,500,700">
  <style type="text/css">
  <style type="text/css">
  <style type="text/css">
  <link rel="stylesheet" href="/resources/templating-kit/themes/unibas/css/main_icons.css" media="all">
  <script type="text/javascript" defer="" async="" src="https://piwik.unibas.ch/piwik.js" style="">
  <script async="" src="https://www.unibas.ch/resources/templating-kit/themes/unibas/1.0.4074/js/shims/combo/23.js">
  <script async="" src="https://www.unibas.ch/resources/templating-kit/themes/unibas/1.0.4074/js/shims/.../_enhanced-modules.js">
  <script async="" src="https://www.unibas.ch/resources/templating-kit/themes/unibas/1.0.4074/js/shims/mediaelement-jaris.js">
  <script async="" src="https://www.unibas.ch/resources/templating-kit/themes/unibas/1.0.4074/js/shims/dom-extend.js">
  <script async="" src="https://www.unibas.ch/resources/templating-kit/themes/unibas/1.0.4074/js/shims/track-ui.js">
  <script async="" src="https://www.unibas.ch/resources/templating-kit/themes/unibas/1.0.4074/js/shims/range-ui.js">
  <script src="https://maps.googleapis.com/maps/api/js?key=AIzaSyCWrNkhNj4V6r9hYrUr6lxG8MgzgNx7xQc&sensor=false" type="text/javascript">
  <script src="https://maps.gstatic.com/maps-api-v3/api/js/20/4/intl/de_ALL/main.js">
  <link rel="stylesheet" href="https://www.unibas.ch/resources/templating-kit/themes/unibas/1.0.4074/js/shims/styles/shim.css">
```

Abbildung 14: Sourcecode (Auszug) der Webseite der Universität Basel, Einbettung Google Maps markiert (Universität Basel)

⁵⁹⁵ Bildzitat Universität Basel (URL: <https://www.unibas.ch/de>). Kartenausschnitt als Bildzitat von Google Maps/Google Earth (Google LLC, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA) unter (teilweiser) Verwendung von Bildern und (Geo-)Informationen von © Swisstopo, © BSF Swissphoto und Swiss Postal Data.

⁵⁹⁶ Bei Google Maps ist die externe Einbindung durch den Bekanntheitsgrad des Produkts offensichtlicher als bei anderen Inhalten.

⁵⁹⁷ Quelltext untersucht mit Mozilla Firefox 40.0.3 und Add-on Firebug 2.0.12. Zeilenumbrüche sind zu entfernen. Quelle: Webseite der Universität Basel (URL: <https://www.unibas.ch/de>).

Weil die aus externen Quellen eingebundenen Elemente einer Webseite nicht offensichtlich sind, sollte vor der Sicherung einer Website der Quelltext überprüft werden.

Eingebundene Objekte können im Gegensatz zu Hyperlinks stets als Aneignung des externen Inhalts und somit als Teil der dargestellten Webseite angesehen werden, weshalb aus pragmatischer und praktischer Sicht weder die Zitierweise anzupassen, noch zwingend eine Überprüfung auf der Quellseite durchzuführen ist.⁵⁹⁸ Sollte aber ein Element klar als externes Objekt erkennbar sein, ist eine Überprüfung auf dessen Quellseite dennoch zu empfehlen.

Bei digitalen, dynamischen Objekten ist der Nachweis der Konsistenz problematisch, weil sich das Objekt schnell und nicht nachvollziehbar verändern lässt oder sich ständig verändert. Für die Wiederverwendung eines Objekts müssen alle dafür nötigen Objekte vorliegen und gemeinsam gesichert werden. Ist jedoch nur die Darstellung des Objekts vorhanden, wie beispielsweise eine dynamische Graphik auf einer Webseite, dann ist eine Sicherung in der vom Autor intendierten Form kaum möglich. Die Dynamik basiert auf einem Algorithmus, der durch ein Bildschirmbild oder durch *Abfilmen*⁵⁹⁹ weder nachvollzogen noch adäquat wiedergegeben werden kann. Deshalb ist auf einer Website stets im Quellcode zu überprüfen, wie das dynamische Objekt eingebunden und wie es gespeichert oder sein Inhalt exportiert werden kann.

5.3.3. Sekundärobjekt und Surrogat

Die Sicherung eines physischen oder digitalen Objekts bedeutet einen Informationsverlust. Bei digitalen Objekten entsteht durch die Sicherung sogar ein zusätzliches, reborn-digitales Objekt, das so bisher nicht existiert hat. Der Nutzer hat insbesondere bei komplexen und hybriden digitalen Objekten zu entscheiden, in welcher Form das Objekt gesichert werden soll, damit es möglichst aussagekräftig und nachvollziehbar bleibt.

Grundsätzlich sollte das ursprüngliche digitale Objekt in der ursprünglich dargestellten Form und Funktionalität gesichert werden. In vielen Fällen ist ein Zugriff auf die relevanten Daten nicht möglich und es sind technische Lösungen einzusetzen, um das Objekt trotzdem in einer möglichst äquivalenten Form zu sichern. Der Prozess der Sicherung eines oder mehrerer Objekte mit geeigneten Mitteln zur Erstellung eines (gemäß den Anforderungen der Langzeitarchivierung dauerhaft) weiterverwendbaren Objekts wird als *digitale Hilfssicherung* bezeichnet. Dadurch entsteht ein neues und einzigartiges digitales Objekt, das sich vom

⁵⁹⁸ Nur wenn der Autor klar auf eine Distanzierung vom gezeigten Inhalt hinweist, kann *eventuell* angenommen werden, dass er sich diesen nicht aneignen will.

⁵⁹⁹ Abfilmen: Programme (Bildschirm-Recorder) wie CamStudio (URL: <http://camstudio.org/>) können Bildschirm-/Desktopinhalte audiovisuell festhalten, inkl. Videos, Mauszeigerbewegungen etc.

ursprünglichen Objekt teilweise massiv unterscheidet.⁶⁰⁰ Das ursprüngliche Objekt kann im intellektuellen Sinn als Primärobjekt, das davon durch eine Hilfssicherung abgeleitete Objekt als reborn-digitales *Sekundärobjekt* oder *Surrogat* bezeichnet werden.⁶⁰¹

Die mit einer *digitalen Hilfssicherung* generierten Objekte werden (im intellektuellen Sinn) als neue, einzigartige, reborn-digitale *Sekundärobjekte* oder *Surrogate* bezeichnet.

Die Hilfssicherung ist besonders bei Objekten im Internet problematisch. Da im Voraus nicht bekannt ist, wann ein Objekt aktualisiert wird, kann es unter Umständen während der Sicherung verändert werden.⁶⁰² Vor allem komplexe Objekte wie ein Webauftritt, das viele unterschiedliche Elemente enthält, benötigen zur Sicherung eine gewisse Zeit, während der ein Administrator, Editor oder auch Nutzer Änderungen vornehmen kann. Es kann deshalb nicht sichergestellt werden, dass komplexe Objekte nach Abschluss der Sicherung so vorliegen, wie sie zu Beginn waren. Für Niels Brügger, der sich in seinen Ausführungen allerdings nur auf Webseiten bezieht, ergeben sich unter diesen Voraussetzungen zwei Konsequenzen:⁶⁰³

1. Aufgrund der asynchronen Beziehung zwischen Aktualisierung und Sicherung geht immer etwas verloren.
2. Es besteht nicht nur die Gefahr, dass etwas fehlt, sondern auch, dass etwas gesichert wird, das so nie vorhanden war – etwas, das sich vom tatsächlich vorhandenen Gewesenen unterscheidet.

Zu Brüggers erster Konsequenz ist anzumerken, dass nicht zwingend etwas verloren geht, wenn der Sicherungsvorgang schnell genug oder in einer kontrollierten Umgebung ablaufen kann. Beispielsweise kann ein einfaches oder einzelnes Objekt sehr schnell gesichert und der Zugang und Zugriff auf einen Webauftritt kann gesperrt werden. Komplexe digitale Objekte sind während des Sicherungsprozesses vor Manipulationen zu schützen, andernfalls ist mit den erwähnten Konsequenzen zu rechnen. Dies betrifft besonders Objekte in Netzwerken, in denen Personen und andere Prozesse auf das Objekt zugreifen können.

Brewster Kahle, der Gründer des Internet Archive, ist sich der von Brügger aufgezeigten Problematik ebenfalls bewusst. Auf die Frage, ob eine Sicherung im Internet Archive authentisiert werden kann, meint er:

⁶⁰⁰ Vgl.: Brügger, Niels: Web Archiving, Web History, and the Web in Digital Humanities, Präsentation, Digital Humanities Luxembourg - Symposium, Luxembourg 21.03.2012, Fol. 7.

⁶⁰¹ Surrogat: Stoff, Mittel o. Ä. als behelfsmäßiger, nicht vollwertiger Ersatz. Wermke u. a. (Hg.): Duden - das Fremdwörterbuch, 2001.

⁶⁰² Siehe dazu: Brügger: Web Archiving, 2012, Fol. 24.

⁶⁰³ Siehe: Ebd., Fol. 16.

*"We can say, 'This is what we know. This is what our records say. This is how we received this information, from which apparent Web site, at this IP address.' But to actually say that this happened in the past is something that we can't say, in an ontological way."*⁶⁰⁴

Insbesondere Aussagen zur tatsächlichen *Funktionalität* eines Objekts sind schwierig, denn das Surrogat kann in vielen Fällen nicht mehr so genutzt werden wie das Primärobjekt. Auf whatisdigitalhumanities.com ist weder der Zugriff auf die Datenbank möglich noch kann der ein Nutzer die darin enthaltenen Definitionen von Digital Humanities exportieren. Die Definitionen könnten nur abgeschrieben werden⁶⁰⁵, womit aber die Funktionalität der Webseite später nicht nachvollziehbar ist, selbst wenn diese in den Metadaten möglichst genau beschrieben ist. Zudem ist auch die Vollständigkeit nicht garantiert, weil Definitionen während des Vorgangs ergänzt oder gelöscht werden können, oder weil gewisse Definitionen aufgrund der zufälligen Reihenfolge während des Prozesses nicht angezeigt werden. Mit einem Bildschirm-Recorder⁶⁰⁶, der das Aufnehmen von Ereignissen auf dem Desktop des Betriebssystems ermöglicht, könnten wenigstens die Darstellung und der Inhalt des digitalen, dynamischen Objekts dokumentiert werden. Dieses Surrogat könnte zumindest einen Teil der Funktionalität zwar besser nachvollziehbar, wenn auch nicht ganz erfahrbar machen.

Einfacher ist die (Hilfs-)Sicherung von einzelnen Objekten. Die oben verwendeten Bilder des NZZ-Artikels wurden von der entsprechenden NZZ-Website direkt als Bild abgespeichert.⁶⁰⁷ Anschliessend wurde der dargestellte Ausschnitt mit einem Bildbearbeitungsprogramm⁶⁰⁸ ausgewählt und in den Text eingefügt. Das Ziel war die Veranschaulichung der Konsistenz, weshalb auf Funktionalitäten wie die Hypertextualität oder Vollständigkeit verzichtet werden konnte. Was oben dargestellt ist, war so aber nie auf der Seite der NZZ vorhanden.

Noch etwas komplexer wird eine Sicherung, wenn es sich um *hybride Objekte* wie Laborberichte handelt, die aus voneinander abhängigen physischen und digitalen Teilen bestehen. Bei solchen hybrid angefertigten Objekten mit jeweils unterschiedlichen Datenbeständen ist zusätzlich zur Sicherung der digitalen Elemente des Objekts zu entscheiden, ob und in welcher Form die relevanten physischen Elemente digitalisiert und gesichert werden sollen.⁶⁰⁹ Falls sie digitalisiert werden, sind diese als Digitalisate zu deklarieren und zu dokumentieren.

⁶⁰⁴ Brewster Kahle im Interview mit: Lepore, Jill: The Cobweb. Can the Internet be archived?, The New Yorker, 19.01.2015, URL: <http://www.newyorker.com/magazine/2015/01/26/cobweb>, Stand: 28.01.2015.

⁶⁰⁵ Handschriftlich, Copy-Paste, Bildschirmbild etc.

⁶⁰⁶ Bspw. 'CamStudio (URL: <http://camstudio.org/>).

⁶⁰⁷ Zur Sicherung des Bildschirmausschnittes (Screenshot) wurde das Add-on 'Screengrab!' unter Firefox verwendet. Alternative: GadwinPrintScreen (URL: <http://www.gadwin.com/printscreens/>).

⁶⁰⁸ Bspw. IrfanView (URL: <http://www.irfanview.com/>).

⁶⁰⁹ Siehe dazu: Potthoff, Jan; Rieger, Sebastian; Johannes, Paul C. u. a.: Elektronisches Laborbuch: Beweiserhaltung und Langzeitarchivierung in der Forschung, in: Digitale Wissenschaft: Stand und

Bei jeder Sicherung hat der Nutzer zu entscheiden, in welcher Form das Primärobjekt gesichert werden soll. Er hat diejenigen technischen Methoden zu verwenden, mit denen möglichst aussagekräftige und nachvollziehbare Surrogate erstellt werden können. Ungemein wichtig ist die Dokumentation der gewählten Form in den Metadaten, damit nachvollzogen werden kann, was ursprünglich vorhanden war und aus welchen Gründen die Form des Surrogats gewählt wurde.

5.3.4. In welchem Format – Konversion und Migration

Digitale Objekte können nur dann dauerhaft archiviert werden, wenn sie in einem darauf ausgelegten Datenformat vorliegen und wenn sie regelmässig auf neue Datenträger und in aktuelle Datenformate migriert werden. Nur damit ist es möglich, dass digitale Objekte auch nach längerer Zeit noch nutzbar bleiben.

Konzepte für die Langzeitarchivierung sind bisher vor allem von den Archivwissenschaften für professionell geführte öffentliche oder private Archive erarbeitet worden. Um den Zugriff und vor allem die Nutzung eines im digitalen Langzeitarchiv gesicherten digitalen Objekts über eine möglichst lange Zeit hinweg sicherstellen zu können, werden digitale Objekte im Ingest-Prozess, in dem neue Objekte entgegengenommen und ins Archiv eingebunden werden, unter anderem in ein geeignetes Datenformat konvertiert. Erst danach wird das Objekt mit den zugehörigen Metadaten in das digitale Langzeitarchiv aufgenommen und in der Verwaltungssoftware registriert.

Solche professionellen Objektsicherungsprozesse stehen allerdings zurzeit nur Behörden, grösseren Forschungsinstitutionen oder für grössere Forschungsprojekten zur Verfügung. Individualforscher⁶¹⁰ können sich meist nur auf ihre informationstechnisches Wissen und praktische Erfahrungen verlassen und darauf hoffen, dass ihre Sicherungen in ein paar Jahren noch gelesen werden können. Deshalb sollten auch Individualforscher Dateiformate verwenden, welche für die Langzeitarchivierung geeignet sind.

Quelloffene⁶¹¹ Dateiformate werden dabei proprietären Formaten vorgezogen, denn letztere können durch die Eigner undokumentiert weiterentwickelt oder eingestellt werden. Falls der Eigner den strukturellen Aufbau des Formats nicht bekannt gibt, können proprietäre Formate

Entwicklung digital vernetzter Forschung in Deutschland ; 20./21. September 2010, Köln ; Beiträge der Tagung, Köln 2011, S. 149-156..

⁶¹⁰ Forschende ohne Bindung an ein Projekt oder Institut mit entsprechender Infrastruktur.

⁶¹¹ Quelloffen: Der Quellcode des Formats ist frei zugänglich und kann gelesen und verändert werden. Proprietär: Der Quellcode der das Format erzeugenden Software ist nicht bekannt (meist aus marktwirtschaftlichen Gründen). Beispiele offener Standards sind für Graphiken 'Tagged Image File Format' TIFF oder 'Portable Network Graphics' PNG, für Dokumente 'Extensible Markup Language' XML oder 'Open Document Format for Office Applications' ODF und für Musik 'Free Lossless Audio Codec' FLAC oder 'Musepack' MPC.

durch andere Software nicht oder durch Reverse Engineering⁶¹² nur annähernd dargestellt werden. Bei quelloffenen Formaten hingegen kann die Syntax eingesehen werden und zudem kann unabhängig vom Urheber weitere Software für die Darstellung des Formats verwendet werden. Da die Software durch die Interpretation des Formats für die Darstellung des digitalen Objekts besorgt ist, kann dieses selbst unverändert bleiben. Nur die Software ist an veränderte technische Rahmenbedingungen anzupassen – theoretisch (!), denn auch Veränderungen an der interpretierenden Software beeinflussen die Darstellung und Bearbeitung eines Objekts, weshalb die korrekte Interpretation ausgiebig zu testen ist.

In Zusammenarbeit von Archivwissenschaften und Softwareherstellern wurden Formate entwickelt, die für die Langzeitarchivierung ausgelegt sind. Für Schriftdokumente wird das Containerformat *PDF/A*⁶¹³ eingesetzt, das festlegt, welche Elemente (bspw. der verwendete Schriftsatz) vollständig im Dokument enthalten sein müssen, welche optional sind und welche nicht vorkommen dürfen. Damit soll sichergestellt werden, dass ein digitales Objekt in diesem Format möglichst lange gelesen werden kann. PDF/A ist als ISO-Standard definiert und wurde mit den Versionen PDF/A-2 und PDF/A-3 ergänzt, die zusätzliche Elemente integrieren können. Im Moment wird für die Langzeitarchivierung vor allem der Standard PDF/A-2 empfohlen.⁶¹⁴

Containerformate haben auch für archivische Belange ihre Berechtigung, solange sie dafür konzipiert und standardisiert wurden. Allerdings ist zu beachten, dass auch wenn der Container für eine dauerhafte Archivierung tauglich ist, es die darin enthaltenen Objekte allenfalls nicht sind. Beispielsweise wurde mit ISO 9660:1988 ein Standard für ein Dateisystem von optischen Datenträgern definiert.⁶¹⁵ Darin wurde 1988 festgelegt, wie auf optischen Datenträgern vorhandene Daten in einer Containerdatei im ISO-Format auf eine Festplatte kopiert werden können. Die ISO-Datei kann mit entsprechender Software geöffnet und die darin enthaltenen Dateien können wie ein Datenträger in das Dateisystem eingebunden werden. Mit diesem Verfahren wurden viele Datenträger erfolgreich archiviert, um eventuellem Datenverlust durch *Datenträgerverlust* vorzubeugen. Wer sich aber erst nach der Revision des ISO-Formats 2013⁶¹⁶ wieder um die Dateien kümmerte, die Ende der 1980er und zu Beginn der 1990er Jahre angefertigt wurden, wird die darin enthaltenen Objekte kaum mehr lesen können, weil deren Dateiformate obsolet wurden. Auch ein

⁶¹² Reverse Engineering: Verfahren, in dem versucht wird, aus dem Resultat das Quellmaterial zu generieren. Fischer; Hofer: Lexikon der Informatik, 2011.

⁶¹³ PDF/A: Portable Document Format für Langzeitarchivierung (A); ISO 19005-1:2005 (URL: http://www.iso.org/iso/catalogue_detail?csnumber=38920).

⁶¹⁴ URL: http://www.iso.org/iso/catalogue_detail.htm?csnumber=57229. Version 3 wird kritisiert, weil darin beliebige Inhalte in Formaten gespeichert werden können, die den Ansprüchen der Langzeitarchivierung nicht genügen.

⁶¹⁵ URL: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=17505.

⁶¹⁶ URL: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=61625.

standardisiertes Containerformat kann nicht garantieren, dass die darin enthaltenen Daten langfristig verwendet werden können.

Da digitale Objekte durch ihr Format definiert sind, ist es auch nicht sinnvoll, nur den Binärcode eines Objekts zu sichern. Dieser muss von einem digitalen Informationssystem so decodiert werden, dass das Objekt für einen Nutzer wieder lesbar ist. Für eine erfolgreiche Interpretation eines Bitcodes werden aber unter anderem im Format definierte Spezifikationen für Hard- und Software vorausgesetzt, damit ein Nutzer die Daten in der ursprünglich erstellten Weise betrachten kann. Sind diese Informationen nicht vorhanden, kann auch der Binärcode nicht interpretiert werden, wie bei der Viking-Sonde der NASA von 1976. Deren Binärcode ist nicht mehr decodierbar, weil die dafür nötige Hard- und Software sowie Dokumentation fehlt und deshalb die Syntax nicht wieder rekonstruiert werden kann.⁶¹⁷

Das für die Sicherung eines Objekts geeignete Zielformat wird durch die Objektart bestimmt. Eine Übersicht über die empfohlenen, archivtauglichen Formate bieten beispielsweise die *Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (KOST)* im *Katalog archivischer Dateiformate*, das *Deutsche Bundesamt für Sicherheit in der Informationstechnik (BSI)*, das *BAR* und diverse Literatur.⁶¹⁸ Im besten Fall wird ein Objekt bereits in einem archivtauglichen Format erstellt, denn bei der Konversion gehen in den allermeisten Fällen Informationen verloren. Es sollte deshalb möglichst früh im Erstellungsprozess ein geeignetes, langzeitarchivtaugliches Format eingesetzt werden.⁶¹⁹

Mit der Wahl des langzeitarchivfähigen Dateiformats und der Sicherung des Objekts in diesem ist jedoch keine dauerhafte Nutzung garantiert. Ein digitales Objekt unterliegt in zweierlei Hinsicht einem Alterungsprozess: der Datenträger, auf dem es sich befindet, altert physisch und das Datenformat, in dem es vorliegt, wird weiterentwickelt oder obsolet. Um dem entgegenzuwirken, kann das Objekt in einer sogenannten Migration auf aktuelle Datenträger und/oder Datenformate migriert werden.

Bei der *Datenträgermigration* wird der Bitcode eines digitalen Objekts von einem Datenträger auf einen anderen Datenträger beliebigen Typs kopiert.⁶²⁰ Die dem Objekt zugehörigen Metadaten sollten mit den Informationen dieser Migration ergänzt werden und die Metadaten selbst sind ebenfalls zu migrieren. Das Objekt behält dabei seine Datenintegrität, weil an ihm

⁶¹⁷ Siehe: Rosenthaler: *Archivierung im digitalen Zeitalter*, 2007, S. 33.

⁶¹⁸ KOST (URL: <http://www.kost-ceco.ch/wiki/whelp/KaD/index.php>), BSI (URL: https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/Inhalt/_content/m/m04/m04170.html), BAR (URL: <http://www.bar.admin.ch/dienstleistungen/00895/00897/>), Literatur bspw.: Gladney: *Preserving Digital Information*, 2007, S. 139-152.

⁶¹⁹ Besonders aus der Sicht von Archivaren und aus eigenen Erfahrungen als wissenschaftlicher Mitarbeiter im Staatsarchiv Solothurn.

⁶²⁰ Siehe bspw. Margulies: *Digitale Daten als Quelle der Geschichtswissenschaft*, 2009, S. 345-347.

nichts geändert wird. Dies ist aber über jeden Migrationsschritt hinweg zu überprüfen, beispielsweise durch eine Prüfsumme.

Die begrenzte Haltbarkeit und Lesbarkeit eines technischen Mediums zwingt in der Langzeitarchivierung zu regelmässigen Migrationsprozessen. Mit der Entwicklung von *permanenten* Datenträgern wie der Rosetta-Disk⁶²¹, der Speicherung in DNA-Molekülen⁶²² oder der M-Disc⁶²³ sollen die zeitlichen Abstände zwischen den nötigen Migrationen verlängert werden oder Migrationen grundsätzlich unnötig werden lassen. Solche Datenträger haben vor allem drei Bedingungen zu erfüllen:⁶²⁴ Erstens muss der Datenträger selbst sowie die in einem Aufzeichnungsverfahren darauf angebrachten Markierungen dauerhaft und stabil sein. Zweitens muss sowohl das physische Format des Datenträgers als auch das Aufzeichnungsverfahren so einfach und allgemein gehalten sein, dass für die Wiedergabe und -nutzung langfristig Lesegeräte zur Verfügung stehen. Drittens muss auf dem Datenträger selbst ersichtlich sein, dass er ein Datenträger für digitale Daten ist. Angaben zum Aufzeichnungsprozess und wie es wiedergegeben werden kann, sind ebenfalls notwendig.

Zurzeit sind solche Datenträger problematisch, denn sie sind weder wirtschaftlich noch in grösserer Masse einsetzbar, die tatsächliche Langlebigkeit der Datenträger ist nicht bestätigt, die Verfügbarkeit von Lesegeräten ist nicht gewährleistet und mit der Materialisierung geht die Digitalität des Objekts verloren. Digitalität bedeutet, dass die Daten in einer bestimmten Aneinanderreihung der Ziffern 0 und 1 gespeichert sind, die durch ein digitales Informationssystem interpretiert werden. Die mikroskopische Abbildung einer Textseite wie im Rosetta-Projekt ist keine digitale Speicherung, weil sie mit einem entsprechenden physischen Mittel stets gelesen werden kann. Das elektronische Mittel aber unterliegt der Obsoleszenz, auch wenn möglichst einfache Daten- und Dateiformate verwendet werden. Zudem entspricht die Verwendung möglichst einfacher Datenformate nicht den heutigen Gegebenheiten, denn dynamische Objekte oder Datenbanken können damit beispielsweise nicht gesichert werden.

⁶²¹ Das Rosetta-Projekt hat ein Medium produziert, das 10'000 Jahre lesbar sein soll. Siehe URL: <http://rosettaproject.org>.

⁶²² DNA: Desoxyribonukleinsäure. Siehe: Grass, Robert N.; Heckel, Reinhard; Puddu, Michela u. a.: Robust Chemical Preservation of Digital Information on DNA in Silica with Error-Correcting Codes, in: Angewandte Chemie International Edition 54 (8), 16.02.2015, S. 2552-2555. DOI: 10.1002/anie.201411378.

⁶²³ M-Disc: Millennial Disc; widerstandsfähiger, optischer Datenspeicher, der Marktreife erlangt hat (URL: <http://www.mdisc.com>). Siehe Forschungen von Barry M. Lunt und Matthew R. Linford, Brigham Young University (URL: <http://byu.edu/>); Lunt, Barry M.; Davis, Robert; Hansen, Douglas u. a.: Permanent Digital Data Storage: A Materials Approach, in: iPRES-2013 – 10th International Conference on Preservation of Digital Objects, Lissabon 2013. URL: http://www.purl.pt/24107/1/iPres2013_PDF/Permanent%20Digital%20Data%20Storage%20A%20Materials%20Approach.pdf, Stand: 05.12.2015.

⁶²⁴ Nach: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 487 f.

Die *Datenformatmigration* ist eine Transformation des digitalen Objekts in ein anderes Datenformat, es entsteht ein neues, reborn-digitales Objekt. Der Inhalt des Objekts wird dabei von seiner Form, dem Dateiformat getrennt und durch einen nicht einsehbaren, elektronisch-algorithmischen Prozess in einer anderen Form gespeichert. Dabei ändert ein sogenannter *Konverter*⁶²⁵ den Bitcode des Objekts und die Datenintegrität geht verloren. Es ist dabei unerheblich, ob nach der Migration das *Primärobjekt* wiederhergestellt werden kann oder nicht (reversible oder irreversible Migration)⁶²⁶, denn bei jeder Transformation durch eine Formatänderung leidet die Vertrauenswürdigkeit des Objekts. Hilfestellungen für die Auswahl eines geeigneten Archivformats bieten sogenannte *File Format Registries*, in denen Informationen zu den einzelnen Formaten gesammelt werden, um die Programmierung von Konvertern zu erleichtern.⁶²⁷

Ob sich das digitale Objekt nach der Datenformatmigration noch in der intendierten Art und Weise darstellen lässt, ist jeweils zu überprüfen. Beispielsweise ist die *Konvertierung* einer Datei im DOC-Format von Microsoft Word⁶²⁸ in das jüngere DOCX-Format oberflächlich gesehen unproblematisch, da das neue Format die Funktionalität der Datei nur erweitert. Genauer betrachtet ist die DOCX-Datei jedoch ein Containerformat, das verschiedene Dateien enthält – im Gegensatz zur DOC-Datei, die alle Informationen in einer Datei speichert. Bei einer DOCX-Datei werden die Informationen zur Darstellung und zum Inhalt getrennt gespeichert. Eine DOC-Datei wird somit nicht 1:1 überführt. Zudem können beim Verteilen der Informationen auf die neue Struktur Fehler auftreten, insbesondere bei komplexen oder eingebetteten Elementen wie Skripten. Das neue Format ist zudem beschränkt rückwärtskompatibel, was bei einer Rücktransformation ins Originaldateiformat wiederum zu Fehlern führen kann.

Besonders problematisch ist die Konvertierung eines digitalen Objekts in ein für die Langzeitarchivierung geeignetes Format. Die Auswahl an solchen einfach strukturierten Formaten ist gering. Da nicht alle Funktionen des Ursprungformats ins neue Format übernommen werden können, ist grundsätzlich mit Informationsverlusten zu rechnen. Nachfolgend die vom

⁶²⁵ Konverter: Programm, das Daten einer gegebenen Struktur in eine andere Struktur überführt. Fischer; Hofer: Lexikon der Informatik, 2011. Die hier angesprochenen Dateiformatkonverter sollen den Inhalt und die Eigenschaften des Quellformats mit möglichst wenig Informationsverlust in das Zielformat übertragen.

⁶²⁶ Nach: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 360.

⁶²⁷ Weiterführend: Neuroth, Heike; Strathmann, Stefan; Osswald, Achim u. a. (Hg.): Nestor Handbuch: eine kleine Enzyklopädie der digitalen Langzeitarchivierung, Göttingen 2010, Kap. 7.5. URL: <http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/index.php>. Beispiel für Registry: Library of Congress (USA), Sustainability of Digital Formats, URL: <http://www.digitalpreservation.gov/formats/fdd/descriptions.shtml>.

⁶²⁸ Der Source Code von Microsoft Word for Windows Version 1.1a wird auf der Webseite des Computer History Museums in Mountain View, CA, USA zum Download zur Verfügung gestellt. Siehe URL: <http://www.computerhistory.org/atcm/microsoft-research-license-agreement-msword-v-1-1a/>.

BAR und der KOST empfohlenen und in der Forschung eingesetzten oder sich in Entwicklung befindenden Dateiformate.⁶²⁹

Inhalt	Format	Bemerkung
Text unstrukturiert	TXT (UTF-8, UTF-16, ISO-8859-1, ISO 8859-15, US-ASCII) PDF/A	BAR KOST
'Office' Dokumente	PDF/A	BAR
Tabellen	CSV PDF/A, (XLS, ODF ODER OOXML)	BAR, KOST KOST
Relationale Datenbanken	SIARD (CSV) RDF	BAR, KOST KOST Forschung
Rasterbilder	TIFF PDF/A-2, (JPEG2000) (TIFF/A)	BAR, KOST KOST Forschung
Audio	WAVE PCM	BAR, KOST
Video	MPEG-4 (uncompressed video, MPEG-2, MJPEG2000)	BAR KOST

Tabelle 4: Archivtaugliche Dateiformate nach KOST und BAR (zusammenfassende Darstellung: Föhr)⁶³⁰

Bei jeder Datenformatmigration wird der Bitstrom einer digitalen Quelle in eine neue Abfolge von Zeichen transformiert und somit die Form des digitalen Objekts verändert.⁶³¹ Eine authentische Überlieferung ist damit nicht gewährleistet. Ob eine korrekte und vollständige Übertragung der inhaltlichen Information und der Funktionalität erfolgte, kann anhand eines Vergleiches der Objekte im Originalformat und im Zielformat und unter Umständen anhand der Metadaten überprüft werden.

Hier zeigt sich einer der grossen Unterschiede zwischen digitalen und physischen Objekten auf permanenten Datenträgern: physische können in einem geeigneten Archiv eingelagert und ohne weitere Bearbeitung nach langer Zeit wieder hervorgeholt und gelesen werden. Digitale Objekte müssen zwingend regelmässig migriert werden, ansonsten können sie nach relativ kurzer Zeit aufgrund der Obsoleszenz des Datenformats nicht mehr gelesen werden.

⁶²⁹ Vgl.: Schweizerisches Bundesarchiv: Archivtaugliche Dateiformate. Standards für die Archivierung digitaler Unterlagen, 04.2018. URL: https://www.bar.admin.ch/dam/bar/de/dokumente/konzepte_und_weisungen/archivtaugliche_dateiformate.1.pdf, Stand: 06.05.2018; KOST: Katalog archivischer Dateiformate, Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen, 06.2015, URL: <http://kost-ceco.ch/wiki/whelp/KaD/index.php>, Stand: 14.04.2016. In Klammern sind beschränkt einsetzbare Formate genannt.

⁶³⁰ Laut Lukas Rosenthaler wird an der Entwicklung eines TIFF/A-Standards geforscht. Nach: Rosenthaler, Lukas: Gespräch über digitale Langzeitarchivierung, 04.05.2016.

⁶³¹ Nach: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 403 f.

5.3.5. Informationssystem – Emulation

Digitale Objekte sind auf ein spezifisches digitales Informationssystem angewiesen. Für jedes Objekt ist festzuhalten, welche Eigenschaften das Informationssystem aufweisen muss, damit es in der intendierten Form dargestellt werden kann. Da auch die Informationssysteme Alterungsprozessen unterliegen und deshalb die Funktionsfähigkeit begrenzt ist, müssen digitale Objekte entweder in ein anderes Dateiformat migriert werden oder das gesamte Informationssystem wird emuliert.

Digitale Objekte sollten unabhängig von der Sicherungsform als Primär- oder Sekundärobjekt zu verschiedenen Zwecken wiederverwendet werden können. Es ist zu überprüfen und zusätzlich zu antizipieren, welche Komponenten des spezifischen Informationssystems auch in Zukunft für die Nutzung benötigt werden und deshalb zu sichern sind. Besonders bei proprietären und kundenspezifischen Systemen ist eine entsprechende Sicherung sehr wahrscheinlich nötig. Periphere Hardware wie Joysticks oder Graphiktablets sind vermutlich ebenfalls zu archivieren, weil die Nutzung eines Programms wie etwa ein Computerspiel nur mit dieser Hardware möglich ist und nur damit eine authentische Nutzung erreicht werden kann.⁶³²

Ein Informationssystem kann emuliert werden, damit weniger Hardware gehortet werden muss, deren Funktionieren nicht langfristig gewährleistet werden kann. Eine *Emulation* bedeutet, dass die Komponenten eines Informationssystems innerhalb einer spezifischen Softwareumgebung simuliert werden. Die Hardware eines Computers, die Systemsoftware und Programme können als eigenständige Softwareplattform emuliert werden, auf der in einem veralteten Format geschriebene Objekte darstellbar sind. Zur Programmierung der Plattform müssen genaue technische Beschreibungen der zu emulierenden Hardware-Komponenten vorliegen, damit ihre Funktionsweise im neuen Informationssystem nachgebildet werden kann. Zusätzlich ist die ursprünglich verwendete Software zu erhalten, damit sie auf dem emulierten System installiert werden kann. Ziel einer Emulation ist die Darstellung eines digitalen Objekts in seinem archetypischen System, ohne dass dieses physisch vorhanden sein muss.

Der Vorteil einer Emulation liegt vor allem darin, dass die ursprüngliche Form des digitalen Objekts nicht geändert wird und somit dessen Datenintegrität gewährt bleibt.⁶³³ Die zur Darstellung benötigte Emulationssoftware wird so an die verfügbare Hardware angepasst, dass das Objekt *originalgetreu* dargestellt und die Funktionalität erhalten werden kann. Nachteilig

⁶³² Bspw. werden anscheinend im Spiel 'Pac-Man' nur Rekorde anerkannt, die auf einem originalen Spielhallen-Automaten erzielt wurden, weil dessen Haptik als unabdingbares Kriterium angesehen wird. Siehe: Beinecke, Julius: Zahlen, bitte! 3.333.360 Punkte bis zum «Kill-Screen» von Pac-Man, News, heise online, 28.06.2016, URL: <http://www.heise.de/newsticker/meldung/Zahlen-bitte-3-333-360-Punkte-bis-zum-Kill-Screen-von-Pac-Man-3227685.html>, Stand: 17.07.2016.

⁶³³ Nach: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 372.

ist jedoch, dass für jede Hardwarekonfiguration ein separater Emulator zu programmieren ist und dass bei einem emulierten System der Systemkontext des Objekts meist verloren geht. Zudem muss die periphere Hardware weiterhin vorhanden sein. Beispielsweise können Videospiele aus den 1980er Jahren auf einer Webseite zwar emuliert dargestellt werden, aber es ist so nicht möglich, die haptische Bedienung einer Spielkonsole oder eines 'Flipperkastens' nachzuvollziehen.⁶³⁴ Es ist ein Unterschied, ob das vielleicht erste Computerspiel OXO (Drei gewinnt) auf dem EDSAC⁶³⁵ aus dem Jahr 1949 gespielt wird – wofür es programmiert wurde – oder auf einem jüngeren Computer. Trotzdem ist für Jeff Rothenberg die Emulation die beste Lösung für die Archivierung von digitalen Objekten, denn "this is the only reliable way to recreate a digital document's original functionality, look, and feel. Though it may not be feasible to preserve every conceivable attribute of a digital document in this way, it should be possible to recreate the document's behaviour as accurately as desired – and to test this accuracy in advance."⁶³⁶ Dieser Aussage kann aus obigen Gründen nicht zugestimmt werden und auch Lukas Rosenthaler kritisiert implizit Rothenbergs Ansatz als zu komplex für die dauerhafte Archivierung von digitalen Objekten, obwohl auch er Rothenberg bezüglich look & feel zustimmt.⁶³⁷

Durch Emulation wird zwar die Archivierung von Hardware überflüssig, was von den meisten Archiven sowieso grundsätzlich abgelehnt wird, sie ist aber durch die erhöhte Belastung von Prozessor (CPU) und Speichervorgängen nicht sehr effizient. Die Verarbeitungsgeschwindigkeit nimmt dadurch ab und lässt oft zu wünschen übrig.⁶³⁸ Das kann die Darstellung des Objekts insbesondere bei rechenintensiven Objekten wie Computerspielen negativ beeinflussen, was die Nutzung erschwert oder verunmöglicht. Zudem unterliegt das Emulationsprogramm selbst einem Alterungsprozess und muss seinerseits immer wieder auf neue Hard- und Software angepasst oder allenfalls sogar emuliert oder migriert werden.⁶³⁹

Eine spezielle Art der Emulation wurde vom Ingenieur Raymond Lorie, damals Forscher bei der Firma IBM, vorgeschlagen. Mit einem Universal Virtual Computer (UVC) stellte er eine universelle Computerkonfiguration vor, die jederzeit auch in zukünftige Informationssysteme

⁶³⁴ Beispiele siehe URL: <http://pica-pic.com/>; Der Hersteller 8BITDO TECH HK LTD. produziert mit modernen Betriebssystemen kompatible Versionen von Nintendo-Steuerungsgeräten, um das 'Retro-Feeling' erleben zu können. Siehe URL: <http://www.8bitdo.com/>.

⁶³⁵ Electronic Delay Storage Automatic Calculator, eine der ersten elektronischen Rechenmaschinen.

⁶³⁶ Rothenberg, Jeff: Avoiding technological quicksand: finding a viable technical foundation for digital preservation a report to the Council on Library and Information Resources, Washington (D.C.) Amsterdam 1999, VI.

⁶³⁷ Vgl.: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 476 f.; Rosenthaler: Gespräch über digitale Langzeitarchivierung, 2016.

⁶³⁸ Nach: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 476.

⁶³⁹ Emulierte Softwares sind auf verschiedenen Plattformen zu finden. Siehe bspw. auf Archive.org, URL: <https://archive.org/details/historicalsoftware>.

implementiert werden könnte.⁶⁴⁰ Innerhalb des UVC könnten alle digitalen Objekte dargestellt werden, was für deren Langzeitarchivierung von Vorteil wäre. Allerdings wären die Objekte mittels Portierung⁶⁴¹ an diese spezifische Umgebung anzupassen, was wiederum deren Datenintegrität verletzt. Zudem ist fraglich, ob ein UVC definiert werden kann, der auch zukünftige Anforderungen an ein digitales Informationssystem erfüllt. Nur schon ein oberflächlicher Vergleich zwischen einem mittelpreislichen, privaten Informationssystem aus dem Jahre 2004 und einem von 2014 lässt die Hoffnung diesbezüglich schwinden, weil sich die Hardwarearchitektur inkl. Peripherie stark verändert hat.

Mit einer Emulation können zwar viele digitale Objekte ohne Dateimigration gesichert werden, allerdings ist dann das Emulationsprogramm an die sich ständig verändernde Hardware anzupassen. Das ist bei einzelnen Objektarten unter Umständen sinnvoll, aber je länger und je mehr Dateiformate von der Emulationssoftware unterstützt werden müssen, desto komplexer und aufwendiger wird die Erhaltung dieser Programme sein.

5.3.6. Sicherungsort

Für die Weiternutzung des digitalen Objekts, insbesondere zur Gewährleistung einer kritischen Analyse durch Dritte, ist ein Speicherort zu wählen, der das Objekt *dauerhaft*, unter einer *eindeutigen Adresse*, *integer*, *authentisch* und *manipulationssicher* zur Verfügung stellt. Dies erfüllen Trusted Digital Repositories (TDR), sie stehen aber zurzeit nur für einige wenige Forschungsprojekte zur Verfügung oder sind noch im Aufbau begriffen. Für (Individual-) Forschende sollte das Researcher Driven Archiving (RDA) ermöglicht werden, das mit einer manipulationsresistenten Trusted Saving Procedure (TSP) digitale Forschungsressourcen in TDR dauerhaft aufnehmen kann. Als Übergangslösung könnte der Nutzer Objekte in einem Personal Trusted Archive (PTA) sammeln und publizieren.

Digitale Objekte oder Daten werden von Wissenschaftlern oft auf einem persönlichen Computer oder in einem persönlichen Speicherplatz in der *Cloud*⁶⁴², in einem Institutionsnetzwerk oder in einer Forschungsdatenbank gesichert. Eine kritische Überprüfung durch Dritte, Datenintegrität und Langzeitverfügbarkeit sind selten gewährleistet, weil keine Zugriffsmöglich-

⁶⁴⁰ Weitere Ausführungen zum UVC siehe: Gladney: Preserving Digital Information, 2007, S. 289-299; Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 476 f.; Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 369 f.

⁶⁴¹ Portierung: Übertragung eines Codes von einer Systemplattform auf eine andere. Der Code ist entsprechend auf die neue Systemarchitektur anzupassen. Vgl.: Prevezanos: Computer Lexikon 2013, 2012.

⁶⁴² Cloud: Konzept zur Bereitstellung von Software, Rechen- oder Speicherkapazität über das Internet. Die nötige Hardware befindet sich nicht mehr lokal, sondern wird von einem Dienstleister betrieben. Die Hardware kann sich dabei 'irgendwo' in den Strukturen des Anbieters befinden, d.h. in irgendeinem Rechenzentrum in irgendeinem Land. Zur Nutzung ist eine Internetverbindung zwingend nötig. Siehe: Ebd.

keiten von aussen eingerichtet wurden oder werden können und weil die Daten oft nur während der Projektlaufzeit aktualisiert und gewartet, anschliessend aber entweder gelöscht werden oder einen schwer durchschaubaren *Datenfriedhof* bilden.⁶⁴³ Zwar können kleine, nur von einer Forschungsinstitution unterhaltene Infrastrukturen als sogenanntes Small Scale Academic Web Archiving für die Langzeitarchivierung von selektiven, projektbezogenen Objekten sinnvoll sein⁶⁴⁴, aber die nachhaltige Finanzierung und die Wiederauffindbarkeit der Objekte ist in Zeiten kleiner werdenden Budgets für Universitäten und Forschungsprojekte und dem Aufheben von als 'Orchideenfächer' bezeichneten Spezialdisziplinen nicht gewährleisten.

Für eine dauerhafte Sicherung bieten sich institutions- und fachgebietsübergreifende *Forschungsinfrastrukturen* an. Darunter werden vom Wissenschaftsrat der BRD teilweise einzigartige "Einrichtungen, Ressourcen und Dienstleistungen in öffentlicher oder privater Trägerschaft [...] verstanden, die speziell für wissenschaftliche Zwecke errichtet, mittelfristig bis tendenziell permanent bereitgestellt werden und für deren sachgerechte Errichtung, Betrieb und Nutzung spezifische fachwissenschaftliche oder interdisziplinäre (Methoden-) Kompetenzen erforderlich sind. Ihre Funktion ist es, Forschung, Lehre und Nachwuchsförderung zu ermöglichen oder zu erleichtern."⁶⁴⁵ Des Weiteren sind sie "[...] örtlich fixiert, auf mehrere Standorte verteilt oder werden ausschliesslich virtuell bereitgestellt. Sie werden nicht ausschliesslich von einzelnen Personen oder Gruppen genutzt, sondern stehen prinzipiell einer internationalen Fachgemeinschaft oder mehreren Fachgemeinschaften offen."⁶⁴⁶ Von nestor wurde ein entsprechender Kriterienkatalog für digitale Langzeitarchive erarbeitet, in dem auf den organisatorischen Rahmen, den Umgang mit digitalen Objekten, die Infrastruktur und die Sicherheit eingegangen wird.⁶⁴⁷

Digitale Forschungsinfrastrukturen für die Geistes- und Sozialwissenschaften sind noch wenig entwickelt, vor allem weil in den Geisteswissenschaften eine Vielzahl von Methoden angewendet wird, die sich teilweise vom einen Forschenden zum anderen unterscheiden.⁶⁴⁸ Aus diesen individuellen Forschungsansätzen und Datenmodellen resultieren spezifische Datenbanken und Publikationen mit ureigenem Aufbau sowie spezifischen Anforderungen an

⁶⁴³ Vgl.: Vompras, Johanna; Schirrwagen, Jochen; Horstmann, Wolfram: Die Bibliothek als Dienstleister für den Umgang mit Forschungsdaten, in: Digitale Wissenschaft: Stand und Entwicklung digital vernetzter Forschung in Deutschland ; 20./21. September 2010, Köln ; Beiträge der Tagung, Köln 2011, S. 101. URL: <http://pub.uni-bielefeld.de/publication/2422868>.

⁶⁴⁴ Wie bspw. das 'Digital Archive for Chinese Studies'. Siehe: Lecher, Hanno E.: Small Scale Academic Web Archiving: DACHS, in: Web Archiving, Berlin 2006, S. 213-225.

⁶⁴⁵ Wissenschaftsrat der Bundesrepublik Deutschland: Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020, 2359-12, Köln 13.06.2012, S. 15 f. URL: <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>, Stand: 24.01.2014.

⁶⁴⁶ Ebd., S. 16.

⁶⁴⁷ Siehe: Neuroth u. a. (Hg.): Nestor Handbuch, 2010, Kap. 5.4.

⁶⁴⁸ Nach: Hügi, Jasmin; Schneider, René: Digitale Forschungsinfrastrukturen in den Geistes- und Sozialwissenschaften, Haute école de gestion den Genève, Genf 24.01.2013.

Zugang und Nutzung.⁶⁴⁹ Zudem werden in den Geisteswissenschaften, im Gegensatz insbesondere zu den Naturwissenschaften, Forschungsressourcen hauptsächlich *gesammelt* und weniger *generiert*, was andere Ansprüche an die Infrastruktur stellt.⁶⁵⁰ Forschungsumgebungen müssen die nötige Infrastruktur mit den entsprechenden Komponenten zur Verfügung stellen und gewährleisten, dass die enthaltenen Daten über eine längere Zeit hinweg Bestand haben und öffentlich zugänglich sind. Der Kriterienkatalog und der gewünschte Funktionsumfang ist sehr umfangreich, was entsprechende finanzielle Kosten nach sich zieht.⁶⁵¹

Im Folgenden skizziere ich fünf Haupttypen von Forschungsinfrastrukturen, um aufzuzeigen, welcher Typus für die Sicherung von digitalen Quellen von (Individual-)Forschenden am besten geeignet ist. Da für die Quellensicherung eine geeignete Infrastruktur zur Verfügung stehen muss, sollte deren Entwicklung dringend vorangetrieben werden.

Ich unterscheide institutionell-lokale, institutionell-externe, individuell-lokale, individuell-externe und extern-institutionelle Infrastrukturen.

Institutionell-lokal sind informationstechnische Infrastrukturen, die von einer Forschungsgruppe oder -institution aufgebaut und unterhalten werden.⁶⁵² Dies können interne Dateiablagen oder projektbezogene Datenbanken sowie öffentlich zugreifbare Dokumentenrepositories sein. Es ist dabei unerheblich, ob die Daten öffentlich oder nur einem beschränkten Nutzerkreis zugänglich sind. Ausschlaggebend ist, dass eine Forschungsgruppe vollumfänglich für die Infrastruktur verantwortlich ist, wobei für das technische Know-how oft externe Spezialisten hinzugezogen werden, die nicht dauerhaft zur Verfügung stehen. Zudem ist problematisch, dass der Unterhalt der Infrastruktur und des darin enthaltenen Datenbestandes meist aus Zeit- und Kostengründen einen kleinen Stellenwert einnimmt und damit die Forschungsressourcen nach Projektende allenfalls nicht mehr zur Verfügung stehen. Ein Beispiel für eine institu-

⁶⁴⁹ Nach: Sahle, Patrick; Kronenwett, Simone; Blumtritt, Jonathan: Sustainability?! Four Paradigms for Humanities Data Centers (Poster), Poster, digital humanities Lausanne - Switzerland '14, 10.07.2014, URL: <http://dharchive.org/paper/DH2014/Poster-285.xml>, Stand: 01.04.2015.

⁶⁵⁰ Nach: Hügi; Schneider: Digitale Forschungsinfrastrukturen in den Geistes- und Sozialwissenschaften, 2013, S. 17-20. Hügi und Schneider versuchen den Terminus Forschungsdatum für die Geisteswissenschaften zu verwenden, können aber keine klare Definition geben. Der Begriff 'Forschungsressource' wäre zielführend.

⁶⁵¹ Bestehende Forschungsinfrastrukturen bspw.: ARIADNE (URL: <http://www.ariadne-infrastructure.eu/>), The Common Language Resources and Technology Infrastructure CLARIN (URL: <http://clarin.eu/>), The Digital Research Infrastructure for the Arts and Humanities DARIAH (URL: <https://www.dariah.eu/>), Network for Digital Methods in the Arts and Humanities NeDiMAH (URL: <http://www.nedimah.eu/>), System for Annotation and Linkage of Sources in Arts and Humanities SALSAH (URL: <http://www.salsah.org/>), TextGrid (URL: <http://www.textgrid.de/>).

⁶⁵² Vgl. auch: Bargheer, Margo; Bellem, Saskia; Schmidt, Birgit: Open Access und Institutional Repositories – Rechtliche Rahmenbedingungen, in: Spindler, Gerald (Hg.): Rechtliche Rahmenbedingungen von Open Access-Publikationen, Göttingen 2006 (Göttinger Schriften zur Internetforschung), S. 1 f.

tionell-lokale Infrastruktur ist das Internet Archive, das ohne gesetzlich festgelegten Auftrag seine Dienstleistungen für Unterhalt, Datensicherheit oder externen Zugriff jederzeit und aus irgendwelchen Gründen aufgeben kann.

Wenn Forschungsgruppen oder -institutionen ihre Infrastruktur an einen externen Provider auslagern, der für den gesamten Unterhalt der Infrastruktur und der Forschungsressourcen zuständig ist, spreche ich vom *institutionell-externen* Typus. Es ist ebenfalls unerheblich, ob die Daten öffentlich oder nur einem bestimmten Nutzerkreis zugänglich sind. Der Vorteil dieses Typus ist jedoch, dass der Provider mit spezialisiertem Know-how die Infrastruktur unterhält und die Forschungsressourcen sehr wahrscheinlich über lange Zeit erhalten bleiben. Zudem kann der Provider sicherstellen, dass Daten unter bestimmten Bedingungen veröffentlicht werden. Ein externer Provider kann allerdings seine Dienstleistungen einstellen, sobald die vereinbarten Gegenleistungen nicht mehr erfolgen, er die Dienstleistung nicht mehr anbieten will oder wenn die Geschäftstätigkeit aufgegeben wird. Die dauerhafte Verfügbarkeit der digitalen Quellen ist nicht gewährleistet. Deshalb sollten aus der oben genannten Definition des Wissenschaftsrates⁶⁵³ die privaten Trägerschaften daraus gestrichen werden.

Von Forschenden oder Forschungsgruppen, die nicht einer Institution angehören, wird oft eine *individuell-lokale* Infrastruktur aufgebaut.⁶⁵⁴ Diese besteht aus individuell und persönlich angeschaffter Hard- und Software, mit der die Forschungsressourcen erstellt, verarbeitet und gespeichert werden. Die Verantwortung für den Unterhalt liegt bei den einzelnen Personen, die auch die Datensicherheit zu gewährleisten haben. Dritte können selten auf die gesammelten Daten zugreifen und sie nutzen. Die Wahrscheinlichkeit, dass die Forschungsressourcen über längere Zeit erhalten und nutzbar bleiben, ist relativ gering, weil eine Nachnutzung der digitalen Quellen nicht vorgesehen ist und der Unterhalt nach Abschluss eines Projektes daher meist eingestellt wird.

Bei *individuell-externen* Infrastrukturen lagern die nicht in eine Institution eingebundenen Forscher ihre Forschungsressourcen an einen externen Provider aus. Dieser ist wiederum für den Unterhalt und die Veröffentlichung der Daten zuständig. Er kann die Dienstleistungen einstellen, sobald die vereinbarten Gegenleistungen nicht mehr erfolgen, er die Dienstleistung nicht mehr anbietet oder wenn die Geschäftstätigkeit

⁶⁵³ Wissenschaftsrat der Bundesrepublik Deutschland: Empfehlungen zu Informationsinfrastrukturen, 2012, S. 15 f.

⁶⁵⁴ Teilweise werden solche Infrastrukturen auch von Forschenden einer Institution aufgebaut, wenn diese die gewünschten Tools nicht zur Verfügung stellt.

aufgegeben wird. Die dauerhafte Verfügbarkeit der digitalen Quellen ist nicht gewährleistet.

Mit *extern-institutionell* ist eine Forschungsinfrastruktur gemeint, die ein Provider mit einem gesetzlichen Auftrag und staatlichen Mitteln dauerhaft anbietet. Beispielsweise wird Wissenschaftlern im von der EU finanziell geförderten Projekt DARIAH der Zugang zu heterogenen, nachnutzbaren Forschungsdaten und zu Werkzeugen für die gemeinsame Nutzung zur Verfügung gestellt.⁶⁵⁵ Allerdings werden die in den einzelnen Ländern angesiedelten DARIAH-Plattformen unterschiedlich lange gefördert, eine Weiterführung ist nicht garantiert. Ebenfalls nicht garantiert ist die Finanzierung des Projekts nestor, des Deutschen Kompetenzzentrums für digitale Langzeitarchivierung mit internationaler Ausstrahlung.⁶⁵⁶ Es ist nicht klar, ob und wie lange die dort gesicherten Daten zur Verfügung stehen und was mit diesen im Falle der Einstellung der Projekte geschieht.

Als Provider kommen nur Forschungsinfrastrukturen infrage, die "reliable, long-term access to managed digital resources to its designated community, now and in the future"⁶⁵⁷ zur Verfügung stellen. Solche Infrastrukturen werden als *Trusted Digital Repositories (TDR)* bezeichnet. Die Anforderungen an solche Repositories sind vielfältig: die Verwendung des OAIS-Referenzmodells (Open Archival Information System), die Einhaltung von Standards wie ISO 16363 oder DIN 31644⁶⁵⁸, Best Practice-Richtlinien⁶⁵⁹, organisatorische Realisierbarkeit, gesetzlicher Auftrag, nachhaltige finanzielle Stabilität, technische und prozessuale Infrastruktur zur Umsetzung von angemessenen Archivierungsstrategien sowie regelmäßige Überprüfungen und Zertifizierungen wie das *Data Seal of Approval*.⁶⁶⁰ Solche

⁶⁵⁵ URL: <http://dariah.eu/>. Siehe auch: Romary u. a.: DARIAH-EU, 2012, S. 36. Die Europäische Kommission fördert ERIC (European Research Infrastructure Consortia), das wiederum DARIAH alimentiert. Siehe URL: <http://ec.europa.eu/research/infrastructures/>.

⁶⁵⁶ Nach: Kretzschmar, Robert: Archive als digitale Informationsinfrastrukturen. Stand und Perspektiven, in: *Archivar, Zeitschrift für Archivwesen* 2, 2013, S. 150.

⁶⁵⁷ Research Libraries Group (RLG) and Online Computer Library Center (OCLC) Working Group on Digital Archive Attributes 2002, S. 5, in: Day, Andreas: *The Long-Term Preservation of Web Content*, in: Masanès, Julien (Hg.): *Web Archiving*, Berlin 2006, S. 181. Zu Repositories siehe bspw.: Furlough, Mike: *What We Talk About When We Talk About Repositories*, in: *Reference & User Services Quarterly* 49 (1), 01.10.2009, S. 18-32.

⁶⁵⁸ ISO siehe URL: <http://www.iso.org/>; DIN: Deutsches Institut für Normung; siehe URL: <http://www.din.de/>.

⁶⁵⁹ Best Practice: bestmögliche, oft bereits erprobte Methode, Maßnahme o. Ä. zur Durchführung oder Umsetzung von etwas. Nach Duden-Online, URL: <http://www.duden.de/>.

⁶⁶⁰ Vgl. auch: Day: *Long-Term Preservation*, 2006, S. 181 f. Zu Research Infrastructure / Cyberinfrastruktur siehe auch: Makarow, Marja; Moulin, Claudine; Žic Fuchs, Milena: *Research Infrastructures in the Digital Humanities*, Strasbourg 2011 (Science Policy Briefing 42), S. 4-8. URL: http://resaw.eu/wp-content/uploads/2013/06/ESF_Research-Infrastructures-in-the-Digital-Humanities.pdf, Stand: 06.05.2018. Das Data Seal of Approval wird seit dem Jahr 2008 von einem internationalen Expertengremium an digitale Langzeitarchive vergeben, die sich anhand von 16 grundlegenden Merkmalen als besonders vertrauenswürdig und nachhaltig erweisen (siehe URL: <http://datasealofapproval.org/>). Bei nachträglichen Zertifizierungen stellt sich die Frage, wie mit

Infrastrukturen sind aufgrund der Notwendigkeit für die internationale Forschung und den darin enthaltenen, teils unter Datenschutz oder Betriebsgeheimnis stehenden Datenbeständen als kritische Infrastrukturen zu betrachten und sollten deshalb möglichst herstellerunabhängig sein und vollständig durch den Provider kontrolliert werden. Nur mit professionell unterhaltenen Forschungsinfrastrukturen ist es möglich, dass Forschungsressourcen über längere Zeit manipulationssicher und nachvollziehbar zur Verfügung stehen.

Ein TDR ist zwar einem physischen Archiv ähnlich, unterscheidet sich aber davon durch die virtuelle Struktur, viele andersartige Prozesse und die Notwendigkeit ständiger Überprüfung und Anpassung der Bestände. Ein TDR respektive der Provider als Institution sollte deshalb als neuer Typus unter den gesetzlich beauftragten Gedächtnisinstitutionen anerkannt werden:

Ein *Trusted Digital Repository TDR* ist ein neuer Typus von gesetzlich beauftragter Gedächtnisinstitution, in dem digitale Forschungsressourcen verlässlich, dauerhaft und manipulationssicher zur Verfügung gestellt und unterhalten werden.

Individualforscher verwenden für die Sicherung ihrer Forschungsressourcen und -notizen häufig eine individuell-lokale Infrastruktur wie das Literaturverwaltungsprogramm Zotero⁶⁶¹, oder eine individuell-externe Infrastruktur wie das Internet Archive.⁶⁶² Zotero bietet vielfältige Möglichkeiten zur Sicherung von unterschiedlichen digitalen Objekten, die als Forschungsressource infrage kommen. Primärobjekte wie ein Bericht im PDF-Format und Sekundärobjekte wie ein Screenshot einer Webseite können zusammen mit ergänzenden Metadaten automatisch in die Datenbank aufgenommen werden. Mit einem Nutzerkonto kann die Datenbank auf verschiedenen Informationssystemen synchronisiert und als individuell-externes Modul auch kollaborativ genutzt werden. Für einen einzelnen Forscher oder eine Forschungsgruppe ist dieses Tool zwar eine enorme Arbeitshilfe, aber die darin enthaltenen Objekte sind weder öffentlich zugänglich, noch stehen sie dauerhaft zur Verfügung, weil sie nicht professionell unterhalten werden. Die Sammlung selbst gehört zwar dem Nutzer, aber wenn der Software-Anbieter den Online-Zugang sperrt, kann die Sammlung nicht mehr, oder nur noch eine gewisse Zeit lang, lokal bearbeitet werden.

Das Internet Archive bietet die Möglichkeit, spezifische Webseiten in definierten Abständen zu sichern. Allerdings werden solche Sicherungsprojekte nur auf Anfrage eingerichtet und sind von Fall zu Fall mit dem Betreiber zu vereinbaren. Es ist zwar die spontane Sicherung

Objekten umgegangen wird, die vor der Zertifizierung ins Archiv aufgenommen worden sind, denn diese sind unter Umständen nicht korrekt erfasst worden.

⁶⁶¹ URL: <https://www.zotero.org/>. Weitere Beispiele: LitLink (URL: <http://www.litlink.ch/>), Mendely (URL: <https://www.mendeley.com/>), Citavi (URL: <http://www.citavi.de/>).

⁶⁶² URL: <https://archive.org/>.

einer Webseite möglich⁶⁶³, aber nicht die eines einzelnen digitalen Objekts. Zudem können diese Sicherungen nicht in eine individuelle, forschungsthemabezogene Sammlung aufgenommen werden. Eine Sicherung ist auch nicht in jedem Fall möglich, denn ein Webseitenbetreiber kann die Sicherung ins Internet Archive mit entsprechenden Einträgen im Webseitencode blockieren. Individuelle Bedürfnisse von Forschenden können von solchen Digitalarchiven aufgrund ihrer strategischen Ausrichtung und Organisationsstruktur selten erfüllt werden.⁶⁶⁴ Die Historikerin Jill Lepore meint, dass ein globales Internet Archive deshalb scheitern würde, weil es einerseits durch die länderspezifischen Rechtslagen verhindert würde und weil andererseits ein gegenseitiges Misstrauen unter den politischen Akteuren herrscht.⁶⁶⁵ Insbesondere misstrauen die europäischen Institutionen den US-amerikanischen aufgrund der komplexen amerikanischen Urheberrechtsbestimmungen und weil einige US-Unternehmen die Eigentumsrechte an den gesicherten Objekten beanspruchen. Die Diskussionen und rechtlichen Auseinandersetzungen um Datenschutzbestimmungen wie derjenigen im vom Europäischen Gerichtshof aufgehobenen Safe-Harbor-Abkommen verhindern den Aufbau von verlässlichen, vertrauenswürdigen Infrastrukturen für Forschende. Für den ungehinderten Zugang zu und die Nutzung von digitalen Forschungsressourcen sind klare und international anerkannte Rechtsgrundlagen eine der zwingenden Voraussetzungen, ansonsten werden die Forschenden die Objekte nicht zur Verfügung stellen.

Lösungsansatz 'Researcher Driven Archiving'

Weil digitale Objekte aufgrund ihrer Volatilität und Manipulierbarkeit sowie durch fehlende Persistenz nicht langfristig verfügbar sind und nachvollzogen werden können, sollten sie von Forschenden individuell gesichert und zugänglich gemacht werden können. Extern-institutionelle Forschungsinfrastrukturen wie DaSCH der SAGW⁶⁶⁶, DARIAH der EU, Library of Congress der Vereinigten Staaten⁶⁶⁷ oder ScienceCloud der Universität Zürich⁶⁶⁸, die teilweise bereits bestehen oder im Aufbau begriffen sind, sind hauptsächlich auf die Bedürfnisse von mittleren und grossen Forschungsinstitutionen oder staatlichen Verwaltungen ausge-

⁶⁶³ Eingabe der URL der Webseite in das Feld bei 'Save Page Now' unter URL:

<https://archive.org/web/>.

⁶⁶⁴ Nach Brügger. Seine Aussage bezieht sich zwar auf die Archivierung von Webauftritten, sie kann aber allgemein für Digitalobjekte ausgelegt werden. Siehe: Brügger: Web History and the Web as a Historical Source, 2012, Abs. 10.

⁶⁶⁵ Vgl.: Lepore: The Cobweb, 2015.

⁶⁶⁶ Siehe: Immenhauser, Beat: SAGW goes digital, in: Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Digital Humanities und Web 2.0, 2012, S. 31; Schweizerische Akademie der Geistes- und Sozialwissenschaften: Final report for the pilot project «Data and Service Center for the Humanities» (DaSCH), Schlussbericht, Schweizerische Akademie der Geistes- und Sozialwissenschaften, Bern 30.03.2015. URL: www.akademien-schweiz.ch/fr/dms/publikationen/10/report1001.pdf, Stand: 06.05.2018.

⁶⁶⁷ URL: <https://www.loc.gov/>.

⁶⁶⁸ URL: <http://www.s3it.uzh.ch/infrastructure/sciencecloud/>.

richtet. Die Möglichkeit zur Sicherung von digitalen Forschungsressourcen ist aber auch für Individualforscher dringend notwendig. Der Historiker Patrick Sahle sieht eine mögliche Lösung für die langfristige Archivierung von Forschungsdaten in "dedizierten Datenzentren, die als neue institutionelle Rahmenorganisationen die vielfältigen Aufgaben dadurch meistern, dass sie die Kompetenzen der Bestände bewahrenden, der technischen Infrastruktur bereitstellenden und der methodischen und fachwissenschaftlichen Kompetenz produzierenden Akteure zusammenbringen."⁶⁶⁹ Rosenthaler weist ergänzend auf Folgendes hin: "no single, monolithic virtual research environment with a generic user interface can meet the needs of all projects."⁶⁷⁰ Jedes für die digitale Langzeitarchivierung von Forschungsdaten konzipierte Datenzentrum sollte für alle Forschenden offenstehen.

Um die Bedürfnisse der Individualforscher zu erfüllen, schlage ich das *Researcher Driven Archiving (RDA)* vor. Darunter wird ein Sicherungsprozess verstanden, der Forschenden die Sicherung von digitalen Forschungsressourcen unabhängig von Forschungsprojekten, -programmen oder -institutionen in TDR erlaubt. Dabei sollten Forschende ihre Forschungsressourcen aus ihren individuell-lokalen in extern-institutionelle Infrastrukturen überführen und anschliessend auch veröffentlichen können. Erst dann ist es möglich, dass Dritte auf die verwendeten digitalen Quellen zugreifen und deren Verwendung nachvollziehen können. Durch die Sicherung der Quellen wird der Forschende gleichzeitig zum Archivar, denn er wählt gezielt aus, was dauerhaft gespeichert wird. Das ist mit ein Grund, warum sich Forschende mit den Prinzipien der Archivwissenschaften vertraut machen sollten.

Mit RDA kann der Forscher das für ihn Wichtige vom Unwichtigen trennen und trägt zur Verringerung des Datenvolumens bei. Institutionen wie das Internet Archive können nur sehr spezifische Sicherungen durchführen, weil die Speicherung der Gesamtheit an vorhandenen Daten nicht sinnvoll, teuer und grundsätzlich gar nicht möglich ist. Der Historiker Mark Sandle nennt die Absicht, alles speichern zu wollen, ein 'Hirngespinnst von totaler Forschung': "the search is continuous and endless, Sisyphean in many ways."⁶⁷¹ RDA trägt dazu bei, dass nur diejenigen Objekte gesichert werden, die für den Forscher auch relevant sind.

⁶⁶⁹ Sahle, Patrick: Forschungsdaten in den Geisteswissenschaften, in: Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Geisteswissenschaftliche Forschung und ihre Daten, Bern 2015 (SAGW Bulletin), S. 45. URL: http://www.sagw.ch/de/dms/sagw/bulletins_sagw/bulletins_2015/Bulletin_4_15/Dossier-SAGW_Bulletin_4_15.pdf. Sahle verweist auf das DaSCH.

⁶⁷⁰ Schweizerische Akademie der Geistes- und Sozialwissenschaften: Final Report, 2015, S. 2.

⁶⁷¹ Siehe: Sandle, Mark: Studying the past in the digital age. From tourist to explorer, in: Weller, Toni (Hg.): History in the digital age, London 2013, S. 137.

Unter *Researcher Driven Archiving (RDA)* wird ein Sicherungsprozess verstanden, der Forschern die Sicherung von digitalen Forschungsressourcen unabhängig von Forschungsprojekten, -programmen oder -institutionen in verlässlichen Forschungsinfrastrukturen erlaubt.

Infrastrukturen für das RDA sollten genügend Flexibilität für die Sicherung von diversen digitalen Objektarten bieten, individualisierbar sein, damit nutzerspezifische Sicherungen und Veröffentlichungen erfolgen können, und als TDR alle nötigen Prozesse bieten, die zur Langzeitarchivierung solcher Objekte auch unter Wahrung des Copyrights nötig sind. Zusätzlich sollten ergänzende Informationen wie die Beschreibungen der angewendeten Methoden und Arbeitsweisen oder der Lebenslauf des Autors darin aufgenommen werden können, um den Nutzern möglichst viele Metadaten zur Verfügung stellen zu können. Je mehr überprüfbare Metadaten vorhanden sind, desto vertrauenswürdiger wird die Quelle.⁶⁷²

Produkte wie *Archive-in-a-Box (AiB)*⁶⁷³ von DARIAH-DE, das *Humanities Data Centre*⁶⁷⁴ oder das bereits erwähnte *DaSCH* der SAGW weisen zwar in diese Richtung, sind aber noch zu wenig auf den Individualforscher ausgerichtet. Die Diskussion dreht sich zurzeit immer noch um persönliche, lokale und/oder nicht persistente oder spezifisch forschungsprojektorientierte Archive, die nur begrenzten Nutzerkreisen zugänglich sind.⁶⁷⁵ Der Aufbau von privaten digitalen Archiven wird zwar von verschiedenen Institutionen und Programmen unterstützt, diese bieten aber meist nur Anleitungen und Hilfestellungen, wie beispielsweise im *National Information Infrastructure and Preservation Program (NDIIPP)* der Library of Congress (US)⁶⁷⁶, und stellen keine für wissenschaftliche Belange ausreichende Infrastruktur mit den genannten Merkmalen zur Verfügung.

Lösungsansatz 'Trusted Saving Procedure'

Wissenschaftliche Forschungsinfrastrukturen sollten nicht nur die Sicherung, Verwaltung und Langzeitarchivierung von digitalen Objekten ermöglichen, sondern auch die *Integrität der*

⁶⁷² Weiterführend in Kap. 5.3.7. Metadaten.

⁶⁷³ URL: <https://de.dariah.eu/archive-in-a-box/>; hätte im Frühjahr 2014 bereitgestellt werden sollen, bis März 2015 nur Testzugang verfügbar, kurz darauf abgeschaltet. Dokumentation zu AiB: Becker, Rainer; Castilho, Richard Eckart de: *Archive-in-a-box (Internal Demonstrator) – Service-Concept M 1.4.1.1*, 31.08.2012. URL: https://dev2.dariah.eu/wiki/download/attachments/14651583/M1411_Archive_in_a_box.pdf?version=2&modificationDate=1366904206713&api=v2. Zum Aufbau von DARIAH siehe: DARIAH-DE: *DARIAH-DE – Aufbau von Forschungsinfrastrukturen für die eHumanities*, in, Göttingen 2012. DOI: 10.3249/webdoc-3589.

⁶⁷⁴ URL: <http://humanities-data-centre.org/>.

⁶⁷⁵ Siehe bspw. Beiträge an der Personal Digital Archive Conference 2015 in New York. Siehe: Ashenfelder, Mike: *The Personal Digital Archiving 2015 Conference*, *The Signal*, 03.08.2015, URL: <https://blogs.loc.gov/thesignal/2015/08/report-on-the-personal-digital-archiving-2015-conference/>, Stand: 06.05.2018.

⁶⁷⁶ URL: <http://www.digitalpreservation.gov/>.

gesicherten Objekte gewährleisten. Dies ist sowohl bei bereits gesicherten als auch bei zu importierenden Objekten sicherzustellen, um vertrauenswürdige, authentische und integre Objekte zu generieren. Es sollten demnach softwaretechnische Methoden zur Verfügung gestellt werden, die einen manipulationssicheren Import von digitalen Objekten durch (Individual-)Forschende ermöglichen. Solche Methoden sind jedoch noch nicht bekannt, weshalb ich die Entwicklung einer Software⁶⁷⁷ vorschlage, die die manipulationsresistente Sicherung von Forschungsressourcen in TDR gewährleistet.⁶⁷⁸ Ein solcher Sicherungsprozess für das Researcher Driven Archiving wird im Folgenden als *Trusted Saving Procedure (TSP)* bezeichnet. Mit der Entwicklung einer nutzerfreundlichen Software sollte es möglich sein, auch informationstechnisch wenig versierten Geschichtswissenschaftlern ein manipulationssicheres, nachvollziehbares und integriertes Repository für digitale Forschungsressourcen zur Verfügung zu stellen.

Die *Trusted Saving Procedure (TSP)* ist ein manipulationssicherer Prozess, der es einem Nutzer ermöglicht, mit einer Software ein digitales Objekt in einem vertrauenswürdigen Repository (TDR) zu sichern.

Folgendes Schema verdeutlicht die Funktionalität der TSP:

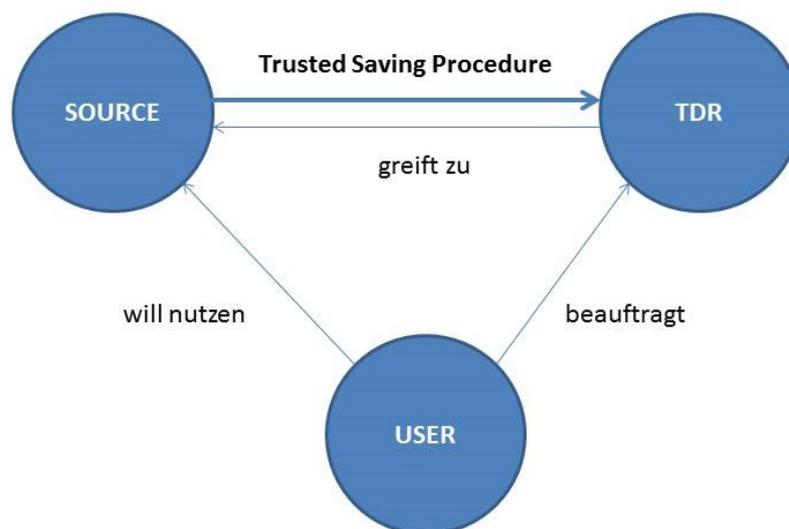


Abbildung 15: Schema Trusted Saving Procedure (Föhr)

Da digitale Objekte per Definition manipulierbar sind, hat der Import eines Objekts in ein TDR direkt zwischen dem Objektprovider und dem TDR zu erfolgen. Der Nutzer darf in den Sicherungsprozess nicht eingreifen können, weil jegliche Zwischenspeicherung auf einem durch den Nutzer beeinflussbaren Medium die Wahrscheinlichkeit einer Manipulation erhöht und

⁶⁷⁷ Software als eigenständiges Programm (Tool), als Add-on, das im Browser integriert werden kann, als webbasierter Dienst etc.

⁶⁷⁸ Siehe dazu: Collins u. a.: *Going Digital*, 2015, S. 27.

damit die Vertrauenswürdigkeit des Objekts mindert. Die Verwaltung mit Tätigkeiten wie Ordnen, Löschen oder Verschieben, das Anbringen von Annotationen wie Notizen, Tags oder definierte Metadaten und Korrekturen sollen vom Nutzer direkt auf der TDR-Infrastruktur durchgeführt werden können. Nur bestimmte Sicherheitsmerkmale wie Hashwert, Zeitstempel oder digitale Signaturen sollen nicht durch den Nutzer bearbeitbar sein, um eine möglichst hohe Authentizität zu gewährleisten.⁶⁷⁹

Der Forschende kann in einer dafür ausgelegten TDR-Forschungsinfrastruktur ein Projekt eröffnen, in dem alle seine Forschungsressourcen gesichert werden. Jedes darin enthaltene digitale Objekt wird mit einer eindeutigen URI respektive einem Permalink und einem Zeitstempel versehen.⁶⁸⁰ Während der aktiven Zeit des Projekts kann der Nutzer alle enthaltenen Objekte bearbeiten. Vor Abschluss des Projekts definiert er diejenigen Objekte, die veröffentlicht werden sollen, alle anderen werden gelöscht oder in eine andere Sammlung verschoben. Schliesst der Nutzer die Sammlung mit einem nicht mehr widerrufbaren Befehl, werden alle darin enthaltenen Daten durch das TDR vor Manipulation geschützt und alle weiteren Änderungen werden nachvollziehbar protokolliert. Das TDR stellt nun mit persistenten Adressen und einem Langzeitarchivierungskonzept sicher, dass auf die Objekte dauerhaft zugegriffen werden kann.

Beispielprozess

Der Nutzer möchte einen Online-Artikel einer Zeitung sichern und in seinem Projekt als digitale Quelle verwenden.⁶⁸¹

1. Der Nutzer gibt die URL des Artikels in das Adressfeld des Browsers ein.
2. Er öffnet die TSP-Software.
3. Der Nutzer wählt darin aus, welche Elemente der Seite wie gesichert werden sollen; beispielsweise als HTML-File oder/und als PDF/A-2 oder/und als Bild – die Software schlägt unter Umständen ein ideales Format vor.
4. Die mit dem TDR verbundene Software initiiert die TSP.
5. Die Sicherung erfolgt direkt zwischen dem Speicherort des Objekts (URL) und dem TDR unter Einbezug vorhandener Metadaten. Dem Objekt wird durch das TDR eine eindeutige URI, einen Permalink und Sicherheitsmerkmale wie ein Zeitstempel und ein Hash-Wert zugewiesen.

⁶⁷⁹ Hashwert, Zeitstempel und digitale Signatur werden später detaillierter beschrieben.

⁶⁸⁰ Dies ist bspw. im DaSCH vorgesehen: Rosenthaler, Lukas; Fornaro, Peter; Clivaz, Claire: DASCH: Data and Service Center for the Humanities, in: Digital Scholarship in the Humanities. The Journal of the Alliance of Digital Humanities Organizations 30 (Supplement 1), 2015, S. 45.

⁶⁸¹ Prozess ohne Berücksichtigung von allfälligen rechtlichen Problemstellungen.

6. Der Nutzer kann die Sicherung und die zugehörigen, ungeschützten Metadaten in der Software bearbeiten.
7. Der Nutzer kann den/die eindeutige(n) URI/URL in seiner Publikation verwenden.
8. Der Nutzer schliesst das Projekt ab, die darin enthaltenen Objekte können veröffentlicht werden.
9. Jeder Leser der Publikation kann selbständig auf den/die darin enthaltenen URI/URL zugreifen.
10. Der Provider stellt sicher, dass das TDR den unter dem/der URI/URL eingetragenen Artikel dauerhaft zur Verfügung stellt.

Das von amerikanischen Bibliotheken ins Leben gerufene Projekt *perma.cc* verfolgt eine der propagierten sehr ähnliche Strategie.⁶⁸² Aus einer eingegebenen URL sichert *perma.cc* den Inhalt der Adresse und erstellt einen Permalink darauf. Allerdings ist es noch nicht möglich, Objekte zu archivieren, die nicht unter einer URL aufzufinden sind. Zudem ist das Projekt noch in einer Beta-Phase und wird hauptsächlich deshalb nur von US-amerikanischen Institutionen unterstützt. Der vielversprechende Ansatz sollte weiterverfolgt werden.

Lösungsansatz 'Personal Trusted Archive'

Der Aufbau von TDR-Infrastrukturen für das RDA hat aufwendige Anpassungen an bestehenden sowie einen kostenintensiven und zeitaufwendigen Aufbau von neuen Systemen zur Folge. Als *Zwischenlösung* käme die Verwendung einer Software infrage, die digitale Objekte in vertrauenswürdigen, aber persönlichen und lokalen Archiven von Forschenden sichert. Diese Archivdatei könnte anschliessend in einem extern-institutionellen Repository veröffentlicht werden.

Diese Software sichert ähnlich wie die der TSP die gewünschten digitalen Objekte in einem separaten, vertrauenswürdigen Verzeichnis, dem *Personal Trusted Archive (PTA)*. Dieses Container-Archiv könnte sich lokal oder in einem virtuellen Sicherungsort in der Cloud befinden, ist verschlüsselt und protokolliert alle Vorgänge. Der Nutzer sichert ein Objekt durch eine darauf angepasste TSP, die es zusammen mit Sicherheitsmerkmalen und entsprechenden Metadaten im definierten PTA sichert – in langzeitarchivfähigen Formaten. Wiederum werden vor dem Abschluss des Projekts die verwendeten und zur Publikation vorgesehenen Objekte definiert. Durch einen nicht widerrufbaren Befehl werden sie in einem manipulationssicheren Container abgelegt und könnten auf DVD, im Web, in einem Repository oder

⁶⁸² URL: <https://perma.cc/>.

anderen Datenträgern publiziert werden. Der Container muss zwingend in einem offenen Format programmiert sein, damit er auch in weiter Zukunft geöffnet werden kann.⁶⁸³

Das *Personal Trusted Archive (PTA)* ist ein vertrauenswürdiger, persönlicher digitaler Container, in dem ein Nutzer in einem manipulationssicheren Prozess (TSP) eine Sammlung von digitalen Objekten erstellen und publizieren kann.

Im Gegensatz zum TDR ist bei diesem Verfahren weder die Langzeitarchivierung noch der Zugriff auf die Objekte gewährleistet. Zudem kann der Nutzer während des Sammelns direkt auf die gesicherten Objekte zugreifen, was den Schutz vor Manipulationen mindert. Essenziell ist, dass die Software durch den Nutzer ständig aktuell gehalten wird, damit auch immer die aktuellsten Verschlüsselungsverfahren verwendet werden. Sobald ein Verfahren nicht mehr sicher ist respektive geknackt wurde, sind alle bisher erstellten Container mit einem aktuellen Verfahren neu zu verschlüsseln, um die Authentizität der Inhalte zu gewährleisten.

Bereits eingesetzt werden Kopier- oder Archivierungsprogramme wie HTTrack.⁶⁸⁴ Im Unterschied zu diesen verwendet ein PTA eine TSP und ausschliesslich langzeitarchivierungsfähige Formate. Nur damit lassen sich vertrauenswürdige und über längere Zeit nutzbare digitale Quellen generieren, die den Anforderungen des wissenschaftlichen Arbeitens genügen.

5.3.7. Metadaten

*"The digital age is the age of standardisation."*⁶⁸⁵ (Collins, 2015)

Metadaten bieten eine *Dokumentation zu Inhalt, Kontext und Überlieferungsgeschichte* eines Objekts oder einer Sammlung.⁶⁸⁶ Metadatenstandards unterstützen den automatisierten Austausch dieser Informationen über verschiedene Plattformen hinweg. Die Beurteilung der Authentizität eines Objekts hängt in grossem Masse von den zur Verfügung stehenden Metadaten ab. Trotz dieser wichtigen Aufgabe werden Metadaten bei digitalen Publikationen oft nachlässig oder gar nicht gepflegt. Dies trifft auch auf wissenschaftliche Publikationen zu, die eigentlich den Richtlinien des wissenschaftlichen Arbeitens verpflichtet sind und deshalb aussagekräftige Metadaten zur Verfügung stellen sollten. Um diesen Missstand zu beheben, sollte eine standardisierte Metadata Information Section (MIS) eingeführt werden, die alle relevanten bibliographischen Metadaten enthält.

⁶⁸³ Offenes Format und starke Verschlüsselung schliessen sich nicht aus, wie das Programm Truecrypt zeigte. Bundesamt für Sicherheit in der Informationstechnik: Sicherheitsanalyse TrueCrypt, Publikationen, Bundesamt für Sicherheit in der Informationstechnik, 16.11.2015, URL: <https://www.bsi.bund.de/DE/Publikationen/Studien/TrueCrypt/truecrypt.html>, Stand: 18.04.2016.

⁶⁸⁴ URL: <http://www.httrack.com/>.

⁶⁸⁵ Collins u. a.: *Going Digital*, 2015, S. 19.

⁶⁸⁶ Siehe: Margulies: *Digitale Daten als Quelle der Geschichtswissenschaft*, 2009, S. 375; Miller: *Metadata for Digital Collections*, 2011, S. 9.

Verschiedene Aufgaben können von Metadaten übernommen werden.⁶⁸⁷ die eindeutige Identifikation eines Objekts, eine Objektbeschreibung innerhalb eines Bestands, die Einbettung in einen Kontext, eine Unterstützung für das Wiederauffinden und für die Identifikation des Objekts, die Gliederung in Sammlungen innerhalb eines Repository, ein Authentizitätsnachweis mit Historiendokumentation, eine Unterstützung für die Sicherstellung der Objektintegrität, das Erfassen von Zugriffs-, Zugangs- und anderen Rechten, die Vereinfachung des Informationsaustausches und das Erfassen von technischen Parametern, die die Darstellungsweise des Objekts beschreiben. Ein Metadatum ist aber auch ein Objekt und deshalb Ausgangspunkt zur Sortierung von Objekten, denen es selbst angehört.⁶⁸⁸ Der Inhalt eines Metadatums muss nicht zwingend überprüfbar sein, wie vom Soziologen Dirk Baecker gefordert⁶⁸⁹, sondern kann auch zusätzliche Informationen zum beschriebenen Objekt enthalten, die sonst unauffindbar sind. Das bedeutet allerdings auch, dass Metadaten nicht nur rein deskriptiv sind, sondern auch interpretative Tendenzen aufweisen können. Beispielsweise können Objekte von individuellen Nutzern durch freie Verschlagwortung mit Tags markiert werden, die teilweise nur für den spezifischen Nutzer Sinn ergeben. Die Beziehung zwischen dem Buch *Es* von Stephen King und dem Tag *Madagaskar* kann beispielsweise erst dann nachvollzogen werden, wenn auch bekannt ist, dass der Nutzer das Buch auf einer Reise nach Madagaskar gelesen hat. Dieser Folksonomie⁶⁹⁰ entgegengesetzt ist die professionelle Verschlagwortung (Taxonomie) nach strengen Regeln durch qualifizierte Personen wie Mitarbeiter einer Bibliothek.

Für die quellenkritische Überprüfung digitaler Objekte spielen Metadaten eine zentrale Rolle.⁶⁹¹ Der Nutzer kann nur auf den Inhalt des Objekts zugreifen, wenn die in den Metadaten enthaltenen Informationen, die Dateiformatspezifikationen, korrekt und vollständig sind und vom Informationssystem interpretiert und dargestellt werden können. Anschliessend kann das Objekt mit allenfalls zusätzlich vorhandenen Kontextinformationen kritisch untersucht werden. Auf allen vier aufeinander aufbauenden Ebenen Informationssystem, Dateninterpretation, Inhalt und Kontext sind Metadaten notwendig, wobei die Darstellbarkeit des Objekts nicht seine Interpretierbarkeit garantiert. Deshalb ist die Sicherung eines Bitcodes

⁶⁸⁷ Nach: Gladney: Preserving Digital Information, 2007, S. 129.

⁶⁸⁸ Nach: Baecker, Dirk: Metadaten. Eine Annäherung an Big Data, in: Geiselberger, Heinrich; Moorstedt, Tobias (Hg.): Big Data: das neue Versprechen der Allwissenheit, Berlin 2013 (edition unseld), S. 165.

⁶⁸⁹ Ebd.

⁶⁹⁰ Folksonomy: Vergabe von freien, nicht kontrollierten Stichworten (Tags) durch Nutzer. Siehe auch: König, René; Nentwich, Michael: Cyberscience 2.0, in: kommunikation @ gesellschaft, 2014, S. 5 und 16.

⁶⁹¹ Siehe auch: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 377-383.

ohne Metadaten sinnlos, weil diese erst die Decodierung erlauben, wie das oben erwähnte Beispiel zur Viking Sonde zeigt.⁶⁹²

Für die Langzeitarchivierung sind regelmässige Datenformatmigrationen zurzeit noch notwendig. Die Änderungen werden protokolliert und die Informationen werden dem Nutzer zur Verfügung gestellt. Dieser kann somit nachvollziehen, warum der Prozess eingeleitet worden ist, welches die Eigenschaften des neuen Formats sind und welche Unterschiede zwischen dem Primär- und dem Sekundärobjekt zu erwarten sind. Die kontrollierte Datenformatmigration erlaubt es, ein Objekt kritisch zu beurteilen, weil dessen Überlieferungsgeschichte, zumindest in einem professionellen Archiv, nachvollzogen werden kann.

Metadaten über den Kontext vermitteln Informationen über inhaltliche und technische Beziehungen wie auch Abhängigkeiten zu anderen sich im Bestand befindenden Objekten. Solche Beziehungen, und vor allem der Grund oder die Absicht für deren Entstehung, die sogenannten Provenienzinformationen⁶⁹³, können wichtige Anhaltspunkte zur ursprünglichen Verwendung und Funktionalität des Objekts geben.⁶⁹⁴ Soll beispielsweise ein elektronisches Kunstwerk in Abwesenheit des Künstlers dargestellt werden, sind detaillierte Anleitungen nötig, um das Werk in der vom Künstler intendierten Weise aufzubauen und korrekt darzustellen. So stellen sich insbesondere bei Netzkunstwerken die Fragen, welche externen Ressourcen einzubinden sind und wie lange diese zur Verfügung stehen. Allenfalls ist auch zu prüfen, ob nur die vorhandenen oder ob neue oder aktualisierte Objekte einzubinden sind. Die Kunsthistorikerin und Medientheoretikerin Tabea Lurk spricht von der *Werklogik*, die zu identifizieren und festzuhalten ist, damit das Werk erneut dargestellt werden kann.⁶⁹⁵

Die für die Nutzung und das rasche Wiederauffinden, das sogenannte Retrieval, benötigten Metadaten werden oft in einer Datenbankanwendung gespeichert und darin mit dem entsprechenden Objekt verknüpft. Falls diese Verknüpfungen zwischen Metadaten und Objekten gelöst werden, werden beide Bestände meistens nahezu wertlos, weil damit für Forschende grundlegende Informationen darüber fehlen, unter welchen Bedingungen der Datenbestand genutzt werden kann.⁶⁹⁶ Aus Sicht der Langzeitarchivierung sollten Metadaten deshalb direkt ins digitale Objekt *integriert* werden. Aus diesem sollten die Metadaten

⁶⁹² Siehe: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 33.

⁶⁹³ Provenienzinformation: Grund oder Absicht für die Entstehung eines Bestandes.

⁶⁹⁴ Siehe: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 397.

⁶⁹⁵ Nach: Lurk, Tabea (Hg.): ComputerKunstGeschichte Schweiz, in: Computergeschichte Schweiz: eine Bestandesaufnahme = Histoire de l'ordinateur en Suisse: un état des lieux, Zürich 2009 (Geschichte und Informatik 17), S. 204.

⁶⁹⁶ Vgl.: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 411; Leonelli, Sabina: Data Interpretation in the Digital Age, in: Perspectives on Science 22 (3), 06.03.2014, S. 19. DOI: 10.1162/POSC_a_00140.

automatisch aus- und in die externe Datenbank eingelesen werden.⁶⁹⁷ Die Datenpflege kann somit auf das Objekt beschränkt werden und die Metadaten bleiben immer mit dem Objekt verknüpft. Es ist allerdings zu berücksichtigen, dass bei der Integration von Metadaten ins Objekt das Objekt selbst verändert wird und damit seine Integrität verliert.⁶⁹⁸ Je nach Objekt ist zu entscheiden, ob die ergänzenden Metadaten integriert werden dürfen und damit als *interne Metadaten* den Verlust der Integrität verursachen oder ob sie nur zusätzlich gesichert und mit dem Objekt verknüpft werden sollen (*externe Metadaten*). Speziell sind Metadaten, die nicht integriert werden können, wie das Metadatum *Hashwert*.⁶⁹⁹ Der Hashwert stellt eine Prüfsumme dar, mit der die Datenintegrität des digitalen Objekts festgestellt werden kann. Dazu wird aus dem Objekt durch einen Algorithmus wie dem SHA-3 in einem Programm wie HashCalc⁷⁰⁰ eine alphanumerische Ganzzahl errechnet, die sich bei jeder Mutation des Objekts verändert.⁷⁰¹ Wird ein Objekt mit dem Hashwert publiziert, kann es auf Manipulationen überprüft werden, indem der Nutzer vom verwendeten Objekt ebenfalls eine Prüfsumme erstellt⁷⁰² und anschliessend die beiden Prüfsummen vergleicht. Sind sie identisch, ist das Objekt integer, aber falls sie sich unterscheiden wurde mindestens eines verändert. Es kann allerdings keine Aussage darüber gemacht werden, was und in welchem Umfang manipuliert wurde. Möglicherweise wurde nur wenig verändert. Das beeinträchtigt zwar die Datenintegrität, aber die Informationsintegrität könnte weiterhin intakt sein.

Um die Verwendung von Metadaten zu vereinfachen, wurden *Metadatenstandards* definiert. Allgemeine Standards wie das Dublin Core Metadata Element Set (DCMS oder Dublin Core), The Metadata Object Description Schema (MODS), The Visual Resources Association Core Categories (VRA Core), Preservation Metadata: Implementation Strategies (PREMIS) und Metadata Encoding and Transmission Standard (METS) erfüllen viele der oben genannten Anforderungen und werden auch in der digitalen Langzeitarchivierung eingesetzt.⁷⁰³ Meta-

⁶⁹⁷ Bspw. erlaubt für Audioobjekte das MP3-Format die Integration von Metadaten, wohingegen das WAV (Waveform Audio File Format) diesen Vorgang nur bedingt unterstützt.

⁶⁹⁸ Vorausgesetzt, dass ein entsprechendes Informationssystem zur Verfügung steht, mit dem die Integration in das Objekt erfolgen kann.

⁶⁹⁹ Hashwert: Zusammengesetzt aus dem engl. 'hash' (zerkleinern) und dem dt. 'Wert'. Hash ist informationstechnisch eine Einwegfunktion (Hashfunktion), die jedem Wert aus einer Definitionsmenge eindeutig und nicht umkehrbar einen Wert aus einer deutlich kleineren Wertemenge zuordnet (bspw. Quersumme). Der Hashwert ist eine kryptisch generierte, alphanumerische Ganzzahl. Nach: Fischer; Hofer: Lexikon der Informatik, 2011. In der Informationstechnik werden Prüfsummen wie der Hashwert eingesetzt, um Übertragungs- und Speicherfehler zu eruieren. Prüfsummenverfahren zur Fehlerkorrektur (bspw. Check-Bit) werden eingesetzt, um fehlerbehafteten digitalen Code korrekt interpretieren zu können.

⁷⁰⁰ URL: <http://www.slavasoft.com/hashcalc/>.

⁷⁰¹ Zur Berechnung der Checksumme siehe bspw. Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 359-363. Zu Hashfunktionen siehe: Poguntke, Werner: Basiswissen IT-Sicherheit: Das Wichtigste für den Schutz von Systemen und Daten, Dortmund 2013 (Informatik), S. 91-96.

⁷⁰² Es ist nicht das gleiche Programm, sondern nur derselbe Algorithmus zu verwenden.

⁷⁰³ Vertiefende Einblicke in Metadaten (Metadatensysteme, Datenmodelle, Struktur, Kategorisierung Syntax etc.) können hier nicht geboten werden. Siehe dazu bspw.: Miller: Metadata for Digital Collections, 2011; Eden, Bradford Lee (Hg.): Twenty-first century metadata operations challenges,

datenstandards gewährleisten die Interoperabilität unter verschiedenen Systemen sowie eine gewisse Flexibilität in der Anzahl von Elementen und ihrer Detaillierungsgrad und Extensibilität, womit die Möglichkeit, weitere Elemente oder Abfragekriterien zu ergänzen, um spezifische Bedürfnisse der Anspruchsgruppe zu befriedigen, gemeint ist, um auch zukünftige Ansprüche an die gesammelten Objekte erfüllen zu können.⁷⁰⁴ Das bedeutet aber auch, dass immer wenn Standards aktualisiert werden, auch die bisherigen Metadaten mit zusätzlichen Metadaten über diese Änderungen zu versehen sind, um deren eigene Überlieferungsgeschichte zu dokumentieren.⁷⁰⁵

Die Standards bieten durch ihre festgelegten Parameter einerseits eine Hilfestellung für die Eingabe aller erforderlichen Informationen, können aber andererseits durch die vorgegebenen Felder die Dokumentation eines Objekts auch einschränken. Bei der Auswahl oder Weiterentwicklung des oder der Standards ist deshalb auch immer die informationstechnische Weiterentwicklung mitzuverfolgen und falls nötig sind entsprechende Anpassungen im Schema vorzunehmen. Diese Anpassungen sind dabei von allen Nutzern zu übernehmen, individuelle Entwicklungen sind zu unterbinden.⁷⁰⁶

Sowohl der Historiker und jetzige Vize-Direktor der Europeana, Harry Verwayen, als auch die Informationswissenschaftlerin Maura Valentino haben aufgezeigt, dass die Einigung der verschiedenen, interessierten Partner auf einen oder wenige Standards und die Publikation von Metadaten für alle Beteiligten von grossem Vorteil ist.⁷⁰⁷ Valentino kommt zum Schluss, dass Universitätsbibliotheken mit möglichst wenigen Standards digitale Sammlungen entwickeln können, die den Bedürfnissen von Forschern und der Öffentlichkeit viel besser entsprechen.⁷⁰⁸ Deshalb ist es nötig, dass jeder Wissenschaftsbereich seine eigenen Anforderungen an einen Metadatenstandard definiert und dass anschliessend ein oder möglichst wenige Standardformat(e) für den Austausch von Metadaten unter allen Bereichen

opportunities, directions, London 2012; Haynes, David: Metadata for information management and retrieval, London 2004; Verwayen, Harry; Arnoldus, Martijn; Kaufman, Peter B.: The Problem of the Yellow Milkmaid, White Paper 2, Europeana, 2011. URL: <http://pro.europeana.eu/publications>, Stand: 17.03.2014; Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 399-416.

⁷⁰⁴ Siehe: Taylor; Joudrey: The organization of information, 2009, S. 92 f.

⁷⁰⁵ Zur rekursiven Abhängigkeit von Datenbeschreibungen (Metadaten) siehe: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 382 f.

⁷⁰⁶ Bspw. wird die individuelle Anpassung der Guidelines for Electronic Text Encoding and Interchange der Text Encoding Initiative gefördert, was zwar die Anforderungen der einzelnen Institution erfüllen hilft, aber den Datenaustausch erschwert. Siehe: Text Encoding Initiative: TEI: Customization, Guidelines, TEI, 07.04.2015, URL: <http://www.tei-c.org/Guidelines/Customization/>, Stand: 09.04.2015. Dem entgegengesetzt heben Collins et al. TEI als gutes Beispiel für Standardisierung auf internationaler Ebene hervor: Collins u. a.: Going Digital, 2015, S. 19.

⁷⁰⁷ Siehe dazu: Verwayen; Arnoldus; Kaufman: Milkmaid, 2011; Leon, Sam: The Revenge of the Yellow Milkmaid: Cultural Heritage Institutions open up dataset of 20m+ items, Open Knowledge Foundation, 17.09.2012, URL: <http://blog.okfn.org/2012/09/17/the-revenge-of-the-yellow-milkmaid-cultural-heritage-institutions-open-up-dataset-of-20m-items/>, Stand: 14.08.2014; Valentino, Maura L.: Integrating Metadata Creation into Catalog Workflow, in: Eden, Bradford Lee (Hg.): Twenty-first century metadata operations challenges, opportunities, directions, London 2012, S. 40-49.

⁷⁰⁸ Valentino: Integrating Metadata Creation into Catalog Workflow, 2012, S. 48.

festgelegt wird beziehungsweise werden.⁷⁰⁹ Die Partner im wissenschaftlichen Archivierungsprozess sollten sich deshalb auf möglichst wenige und flexible Standards einigen, die den reibungslosen Austausch unter den Gedächtnisinstitutionen und Forschenden garantieren. Diese Standards sind von den Forschenden, Institutionen und den wissenschaftlichen Publikationsorganen, insbesondere den Repositories, zwingend einzuhalten.

Unabhängig vom Metadatenstandard und von der Integration und/oder Verknüpfung der Metadaten eines digitalen Objekts ist festzulegen, was die Elemente eines Standards, das Metadatenset, enthalten. Der Informationswissenschaftler Stephen Miller hat unter Verwendung bestehender Metadaten schemata eine Anleitung zu *Metadata for Digital Collections* erstellt.⁷¹⁰ Weil aber digitale Objekte andere Eigenschaften als bisher bekannte Objekte haben, wäre die Definition von Elementen unabhängig von bestehenden Metadaten sets sinnvoller. Erst anschliessend wäre zu überprüfen, ob ein bestehendes Schema verwendet werden könnte oder ob ein oder mehrere Schema(ta) ergänzt werden sollten. Beispielsweise wären bei Hilfssicherungen Angaben über das Primärobjekt wichtig, ebenso über den Grund, warum ein Sekundärobjekt in der vorliegenden Form gesichert wurde. Anhand dieser Informationen könnte das Ziel der Sicherung und weshalb das spezifische Format verwendet wurde, nachvollzogen werden. Beispielsweise ist in einem Online-Artikel nur der Text und nicht dessen Darstellung relevant, weshalb dieser in PDF/A und nicht in HTML gesichert wurde. Besonders bei digitalen dynamischen oder graphischen Objekten ist es eminent wichtig, dass Informationen über allfällig erfolgte Darstellungsänderungen festgehalten werden, ansonsten die Verwendbarkeit und den Aussagewert des Objekts gemindert werden.

Solche Informationen können zwar in den meisten bestehenden Standards in einem beschreibenden Textelement verzeichnet werden, aber es wären spezifische Elemente sinnvoll, um das automatische Auslesen der Metadaten und einen einfachen, ebenfalls automatischen Datenaustausch unter verschiedenen Plattformen und Institutionen sicherzustellen. Hier kann aufgrund der Fragestellung keine tiefgreifende Untersuchung zur Entwicklung eines neuen oder angepassten Metadaten schemas erfolgen. Ich rege aber an, dass die Aufnahme von spezifischen Elementen für digitale Objekte in ein Metadaten schema diskutiert wird. Nachfolgend eine nicht abschliessende Liste der zu diskutierenden Elemente mit einer kurzen Beschreibung:

⁷⁰⁹ Damian Hess schlägt mit dem Digital Asset Management DAM ein flexibles Metadatenmodell für digitale Sammlungen vor, mit dem die Anzahl der verwendeten Standards reduziert werden könnte. Bisher wurde der Ansatz noch nicht breit diskutiert und kann auch hier nicht ausführlich dargestellt werden. Siehe: Hess, Demian: DAM and the Need for Flexible Metadata Models, Digital Asset Management, Avalon Consulting, LLC – Blogs, 17.11.2014, URL: <http://blogs.avalonconsult.com/blog/generic/dam-and-the-need-for-flexible-metadata-models/>, Stand: 21.08.2015.

⁷¹⁰ Siehe: Miller: Metadata for Digital Collections, 2011.

Name	Beschreibung
identifier	Persistenter Identifikator ⁷¹¹
creator(object)	Urheber des digitalen Objekts
creator(collection)	Urheber der Sammlung, in die das Objekt aufgenommen wurde
date(creation)	Angabe des Entstehungszeitpunkts resp. der letzten Manipulation des Objekts
date(collection)	Zeitpunkt der Integration in eine Sammlung – Zeitstempel der TSP-Software
date(publication)	Datum der Erstpublikation eines Beitrags (bspw. in Sammelbänden, auf anderen Webseiten)
description(object)	Beschreibung des Objekts
description(collection)	Beschreibung der Sammlung, mit der das Objekt verknüpft wurde
digital.type	Digitalisat, genuin digital oder reborn-digital
hash	Hashwert des Objekts mit der Angabe des Berechnungsalgorithmus
history	Angaben über Änderungen am Objekt (Migrationen etc.)
infosys	Angaben über das gesamte Informationssystem (Hard- und Software), mit dem das Objekt erstellt wurde
primary.secondary	Angabe, ob das Objekt ein Primär- oder Sekundärobjekt ist (im intellektuellen Sinn)
tags	Schlüsselwörter für das Wiederauffinden oder die Sammlungsordnung
...	...

Tabelle 5: Beispiele von spezifischen Elementen für digitale Objekte zur Ergänzung eines Metadatenschemas (Föhr)

Wenn solche Metadaten und die Information über den Metadatenstandard zusammen mit einem digitalen Objekt zur Verfügung gestellt werden, kann eine TSP-Software die Informationen automatisch auslesen und in ein Verzeichnis übertragen. Elemente wie creator(object), creator(collection), date(object) und date(collection) sollten vor Manipulation durch den Nutzer geschützt werden, alle anderen sind vom Nutzer zu überprüfen und gegebenenfalls zu aktualisieren, zu ergänzen oder zu löschen. Beim Abschluss des Projekts sollten die Metadaten in die zugehörigen Objekte integriert werden, bevor keine Manipulation an diesen mehr möglich ist. Die TSP-Software wäre um entsprechende Module zu ergänzen. Erst anschließend sollten die einzelnen Objekte veröffentlicht werden.

⁷¹¹ Siehe dazu Lösungsansatz Persistentes Identifikations- und Adressierungssystem im TDR in Kap. 5.4.3. Persistenz.

Das automatisierte Auslesen von Metadaten ist beispielsweise bei vielen Literaturverwaltungssoftwares integriert. Damit soll dem Nutzer die Arbeit erleichtert und der Import von möglichst korrekten und einheitlichen Metadaten ermöglicht werden. Letzteres ist zurzeit noch relativ schwierig, weil bei vielen digitalen Objekten keine Metadaten zur Verfügung stehen oder diese nicht in einer standardisierten Form vorliegen. Die Software muss die Informationen aus verschiedenen Quellen zusammensuchen, wie durch das Auslesen der Standardmetadatenformate aus dem Objekt selbst, durch Abfrage von Datenbanken wie WorldCat oder Scannen eines Textes mittels OCR-Modulen.⁷¹² Keines der Verfahren liefert durchgehend befriedigende Ergebnisse und der Nutzer muss die eingelesenen Metadaten oft korrigieren oder ergänzen. Beim Import von Webseiten beispielsweise werden die Metadaten aus dem im HTML Code hinterlegten Tag Haeder ausgelesen, was zwar bei einer Seite wie der NZZ gute Resultate liefert, bei der Seite eines Vereins und auch bei einem Workshopbericht auf einer Universitätswebseite aber kaum verwendbare Resultate ergibt.

```
<!DOCTYPE html>
<html class=" js" lang="de" xmlns:og="http://ogp.me/ns#" xmlns:fb="http://www.facebook.com/2008/fbml">
  <head>
    <meta content="Neue Gastkommentar zum Schweizer «Historikerstreit», Nachrichten Gastkommentar zum Schweizer «Historikerstreit», Debatte" name="Keywords">
    <meta content="Gastkommentar zum Schweizer «Historikerstreit" name="Description">
    <meta content="Neue Zürcher Zeitung AG, Switzerland" name="Publisher">
    <meta content="online.redaktion@nzz.ch" name="email">
    <meta content="Copyright ©copy; 2015 Neue Zürcher Zeitung AG, Switzerland, http://www.nzz.ch/agb" name="copyright">
    <meta content="de" http-equiv="content-language">
    <meta content="-1" http-equiv="expires">
    <meta content="Impressum Neue Zürcher Zeitung, Switzerland, http://www.nzz.ch/impressum" name="Author">
    <meta content="global" name="distribution">
    <meta content="index, follow, nosrchive, noodp" name="robots">
    <meta content="1 hour" name="revisit-after">
    <meta content="no-cache" http-equiv="pragma">
    <meta content="138008459595436" property="fb:app_id">
    <meta content="Gastkommentar zum Schweizer «Historikerstreit»: Die Debatte über die Schweizer Geschichte wird fahrlässig" property="og:title">
    <meta content="Gastkommentar zum Schweizer «Historikerstreit" property="og:description">
```

Abbildung 16: Metadaten zum Artikel 'Die Debatte über die Schweizer Geschichte wird fahrlässig' auf der Webseite der NZZ (NZZ Online)⁷¹³

```
<!DOCTYPE html>
<html id="yui3_17_2_1_1524170946470_418" class="yui3-js-enabled js flexbox canvas canvastext webgl no-touch _active wf-futurapt-n8-active wf-futurapt-i8-active wf-active" xmlns:og="http://opengraphprotocol.org/schema/" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:website="http://ogp.me/ns/website" itemscope="" itemtype="http://schema.org/WebPage" style="" lang="de-DE">
  <div id="yui3-css-stamp" style="position: absolute !important; visibility: hidden !important"></div>
  <script async="" src="//www.google-analytics.com/analytics.js"></script>
  <script type="text/javascript"></script>
```

Abbildung 17: Metadaten zur Homepage der AB Glanzenburger (AB Glanzenburger)⁷¹⁴

```
<!DOCTYPE html>
<html class=" js " lang="de-DE">
  <head>
    <meta charset="utf-8">
    <base href="https://europa.unibas.ch/">
    <link type="image/x-icon; charset=binary" href="https://europa.unibas.ch/fileadmin/webbaukasten/templates/images/favicon.ico" rel="shortcut icon">
    <link type="image/x-icon; charset=binary" href="https://europa.unibas.ch/fileadmin/webbaukasten/templates/images/favicon.ico" rel="icon">
    <title>Europainstitut: Workshop: Digital Humanities for European Global Studies</title>
    <meta content="TYPO3 4.5 CMS" name="generator">
    <meta content="7 days" name="revisit-after">
    <meta content="index, follow" name="robots">
    <meta content="2015/04/11" name="lastmodified">
    <meta content="IE=edge, chrome=1" name="X-UA-Compatible">
    <meta content="width=974" name="viewport">
```

Abbildung 18: Metadaten zur Webseite des Workshops DH for European Global Studies (Universität Basel)⁷¹⁵

⁷¹² OCR: Optical Character Recognition. Überbegriff für optisches Erkennen von (Klar-)Schriften. Nach: Fischer; Hofer: Lexikon der Informatik, 2011.

⁷¹³ Sourcecode zitiert nach URL: <http://www.nzz.ch/meinung/debatte/falsche-faehrten-1.18519066>, Stand: 14.04.2015.

⁷¹⁴ Sourcecode zitiert nach URL: <http://www.glanzenburger.ch/>, Stand: 19.04.2018.

Besonders bei Publikationsplattformen von wissenschaftlichen Institutionen sowie wissenschaftlichen und wissenschaftsnahen Publikationsgesellschaften ist das Fehlen von standardisierten und vollständigen Metadaten ein Ärgernis. Würden sich diese Institutionen auf einen bestimmten Metadatenstandard⁷¹⁶ einigen und die Verwendung mit geeigneten Mitteln wie Records Management Systemen sicherstellen, könnte der Informationsaustausch unter den Institutionen vereinfacht und die Authentizität des einzelnen Objekts erhöht werden. Gerade diese Institutionen würden von einer konsequenten Anwendung und Publikation von standardisierten Metadaten profitieren.

Bei physischen Publikationen hat die konsequente Anwendung von Paratexten viele Vorteile gebracht. Bei einem Impressum eines Buches beispielsweise stellt der Verlag sicher, dass die dort angegebenen Informationen der Wahrheit entsprechen. Anhand dieser können die Auflage des Buches und all seine Kopien eindeutig bestimmt werden. Das Impressum ist unveränderlich, was das einfache Recherchieren und das eindeutige Zitieren ermöglicht. Das Impressum eines digitalen Objekts ist hingegen veränderlich und die Inhalte sind nicht zwingend eindeutig. Digitale Metadaten können jedoch im Gegensatz zu physischen automatisch ausgelesen, verarbeitet und verbreitet werden. Damit kann die Distribution einer Publikation und deren Registrierung in Bibliotheks- oder anderen Verzeichnissen schneller, individueller und kostengünstiger erfolgen.

Lösungsansatz 'Metadata Information Section'

Um den Austausch verlässlicher Metadaten zu fördern und damit insbesondere die Authentizität von wissenschaftlichen digitalen Objekten zu erhöhen, wird die Einführung einer *Metadata Information Section (MIS)* vorgeschlagen, die alle relevanten Metadaten enthält und durch entsprechende TSP-Software automatisch ausgelesen werden kann.⁷¹⁷ Die MIS kann sowohl als 'Mehr an besonderer Informationen' zur Reduktion des Information Overload aufgefasst werden, wie es David Weinberger und Klaus Merten fordern⁷¹⁸, als auch als Paratext, der zusätzliche Informationen über das Objekt liefert.

Ich unterscheide zwei Arten der MIS: Die eine ist in das digitale Objekt integriert, die andere wird auf der Publikationsplattform des Objekts als ergänzende Information zur Verfügung gestellt.

⁷¹⁵ Sourcecode zitiert nach URL: <https://europa.unibas.ch/aktuelles/nachrichtendetails/article/workshop-digital-humanities-for-european-global-studies/>, 14.04.2015.

⁷¹⁶ Oder möglichst wenigen, untereinander kompatiblen Metadatenstandards.

⁷¹⁷ Vergleichbar mit den Hinweisen für die Zitierung eines Online-Artikels. Siehe bspw. Zitierungsbox auf 'British History Online' mit verschiedenen Formaten für jeden Artikel. URL: <http://www.british-history.ac.uk/>.

⁷¹⁸ Siehe Kap. 5.1.3. Suchtechniken.

Die *im Objekt integrierte MIS* wird bei Dokumentenformaten wie PDF verwendet und wird zusätzlich zu den versteckt im Objektformat enthaltenen Metadaten am Schluss des Objekts im Klartext dargestellt. Die MIS kann aus Platzgründen auf bibliographische und identifizierende Informationen beschränkt werden, die aber zwingend vorhanden sein sollten (MUSS-Elemente). Es ist beispielsweise nicht sinnvoll, das Transkript eines Manuskripts in die MIS zu integrieren. Deshalb sollte auch die Zeichenzahl in bestimmten Elementen wie der description sinnvoll beschränkt werden.

Die *das Objekt ergänzende MIS* enthält alle Informationen der im Objekt integrierten MIS. Ergänzt werden Informationen wie Sicherheitsmerkmale, die mit dem Objekt nur verknüpft, aber nicht in dasselbe integriert werden können. Beispielsweise kann der Hashwert auf Webseiten, in Repositories oder in Datenbanken nur zusätzlich zum Objekt und mit diesem verknüpft zur Verfügung gestellt werden.

Die Anwendung der beiden MIS veranschauliche ich anhand des Posters *Historische Quellenkritik im Digitalen Zeitalter*. Die Metadaten der integrierten MIS (ausgewählte Elemente) können im Dokument 'Poster – MIS.pdf' wie folgt dargestellt werden:

creator(object): Föhr, Pascal; date(collection): 2015-01-15; language: german; title: Poster: Historische Quellenkritik im Digitalen Zeitalter; description: Das Poster wurde anlässlich des Workshops Digital Humanities for European Global Studies, Basel, 27.03.2015 präsentiert.; ...
--

Historische Quellenkritik im Digitalen Zeitalter

Prozess der Objektbeurteilung:



Wie verändert sich die historisch-kritische Methode durch die Verwendung von digitalen Objekten als Forschungsressource sowie von digitalen Informations- und Kommunikationsmedien?

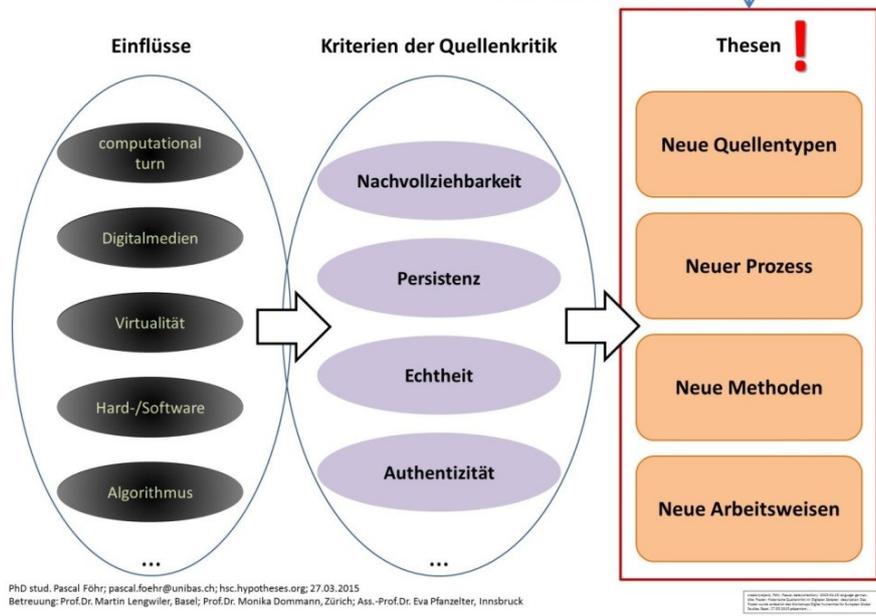


Abbildung 19: Poster 'Historische Quellenkritik im Digitalen Zeitalter' mit Metadata Information Section (Föhr, 15.01.2015)



Abbildung 20: Vergrößerung der MIS im Poster (Föhr)

Die auf einer Webseite zu publizierende, ergänzende MIS würde wie folgt lauten (Ergänzung in Fettschrift):

```
creator(object): Föhr, Pascal; date(collection): 2015-01-15; language: german; title: Poster: Historische Quellenkritik im Digitalen Zeitalter; description: Das Poster wurde anlässlich des Workshops Digital Humanities for European Global Studies, Basel, 27.03.2015 präsentiert.; ...  
hash - MD5: 7c28f3885ace23985dfa03c2c0aa13db
```

Abbildung 21: Ergänzende MIS auf der Webseite (Föhr)

Die wichtigsten Metadaten könnten so automatisch mit einer TSP in einem TDR gesichert werden. Dies wäre ein wichtiger Schritt für den Austausch von verlässlichen, wissenschaftlichen Informationen. Für die Erstellung der MIS könnten Autoren eines digitalen Objekts mit entsprechender Software unterstützt werden. Ein Publishingtool⁷¹⁹ für Blogs oder Online-Artikel könnte mit einem entsprechenden Modul erweitert werden, das die Eingabe der geforderten Informationen für die MIS erzwingt und die Publikation des Objekts nur mit einer vollständigen MIS ermöglicht. Bei der Erstellung von Dokumenten könnte ein -Add-on in der Textverarbeitung die MIS ebenfalls automatisch in das Dokument einfügen und fehlende Angaben monieren. Wenn das mit einer MIS versehene Objekt auf einer Fachplattform publiziert wird, können die Metadaten automatisch ausgelesen und mit entsprechenden Informationen wie dem MD5-Hashwert automatisch ergänzt werden. Der Aufwand eines Autors für die Erstellung einer MIS wäre damit relativ gering.

Es wird für wissenschaftliche und wissenschaftsnaher Publikationsformen (digital und analog) die Einführung einer **Metadata Information Section (MIS)** vorgeschlagen, die alle mit dem Objekt verknüpften Metadaten zur Verfügung stellt. Für die Erstellung und Verwendung der MIS sollten entsprechende Softwares entwickelt werden, die eine automatisierte Verarbeitung sicherstellen.

Mit den Zotero-COinS besteht bereits ein System, mit dem bibliographische Angaben auf einer Webseite hinterlegt werden können.⁷²⁰ Allerdings werden die Informationen im HTML Code zwischen den Tags URL: span und URL: /span hinterlegt und damit vom Browser nicht als Text interpretiert. Es ist jedoch zusätzlich notwendig, dass der Nutzer die Angaben als Klartext vorfindet, um sie überprüfen zu können. Zudem korrespondiert COinS (noch) nicht mit den in einem Objekt integrierten Informationen.

Das Erstellen, Einfügen und Pflegen einer MIS sollte als Pflicht des wissenschaftlichen Arbeitens aufgefasst werden, denn es werden damit nicht nur die verwendeten Quellen nachgewiesen, sondern auch die zugehörigen Metadaten zur Verwendung durch Dritte bereitgestellt. Insbesondere im wissenschaftlichen Umfeld, das verlässliche Metadaten verlangt,

⁷¹⁹ Bspw. ein Content Management System (CMS), eine Software zur Verwaltung und Publikation von (Online-)Inhalten. Siehe: Prevezanos: Computer Lexikon 2013, 2012.

⁷²⁰ Siehe URL: https://www.zotero.org/support/dev/exposing_metadata/coins, Stand: 05.06.2018.

wäre eine MIS sehr vorteilhaft. Der für Forschende zur Erstellung nötige, relativ kleine Aufwand lässt sich durch den grossen Informationsgewinn rechtfertigen und mit informationstechnischer Unterstützung minimieren. Die Verwendung einer MIS und zusätzlicher Sicherheitsmerkmale erhöht die Authentizität eines Objekts und kann damit auch die Reputation der Autorschaft steigern. Was jedoch auch klar strukturierte und standardisierte Metadaten nicht verhindern können, ist die je nach Erfahrung und Forschungsfrage individuelle Interpretation der vorhandenen Informationen durch den einzelnen Forschenden.⁷²¹

Zum Schluss sei noch die Frage gestellt, was mit all den angesammelten Metadaten geschieht? Werden sie tatsächlich ausgewertet oder sind wir schon bei 'too much metadata' angelangt, wie die Digitalbyzantinistin Tara Andrews bemerkte?⁷²² Um die Flut an Metadaten zu kontrollieren, ist die Verwendung von klar definierten, standardisierten und kompatiblen Metadatenschemata sowie von Software als Hilfsmittel zur korrekten und vollständigen Darstellung der Metadaten nötig. Nur damit lassen sich die Informationen filtern und sinnvoll nutzen. Literaturverwaltungs-, Archiv- und Bibliothekssoftware sollten deshalb auch alle von den Schemata verwendeten Elemente enthalten und zuordnen können.⁷²³ Die Geschichtswissenschaft ist auf die Verwendung von Software zur Bewältigung der Informationsflut angewiesen, weshalb ein Grundverständnis über Informationstechnik eine Voraussetzung für korrektes wissenschaftliches Arbeiten sein muss.

5.3.8. Zusammenfassung

Die Quellensicherung von digitalen Forschungsressourcen ist ein wichtiger Schritt der historisch-kritischen Methode, weil digitale Objekte grundsätzlich volatil sind und manipuliert werden können. Um eine digitale Forschungsressource wissenschaftlich nutzen zu können, ist diese in einer verlässlichen, manipulationssicheren Forschungsinfrastruktur zu sichern, damit sie auch Dritten zur kritischen Beurteilung zur Verfügung gestellt werden können.

Der Prozess der traditionellen historischen Methode wird um den Schritt *Quellensicherung* erweitert, um den Zugriff auf eine digitale Forschungsressource sowie deren Verfügbarkeit dauerhaft zu gewährleisten.

⁷²¹ Siehe dazu bspw.: Leonelli: *Data Interpretation in the Digital Age*, 2014, S. 15.

⁷²² Nach Tara Andrews im Workshop *Digital Humanities for European Global Studies*, Basel, 27.03.2015. Persönliche Webseite siehe URL: <http://www.digitalbyzantinist.org/>.

⁷²³ Wenn mehrere Standardschemata Verwendung finden, sollten diese von der Software einheitlich dargestellt und auch in das vom Nutzer gewählte Format exportiert werden können. D.h. die Software sollte alle Elemente aller Standards berücksichtigen und entsprechend darstellen und umformatieren (sic!) können.

In diesem Prozessschritt sind folgende Fragen zu klären: Welches Objekt wird gesichert? In welcher Form? In welchem Format? Mit welchen Hard- und Softwarekomponenten? An welchem Speicherort? Mit welchen standardisierten Metadaten?

Zuerst ist festzulegen, welche(s) Objekt(e) zu sichern sind. Die Auswahl hat alle voneinander abhängigen Objekte zu umfassen, die für eine erneute, korrekte Darstellung benötigt werden, was besonders bei reborn-digitalen Objekten problematisch sein kann. Deshalb sollte bei Webseiten auch der Quellcode überprüft werden, um eventuelle Abhängigkeiten feststellen zu können. Der Einsatz von technischen Hilfsmitteln kann die Anfertigung einer nachvollziehbaren Dokumentation von dynamischen Objekten erleichtern.

In einem weiteren Schritt ist festzulegen, in welcher Form und in welchem Format oder gegebenenfalls in welchen Formaten das Objekt gesichert werden soll. Grundsätzlich ist das Objekt in der ursprünglichen Form und im ursprünglichen Format zu sichern. Wenn jedoch kein Zugriff auf seine archetypische Form möglich ist, oder wenn das Dateiformat des Objekts für eine spätere Wiederverwendung eventuell nicht geeignet ist, sollte eine Hilfssicherung erfolgen. Dabei entsteht ein vom Primärobjekt, so wie es archetypisch vorliegt und dargestellt wird, unterschiedliches und eigenes, reborn-digitales Sekundärobjekt. Dieses Surrogat kann in seiner Funktionalität eingeschränkt sein. Bei hybriden Objekten ist zu entscheiden, welche Sicherungsmassnahmen zu ergreifen sind, das heisst, ob es weiterhin hybrid bestehen bleibt, ob die physischen Elemente digitalisiert oder ob allenfalls die digitalen Objekte 'materialisiert' werden. Es ist zu beachten, dass für die Langzeitarchivierung standardisierte und quelloffene Formate gewählt werden sollten, die bei einer Konversion möglichst alle Eigenschaften des Ursprungformats übernehmen, um die Vollständigkeit des Inhalts und der Funktionalität zu gewährleisten. Dem Alterungsprozess von Objekten ist durch eine systematische Kombination von Datenträger- und Dateiformatmigration entgegenzuwirken.

Des Weiteren ist zu untersuchen, ob Hard- und/oder Softwarekomponenten mitgesichert werden sollten, damit das digitale Objekt wieder in seiner archetypischen Form dargestellt werden kann. Diese sind allenfalls separat zu archivieren oder zu emulieren. Bei der Anwendung von Emulatoren ist zu überprüfen, ob das digitale Objekt tatsächlich korrekt dargestellt wird und seine Funktionalität behält. Da Emulatoren selber einem Alterungsprozess unterliegen, sind sie in regelmässigen, kurzen Abständen anzupassen.

Der Sicherungsprozess eines digitalen Objekts sollte unabhängig vom Nutzer durchgeführt werden können, um Manipulationen auszuschliessen und den Zugriff auf das Objekt langfristig sicherzustellen. Es sollten deshalb extern-institutionelle Forschungsinfrastrukturen, die sogenannten *Trusted Digital Repositories (TDR)*, für die Langzeitarchivierung digitaler Objekte aufgebaut werden.

Mit dem *Researcher Driven Archiving (RDA)* könnte die individuelle Sicherung digitaler Objekte in TDR ermöglicht werden. Dafür sollte eine nutzerfreundliche Software entwickelt werden, die jedes Objekt durch eine *Trusted Saving Procedure (TSP)* in einem TDR eines verlässlichen Providers sichert. Wird die digitale Forschungsressource als Quelle verwendet, kann sie unter einer/einem eindeutigen und dauerhaften URI/URL veröffentlicht werden.⁷²⁴ Jede nachfolgende Änderung oder Ergänzung wird protokolliert und ist von Dritten einsehbar. Der Provider ist per gesetzlichem Auftrag zu verpflichten, die Langzeitarchivierung sicherzustellen. Alternativ oder zusätzlich sollte eine Software entwickelt werden, die dem Forschenden durch einen TSP die vertrauenswürdige Sicherung seiner Forschungsressourcen in einem lokalen *Personal Trusted Archive (PTA)* erlaubt.

Schliesslich ist das gesicherte Objekt mit *standardisierten Metadaten* zu versehen, die sowohl durch den Sicherungsprozess entstehen, als auch vom Nutzer nachträglich hinzugefügt werden können. Es wird die Einführung einer *Metadata Information Section (MIS)* vorgeschlagen, die in alle digitale und analoge Objekte *integriert* wird, alle relevanten bibliographischen Metadaten enthält und automatisch ausgelesen werden kann. Bestimmte Metadaten wie Zeitstempel oder Provenienz sollten vom Nutzer nicht verändert werden können. Eine ergänzende MIS enthält zusätzliche Informationen wie Sicherheitsmerkmale, die nicht ins Objekt integriert werden können, und wird zusammen mit dem Objekt und mit diesem verknüpft publiziert. Die MIS sollte in alle wissenschaftlichen Publikationen und Publikationsplattformen implementiert werden, um den wissenschaftlichen Austausch mit verlässlichen Metadaten zu fördern.

5.4. Quellenkritik

Die quellenkritische Überprüfung bedeutet die Prüfung der Authentizität einer Quelle und beinhaltet das Erfassen und Erklären der besonderen räumlichen, zeitlichen, subjektiven und formalen Bedingungen einer Quelle sowie die Prüfung ihres Wahrheitsgehalts und ihrer Glaubwürdigkeit.⁷²⁵ Wie ich bereits im dritten Kapitel Traditionelle Methode und digitale Objekte festgestellt habe, bleiben die in der traditionellen historischen Quellenkritik an ein Objekt zu stellenden Fragen und die Grundlagen der historisch-kritischen Methode dieselben.⁷²⁶ Allerdings ist die Authentizität digitaler Objekte aufgrund ihrer spezifischen Eigenschaften mit Verfahren aus der Informationstechnik zu untersuchen, die in den quellenkritischen Untersuchungsprozess zu integrieren sind. Die historisch-kritische Methode verändert sich durch die Verwendung von digitalen Objekten als Forschungsressource sowie

⁷²⁴ Rechtliche Einschränkungen werden hier nicht beachtet.

⁷²⁵ Aus Kap. 3.1.3. Quellenkritik und nach: Fuchs; Raab: Quellenkritik, 2002, S. 653.

⁷²⁶ Im Folgenden werden diese Grundlagen vorausgesetzt und nur auf die spezifischen Probleme und Lösungsansätze im Umgang mit digitalen Objekten eingegangen.

von digitalen Informations- und Kommunikationsmedien sowohl auf methodologischer, als auch auf intellektueller Ebene. Trotz der bei einer Untersuchung von grossen digitalen Datenbeständen, Schlagwort 'Big Data', entstehenden methodologischen Fragen wird auf eine vertiefte Erläuterung derselben verzichtet, weil die Problematik im Kontext der algorithmischen Auswertungen grundsätzlich besprochen, weil das Thema bereits von anderen Autoren eingängig untersucht wird und es nicht Kernthema dieser Arbeit ist.⁷²⁷

Als Erstes ist der Begriff der Authentizität zu klären, weil dieser in der Geschichtswissenschaft und in der Informationstechnik unterschiedlich definiert wird. Anschliessend ist festzuhalten, was unter Integrität eines Objekts zu verstehen ist, da sie ein wichtiges Kriterium zur Beurteilung der Authentizität darstellt. Neben Authentizität und Integrität wird als drittes Grundkonzept die Persistenz erläutert, die vor allem für die Nachvollziehbarkeit einer Quelle eine grosse Rolle spielt.

Antworten auf die Fragen einer quellenkritischen Untersuchung können nur gefunden werden, wenn Hinweise aus verschiedenen Perspektiven zusammengefügt werden. Beispielsweise sind für die Bestimmung des Autors eines digital vorliegenden Textes Informationen aus dem Dateinamen, dem Textinhalt und den Metadaten wie auch Informationen über die Beziehungen zu anderen Dokumenten wertvoll. Die traditionelle Unterscheidung zwischen der inneren und der äusseren Quellenkritik ist nicht mehr sinnvoll, weil diese ineinander übergehen und deren Fragen nur übergreifend beantwortet werden können. Dieser Ansicht ist beispielsweise auch der Digitalhistoriker Andreas Fickers:

*"[...] in order to offer a historical interpretation of the function, role or importance of a specific source based on a specific historical question, 'internal' and 'external' source criticism need to go hand in hand."*⁷²⁸

Im quellenkritischen Prozess sollten deshalb nicht einzelne, einer inneren und äusseren Quellenkritik zugeordnete Fragen gestellt werden. Besser wäre die Definition von *Fragekomplexen*, die durch eine gesamthafte Untersuchung des Objekts beantwortet werden können. Die nach der Klärung von Authentizität, Integrität und Persistenz folgenden Unterkapitel sind nach solchen Fragekomplexen aufgebaut. Diese werden im quellenkritischen Prozess nicht sequenziell, sondern iterativ untersucht, weil sie sich gegenseitig ergänzen und Erkenntnisse aus einem Komplex auch neue Fragen in einem anderen aufwerfen können. Innerhalb dieser

⁷²⁷ Siehe weiterführend bspw.: Reichert (Hg.): Big Data, 2014; Geiselberger, Heinrich; Moorstedt, Tobias (Hg.): Big Data: das neue Versprechen der Allwissenheit, Berlin 2013 (edition unseld); Strasser, Bruno J.; Edwards, Paul N.: Big data is the answer... but what is the question?, unveröffentlichtes Paper [zugestellt per E-Mail am 18.10.2015], infoclio.ch-Tagung: Daten und Geschichtswissenschaften, Bern 16.10.2015; boyd, danah; Crawford, Kate: Six Provocations for Big Data, SSRN Scholarly Paper ID 1926431, Social Science Research Network, Rochester, NY 21.09.2011, S. 17. URL: <http://papers.ssrn.com/abstract=1926431>, Stand: 24.02.2015; Burkhardt: Digitale Datenbanken, 2015.

⁷²⁸ Fickers, Andreas: Towards A New Digital Historicism? Doing History In The Age Of Abundance, in: Journal Of European Television History & Culture (1), 2012, S. 25.

Fragekomplexe können digitale Objekte nur exemplarisch untersucht werden, da sie in sehr unterschiedlichen Formen und Formaten vorkommen. Zudem entstehen laufend neue Objektarten und bestehende werden weiterentwickelt oder eingestellt.

5.4.1. Historische versus informationstechnische Authentizität

*"Authentizität ist ein Schlüsselbegriff im Umgang mit Geschichte"*⁷²⁹
(Pirker und Rüdiger, 2010)

Es ist zu unterscheiden zwischen der historischen und der informationstechnischen Authentizität. Wann ein Objekt als historisch authentisch bezeichnet werden kann, ist nicht genau definiert. Der Wissenschaftler hat mit verschiedenen Methoden herauszufinden, ob das Objekt das ist, was es vorgibt zu sein. In der Informatik erfolgt die Feststellung der Authentizität über genaue Identifikationsmerkmale, die vom Objekt erfüllt werden müssen. Bei digitalen Objekten sind zur kritischen Beurteilung beide Aspekte von Interesse.

Zur sprachlichen Präzisierung ist zunächst festzuhalten: *Authentizität*⁷³⁰ wird als Aussage über die Echtheit, Zuverlässigkeit oder Glaubwürdigkeit einer Person oder Sache verstanden. Die *Authentifikation* ist der Prozess der Identitätsüberprüfung mittels *Authentifizierung* (Überprüfung der Person oder Sache durch Externe) und/oder *Authentisierung* (Nachweis der Identität durch die Person oder Sache selbst).⁷³¹ Die Identitätsprüfung ist im Kontext der historisch-kritischen Methode die Frage an das Objekt, ob es das ist, was es zu sein scheint. Von diesen Begriffen abzugrenzen ist die *Autorisierung*, die die Berechtigung oder Ermächtigung einer Person oder Sache für einen Zugang oder Zugriff auf eine Person oder Sache beschreibt.

Aus *historischer* Sicht wird Authentizität mehrdeutig und mit Begriffen wie Wahrheit, Echtheit, Original, Faktizität, Tradition oder Ritual synonym oder überlappend verwendet.⁷³² Authentizität wird gleichgesetzt mit *tatsächlich* sowie dem Anspruch, dass die vorliegende Quelle das ist, was sie zu sein scheint. Gleichzeitig wird Authentizität mit dem Wahrheitsanspruch verknüpft, dass der Inhalt die historische Begebenheit so wiedergibt, wie es tatsächlich gewesen ist. Dabei wird diese historische Authentizität nur durch die

⁷²⁹ Pirker, Eva Ulrike; Rüdiger, Mark: Authentizitätsfiktionen in populären Geschichtskulturen: Annäherungen, in: *Echte Geschichte: Authentizitätsfiktionen in populären Geschichtskulturen*, Bielefeld 2010 (Historische Lebenswelten in populären Wissenskulturen), S. 14.

⁷³⁰ Zur Dekonstruktion des Begriffs Authentizität siehe bspw.: Uhl, Heidemarie; Rössner, Michael: *Renaissance der Authentizität? : Über die neue Sehnsucht nach dem Ursprünglichen*, Bielefeld 2012 (Kultur- und Medientheorie). Siehe auch: InterPares Project 1 (URL: <http://www.interpares.org>).

⁷³¹ Im Englischen wird der Begriff 'authentication' sowohl für Authentifizierung als auch für Authentisierung verwendet.

⁷³² Pirker; Rüdiger: *Authentizitätsfiktionen*, 2010, S. 13. Siehe auch Kap. 3.1.3. Quellenkritik.

Untersuchung des die Quelle umgebenden Kontextes erreicht werden, denn die Authentifizierung bedeutet, dass Klarheit über die Herkunft einer Aussage und den Blickwinkel des Aussagens verschafft wird, wie Philipp Sarasin unter Berufung auf Michel Foucault ausführt.⁷³³ Bei einer Authentifizierung ist Kontextualisierung immer notwendig, denn es sind nach Foucault zusätzlich "die medialen, diskursiven, politischen und ökonomischen Verhältnisse, die eine Aussage möglich gemacht haben"⁷³⁴ zu untersuchen. In einen Kontext eingebettete Fakten und Tatsachen sind Gradmesser der historischen Authentizität oder, anders ausgedrückt, der historischen Wahrheit.⁷³⁵ Mit Authentizität ist die Frage nach der Echtheit der historischen Quelle verknüpft. Echtheit wiederum wird unter Berücksichtigung von Datierung, Entstehung, Autorschaft, Irrtum oder Fälschung beurteilt.

Authentizität basiert auf einem Vergleich: Ist das, was vorliegt, gleich einem vorherigen Zustand desselben Objekts? Für digitale Objekte gilt, dass eine Kopie exakt gleich wie das Original ist – beide werden zu Klonen.⁷³⁶ Im täglichen Gebrauch wird auch dann von einer exakten Kopie gesprochen, wenn sie für die menschliche Wahrnehmung nicht vom Original zu unterscheiden ist, wie beispielsweise eine unkomprimierte von einer wenig komprimierten Musikdatei mit demselben Inhalt. Die *Originalität* eines physischen Objekts kann nachgewiesen werden, jene eines digitalen Objekt jedoch nicht. Roy Rosenzweig meint:

*"But digital information – because it is so easily altered and copied, lacks physical marks of its origins, and, indeed, even the clear notion of an 'original' – cannot be authenticated as physical documents and objects can."*⁷³⁷

Die Literaturwissenschaftlerin Eva Pirker und der Historiker Mark Rüdiger führen unter Berufung auf Walter Benjamin auf, dass sich das Original auch durch "auratische Erfahrungen"⁷³⁸ manifestiert. Sie argumentieren, dass einem physischen Objekt wie einer Burgruine oder dem Bundesbrief der Schweizerischen Eidgenossenschaft eine bestimmte Aura inneohnt, deren Ausstrahlung die Originalität spürbar macht. Dies sei bei einem digitalen Objekt noch⁷³⁹ nicht möglich, denn es würden sowohl das *Hier* als auch das *Jetzt* des Originals fehlen, und es kann aufgrund seiner Datenträgerlosigkeit nicht 'erfahrbar' gemacht werden. Wird die Aussage eng auf das digitale Objekt bezogen, stimme ich ihr zu. Wenn jedoch das Informationssystem mit einbezogen wird, ohne das ein digitales Objekt nicht erfahrbar gemacht

⁷³³ Nach: Sarasin: *Schlaue Maschinen*, 2013, S. 196. Sarasin ist Mitbegründer und Mitglied des interdisziplinären 'Zentrums Geschichte des Wissens' (URL: <https://www.zgw.ethz.ch/>).

⁷³⁴ Ebd.

⁷³⁵ Nach: Pirker; Rüdiger: *Authentizitätsfiktionen*, 2010, S. 14.

⁷³⁶ Siehe Kap. 2.1. Eigenschaften eines Digitalen Objekts.

⁷³⁷ Rosenzweig, Roy: *Scarcity or Abundance? Preserving the Past*, in: *Clio wired – The future of the Past in the Digital Age*, 2011, S. 10.

⁷³⁸ Pirker; Rüdiger: *Authentizitätsfiktionen*, 2010, S. 18-20, insb. S. 19. Zum Begriff 'Aura' siehe: Benjamin: *L'œuvre d'art à l'époque de sa reproduction mécanisée*, in: *Zeitschrift für Sozialforschung* (5), 1936, S. 40-68.

⁷³⁹ Zu Systemen, die Gerüche verbreiten oder erkennen können, siehe: Turkel: *Intervention*, 2012.

werden kann, können gewisse digitale Objekte durchaus eine gewisse Aura ausstrahlen. Dies wäre beispielsweise bei Computerspielen wie Pac-Man der Fall, das in einer 'Gamehalle' am 'Gamekasten', der Arcade Video Game Machine, gespielt wird. Damit wäre Benjamins auratische Erfahrung auch bei digitalen Objekten möglich, wie nostalgische Gamer der 1980er Jahre sicher bestätigen würden.⁷⁴⁰ Eine diesbezügliche vertiefende Untersuchung kann hier jedoch nicht geleistet werden.

Authentizität im historischen Sinne wird jedoch problematisch, wenn Wahrheit oder Echtheit in der virtuellen Realität / Virtual Reality beurteilt werden soll. Beispielsweise schaffen sich die Nutzer der virtuellen Welt *Second Life*⁷⁴¹ einen künstlichen Lebensraum, in dem sie durch ihre *Avatare*⁷⁴² miteinander interagieren. Der Nutzer ist physisch echt und steuert seinen computergenerierten Avatar. Es wird beim Nutzer ein "Gefühl der Präsenz im digitalen Raum und die Empfindung des Eintauchens in virtuelle Welten, die sog. Immersion"⁷⁴³ erzeugt. Die Interaktion mit einem weiteren Avatar geschieht durch manuelle Steuerung und Kommunikation durch die Nutzer, wird jedoch virtuell dargestellt. Welcher Bereich nun als authentisch bezeichnet werden kann, ist hier nicht klar, denn eine physische Aktion wird zwar durch die Virtualisierung verfälscht, stellt aber dennoch den echten Willen des Nutzers dar.

Ein ähnliches Problem tritt bei der Verwendung von aktuellen Kommunikationsmitteln für das Nachzeichnen historischer Begebenheiten dar, wie im erwähnten Facebook-Projekt über Léon Vivien.⁷⁴⁴ Viviens Briefe werden als authentische Kommunikationsmedien durch die Verschiebung der Zeitkomponente und die Migration in ein anderes Medium verfälscht, ihr Inhalt jedoch bleibt authentisch.

Der Begriff Authentizität bezeichnet in der Geschichtswissenschaft präzise, aber gleichzeitig unterschiedliche Dinge in unterschiedlichem Kontext und in verschiedenen Gemeinschaften. Er kann *original* bedeuten, aber auch *originalgetreu*; das Objekt kann beschädigt oder unbeschädigt sein, von bekannter oder unbekannter Herkunft stammen.⁷⁴⁵

⁷⁴⁰ Annahme, aufgrund von persönlichen Gesprächen mit verschiedenen Personen, die als Jugendliche in den 1980er Jahren in den Gamehallen ihre Zeit verbrachten und sich heute gerne daran erinnern.

⁷⁴¹ URL: <http://secondlife.com/>.

⁷⁴² Avatar: virtuelle Gestalt als spielbare Figur oder virtuelle Person. In *Second Life* menschähnlich, in anderen 'Welten' können auch andere Lebensformen und (mystische) Gestalten angenommen werden. Siehe weiterführend: Ackermann, Judith: Avatars: Identitäten und digitale (Ab-)Bilder im Netz, in: Stiegler, Christian; Breitenbach, Patrick; Zorbach, Thomas (Hg.): *New Media Culture: Mediale Phänomene der Netzkultur*, Bielefeld 2015 (Digitale Gesellschaft), S. 51-65.

⁷⁴³ Ebd., S. 53.

⁷⁴⁴ Siehe Kap. 4. Digitalmedien.

⁷⁴⁵ Nach: Smith, Abby: Introduction, in: Cullen et al. (Hg.): *Authenticity in a Digital Environment*, Washington, D.C 2000, S. VI.

Im Gegensatz zur historischen Methode wird in der *Informatik* die Authentizität nicht anhand des Inhalts, der Form oder des Kontextes bewertet⁷⁴⁶, sondern nur über eine eindeutige, einseitige Identifikation respektive durch die Überprüfung der Integrität. Das physische oder digitale Objekt muss *alle* erforderlichen Identifikationsmerkmale aufweisen, um für das Informationssystem autorisiert zu werden. Theoretisch können unendlich viele Merkmale festgelegt werden, in der Praxis werden jedoch meistens nur eines bis drei verwendet.⁷⁴⁷ Beispielsweise kann die Identifikation eines Nutzers über ein Passwort, einen Loginnamen plus Passwort oder durch Loginnamen, Passwort und eine per SMS zugestellte Zahl erfolgen.

Zur Beurteilung der Authentizität eines digitalen Objekts ist vor allem die ständige Weiterentwicklung von Computersystemen problematisch. Wenn ein Computersystem nicht mehr im Originalzustand nutzbar ist, können digitale Objekte meistens auch nicht mehr in ihrem archetypischen Zustand dargestellt werden.⁷⁴⁸ Änderungen an einem System sind deshalb hinsichtlich ihrer Beeinflussung des Inhalts und die Repräsentation des Objekts zu untersuchen. Die Frage nach der Echtheit ist an die Frage nach der Authentizität des digitalen Objekts zu koppeln⁷⁴⁹, denn es gilt nachzuweisen, dass es unverändert überliefert wurde und wird. Da Objekte datenträgerunabhängig sind, ist der Nachweis hauptsächlich aus dem Objekt selbst und den zugehörigen Metadaten zu erbringen.⁷⁵⁰ Physische Komponenten wie USB-Sticks können eventuell anhand der Etikette oder von Gebrauchsspuren zusätzliche Hinweise liefern.

Zur Beurteilung der Authentizität eines digitalen Objekts liefern die informationstechnischen Identifikationsmerkmale zwar eindeutige Hinweise, das Objekt darf aber nicht allein daraufhin überprüft werden. Auch ein Objekt ohne solche Merkmale kann als potenzielle historische Quelle in Betracht gezogen werden, weil selbst Fälschungen in gewissen Kontexten und für bestimmte Fragestellungen wertvolle Informationen beinhalten können. Die Feststellung der informationstechnischen Authentizität dient der Beurteilung der Authentizität des gesamten Objekts.⁷⁵¹

5.4.2. Integrität

Es ist wie beim Begriff Authentizität zwischen einer historischen und einer informationstechnischen Auslegung des Begriffes Integrität zu unterscheiden. Aus intellektueller Sicht kann

⁷⁴⁶ Siehe: Lynch, Clifford A.: *Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust*, in: Cullen et al. (Hg.): *Authenticity in a Digital Environment*, Washington, D.C 2000, S. 35 f.

⁷⁴⁷ Es wird von Ein- bis Mehrfaktor-Authentisierung gesprochen.

⁷⁴⁸ Siehe: Margulies: *Digitale Daten als Quelle der Geschichtswissenschaft*, 2009, S. 326-328.

⁷⁴⁹ Vgl.: Ebd., S. 326.

⁷⁵⁰ Siehe auch Diskussion um den Begriff Originalkopie in Kap. 3.1.2. Quellenkunde.

⁷⁵¹ Siehe Kap. 5.4.12. Beziehungen zu Objekten, Lösungsansatz Authentizitätsapproximation.

von verschiedenen Arten von Integritäten gesprochen werden. Für die quellenkritische Untersuchung sind die inhaltliche, die Darstellungs- und die Nutzungsintegrität sowie die Integrität der involvierten Informationssysteme von Interesse. Wissenschaftliche oder wissenschaftsnahe Publikationen sollten mit einer qualifizierten digitalen Signatur mit Zeitstempel versehen werden, um den Integritätsnachweis zu erbringen.

Obwohl der Begriff *Integrität*⁷⁵² als wichtiges Kriterium für die Beurteilung der Authentizität eines Objekts gilt, ist dessen Bedeutung in der Geschichtswissenschaft nicht eindeutig geklärt. In der Fachliteratur wird er zwar oft verwendet, aber weder in einschlägigen Lexika noch in Einführungsliteratur zum Studium genauer untersucht oder definiert.⁷⁵³

Eine *historisch-quellenkritische Integrität* bezieht sich auf ein physisches oder virtuelles Objekt, dem kein eigenes moralisch-ethisches Wertesystem zugeschrieben werden kann. Ein Objekt kann materiell und inhaltlich vollständig integer sein, wenn angenommen werden kann, dass es unbeschädigt und unversehrt, aber nicht unberührt, oder unverändert und vollständig ist, insbesondere bei mündlichen Überlieferungen. Wenn es materiell beschädigt ist oder verändert wurde, ist es nicht mehr in seiner Gesamtheit integer, kann aber trotzdem aus intellektueller Sicht weiterhin integrale Aspekte aufweisen. Beispielsweise kann ein Buch durch Gebrauch verschmutzt und zerknittert sein und eingerissene Seiten aufweisen, die inhaltliche Integrität ist aber dennoch gegeben.

Die *informationstechnische Integrität* eines digitalen Objekts⁷⁵⁴ umfasst die Unversehrtheit des Bitcodes und die semantische Korrektheit. Jede Änderung am Bitcode eines Objekts verstößt gegen dessen Integrität. Übermittlungsfehler sowie Format- oder Inhaltsänderungen verändern das Objekt, wodurch es seine Integrität verliert. Ein Objekt ist semantisch korrekt und integer, wenn der reale oder modellierte Sachverhalt wahrheitsgetreu wiedergegeben wird. Kann er durch das Informationssystem nicht oder nicht korrekt dargestellt werden, ist die semantische Integrität nicht mehr gegeben. Das Objekt ist nur dann integer, wenn sowohl der Bitcode unversehrt, als auch die Semantik korrekt ist. Für den Informatiker Clifford Lynch bedeutet Integrität eines digitalen Objekts, dass es über die Zeit und auf dem

⁷⁵² Aus dem Lateinischen von 'tangere' = berühren. Integrität kann mit *Unversehrtheit* oder *Unberührt-heit* übersetzt werden. Siehe zur soziologischen Unterscheidung von Integrität bspw.: Kesselring, Thomas: Mens sana in corpore sano. Integrität aus ethischer Sicht, in: Tagung «Wieviel Körper darf es sein?», Bern 2007, S. 1-4. URL: <http://docplayer.org/32570566-Thomas-kesselring-mens-sana-in-corpore-sano-integritaet-aus-ethischer-sicht.html>, Stand: 06.05.2018.

⁷⁵³ Für Lexika siehe bspw.: Fuchs; Raab: Wörterbuch zur Geschichte, 2002. Für Einführungsliteratur siehe bisher genannte.

⁷⁵⁴ Es wird hier nur auf ein digitales Objekt Bezug genommen, nicht auf Domänen-, Entitäts- oder Tupelintegritäten oder auf referenzielle Integrität oder Integritätsregeln. Siehe für den ganzen Absatz: Fischer; Hofer: Lexikon der Informatik, 2011.

Transportweg nicht beschädigt wurde respektive, dass die genau gleiche Bitfolge vorliegt, die bei der Erstellung des Objekts generiert wurde."⁷⁵⁵

Bei digitalen Objekten bedeutet die Beeinträchtigung der Integrität einen Informations- und/oder einen Datenverlust. Bei einem *Informationsverlust* sind zwar noch Daten vorhanden, sie können aber nicht mehr vollständig interpretiert werden. Bei einem *Datenverlust* sind keine oder keine verwertbaren Daten mehr vorhanden. Wenn vorhandene Daten korrekt dargestellt werden können, aber die falschen Informationen enthalten, spricht man von *manipulierten oder korrupten Daten*. Werden Daten, die willentlich oder unwillentlich⁷⁵⁶ manipuliert wurden, nicht entdeckt, kann dies grössere Auswirkungen nach sich ziehen. Beispielsweise wurde 2011 die Webseite des US-Senats von der Gruppe LulzSec gehackt.⁷⁵⁷ In einem solchen Fall müssen die verantwortlichen Administratoren, um sicherzugehen, dass sich keine manipulierten Daten auf der Webseite mehr befinden, den gesamten Webauftritt aus einem Backup⁷⁵⁸ mit integren Daten vollständig wiederherstellen. Dies ist allerdings nur dann erfolgreich, wenn integrale Daten noch verfügbar sind, d.h. ein Backup vor dem Einbruch vorhanden ist.⁷⁵⁹ Für die Nutzer der Seite bestehen kaum Möglichkeiten zu überprüfen, ob die von ihnen entnommenen Daten integer sind oder bereits kompromittiert wurden.

In den Archivwissenschaften ist die Integrität eines digitalen Objekts von grosser Wichtigkeit, denn:

*"we want reliable evidence that a digital object was not later modified even slightly, because in some cases even a single changed bit can radically alter the meaning of text or the behaviour of a program."*⁷⁶⁰

Henry Gladney unterscheidet zwischen *Integrität* und *wahrer Herkunft*:

*"To be authentic a document must be whole and not disturbed too much – that is integrity. It must have originated with the purported author as part of the purported event – that is true provenance."*⁷⁶¹

⁷⁵⁵ Lynch: *Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust*, 2000, S. 38.

⁷⁵⁶ Unwillentlich manipulierte Daten (Fachtermini: bit rot, bit decay, (silent) data corruption) entstehen bspw. durch Veränderungen während des Kopiervorganges von digitalen Daten. Nach einer Untersuchung von Bernd Panzer-Steindel von 2007 ist eine von 1'500 Dateien davon betroffen. Selbst bei modernen Datenspeichersystemen treten solche Fehler weiterhin häufig auf (die Ausnahme ist das auf diese Problematik entwickelte ZFS). Siehe: Panzer-Steindel, Bernd: *Data integrity*, CERN, 08.04.2007. URL: http://indico.cern.ch/event/13797/contributions/1362288/attachments/115080/163419/Data_integrity_v3.pdf, Stand: 19.04.2016.

⁷⁵⁷ Siehe: Morse, Andrew; Sherr, Ian: *Senate Website Gets Hacked*, in: *Wall Street Journal*, 14.06.2011. URL: <http://online.wsj.com/article/SB10001424052702303848104576383970053018848>, Stand: 13.11.2012.

⁷⁵⁸ Backup: Datensicherung auf externen Datenträgern oder Serversystemen.

⁷⁵⁹ Die Entdeckung des Einbruchs ist nicht zwangsläufig zeitgleich mit dem tatsächlich erfolgten. Wann ein Einbruch stattgefunden hat, ist unter Umständen schwierig festzustellen.

⁷⁶⁰ Gladney: *Preserving Digital Information*, 2007, S. 165 f.

⁷⁶¹ Ebd., S. 106. Schriftschnitt im Original invers.

Allerdings widerspricht er mit 'not disturbed too much' seiner vorhergehenden Aussage, dass jegliche nachträgliche Manipulation nicht akzeptabel sei. Die gängige Archivpraxis fordert, dass die physische Integrität nachweisbar ist und die Vollständigkeit und Unbeschädigtheit des Bitcodes kontinuierlich gewährleisten muss.⁷⁶² Folglich ist jede Veränderung des Bitcodes eine Verletzung der Datenintegrität.

Als Integritätsverlust ist auch die Verwendung von *DRM-Verfahren* zu werten. Wird ein digitales Objekt beispielsweise mit einem digitalen Wasserzeichen versehen, wird das Objekt um Elemente ergänzt, die theoretisch ohne Kenntnis des verwendeten Algorithmus nicht wieder entfernt werden können.⁷⁶³ Format- und Formänderungen sind zwar weiterhin möglich, das Wasserzeichen bleibt jedoch in allen erhalten. Mit einem solchen Prozess wird ein Objekt willentlich und systematisch verändert und damit auch dessen Integrität verletzt, was unter anderem auch Auswirkungen auf die Langzeitarchivierung und Funktionalität des Objekts haben kann. Eventuell ist es mit diesem zusätzlichen Element nicht mehr auf den dafür vorgesehenen Systemen darstellbar, oder die Sicherung in einem langzeitarchivfähigen Format ist nicht möglich.⁷⁶⁴ Nur wenn das daraus entstandene Objekt vom Autor oder Rechteinhaber als Original definiert und akzeptiert wird, kann es wieder als integer bezeichnet werden.⁷⁶⁵

Die Datenintegrität gibt einen wichtigen Hinweis darauf, ob das digitale Objekt manipuliert worden ist. Zur Überprüfung eignet sich beispielsweise der bereits besprochene Hashwert oder die *digitale Signatur*, die als eine der Schlüsseltechniken zur Ermittlung der Integrität digitaler Objekte gilt.⁷⁶⁶ Ein signiertes Objekt entsteht, indem es in einem asymmetrischen, kryptographischen Verfahren mit einem kryptischen, privaten Schlüssel einer Person oder einer Institution verbunden respektive signiert wird. Eine *fortgeschrittene Signatur* erfüllt zusätzlich gesetzliche Anforderungen wie die eindeutige Identifizierung des Signaturinhabers und der Nachweisbarkeit von Manipulationen an den signierten Daten, indem (es vor der Signierung aus dem Objekt ein Hashwert erzeugt und mit der Signatur verbunden wird).⁷⁶⁷

⁷⁶² Nach: Keller-Marxer, Peter: Archivierung elektronischer digitaler Daten und Akten (ARELDA) im Bundesarchiv, VSA/AAS-Arbeitstagung «Elektronische Akten – eine Strategie. Und dann?», Bern 22.11.2002, Fol. 15; Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 450.

⁷⁶³ Vgl.: Lynch: Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust, 2000, S. 42 f.

⁷⁶⁴ Siehe weiterführend: Ebd., S. 43.

⁷⁶⁵ Der Autor, der Rechteinhaber und der Nutzer sind oft unterschiedlicher Ansicht, was als 'Original' akzeptiert werden soll.

⁷⁶⁶ Day: Long-Term Preservation, 2006, S. 190. Zur Funktionsweise siehe: Hühnlein, Detlef; Korte, Ulrike: Grundlagen der elektronischen Signatur: Recht, Technik, Anwendung, Ingelheim 2006.

⁷⁶⁷ Für die Schweiz siehe: Bundesversammlung der Schweizerischen Eidgenossenschaft: Bundesgesetz über Zertifizierungsdienste im Bereich der elektronischen Signatur (Bundesgesetz über die elektronische Signatur, ZertES), 01.08.2008. URL: <http://www.admin.ch/opc/de/classified-compilation/20011277/index.html>, Stand: 24.04.2015, Art. 2 lit. b. Mit der am 18.03.2016 beschlossenen Totalrevision des ZertES werden künftig eine qualifizierte elektronische Signatur (der handschriftlichen Unterschrift gleichgestellt), eine elektronische Signatur und ein elektronisches Siegel für juristische

Wenn für die Signierung ein qualifiziertes Zertifikat einer staatlich autorisierten Signaturstelle verwendet wird, spricht man von einer *qualifizierten Signatur*.⁷⁶⁸ Zusätzlich kann die digitale Signatur mit einem Zeitstempel kombiniert werden, was die eindeutige Datierung eines digitalen Objekts erlaubt.⁷⁶⁹

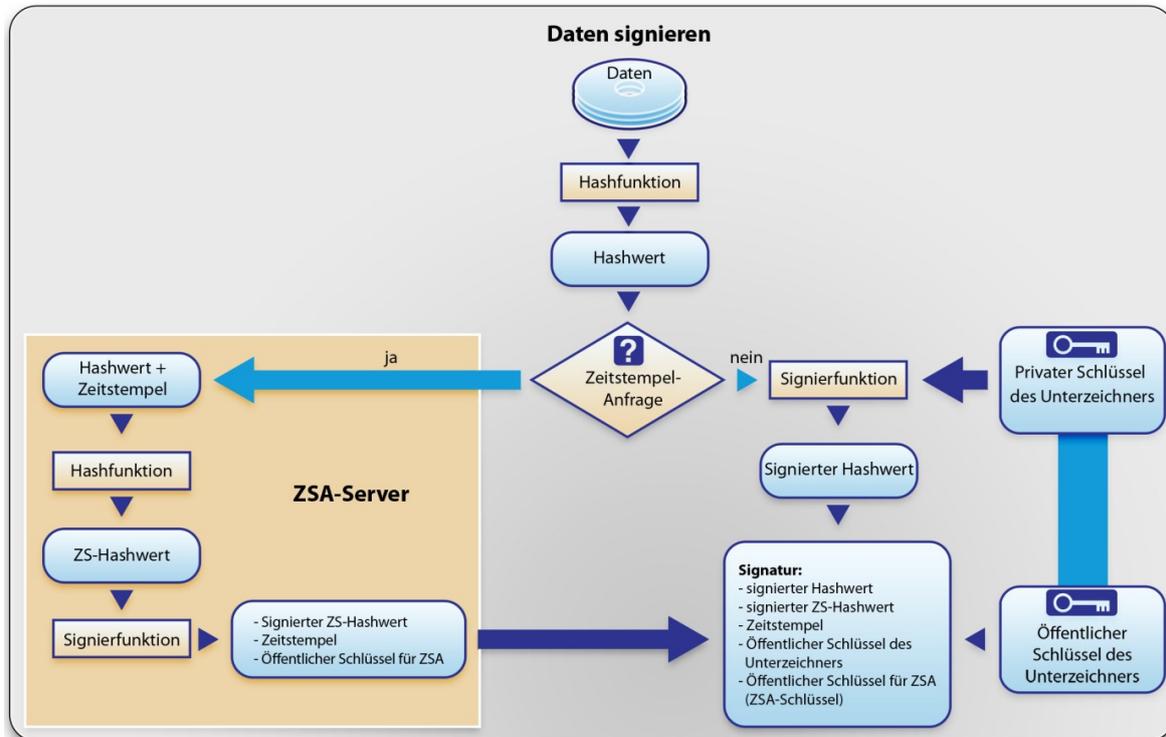


Abbildung 22: Signatur erstellen (Soft Xpansion)⁷⁷⁰

Durch das Auslesen der Signatur mit dem öffentlichen Schlüssel des Signierenden und mit dem erneut aus dem Objekt erzeugten Hashwert werden die vorhandenen Informationen verglichen. Stimmen diese überein, ist das Objekt integer und stammt vom angegebenen Signierenden – unter der Voraussetzung, dass dessen privater Schlüssel immer noch geheim ist. Andernfalls wurde es manipuliert oder stammt nicht vom Signierenden.

Personen und Behörden zur Verfügung stehen (Botschaft zur Totalrevision des Bundesgesetzes über die elektronische Signatur, ZertES vom 15.01.2014. URL: <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20140015>.

⁷⁶⁸ Ebd., Art. 6 und 7.

⁷⁶⁹ Bspw. erlaubt 'ArchiSig' (ArchiSoft) des Fraunhofer-Instituts die erwähnte Kombination. Siehe URL: <https://www.sit.fraunhofer.de/de/angebote/projekte/archisoft/>.

⁷⁷⁰ Abbildungen 22 und 23 mit freundlicher Genehmigung der soft Xpansion GmbH & Co. KG, URL: <http://www.soft-xpansion.eu> (E-Mail vom 10.11.2015, Marius Trojca); Bildzitate: soft Xpansion GmbH & Co. KG: Sichere PDF-Dokumente durch digitale Signaturen, 2011, S. 5. URL: <http://soft-xpansion.de/utills/csp/sichere%20pdf-dokumente%20durch%20digitale%20signaturen.pdf>, Stand: 05.11.2015.

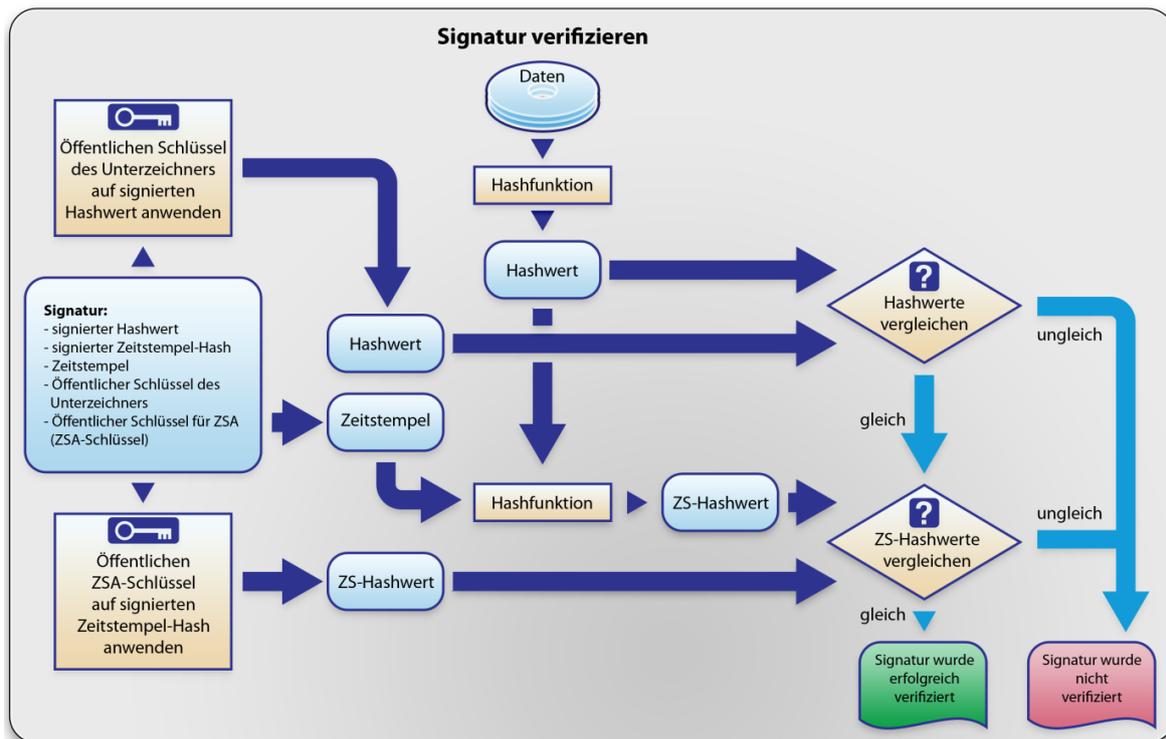


Abbildung 23: Signatur verifizieren (Soft Xpansion)

Ein mit einer digitalen Signatur versehenes Objekt befand sich zum Zeitpunkt der Signierung im ursprünglichen Zustand. Über vorausgehende Änderungen kann keine Aussage gemacht werden. Das Signieren kann und sollte in regelmässigen Abständen auf Korrektheit überprüft und wiederholt werden, woraus eine Überlieferungsgeschichte entsteht. Mit jeder Signatur wird bestätigt, dass das Objekt zu diesem Zeitpunkt integer war und ist.

Bei den algorithmischen Verfahren zur Integritätssicherung muss die Korrektheit der vom Informationssystem erzeugten Hashwerte und Signaturen vorausgesetzt werden, wenn auch diese vom Nutzer selbst nicht zweifelsfrei überprüft werden können.⁷⁷¹ Es ist deshalb darauf zu vertrauen, dass die eingesetzten informationstechnischen Komponenten grundsätzlich Daten korrekt verarbeiten.⁷⁷² Resultate aus Forschungen und Weiterentwicklungen in der Kryptologie sind stets zu verfolgen, um allenfalls weiterentwickelte oder neuartige Verfahren einsetzen zu können.

Unabhängig von der informationstechnischen Feststellung der Integrität eines digitalen Objekts ist auch dessen *inhaltliche Integrität* zu überprüfen. Wie bei einem physischen Objekt kann es zwar manipuliert worden sein, die Informationsintegrität aber kann weiterhin bestehen, wenn die inhaltliche Aussage und der Wahrheitsgehalt nicht verändert worden sind. Dies ist beispielsweise bei jeder Datenformatmigration der Fall, bei der ein neues und informationstechnisch gegenüber dem vorherigen nicht integriertes Objekt entsteht. Der Inhalt ist mit

⁷⁷¹ Siehe: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 398-400.

⁷⁷² Siehe: Lynch: Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust, 2000, S. 38.

grosser Wahrscheinlichkeit jedoch der gleiche, ist aber zu überprüfen. Bei der Konversion können Zeichensätze falsch verwendet werden, wodurch ein Text nicht mehr lesbar sein kann.

Informationsintegrität ist allerdings nur dann feststellbar, wenn der Aussagegehalt der bisherigen Information bekannt und/oder das Primärobjekt noch vorhanden ist sowie wenn der Konversionsprozess offensichtlich fehlerhaft war. In den ersten beiden Fällen kann das Sekundärobjekt mittels verschiedener algorithmischer Verfahren mit dem Primärobjekt verglichen werden. Beispielsweise können durch textanalytische Verfahren die Buchstabenfolgen zwischen zwei Textdokumenten untersucht werden.⁷⁷³ Solche Verfahren können mit entsprechender Software auch zwischen Dokumenten unterschiedlichen Formats durchgeführt werden.

Wenn nur das digitale Objekt selbst vorhanden ist oder wenn Objekte mit unterschiedlicher Darstellung vorliegen, können keine automatisierten Vergleiche durchgeführt werden. Im ersten Fall stehen beispielsweise keine oder nur verkürzte Angaben zu seinem Inhalt zur Verfügung. Im zweiten Fall wird beispielsweise ein Bild in seinen Metadaten in Textform beschrieben. Der Nutzer muss versuchen, die inhaltliche Integrität des Objekts anhand der zur Verfügung stehenden Informationen und durch seine Erfahrung zu beurteilen.

Vor allem in der Kunstgeschichte ist auch die *Darstellungsintegrität* eines digitalen Objekts von grosser Wichtigkeit. Ein digitales Kunstobjekt wie ein Bild oder eine Projektion wird von einem Künstler in einer bestimmten Darstellungsform entworfen. Wird es erneut dargestellt, so ist nur unter Verwendung desselben Informationssystems die genau gleiche Darstellung möglich – Fehler sind nicht ausgeschlossen. Bei jedem anderen Informationssystem wird eine intendierte Darstellung aber umso unwahrscheinlicher, je stärker sich dessen Komponenten von jenen des ursprünglich verwendeten Informationssystems unterscheiden. Wird beispielsweise ein Bild an einem Bildschirm entworfen, verändert es sich bei einer Projektion mit einem Beamer auf eine für den Urheber unter Umständen nicht akzeptable Art und Weise.⁷⁷⁴ Für die Beurteilung der Darstellungsintegrität ist auch der Darstellungsort zu berücksichtigen. Die Hinweise oder Anleitungen zur intendierten Darstellung sollten deshalb in die Metadaten aufgenommen werden.

Mit der *Nutzungsintegrität* wird nachgewiesen, ob das digitale Objekt in der vom Urheber bestimmten Art und Weise genutzt werden kann. Beispielsweise ist die Nutzungsform eines Bildes die visuelle Betrachtung und ist unabhängig von der Darstellungsintegrität. Ein Computerspiel hingegen wurde für die Nutzung mit bestimmten peripheren Komponenten wie Joysticks entwickelt, die grundsätzlich bei jeder Verwendung zur Verfügung stehen müssen.

⁷⁷³ Mit geeigneten Programmen wie Microsoft Word (URL: <https://products.office.com/de-de/word>) möglich.

⁷⁷⁴ Weitere Fehlerquellen wurden bereits angesprochen.

Wird beispielsweise durch eine Emulation davon abgewichen, wird die Nutzungsintegrität verletzt und das Objekt kann allenfalls gar nicht genutzt werden.

Nicht nur die Integrität des Objekts ist zu überprüfen, sondern auch die *Integrität der verwendeten Informationssysteme*. Ein elektronisches Informationssystem kann beispielsweise durch eine extern eingeschleuste Software wie einem Trojaner oder Virus manipuliert worden sein, die Daten löscht oder die korrekte Ausführung von Programmen verhindert.⁷⁷⁵ Auch Hardware kann manipuliert werden, wie das Team um den Informatiker Samuel T. King der University of Illinois (USA) demonstrierte.⁷⁷⁶ Mit nur wenigen zusätzlichen, manipulierten Mikroprozessoren in einem Informationssystem könnte sich ein Angreifer durch eine sogenannte 'Backdoor' (Hintertür) vollen Zugriff auf das System verschaffen und könnte Daten und Software manipulieren oder löschen. Die Aufdeckung von Hardwaremanipulationen ist selbst für Informationstechniker extrem schwierig, Laien entdecken solche Manipulationen nur zufällig. Nutzer sollten sich zumindest vergewissern, dass die für die korrekte Darstellung und Verarbeitung des Objekts nötigen Hard- und Softwarekomponenten vorhanden, aktuell und funktionstüchtig sind und das Informationssystem auf Schadsoftware geprüft worden ist.

Der Integritätsnachweis eines digitalen Objekts ist auch *juristisch* von Belang.⁷⁷⁷ Je überzeugender dessen Integrität nachgewiesen werden kann, desto beweiskräftiger ist es in juristischen Verfahren. Da beim Geschäftsverkehr mittels elektronischer Informations- und Kommunikationstechnologien Manipulationen nicht festgestellt werden können und auch kein vertrauenswürdiger Rückschluss auf den Aussteller gezogen werden kann, sind Techniken zu verwenden, die die Integrität des Objekts beweisen und auch den Aussteller identifizieren.⁷⁷⁸ Dazu geeignet ist die oben erwähnte digitale Signatur, die im sicheren elektronischen Rechtsverkehr Beweismittelfunktion übernehmen kann. Damit versehen, werden digitale Dokumente zu Urkunden⁷⁷⁹, deren Echtheit wie folgt definiert wird:

*"Eine Urkunde ist dann echt, wenn sie in ihrer gegenwärtigen Erscheinungsform und ihrem Inhalt nach vom Aussteller stammt oder zumindest dessen Willen ausdrückt und somit die Gewissheit besteht, dass sie das ist, wofür sie ausgegeben wird."*⁷⁸⁰

⁷⁷⁵ Trojaner oder Trojanisches Pferd: Software, welche mit dem Ziel der Weiterleitung von Daten dem Wirtsystem falsche Informationen vorspiegelt. Virus: Software mit meist destruktivem Ziel. Siehe weiterführend: Fischer; Hofer: Lexikon der Informatik, 2011; Prevezanos: Computer Lexikon 2013, 2012.

⁷⁷⁶ Nur 1'341 manipulierte von rund 1,8 Mio. vorhandenen Transistoren (0.08%). Siehe: King, Samuel T.; Tucek, Joseph; Cozzie, Anthony u. a.: Designing and implementing malicious hardware, in: Firt Usenix Workshop on Large-Scale Exploits and emergent Threats, San Francisco, CA, USA 2008. URL: https://www.usenix.org/legacy/events/leet08/tech/full_papers/king/king.pdf, Stand: 17.09.2015.

⁷⁷⁷ Siehe zu digitalen Objekten als juristischen Beweisen: Fanger, Reto: Digitale Dokumente als Beweis im Zivilprozess, Basel 2005 (Basler Studien zur Rechtswissenschaft).

⁷⁷⁸ Vgl.: Ebd., S. 56.

⁷⁷⁹ Siehe: Ebd., S. 132.

⁷⁸⁰ Ebd., S. 146.

In der Schweiz wurden die Rahmenbedingungen für die elektronische Geschäftsführung und die technischen, organisatorischen und verfahrenstechnischen Anforderungen an die Beweiskraft und die Kontrolle von elektronisch oder in vergleichbarer Weise erzeugten Daten und Informationen per Verordnung festgelegt.⁷⁸¹ Als Hauptbestandteil eines beweiskräftigen Dokuments wird die elektronische Signatur angesehen, deren Anforderungen im *Bundesgesetz über die elektronische Signatur (ZertES)* ausgeführt sind.⁷⁸² Es verpflichtet Unternehmen zum Einsatz der elektronischen Signatur, wenn digitale Objekte rechtsgültig aufbewahrt und verwendet werden sollen:

*"Die in Artikel 122 Absatz 1 MWSTV verlangten Voraussetzungen für die Beweiskraft elektronischer Daten sind erfüllt, sofern: [...] die elektronischen Daten bis zum Ablauf der Aufbewahrungsdauer nach Artikel 11 im notwendigen Umfang mittels Verifikation der elektronischen Signatur auf Integrität, Authentizität und Signaturberechtigung geprüft werden und bei automatisierter Verarbeitung die Verifikation der elektronischen Daten systematisch nach abgeschlossener Übermittlung, spätestens aber vor ihrer Verwendung, stattfindet und das Ergebnis dokumentiert ist,"*⁷⁸³

Der Gesetzgeber verlangt zur Erlangung der Beweiskraft zwingend die Prüfung auf Integrität und Authentizität mittels digitaler Signatur. Damit kann der Autor auch rechtlich haftbar gemacht werden, denn er kann die Signierung weder abstreiten noch leugnen.

Dass Datenintegrität von grosser Wichtigkeit ist, wurde mittlerweile auch auf höchster *politischer Ebene* erkannt. Der damalige estnische Staatspräsident Toomas Hendrik Ilves sagte 2013 in einer Rede vor dem Steering Board of the European Cloud Partnership:

*"[...], violating the integrity of data is even more dangerous, threatening lives and infrastructure essential to the functioning of society – we can for example imagine the damage that can be done by changing people's blood groups or the cycles of traffic lights by simple changes of the data in databases."*⁷⁸⁴

⁷⁸¹ Siehe: Schweizerischer Bundesrat: Verordnung über die Führung und Aufbewahrung der Geschäftsbücher (Geschäftsbücherverordnung; GeBüV), 01.01.2013. URL: <http://www.admin.ch/opc/de/classified-compilation/20001467/index.html>; Eidgenössisches Finanzdepartement EFD: Verordnung des EFD über elektronische Daten und Informationen (EIDI-V), 01.01.2010. URL: <http://www.admin.ch/opc/de/classified-compilation/20092054/index.html>. Andere Länder wie bspw. Deutschland oder Österreich kennen ähnliche Verordnungen zur elektronischen Geschäftsführung.

⁷⁸² Siehe: Bundesversammlung der Schweizerischen Eidgenossenschaft: ZertES, 2008. Andere Staaten kennen ähnliche Gesetze, bspw. Österreich: Bundesgesetz über elektronische Signaturen (Signaturgesetz – SigG).

⁷⁸³ Eidgenössisches Finanzdepartement EFD: EIDI-V, 2010, Art. 3, Ziff. 1, Lit. c. MWSTV: Mehrwertsteuerverordnung.

⁷⁸⁴ Ilves, Toomas Hendrik: The greatest cyber threat comes from violation of data integrity, Steering Board of the European Cloud Partnership, Berlin 14.11.2013. URL: <https://vp2006-2016.president.ee/en/media/press-releases/9604-president-ilves-in-berlin-the-greatest-cyber-threat-comes-from-violation-of-data-integrity/layout-visit.html>, Stand: 06.05.2018. Estland gilt zurzeit als informationstechnisch weit entwickelt. Siehe dazu: Langer, Marie-Astrid: Zu Besuch in der Zukunft: Fortschrittliches Estland, Europa, Neue Zürcher Zeitung, 27.02.2015, URL:

Manipulierte Datenbanken weisen falsche Informationen auf, die zu Falschinterpretationen und -darstellungen führen. Wird die Manipulation professionell durchgeführt, das heisst wenn alle für das gewünschte Resultat nötigen Daten entsprechend abgeändert werden, ist es kaum mehr möglich, die 'Wahrheit' nachzubilden.

Lösungsansatz Digitale Signatur und MIS

Die Erkenntnis der Wichtigkeit eines Integritätsnachweises bei digitalen Objekten haben Politiker, Juristen, Informatiker und auch Wissenschaftler erkannt. Mit der digitalen Signatur wurde ein System entwickelt, das die Integrität und Authentizität von digitalen Objekten, insbesondere von wissenschaftlichen und wissenschaftsnahen Publikationen gewährleisten könnte. Dabei können nicht nur das Objekt selbst, sondern auch die zugehörigen Metadaten beglaubigt werden, die dann durch Dritte unabhängig überprüft werden können. Die Verwendung der fortgeschrittenen oder qualifizierten digitalen Signatur würde auch den Urheber eindeutig identifizieren, was die Wahrscheinlichkeit von willentlich fehlerhaft erstellten Angaben reduziert. Mit einem Zeitstempel versehen, würde das Objekt auch eine eindeutige, verlässliche und damit zitierfähige Zeitangabe aufweisen.

Ich rege an, dass in der Wissenschaft die qualifizierte Signatur mit Zeitstempel als Integritätsnachweis digitaler Objekte, insbesondere von wissenschaftlichen und wissenschaftsnahen Publikationen, eingeführt wird. Zusammen mit der Metadata Information Section MIS würden zwei Möglichkeiten zum automatischen Austausch von beglaubigten Informationen und Objekten zur Verfügung stehen.

5.4.3. Persistenz

*"But perhaps the biggest problem with digital sources, though, is that like so many things in the digital world, they somehow have the ability to vanish completely when you really want or need them."*⁷⁸⁵ (Wellerstein, 2013)

Noch können wenige Einrichtungen digitale Quellen dauerhaft speichern und auch den Zugriff auf diese sicherstellen. Zwar versuchen verschiedene Dienstleister digitale Objekte langfristig zu archivieren, aber ob und welche Unternehmen oder Institutionen dauerhaft bestehen werden, ist fraglich. Institutionen wie das Internet Archive können ihre Dienstleistungen ohne weiteres und ohne Vorankündigung einstellen, weil sie weder zur permanenten Fortführung verpflichtet noch über lange Zeit finanziert sind. Des Weiteren ist die

http://www.nzz.ch/international/europa/zu-besuch-in-der-zukunft-1.18491451?extcid=Newsletter_27022015_Top-News_am_Morgen, Stand: 27.02.2015.

⁷⁸⁵ Wellerstein, Alex: The year of the disappearing websites, Restricted Data: The Nuclear Secrecy Blog, 27.12.2013, URL: <http://blog.nuclearsecrecy.com/2013/12/27/year-disappearing-websites/>, Stand: 01.08.2014.

Adressierung eines digitalen Objekts ebenfalls nicht persistent. Techniken wie DOI und Permalinks suggerieren eine dauerhafte Adressierung, können dies aber nur in einem verlässlichen Langzeitarchiv garantieren.

Der Begriff *Persistenz* wird sowohl in der Geschichtswissenschaft als auch in der Informatik im Sinne von *dauerhaft existierend* verwendet.⁷⁸⁶ Die Persistenz eines Objekts hängt stark von seinem physischen Träger und der Behandlung desselben ab, was bei digitalen Objekten nur teilweise eine Rolle spielt. Es ist zwar möglich, einen Bitcode auf grundsätzlich persistenten Medien aus Glas, Stein, Metall oder Keramik zu speichern⁷⁸⁷, jedoch wird erfahrungsgemäss nach etwa einer Dekade das Informationssystem zu seiner Darstellung nicht mehr zur Verfügung stehen und damit auch die mit dem Objekt verbundene Digitalität. Der oft angeführte Vergleich mit dem Stein von Rosette, der die Entschlüsselung der ägyptischen Hieroglyphen massgeblich beeinflusste, kann bei digitalen Objekten nicht herbeigezogen werden. Auf dem Stein wurde derselbe Text in den drei Sprachen Ägyptisch (Hieroglyphen), Demotisch und Altgriechisch eingemeisselt und ist aus dem Objekt selbst interpretierbar. Aus dem Bitcode 0 und 1 können hingegen keine Rückschlüsse auf seinen Inhalt oder die Darstellung erfolgen.

Ein weiteres Problem ist, dass ein digitales Objekt innert Sekundenbruchteilen gelöscht werden kann; mit entsprechendem Vorwissen und unter Verwendung spezieller Programme ist es auch nicht wiederherstellbar.⁷⁸⁸ Werden beispielsweise während des Löschprozesses alle vom Objekt genutzten Speichersektoren überschrieben, lassen sich auch mit Methoden der digitalen Forensik kaum mehr verwertbare Daten wiederherstellen. Selbst wenn solche noch vorhanden sein sollten, ist das Objekt höchstwahrscheinlich unvollständig und eine Interpretation kann nur hypothetisch erfolgen. Solche Löschvorgänge können in vernetzten Informationssystemen auch aus weiter Entfernung und von eigentlich nicht autorisierten Personen oder Prozessen ausgelöst werden. Ein im schweizerischen Herisau sitzender Hacker⁷⁸⁹ kann theoretisch den gesamten aktiven Datenbestand der Central Intelligence Agency (CIA) der

⁷⁸⁶ Siehe zum Begriff 'persistent, Persistenz': Fischer; Hofer: Lexikon der Informatik, 2011.

⁷⁸⁷ Zu Glas siehe bspw.: Zhang, Jingyu; Gecevičius, Mindaugas; Beresna, Martynas u. a.: 5D Data Storage by Ultrafast Laser Nanostructuring in Glass, University of Southampton, 2013. URL: <https://eprints.soton.ac.uk/364916/1/5973.pdf>, Stand: 06.05.2018. Zu Metall siehe Anmerkungen zum Rosetta-Projekt von: Rosenthaler: Archivierung im digitalen Zeitalter, 2007, S. 488-490.

⁷⁸⁸ Bspw. das vom Deutschen Bundesamt für Informatik empfohlene Programm 'Darik's Boot And Nuke' (DBAN). URL: <http://www.dban.org/>. Bei einem einfachen Löschvorgang wird in der 'Speicherverwaltung' des Informationssystems der Verweis auf die Speichersektoren der Daten gelöscht, die Daten sind aber weiterhin vorhanden, können ausgelesen werden und sind wiederherstellbar.

⁷⁸⁹ Unter dem Begriff werden allgemein Personen verstanden, die sich unberechtigt Zugriff auf geschützte, digitale Datenbestände verschaffen. Es wäre zu unterscheiden: Knacker: unbefugter Eindringling in geschützte Datenbestände; Hacker oder 'Grey Hats': Knacker, der auf Lücken in der Datensicherheit aufmerksam macht; Cracker: Knacker, der Lücken in der Datensicherheit zur Manipulation von Daten und/oder 'Bereicherung' im weiteren Sinne ausnutzt. Siehe bspw.: Fischer; Hofer: Lexikon der Informatik, 2011.

Vereinigten Staaten löschen. Alle Daten, die zwischen dem letzten Backup-Prozess und dem Löschvorgang erstellt wurden, würden dabei verloren gehen.

Informationssysteme basieren auf eindeutigen *Datenpfaden* (Adressen), die jedem digitalen Objekt einen bestimmten Namen zuweisen. In einem Ordner einer Ordnerstruktur ist nur ein Objekt mit dieser Bezeichnung erlaubt. Soll ein anderes Objekt mit derselben Bezeichnung in diesem Ordner gespeichert werden, wird entweder das bereits vorhandene ersetzt oder das neue wird mit einer anderen, noch nicht verwendeten Bezeichnung gespeichert. Beim Ersetzen wird das vorhandene Objekt gelöscht, worauf das neue gespeichert wird. Danach steht nur noch das neue Objekt zur Verfügung. Auf demselben Prinzip basiert der Manipulationsprozess: ein bestehendes wird durch ein neues Objekt ersetzt, das unter derselben Adresse wie das ursprüngliche erreichbar ist.

Die Adressierung oder Identifikation mittels Verlinkung ist besonders im Internet ein wichtiges Merkmal digitaler Objekte. Unter Angabe von standardisierten Linkmustern wie URI, URL, URN oder URC⁷⁹⁰ können die Objekte untereinander aufgerufen werden. Die Persistenz solcher Links, die ich hier *Adresspersistenz* nenne, ist aber nicht gewährleistet. Populäre Online-Plattformen wie YouTube sind sehr dynamisch und verfügen über keine Persistenz, weil sowohl das Material im Einzelnen als auch die Sammlungen als Ganzes sehr volatil sind.⁷⁹¹ Die Informatiker Hany SalahEldeen und Michael Nelson haben in ihrer Untersuchung über in Social Media verlinkte digitale Quellen gezeigt, dass nach rund zweieinhalb Jahren 27% der darin verlinkten Quellen unwiederbringlich verloren sind.⁷⁹² Bereits nach einem Jahr waren über 10% nicht mehr auffindbar. Auch im wissenschaftlichen Umfeld stellt die Adresspersistenz ein grosses Problem dar, wie Peter Haber anhand verschiedener Beispiele verdeutlicht hat.⁷⁹³ Links und Zusatzmaterialien, beispielsweise als Bestandteile eines mit digitalen Publikationssystemen erstellten Werks, können aus diversen Gründen plötzlich nicht mehr abrufbar sein, was das Angebot ab- oder sogar entwertet. Der Informatiker Steve Lawrence hat 2001 festgestellt, dass eine wissenschaftliche Webseite nach durchschnittlich 75 Tagen verschwindet.⁷⁹⁴ Diese relativ alte Untersuchung wurde leider bisher nicht wiederholt, aber weitere Autoren berichten Ähnliches über nicht mehr funktionierende Links

⁷⁹⁰ Zur Unterscheidung zwischen Uniform Resource Identifier (URI), UR Locator (URL), UR Name, UR Citation (URC) u.a. siehe bspw. W3C/IETF URI Planning Interest Group: URIs, URLs, and URNs: Clarifications and Recommendations 1.0, World Wide Web Consortium (W3C), 21.09.2001, URL: <http://www.w3.org/TR/uri-clarification/>, Stand: 30.07.2014.

⁷⁹¹ Nach: Haber, Peter: Sprung in eine andere Welt? Mediengeschichte im Zeichen von Digitalität und Remediation, in: Schweizerische Zeitschrift für Geschichte 60 (1), 2010, S. 132. DOI: <http://dx.doi.org/10.5169/seals-109698>.

⁷⁹² Vgl.: SalahEldeen, Hany M.; Nelson, Michael L.: Losing My Revolution: How Many Resources Shared on Social Media Have Been Lost?, arXiv:1209.3026 [cs.DL], 13.12.2012. URL: <http://arxiv.org/abs/1209.3026>, Stand: 06.05.2018.

⁷⁹³ Vgl.: Haber: Digital Past, 2011, S. 146.

⁷⁹⁴ Siehe: Lawrence, Steve; Pennock, David M.; Flake, David William u. a.: Persistence of Web References in Scientific Research, in: Computer (34), 02.2001, S. 26-31, hier S. 30.

in ihren aktuellen Sammlungen, was angesichts der grossen technischen Entwicklungsschritte der letzten Jahre nicht erstaunt.⁷⁹⁵ Die Kommunikationswissenschaftlerin Megan Ankerson meinte sogar, es sei einfacher, einen Film von 1924 zu finden als eine Webseite von 1994, denn:

*"Domain names were resold, repurposed, or abandoned, only to re-emerge a short time later as one of those generic directories of sponsored links that thrive on the expired traffic of an outdated web."*⁷⁹⁶

Wenn die im Link verwendete URL nicht oder nicht mehr vorhanden ist, erscheint oft der berühmte Fehler '404: Site not found'. Dies stellt allerdings den besten Fall dar, denn ein nicht vorhandenes Objekt kann auch nicht manipuliert worden sein.

Vergisst das Netz schnell, wie der Journalist Konrad Lischka unter Berufung auf SalahEldeen und Nelson formulierte und wie es Eric Schmidt, Verwaltungsratsvorsitzender von Google, gar forderte?⁷⁹⁷ Wenn unter dem angegebenen URI ein digitales Objekt nicht mehr auffindbar ist, bedeutet dies nicht zwangsläufig den Verlust desselben, weil es unter einer anderen Adresse weiterhin abrufbar sein könnte. Oder das Objekt ist noch unter derselben Adresse erreichbar, nur wurde in einer Datenbank der darauf verweisende Hyperlink gelöscht. Das Netz vergisst, aber nicht wegen fehlerhafter Links, sondern nur wenn Objekte gelöscht worden sind. Dieser Unterschied wird oft nicht beachtet, so auch in den Diskussionen nach dem Urteil des Europäischen Gerichtshofs, das Google verpflichtete, Verweise auf Webseiten mit persönlichen Daten aus seiner Resultatliste zu streichen.⁷⁹⁸ Dies hat unabhängig davon zu erfolgen, ob die Daten auf der sie publizierenden Webseite weiterhin zur Verfügung stehen. Die Daten selbst werden nicht gelöscht, weil sie von Externen meist auch nicht gelöscht werden können, und sind eventuell mit einer anderen Suchmaschine weiterhin auffindbar. Die Kritik, dass das Urteil gegen die Presse- und Meinungsfreiheit verstosse, ist jedoch nicht haltbar. Da das Objekt selbst noch vorhanden ist oder sein kann, wird mit der Löschung des Links aus der beliebten Suchmaschine höchstens die Suche nach den gewünschten Informationen erschwert. Freiheiten werden damit nicht beschnitten.

⁷⁹⁵ Siehe dazu: Weigert, Martin: Onlinepräsenzen gehen verloren: Die mobile Revolution und das vergessliche Web, netzwertig, 15.01.2014, URL: <http://netzwertig.com/2014/01/15/onlinepraesenzen-gehen-verloren-die-mobile-revolution-und-das-vergessliche-web/>, Stand: 06.03.2014.

⁷⁹⁶ Ankerson, Megan: Writing web histories with an eye on the analog past, in: New Media & Society (May 2012), 12.04.2012, S. 1 f. DOI: 10.1177/1461444811414834.

⁷⁹⁷ Siehe: Lischka, Konrad: Studie zur Web-Haltbarkeit: Das Netz vergisst schnell -, SPIEGEL ONLINE – Nachrichten – Netzwelt, URL: <http://www.spiegel.de/netzwelt/web/web-archiv-studie-zur-haltbarkeit-von-online-quellen-a-856936-druck.html>, Stand: 25.09.2012; Furness, Hannah: Hay Festival 2013: Teenagers' mistakes will stay with them forever, warns Google chief Eric Schmidt, in: The Telegraph, 25.05.2013. URL: <http://www.telegraph.co.uk/technology/eric-schmidt/10080596/Hay-Festival-2013-Teenagers-mistakes-will-stay-with-them-forever-warns-Google-chief-Eric-Schmidt.html>, Stand: 01.08.2014.

⁷⁹⁸ Siehe: EuGH: Urteil des Europäischen Gerichtshofes; Rechtssache C-131/12, 2014.

Die Adresspersistenz wird auch bei der Verwendung von *URL-Shortener* beeinträchtigt. Ein URL-Shortener ist ein Dienst zur Verkürzung von URL-Adressen, der einen kurzen Link auf den ursprünglichen Link erstellt, der auf das digitale Objekt verweist.⁷⁹⁹ Beispielsweise lautet die vollständige URL des Studienbereichs *Neuere und neueste Geschichte* des Departments Geschichte der Universität Basel:

<https://dg.philhist.unibas.ch/bereiche/neuereneueste-geschichte/>.

Mit dem zwischengeschalteten Dienst Bitly wird die URL auf <http://bit.ly/1EuPp3T> verkürzt. Die Zielwebseite kann direkt mit der Bitly-URL aufgerufen werden. Problematisch ist jedoch, dass der Dienst die Weiterleitung dauerhaft zur Verfügung zu stellen hat. Die Einstellung des Shortener-Dienst *Cli.gs* im Juni 2009 nach einem erfolgreichen Angriff zeigt auf, dass dies kaum möglich ist. Ein Hacker leitete damals mehr als 2,2 Millionen Kurzlinks von *Cli.gs* auf eine andere Webseite um, worauf der Dienst zuerst temporär und kurz darauf ganz offline ging.⁸⁰⁰ Weil bei solchen Diensten die Adresspersistenz nicht gewährleistet ist und sie zudem ein Sicherheitsrisiko darstellen⁸⁰¹, sollten URL-Shortener nicht verwendet werden. Wenn eine potenzielle Quelle verwendet wird, die wie in einem Tweet gekürzte URL aufweist, ist die Quelle mit aufgeschlüsselter URL zu sichern.

Um die Adresspersistenz zu gewährleisten, wurden verschiedene Techniken entwickelt, wie beispielsweise *Permalinks*.⁸⁰² Darunter wird ein gewissermassen ewig gültiger Link auf ein einzelnes Objekt wie ein Blog-Post oder ein Newseintrag verstanden, der den dauerhaften Zugriff unter dieser URL gewährleisten soll.⁸⁰³ Aber der Permalink suggeriert nur eine Persistenz, die weder die permanente URL noch das digitale Objekt erfüllen kann. Wenn einem Objekt ein Permalink zugewiesen wird, ist es zwar unter dieser eindeutigen Adresse erreichbar, es kann aber weiterhin mutiert oder gelöscht werden. Der Permalink selbst kann verändert werden, beispielsweise durch den Autor oder durch den Host.

⁷⁹⁹ Dienstanbieter sind bspw. Google (URL: <http://goo.gl/>), Bitly (URL: <http://bitly.com/>), TinyURL (URL: <http://tinyurl.com/>).

⁸⁰⁰ Nach: Bachfeld, Daniel: 2,2 Millionen URLs bei URL-Verkürzerdienst manipuliert, Newsticker, heise Security, 17.06.2009, URL: <http://www.heise.de/security/meldung/2-2-Millionen-URLs-bei-URL-Verkuerzerdienst-manipuliert-182569.html>, Stand: 30.04.2015.

⁸⁰¹ Die Kurz-URL kann auf eine mit Malware verseuchte Webseite oder einen Webauftritt eines Onlinebetrügers verweisen. Viele Nutzer klicken diese URLs nicht mehr an und in vielen Institutionen und Unternehmen ist die Nutzung solcher Dienste nicht möglich.

⁸⁰² Permalink: Zusammensetzung aus 'permanent' und 'Link'. Definition von Permanenz: "The ability of paper to last at least several hundred years without significant deterioration under normal use and storage conditions in libraries and archives." National Information Standards Organization: Permanence of Paper for Publications and Documents in Libraries and Archives, Standard ANSI/NISO Z39.48-1992 (R2009), Baltimore (USA) 2009, S. 2. URL: <https://www.niso.org/publications/ansiniso-z3948-1992-r2009-permanence-paper-publications-and-documents-libraries-and>, Stand: 06.05.2018.

⁸⁰³ Nach: Fischer; Hofer: Lexikon der Informatik, 2011. Beispiel: unter der URL hsc.hypotheses.org sind mehrere Beiträge zu finden, 'rutschen' jedoch nach einer gewissen Zeit auf eine Unterseite. Der direkte Zugriff auf den Beitrag 'just a historian' ist durch den Permalink hsc.hypotheses.org/241 möglich.

Valentin Groebners Hoffnung, dass der 2005 eingeführte *Digital Object Identifier (DOI)* das Persistenzproblem dauerhaft lösen werde, dürfte für die meisten digitalen Objekte ein frommer Wunsch bleiben.⁸⁰⁴ DOIs sind ein Identifikationsmerkmal und funktionieren analog einer ISBN-Nummer⁸⁰⁵ für Bücher. Mit einem DOI verknüpfte Objekte sind unabhängig vom Speicherort aufrufbar. Dieser kann vom Autor frei gewählt werden. Die *International DOI Foundation*⁸⁰⁶ stellt mit einer Datenbank den entsprechenden Informationsdienst zur Verfügung, der ähnlich einem URL-Shortener die Aufrufe auf den momentanen Speicherort umleitet. Zwar könnte für jedes digitale Objekt ein DOI beantragt werden⁸⁰⁷, aber der Prozess ist aufwendig und kann nicht für jedes einzelne Objekt, das möglicherweise in einem Forschungsprojekt verwendet werden könnte, durchgeführt werden. Würde nur die Homepage aller Webseiten als PDF gespeichert und mit einem DOI versehen werden, wären knapp eine Milliarde Einträge⁸⁰⁸ nötig, was kaum zu bewältigen ist. DOIs können zwar den Zugriff auf ein Objekt unabhängig seines Speicherorts sicherstellen, aber nicht die Persistenz des Objekts selbst. Wenn es gelöscht wird, kann es nicht mehr aufgerufen werden. Trotzdem wird der DOI als *persistentes Identifikations- und Adressierungssystem*⁸⁰⁹ bezeichnet.

Ein weiteres persistentes System ist der vor allem in der Wissenschaft verwendete Uniform Resource Name URN.⁸¹⁰ Auch bei diesem wird dem Objekt eine eindeutige Bezeichnung zugeordnet, mit der es unabhängig vom Speicherort aufgerufen werden kann. Der URN ist zudem in verschiedene Namensräume für Institutionen und Subräume wie der ISBN für Bücher aufgeteilt, um die Objekte strukturierter verzeichnen zu können. URN werden hauptsächlich von Bibliotheken oder ähnlichen Institutionen vergeben, die mit dem Eintrag des URN meist auch eine Archivkopie respektive einen Klon anfertigen. Der Eigentümer kann dann zwar sein Objekt auf seiner Infrastruktur löschen, es steht aber unter der URN in der Infrastruktur der Institution weiterhin zur Verfügung.

⁸⁰⁴ Groebner spricht von einem 'Domain Object Identifier', gemeint ist jedoch der Digital Object Identifier. Groebner: *Wissenschaftssprache digital. Die Zukunft von gestern*, 2014, S. 104. Weitere Informationen zum DOI siehe URL: <http://www.doi.org>.

⁸⁰⁵ ISBN: International Standard Book Number. Beispiel einer Adressierung: urn:isbn:978-3-486-70704-5 für Peter Habers 'Digital Past'.

⁸⁰⁶ URL: <http://www.doi.org/>.

⁸⁰⁷ Siehe zum Verfahren mit 'Zenodo': URL: <https://guides.github.com/activities/citable-code/>.

⁸⁰⁸ Stand Februar 2015. Netcraft zählt in der monatlichen Web Server Survey knapp eine Milliarde Webseiten. URL: <http://news.netcraft.com/archives/category/web-server-survey>; Stand: 30.7.2014. Netcraft ist ein renommiertes Unternehmen für die Erhebung von Forschungsdaten zum Internet.

⁸⁰⁹ Siehe bspw.: Runkehl, Jens; Siever, Torsten: Zitieren und Belegen, in: Gasteiner, Martin; Haber, Peter (Hg.): *Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften*, Wien, Köln, Weimar 2010, S. 135 f. Siehe auch: URL: <http://www.persistent-identifier.de>.

⁸¹⁰ Der URN scheint sich aber noch nicht durchgesetzt zu haben. Siehe: Schweizerische Nationalbibliothek: e-Helvetica, URN-Handbuch, 04.2014. URL: <https://www.nb.admin.ch/snl/de/home/nb-professionell/e-helvetica/dokumente/dokumente-zur-urn--uniform-resource-name-.html>, Stand: 06.05.2018.

Sowohl beim DOI als auch beim URN müssen die zugehörigen Informationen von einem zwischengeschalteten Resolver-Dienst einer vertrauenswürdigen und dauerhaft betriebenen Institution verwaltet werden, der die Anfragen auf entsprechende URL umleitet. Jedes Objekt muss darin registriert und jegliche Änderung seines Speicherplatzes muss gemeldet werden. Allerdings stellt sich die Frage wie der Dienst Kenntnis von der neuen Adresse des Objekts erhält, damit die Datenbank des Resolvers aktualisiert werden kann.

Lösungsansatz Persistentes Identifikations- und Adressierungssystem im TDR

Für digitale Forschungsressourcen wird ein grosser Umfang an *persistenten Identifikatoren (PI)* benötigt. Zudem ist die Verwendung der Identifikatoren von Resolver-Diensten loszulösen, um nicht von diesen abhängig zu sein. Diese Anforderungen könnten mit gesetzlich beauftragten TDR umgesetzt werden, die digitale Objekte mit einem PI versehen und unter einer permanenten Adresse publizieren. Der PI ist entsprechend auch in den Metadaten aufzuführen.

Der PI könnte wie folgt aufgebaut sein: ISIL-Nummer⁸¹¹ des Repository, Nutzernummer, Projekt- / Sammlungsnummer, Laufnummer.

Fiktives Beispiel PI: CH-101111-1.00120.A0Y4.0218

Um keinen Resolver-Dienst zu benötigen, wird die URL des TDR der PI vorangestellt:

URL (fiktiv): <https://www.tdr.bar.ch/CH-101111-1.00120.A0Y4.0218>

Das Objekt kann damit eindeutig identifiziert werden, ist unabhängig vom Speicherort und kann ohne Resolver-Dienst direkt aufgerufen werden. Das TDR ist per definitionem dafür besorgt, dass die Adressierung und Identifikation tatsächlich persistent bleibt. Um dies gewährleisten zu können, sollte die *Domain eines gesetzlich beauftragten TDR für den Verkauf gesperrt* werden. Damit kann die Adresspersistenz dauerhaft sichergestellt werden, auch wenn der Provider wechselt oder eine andere Informationssystemstruktur verwendet wird.

5.4.4. Objektbeschreibung

Bei digitalen Objekten umfasst die Objektbeschreibung nicht nur eine Beschreibung seiner dargestellten Form, sondern auch die informationstechnische Beschreibung seines Formats: Dateien benötigen Formate. Diese erst ermöglichen die Darstellung des Objekts, weshalb das Dateiformat vor der Interpretation zu überprüfen ist. Bei digitalen Objekten ist zudem zu untersuchen, ob es eine Inszenierung oder Teil einer solchen ist.

⁸¹¹ ISIL: International Standard Identifier for Libraries and Related Organizations. Standard: ISO 15511:2011, siehe URL: <http://www.iso.org/>.

Die Objektbeschreibung nimmt auch bei digitalen Objekten eine zentrale Stellung ein. Es ist zwischen der Beschreibung der dargestellten Form mit Elementen wie Inhalt, Datierung, Überlieferungsgeschichte oder Quellengattung und der informationstechnischen Beschreibung zu unterscheiden.⁸¹² Da das Objekt nur mit einem elektronischen Informationssystem dargestellt werden kann, muss dieses über alle nötigen Informationen und Komponenten verfügen. Befindet sich das Objekt beispielsweise auf einem Datenträger, muss dieser zuerst ausgelesen werden können. Das Lesegerät muss dabei die vom Datenträger geforderten technischen Spezifikationen erfüllen, ansonsten das Objekt vom Informationssystem nicht gelesen und verarbeitet werden kann. Zudem sind Kenntnisse über den Inhalt des digitalen Objekts und über Dateiendungen von Nutzen, denn sie können Hinweise auf das einzusetzende Informationssystem geben. Ist beispielsweise bekannt, dass es sich bei einem digitalen Objekt um ein Textdokument handelt, können Worterkennungsprogramme für dessen Darstellung eingesetzt werden. Ein Programm für das Auslesen von Bild-Metadaten (EXIF-Format) wäre dagegen nicht nötig. Dateinamen und -endungen sind jedoch nicht verlässlich, weil sie sehr einfach verändert werden können. Eine Datei mit dem Namen 'Katzenbild.pdf' beinhaltet nicht immer ein (Katzen-)Bild und ist nicht immer im PDF-Format codiert. Wenn das Objekt mit dem dafür vorgesehenen Programm nicht geöffnet werden kann, kann dies verschiedene Ursachen haben: die Dateiendung kann falsch, das Dateiformat obsolet oder der Bitcode korrupt sein.

Das Erkennen des korrekten Dateiformats⁸¹³ erleichtern diverse Softwares wie *DROID* oder *JHOVE*.⁸¹⁴ Der Algorithmus der Software untersucht den Code des Objekts auf Merkmale ihm bekannter Dateiformate. Als Resultat dieser automatisierten Untersuchung wird das wahrscheinlichste Format des Objekts ausgegeben, was nicht zwangsläufig dem tatsächlichen entsprechen muss. Ohne weiter auf Details einzugehen, ist hier zu erwähnen, dass dieser Vorgang als eine Form von digitaler Quellenkritik darstellt, der die weitere kritische Überprüfung massgeblich beeinflusst. Dem Resultat entsprechend wird für die weitere Untersuchung eine für die Darstellung geeignete Software verwendet. Es lässt sich jedoch in den meisten Fällen nicht nachvollziehen, ob das Resultat korrekt und ob es für die Darstellung eine geeignetere Software gäbe. Eine falsche oder unpräzise Reproduktion ist quellenkritisch sehr problematisch, besonders wenn sie nicht als solche erkannt wird.

In der Objektbeschreibung sind eventuell auch Informationen enthalten, die das Objekt als *Inszenierung* oder als Teil einer Inszenierung identifizieren. Besonders Animationen oder

⁸¹² Siehe dazu auch: Schassan, Thorsten: Digitale Quellen: Datei- und Datenformate, in: Busse, Laura; Enderle, Wilfried; Hohls, Rüdiger u. a. (Hg.): Clio Guide. URL: <http://guides.clio-online.de/guides/arbeitsformen-und-techniken/digitale-quellen-dateiformate/2016>, Stand: 10.07.2016.

⁸¹³ Zu Daten- und Dateiformaten vgl. auch: Ebd.

⁸¹⁴ DROID siehe URL: <http://digital-preservation.github.io/droid/>. JHOVE siehe URL: <http://jhove.openpreservation.org/>.

audiovisuelle Darstellungen werden inszeniert.⁸¹⁵ Erkennt ein Nutzer die Inszenierung nicht als solche, kann diese eine Authentizität suggerieren, die den Nutzer in seiner Realitätswahrnehmung täuscht und zu falschen Aussagen oder ahistorischen Darstellungen verleiten kann. Das Erkennen von für die erneute Darstellung zulässigen oder vernachlässigbaren Elementen, die das Objekt verändern oder verfälschen, ist jedoch nicht einfach. Die Geschichtsdidaktikerin Astrid Schwabe stellte aus Nutzersicht fest:

"Die Grenzen zwischen Rekonstruktionen mit zulässigen Ergänzungen, Komplexitätsreduktionen oder 'Stilisierungen' und inadäquater Vereinfachungen oder verschleierten Inszenierungen sind schwer auszumachen."⁸¹⁶

Die Objektbeschreibung kann oder sollte nicht nur Hinweise dazu geben, welche Komponenten eines Informationssystems zwingend benötigt werden, sondern auch, worauf der Autor besonders Wert legte und welche Elemente vernachlässigt werden können. Diese Informationen sind nicht nur für die Interpretation von virtuellen, sondern auch von analogen Objekten nützlich. Die dargestellte Form ist durch den Betrachter zu reflektieren und auf Konstruktionen und Rekonstruktionen zu überprüfen. Die Schwierigkeit bei digitalen Objekten ist vor allem durch ihre fehlende Materialität bedingt, die nicht oder nur schwerlich überprüft werden kann.

5.4.5. Echtheit

Die Überprüfung der Echtheit eines digitalen Objekts erfolgt durch die Betrachtung seiner *Überlieferungsgeschichte*. Da die Materialität nicht untersucht werden kann, ist die Unterscheidung zwischen Echtheit oder Fälschung beziehungsweise Irrtum anhand der allenfalls vorhandenen Objektvarianten respektive -versionen und Metadaten zu beurteilen. Hierbei spielen die Vertrauenswürdigkeit der das Objekt publizierenden Institution und das Verhältnis zwischen Objekt und Institution wichtige Rollen.

Für die Untersuchung der Überlieferungsgeschichte unterscheide ich drei Ausgangslagen: es liegt a) ein singuläres Objekt, b) ein Objekt in mehreren Varianten, c) ein archiviertes Objekt vor.

a) Singuläres digitales Objekt

Unter *singulär* wird ein Objekt verstanden, bei dem keine Varianten oder Klone bekannt sind. Aufgrund der fehlenden Materialität können weder Alter, Ort noch zeitliche

⁸¹⁵ Siehe dazu: Schwabe: Historisches Lernen im World Wide Web, 2012, S. 140.

⁸¹⁶ Ebd.

Veränderungen festgestellt werden.⁸¹⁷ Eine Überlieferungsgeschichte kann deshalb nur anhand der vorhandenen Metadaten oder inhaltlicher Informationen rekonstruiert werden. Dies ist nur möglich, falls das Objekt unmittelbar oder zufälligerweise dargestellt werden kann. Beispielsweise wird bei einem unbekanntem Dateiformat durch die Trial-and-Error-Methode⁸¹⁸ versucht eine darstellbare Form zu finden, um Informationen über das konkrete Format auszulesen.

Einige Datenformate erlauben die Protokollierung einer Art von Überlieferungsgeschichte innerhalb des Objekts. Beispielsweise werden in Microsoft Word-Dokumenten Zwischenstände des Manuskripts mit den Funktionen *AutoWiederherstellen* und *Schnellspeicherung* erstellt, die bei einem allfälligen Informationssystemabsturz wieder aufgerufen werden können. Solche versteckten Informationen (Hidden Data) können mit entsprechenden Programmen ausgelesen, aber auch manipuliert und gelöscht werden.

b) Objektvarianten

Liegt ein digitales Objekt in mehreren Varianten vor, ist zuerst zu überprüfen, ob es sich tatsächlich um Varianten, Variantencluster oder allenfalls um eigene Objekte handelt. Das ist bei einem Textdokument wahrscheinlich offensichtlicher erkennbar als bei künstlerischen Objekten. Beispielsweise hat der Nutzer oder Betrachter von ähnlichen Bildobjekten für deren Überlieferungsgeschichte in Erfahrung zu bringen, ob der Autor verschiedene Varianten oder singuläre Objekte schaffen wollte. Wenn die Intention des Autors nicht vorher bekannt war, sind eventuell in den Metadaten oder im Objekt selbst Hinweise darüber zu finden. Es sind deshalb die Metadaten der Objektvarianten und deren inhaltliche Informationen zu überprüfen, um die Überlieferungsgeschichte erarbeiten zu können. Es ist dabei zu beachten, dass das jüngste Objekt nicht zwingend das relevante darstellt, wenn beispielsweise aus mehreren Entwürfen eines Briefes, eines Gesetzentwurfs oder eines Plakates einer ausgewählt wird.

c) Archiviertes Objekt

Ein archiviertes digitales Objekt wurde von einer vertrauenswürdigen und dauerhaften Institution wie dem Landesarchiv zusammen mit Metadaten gesichert und zur Nutzung zur Verfügung gestellt. Ob das Objekt und dessen Metadaten dargestellt wer-

⁸¹⁷ Nach: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 405.

⁸¹⁸ Trial-and-Error-Methode: Versuch und Irrtum. Dabei werden Lösungsmöglichkeiten bis zu einem Erfolg abgearbeitet resp. durchprobiert. Das Problem ist jedoch, dass ein 'Erfolg' nicht die gesamte Funktionalität des Objekts zur Verfügung stellen muss. Enthält bspw. ein Textdokument 'anklickbare' Hyperlinks, kann das Dokument evtl. in einem Texteditor dargestellt werden, der aber diese Funktion nicht zur Verfügung stellt.

den können, ist allerdings zunächst nicht sicher. Erst wenn eine korrekte Darstellung des Objekts und der zugehörigen Metadaten möglich ist, kann die allenfalls vorhandene Überlieferungsgeschichte überprüft werden. In vielen Fällen ist ein archiviertes Objekt singular und ist das relevante von allenfalls mehreren Versionen, die nicht gespeichert wurden.⁸¹⁹

Unabhängig von der Ausgangslage sind für die Prüfung der Überlieferungsgeschichte eines digitalen Objekts vertrauenswürdige Metadaten und die Informationen aus der dargestellten Form essenziell.

Ein wichtiges Indiz für die Echtheit der Informationen ist wie bei physischen Objekten die *Vertrauenswürdigkeit der Institution*, von der das Objekt bezogen wird.⁸²⁰ Institutionen, die die langzeitliche Aufbewahrung von digitalen Objekten nach archivischen Standards sicherstellen, kommt eine zentrale Bedeutung für die Beurteilung der Echtheit zu. Beispielsweise kann ein Nationalarchiv, das mit entsprechenden Prozessen die Langzeitarchivierung von digitalen Objekten garantiert, als vertrauenswürdig angesehen werden. Dagegen ist die Webseite eines kleinen Handwerksbetriebs weniger vertrauenswürdig, weil die Unternehmung vermutlich, sofern nicht anderweitig verifiziert, keine die langzeitliche Datenintegrität sicherstellenden Prozesse sowie überlieferungsgeschichtliche Informationen zur Verfügung stellt. Für die Beurteilung der Echtheit sollten alle Archivierungsschritte wie Migrationen nachvollzogen sowie alle Integritätsnachweise überprüft werden können, um den archetypischen Zustand des Objekts zu beurteilen.⁸²¹ Dies ist vor allem bei den erwähnten archivierten Objekten möglich, deren Überlieferungsgeschichte mit vertrauenswürdigen Metadaten durch die Institution sichergestellt wird.

Ein weiteres Indiz ist das *Verhältnis des Objekts zur Institution*. Wenn das Objekt inhaltlich mit dem Fokus der Institution übereinstimmt, werden die Informationen vertrauenswürdiger sein, als wenn sich die Themenbereiche von Objekt und Institution stark unterscheiden. Beispielsweise wird die Dokumentation über die Standards von sanitären Anlagen auf einer Webseite eines Sanitärinstallationsbetriebs wahrscheinlich 'echter' sein als im Repository eines universitären historischen Seminars.

Wie bei physischen Quellen sind auch bei digitalen Quellen die vorhandenen Informationen über die Quelle zu verifizieren. Deren Echtheit kann erst nach der vollständig abgeschlossenen Quellenkritik beurteilt werden.

⁸¹⁹ Behördliche Archive übernehmen fast ausschliesslich Endversionen von Dokumenten. Nur in Ausnahmefällen werden Varianten übernommen, um bspw. den Entstehungsprozess zu dokumentieren (in den Metadaten nachgewiesen).

⁸²⁰ Vgl.: Margulies: *Digitale Daten als Quelle der Geschichtswissenschaft*, 2009, S. 407.

⁸²¹ Vgl.: Ebd., S. 406.

5.4.6. Datierung

*"Wenn es also bei digitalen Quellen medial bedingt keine dateninhärente chronologische Unterscheidungsmöglichkeit von 'früher' und 'später' gibt, müssen andere, an die medialen Gegebenheiten angepassten Kriterien entwickelt werden."*⁸²² (Haber, 2011)

Eine Datierung eines singulären digitalen Objekts und die damit zusammenhängende Frage nach dem Früheren und dem Späteren kann ohne zugehörige und vertrauenswürdige Metadaten nicht eindeutig vorgenommen werden. Wenn Änderungen an einem Objekt nicht protokolliert und mit ihm verknüpft werden, wird ein diakritisches Verfahren ergebnislos bleiben.⁸²³ Das Vertrauen in die ein Objekt publizierende Institution ist zentral, weil diese zur Integrität des Objekts auch integrale und verlässliche Metadaten mitliefert. Dafür können verschiedene technische Systeme eingesetzt werden, die mit Zeitstempel, Protokollierung und Versionierung die Überlieferungsgeschichte eines Objekts darstellen können. Wenn unterschiedliche plausible Datumsangaben vorliegen, hat der Nutzer sich für eine zu entscheiden und den Entscheid zu begründen. Digitale Objekte mit einer digitalen Signatur und Zeitstempel können eindeutig datiert werden.

Ebenso wie das digitale Objekt manipuliert werden kann, können auch gezielt Metadaten wie Datums- und Zeitangaben manipuliert werden. Beispielsweise kann der Aufnahmezeitpunkt eines Bildes im Microsoft Datei-Explorer oder einem Programm wie ExifTool⁸²⁴ geändert werden. Eine Rückführung des vorliegenden Objekts in den Archetypus ist deshalb nur dann möglich, wenn die Überlieferungsgeschichte des Objekts vollständig und vertrauenswürdig ist.

Für die Datierung ist wiederum das *Vertrauen in die Institution* zentral, die die Informationen zur Verfügung stellt. Wenn angenommen werden kann, dass diese eine adäquate Prüfung des digitalen Objekts vor dem Ingest ins Archiv durchgeführt hat und die archivischen Prozesse korrekt ablaufen, kann mittels der vorhandenen Metadaten eine Datierung mit hohem Wahrscheinlichkeitsgrad erfolgen – zumindest seit dem Zeitpunkt der Archivierung durch die Institution. Auch ist eine Rückführung auf den Archetypus eines Objekts möglich, weil alle nötigen Informationen in den Metadaten und in den Objektversionen zur Verfügung stehen. Die Metadaten sind auf entsprechende Datierungselemente zu durchsuchen. Die vorhandenen Informationen sind auf ihre Plausibilität hin zu überprüfen und entsprechend zu

⁸²² Haber: Digital Past, 2011, S. 107 f.

⁸²³ Margulies geht davon aus, dass bei jedem Objekt der Archetypus festgestellt werden kann, was aber nur möglich ist, wenn die dafür nötigen Informationen überhaupt vorhanden sind. Siehe: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 401.

⁸²⁴ URL: <http://www.sno.phy.queensu.ca/~phil/exiftool/>.

interpretieren. Beispielsweise könnte die Integrität des Objekts kompromittiert sein, wenn das Datum von `date(creation)` nach demjenigen von `date(collection)` liegt.

Informativer und verlässlicher sind Datumsangaben, welche vom erstellenden System automatisch erzeugt und ins digitale Objekt integriert, oder im System hinterlegt werden. Einige Records / Data Management Systeme, die hauptsächlich in grösseren Unternehmen oder staatlichen Verwaltungen eingesetzt werden, erfassen genaue Zeitangaben mittels *Zeitstempel und/oder Protokolleinträgen*.⁸²⁵ Damit wird innerhalb des Systems sichergestellt, dass jede Änderung an einem Objekt nachvollzogen werden kann. Die Einbettung eines Objekts in ein solches System ist aber für eine quellenkritische Überprüfung problematisch, denn es ist fraglich, ob auf das Objekt und die dabei entstehenden Metadaten zugegriffen werden kann. Zudem verwenden solche Systeme oft proprietäre Datenformate, die bei der Überführung in ein Langzeitarchiv migriert werden müssen, was meistens mit einem Datenverlust verbunden ist. Beim Herauslösen des Objekts aus dem System sind auch die zugehörigen Metadaten zu exportieren und mit dem Objekt zu verknüpfen. Zudem sind diese Daten, da sie sich nun nicht mehr im abgeschotteten System befinden, vor Manipulationen zu schützen.

Obwohl mit RM und DM Systemen qualitativ hochwertige Metadaten erzeugt werden können, ist das Vertrauen in deren Echtheit von der Institution abhängig. Amtlichen Verwaltungseinrichtungen wird konventionell eher vertraut als einer wirtschaftlich orientierten Unternehmung, die mit einem proprietären RMS ein Objekt mit zugehörigem Änderungsprotokoll zur Verfügung stellt. Der Einsatz von Standards wie ISO 15489-1/-2⁸²⁶ und unabhängige, regelmässige (Re-)Zertifizierungen stärken das Vertrauen in eine Institution, wenn dies für den Nutzer klar erkenn- und nachprüfbar ist.⁸²⁷ Informationen darüber sollte die Institution deshalb aus Eigeninteresse veröffentlichen.

Datierungsinformationen auf Webseiten und Blogs oder in Newslettern sind wenig vertrauenswürdig, weil diese bei den meisten Systemen mit wenig Aufwand manipuliert werden können und bei den wenigsten Systemen eine Versionierung mit Protokollierung erfolgt und publiziert wird. Hinweise auf eine erfolgte Aktualisierung, beispielsweise mit [UPDATE – Datum], sind freiwillige Zusatzinformation des (Erst-)Verfassers und sollen die Seriosität der Publikation unterstreichen oder vortäuschen. Was genau aktualisiert wurde, kann der Nutzer trotzdem nicht nachvollziehen, denn der Autor oder ein Editor kann nach dem Hinweis eine

⁸²⁵ Eingehender: Haber: digital past (unv. Habil.), 2009, S. 197; Gladney: Preserving Digital Information, 2007, S. 171 f.

⁸²⁶ ISO 15489-1/-2: Information and documentation -- Records management -- Part 1: General; Part 2: Guidelines. Siehe URL: <http://www.iso.org>.

⁸²⁷ Bspw. durch einen Zertifizierungshinweis auf der Webseite der Unternehmung. Die Überprüfung der Zertifizierung sollte möglich sein, weil ein Zertifikat o.ä. auch gefälscht werden kann (insb. bei Online-Shops werden gerne gefälschte Qualitätszertifikate verwendet, um Seriosität vorzutäuschen).

Ergänzung eingefügt und/oder am vorher bestehenden Objekt Mutationen vorgenommen haben. Für wissenschaftliche und wissenschaftsnahe Online-Plattformen wäre die Verwendung eines Publishingtools sinnvoll, das jede Änderung an einem Objekt protokolliert und versioniert, und diese Informationen dem Nutzer wie in der Wikipedia zur Verfügung stellt. Mit der damit gewährleisteten Nachvollziehbarkeit aller Veränderungen am publizierten Objekt wird die Überlieferungsgeschichte plausibler und würde das Vertrauen in die Institution weiter stärken.

Die Datierung eines Objekts kann unter Umständen auch durch *inhaltliche Hinweise oder durch verknüpfte Objekte* erfolgen. Dabei wird es wie ein physisches Objekt auf inhaltliche Informationen hin untersucht, die Rückschlüsse auf die Überlieferungsgeschichte erlauben. Verknüpfungen zu anderen Objekten, die sich innerhalb des Informationssystems, oder in einem anderen Informationsraum wie dem Web oder der Gedächtnisinstitutionen befinden, können durch die darin enthaltenen Informationen eine Datierung des vorliegenden Objekts erlauben. Hierbei sollten auch Informationen aus physischen Beziehungsobjekten für eine Datierung des digitalen Objekts hinzugezogen werden, womit wieder hybrides Recherchieren gefordert ist. Beziehungsobjekte in virtuellen Räumen weisen die bekannten Probleme wie Manipulation oder fehlende Metadaten auf und sind entsprechend auf die Plausibilität ihrer Angaben zu überprüfen. Da digitale Objekte durch ihre Hypertextualität fast unendlich miteinander verknüpft sein können, muss die Überprüfung zwangsläufig nach einem verhältnismässigen Aufwand, der sich nach der Erfahrung des Recherchierenden und der Plausibilität der vorhandenen Informationen ermisst, abgebrochen werden.

Die Suche nach verlässlichen Datumsangaben kann unterschiedliche, dennoch plausible Informationen ergeben. Der Nutzer hat nun zu entscheiden, welches der vorgefundenen Daten verwendet werden soll. Auf einer jede Änderung protokollierenden Kollaborationsplattform kann dies offensichtlich sein, bei einem Dokument in einem Repository ist die Entscheidung aber schwieriger. Unterscheiden sich beispielsweise das Publikations- und das im Dokument enthaltene Erstelldatum, ist nach dem Grund dafür zu suchen. Oft wird das Erstelldatum verwendet, das aber aus diversen Gründen vom Autor manipuliert worden sein kann. Nimmt nun der Nutzer dieses Dokument in seine digitale Sammlung auf, entsteht zusätzlich das Datum der Quellensicherung. Falls dieses die erste, verlässliche Angabe darstellt, weil der Nutzer einzig der Sicherungsplattform vertrauen kann, wäre dieses Datum zu verwenden. Wird das Dokument als Quelle verwendet, hat der Nutzer auf das Zustandekommen der Datierung hinzuweisen.

Eine eindeutige Datierung kann nur dann erfolgen, wenn vertrauenswürdige Institutionen und Mechanismen die Korrektheit der Angaben garantieren. Eine der zurzeit verlässlichsten Techniken ist die *fortgeschrittene oder qualifizierte elektronische Signatur mit Zeitstempel*. Einem damit versehenen Objekt kann das nötige Vertrauen sogar unabhängig von der Insti-

tution entgegengebracht werden, weil die Informationen eindeutig und zertifiziert sind. Das Signaturdatum bezeugt, dass danach keine Mutation mehr vorgenommen worden ist.

Lösungsansätze Datierung

Wissenschaftliche und wissenschaftsnahe Publikationen sollten mit einer *qualifizierten oder fortgeschrittenen elektronischen Signatur mit Zeitstempel* versehen werden. Der Zeitstempel ist eine eindeutige, den wissenschaftlichen Ansprüchen entsprechende und vertrauenswürdige Angabe.

Publikationsplattformen für wissenschaftliche oder wissenschaftsnahe Informationen sollten Publishingtools verwenden, die Änderungen an den publizierten Objekten wie Beiträge, Dokumente oder Bilder protokolliert, versioniert und diese Informationen den Nutzern zur Verfügung stellt. Damit könnte die Überlieferungsgeschichte der potenziellen Quelle nachvollzogen und das Vertrauen in die Institution gestärkt werden.

5.4.7. Entstehungs- / Erscheinungsort

Die Mobilität der Autoren und die Entwicklung von elektronischen Publikationsmedien haben die Definition von Entstehungs- und Erscheinungsorten erschwert. Kollaborative Arbeiten verschiedener Autoren aus mehreren Kontinenten an einem einzelnen digitalen Objekt sind keine Seltenheit mehr. Die Verwendung des juristischen Sitzes einer Publikationsgesellschaft als Erscheinungsort wird hinfällig, weil das Objekt mehrfach, an verschiedenen Orten gleichzeitig sowie verlinkt publiziert werden kann. Deshalb sollte auf die Angabe der Entstehungs- und Erscheinungsorte verzichtet werden und an deren Stelle nur noch der persistente Identifikator verwendet werden.

Der Entstehungsort eines Objekts ist selten eindeutig und wird durch Verhandlung oder Konvention festgelegt. Beispielsweise wird ein Buch von einem Autor an seinem physischen Arbeitsort geschrieben, die Druckerei aber befindet sich an einem anderen Ort und der juristische Sitz des herausgebenden Verlags befindet sich wieder woanders. Per Konvention wird bei physischen Publikationen der juristische Sitz des Verlags als zitierfähiger Erscheinungsort im Impressum angegeben, ohne Rücksicht auf den oder die tatsächlichen Entstehungsort(e). Bei Objekten ohne Angabe eines vertrauenswürdigen Erscheinungsorts ist objektspezifisch zu entscheiden, ob die Fundstelle, die Sammlung, der im Dokument erwähnte oder ein anderer Ort angegeben wird, oder ob auf die Tatsache des fehlenden Erscheinungsorts hingewiesen wird.

Bei einem digitalen Objekt ist der Entstehungsprozess ähnlich, denn es kann vom Autor 'irgendwo' generiert und bearbeitet werden. Heutige Informationssystemen und gute Internetabdeckung erlauben eine hohe Mobilität. Dem Autor stehen verschiedene Online-Plattformen⁸²⁸ für die Arbeit zur Verfügung, die sich auf einem Informationssystem irgendwo auf der Welt befinden. Der Autor weiss meistens nicht, in welchem oder welchen physischen Datencenter(n) seine Eingaben verarbeitet und wo seine Daten gespeichert werden. Die Bestimmung eines Entstehungsorts wird bei kollaborativer Arbeit mehrerer Autoren, die gleichzeitig auf dasselbe Objekt von mehreren physischen Orten aus zugreifen, weiter erschwert.

Auch der Erscheinungsort ist oft nicht klar zu definieren. Die Publikationsgesellschaft hat ihren physischen Sitz an einem bestimmten Ort, aber die Publikation erfolgt eventuell durch eine Unternehmung an einem anderen Ort, deren Publikationsplattform sich in einem Datencenter irgendwo auf der Welt befindet. Zudem kann das Objekt gleichzeitig auf verschiedenen Plattformen publiziert und mit anderen Objekten verlinkt werden. Trotzdem wird bei digitalen Objekten für die Angabe des Erscheinungsorts meist auf die bestehenden Konventionen aus analoger Zeit zurückgegriffen. Häufig ist dies der *juristische Sitz der publizierenden Institution*. Beispielsweise ist der Erscheinungsort des Beitrages *Serf and Turf*⁸²⁹ abhängig von der publizierenden Institution. Das Dokument wurde von Mitgliedern der University of California Santa Barbara in Santa Barbara (USA) erstellt und auch dort publiziert.⁸³⁰ Zusätzlich findet man es auch auf der Publikationsplattform *arXiv.org* der Cornell University Library in Ithaca (US). Das Dokument wurde somit mindestens an zwei Orten publiziert und hat damit auch mindestens zwei Erscheinungsorte und Publikationsdaten.

Problematischer sind digitale Objekten, die nicht als Dokument publiziert werden. Bei einer Webseite könnte eventuell auf die Angaben im Domain-Whois⁸³¹ zurückgegriffen werden, in dem der Halter der Domain aufgeführt ist. Doch ein per E-Mail zugestelltes Bild hat nur einen Herkunfts-, aber keinen Erscheinungsort. Üblicherweise wird bei fehlenden Informationen zum Erscheinungsort der momentane *Speicherort*, das heisst die URL des Objekts verwendet. Da jedoch Speicher- und Erscheinungsort nicht gleichbedeutend und nicht voneinander abhängig sind, sollte eine URL nicht als Erscheinungsort verwendet werden.

Noch immer wird lieber auf die physische Publikation mit einem bestimmten Erscheinungsort verwiesen, auch wenn sie nur digital konsultiert wurde. Bei diesem wäre nicht der physische

⁸²⁸ Solche Dienste bieten bspw. Wikipedia (URL: <https://www.wikipedia.org/>), Etherpad (URL: <http://etherpad.org/>), Google Docs (URL: <https://docs.google.com>).

⁸²⁹ Wang, Gang; Wilson, Christo; Zhao, Xiaohan u. a.: Serf and Turf: Crowdturfing for Fun and Profit, arXiv:1111.5654 [cs.SI], 23.11.2011. URL: <http://arxiv.org/abs/1111.5654>, Stand: 03.09.2014.

⁸³⁰ URL: <http://www.cs.ucsb.edu/~gangw/crowdturf-www12.pdf>.

⁸³¹ Das Domain-Whois enthält Angaben über Halter des Domainnamens (Person, Institution, Firma). Siehe bspw. URL: <https://www.whois.net/>.

Erscheinungsort massgebend, sondern der digitale Publikationsort. Da aber qualitätssichernde und vertrauenswürdige Instanzen fehlen und eine Adresspersistenz selten gegeben ist, wird behelfsmässig auf den juristischen Erscheinungsort der Publikation zurückgegriffen, um eine vertrauenswürdige physische Quelle vorzutäuschen – oft ohne diese konsultiert zu haben.

Da aber der physische Erscheinungsort für digitale Objekte irrelevant ist, wird die *URL zur symbolischen Währung für deren Reputation*.⁸³² Wie schon mehrfach darauf hingewiesen, ist das Vertrauen in die publizierende Institution ausschlaggebend für die kritische Beurteilung der vorhandenen Angaben. Beginnt eine URL beispielsweise mit *http://unibas.ch/...* ist sie aus wissenschaftlicher Sicht vertrauenswürdiger als eine mit *http://bluewin.ch/...*, die vor allem dem Hosting von Webseiten von Privatpersonen dient.⁸³³ Für die Einschätzung der Vertrauenswürdigkeit ist der Halter der URL im Domain-Whois nachzuschlagen. Je vertrauenswürdiger der Halter ist, desto vertrauenswürdiger sind auch die unter seiner URL publizierten Dokumente.

Fach- und Themenportale, die durch den Meinungs austausch von Fachpersonen einer ständigen Qualitätskontrolle unterliegen und von vertrauenswürdigen Haltern wie Universitäten betrieben werden, werden im wissenschaftlichen Umfeld oft als gehaltvollste und vertrauenswürdigste Plattformen genannt.⁸³⁴ Allerdings ist unklar, welche Webauftritte als Fach- und Themenportal bezeichnet werden können und wie die Aufnahme in diesen Kreis erfolgt. Dies ist stark vom Vertrauen der Wissenschaftsgemeinschaft gegenüber dem Betreiber und seines Webauftritts abhängig. Das Vertrauen kann er mit entsprechenden wissenschaftlichen Leistungen erarbeiten. Beispielsweise hat sich *hypotheses.org* aus einem Blogportal unter vielen zu einem anerkannten Fachportal⁸³⁵ entwickelt, unter anderem weil nur auf berechtigten Antrag hin ein Blog eröffnet werden konnte und eine wissenschaftliche Trägerschaft und Redaktion für die Qualität der Inhalte verantwortlich zeichnet.⁸³⁶

⁸³² Nach: Haber: *digital past* (unv. Habil.), 2009, S. 72.

⁸³³ Domaininhaber von *unibas.ch* ist die Universität Basel; von *bluewin.ch* der Hoster Swisscom (Schweiz) AG.

⁸³⁴ Beispiele: *Clio-online* (<http://www.clio-online.de/>), *historicum.net* (<http://www.historicum.net/>), *infoclio* (<http://infoclio.ch/>).

⁸³⁵ Unter einem Fachportal wird ein Webauftritt für den Zugang zu einem wissenschaftlichen Themengebiet unter der Leitung eines wissenschaftlichen Fachgremiums verstanden. Welche Anforderungen gelten und welche Webauftritte als Fachportal bezeichnet werden können, wird aus eigener Erfahrung sehr unterschiedlich beurteilt.

⁸³⁶ Siehe URL: <http://hypotheses.org/about/academic-committees>. Weitere deutschsprachige Fachportale: *Clio-online* (URL: <http://www.clio-online.de/>) oder *historicum* (URL: <https://www.historicum.net/>).

Lösungsansatz Erscheinungsort

Da für eine digitale Quelle der Erscheinungsort irrelevant und der Speicherort massgebend für die Qualität des Objekts ist, kann auf die Angabe des Erscheinungsortes verzichtet werden. Aber auch die Angabe des Speicherortes ist den vorhergehenden Ausführungen folgend nicht sinnvoll. Ich schlage deshalb vor, auf beide Angaben zu verzichten und ausschliesslich den eindeutigen und persistenten Identifikator des Objekts, ergänzt um die URL eines (gesetzlich beauftragten) TDR anzugeben:

Fiktives Beispiel PI: CH-101111-1.00120.A0Y4.0218

URL (fiktiv): <https://www.tdr.bar.ch/CH-101111-1.00120.A0Y4.0218>

Objekte, deren Adresse nicht vertrauenswürdig und/oder persistent ist, erhalten durch die Sicherung in ein TDR wissenschaftliche Qualitätsmerkmale wie PI, Datum oder Nachvollziehbarkeit und können anschliessend aufgrund der Reputation des TDR für wissenschaftliche Zwecke verwendet werden.

5.4.8. Autorschaft

Als Autor eines digitalen Objekts können nicht nur natürliche Personen identifiziert werden, sondern auch Programme, die automatisch und autonom digitale Objekte erstellen.⁸³⁷ Das Feststellen der Autorschaft ist mit Problemen wie künstlicher Intelligenz, falscher Identitäten, Anonymität oder der Verwendung eines Pseudonyms verbunden. Jeder Autor verfolgt jeweils eigene Intentionen, die auch in der Wahl seiner Identitätsform zum Ausdruck kommen. Die Verwendung von fortgeschrittenen oder qualifizierten digitalen Signaturen würden verlässliche Informationen über den oder die Autor(en) liefern und sollten dringend eingesetzt werden.

Dem Themenkomplex *Autorschaft* werden einerseits die Feststellung der Autorschaft aus der äusseren Quellenkritik und andererseits das Nachvollziehen der *Intention* oder *Tendenz* des Autors aus der inneren Quellenkritik zugerechnet. Kritik am Konzept der Autorschaft wird hier ausgeklammert.⁸³⁸ Die Intention und das Motiv des Autors wie Unterhaltung, Überzeugung oder politische Einstellung lassen sich eventuell bereits durch die traditionelle Analyse des

⁸³⁷ Ich führe hier keine Diskussion, ob eine natürliche Person angesichts der informationstechnischen Verarbeitung überhaupt als Autor genannt werden kann. Ich nehme an, dass Hard- und Software als Hilfsmittel für die Umsetzung eines intellektuellen Prozesses eines Menschen eingesetzt werden und deshalb dieser als Autor bezeichnet wird.

⁸³⁸ Bspw. befürwortet Daniel O'Donnell die Einführung einer *Creditline* anstelle der Autorschaft, um alle Beteiligten eines Forschungsprojekts vom Programmierer zum Verfasser eines Werks aufzuführen. Siehe dazu: O'Donnell, Daniel Paul: The credit line, 16.08.2014, URL: <http://people.uleth.ca/~daniel.odonnell/Blog/firs-thing-we-do-let-s-kill-all-the-authors>, Stand: 19.11.2014.

dargestellten Objekts genauer bestimmen. Es kann unter anderem möglich sein, den Bildungs- und Informationsstand des Autors oder seine Nähe zum tatsächlichen Geschehen festzustellen. Ein in eine Webseite eingebettetes, dynamisches Objekt, das sich nach der vom Nutzer eingespielten Musik verändert, dient eventuell nur der Unterhaltung. Aber es lässt sich eventuell aus der Komplexität des Objekts schliessen, ob der Autor einen höheren Bildungsstand hat, weil er sich die dafür nötigen Programmierkenntnisse aneignete.⁸³⁹

Bei digitalen Objekten sollte allerdings zuerst festgestellt werden, ob es sich beim Autor um eine *natürliche Person* handelt⁸⁴⁰, oder ob das Objekt unabhängig algorithmisch generiert worden ist. Diese Unterscheidung ist mittlerweile nicht mehr so einfach und eindeutig. Es werden bereits heute erfolgreich Systeme für die Erstellung algorithmischer Textobjekte (sog. software-generated content) aus Daten unterschiedlicher Herkunft eingesetzt, so genannte *Softbots* oder Roboterjournalisten.⁸⁴¹ Weil die meisten Menschen mit reinen Datenansammlungen nicht umgehen können, werden narrative Berichte algorithmisch generiert, die den Sinn der Daten erklären und zu neuen Erkenntnissen führen, denn Daten sind weder Informationen noch Erkenntnis, wie der Informatiker Kris Hammond pointiert meint.⁸⁴² Zur Unterstützung des Menschen wurden und werden deshalb Computersysteme entwickelt, die Daten analysieren und anschliessend Ereignisse ableiten, diese gruppieren und darin sinnvolle Muster erkennen, deren Wichtigkeit abwägen und schliesslich die Resultate in einem Sprachsystem formulieren. Was zurzeit bereits bei Börsen- oder Sportinformationen sehr gut funktioniert, wird weiterentwickelt und soll bis 2025 für mehr als 90% der Publikationen von aktuellen Nachrichten in allen Sparten eingesetzt werden, behauptet der Technikjournalist Steven Levy.⁸⁴³ Im E-Mail Programm Google Mail generiert die Funktion *Smart Reply* eine automatische Antwort auf eine erhaltene E-Mail. Damit wird der Softbot zum Chatbot.⁸⁴⁴ Zwar haben die Teilnehmer einer Studie des Medien- und Kommunikationswissenschaftlers Christer Clerwall die Texte von Softbots als langweilig zu lesen bezeichnet, aber sie werden als objektiver, informativer und vertrauenswürdiger

⁸³⁹ Dass dies nicht zwingend so sein muss, zeigen Programme, mit denen ein Nutzer komplexe digitale Objekte entwickeln kann, ohne dass Programmierkenntnisse erforderlich wären. Ein Beispiel sind Bildbearbeitungsprogramme wie The GNU Image Manipulation Programm Gimp (URL: <http://www.gimp.org/>) oder Adobe Photoshop (URL: <http://www.photoshop.com/>).

⁸⁴⁰ Eine oder mehrere natürliche Personen haben durch Eingaben an einem elektronischen Informationssystem das digitale Objekt erstellt.

⁸⁴¹ Bspw.: Narrative Science (URL: <https://www.narrativescience.com/>); *Wordsmith* von Automated Insights (URL: <http://automatedinsights.com/>).

⁸⁴² Nach: Hammond, Kris: The End of Big Data: AI and the Rise of the Narrative, Blog, NarrativeScience, 06.03.2015, URL: <http://www.narrativescience.com/blog/big-data-ai-rise-narrative/>, Stand: 23.05.2015.

⁸⁴³ Levy, Steven: Can an Algorithm Write a Better News Story Than a Human Reporter?, Blog, WIRED, 24.04.2012, URL: <http://www.wired.com/2012/04/can-an-algorithm-write-a-better-news-story-than-a-human-reporter/>, Stand: 23.05.2015.

⁸⁴⁴ Siehe: Corrado, Greg: Computer, respond to this email., Research Blog, 03.11.2015, URL: <http://googleresearch.blogspot.com/2015/11/computer-respond-to-this-email.html>, Stand: 03.11.2015.

eingeschätzt als der Text eines Journalisten.⁸⁴⁵ Diese Entwicklung wird kaum bei Blog- oder Newsbeiträgen und Kurzberichten stehen bleiben, sondern es werden zukünftig durch künstliche Intelligenz wohl auch Literatur oder Forschungsberichte erstellt werden können.⁸⁴⁶ Bereits möglich ist die automatisierte Erstellung von Wikipediaartikel. Das Team um Peter Liu hat (experimentell) modelliert, wie ein Programm aus mehreren oder vielen Artikeln eine sinnvolle Zusammenfassung erstellen kann.⁸⁴⁷ Von Steven Levys Prophezeiung sind wir nicht mehr weit entfernt.

Eine Intention oder ein Motiv bei Softbots festzustellen, ist kaum möglich. Allerdings darf auch nicht von einer grundsätzlichen Objektivität oder Meinungslosigkeit solcher Programme ausgegangen werden. Da der Algorithmus für die Auswahl und Gewichtung der Daten verantwortlich ist, wäre dieser auf tendenziöse Auswahlkriterien hin zu untersuchen. Der Zugriff auf den Algorithmus ist in den meisten Fällen jedoch nicht möglich, weil beispielsweise nur das digitale Objekt zur Verfügung steht oder er als Geschäftsgeheimnis eingestuft wird, oder eine Beurteilung nur mit spezialisiertem Wissen erfolgen kann. Deshalb sollten aus Softbots generierte Objekte nur mit grosser Vorsicht verwendet werden. Allerdings wird ein Softbot nicht immer als solcher erkannt, wie Clerwall ebenfalls nachweisen konnte. Deshalb wäre die eindeutige Kennzeichnung eines durch einen Roboterjournalisten erstellten Artikels notwendig. Als Autor dieser *Quartärquelle* sollten der Name und die Version des Programms sowie die für die Programmierung verantwortliche Person oder das Programmiererteam als Autoren aufgeführt werden.

Zu entscheiden ist auch, wer wie im letzten Beispiel der Autorschaft zuzurechnen ist. Beispielsweise werden in den aktuellen Webpublishing Tools die Darstellung und der Inhalt voneinander getrennt. Bei meinem eigenen Blog hsc.hypotheses.org bin ich der alleinige Autor der Texte, aber für die Darstellung wird das Produkt Wordpress verwendet und dafür ist ein anderes Autorenteam zuständig. Verweist jemand auf diesen Blog, ist der Autor eindeutig, weil hypotheses.org mit Wordpress nur ein Hilfsmittel zur Publikation meiner Inhalte zur Verfügung stellt. Wenn aber Inhalt und Form so ineinander verschmelzen, dass es als

⁸⁴⁵ Siehe dazu die Studie von Christer Clerwall mit zwar relativ kleiner Stichprobe, die aber dennoch das Entwicklungspotenzial solcher Software aufzeigt: Clerwall, Christer: Enter the Robot Journalist, in: *Journalism Practice* 8 (5), 03.09.2014, S. 519-531. DOI: 10.1080/17512786.2014.883116. Eine neue Studie stützt die Erkenntnisse von Clerwall, siehe: Graefe, Andreas; Haim, Mario; Haarmann, Bastian u. a.: Readers' perception of computer-generated news: Credibility, expertise, and readability, in: *Journalism*, 17.04.2016. DOI: 10.1177/1464884916641269.

⁸⁴⁶ 2014 hat das KI-Programm 'Eugene' den Turing-Test bestanden, in dem ein Mensch versucht, per schriftlicher Kommunikation herauszufinden, ob das Gegenüber ein Mensch ist oder nicht. Mit dem Bestehen des Tests hat das Programm bewiesen, dass es die Intelligenz eines 13-Jährigen simulieren kann.

⁸⁴⁷ Siehe: Liu, Peter J.; Saleh, Mohammad; Pot, Etienne u. a.: Generating Wikipedia by Summarizing Long Sequences, arXiv:1801.10198 [cs.CL], 30.01.2018. URL: <http://arxiv.org/abs/1801.10198>, Stand: 19.02.2018.

ein integrales Objekt angesehen werden kann oder muss, wären auch die Autoren des darstellenden Systems zu nennen.

Grundsätzlich sind auch bei der Feststellung der Autorschaft eines jeden digitalen Objekts die oben behandelten Themen Datierung und Entstehungsort sowie Informationen zur Überlieferungsgeschichte, über die Publikationsgesellschaft oder digitale Signaturen zu berücksichtigen. Die Angaben zum Autor sind zu überprüfen und allenfalls zu ergänzen, weil sie auch fehlen, falsch angegeben oder manipuliert worden sein können. Ein falscher Autorname kann beispielsweise durch *Spoofing*⁸⁴⁸ vorgetäuscht werden, Kenntnisse von entsprechenden Programmen und Tools vorausgesetzt. Digitale Identität⁸⁴⁹ ist das, was der Nutzer als Identitätsgeber behauptet, was sie sei. Identität ist im Cyberspace volatil, denn der Nutzer selbst bestimmt, wen er darstellen will. Jede Art von Identität kann angenommen, genutzt oder wieder abgelegt werden – die Medientheoretikerin Stacey Koosel spricht vom "Private Public Paradox".⁸⁵⁰ Jeder Nutzer legt das Verhältnis zwischen Privatheit und Öffentlichkeit selbst fest. Oft werden auch mehrere Identitäten verwendet, die nicht zwingend Eigenschaften des Nutzers aufweisen: "they can be a subtype of a public persona, an extension of our 'true' selves, or they can be completely fabricated and fantastical, to function as a mask to hide the identity of an Internet user from the rest of the world."⁸⁵¹ Akademische Titel wie Doktor oder Professor, Institutionsnamen wie Universität oder Fachhochschule und Zertifikatslogos sind keine verlässlichen Qualitätsmerkmale, weil auch diese frei erfunden oder gestaltet werden können.⁸⁵² Es ist auch möglich, die digitale Identität einer Institution oder einer Person vorzutäuschen, wie es besonders bei Phishing-Angriffen⁸⁵³ gemacht wird. Dabei werden E-Mails, Dokumente und vor allem Webseiten mit bekannten Namen und Logos äusserst originalgetreu nachgebaut, um den Anschein der Authentizität zu erwecken. Der Nutzer, der die Fälschung nicht erkennt, gibt dem Identitätsfälscher eventuell sensible Daten wie Kreditkartennummern und Codes bekannt. Phishing ist nicht nur für Bankkonteninformationen interessant, sondern kann auch im wissenschaftlichen Umfeld

⁸⁴⁸ Spoofing: Vorspiegelung einer falschen Identität durch aktiven Eingriff in ein IT-System. Siehe: Poguntke: Basiswissen IT-Sicherheit, 2013, S. 16 f.

⁸⁴⁹ Verwandte Begriffe: digital identity, online identity, online personality, digiSelf, virtual identity, avatar, online persona.

⁸⁵⁰ Siehe dazu: Koosel, Stacey M.: Exploring Digital Identity: Beyond the Private Public Paradox, in: Runnel, Pille; u.a. (Hg.): The digital turn – user's practices and cultural transformations, Frankfurt a.M. 2013, S. 161-173, hier S. 164.

⁸⁵¹ Ebd., S. 161.

⁸⁵² Titel u.ä. werden bspw. von Fritch und Cromwell unter Hinweis auf eine skeptische Haltung gegenüber Angaben im Internet (WWW) als vertrauensstiftende Merkmale vorgeschlagen. Siehe: Fritch; Cromwell: Evaluating Internet resources, 2001, S. 501.

⁸⁵³ Phishing: aus dem Engl. 'to fish'; Trickbetrug. Die Zielperson soll sicherheitsrelevante Informationen auf einer Webseite eingeben oder per Mail zusenden. Siehe: Andress, Jason (Hg.): The basics of information security: understanding the fundamentals of InfoSec in theory and practice, Amsterdam 2011, S. 123 f.

eingesetzt werden. Wissenschaftsspionage ist eine reale, leider zu wenig beachtete Bedrohung.

Als Massnahmen gegen falsche Identitäten bietet sich bei Objekten im Internet wiederum ein Blick auf die Reputation der publizierenden Institution und deren URL an. Sind diese bekannt und vertrauenswürdig, kann auch dem Inhalt vertraut werden. Bei einzelnen Objekten kann das Vertrauen in die Korrektheit der Informationen mit einer digitalen Signatur gestärkt werden. Fortgeschrittene und qualifizierte Signaturen beweisen zudem die Identität des Urhebers. Mit einem Zeitstempel kann zudem auch der Entstehungszeitraum überprüft werden.

Als Schutz vor Fälschungen respektive für den Identitätsnachweis eines Objekts werden unter anderem diverse *DRM-Verfahren* eingesetzt, insbesondere bei Multimediaobjekten. Das Objekt wird verschlüsselt und kann nur mit entsprechender Lizenz und mit den vorgeschriebenen Komponenten eines Informationssystems genutzt werden.⁸⁵⁴ Es kann so durch den Rechteinhaber identifiziert werden, was ihm aber auch ermöglicht, die Nutzung des Objekts in bestimmten Systemen zu verhindern. DRM-Systeme werden hauptsächlich als technische Massnahme zum Schutz geistigen Eigentums eingesetzt. Die Umgehung dieser Verfahren ist in einigen Ländern gar strafbar.⁸⁵⁵

Der Einsatz von DRM-Verfahren dient vordergründig der Identitätssicherung, lässt aber zusätzlich auf die Intention des Autors respektive des Rechteinhabers schliessen. Die Verwendung eines restriktiven Verfahrens, das die Nutzung stark einschränkt oder deutliche Signaturen wie Wasserzeichen zeigt, lässt auf die Intention der Sicherstellung der Integrität des Objekts schliessen. Bei audiovisuellen Objekten kann zusätzlich die Feststellung der Nutzung beabsichtigt sein⁸⁵⁶, um eine entsprechende finanzielle Nutzungsgebühr zu berechnen. Bei einem administrativen Dokument kann DRM zur Sicherstellung rechtlicher Beweisfähigkeit eingesetzt werden.

Bei der Untersuchung der Autorschaft ist eine skeptische Haltung angebracht, weil eine verlässliche, vorgängige Authentizitätsprüfung durch Verlage oder ähnliche Institutionen nicht erfolgt. Digitale Objekte und Informationen können unter anonymen, pseudonymen oder falschen Autorangaben von fast jeder Person publiziert und instantan wiederverwendet wer-

⁸⁵⁴ Die Nutzung DRM-geschützter Objekte stellt Gedächtnisinstitutionen vor grosse Probleme: (teure) Lizenzen müssen eingekauft werden, die Komponenten des vorgeschriebenen Informationssystems müssen vorhanden sein etc. Siehe: Gladney: *Preserving Digital Information*, 2007, S. 172 f.

⁸⁵⁵ Verhandlungen darüber, was mit dem Erwerb einer Lizenz erworben wird (Lesen, Kopieren, Weiterverkauf etc.) und was ein 'gerechte' Abgabe für Autoren geistiger Werke ist, werden zwischen Konsumenten, Gesetzgebern und Rechteinhabern resp. -verwertern verbissen geführt, eine Einigung ist in weiter Ferne. Siehe zu den verschiedenen Positionen bspw.: Arbeitsgruppe zum Urheberrecht: Schlussbericht AGUR12, 2013.

⁸⁵⁶ Bspw. mit Pay-per-View. Darunter werden verschiedene Dienste verstanden, die den Zugriff auf ein digitales Objekt nur durch vorgängige Bezahlung (einmalig, Abo) ermöglichen.

den.⁸⁵⁷ Die Anonymität⁸⁵⁸ und – als ihre spezielle Ausprägung – die Verwendung eines Pseudonyms⁸⁵⁹ dienen dem Persönlichkeitsschutz eines Autors und können teilweise sehr schwer oder gar nicht aufgelöst werden.

Bei einer *anonymen Autorschaft* sind wie bei physischen Objekten die verschiedenen Indizien zur Überlieferungsgeschichte wie Inhalt, Darstellung oder Herkunft zusammenzutragen, mit denen trotz fehlender Autorangaben möglicherweise eine bestimmte Person als Autor identifiziert werden kann. Ist dies nicht möglich, so ist dies entsprechend zu vermerken. Der Wert des Objekts als potenzielle Quelle nimmt dadurch stark ab.

Die Verwendung eines *Pseudonyms* kann als Spezialfall der Anonymität betrachtet werden. In beiden Fällen bleibt die Identität des Autors verborgen, ein Pseudonym ermöglicht dem Autor sich hinter einem falschen, aber trotzdem ihm zugehörigen Namen oder Symbol zu verstecken.

Ein Pseudonym muss allerdings als solches erkannt werden. Wird der Autorname beispielsweise mit 'Harry Potter'⁸⁶⁰ angegeben, handelt es sich dabei sehr wahrscheinlich um ein Pseudonym. Hingegen ist die Angabe 'C. G. Jung' zwar aufgrund der Ähnlichkeit zum bekannten Psychologen verdächtig, kann aber durchaus korrekt sein (Christian Gerhard Jung). Auch die Verwendung eines bekannten Pseudonyms wie Woody Allen⁸⁶¹ durch eine andere Person ist möglich. Der offensichtlich korrekte Name 'Claudia Schindler' muss nicht zutreffen, denn der Name könnte von einer Person als Pseudonym verwendet werden. Eventuell wird dieses Pseudonym dazu verwendet, um einer andern Person in einer elektronischen Konversation vertrauliche Informationen zu erhalten, weil sie diese Claudia Schindler vermeintlich kennt.⁸⁶² Deshalb sollte auch Kontakten in Digitalmedien nicht von vornherein vertraut werden. Twitter beispielsweise beziffert die vermutlich mittels Crowdfunding automatisiert erstellten Accounts im August 2014 auf rund 8,5% von total über einer Milliarde registrierten Nutzern.⁸⁶³ Dasselbe Problem haben auch andere Social Media – Plattformen.

⁸⁵⁷ Vgl.: Fritch; Cromwell: Evaluating Internet resources, 2001, S. 501.

⁸⁵⁸ Anonymität: das Nichtbekanntsein, Nichtgenanntsein; Namenlosigkeit. Nach: Wermke u. a. (Hg.): Duden - das Fremdwörterbuch, 2001.

⁸⁵⁹ Pseudonym: angenommener, nicht richtiger Name; Deckname. Nach: Ebd.

⁸⁶⁰ Fiktive Figur aus den Geschichten von Joanne K. Rowling.

⁸⁶¹ Woody Allen alias Allan Stewart Konigsberg.

⁸⁶² Zum sogenannten Social Engineering siehe Literatur in FN 430.

⁸⁶³ Siehe: Twitter Inc.: Quarterly Report, Form 10-Q, United States Securities and Exchange Commission, Washington (D.C.) 11.08.2014, S. 4. URL: http://www.sec.gov/Archives/edgar/data/1418091/000156459014003474/twtr-10q_20140630.htm, Stand: 30.06.2015.

Crowdfunding: Natürliche Personen erstellen auf Auftrag einer Organisation Accounts und 'verlinken' diese mit dem/den vereinbarten digitalen Objekt(en). Siehe: Wang u. a.: Serf and Turf, 2011 ; De Cristofaro, Emiliano; Friedman, Arik; Jourjon, Guillaume u. a.: Paying for Likes? Understanding Facebook Like Fraud Using Honeypots, in: ACM Internet Measurement Conference 2014, arXiv:1409.2097 [cs.SI], Vancouver 2014. URL: <http://arxiv.org/abs/1409.2097>, Stand: 30.06.2015.

Die Verwendung eines Pseudonyms oder Avatars ist nicht nur in Foren, virtuellen Welten oder im Darknet üblich. Gerne werden Pseudonyme auch für Kommentare zu Artikeln auf als seriös geltenden Webangeboten wie der NZZ online verwendet. Die Autorangaben können während der Erstellung eines neuen Nutzerkontos frei gewählt werden, eine Identitätsprüfung ist bei solchen Angeboten kaum möglich. Sicherheitsfunktionen wie CAPTCHA⁸⁶⁴ oder Rechenaufgaben und die Bestätigung durch einen in einer funktionierenden E-Mail-Adresse zugestellten Hyperlink sind keine Identitätskontrolle und werden nur zur Spam-Abwehr eingesetzt. Nutzer können beliebig viele virtuelle Identitäten mit gleichen oder unterschiedlichen Pseudonymen erstellen, die untereinander verknüpft sein können, aber nicht müssen. Jede virtuelle Identität kann sich von der anderen unterscheiden und kann Eigenschaften des Nutzers annehmen. Der Nutzer kann mit seinem Pseudonym sehr öffentlichkeitswirksam agieren, während er persönliche Informationen des realen Lebens sehr gut schützt.⁸⁶⁵ Die meisten Nutzer schützen sich mit einem Pseudonym vor Überwachung, Identifikation und Werbung.

Wird nun festgestellt, dass der Autor ein Pseudonym verwendet hat oder dass er anonym bleiben will, ist dies ein Hinweis auf seine Intention. Der Autor schützt sich beispielsweise durch Anonymität nicht nur vor der Einflussnahme und im Extremfall vor Repression Dritter auf sein Werk, sondern will auch verhindern, dass seine Aktivitäten öffentlich bekannt werden. Beispielsweise werden Änderungen an einem Artikel in Wikipedia mit einem Zeitstempel und unter Angabe des Autors veröffentlicht, wodurch nachvollziehbar wird, wer wann was geschrieben hat und wie oft und wie lange jemand sich mit Wikipedia-Einträgen beschäftigt. Solche Informationen können einem Autor beruflichen und privaten Schaden zufügen, wenn ihre Aktivitäten bekannt werden. Beispielsweise steht für einige Universitätsprofessoren, die auf Wikipedia aktiv sind, die Wissensverbreitung im Vordergrund. Sie wollen aber anonym bleiben, um sich von anderen nicht beeinflussen zu lassen.⁸⁶⁶

Pseudonyme können auch zum Ausdruck bringen, dass sich ein Autor als Sprecher einer bestimmten Gruppe versteht.⁸⁶⁷ Veröffentlicht 'Velofahrerfürst' einen Online-Kommentar zu einem Zeitungsartikel über Doping im Radsport, so gibt er nicht nur seine persönliche Meinung kund, sondern will durch die Wahl des Pseudonyms eventuell auch als Sprachführer

⁸⁶⁴ CAPTCHA: Completely Automated Public Turing test to tell Computers and Humans Apart. Turing Test, dessen Bestehen menschliche Intelligenz voraussetzt. Nach: Fischer; Hofer: Lexikon der Informatik, 2011.

⁸⁶⁵ Für Stacy Koosel ein Paradoxon. Siehe: Koosel: Exploring Digital Identity: Beyond the Private Public Paradox, 2013, S. 164.

⁸⁶⁶ Siehe dazu Wiegand, Dorothee: Wikipedia: Zu gut, um nicht zitiert zu werden?, Newsticker, heise online, 04.10.2014, URL: <http://www.heise.de/newsticker/meldung/Wikipedia-Zu-gut-um-nicht-zitiert-zu-werden-2411605.html>, Stand: 07.10.2014.

⁸⁶⁷ Weiterführend: Evens, Aden: Web 2.0 and the Ontology of the Digital, in: DHQ Digital Humanities 6 (Futures of Digital Studies: 2) (2), 2012. URL: <http://www.digitalhumanities.org/dhq/vol/6/2/000120/000120.html>, Stand: 17.08.2012.

einer Velofahrgemeinschaft wahrgenommen werden. Die tatsächlichen Machtverhältnisse zwischen dem Autor und der von ihm vermeintlich vertretenen Gruppe können jedoch ohne zusätzliche Informationen kaum nachvollzogen werden.

Als spezielle Form des Pseudonyms kann das 'Face Reenactment' angesehen werden. Hierbei wird in einem Video die Mimik einer Person durch die einer anderen verfremdet. Der Informatiker Justus Thies hat aufgezeigt, wie es in Echtzeit möglich ist, dass sich in einem Film Arnold Schwarzeneggers Gesicht sehr realistisch nach der Vorgabe des Gesichtsausdrucks einer Frau bewegt.⁸⁶⁸ Die Frau als Autorin nutzt eine visuelle Darstellung von Arnold Schwarzenegger als Pseudonym, mit dem sie ihre Aussagen kommuniziert. Mit der Weiterentwicklung und Perfektionierung dieser Techniken⁸⁶⁹ wird die quellenkritische Überprüfung von Filmmaterial massiv erschwert, falls keine Authentizitätsstiftenden Merkmale wie eine digitale Signatur für solche Objekte eingesetzt werden.

Der Autor hat in der virtuellen Welt nicht nur sehr vielfältige Möglichkeiten seine Identität aktiv zu gestalten, sondern er kann zusätzliche Funktionen übernehmen. Eine solche erhält der Autor beispielsweise beim Schreiben von Hypertexten. Indem er seine eigenen, bereits publizierten Texte immer wieder verändern kann, wird er zum eigenen *Editor*. Der publizierte Hypertext erhält damit eine "quasi-auktoriale Rahmungsfunktion"⁸⁷⁰, die bei Bedarf instantan umgestaltet und umgedeutet werden kann. Der Autor erhält damit auch einen enormen Machtzuwachs, denn das publizierte Objekt kann nachträglich ergänzt, angepasst oder auch gelöscht werden. Dies ist bei gedruckten Texten nur durch eine Neuauflage möglich und mit hohem zeitlichem, finanziellem und organisatorischem Aufwand verbunden.

Besonders in digitalen, kollaborativen Arbeitsumgebungen wie Wikipedia oder Etherpad sind Änderungsprozesse vorgesehen und erwünscht. Werden diese protokolliert, kann der gesamte Erarbeitungsprozess eines Objekts nachvollzogen werden. Zudem wird sichtbar, welche Autoren wie stark mitgewirkt haben. Daraus lassen sich teilweise auch die Intentionen der einzelnen Autoren ableiten. Beispielsweise lässt sich die Intention eines Autors, der eine formale Anpassung eines Textes vornimmt, von derjenigen eines Administrators unterscheiden, der eine Textpassage löscht oder löschen muss. Erster ist anscheinend mit dem Inhalt einverstanden und nimmt nur eine marginale Änderung vor. Der Administrator löscht die Textpassage aufgrund bestimmter Richtlinien oder aus eigenem Antrieb, weil sie zum

⁸⁶⁸ Thies, Justus; Zollhöfer, Michael; Stamminger, Marc u. a.: Face2Face: Real-time Face Capture and Reenactment of RGB Videos. URL: <http://www.graphics.stanford.edu/~niessner/thies2016face.html>, Stand: 26.06.2016.

⁸⁶⁹ Ende 2017 wurde die Weiterentwicklung des 'Face Reenactments' als 'Deepfakes' bekannt. Gesichter von Pornodarstellerinnen wurden mittels intelligenter Software mit denen bekannter Schauspielerinnen wie Natalie Portmann oder Gal Gadot ersetzt. Siehe bspw.: Kühl, Eike: Künstliche Intelligenz: Auf Fake News folgt Fake Porn, Zeit Online, 26.01.2018, URL: <http://www.zeit.de/digital/internet/2018-01/kuenstliche-intelligenz-deepfakes-porno-face-swap/komplettansicht?print>, Stand: 07.02.2018.

⁸⁷⁰ Wirth, Uwe: Hypertext, in: Grundbegriffe der Medientheorie, Paderborn 2005, S. 90.

Beispiel seiner persönlichen Einstellung widerspricht und weil er die Macht dazu hat.⁸⁷¹ Damit entsteht eine neue Form von kollaborativer Autorschaft, weil sowohl Autoren und Editoren, die schreiben und ergänzen, als auch solche, die mutieren und löschen beteiligt sind.

Der Autor gibt mit den gesetzten Hyperlinks auf andere Objekte weitere Hinweise auf seine Intention. Aus den mit dem Objekt verknüpften oder nicht verknüpften Objekten oder Elementen werden die vom Autor verfolgten Interessen aufgezeigt. Allerdings ist bei der Interpretation solcher Informationen Vorsicht geboten, denn insbesondere bei kollaborativen Plattformen wie Wikipedia ist ein Ersteintrag als erster Entwurf oder Anstoss für die weitere Bearbeitung zu betrachten, der von der *Community* ergänzt und mit anderen Objekten verknüpft werden *soll*. Die Setzung von Hyperlinks untersteht eventuell einer klaren Regelung oder erfolgt sogar automatisiert.

Unabhängig davon, was der Autor mit seiner Linksetzung intendiert hat, wird der Nutzer durch das (Nicht-)Verwenden dieser Hyperlinks in nicht vorbestimmter Reihenfolge selber zum Autor, indem er sich sein eigenes, nicht-lineares Narrativ erarbeitet.⁸⁷² Diese Art von Autorschaft wird jedoch von der Quellenkritik nicht erfasst und deshalb hier nicht weiter besprochen.

Anonyme und pseudonyme Identitätsformen sind bereits in der analogen Welt bekannt. Verlässliche Angaben über einen Autor werden durch Kontrollinstanzen wie Behörden oder Verlage zur Verfügung gestellt. Die Reputation der publizierenden Institution ist ausschlaggebend für das Vertrauen in die Korrektheit dieser Angaben. Bei digitalen Publikationen fehlen zwar oft solche Kontrollinstanzen, aber ähnlich wie in analogen Medien vertraut der Nutzer der Reputation des Hosters. Gänzlich neu ist das selbständig algorithmische Erstellen von digitalen Objekten durch Systeme der künstlichen Intelligenz, die die Rolle des Autors übernehmen können. Der Autor erhält auch neue Funktionen und einen Machtzuwachs, denn er kann sein Objekt selbständig publizieren und nach der Publikation sein Objekt ohne hohen zeitlichen, finanziellen, und organisatorischen Aufwand editieren und erneut publizieren. Mit in das Objekt integrierten Schutzmechanismen kann er sehr spezifische Nutzungsformen bestimmen, was bei physischen Publikationen nicht möglich ist.

⁸⁷¹ In diversen Beiträgen wird die Macht der Administratoren diskutiert. Sie zur Deutschen Wikipedia bspw.: Kleinz, Torsten; Windeck, Christof: WikiCon: Konfliktfeld Wikipedia, Newsticker, heise online, 05.10.2014, URL: <http://www.heise.de/newsticker/meldung/WikiCon-Konfliktfeld-Wikipedia-2411752.html>, Stand: 27.05.2015.. Untersuchung zu konkretem Beispiel 'Wikipedia Art' siehe: Tkacz, Nathaniel: Wikipedia and the politics of openness, Chicago 2015, S. 50-61.

⁸⁷² Siehe bspw.: Haber: digital past (unv. Habil.), 2009, S. 202 f.

5.4.9. Adressat

Digitale Publikationsformen erreichen einen viel grösseren und damit heterogeneren potenziellen Nutzerkreis als physische Medien. Durch Verlinkung wird auch relevant, auf welchen Sekundärplattformen das digitale Objekt eingebettet wurde. Jedoch können nicht alle potenziellen Adressaten erreicht werden, sei es, weil diese technisch, willentlich oder intellektuell ausgeschlossen sind, oder weil der Autor eigenmächtig einen Teil der Adressaten ausschliesst. Erstautor und Adressat können durch digitale Publikationsmedien nicht nur in einen Dialog treten, sondern auch gemeinsam an einem Objekt arbeiten. Bei solchen unmittelbaren Dialogen wächst die argumentatorische Macht des Adressaten. Da der Adressat das publizierte Objekt manipulieren kann, geht er mit dem Autor eine manuskriptive Kollaboration ein.

Wie in der traditionellen Quellenkritik ist der soziale, kulturelle und politische Hintergrund des Adressaten sowie das Verhältnis zwischen Autor und Adressat zu untersuchen. Bei digitalen Objekten bedürfen zwei Punkte besonderer Aufmerksamkeit: Erstens wird mit digitalen Objekten und Inhalten tendenziell ein viel grösserer und heterogenerer Adressatenkreis erreicht als mit gedruckten Publikationen, und zweitens können mannigfaltige Interaktionen zwischen (Erst-)Autor und Adressat aufgebaut werden, die vor allem bezüglich Schnelligkeit und direkter Verknüpfung bisher nicht zur Verfügung standen.

Wird ein digitales Objekt einem bestimmten und kontrollierten Nutzerkreis zur Verfügung gestellt, kann der Hintergrund der oder des Adressaten spezifisch bestimmt werden, sofern die entsprechenden Informationen zur Verfügung stehen. Je grösser der Nutzerkreis jedoch wird, desto allgemeiner werden die Aussagen oder es lassen sich überhaupt keine Aussagen mehr machen. Dies ist nicht nur bei Projekten wie Wikipedia der Fall, bei dem "jeder Internetnutzer"⁸⁷³ angesprochen wird, sondern auch bei auf Internetportalen veröffentlichten Blog-Posts, Zeitungsartikeln, Newsbeiträgen oder Videos, die ebenfalls von jedem Internetnutzer eingesehen werden können – weltweit! Durch algorithmisch generierte Mashups, beispielsweise von Newsaggregatoren wie Newscron, werden digitale Objekte und Inhalte von unterschiedlichen Webseiten teilweise nutzerspezifisch neu zusammengestellt und an ganz unterschiedliche Nutzer und Nutzergruppen adressiert. Es ist nicht nur aus der Sicht des Autors zu beurteilen, für wen das Objekt zur Verfügung gestellt wird, sondern auch aus der Sicht der potenziellen Nutzer. Diese fragen, welche individuellen Anpassungen durch und für sie in der Publikationsform möglich sind. Dabei kann auch relevant sein, welche primäre Plattform vom

⁸⁷³ Wikipedia über Wikipedia: «Vanellus»: Wikipedia, in: Wikipedia, 04.05.2015. URL: <https://de.wikipedia.org/w/index.php?title=Wikipedia&oldid=141752818>, Stand: 08.05.2015.

Autor für die Erstpublikation gewählt wurde und welche sekundäre(n) Plattform(en) das Objekt weiterverbreiten.

Allerdings ist bei aller Euphorie über die Masse der potenziell zu erreichenden Adressaten auch zu beachten, dass digitale Daten Personen, denen moderne Kommunikationstechnologien nicht zur Verfügung stehen, ausschliessen. Ältere, ärmere und zuweilen reichere Personen sowie bildungsfernere Bevölkerungsschichten und in verschiedener Hinsicht behinderte Personen können oder wollen nicht teilnehmen. Diese Kluft wird als *Digital Divide* bezeichnet, der besonders dann problematisch wird, wenn (lebens-)wichtige Informationen nur noch digital zur Verfügung stehen.⁸⁷⁴

Aber auch innerhalb der Gruppe der Informationstechniknutzenden besteht ein Digital Divide. Personen, denen im Umgang mit Informationstechnik die nötigen Kompetenzen fehlen, finden sich in der heutigen, technologielastrigen Umwelt weniger zurecht.⁸⁷⁵ Bemühungen für die Etablierung von IT-basierten Lernformen wie Massive Open Online Course (MOOC) können deshalb nicht nur an fehlender Infrastruktur scheitern, sondern auch an den Personen, die die Technik nicht nutzen können oder wollen.

Deshalb ist auch zu überprüfen, ob der Adressat mit der gewählten Publikationsform überhaupt erreicht wird. Beispielsweise wird ein Fachartikel, der auf einem Fachportal mit Paywall publiziert wird, weniger Personen erreichen als auf einer frei zugänglichen Fachplattform. Ähnlich steht es mit einem Dokumentarfilm, der DRM-geschützt zur Miete angeboten oder DRM-frei auf dem Videoportal youtube⁸⁷⁶ publiziert wird. Beide Beispiele richten sich an dasselbe Publikum, aber die gewählte Publikations(platt)form ist entscheidend, wie viele Nutzer erreicht werden.

Aus der gewählten Publikationsform und Publikationsart lassen sich schon Hinweise auf das *Verhältnis zwischen Autor und Adressat* entnehmen. Je mehr der Adressat an der Erstellung, Verwendung oder Manipulation des Objekts partizipieren kann, desto mehr wird er selbst zum Editor oder Autor und kommuniziert oder kollaboriert mit dem Erstautor. Der Adressat kann beispielsweise durch Kommentare oder Anmerkungen das vom Autor publizierte digitale Objekt ergänzen. Der Kommentar wird diesem dabei direkt angefügt und beeinflusst dessen Aussage unmittelbar. Autor und Adressat treten in einen direkt sichtbar verknüpften Dialog. Dies ist bei gedruckten Zeitungsartikeln nicht möglich, denn Leserbriefe werden

⁸⁷⁴ Ende 2015 nutzten laut Schätzungen der ITU rund 3,2 Mia. Menschen das Internet, rund 4,2 Mia. (noch) nicht. Siehe: Broadband Commission for Digital Development: The State of Broadband 2015: Broadband as a Foundation for Sustainable Development, ITU & UNESCO, 21.09.2015, S. 19. URL: <http://www.broadbandcommission.org/publications/Pages/SOB-2015.aspx>, Stand: 06.05.2018.

⁸⁷⁵ Siehe dazu bspw.: Centre for Educational Research and Innovation of the OECD: Are the New Millennium Learners Making the Grade? Technology Use And Educational Performance in PISA, Paris 2010. URL: <http://www.oecd-ilibrary.org/content/book/9789264076044-en>, Stand: 08.05.2015.

⁸⁷⁶ URL: <https://www.youtube.com/>. Laut eigenen Angaben hat youtube mehr als 1 Mia. Nutzer (URL: <https://www.youtube.com/yt/press/de/statistics.html>).

durch die Redaktion explizit ausgewählt respektive kuratiert und können erst später veröffentlicht werden, wodurch eine klare Trennung von Artikel und seinen Ergänzungen geschaffen wird. Bei einem Online-Artikel stehen alle Kommentare unabhängig vom Zeitpunkt der Eingabe sofort zur Verfügung und sind zusammen mit dem Artikel sichtbar.⁸⁷⁷ Dies trifft auf alle Publikationsformen zu, die eine Kommentarfunktion zur Verfügung stellen. Damit ändert sich auch das Verhältnis zwischen Autor und Adressat, denn der Autor kann seinen Beitrag nicht mehr von den Rückmeldungen der Adressaten separieren. Zwar hat auch der Autor weiterhin die Möglichkeit, auf Beiträge mit einer Replik zu antworten, dennoch bleibt die Argumentation des Adressaten bestehen. Diese Zunahme der Einflussmöglichkeiten des Adressaten bezeichne ich als *argumentatorischen Machtzuwachs*.

In anderen Publikationsformen kann ein Autor Beiträge verfassen, die explizit zur Bearbeitung durch den Adressaten konzipiert sind, wie beispielsweise in Wikis.⁸⁷⁸ Jeder Interessierte kann das ursprünglich publizierte digitale Objekt ändern, womit der Erstautor nicht nur zum Editor, sondern auch zum Adressaten wird. Durch die von anderen durchgeführten Manipulationen kann er zu weiteren Ergänzungen, Änderungen oder Löschungen angeregt werden. Der argumentatorische Machtzuwachs des Adressaten ist bei solchen Formen noch grösser als beim erstgenannten Beispiel. Autor und Adressat treten hier nicht nur in einen Dialog, sondern arbeiten gemeinsam an einem Manuskript. Sie treten in die von mir so genannte *manuskriptive Kollaboration*.

Adressaten haben mit den zur Verfügung stehenden Plattformen auch die Möglichkeit ihre Meinung zu einem abgeschlossenen Werk zu publizieren und treten so mit dem Autor und anderen Adressaten oder Interessenten in Kontakt. Das Online-Rezensieren hat mittlerweile eine hohe Akzeptanz erreicht⁸⁷⁹ und ermöglicht nicht nur eine im Gegensatz zu Papiermedien sehr schnelle und unkomplizierte Verbreitung der Rezension, sondern durch Verlinkung und Kommentare auch eine weiterführende, transparente (Fach-)Diskussion.⁸⁸⁰

Des Weiteren ist zu beachten, dass sich Hypertexte durch die Interaktion mit dem Adressaten verändern. Durch den *hypertextuellen Leseprozess* werden Objekte – insbesondere Texte – zu anderen Objekten.⁸⁸¹ Der Leseprozess ist im Einzelfall kaum nachzuvollziehen,

⁸⁷⁷ Vorausgesetzt wird, die Kommentarfunktion wurde nicht deaktiviert und der Autor oder Administrator hat keine Kommentare gelöscht oder aus gesetzlichen oder ethisch-moralischen Gründen löschen müssen.

⁸⁷⁸ Wiki: haw. für schnell; Sammlung von Webseiten, die durch ein Content Management System verwaltet werden, das die Erstellung, Ergänzung und Manipulation von Einträgen erlaubt sowie eine Versionierung zur Verfügung stellt. Nach: Prevezanos: Computer Lexikon 2013, 2012.

⁸⁷⁹ Renommierete Plattformen sind bspw. H/Soz/Kult (<https://www.hsozkult.de/review/page>) oder Historische Rezensionen Online von Clio-Online (<http://beta.clio-online.de/hro/page>).

⁸⁸⁰ Vgl.: Pfanzelter, Eva: Rezensieren, kommentieren, bloggen: Wissenschaftliche Onlinekommunikation, in: Bernsen, Daniel; Kerber, Ulf (Hg.): Praxishandbuch Historisches Lernen und Medienbildung im digitalen Zeitalter, Opladen Berlin Toronto 2017, S. 159 f.

⁸⁸¹ Nach: Poster, Mark: Whats the Matter with the Internet?, Minneapolis 2001, S. 68.

denn je nach Leser ist mit einer anderen Inhaltsüberlieferung zu rechnen. Es wäre von Interesse zu wissen, welche Informationen der Leser aufgenommen oder ausgelassen hat. Dies lässt sich bei einem in sich geschlossenen Korpus, in dem Links nur auf Elemente des Korpus verweisen, möglicherweise feststellen, aber sobald Links auf externe Objekte verweisen, kann dies nicht mehr nachvollzogen werden.

Die Vorstellung einer Universalität und globalen Öffentlichkeit durch Digitalmedien ist allerdings noch utopisch⁸⁸², denn einerseits werden durch den Digital Divide viele von der Partizipation an der digitalen Welt ausgeschlossen, andererseits befinden sich in der Personengruppe mit entsprechendem Zugang viele Nutzer, die durch Sprache, (staatliche) Restriktionen und Interessen voneinander getrennt sind. Theoretisch könnten zwar alle miteinander kommunizieren, wenn alle dieselbe Sprache wie beispielsweise Englisch und dasselbe Tool wie beispielsweise Facebook verwenden würden. Das Resultat eines Ständig-mit-allem-Kommunizierens wäre aber reinste Kakophonie, wie es der Soziologe Hans Geser ausdrückt.⁸⁸³ Die bestehende Zersplitterung respektive Balkanisierung⁸⁸⁴ ist folglich kein Übel, sondern Notwendigkeit.

5.4.10. Darstellung

Ein digitales Objekt kann nur dann dargestellt werden, wenn ein spezifisches Informationssystem zur Verfügung steht, das bestimmte Parameter wie Hard- und Software oder Lizenzen einhält. Die Parameter können aus den Metadaten oder allenfalls aus einer Hilfsdarstellung ausgelesen werden. Besonders problematisch sind Sicherheitsmechanismen wie DRM oder digitale Signaturen, die eine Darstellung verfälschen oder verhindern können.

Die Darstellung eines digitalen Objekts hängt von dem Informationssystem ab, mit dem es erstellt wurde, und zusätzlich auch von dem, mit dem es erneut dargestellt wird. Je stärker sich die beiden Informationssysteme in ihren Komponenten unterscheiden, desto wahrscheinlicher ist eine verfälschte, das heisst nicht vom Urheber intendierte Darstellung. Deshalb ist vor allem die Frage zu klären, ob die *vorliegende Darstellung vom Urheber intendiert* wurde⁸⁸⁵, und zwar bei allen Objektformen.

⁸⁸² Dazu bspw.: Groebner: Wissenschaftssprache digital. Die Zukunft von gestern, 2014, S. 122.

⁸⁸³ Nach Hans Geser, in: Nickl, Roger; Gull, Thomas: Wer herrscht im Internet? Interview mit Hans Geser und Michael Latzer, in: magazin – Universität Zürich (4), 2011, S. 38.

⁸⁸⁴ Balkanisierung ist nach Van Alstyne und Brynjolfsson eine interessenengebundene Fragmentierung der Interaktionsgruppen, die nicht entlang geographischer Grenzen verläuft. Van Alstyne, Marshall; Brynjolfsson, Erik: Global village or cyber-balkans? Modeling and measuring the integration of electronic communities, in: Management Science 51 (6), 2005, S. 3.

⁸⁸⁵ Vgl. Kap. 2.1. Eigenschaften eines Digitalen Objekts.

Für die korrekte Darstellung und Rezeption durch den Nutzer muss das Informationssystem über alle nötigen Komponenten verfügen oder diese korrekt emulieren können. Angaben in den Metadaten können helfen, die nötige Konfiguration des Informationssystems vorzunehmen. Allenfalls können mit einer behelfsmässigen Darstellung des Objekts durch Try-and-Error-Verfahren oder Emulation fehlende Informationen ausgelesen werden, die dann für eine korrekte Darstellung verwendet werden können. Wenn die nötigen Metadaten nicht vorhanden sind, ist stets von einer behelfsmässigen und nicht authentischen Darstellung auszugehen und es ist abzuwägen, ob die Darstellung den Ansprüchen (der Fragestellung) genügt. Beispielsweise ist die Darstellung eines mit Microsoft Word erstellten Dokuments unter anderem von der verwendeten Schrifttype abhängig.⁸⁸⁶ Ist diese im darstellenden Word nicht vorhanden, zeigt es unter Umständen falsche Zeichen an und der Inhalt wird für den Nutzer unter Umständen unverständlich. Mit der Wahl einer anderen Schrifttype wie Arial kann eventuell der Inhalt trotzdem gelesen und interpretiert werden. Zur Vermeidung dieses Problems wurde das PDF entwickelt, mit dem ein Objekt auf unterschiedlichen Informationssystemen korrekt dargestellt werden kann. Obwohl dem nicht so ist, wie ich bereits aufzeigte, sind die Unterschiede in der Darstellung in den meisten Fällen vernachlässigbar. Bei einem Textdokument ist meistens nur der Text von Interesse; ob der Hintergrund weiss oder gelblich ist, kann irrelevant sein. Bei einem Bildobjekt hingegen ist das Bild mit all seinen Parametern wie Dimension oder der Farbsättigung von Interesse. Problematisch kann beispielsweise sein, wenn die Darstellung nicht in der intendierten Grösse auf einer Kinoleinwand möglich ist oder das Objekt durch einen Bildschirmfehler bläulich dargestellt wird.

Je komplexer das Dateiformat eines digitalen Objekts aufgebaut ist, desto mehr Parameter müssen für seine Darstellung eingehalten werden. Aber auch die Darstellung eines Objekts in einem einfachen Dateiformat kann problematisch sein. Beispielsweise kann ein Textelement einer Webseite in einem Browser korrekt, durch fehlerhafte Programmierung einer Komponente in einem anderen jedoch unvollständig dargestellt werden. Wird der Text beispielsweise seitlich beschnitten, ist das Darstellungsproblem offensichtlich. Ist er aber zufälligerweise nach einem Absatz gekürzt, wird unter Umständen die Aussage des Textes verfälscht. Dies kann zu Fehlinterpretationen führen, insbesondere wenn keine aussagekräftigen Metadaten oder kein zusätzliches Objekt mit vollständigem Inhalt zur Verfügung stehen.

Dieses Problem wird bei Emulationen vergrössert, wo das Informationssystem um eine softwaretechnische Ebene ergänzt wird, die ebenfalls auf korrekte Datenverarbeitung hin zu überprüfen ist. Bei Emulationen treten beispielsweise Probleme mit schnellerer oder langsamerer Verarbeitung sowie mit fehlerhafter Darstellung aufgrund von Programmierfehlern im Emulator auf. Der Aussagewert einer digitalen Quelle kann sich dadurch ändern,

⁸⁸⁶ Siehe weiterführend: Schassan: Digitale Quellen: Datei- und Datenformate, 2016.

insbesondere "wenn ihre inhaltlichen Informationen von der Interaktion mit ihrem Betrachter über die Ein- und Ausgabegeräte abhängig sind."⁸⁸⁷ Die für eine Darstellung nötigen Komponenten müssen zur Verfügung stehen. Zudem ist das Zusammenspiel zwischen dem verwendeten Informationssystem und dem emulierten Objekt auf dessen vollständige Funktionalität hin zu überprüfen. In der Überlieferungsgeschichte oder zusätzlichen Erfahrungsberichten aus früheren Verwendungen können Hinweise über das korrekte Funktionieren des Emulators enthalten sein. Daraus lässt sich schliessen, dass die erneute Darstellung mit hoher Wahrscheinlichkeit korrekt ist, wenn die verlangten Parameter eingehalten werden.

Sicherheits- und Schutzmechanismen wie DRM verhindern oft eine korrekte Darstellung. Solche Systeme verwenden zum Schutz der Integrität und der Identität des digitalen Objekts komplexe Algorithmen und benötigen meistens spezifische Komponenten und Lizenzen. Sind diese nicht vorhanden, kann das Objekt nicht oder nicht korrekt dargestellt werden. Objekte mit Schutzmechanismen werden deshalb nur selten in Langzeitarchive aufgenommen. Geschützte Objekte gehen vermutlich verloren, wenn sie nicht in ein ungeschütztes Format migriert und/oder wenn die entsprechend nötigen Komponenten nicht mitgesichert werden.⁸⁸⁸ Auch die digitale Signatur ist keine Ausnahme. Beim Signieren wird aus dem zu signierenden Objekt mit der Signatur ein zusätzlicher Datensatz generiert, der entweder zusätzlich zum Objekt vorzuhalten oder in dieses zu integrieren ist. Letzteres ist beispielsweise bei Dateiformaten wie PDF oder XPS sowie XML-basierten Formaten möglich. Die Darstellung kann allerdings fehlerhaft sein, wenn das Informationssystem die dafür nötigen Komponenten nicht zur Verfügung stellt und deshalb die sicherheitsrelevanten Elemente des Objekts falsch interpretiert werden, oder wenn die Signatur nicht korrekt implementiert wird.

Der Problemkomplex Darstellung umfasst nicht nur das *Darstellen-Können*, sondern auch die Art und Weise der Darstellung. Es ist besonders darauf zu achten, ob und wie das dargestellte Objekt inszeniert wurde.⁸⁸⁹ Wenn dies aus den Metadaten nicht ersichtlich ist, "können nicht wahrgenommene und nicht kritisierbare, folglich täuschende Realitätswahrnehmungen produziert werden, die dem Kriterium der Wirklichkeit widersprechen würden."⁸⁹⁰ Insbesondere virtuelle Räume sowie audiovisuelle Darstellungen können einem Nutzer einen falschen Eindruck von Objektivität durch "mangelnde Authentizität, Manipulation, ja sogar Fälschung"⁸⁹¹ vermitteln. Es ist ohne Zusatzinformationen kaum möglich, zulässige von

⁸⁸⁷ Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 405.

⁸⁸⁸ Die Verwendung von resp. Konvertierung in Open Access Formate wird zwar als 'Lösung' propagiert, wird aber oft aus wirtschaftlichen Gründen seitens der Urheberrechtsinhaber nicht umgesetzt oder sogar verhindert.

⁸⁸⁹ Siehe auch Kap. 5.4.4. Objektbeschreibung.

⁸⁹⁰ Schwabe: Historisches Lernen im World Wide Web, 2012, S. 140.

⁸⁹¹ Ebd.

unzulässigen Ergänzungen, Manipulationen, Vereinfachungen oder künstlerischem Ausdruck zu unterscheiden.

Die Darstellung eines digitalen Objekts ist, im Gegensatz zu physischen Objekten, von einem Informationssystem abhängig, das spezifische Parameter erfüllen muss. Physische Objekte können dagegen unmittelbar verwendet werden. Der Nutzer hat bei dargestellten, digitalen Objekten zuerst zu beurteilen und zu entscheiden, ob die ihm sich präsentierende Darstellung korrekt ist, ob erkannte Fehler vernachlässigt werden können oder ob das Objekt aufgrund dieser Fehler nicht verwendet werden kann.

5.4.11. Inhalt

"Gründlich recherchierte historische Forschungsergebnisse stehen im World Wide Web nur einen Klick entfernt von interessegeleiteten manipulierten Aussagen [...]"⁸⁹² (Schwabe, 2012)

Der Inhalt eines physischen oder digitalen Objekts ist auf seine Plausibilität und Vollständigkeit hin zu überprüfen, wobei die bei digitalen Objekten auch im Code versteckten Informationen überprüft werden sollten. Da ein digitales Objekt manipuliert werden kann, ist es als Manuskript aufzufassen, das nur mit Sicherheitsmerkmalen wie der digitalen Signatur versioniert werden kann. Inhalte können im Web sehr schnell und nutzerspezifisch publiziert werden. Der Nutzer kann bei Hyperlinks selten zwischen einem Hinweis, den er lesen *kann*, und einer Ergänzung, die er lesen *muss*, unterscheiden. Deshalb sollten die in einem Objekt enthaltenen Links jeweils überprüft werden.

Eine unter Umständen mit vernachlässigbaren Abweichungen korrekte Darstellung des digitalen Objekts ermöglicht erst die Untersuchung seines Inhalts. Dieser gibt eventuell Aufschluss über Autorschaft, Adressat, Überlieferungsgeschichte, Motiv, Datierung und Beziehungen zu anderen Objekten. Formalien sowie sprachliche und sachliche Aufschlüsselung sind weitere Bestandteile der kritischen Untersuchung.⁸⁹³

Da digitale Objekte als *Code* vorliegen, erfahren sie bereits eine technische Interpretation durch das darstellende Informationssystem. Der Code ist im Normalfall nicht als Inhalt aufzufassen, sondern ist nur ein notwendiges Hilfsmittel zur Darstellung des Objekts, der zur inhaltlichen Aussage nichts beiträgt. John Cayley bringt dies mit folgender Formulierung auf den Punkt: "The Code is not the Text (unless it is the Text)."⁸⁹⁴ Allerdings kann ein Programm-

⁸⁹² Ebd., S. 143. Schriftschnitt im Original invers.

⁸⁹³ Siehe dazu Kap. 3.1.3. Quellenkritik.

⁸⁹⁴ Cayley: *The Code is not the Text (unless it is the Text)*, 2002. Siehe auch Kap. 3.3. Objekt, Forschungsressource, Quelle.

code zusätzliche inhaltliche Informationen enthalten (Quellentypus der unsichtbar codierten Information), die auch für die Interpretation von Interesse sein können. Deshalb ist die Untersuchung von im Code enthaltenen Kommentaren wichtig, die in fast jeder Programmiersprache eingefügt werden können. Beispielsweise könnte eine HTML-Seite mit einem Text den folgenden Kommentar enthalten:

'URL: !--- aus digital past kopiert; mit Brüggers Paratext ergänzen & umformulieren! ---'

Diese inhaltliche Ergänzung beeinflusst wahrscheinlich auch die Interpretation des Textes, weil der Abschnitt plagiiert sein könnte.

Der Inhalt digitaler und analoger Forschungsressourcen ist auf seine *Plausibilität* hin zu überprüfen, obwohl Daniel Cohen digitalen historischen Dokumenten einen sehr hohen Wahrheitsgehalt attestiert:

*"Concern about the falsification of digital historical documents and materials, we believe, has mostly turned out to be a phantom problem. [...] In one study, the Philadelphia Inquirer discovered that only about 10 to 15 percent of their 300'000 registered users had entered bad email addresses [...]."*⁸⁹⁵

Dass Cohen von korrekten E-Mail-Adressen auf die Korrektheit des Inhalts schliesst, kann nicht nachvollzogen werden, zumal auch eine Abweichung von 15% für wissenschaftliche Zwecke ein eher hoher Wert darstellt. Vor allem zwei Gründe sprechen gegen Cohens These: Erstens können Nutzer zwar inhaltlich korrekte Angaben machen, wollen aber beispielsweise zur Vermeidung von Spam ihre korrekte E-Mail-Adresse nicht preisgeben. Zweitens ist eine korrekte E-Mail-Adresse auch keine Garantie für inhaltlich korrekte Angaben. Wird beispielsweise eine Umfrage mit einem Wettbewerb verbunden, wird ein Teilnehmer sehr wahrscheinlich seine korrekte E-Mail-Adresse hinterlegen. Aber er muss sich nicht zwingend eingehend mit den Fragen beschäftigen, sondern kann sich auch 'durchgeklickt' haben – seine Motivation ist der potenzielle Gewinn.

Aufgrund der Möglichkeiten, digitale Objekte nachträglich zu manipulieren oder anzupassen und informationstechnische Systeme zu vernetzen, sowie des Wegfalls vertrauenswürdiger Produktions- und Publikationsprozesse, sind die Inhalte digitaler Objekte noch kritischer zu überprüfen als jene von physischen Objekten. Bei analogen Publikationen können neue Erkenntnisse nur bis zu einem gewissen Zeitpunkt berücksichtigt werden, danach wird der kommunikative Fluss unterbrochen: ""Stopp, jetzt nur das!""⁸⁹⁶, wie Valentin Groebner meint. Hingegen kann beispielsweise ein Blogbeitrag auch nach der Publikation verändert, angepasst oder ergänzt werden, unabhängig vom Publikationsdatum und ohne Kenntnis des

⁸⁹⁵ Cohen, Daniel J.: Collecting History Online, in: Clio wired – The future of the Past in the Digital Age, 2011, S. 182. Siehe dazu auch: CNN: Web newspaper registration stirs debate, CNN.com, 14.06.2004, URL: http://chnm.gmu.edu/digitalhistory/links/cached/chapter6/6_28a_registration.htm, Stand: 17.04.2013.

⁸⁹⁶ Groebner: Wissenschaftssprache digital. Die Zukunft von gestern, 2014, S. 121.

Nutzers. Die manuskriptive Kollaboration zwischen Autor und Adressat kann den Inhalt weiter verändern. Deshalb ist das digitale Objekt stets als *Manuskript* aufzufassen. Dies ist insbesondere bei Objekten im Web zu berücksichtigen, denn auch bei einer vertrauenswürdigen URL oder bekannten Autoren können sich Inhalte einer Publikation absichtlich oder unabsichtlich nachträglich ändern. Dazu ein Beispiel zur Veranschaulichung:

Ein Dozent einer Universität veröffentlicht auf seiner Webseite einen Fachartikel, der von Kollegen und Studenten gelesen und weiterverwendet wird. Nach wenigen Tagen fällt dem Dozenten auf, dass in einem Absatz ein 'nicht' fehlt. Er korrigiert dies und ersetzt die publizierte Datei durch die angepasste. Angaben zur veränderten Version vergisst er aus Zeitgründen und Unachtsamkeit. Auf der Webseite ist deshalb auch nicht ersichtlich, dass nun eine neue Version mit einer inhaltlichen Veränderung publiziert ist. Wer zufälligerweise die alte und falsche Formulierung zitiert hatte und seine Argumentation darauf stützte, der kann sich nun nicht mehr darauf berufen, denn es steht ausschliesslich die angepasste Datei zur Verfügung. Der Zitierende gerät in Erklärungsnot, denn er kann die ursprüngliche Formulierung nicht von sich aus beweisen. Ein allenfalls von ihm gesichertes Surrogat der ersten Version ist ein schwacher Beweis, denn es wäre einerseits irgendwie zu publizieren und andererseits könnte die Datei theoretisch durch ihn manipuliert worden sein.

Mit Identifikations- und Sicherheitsmerkmalen wie der qualifizierten digitalen Signatur versehene Objekte sind jedoch als Beweis verwendbar. Damit kann nicht nur die neue Version erkannt, sondern auch die alte Version könnte gegenüber Dritten legitimiert werden. Der Zitierende könnte damit nachweisen, dass die von ihm verwendete Version tatsächlich existierte. Mögliche Vorwürfe wie 'Das hast Du Dir ausgedacht!' oder 'Du hast falsch abgeschrieben!' könnten entkräftet werden.

Eine Veränderung des Inhalts kann auch durch technische Massnahmen erfolgen. So erlauben beispielsweise zeitverzögerte Übertragungen von als 'live' bezeichneten Produktionen Eingriffe in den Inhalt und "machen das Unmanipulierbare manipulierbar"⁸⁹⁷ und täuschen die dem 'live' eigentlich inhärente Unmittelbarkeit nur vor. Wenn nun tatsächlich eingegriffen wird, wobei der Grund unerheblich ist, dann wird der übertragene Inhalt manipuliert und seine Aussage(n) möglicherweise verändert. Die Manipulation ist dann problematisch, wenn sie nicht vom Urheber selbst vorgenommen wird und das geänderte Objekt deshalb nicht seiner intendierten Darstellung und unter Umständen auch nicht seiner inhaltlichen Aussage entspricht.

⁸⁹⁷ Kittler, Friedrich A.: Grammophon, Film, Typewriter, Berlin 1986, S. 166. Bspw. Zeitverzögerung bei Grossveranstaltung zur Verhinderung von sexuellen oder gewalttätigen Darstellungen. Beispiel: die Übertragung des jährlichen Finalspiels der American Super Bowl der National Football League (USA) erfolgt seit dem 'Nippelgate' (Janet Jackson und Justin Timberlake, 1. Februar 2004, 38. Super Bowl) nur noch mit einer Verzögerung von fünf Sekunden.

Vor allem im Web können mit verschiedenen Verfahren *domain-, orts- oder browserspezifische Informationen* bereitgestellt werden. Je nach aufgerufener Domain und verwendetem Browser werden andere Informationen präsentiert. Beispielsweise zeigt Google Maps je nach verwendeter URL unterschiedliche Grenzverläufe in der umstrittenen Region Kaschmir. Der Grenzverlauf ist bei Google Maps Indien eindeutig Indien zugeschlagen, während bei Google Maps Schweiz die Region mit gestrichelten Grenzverläufen als nicht eindeutiges Hoheitsgebiet eingezeichnet ist.⁸⁹⁸ Ohne Hintergrundwissen sind solche inhaltlichen Unterschiede nicht zu erkennen.



Abbildung 24: Grenzen Kaschmirs, Ausschnitt Google Maps Schweiz (Google Maps)⁸⁹⁹

⁸⁹⁸ Gründe sind v.a. lokale rechtliche Vorschriften. Für das aufgeführte Beispiel wurde im Firefox 39.0 die Ortungsfunktion (geo.enabled) ausgeschaltet. URL für die Schweiz (<https://www.google.ch/maps>), Indien (<https://www.google.co.in/maps>); Stand: 17.07.2015 – Suchwort jeweils 'Indien'. Kartenausschnitte (Abbildungen 24 und 25) als Bildzitate von Google Maps/Google Earth (Google LLC, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA). Kartenanbieter 'Worldwide': Unicode Common Locale Data Repository © Unicode, Inc. (<http://www.unicode.org/copyright.html>) All rights reserved und © 2010 Europa Technologies.

⁸⁹⁹ Zum Bildzitat siehe FN 898.



Abbildung 25: Grenzen Kaschmirs, Ausschnitt Google Maps Indien (Google Maps)⁹⁰⁰

Browserspezifische (oft inkorrekt als personenspezifisch bezeichnete) Inhalte können beispielsweise mittels *Browser-Fingerprinting*⁹⁰¹ abgerufen werden. Bei diesem Verfahren ruft ein System verschiedene Merkmale wie Land, Sprache oder installierte Ad-ons des eingesetzten Browsers ab, womit in den meisten Fällen ein Browser eindeutig identifiziert werden kann. Das ist nicht nur für die Verbreitung von Werbung interessant, sondern auch für die gezielte Verbreitung von Inhalten. Inwiefern Informationen auf den Nutzer abgestimmt präsentiert werden oder der Nutzer mit den präsentierten Informationen gezielt manipuliert wird, ist oft schwierig zu beurteilen. Manipulative Methoden, im paternalistischen Sinne oft als *nudging*⁹⁰² bezeichnet, beeinträchtigen allerdings die Art und Weise wie der Nutzer seine Entscheidung trifft, was aus wissenschaftlicher Sicht nicht statthaft ist.⁹⁰³

Die Verbreitung von Informationen hat durch die Kommunikationsmöglichkeiten des Internets massiv an Geschwindigkeit zugenommen. Um Nutzer zufrieden stellen zu können, *müssen* interessante *News* unmittelbar publiziert werden. (Online-)Publikationsmedien und auch Fachpublikationen stehen unter grossem Druck, diesen Ansprüchen gerecht zu werden, weil

⁹⁰⁰ Zum Bildzitat siehe FN 898.

⁹⁰¹ Zur persönlichen Informationssicherheit siehe FN 430.

⁹⁰² Vgl.: Helbing, Dirk; Frey, Bruno; Zicari, Roberto u. a.: Digitale Demokratie statt Datendiktatur, in: Das Digital-Manifest, 2015, S. 5-19; Helbing, Dirk: »BIG NUDGING« – zur Problemlösung wenig geeignet, in: Das Digital-Manifest, 2015, S. 37-39.

⁹⁰³ Vgl.: Helbing u. a.: Digitale Demokratie statt Datendiktatur, 2015, S. 11.

sonst Nutzer und damit Einnahmen verloren gehen und der Marktwert der Unternehmung schwindet. Nutzerzahlen sind die wichtigste Währung für den finanziellen Erfolg eines Publikationsmediums, besonders bei kostenlosen Informationsplattformen. Aber die "Geschwindigkeit und hohe Sendefrequenz, die für schnelle Reaktionen und Unmittelbarkeit sorgen, machen solche digitalen Plattformen zum Medium für immer neue aufgeregte Debatten. Und für beschleunigtes Vergessen: Fertiges, Konzentriertes, Abgeschlossenes geht darin unter."⁹⁰⁴ Eine seriöse und rechercheintensive Berichterstattung ist nur noch bei umfangreicheren Reportagen möglich und gewünscht, die allerdings oft in anderer Form publiziert werden, beispielsweise gedruckt. Auch analoge Zeitungen schreiben und schreiben voneinander ab, indem sie ihre Geschichten nur leicht abändern und mit etwas Kaffeeklatsch anreichern.⁹⁰⁵ Oft geht die journalistische Berufsethik vergessen und Falschmeldungen werden verbreitet. Informationen waren und sind niemals stabil, wie auch der Historiker und Bibliothekar Robert Darnton meint:

*"Informationen sollten nicht als harte Fakten oder als Brocken von Wirklichkeit begriffen werden, [...] sondern als Botschaften, die im Zuge ihrer Übermittlung ständig verändert werden. An Stelle von fest gefügten Dokumenten haben wir es mit multip-len, veränderbaren Texten zu tun."*⁹⁰⁶

Bei Online-Publikationsplattformen sind oft nicht nur Berufs-, sondern auch sogenannte Bürgerjournalisten tätig, was die Gefahr von tendenziösen oder auch falschen Darstellungen massiv erhöht. Die Manipulation der und durch die Massenmedien ist an sich nichts Neues, hat aber in ihrer Unmittelbarkeit eine neue Dimension erreicht:⁹⁰⁷ Eine Zeitung muss heute ihre Inhalte einmal pro Tag veröffentlichen, ein Nachrichtenfernsehsender 24 Stunden an 365 Tagen pro Jahr. Ein Online-Publikationsformat hat hingegen einen fast unendlichen Platz zu füllen, weshalb gemäss dem selbsternannten 'Medienmanipulator' Ryan Holiday nur die News-Website gewinnt, "that covers the most stuff."⁹⁰⁸ Solche Webseiten benötigen für das Überleben genügend Besucher und wer zuerst publiziert, wird eher beachtet. Deshalb wird eine äusserst interessante Information trotz der fehlenden Überprüfbarkeit ihres Wahrheitsgehalts eher publiziert als zurückgehalten. Falschmeldungen können später berichtigt werden, was Holidays Taktiken *Give them what spreads, not what's good* und *Just make stuff up (everyone else is doing it)* in die Hände spielt.⁹⁰⁹

⁹⁰⁴ Groebner: Wissenschaftssprache digital. Die Zukunft von gestern, 2014, S. 116 f.

⁹⁰⁵ Nach: Darnton, Robert: Die Bibliothek im Informationszeitalter. 6000 Jahre Schrift, in: Becker, Konrad (Hg.): Deep Search: Politik des Suchens jenseits von Google, Innsbruck 2009, S. 41 f.

⁹⁰⁶ Ebd., S. 43.

⁹⁰⁷ Zu Medienkritik siehe bspw.: Krüger, Uwe: Meinungsmacht: der Einfluss von Eliten auf Leitmedien und Alpha-Journalisten – eine kritische Netzwerkanalyse, Köln 2013 (Reihe des Instituts für Praktische Journalismus- und Kommunikationsforschung 9).

⁹⁰⁸ Holiday, Ryan: Trust me, I'm lying. Confessions of a media manipulator, New York 2013, S. 14. Holiday bezeichnet jede Online-Publikationsform als Blog (Ebd. S. 13).

⁹⁰⁹ Siehe: Ebd., Kap. 6 und 12.

Es bestehen allerdings erste Ansätze, um die Vertrauenswürdigkeit und Qualität von Webinhalten zu bewerten. Die Wissensforscherin Xin Luna Dong hat ein Modell entwickelt, mit dem ein sogenannter "trustworthiness score" namens Knowledge-Based Trust (KBT) algorithmisch errechnet wird.⁹¹⁰ Erste Tests seien vielversprechend verlaufen und würden Fortschritte in der Qualifizierung von Quellen aus dem Web aufzeigen. Dies ist zwar ein guter Ansatz, denn mit dem KBT würde die quellenkritische Analyse durch den Nutzer vereinfacht werden. Allerdings wäre auch dieser einer kritischen Analyse zu unterwerfen und bei jeder Änderung des Algorithmus neu zu überprüfen.

Die Darstellung eines digitalen Objekts ist nicht nur auf deren Korrektheit, sondern auch auf fehlende Elemente wie Text oder Bilder hin zu untersuchen. Ein grundsätzliches Problem zur Feststellung der *Vollständigkeit* des Inhalts stellt das Hypertextkonzept an sich dar, weil sich der Leser nicht sicher sein kann, ob er alle relevanten Inhalte gelesen hat oder nicht⁹¹¹, denn "das Netz sinnvoller Verknüpfungen und Beziehungen endet nie."⁹¹² Zu unterscheiden sind *referenzielle* und *typisierte* Links.⁹¹³ Während referenzielle Links auf "formalen, syntaktischen (vorwärts, rückwärts etc.) beziehungsweise assoziativen Prinzipien"⁹¹⁴ beruhen, spezifizieren im Gegensatz dazu typisierte Links deren Beziehung zu den verknüpften Inhalten. Assoziativ-referenzielle Links sind den geläufigen Fussnoten ähnlich, die auf Quellen und weiterführende Literatur oder Objekte verweisen. Typisierte Hyperlinks werden vom Autor eingesetzt, um auf alles zu verweisen, was er nicht mehr selbst zu erklären braucht. Das Objekt, auf das verwiesen wird, ist so zwar nicht im vorliegenden integriert, es kann aber Kohärenzen und Relationen aufzeigen und deshalb dennoch Teil des Inhalts sein, insbesondere bei Textobjekten.⁹¹⁵

Wie bereits erwähnt, kann der Nutzer bei einem Hyperlink selten zwischen einem schlichten Hinweis, den er weiterverfolgen *kann*, und einer Aneignung des Inhalts als Aussage und Teil des vorliegenden Objekts, den er lesen *muss*, unterscheiden. Dazu müsste er jeden Link nachverfolgen und die darin befindlichen Inhalte überprüfen. Die verknüpften Objekte können

⁹¹⁰ Siehe: Dong, Xin Luna; Gabrilovich, Evgeniy; Murphy, Kevin u. a.: Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources, arXiv:1502.03519 [cs.DB], 11.02.2015. URL: <http://arxiv.org/abs/1502.03519>, Stand: 03.03.2015.

⁹¹¹ Siehe dazu: Haber, Peter: Geschichte schreiben im Digitalen Zeitalter, in: Digitale Wissenschaft: Stand und Entwicklung digital vernetzter Forschung in Deutschland ; 20./21. September 2010, Köln ; Beiträge der Tagung, Köln 2011, S. 23.

⁹¹² Schmale: Digitale Geschichtswissenschaft, 2010, S. 110.

⁹¹³ Vertiefend: Kuhlen: Hypertext, 1991, S. 113-123.

⁹¹⁴ Krameritsch: Geschichte(n) im Netzwerk. Hypertext und dessen Potenziale für die Produktion, Repräsentation und Rezeption der historischen Erzählung, 2007, S. 140. Krameritsch reduziert Kuhlens Typologisierung von Verknüpfungen auf die zwei Grundtypen 'referenziell' und 'typisiert', was für das Verständnis unterschiedlicher Linktypen ausreicht und so übernommen wird. Deshalb siehe vertiefend: Kuhlen: Hypertext, 1991, Kap. 2.2 (S. 98-123).

⁹¹⁵ Nach: Krameritsch: Geschichte(n) im Netzwerk. Hypertext und dessen Potenziale für die Produktion, Repräsentation und Rezeption der historischen Erzählung, 2007, S. 143-146.

wiederum Hyperlinks enthalten, die Teil des neu aufgerufenen Objekts sind und ihrerseits einer Überprüfung bedürfen.

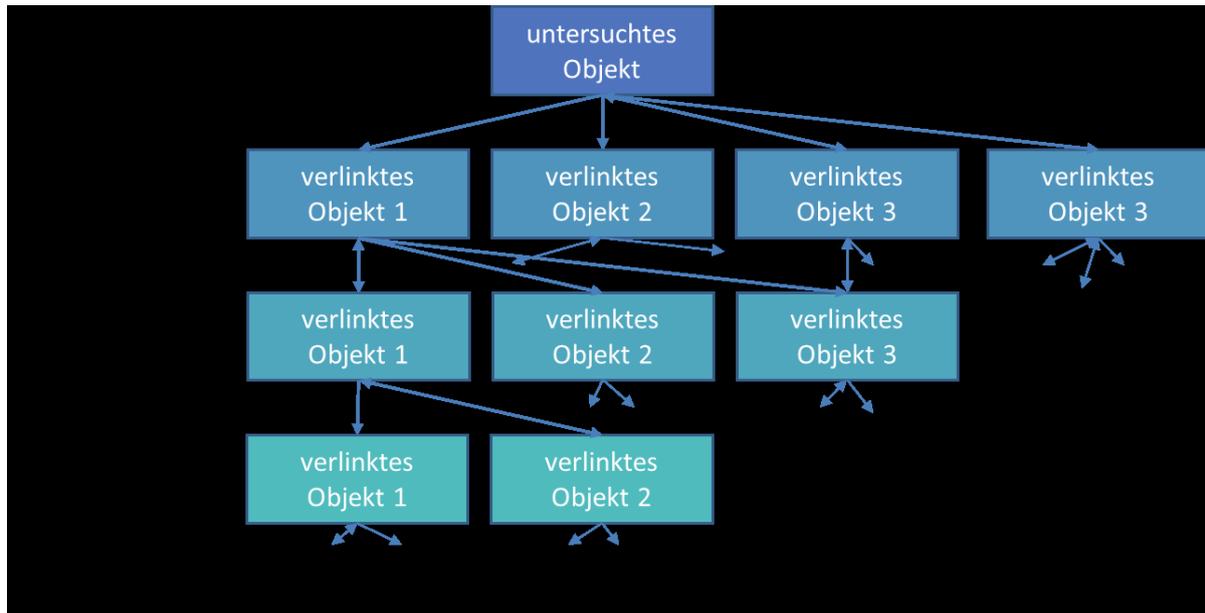


Abbildung 26: Schema 'Linkhierarchie' (Föhr)

Bei diesem nie endenden Netz von Verknüpfungen stellt sich die Frage, wann die Überprüfung abgebrochen werden kann. Aus pragmatischen Gründen wird zwar oft angenommen, dass nur die nächste mit dem Link aufgerufene Ebene in die Interpretation mit einbezogen werden sollte. Dies ist jedoch keine allgemein anerkannte Konvention und entspricht auch nicht der Funktionalität des Hypertextes, der zu "netzwerkartig verknüpften Informationskomplexen"⁹¹⁶ tendiert. Der Nutzer hat aufgrund der Untersuchung des Objekts und seiner eigenen Erfahrung abzuwägen, ob das verlinkte Objekt oder allenfalls auch darin verlinkte Objekte als Teil des Inhalts aufzufassen und damit auch für die Interpretation zu beachten sind.

Die Integration und Verknüpfung von externen Inhalten bildet den grossen Unterschied zwischen digitalen und physischen Objekten. Bei physischen Publikationen ist der Verweis auf andere Inhalte nur dann möglich, wenn diese zum Verständnis nicht benötigt werden oder wenn sich diese Inhalte innerhalb derselben Publikation befinden. Bei digitalen Objekten hingegen können auch Inhalte verlinkt werden, die für das Verständnis vorausgesetzt werden und sich nicht innerhalb des Objekts befinden. Der Autor einer digitalen Publikation kann vom Nutzer verlangen, dass dieser den angegebenen Links folgt, weil die Inhalte meistens auch unmittelbar zur Verfügung stehen. Bei physischen Publikationen kann dies nicht vorausgesetzt werden. Selbst wenn sich der Nutzer an einem Ort wie einer Bibliothek befindet

⁹¹⁶ Lengwiler: Praxisbuch Geschichte, 2011, S. 266.

und alle verknüpften Werke zur Verfügung stehen, ist dies nur bei kurzen Texten mit sehr wenigen Verweisen praktikabel, aber auch dann nicht sinnvoll.

5.4.12. Beziehungen zu Objekten

Zwischen Objekten können Beziehungen entstehen, die durch inhaltliche, hypertextuelle, sprachliche oder symbolische Verweise oder durch ihre Einbettung in eine Sammlung oder andere Verknüpfungen definiert sind. Digitale Objekte können, im Gegensatz zu physischen, sogar so stark von anderen Objekten abhängig sein, dass sie ohne diese nicht dargestellt werden können oder keine sinnvollen Inhalte liefern. Bei digitalen Objekten werden Beziehungen untereinander hauptsächlich mittels Hyperlinks realisiert. Mit einer qualitativen Analyse der Hyperlinks können Beziehungen und Abhängigkeiten festgestellt und beurteilt werden. Der Autor drückt mit dem Verknüpfen und Nicht-Verknüpfen eines Objekts mit einem anderen zudem seine Intention aus. Deshalb sind die in einem Objekt enthaltenen Links kritisch zu überprüfen und zu beurteilen – der Link-Check.

Das digitale Objekt ist, wie in der traditionellen Quellenkritik, in den Kontext anderer Quellen einzubetten und mit diesen zu vergleichen. Es ist zu recherchieren, was in anderen, als autoritativ bekannten Quellen über das vorgefundene Objekt steht⁹¹⁷, und es ist zu überprüfen, worauf sich das Objekt bezieht und ob es gegebenenfalls auf welche weiteren Objekte verweist. Damit können das Verhältnis der Quellen zueinander, der zeitliche und räumliche Abstand zwischen den Quellen sowie eventuell wiederkehrende Objekte oder Motive geklärt werden. Allerdings wird das Einordnen in einen spezifischen Kontext erschwert oder gar verunmöglicht, wenn die Autorschaft und der Entstehungskontext eines Objekts unbekannt sind.⁹¹⁸

Um zu erfahren, was andere Online-Quellen über das vorgefundene Objekt berichten, ist nach Hyperlinks zu suchen, die von der Online-Quelle auf das Objekt verweisen, die sogenannten *Backlinks*. Backlink-Tracker⁹¹⁹ geben Hinweise auf die Anzahl der Backlinks und unter welcher URL diese zu finden sind. Allerdings arbeiten solche Tracker sehr unzuverlässig, wie das Beispiel von 'hsc.hypotheses.org' zeigt: eine Überprüfung ergab Resultate von 0 (analyzebacklinks), über 23 (backlinktest) bis zu mindestens 1'000 Treffern (mehr wird nicht gelistet; Backlinkwatch).⁹²⁰ Beziehungen unter Objekten dürfen deshalb nicht nur mittels

⁹¹⁷ Siehe dazu: Pfanzer: Von der Quellenkritik zum kritischen Umgang mit digitalen Ressourcen, 2010.

⁹¹⁸ Nach: Lengwiler: Praxisbuch Geschichte, 2011, S. 266.

⁹¹⁹ Backlink-Tracker: Spezielle Suchmaschine, die nach Backlinks zur eingegebenen URL sucht.

⁹²⁰ Überprüfung am 25.05.2015. Backlink-Tracker bspw.: Backlinkwatch (URL: <http://www.backlinkwatch.com/>), analyzebacklinks (URL: <http://www.analyzebacklinks.com/>), seo-united (URL: <http://www.seo-united.de/backlink-checker/>), European Business Connect (URL:

Tracker oder Suchmaschinen überprüft werden, sondern jedes vertrauenswürdige (Fach-) Portal ist einzeln auf mögliche Backlinks zu durchsuchen.⁹²¹ Fachportale sind aber auch nicht in jedem Falle zuverlässig, wie das viel zitierte Beispiel des Historikers Stuart Jenks zeigt, dessen Jux-Beitrag in einer Festschrift im Jahre 2000 von der Online-Plattform Infoseek als bester Beitrag zum Thema *Schwarzer Tod* bezeichnet wurde.⁹²² Auch fast 20 Jahre später sind solche Fehler nicht auszuschließen.⁹²³ Trotzdem können Backlinks Hinweise zur Vertrauenswürdigkeit des Objekts geben, abhängig von der Reputation der Institution, auf dessen Webauftritt der Backlink aufgeführt ist.

Bei Objekten, die Hyperlinks enthalten, sollte auch deren *Verwendungsgrund* überprüft werden. Ein Hyperlink wird nicht nur anstelle einer herkömmlichen Fussnote als Quellennachweis verwendet, sondern kann mehrere Funktionen erfüllen wie beispielsweise Navigation oder die Aneignung des darunter aufrufbaren Inhalts.⁹²⁴ Es ist deshalb zu untersuchen, mit welcher Absicht dieser Link gesetzt und damit eine Beziehung zwischen zwei Objekten hergestellt wurde.⁹²⁵ Mit dieser *qualitativen Untersuchung*⁹²⁶ von Hyperlinks, die im Folgenden als *Link-Check (LC)* bezeichnet wird, soll festgestellt werden, welche Links für die kritische Beurteilung relevant sind und mit welcher Intention der Autor diese gesetzt hat. Als relevant sind Links zu bezeichnen, die auf zusätzliche, das Primärobjekt *inhaltlich ergänzende* Informationen verweisen, wie ich im vorangegangenen Kapitel ausführte. Allerdings kann ein Link je nach Fragestellung unterschiedlich beurteilt werden. Beispielsweise ist der Link aus dem Wort *Quelle* auf 'hist.net' für das Thema der Quellenkritik oder für eine Netzwerkanalyse des Autors interessant. Wenn das Wort jedoch in einem Text zur Frage nach der Bedeutung der Schlacht von Morgarten für die junge Schweizerische Eidgenossenschaft enthalten ist, könnte der Link für den Nutzer irrelevant sein. Der Link könnte vom Autor eventuell nur eingefügt worden sein, um einerseits seine persönliche Beziehung zum Autor der genannten Webseite aufzuzeigen oder um auf den aktuellen Quellendiskurs hinzuweisen und damit seine Kenntnisse über den Stand der geführten Diskussionen hervorzuheben.

<http://www.european-business-connect.de/backlinkchecker.php>), backlinktest – Beta (URL: <http://www.backlinktest.com/>).

⁹²¹ Zur Problematik mit Suchmaschinen siehe Kap. 3.1. Traditionelle Prozessschritte und digitale Objekte.

⁹²² Nach: Jenks, Stuart; Marra, Stephanie: Über die Verlässlichkeit von Informationen im Internet, in: Jenks, Stuart; Marra, Stephanie (Hg.): *Internet-Handbuch Geschichte*, Köln 2001, S. 265-271.

⁹²³ Siehe dazu bspw. der 'Test' des Bloggers unter dem Pseudonym Neurosceptic auf einer vertrauenswürdigen Webseite: Neurosceptic: *Predatory Journals Hit By «Star Wars» Sting*, Blog, Neurosceptic, 22.07.2017, URL: <http://blogs.discovermagazine.com/neurosceptic/2017/07/22/predatory-journals-star-wars-sting/>, Stand: 20.02.2018.

⁹²⁴ Vgl.: Kühlen: *Hypertext*, 1991, S. 77-123.

⁹²⁵ Siehe auch Kap. 5.4.11. Inhalt.

⁹²⁶ Im Gegensatz dazu untersucht Webometrics oder Cybermetrics Hyperlinks quantitativ, insb. als Netzwerkanalyse. Siehe dazu bspw.: Thelwall, Michael: *Introduction to Webometrics: Quantitative Web Research for the Social Sciences*, in: *Synthesis Lectures on Information Concepts, Retrieval, and Services* 1 (1), 01.01.2009, S. 1-116. DOI: 10.2200/S00176ED1V01Y200903ICR004.

Lösungsansatz Link-Check Element

Die in einem digitalen Objekt enthaltenen Hyperlinks sollten auf zwei Aspekte hin untersucht werden: Erstens auf inhaltliche Verweise, die entweder nur als Hinweis oder als wichtige Ergänzung auftreten, und zweitens auf die Intention der Linksetzung. Ich schlage die Einführung eines *Link-Check Elements LC* vor, mit dem der Nutzer in den Quellenangaben zu digitalen Objekten verzeichnet, ob und wie tief die in einem digitalen Objekt enthaltenen Links überprüft wurden. Dies soll mit dem Kürzel LC und einer sogenannten *Link-Check-Ziffer* ausgedrückt werden. Die Link-Check-Ziffer 0 (LC 0) besagt, dass kein Link des Objekts nachverfolgt wurde. LC 1 bedeutet, dass alle oder definierte Hyperlinks eines Objekts untersucht wurden, indem der Link aufgerufen wurde (1. Ebene). Mit LC 2 wird ausgedrückt, dass auch Hyperlinks der Objekte aus der 1. Ebene untersucht wurden (folglich die 2. Ebene).⁹²⁷ Aufgrund der Fragestellung können gewisse verlinkte Elemente von einer Nachverfolgung als irrelevant angesehen und ausgeschlossen werden, beispielsweise die auf einer Webseite verlinkten Werbebanner. Es ist deshalb auch zu erwähnen, auf welche Objektarten wie Text oder Graphiken sich der Link-Check bezieht und bei welchen gegebenenfalls von einer Überprüfung abgesehen werden kann.

Beispiel:

Föhr, Pascal: 'just a historian', Historical Source Criticism, 09.08.2013, URL: <http://hsc.hypotheses.org/241>, LC 1 (nur Beitragstext), Stand: 06.11.2014.

Beziehungen zu anderen Objekten entstehen nicht nur durch Hyperlinks, sondern auch bei Digitalisaten, vor allem zu seiner physischen Repräsentation, und bei reborn-digitalen Objekten, die beispielsweise aus einer Datenbank generiert wurden. Reborn-digitale Objekte sind teilweise sogar von anderen digitalen Objekten direkt und indirekt abhängig. Eine direkte Abhängigkeit besteht, wenn das Objekt ohne ein anderes nicht dargestellt werden kann. Indirekt abhängig ist es, wenn es zwar dargestellt werden kann, aber Inhalte fehlen, weil diese aus dem anderen Objekt oder aus anderen Objekten stammen und für die Darstellung nur aufbereitet werden.

Auch physische Quellen sollten auf eine mögliche Beziehung zu digitalen Objekten hin überprüft werden. Laut Groebner sind physische Publikationen "gedruckte Informationsspeicher, mit denen die digital verfügbar gemachten Informationen verifiziert und stabilisiert werden."⁹²⁸ Dies stimmt insofern, als dass sich vor allem Digital Immigrants – teilweise berechtigt – auf bewährte Forschungsmethoden und Publikationsmedien stützen, weil digitale Stabilisierungsmechanismen noch zu wenig verlässlich sind. Mit der Publikation in

⁹²⁷ Siehe dazu das Schema 'Linkhierarchie' in Kap. 5.4.11. Inhalt.

⁹²⁸ Groebner: Wissenschaftssprache digital. Die Zukunft von gestern, 2014, S. 120.

einem physischen Medium werden digitale Objekte nicht nur stabilisiert, sondern auch legitimiert. Nicht weil die Informationen damit per se richtiger wären, sondern weil physische Publikationen weiterhin einen hohen Wahrheitsanspruch suggerieren und hohes Vertrauen genießen. Researcher Driven Archiving in Trusted Digital Repositories und mit Sicherheitsmerkmalen versehene digitale Objekte könnten Vorbehalte auf mangelnde Verifikationsmöglichkeiten entkräften.

5.4.13. Authentizitätsapproximation

Die Ausführungen im vorangegangenen Kapitel haben gezeigt, dass die Authentizität eines digitalen Objekts aufgrund seiner Eigenschaften und seiner 'informationstechnischen Umwelt', wozu Informationssysteme, Internet oder Datenbanken gehören, schwierig festzustellen ist. Vor allem die nicht nachvollziehbare Manipulierbarkeit, Volatilität, verlustfreie Vielfältig- und (gleichzeitige) Wiederverwendbarkeit und die Abhängigkeit von einem spezifischen, digitalen Informationssystem sind problembehaftete Eigenschaften eines digitalen Objekts. Problemfelder, die in der informationstechnischen Umwelt auftreten, sind unter anderem die Integrität des Bitcodes, der Informationen, der Darstellung und der Nutzung sowie die (Adress-)Persistenz, die Überlieferungsgeschichte, die Datierung, der Publikations- und Speicherort, die Feststellung der Autorschaft und der Adressaten respektive Nutzer und die Hypermedialität. Für eine adäquate Untersuchung sind vielfältige Methoden mit traditionellen und informationstechnischen Mitteln anzuwenden, was umfassende Medien- und Informationskompetenzen voraussetzt.

Wenn physische Objekte wie Bücher oder Zeitschriften bibliographische Herkunftsangaben aus einem anerkannten Verlag aufweisen, kann diesen per Konvention vertraut werden und nur im Verdachtsfalle sind sie näher zu untersuchen und zu falsifizieren. Hingegen sind die bei digitalen Objekten enthaltenen Informationen zu verifizieren oder zu beglaubigen, wenn keine Sicherheitsmerkmale vorhanden sind. Die Ausgangslage zur Untersuchung eines Objekts wird demnach invertiert: da qualitätssichernde Organe fehlen und oft keine authentizitätsfördernden Merkmale wie Signaturen eingesetzt werden, können die Informationen können nicht mehr einfach übernommen werden, sondern sie sind zuerst zu verifizieren. Mit jeder Bestätigung der Korrektheit einer Angabe steigt die Wahrscheinlichkeit, dass das Objekt authentisch ist, jeder entdeckte Fehler oder jede fehlende Angabe lässt sie sinken. Je mehr Informationen über das Objekt verifiziert werden können, desto grösser ist die Wahrscheinlichkeit seiner Authentizität. Insbesondere Informationen aus dem Internet sollten kritisch und skeptisch auf ihren Wahrheitsgehalt hin überprüft werden, wie John Fritch und Robert Cromwell ausführen:

*"It is better to be overly skeptical of the authority of Internet information than to be overly naive and trusting of Internet information. The process of ascribing cognitive authority should take the form of disproving initial skepticism, thus ensuring that proper evaluation occurs."*⁹²⁹

Die Überprüfung eines digitalen Objekts kann als Erarbeitung eines subjektiven Näherungswerts aufgefasst werden. Dieser drückt den Wahrscheinlichkeitsgrad aus, für wie authentisch das Objekt vom Urteilenden eingeschätzt wird.

Lösungsansatz Authentizitätsapproximation

Wissenschaftler schätzen aufgrund ihrer Erfahrung viele Objekte auf ihre Authentizität und Verwendbarkeit als Quelle ein, sie erarbeiten quasi intuitiv einen Näherungswert. Dieser Prozess könnte durch eine mathematische Berechnung unterstützt werden, die zu einer strukturierten und für Dritte nachvollziehbaren Beurteilung verhilft. Das Verfahren wird idealerweise bei allen verwendeten Quellen angewendet, aber in vielen Fällen ist der Quellenkorpus zu umfangreich, um alle mit diesem Verfahren zu beurteilen. Bei schwierig zu beurteilenden Quellen oder für Schlüsselquellen wäre die transparente Berechnung jedoch hilfreich, um die Beurteilung der Authentizität dieser Quelle(n) nachvollziehbar darzustellen.

Das Ergebnis der mathematischen Berechnung drückt in einer Prozentzahl die subjektive Einschätzung des Nutzers über die Authentizität eines bestimmten digitalen oder analogen Objekts aus. Diesen Näherungswert bezeichne ich als *Authentizitätsapproximation*.

Mit der *Authentizitätsapproximation* wird ein subjektiver Näherungswert in Prozenten angegeben, der die Wahrscheinlichkeit der Authentizität eines Objekts nach dessen Verifikation angibt.

Merkmale

Da die Erfassung und Bewertung eines jeden Merkmals einen unverhältnismässigen Arbeitsaufwand bedeuten würde, schlage ich die Verwendung von sieben Fragekomplexen vor, die aus der Quellenkritik bekannt sind. Die vielfältigen Probleme können in diesen subsumiert werden:⁹³⁰

⁹²⁹ Fritch; Cromwell: Evaluating Internet resources, 2001, S. 501.

⁹³⁰ Die Anzahl der Fragekomplexe basiert auf der Miller'schen Zahl, siehe: Miller, George A.: The Magical Number Seven, Cogprints, o. J., URL: <http://cogprints.org/730/1/miller.html>, Stand: 23.08.2015, Original in: The Psychological Review, 1956, Bd. 63, S. 81-97.

Fragekomplex	Inhalt / Merkmale
Integrität	Bitcode, Information, Darstellung, Nutzung; Sicherheitsmerkmale; informationstechnische Angaben
Persistenz	Publikations-/Speicherort (Institution, URL/URI/...); Langzeitverfügbarkeit (Datenformat etc.)
Datierung	Publikations-/Änderungsdatum, Überlieferungsgeschichte
Autorschaft	Bekanntheit (anonym, Pseudonym, Softbot), Intention
Adressat	Bekanntheit, Verhältnis zum Autor
Inhalt	inhaltliche Hinweise auf Authentizität
Beziehungen	Verweise, Hinweise, Hyperlinks

Tabelle 6: Beurteilungsmerkmale Authentizitätsapproximation (Föhr)

Berechnung

Zur Berechnung der Authentizitätsapproximation werden für die obigen Merkmale jeweils zwei Werte erarbeitet. Der erste Wert ist das Resultat der Untersuchung der vorhandenen Informationen über und aus dem Objekt. Der Nutzer beziffert subjektiv mit einem prozentualen Wert, für wie plausibel und authentisch er diese Informationen einschätzt. Mit dem zweiten Wert drückt der Nutzer in einer vordefinierten Skala aus, für wie wichtig er die Information in Bezug auf die Fragestellung hält (Gewichtung). Je kleiner der Prozentwert, desto weniger authentisch und plausibel ist die Information und je kleiner der Skalenwert, desto weniger wichtig ist das Kriterium.

Die Authentizitätsapproximation wird durch die Addition der Resultate aus den Multiplikationen der einzelnen Authentizitätswerte mit ihrer Gewichtung berechnet. Diese Summe wird durch die Summe der Gewichtungen dividiert. Der Wert der Authentizitätsapproximation bewegt sich zwischen 0 (keinesfalls authentisch) und 100 (absolut authentisch). Weil es sich bei der Authentizitätsapproximation um eine subjektive Bewertung handelt, sollte die Bewertung mit entsprechenden Bemerkungen begründet werden. Eine Tabelle erleichtert das Festhalten der Hinweise zu den Inhalten und Merkmalen sowie die Berechnung des Näherungswerts.

Abkürzungen:

- A: Authentizität (zwischen 0 und 100%; je höher der Wert, desto authentischer)
- G: Gewichtung (zwischen definierten Werten; je höher der Wert, desto wichtiger)
- M: Multiplikationswert ($A \times G$)

Fragekomplex	Bemerkungen	A	G	M
Integrität	• ...	XX	X	...
Persistenz	• ...			
Datierung				
Autorschaft				
Adressat				
Inhalt				
Beziehungen				
Summe (Werte aller Merkmale)				
Authentizitätsapproximation (Summe M geteilt durch Summe G), gerundet		%		

Tabelle 7: Bewertungstabelle (Föhr)

Beispiele

Anhand von drei Beispielen soll die Berechnung der Authentizitätsapproximation verdeutlicht werden.

Beispiel 1: PDF-Dokument in einem Repository

- Dokument: 'Turner et al.: White Paper: The Digital Universe of Opportunities'⁹³¹
- Gewichtung für die Fragestellung dieser Arbeit wird in einer Skala von 1 (unwichtig) bis 5 (sehr wichtig) bewertet

Fragekomplex	Bemerkungen	A	G	M
Integrität	<ul style="list-style-type: none"> • Keine Integritätsmerkmale vorhanden • Inhalte online (URL: http://www.emc.com/leadership/digital-universe/index.htm) und im Dokument stimmen überein • emc.com: vertrauenswürdig (whois) • freier Zugriff • DRM nicht erkennbar 	30	5	150
Persistenz	<ul style="list-style-type: none"> • URL kaum persistent • private Unternehmung • PDF 	60	4	240

⁹³¹ Turner u. a.: The Digital Universe of Opportunities, 2014.

Fragekomplex	Bemerkungen	A	G	M
Datierung	<ul style="list-style-type: none"> • April 2014 • Unter http://www.emc.com/leadership/digital-universe/index.htm kann zugehöriger Inhalt aufgerufen werden, Datum stimmt überein 	90	3	270
Autorschaft	<ul style="list-style-type: none"> • 4 Autoren, Adresse der Institution vorhanden • Autoren haben auf dem Webauftritt von idc ein 'analyst profile' • idc.com: vertrauenswürdig (whois) • Titel: Turner (Senior Vice President of Enterprise Systems and IDC Fellow for The Internet of Things), Gantz (Senior Vice President), Reinsel (Group Vice President, Storage, Semiconductors, GRC, and Pricing), Minton (vice president) – Ghostwriter (Mitarbeiter) möglich, aber Mutmassung 	100	4	400
Adressat	<ul style="list-style-type: none"> • technikaffine, interessierte Personen 	100	3	300
Inhalt	<ul style="list-style-type: none"> • Herkunft von der idc/emc plausibel • technischer Schreibstil 	90	2	180
Beziehungen	<ul style="list-style-type: none"> • Multimedialer Inhalt unter http://www.emc.com/leadership/digital-universe/index.htm; Inhalte stimmen überein 	80	2	160
Summe (Werte aller Merkmale)			23	1700
Authentizitätsapproximation (Summe M geteilt durch Summe G), gerundet		73.91%		

Tabelle 8: Authentizitätsapproximation 'Turner et al.: White Paper' (Föhr)

Eine Authentizitätsapproximation von 73.91% zeigt auf, dass das Objekt nicht sehr vertrauenswürdig ist, weil mehr als ein Viertel der authentifizierenden Merkmale fehlen. Insbesondere die tiefe Beurteilung der Integrität, verbunden mit hoher Gewichtung führt zu einem tiefen Approximationswert. Wären Sicherheitsmerkmale wie der Hashwert oder die digitale Signatur im Dokument implementiert, hätte die Bewertung auf 90% erhöht werden können und der Gesamtwert wäre auf 86.96% gestiegen. Da die Integrität bei digitalen Objekten aufgrund ihrer Volatilität und Manipulationsfähigkeit meistens sehr hoch gewichtet wird, haben vorhandene Sicherheitsmerkmale eine starke Auswirkung auf das Vertrauen und damit auch auf das Ergebnis der Authentizitätsapproximation.

Beispiel 2: Tweet



Abbildung 27: Tweet von David M. Berry (Föhr)

- Tweet: David M. Berry, @berrydm
- Gesichert am 07.10.2015, 11.02 Uhr als Bilddatei in JPG
- Gewichtung für die Fragestellung dieser Arbeit wird in einer Skala von 1 (unwichtig) bis 5 (sehr wichtig) bewertet
- Abkürzungen gemäss erstem Beispiel

Fragekomplex	Hinweis	A	G	M
Integrität	<ul style="list-style-type: none"> • Keine Integritätsmerkmale, Angaben nur von Twitter • Tweets nach Versand grundsätzlich nicht veränderbar • twitter.com: vertrauenswürdig (whois) • freier Zugriff • DRM nicht erkennbar 	80	5	400
Persistenz	<ul style="list-style-type: none"> • Im Archiv von Twitter verfügbar • private Unternehmung • Tweet bei Suche im Archiv nicht wiedergefunden 	60	1	60
Datierung	<ul style="list-style-type: none"> • 07.10.2015, 03.10 Uhr • Jeder Tweet wird vom Dienst automatisch mit dem Publikationsdatum versehen. 	100	4	400
Autorschaft	<ul style="list-style-type: none"> • David Berry ist bekannt als 'digital humanist' – plausibel • Herkunft von Berry plausibel (Bild stimmt mit Person überein – persönlich getroffen) • Berry kennt pfhist persönlich – plausibel 	100	4	400
Adressat	<ul style="list-style-type: none"> • Twitterer direkt (pfhist) – plausibel, da persönlich bekannt • interessierte Personen (Follower von berrydm und am Thema Interessierte postdigital) – Thema plausibel • Twitternutzer weltweit 	100	1	100
Inhalt	<ul style="list-style-type: none"> • kritische Aussage von Berry plausibel • Unter der URL befindet sich eine nicht direkt mit dem Thema verwandte Webseite (falsch kopiert?): https://www.zotero.org/support/quick_start_guide 	70	3	210

Fragekomplex	Hinweis	A	G	M
Beziehungen	<ul style="list-style-type: none"> • Verwendung eines URL-Shorteners: selten von Berry genutzt • Beziehungen zu anderen Nutzern (Retweets) plausibel, Berry hat über dreitausend Follower • verlinkte URL verdächtig 	70	1	70
Summe (Werte aller Merkmale)			19	1640
Authentizitätsapproximation (Summe M geteilt Summe G), gerundet		86.32%		

Tabelle 9: Authentizitätsapproximation 'Tweet von David M. Berry' (Föhr)

Eine Authentizitätsapproximation von 86.32% zeigt auf, dass das Objekt vertrauenswürdig ist, weil nur etwa ein Siebtel der das Objekt authentifizierenden Merkmale fehlt. Die vom Dienst automatisch eingefügten und von Nutzern nicht mehr veränderbaren Angaben zu Autor, Datierung und Adressaten tragen zur guten Bewertung bei. Der Tweet kann als vertrauenswürdige Quelle verwendet werden.

Allerdings: der Tweet von David Berry wurde durch Pascal Föhr mit dem Dienst tweetfake⁹³² am 07.10.2015, um 10.58 Uhr erstellt und mit dem Logo von Twitter ergänzt. Somit ist der Tweet und die Darstellung gefälscht (folglich ist der Urheber des Tweets Pascal Föhr und nicht David Berry). Wäre dies vor der Berechnung bekannt gewesen, wären die A-Werte von Datierung, Autorschaft und Inhalt auf 0 gesunken und auch die anderen Werte würden nicht mehr so hoch ausfallen.

⁹³² URL: <http://tweetfake.com/>.

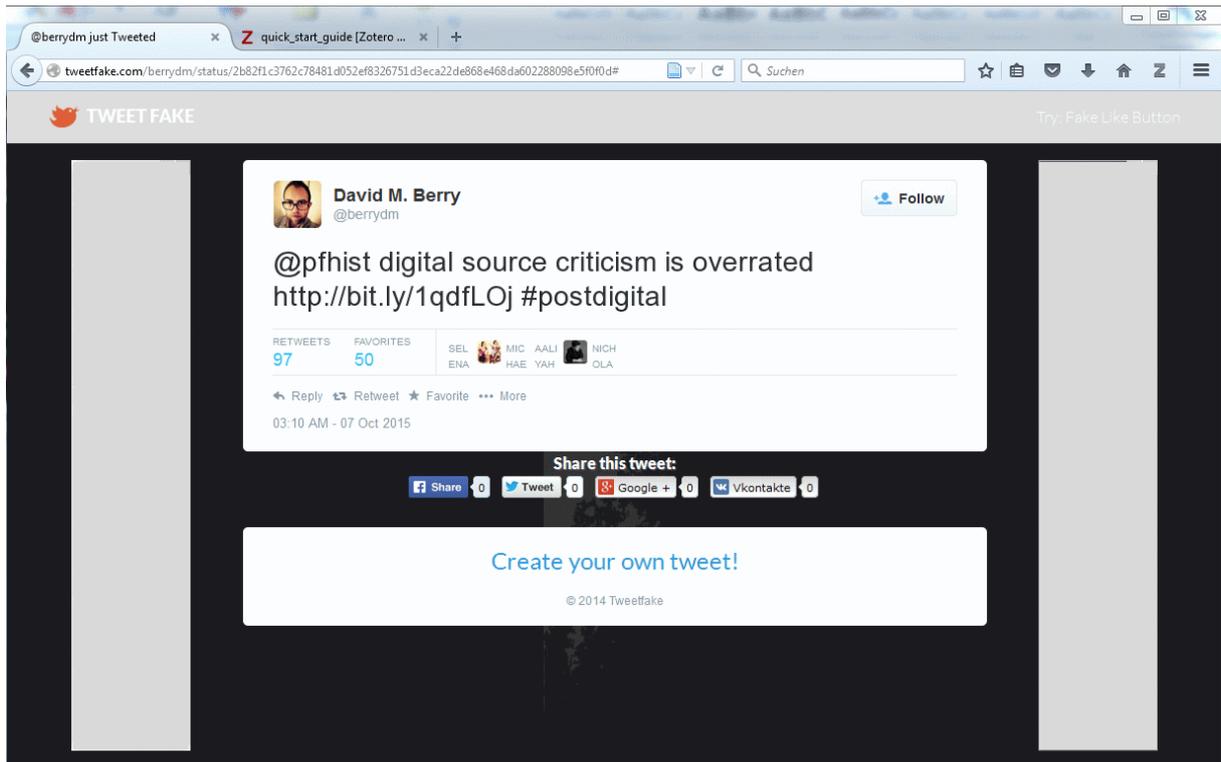


Abbildung 28: Tweet von David M. Berry, mit tweetfake erstellt (URL: <http://tweetfake.com/>)⁹³³

Beispiel 3: Bild aus der Europeana



Abbildung 29: Aerial picture of Princes Street, Edinburgh (Europeana)

⁹³³ Aus rechtlichen Gründen sind die Werbeanzeigen eingegraut.

- Bild: Aerial picture of Princes Street, Edinburgh⁹³⁴
- Archiv: Europeana
- Gewichtung zur Fragestellung: 'Wandel des Stadtbildes von Edinburgh 1900-1930'
- Abkürzungen gemäss erstem Beispiel

Fragekomplex	Bemerkungen	A	G	M
Integrität	<ul style="list-style-type: none"> • keine Integritätsmerkmale vorhanden • Titel entspricht verknüpftem Bild • Digitalisat • europeana.eu: vertrauenswürdig (whois: eurid.eu) <ul style="list-style-type: none"> • Institution vertrauenswürdig (Netzwerk von Gedächtnisinstitutionen und mitfinanziert von der EU; https://www.europeana.eu/portal/de/about.html) • freier Zugriff • DRM nicht erkennbar 	90	5	450
Persistenz	<ul style="list-style-type: none"> • Institution ist für Persistenz besorgt, auf langzeitliche Archivierung ausgerichtet • URL: wahrscheinlich • JPG-File – kein langzeitarchivwürdiges Format⁹³⁵ 	90	3	270
Datierung	<ul style="list-style-type: none"> • 13.04.14 – handdatiert, nicht Datum der Bildaufnahme – aus Metadaten und aus der wahrscheinlich zum Bild gehörigen Rückseite in der Kollektion⁹³⁶ • Bild älter als 14 Jahre unwahrscheinlich 	80	5	400
Autorschaft	<ul style="list-style-type: none"> • Photograph unbekannt • Absender: Harry Stanley Green (Metadaten, Rückseite) • Beitragende: Annelies Dogterom (Metadaten), unbekannt 	70	1	70
Adressat	<ul style="list-style-type: none"> • Muriel Green, Ehefrau (Metadaten) 	100	1	100
Inhalt	<ul style="list-style-type: none"> • Vorderseite einer Postkarte (Rückseite verlinkt) 	100	5	500
Beziehungen	<ul style="list-style-type: none"> • Teil einer Kollektion (verlinkt) aus der Korrespondenz zwischen dem Ehepaar Green während des 1. Weltkriegs – plausibel 	90	2	180
Summe (Werte aller Merkmale)			22	1970
Authentizitätsapproximation (Summe M geteilt Summe G), gerundet		89.55%		

Tabelle 10: Authentizitätsapproximation 'Europeana: Aerial picture of Princes Street, Edinburgh' (Föhr)

Eine Authentizitätsapproximation von 89.55% zeigt auf, dass das Objekt vertrauenswürdig ist, weil nur noch etwas mehr als ein Zehntel der das Objekt authentifizierenden Merkmale

⁹³⁴ Bildzitat Europeana (URL: <https://www.europeana.eu/portal/de>); Sammlung: Green, Muriel: Harry Stanley Green's letters and medals, URL: https://www.europeana.eu/portal/en/record/2020601/contributions_17034.html#&gid=1&pid=21 Stand: 06.05.2018. CC-Lizenz.

⁹³⁵ Nach KOST. Siehe URL: <http://www.kost-ceco.ch/wiki/whelp/KaD>.

⁹³⁶ URL: https://www.europeana.eu/portal/en/record/2020601/contributions_17034.html#&gid=1&pid=22.

fehlen. Besonders die vertrauenswürdige Institution, die auch für die dauerhafte Archivierung besorgt ist und wichtige Metadaten zur Verfügung stellt, wirkt sich positiv auf den Wert aus.

Interpretation

Die Interpretation des Authentizitätsapproximationswerts ist von der Ausgangslage und der Fragestellung abhängig. Für ein digitales Objekt, das sich als Publikation einer renommierten Institution ausgibt (siehe das Beispiel oben), scheint die Anforderung an einen Wert von mindestens 80% gerechtfertigt zu sein, um es als vertrauenswürdig bezeichnen zu können. Bei einem physischen Grabstein hat der Beurteilende die Anforderungen aufgrund der Ausgangslage sowie seiner Erfahrungen und mit entsprechender Begründung auf das Objekt anzupassen. Dadurch können Quellen trotz sehr unterschiedlichen Voraussetzungen eine ähnlich hohe Authentizitätsapproximation aufweisen und als authentisch gelten respektive kann bereits ein niedriger Wert vom Forschenden als authentisch bezeichnet werden.

Wie das Beispiel des angeblich von David Berry verfassten Tweets zeigt, kann eine als vertrauenswürdig bezeichnete Quelle trotzdem gefälscht sein. Die falsch interpretierte oder vernachlässigte Information, dass der Tweet im Archiv nicht wiederaufgefunden werden konnte, erhält erst mit der Kenntnis seiner Fälschung die entsprechende Gewichtung. Um den Tweet als Fälschung zu entlarven, sind weitere Informationen nötig, wie beispielsweise die persönliche Aussage von Berry, er habe nie so etwas geschrieben. Vielleicht lässt sich aber der Ersteller des falschen Tweets eruieren, hier Pascal Föhr, und dessen Intention feststellen, hier das Problem einer gefälschten Quelle aufzeigen. Die Quelle ist unter diesen Voraussetzungen neu zu beurteilen und ist (vermutlich) wieder als authentisch zu bezeichnen.

Software zur Erfassung

Die Authentizitätsapproximation kann sehr einfach mit einer digitalen Tabelle berechnet werden, die mit dem jeweilig zu untersuchenden Objekt verknüpft wird. Benutzerfreundlich wäre ein eigenständiges Programm oder/und ein Add-on für Literaturverwaltungsprogramme, das die Erarbeitung der Authentizitätsapproximation unterstützt. Der Nutzer könnte darin die Fragekomplexe definieren, Bemerkungen zur Bewertung festhalten, die Authentizität mit einer Prozentangabe bewerten und die Gewichtung hinterlegen. Bekannte Faktoren könnten mit Default-Werten versehen werden, die automatisch eingefügt werden. Beispielsweise wäre eine White- und Blacklist für bekannte URL hilfreich. Die Berechnung würde ebenfalls automatisch erfolgen und der entsprechende Prozentwert würde den Wahrscheinlichkeitsgrad der Authentizität einer verwendeten digitalen oder analogen Quelle mit einem Näherungswert ausdrücken. Je höher der Wert läge, desto authentischer wäre die Quelle

vom Nutzer beurteilt worden. Die Bewertung der Authentizitätsapproximation sollte schliesslich exportiert und in einem Langzeitarchiv gesichert werden können.

Weiterer Ansatz

Einen vergleichbaren Ansatz wie die Authentizitätsapproximation (AP) verfolgt das von Xin Luna Dong postulierte Modell zur Bewertung von Vertrauenswürdigkeit und Qualität von (Web-)Inhalten.⁹³⁷ Im Unterschied zur AP berechnet Dong den Wert ausschliesslich algorithmisch und unabhängig von einer bestimmten Ausgangslage, wohingegen die AP mit subjektiven Werten arbeitet, die je nach Ausgangslage respektive Forschungsfrage unterschiedlich sein können. Da in der Geschichtswissenschaft die potenziellen Quellen abhängig von der Ausgangslage flexibel sollten bewertet werden können, stellt das hier vorgeschlagene Berechnungsmodell der Authentizitätsapproximation eine geeignetere Methode dar.

5.4.14. Zusammenfassung (Fragen und Thesen)

Die kritische Untersuchung eines digitalen Objekts basiert auf den Fragen der traditionellen Methode, deren vollständige Beantwortung nur unter Einbezug von informationstechnischen Untersuchungsmethoden möglich ist. In den vorangegangenen Ausführungen habe ich zuerst die Konzepte der Authentizität, der Integrität und der Persistenz erläutert und anschliessend Fragen nach der Objektbeschreibung, der Echtheit, der Datierung, dem Entstehungsort, der Autorschaft, dem Adressaten, der Darstellung, dem Inhalt und den Beziehungen zu anderen Objekten untersucht. Am Schluss habe ich aufgezeigt, wie mit dem Verfahren der Authentizitätsapproximation die Beurteilung einer Quelle transparent und nachvollziehbar erarbeitet und präsentiert werden kann.

Authentizität bedeutet im historischen und im informationstechnischen Kontext unterschiedliches. Die Geschichtswissenschaft verwendet den Begriff zur Beurteilung der Qualität des Objekts. Der Begriff Authentizität bezeichnet in der Geschichtswissenschaft präzise, aber gleichzeitig unterschiedliche Dinge in unterschiedlichem Kontext und in verschiedenen Gemeinschaften. Er kann *Original* bedeuten, aber auch *originalgetreu*; das Untersuchungsobjekt kann beschädigt oder unbeschädigt sein, von bekannter oder unbekannter Herkunft stammen. In der Informatik hingegen wird ein Objekt mit eindeutigen Identifikationsmerkmalen wie einem Loginnamen verglichen.

Die historische *Integrität* eines Objekts ist gewahrt, wenn es materiell und/oder inhaltlich unbeschädigt sowie unversehrt und unverändert ist. Die Feststellung der Integrität kann aufwendig oder sogar unmöglich sein. Demgegenüber kann die informationstechnische Inte-

⁹³⁷ Dong u. a.: Knowledge-Based Trust, 2015. Siehe auch Kap. 5.4.11. Inhalt.

gritat des digitalen Objekts durch verschiedene Verfahren wie dem Hashwert oder der digitalen Signatur relativ einfach uberpruft werden, weil sich die Unversehrtheit des Bitcodes und die semantische Korrektheit automatisch uberpruft werden lasst. Unterschieden werden konnen Informations-, Darstellungs- und Nutzungsintegritat. Zusatzlich sollte die Integritat der verwendeten elektronischen Informationssysteme uberpruft werden, was jedoch kaum moglich ist. Verlust von Integritat bedeutet einen Informations- und/oder Datenverlust. Dieser Verlust ist je nach Art der Fragestellung unter Umstanden vernachlassigbar. Digitale Objekte sollten fur die wissenschaftliche Verwendung mit einer *Metadata Information Section (MIS)* erganzt und mit einer digitalen Signatur beglaubigt werden.

Bei digitalen Objekten ist zwischen der *Persistenz* von volatilen Objekten und der *Adresspersistenz* zum Wiederauffinden von Objekten zu unterscheiden. Sie bedingen sich gegenseitig nicht. Um beide Persistenzen zu gewahrleisten, sollten digitale Forschungsressourcen in einem *Trusted Digital Repository TDR* mit objektspezifischer, persistenter Adressierung gesichert werden.

Die informationstechnische *Objektbeschreibung* bildet die notige Voraussetzung fur die Darstellung eines digitalen Objekts durch ein elektronisches Informationssystem. Erst wenn dieses dargestellt wird, konnen aus dem Objekt und/oder aus den zugehorigen Metadaten weitere beschreibende Elemente wie Datierung, Uberlieferungsgeschichte, Quellengattung oder allenfalls der Erhaltungszustand des Datentragers ausgelesen werden.

Die *Echtheit* eines singularen digitalen Objekts kann nur anhand der im Objekt enthaltenen Uberlieferungsgeschichte und/oder unter Zuhilfenahme von dessen Metadaten uberpruft werden. Falls ein Objekt in mehreren Varianten vorliegt, kann die Uberlieferungsgeschichte eventuell mit einem diakritischen Verfahren nachvollzogen werden. Bei archivierten Objekten stehen – durch die per Konvention vertrauenswurdige Institution des Archivs – die verlasslichsten Informationen zur Verfugung. Die Uberlieferungsgeschichte eines digitalen Objekts wird anhand der vorhandenen Metadaten, der Vertrauenswurdigkeit der das Objekt archivierenden Institution und dem Verhaltnis des Objekts zu dieser Institution beurteilt.

Auch die Glaubwurdigkeit der *Datierung* eines digitalen Objekts ist von der Vertrauenswurdigkeit der publizierenden Institution abhangig, wenn keine Sicherheitsmerkmale wie Zeitstempel in das Objekt integriert wurden. Eine Ruckfuhrung eines digitalen Objekts auf seinen Archetypus kann nur erfolgen, wenn vertrauenswurdige Institutionen die Prozesse zur Langzeitarchivierung korrekt einhalten und zu jedem Objekt eine vollstandige Uberlieferungsgeschichte zur Verfugung stellen. Fur eine eindeutige Datierung von wissenschaftlichen Publikationen sollten *qualifizierte elektronische Signaturen mit Zeitstempel* verwendet werden.

Der *Entstehungs- / Erscheinungsort* ist vom Speicherort des digitalen Objekts zu unterscheiden. Zwar kann nach bestehender Konvention der juristische Sitz der publizierenden Institution als Erscheinungsort verwendet werden, aber bei digitalen Objekten ist er irrelevant. Besser wäre es, bei der Verwendung von digitalen Objekten als Forschungsressource grundsätzlich auf die Angabe ihres Erscheinungsorts zu verzichten und nur den zugehörigen *persistenten Identifikator* anzugeben, der um die URL des gesetzlich beauftragten TDR ergänzt werden kann.

Bei der *Autorschaft* ist zwischen natürlichen Personen und Algorithmen respektive Softwares mit künstlicher Intelligenz zu unterscheiden. Anonymität und Pseudonyme sollten aufgelöst werden. Mit einer *fortgeschrittenen digitalen Signatur* kann der Autor eindeutig identifiziert werden. Auf die Intention und das Motiv des Autors können verschiedene Merkmale hinweisen, unter anderem die Verwendung von DRM-Verfahren und Pseudonymen, die Wahrung der Anonymität, das Editieren der eigenen Publikation oder gesetzte und nicht gesetzte Hyperlinks.

Die eingesetzte Technik bestimmt das *Verhältnis zwischen Autor und Adressat* mit. Der Adressat erhält durch direkt sichtbare Ergänzungen eines digitalen Objekts mit Annotationen einen argumentativen Machtzuwachs. Er kann zudem mit dem Autor eine manuskriptive Kollaboration eingehen. Jedoch ist zu hinterfragen, ob die gewünschten Adressaten aufgrund des sogenannten Digital Divide überhaupt erreicht werden.

Jedes digitale Objekt soll hinsichtlich seiner intendierten, korrekten *Darstellung* kontrolliert werden, insbesondere wenn Hilfsdarstellungen wie eine Emulation oder Sicherheits- und Schutzmechanismen verwendet werden. Es ist zudem zu untersuchen, ob es sich um eine Inszenierung handelt, die zu einer ahistorischen Rezeption führen kann.

Wird das digitale Objekt korrekt dargestellt, so kann sein *Inhalt* in einem ersten Schritt mit traditionellen Methoden und Techniken überprüft werden. Da das Objekt gleichzeitig in mehreren Ebenen vorliegt, sind diese zusätzlich auf inhaltliche Angaben zu prüfen. Technische Massnahmen wie eine zeitverzögerte Übertragung können zudem Unmittelbarkeit vortäuschen. Inhalte im Web können domain-, orts- oder browserspezifisch generiert werden. Hyperlinks ergänzen das digitale Objekt mit Hinweisen, die der Nutzer lesen kann, und ergänzenden Inhalten, die der Nutzer lesen muss. Objekte im Hypertext werden durch nutzerspezifische Leseprozesse zu anderen Objekten; mit entsprechenden Auswirkungen auf den Inhalt.

Vor allem Hyperlinks setzen digitale Objekte zueinander in *Beziehung*, wobei die im Objekt verwendeten Hyperlinks Hinweise auf die Intention des Autors geben. Backlinks von URLs mit hoher Reputation liefern Hinweise zur Beurteilung der Vertrauenswürdigkeit eines Objekts. Die im Objekt enthaltenen Hyperlinks sind qualitativ zu analysieren. Der *Link-Check*

beinhaltet sowohl eine inhaltliche Überprüfung als auch eine Beurteilung darüber, mit welcher Absicht Hyperlinks gesetzt oder nicht gesetzt wurden. Für Quellenangaben sollte ein *Link-Check Element* mit *Link-Check-Ziffer* ergänzt werden, mit dem der Nutzer die Art und Tiefe der Hyperlinküberprüfung offenlegt.

Für die digitale Langzeitarchivierung stellt das Verfahren der digitalen Signatur eine Herausforderung dar, denn einerseits sind Signaturen nur eine bestimmte Zeit gültig und andererseits verletzen Signaturen die Integrität eines Objekts und können zu Darstellungsproblemen führen. Da eine digitale Signatur die Integrität des Objekts beweisen kann, eine fortgeschrittene und qualifizierte Signatur auch die Identität des Ausstellers belegt und mit einem Zeitstempel auch eine Datierung ermöglichen kann, stellt dieses Verfahren für wissenschaftliche Zwecke die beste technische Lösung zur Authentifizierung eines Objekts dar.

Die quellenkritische Untersuchung eines digitalen Objekts erfordert neue Kompetenzen, Techniken und Methoden. Im Gegensatz zu physischen müssen digitale Objekte zuerst dargestellt werden können, wofür eine vollständige Objektbeschreibung und ein spezifisches Informationssystem vorhanden sein muss. Neu ist, dass die Authentizität eines Objekts mit einer informationstechnischen Überprüfung eindeutig bestimmt werden kann. Digitale Objekte sind volatil und auch die Adressierung ist bisher in den wenigsten Fällen persistent, wohingegen physische Objekte in den meisten Fällen als persistent angesehen werden können. Problematisch sind bei digitalen Objekten auch deren Datierung sowie die Feststellung des Entstehungs- und/oder Erscheinungsortes und der Autorschaft, weil Kontrollinstanzen und verlässliche Metadaten oft fehlen. Als neuer Typ von Autorschaft treten nun auch Softwares auf, die digitale Objekte autonom erstellen. Der Autor erhält zudem neue Funktionen und einen Machtzuwachs, denn er kann sein Objekt selbständig publizieren sowie danach sein Objekt mit geringem zeitlichen, finanziellen und organisatorischen Aufwand editieren und erneut publizieren. Mit in das Objekt integrierten Schutzmechanismen kann er sehr spezifische Nutzungsformen bestimmen, was bei physischen nicht möglich ist. Digitale Objekte können mit Hyperlinks viel enger mit anderen, insbesondere externen Elementen verknüpft werden, die der Autor für das Verständnis voraussetzt. Bei physischen Objekten ist eine solch tiefe Verschachtelung respektive Verknüpfung der Inhalte (meist) weder praktikabel noch sinnvoll.

Die Quellenkritik eines physischen oder digitalen Objekts basiert hauptsächlich auf dem (Vor-)Wissen und den Erfahrungen des Untersuchenden. Die vorhandenen Informationen werden intuitiv bezüglich ihrer Authentizität beurteilt und oft wird die Quelle nur dann weiter untersucht, wenn nicht genügend vertrauenswürdige Informationen zur Verfügung stehen. Ein Objekt kann in unterschiedlichem Masse authentisch sein, abhängig von der Fragestellung, von den zur Verfügung stehenden Informationen, von Sicherheitsmerkmalen und auch von den medien- und informationstechnischen Kompetenzen des Nutzers. Es ist jedoch

unerlässlich, dass die Authentizität einer Quelle beurteilt wird. Ob dies intuitiv oder mit dem transparenten und nachvollziehbaren Verfahren der Authentizitätsapproximation erfolgt, ist unerheblich. Damit ist die fünfte These bestätigt:

T5: Die quellenkritische Überprüfung eines digitalen Objekts bedeutet die Erarbeitung eines Näherungswerts für die Wahrscheinlichkeit der Authentizität des Objekts.

5.5. Interpretation

Nach Simon Margulies unterscheidet sich die inhaltliche Interpretation digitaler Objekte aufgrund ihrer analogen Repräsentation nicht von jener anderer Quellen.⁹³⁸ Die traditionelle Methodik würde demnach genügen. Dem entgegen stehen die vielfältigen medientheoretischen und philosophischen Diskussionen, in denen digitale Informationssysteme als wesentlicher und teilweise untrennbarer Teil der sozio-kulturellen Entwicklung des heutigen Menschen zu bezeichnen sind.⁹³⁹ Daraus lässt sich schliessen, dass die Interpretation des Inhalts eines digitalen Objekts durch digitale Informationssysteme beeinflusst ist. Von besonderem Interesse sind der Digitaldatentyp und das Datenformat eines digitalen Objekts, die eingesetzten Informationssysteme, und die Hypermedialität. Deshalb stimme ich Margulies nur insofern zu, als dass bei der inhaltlichen Interpretation eines digitalen Objekts auch bisher bekannte, 'analoge' Methoden anzuwenden sind.

Der Historiker Leif Scheuermann meint zwar, dass die digitale Hermeneutik ausschliesslich innerhalb von Digitalmedien vollzogen wird. Dies ist nicht korrekt, denn die Repräsentation der Ergebnisse digitaler Prozesse ist analog und wird (oder 'kann nur') vom Forschenden in dieser Form rezipiert werden. Er interpretiert diese Ergebnisse abschliessend. Ein digitales Informationssystem kann nichts verstehen oder erklären – von semantischen Systemen, die Wissen selbst generieren können, sind wir noch weit entfernt. Technische und menschliche Verfahren der Interpretation greifen ineinander, weil der Forschende nur mit einem digitalen Informationssystem überhaupt digitale Objekte bearbeiten kann.⁹⁴⁰ Die Interpretation digitaler Objekte erfolgt hauptsächlich mit informationstechnischen Mitteln, aber nicht ausschliesslich.

Für die Interpretation digitaler Quellen werden drei Schritte vorgestellt, die allerdings je nach Art der Quelle ineinander verwoben sein und nicht scharf abgegrenzt werden können, wie

⁹³⁸ Nach: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 324. Margulies bezieht sich in seiner Argumentation auf eine Aussage von Jan Hodel, die dieser so nicht gemacht hat. Siehe: Hodel: Historische Online-Kompetenz. Überlegungen zu einem hybriden Kompetenzmodell, 2005, S. 153-155.

⁹³⁹ Vgl. bspw. div. Beiträge in: Kelly, Matthew; Bielby, Jared (Hg.): Information Cultures in the Digital Age: A Festschrift in Honor of Rafael Capurro, Wiesbaden 2016, insb. Kap. I (Culture and Philosophy of Information).

⁹⁴⁰ Vgl. dazu auch: Burkhardt: Digitale Datenbanken, 2015, S. 74.

folgend gezeigt wird. Die Schritte folgen dem Schema Objekt (digitales oder analoges Objekt) und Subjekt (der subjektiv interpretierende Forschende) und – für digitale Objekte – ein geeignetes digitales Informationssystem, das zwischen Objekt und Subjekt vermittelt.

Im ersten Schritt der Interpretation werden die zu untersuchenden digitalen Objekte in einer digitalen Forschungsumgebung zusammengefasst, geordnet und aufgearbeitet, was an sich schon einen hermeneutischen Prozess darstellen kann.⁹⁴¹ Dies ist beispielsweise dann der Fall, wenn der Forschende Daten unterschiedlicher Herkunft auswählt und in eine Datenbank aufnimmt, bestimmte Daten miteinander verknüpft oder Objekte kennzeichnet (bspw. durch Tagging), wobei bereits eine interpretatorische Leistung erbracht wird. Daten sinnvoll zu ordnen und sie als Quellen zu nutzen, ist eine grosse Herausforderung für Historiker in Anbetracht der unzähligen Datenformen und -formaten.⁹⁴²

Sachverhalte auf einer digitalen Karte zu visualisieren und diese dann in herkömmlicher Form für seine Arbeit an einem Text zu nutzen, stellt hingegen keinen hermeneutischen Prozess dar, sondern ist nur eine computerunterstützte Darstellung von Daten.⁹⁴³ Deshalb kritisiert Scheuermann Projekte wie die *Venice Time Machine*⁹⁴⁴, weil sie 'nur' die Rekonstruktion und Visualisierung der Vergangenheit zum Ziel hat und keinen digital hermeneutischen Prozess durchläuft. Er übersieht jedoch, dass eine solche Visualisierung zwar nicht das Ergebnis eines hermeneutischen Prozesses ist, aber ein Hilfsmittel sein kann, um neue hermeneutische Prozesse zu ermöglichen. Scheuermann irrt jedoch, wenn er behauptet, dass eine Datenbankabfrage keinen heuristischen Prozess auslöst.⁹⁴⁵ Marcus Burkhardt hat dargestellt, dass Datenbanken das Potenzial haben, die in ihr gespeicherten Informationen zu neuartigen und bisher unbekanntem Ressourcen zu transformieren.⁹⁴⁶

Aus diesen Ausführungen ist deutlich geworden, dass in der digitalen Hermeneutik nicht nur das Ergebnis, sondern auch dessen Entstehung von Bedeutung ist:

"Durch die Integration und freie Kombination unterschiedlichster Daten und Anwendungen kann im digitalen Medium ein hermeneutischer Prozess stattfinden, der nicht mehr implizit im Ergebnis definiert, sondern selbst Teil des Ergebnisses ist."⁹⁴⁷

Deshalb ist es auch immanent wichtig, dass in einem zweiten Schritt nicht nur der zum Ergebnis führende, hermeneutische Prozess mit den verwendeten Algorithmen⁹⁴⁸ nachvoll-

⁹⁴¹ Vgl.: Scheuermann, Leif: Die Abgrenzung der digitalen Geisteswissenschaften, in: Digital Classics Online 2 (1), 18.02.2016, S. 60-62.

⁹⁴² Siehe auch: Gibbs; Owens: The Hermeneutics of Data, 2013, S. 160.

⁹⁴³ Scheuermann: Die Abgrenzung der digitalen Geisteswissenschaften, 2016, S. 61.

⁹⁴⁴ Siehe URL: <https://vtm.epfl.ch/>.

⁹⁴⁵ Scheuermann: Die Abgrenzung der digitalen Geisteswissenschaften, 2016, S. 61.

⁹⁴⁶ Vgl.: Burkhardt: Digitale Datenbanken, 2015, S. 116.

⁹⁴⁷ Scheuermann, Leif: Zu virtuellen Forschungsumgebungen, einer genuin digitalen Hermeneutik sowie deren Visualisierung, Abstract, DHd2016, 12-03.2016, URL: <http://www.dhd2016.de/abstracts/votr%C3%A4ge-024.html>, Stand: 19.02.2018.

ziehbar dokumentiert und in einer geeigneten Visualisierungsform dargestellt wird, sondern auch zur Forschungsfrage führende Vorarbeiten. Für den Historiker Fred Gibbs und den Digitalarchivar Trevor Owens helfen diese Erklärungen dem Leser die Entstehung der Forschungsfrage zu verstehen, mögliche Irrtümer im Bezugssystem zu kennzeichnen und auch Fehler in der Kategorienbestimmung zu entdecken.⁹⁴⁹ Zudem würden sie dazu inspiriert werden, über den Einsatz dieser Techniken und Methoden zum Nutzen ihrer eigenen Forschungstätigkeit nachzudenken.

Scheuermann spricht hierbei von einer Metaebene, die der Forschende zu erarbeiten hat und für die eine eigene Ontologie anzuwenden wäre.⁹⁵⁰ Der Forschende hat den Forschungsprozess selbstreferenziell zu dokumentieren und die vorliegenden Ergebnisse neu zu formulieren und zu visualisieren.⁹⁵¹ Dass dies aber zwingend in einer digitalen Form zu geschehen hat, wie Scheuermann fordert⁹⁵², ist nicht korrekt. Oft können digitale Objekte, worunter auch die Ergebnisse von Auswertungen und Abfragen fallen, nur digital adäquat visualisiert werden. Je nach Forschungsfrage und verwendetem Informationssystem kann auch eine schriftliche Darstellung ausreichen.

Im dritten Schritt interpretiert der Forschende die Ergebnisse seiner Auswertungen und seine digitalen und analogen Objekte. Der Forschende geht bei analogen nach der traditionellen Methodik vor, bei digitalen Objekten sind jedoch mehrere externe Faktoren in die Interpretation miteinzubeziehen, denn sie wird durch die Verwendung von digitalen Informationssystemen beeinflusst. Von besonderem Interesse sind der Digitaldatentyp und das Datenformat eines digitalen Objekts, die eingesetzten Informationssysteme, und die Hypermedialität.

Digitale Informationssysteme können zu einer reichhaltigeren Interpretation von umfangreichem Quellenmaterial führen und Zusammenhänge aufdecken, die sonst verborgen geblieben wären.⁹⁵³ Rege Diskussionen darüber, wie digitale Daten in der Geschichtsforschung genutzt, welchen Gültigkeitsregeln digitale Beweise unterworfen und wie diese interpretiert werden sollen, zeugen von der Spannung zwischen den Ansprüchen an das wissenschaftliche Arbeiten und der Neuartigkeit digitaler Objekte, die als Quellen

⁹⁴⁸ Weder die Auswahl eines Algorithmus, noch ein Algorithmus ist theorie- oder hypothesenfrei.

⁹⁴⁹ Gibbs; Owens: *The Hermeneutics of Data*, 2013, S. 165.

⁹⁵⁰ Scheuermann: *Die Abgrenzung der digitalen Geisteswissenschaften*, 2016, S. 61. Da die geforderte Ontologie nicht im Fokus dieser Arbeit steht, wird diese hier nicht weiter diskutiert.

⁹⁵¹ Dies fordern auch: Gibbs; Owens: *The Hermeneutics of Data*, 2013, S. 168.

⁹⁵² Scheuermann: *Zu virtuellen Forschungsumgebungen, einer genuin digitalen Hermeneutik sowie deren Visualisierung*, 2016.

⁹⁵³ Siehe: Dougherty, Jack; Nawrotzki, Kristen; Rochez, Charlotte D. u. a.: *Conclusions. What We Learned from Writing History in the Digital Age*, in: Dougherty, Jack; Nawrotzki, Kristen (Hg.): *Writing history in the digital age*, Ann Arbor 2013 (Digital humanities), S. 260.

verwendet werden sollen.⁹⁵⁴ Wir sind noch immer nicht mit dem Medium vertraut, das uns massiert.

Im Folgenden werden die bereits erwähnten *externen Faktoren*, die in die Interpretation eines digitalen Objekts einzubeziehen sind, noch einmal zusammengefasst.

5.5.1. Digitaldatentyp

Wenn der Digitaldatentyp eines Objekts festgestellt werden kann, weist dieser bereits auf bei der Interpretation zu berücksichtigende Eigenschaften und Beziehungen zu anderen Objekten hin.

Die Eigenschaften eines Digitalisats, das aus einem physischen Objekt generiert wurde, beeinflussen seinen Aussagewert und damit auch die Interpretation. Es kann beispielsweise von Bedeutung sein, aus welchem Material ein Siegel ist, das als Digitalbild präsentiert wird. Bei reborn-digitalen Objekten ist es beispielsweise von Belang, ob es sich um ein durch eine Hilfssicherung erstelltes Sekundärobjekt handelt, ob es sich aus einem anderen digitalen Objekt generiert, wobei weiterhin eine inhaltliche Abhängigkeit zu einem anderen Objekt besteht⁹⁵⁵ oder ob es generiert wurde, indem alle externen Daten in das Objekt integriert wurden. In den letzten beiden Fällen ist das Objekt zwar als neues Primärobjekt zu betrachten, aber für die Interpretation sollte die Datengrundlage mit einbezogen werden. Ein Sekundärobjekt oder Surrogat wurde als eine Art originalgetreues Abbild des Ursprungsobjekts erstellt, das eventuell so nicht mehr vorhanden ist. Die für den Aussagewert relevanten Elemente des Primärobjekts, insbesondere Hyperlinks, sollten ebenfalls in die Interpretation mit einbezogen werden.

Digital generierte Daten haben weder eine physische noch eine digitale Vorlage. Dennoch können Abhängigkeiten zu anderen Objekten bestehen, insbesondere in der Darstellung einer Webseite mit CSS. Damit stellt sich aber auch die Frage, ob die Intention der Autorschaft dieser Elemente mit in die Interpretation einbezogen werden sollte. Dies kann der Fall sein, wenn die Darstellungsform die Aussage des Objekts massgeblich beeinflusst.

Bereits erwähnt wurde, dass Digitaldatentypen nicht immer eindeutig sind, denn ein digitales Objekt kann auch verschiedene Typen vereinen. Beispielsweise kann ein Digitalisat mit Metadaten aus einer Datenbank automatisch ergänzt werden.

⁹⁵⁴ Siehe bspw. die Veranstaltung 'Beyond Methods of Mining: A Workshop on Doing Historical Research using Digital Data', Universität Utrecht (NL), 14./15.09.2015. URL: <http://asymenc.wp.hum.uu.nl/beyond-methods-of-mining-a-workshop-on-doing-historical-research-using-digital-data/>, Stand: 06.05.2018.

⁹⁵⁵ Das Objekt kann auch aus mehreren Objekten generiert werden und ist von allen abhängig.

5.5.2. Datenform(at)

Simon Margulies geht in seiner obigen Aussage, dass die inhaltliche Interpretation digitaler Objekte aufgrund ihrer analogen Repräsentation sich nicht von jener anderer Quellen unterscheidet, davon aus, dass ein Informationssystem ausschliesslich für den Menschen rezipierbare Objekte anzeigen. Wie in Kapitel 2. 1 und 0 erläutert, liegt ein digitales Objekt gleichzeitig in verschiedenen Datenformen vor, die zwar dasselbe Objekt beschreiben, aber nicht zwingend vom Menschen interpretiert werden können. Binär- oder Quellcodes bedürfen einer Interpretation durch ein Informationssystem, bevor das Ergebnis anschliessend vom Menschen gelesen und interpretiert werden kann.

Zudem können auch Daten dargestellt werden, die vom Menschen zwar gelesen werden können, aber keinen unmittelbaren Sinn ergeben. Beispielsweise sind Datentabellen einer Netzwerkanalyse erst rezipierbar, wenn diese visualisiert werden. Solche Rohdaten bedürfen folglich ebenfalls zuerst einer algorithmischen Interpretation durch ein Informationssystem und sind deshalb nur mittelbar vom Menschen interpretierbar. Ein Objekt kann zudem rezipierbar sein und gleichzeitig auch als Rohdatum Verwendung finden, wie beispielsweise ein einzelner literarischer Text aus dem 18. Jahrhundert. Dieser kann vom Nutzer gelesen und interpretiert werden, ist aber für eine TEI-Untersuchung zur Literatur in der Rokokozeit nur ein Teil der Rohdaten. Erst die algorithmische Verarbeitung bringt diesen Text als Teil des gesamten Textkorpus in eine für die entsprechende Fragestellung rezipier- und interpretierbare Form. Weil Rohdaten für deren Interpretation zuerst in eine rezipierbare Form gebracht werden müssen, sind *in Ergänzung* zur traditionellen Methode informationstechnische Kenntnisse und fundierte Medienkompetenzen über digitale Medien nötig. Die Informationstechnik ist hierbei keine Hilfswissenschaft, sondern die Basis einer wissenschaftlich fundierten, historisch-kritischen Untersuchung von digitalen Daten.

Daten in einer visualisierten Form können für die Interpretation eine weitere Herausforderung darstellen. Beispielsweise besteht bei einer *Infographik* die Gefahr, dass sie nicht gelesen, sondern nur betrachtet wird.⁹⁵⁶ Die Interpretation einer Graphik unterliegt gruppenspezifischen Kommunikationskulturen und Sehkonventionen und ist "in Kontexte sozialer Einschreibungen und historischer Konjunkturen eingebunden."⁹⁵⁷ Das Computerbild ist vielfältig

⁹⁵⁶ Vgl.: Reichert, Ramón: Facebooks Big Data. Die Medien- und Wissenstechniken kollektiver Verdichtung, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 444.

⁹⁵⁷ Ebd.

interpretierbar und steht einer regelkonformen Dateninterpretation entgegen.⁹⁵⁸ Der Historiker hat deshalb den Blick für Datenqualitäten und Relationen zu schulen.⁹⁵⁹

5.5.3. Informationssysteme

Da die Erstellung und die erneute Darstellung eines digitalen Objekts nur durch ein Informationssystem möglich sind, ist dieses quellenkritisch zu überprüfen. Hier sind unter anderem limitierende Faktoren bei der Erstellung, der Verarbeitung und der Darstellung des Objekts zu untersuchen. Kritisch zu hinterfragen sind insbesondere Änderungen, die durch technologische Anpassungen einer Software wie einem Archivinformationssystem, die Trägerschaft der publizierenden Institution, das Geschäftsmodell, die rechtlichen Rahmenbedingungen oder Formatänderungen eintreten können.

Insbesondere bei der Verwendung von Rohdaten sollte der Nutzer vertiefte Kenntnisse über die Funktionalitäten der Systeme haben, die diese Daten generiert haben und zur Verfügung stellen.⁹⁶⁰ Beispielsweise ist bei computerunterstützter Forschung, von der Wissenschaftsphilosophin Sabina Leonelli als "machine-driven discovery"⁹⁶¹ bezeichnet, zu überprüfen, ob die eingesetzten Systeme und Algorithmen für die Forschungsfrage verwendet und ob die infrastrukturellen Rahmenbedingungen akzeptiert werden können. Es ist auch nachzufragen, welche Abhängigkeiten bei der Datengenerierung bestehen, vor allem "welches Interesse unterschiedliche Förderinstitutionen mit ihrem finanziellen Engagement verbinden."⁹⁶² Dadurch lassen sich eventuell Intentionen feststellen, wodurch die vorliegenden Objekte anders zu interpretieren sind.

Es ist zu präzisieren, dass unter Rohdaten nicht ein unstrukturierter Datenberg zu verstehen ist, sondern dass es Primärdaten sind, die nach bestimmten algorithmischen Regeln in einer bestimmten Form und einem bestimmten Format erstellt wurden. Im Algorithmus festgelegte Parameter bestimmen, welche Daten registriert werden. Was nicht in diesen Parametern enthalten ist, wird auch nicht festgehalten. Das Informationssystem 'versteh' Daten weder bei der Generierung noch bei der Verarbeitung. Einem Menschen können hingegen nicht vorgesehene Anomalien oder die Abwesenheit von Daten auffallen, der vielleicht auch gleichzeitig den Grund dafür erkennt. Deshalb ist die computergestützte Analyse von digitalen Datenbeständen weder objektiv noch neutral:

⁹⁵⁸ Vgl.: Ebd., S. 445.

⁹⁵⁹ Vgl.: Richterich, Annika: Infodemiologie – von «Supply» zu «Demand». Google Flu Trends und transaktionale Big Data in der epidemiologischen Surveillance, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 357.

⁹⁶⁰ Vgl.: Leonelli: Data Interpretation in the Digital Age, 2014. Leonelli bezieht sich hauptsächlich auf das Wissenschaftsgebiet der Biologie, ihre Argumentation kann aber für die gesamte Wissenschaft geltend gemacht werden.

⁹⁶¹ Ebd., S. 8.

⁹⁶² Röhle: Big Data – Big Humanities? Eine historische Perspektive, 2014, S. 169.

*"What is too often forgotten, though, is that our digital helpers are full of 'theory' and 'judgement' already. As with any methodology, they rely on sets of assumptions, models, and strategies. Theory is already at work on the most basic level when it comes to defining units of analysis, algorithms, and visualisation procedures."*⁹⁶³

Vorhandene Datenbestände verleiten unter Umständen auch dazu, sie *nutzen zu müssen*. Statt dass zuerst die theoretischen Grundlagen einer Analyse festgelegt werden und dann der Datenbestand auf Verwendungsmöglichkeiten geprüft wird, wird lieber vom Datenbestand ausgegangen und "ex post Begründungen oder sogar Theoretisierungen"⁹⁶⁴ werden nachgeliefert. Der Journalist Chris Anderson meinte in einem viel beachteten, provokativen und umstrittenen Artikel sogar, dass leistungsfähige Computersysteme Muster erkennen können, die Wissenschaftler nie sehen würden und deshalb Datenmodellierung obsolet sei:

*"We can stop looking for models. We can analyze the data without hypotheses about what it might show."*⁹⁶⁵

Dies ist meines Erachtens kein Aufruf zur Abschaffung von Theorien und Modellen, sondern ein sarkastischer Vorwurf an Wissenschaftler und Wissenschaftsdisziplinen wie *Data Driven Science* und *Data Driven Research*, die sich an den bereits vorhandenen Daten orientieren. So unterstützenswert neue Forschungsansätze und -methoden sind, nachträgliches Theoretisieren verstößt weiterhin gegen die gängigen Regeln des wissenschaftlichen Arbeitens.

Des Weiteren sind Informationssysteme problematisch, die *Daten unterschiedlicher Herkunft* zusammenfassend darstellen. Gemeint sind beispielsweise Records Management Systeme, die alle für eine Sachbearbeitung nötigen Informationen anzeigen, die jedoch in verschiedenen Datenbanken des Systems abgelegt sind und deren Dateneigner andere Stellen der Institution sind. Unter Umständen kann die Generierung des digitalen Objekts nur dann nachvollzogen werden, wenn die Herkunft der einzelnen Elemente und die Funktionalität dieses Moduls des Informationssystems bekannt ist.

Objekte, die durch Informationssysteme aus dem Feld der *künstlichen Intelligenz* erstellt wurden, sind durch den Menschen kaum mehr interpretierbar. Der Programmierer Alexander Mordvintsev hat als Weiterführung der Bilderkennung demonstriert⁹⁶⁶, wie ein künstliches neuronales Netz durch Training und mit vorgegebenen Parametern ein Bild interpretiert und

⁹⁶³ Rieder; Röhle: Digital Methods, 2012, S. 70.

⁹⁶⁴ Mahrt, Merja; Scharkow, Michael: Der Wert von Big Data für die Erforschung digitaler Medien, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 227.

⁹⁶⁵ Anderson, Chris: The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, WIRED MAGAZINE, 27.06.2008, URL: http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory, Stand: 07.06.2014.

⁹⁶⁶ Siehe: Mordvintsev, Alexander; Olah, Christopher; Tyka, Mike: Inceptionism: Going Deeper into Neural Networks, Google Research Blog, 17.06.2015, URL: <http://googleresearch.blogspot.com/2015/06/inceptionism-going-deeper-into-neural.html>, Stand: 24.07.2015.

zu einem neuen Bild als Ergebnis kommt, was ich als Quartärquelle bezeichnet habe.⁹⁶⁷ Ohne Kenntnis der zugrundeliegenden Konzepte der Neuroinformatik ist es nicht möglich, eine kritische Analyse solcher Objekte vorzunehmen. Dasselbe gilt für die Interpretation eines durch ein Informationssystem der künstlichen Intelligenz selbständig generierten Textes. Auch hier ist es kaum mehr möglich, eine kritische Analyse vorzunehmen, denn Fragen nach Autor, Adressat oder Intention können ohne Kenntnis der Konzepte und Algorithmen nicht beantwortet werden.

5.5.4. Hypermedialität

Beim Lesen von physischen Objekten können Hinweise oder Verweise auf andere Objekte mit verschiedenen Techniken wie Klammern oder Fussnoten an das Objekt angefügt werden. Um diesen Referenzen zu folgen, sind die referenzierten Objekte zu beschaffen und parallel zu lesen. Demgegenüber kann eine Anmerkung oder ein Verweis im digitalen Objekt integriert und fest verknüpft werden. Damit wird nicht nur referenziert, sondern die Verknüpfung wird (bequem) realisiert.⁹⁶⁸ Hyperlinks können deshalb nicht nur zum Zweck einer traditionellen Zitierung verwendet werden, sondern verschiedene Funktionen übernehmen.⁹⁶⁹ Bei vorhandenen Links ist einerseits zu entscheiden, ob und bis zu welcher Tiefe diese überprüft werden sollen, andererseits sollte nachvollzogen werden können, aus welcher Intention der Autor diese Hyperlinks eingefügt hat.⁹⁷⁰ Navigationselemente oder Hinweise können vernachlässigt werden, bei anderen Linkarten wäre die Beziehung zwischen den zwei Objekten ist genauer zu untersuchen. Besonders wenn unter dem Link relevante inhaltliche Ergänzungen zum Primärobjekt enthalten sind, ist eine qualitative Untersuchung notwendig. Das Link-Check-Element in der Belegkette könnte dem Nutzer hilfreiche Hinweise und Informationen bieten.

5.5.5. Zusammenfassung

Die Formatänderung von einem physischen zu einem digitalen Objekt verändert auch den Interpretationsprozess.⁹⁷¹ Die inhaltliche Interpretation digitaler Objekte kann nicht wie bis

⁹⁶⁷ Siehe Kap.5.1.3. Suchtechniken.

⁹⁶⁸ Nach: Krameritsch: Geschichte(n) im Netzwerk. Hypertext und dessen Potenziale für die Produktion, Repräsentation und Rezeption der historischen Erzählung, 2007, S. 139. Krameritsch spricht nur von linearen Texten, aus obigen Ausführungen lässt sich jedoch ableiten, dass dies für jede Objektart gilt.

⁹⁶⁹ Siehe bspw.: Weller, Katrin; Dröge, Evelyn; Puschmann, Cornelius: Citation Analysis in Twitter: Approaches for Defining and Measuring Information Flows within Tweets during Scientific Conferences, in: #MSM2011, 1st Workshop on Making Sense of Microposts, Heraklion, Kreta 2011, S. 11. URL: http://ceur-ws.org/Vol-718/paper_04.pdf, Stand: 06.05.2018.

⁹⁷⁰ Siehe Kap. 5.4.11. Inhalt und Kap. 5.4.12. Beziehungen zu Objekten.

⁹⁷¹ Vgl.: Weller: Conclusion. A changing field, 2013, S. 201.

anhin mit traditionellen Methoden erfolgen, sondern es ist auch das für die Erstellung und das für die erneute Darstellung verwendete elektronische Informationssystem mit einzubeziehen. Werden die zu untersuchenden digitalen Objekte in einer digitalen Forschungsumgebung zusammengefasst, geordnet und durch Neukombinieren und Manipulieren aufgearbeitet, stellt dies bereits einen hermeneutischen Prozess dar und ist entsprechend zu dokumentieren und adäquat zu visualisieren. Die Digitaldatentypen bieten Hinweise für die Interpretation, insbesondere auf Beziehungen zu und Abhängigkeiten von anderen Objekten. Eventuell sind die dargestellten Inhalte vorab durch ein algorithmisches Verfahren in eine rezipierbare Form zu bringen. Es ist abzuwägen, welche der durch Hyperlinks eingebundenen, aus externen Quellen stammenden, ergänzenden Inhalte in die Interpretation mit einbezogen werden sollten. Besondere Vorsicht ist geboten bei der Interpretation von grossen Datenbeständen respektive Big Data, Visualisierungen und algorithmisch erzeugten Inhalten durch Systeme der künstlichen Intelligenz. Die Aussagekraft, Relevanz und Relationen der von Systemen der künstlichen Intelligenz generierten Inhalte sind besonders kritisch zu hinterfragen. Bei Visualisierungen sind zudem die gruppenspezifischen Kommunikationskulturen und Sehkonventionen zu beachten. Erst mit Kenntnissen über die Rohdaten und die verwendeten Algorithmen kann eine fundierte Interpretation erfolgen.

5.6. Darstellung und Publikation

*"Die Publikationskultur in der Geschichtswissenschaft steckt in einem grundlegenden Veränderungsprozess."*⁹⁷² (Hecker-Stampehl, 2013)

Grundsätzlich werden die Forschungsergebnisse in einer schriftlichen Darstellung publiziert, ansonsten der Forschungsprozess nicht abgeschlossen ist.⁹⁷³ Das idealtypische Verfahren innerhalb des wissenschaftlichen Publikationswesens⁹⁷⁴ startet mit der Einreichung eines Manuskripts (Input), gefolgt von einem Gutachten (Control), der anschliessenden Publikation (Output) und schliesslich der Rezension (Feedback).⁹⁷⁵ Digitale Publikationsverfahren bieten nun neue Möglichkeiten, wie das Manuskript erstellt und eingereicht, wie es begutachtet, überarbeitet und anschliessend publiziert und rezensiert werden kann.⁹⁷⁶ In den Digital Humanities und der Digital History wird nun deshalb diskutiert, ob neue Formen der

⁹⁷² Hecker-Stampehl, Jan: Bloggen in der Geschichtswissenschaft als Form des Wissenstransfers, in: Haber, Peter; Pfanzelter, Eva (Hg.): *historyblogosphere. Bloggen in den Geschichtswissenschaften*, München 2013, S. 37.

⁹⁷³ Siehe Kap. 3.1.5. Darstellung.

⁹⁷⁴ Zur Zukunft der wissenschaftlichen Publikation siehe bspw.: Cope, Bill; Phillips, Angus (Hg.): *The future of the academic journal*, Oxford 2014 (Chandos information professional series).

⁹⁷⁵ Vgl.: Haber: *Digital Past*, 2011, S. 143.

⁹⁷⁶ Siehe weiterführend bspw.: Lengwiler: *Praxisbuch Geschichte*, 2011, S. 267-270.

Darstellung in die Geschichtsschreibung Einzug halten beziehungsweise welche Beiträge als akademisch gelten sollen, wenn sie in einer neuen Darstellungsform präsentiert werden.⁹⁷⁷

Der mediale Wandel hat die Aufmerksamkeit für das *Wie* der Darstellung verstärkt und zu neuen Lesegewohnheiten geführt. Physische Narrative präsentieren sich hauptsächlich in einer chronologischen Ordnung und erzählen eine einzige, kohärente Geschichte. Hypertextualität bricht mit dieser stringenten, bindenden Ordnung und stellt damit die Monographie als wichtigste wissenschaftliche Darstellungsform infrage, insbesondere zur Erlangung akademischer Würden.⁹⁷⁸ Neue Formen wie Webseiten, Blogs oder kollaborativ erarbeitete wissenschaftliche Texte sollten als wissenschaftliche Publikation Anerkennung finden. Dies "rüttelt aber auch am Selbstbild von Historikern, das immer noch gerne davon ausgeht, dass ein Text das Ergebnis eines solitären, individuell erarbeiteten Forschungsprozesses ist."⁹⁷⁹ Die Akzeptanz neuer Darstellungs- und Kommunikationsformen⁹⁸⁰ hängt zurzeit noch stark von Einzelpersonen, insbesondere den Professoren ab. Mit der Übernahme von Lehrstühlen durch Digital Natives und technikaffinen Digital Immigrants nimmt deren Anzahl langsam zu, womit auch eine neue Publikationskultur in der Wissenschaft Einzug hält.⁹⁸¹ Damit wird sich die Beschränkung oder vielmehr die Selbstbeschränkung auf gedruckte Journalartikel oder Buchformen beziehungsweise auf die monographische Publikationsformen vermutlich nicht mehr lange halten können.⁹⁸²

Publikation dürfe nicht mit Kommunikation verwechselt werden, meint der Philosoph und Digitalwissenschaftler Niels-Oliver Walkowski in seiner Verteidigung der Monographie, denn eine Publikation unterscheidet sich von einer Kommunikationssituation durch "Abgeschlossenheit, Finalität und die Herstellung eines autodiskursiven Raumes."⁹⁸³ Er steht beispielhaft für jene, für die eine Publikation ""Stopp, jetzt nur das!"⁹⁸⁴ bedeuten muss. Für

⁹⁷⁷ Siehe Kap. 1.1. Forschungsstand.

⁹⁷⁸ Nach: Haber: Digital Past, 2011, S. 116 f. Die meisten physischen Narrative müssen nicht zwingend linear rezipiert werden, eine nichtlineare Lektüre macht aber keinen Sinn.

⁹⁷⁹ Haber: Zeitgeschichte und Digital Humanities, 2012.

⁹⁸⁰ Siehe zu digitaler Fachkommunikation: Prinz: Kommunikation im digitalen Raum, 2016.

⁹⁸¹ Siehe Aussage von PD Dr. Jochen Strobel: Ostendorf-Rupp, Sonja: Kurz nachgefragt: Akzeptanz digitaler wissenschaftlicher Publikationen, Dresden Summer School 2012, 02.10.2012, URL: <http://dss.hypotheses.org/455>, Stand: 04.03.2014. Zum Wandel der Publikations- und Forschungskultur siehe bspw.: Haber, Peter; Pfanzelter, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013; Nentwich: Cyberscience 2.0, 2012; Wulf, Tobias; Foerster, Sascha: Wissenschaftliches Bloggen: Digitale Miszelle oder kommunikativer "Turn"? Vortrag von Tobias Wulf und Sascha Foerster #digigw14, Digitale Geschichtswissenschaft, 16.09.2014, URL: <http://digigw.hypotheses.org/838>, Stand: 18.11.2014; Bruns, Axel: Blogs, Wikipedia, Second Life, and Beyond from Production to Produsage, New York 2008 (Digital formations).

⁹⁸² Verfechter der physischen Publikation sind bspw.: Groebner: Wissenschaftssprache digital. Die Zukunft von gestern, 2014; Koller: Geschichte digital, 2016, S. 37.

⁹⁸³ Walkowski, Niels-Oliver: Text, Denken und E-Science, Eine intermediale Annäherung an eine Konstellation, in: Gugerli et al. (Hg.): Digital Humanities, Zürich 2013 (Nach Feierabend 9), S. 48.

⁹⁸⁴ Groebner: Wissenschaftssprache digital. Die Zukunft von gestern, 2014, S. 121.

diese Gruppe ist der Unterbruch des kommunikativen Flusses nötig, damit eine intensivere Auseinandersetzung mit dem Inhalt erfolgen könne.

Viele der heutigen Möglichkeiten der Digitalmedien sind jedoch Mischformen.⁹⁸⁵ Beispielsweise kann ein gedruckter Zeitschriftenartikel als Publikation eingeordnet werden. Wird er auch als Online-Artikel publiziert, kann er mit den üblichen⁹⁸⁶ digitalen Kollaborations- und Distributionsmöglichkeiten ergänzt werden und ist damit eher eine Publikationsform, die ich als *kommunikative Publikation* bezeichne. Noch komplexere Formen wie das interaktive Romanprojekt *Beside Myself*⁹⁸⁷ des Schriftstellers Jeff Gomez, in dem der Leser mehreren Erzählsträngen folgen oder den Charakteren und dem Autor aus dem Text heraus E-Mails zusenden kann, harren noch der korrekten Bezeichnung.⁹⁸⁸

Gegen eine Veränderung der Publikationskultur wehren sich auch Verlage, deren Geschäftsmodell durch die neuen, digitalen Publikationsformen ins Wanken gerät. Die Zusammenarbeit zwischen Autoren und Verlagen ist teilweise vorteilhaft, beispielsweise für die Reputation des Autors oder eine Authentizitätskontrolle für den Leser. Nachteile wie hohe Kosten für Zeitschriftensubskriptionen oder Beiträge ohne Vergütung animieren jedoch Wissenschaftler und Gedächtnisinstitutionen vermehrt dazu, die neuen Distributionskanäle interessiert zu beobachten und immer stärker zu nutzen. Die Vorteile der neuen Publikationskanäle scheinen zu überwiegen, wie auch Hubertus Kohle feststellte, der selber einen Lehrstuhl für Mittlere und Neuere Kunstgeschichte an der Ludwig-Maximilians-Universität München innehat:

*"Das digitale Publizieren wird nicht zurückgehen. Das behaupten nur Verlage und Professoren."*⁹⁸⁹

Auch einige geisteswissenschaftliche Gremien behaupten dies, insbesondere wenn es um die Publikation von Forschungsinhalten geht. Beispielsweise werden im SAGW-Positionspapier von 2012 die digitalen Formate nicht einmal erwähnt und die Monographie scheint noch immer das Richtmass für akademische Qualifikationsschriften zu sein.⁹⁹⁰ Die angekündigte

⁹⁸⁵ Eindeutigkeit ist nicht möglich, denn jede Publikation ist auch eine Art von Kommunikation.

⁹⁸⁶ Bspw. Button, der die URL des Artikels als Mail oder Tweet versendet, Kommentarfunktion etc.

⁹⁸⁷ Gomez, Jeff: *Beside Myself: An Interactive novel for the iPad*, Beside Myself, 2016, URL: <http://www.besidemyself.com/>, Stand: 24.07.2016.

⁹⁸⁸ Siehe auch: Gomez, Jeff: Die erzählerische Singularität: Geschichten erzählen im digitalen Zeitalter, *Aus Politik und Zeitgeschichte*, 02.10.2012, URL: <http://www.bpb.de/apuz/145376/geschichten-erzaehlen-im-digitalen-zeitalter?p=all>, Stand: 11.02.2014.

⁹⁸⁹ Prof. Dr. Hubertus Kohle, Anmerkung im Vortrag: Die Gedächtnisinstitution im digitalen Zeitalter. *Perspektiven und Chancen*, 2.10.2012, in: Ostendorf-Rupp: *Kurz nachgefragt*, 2012.

⁹⁹⁰ Böhler, Fritz; Engler, Balz; Glauser, Jürg u. a.: *Positionspapier – Für eine Erneuerung der Geisteswissenschaften. Empfehlungen der SAGW zuhanden der Leitungsorgane der Hochschulen, der Lehrenden, der Förderorganisationen und des Staatssekretariats für Bildung und Forschung*, 05.12.2012, insb. S. 24. URL: www.sagw.ch/dms/sagw/schwerpunkte/wissenschaft_und_technikgeschichte/publikationen/Positionspapier_mit_Ruecken, Stand: 06.05.2018. Im Positionspapier werden die Begriffe 'digital', 'online' oder 'Web' nicht verwendet.

Erneuerung der Geisteswissenschaften soll noch im traditionellen Sinne erfolgen.⁹⁹¹ Dennoch öffnete sich die SAGW gegenüber Digitalmedien, indem sie moderne Methoden der Wissensproduktion, -vermittlung und -nutzung unterstützt.⁹⁹²

Es ist nicht mehr zu diskutieren, ob neue Darstellungsformen anerkannt werden sollen, sondern unter welchen Rahmenbedingungen sie anzuerkennen sind.⁹⁹³ Ich diskutiere im Folgenden fünf Themenbereiche, die sich im Kontext der Digitalität verändern. Anhand von Blogs zeige ich Möglichkeiten neuer Darstellungsformen auf und dass sie auch als wissenschaftliche Publikations- und Kommunikationsform wahrgenommen werden sollten. Die Verwendung solcher Formen verändert auch den Schreibprozess inklusive der wissenschaftlichen Belegkette und die Visualisierung der Erkenntnisse. Open Access und Open Data sind neue Konzepte, die den wissenschaftlichen Publikationsprozess stark beeinflussen und neu gestaltet.

5.6.1. Neue Darstellungsformen

Neue Darstellungsformen wie Webauftritte oder Blogs erlauben es Wissenschaftlern nicht nur, ihre Forschungsergebnisse einem grossen Publikum zu präsentieren oder zur Diskussion zu stellen, sondern ermöglichen erst die adäquate Darstellung digitaler Objekte. Da diese nicht materialisiert werden können, sind sie auf eine digitale Publikationsform angewiesen.

Der Sinn des Einsatzes von Digitalmedien zur Publikation wissenschaftlicher Arbeiten wurde bereits von verschiedenen Wissenschaftlern erkannt. Einer davon ist Patrick Sahle, der sogar meint:

*"Wenn es darum geht, für die Forschung inhaltsreiche, anspruchsvolle und zeitgemässe Editionen bereitzustellen, dann führt an digitalen Formen kein Weg vorbei."*⁹⁹⁴

Allerdings scheinen grosse Wissenschaftsinstitutionen solche Projekte trotzdem zu scheuen und führen als Gegenargumente "technische Hürden, inhaltliche Komplexität, mangelnde Erfahrung und beschränkte Möglichkeiten, digitale Editionen zu nutzen, befürchteter Qualitätsverlust, Sorge um nicht sichergestellte langfristige Zugriffsmöglichkeit und Datenpflege

⁹⁹¹ Siehe zur kritischen Beurteilung des Positionspapieres auch: Wampfler: «online fist». Geisteswissenschaften als Social Media, 2013, S. 92 f.

⁹⁹² Siehe Schwerpunkt 'Wissenschaft im Wandel' unter URL: <http://www.sagw.ch/sagw/schwerpunkte/wi-wandel.html>.

⁹⁹³ Die bereits mehrmals geforderte Verwendung von Digitalen Signaturen und verlässlichen Repositories können hierbei eine wichtige Rolle spielen.

⁹⁹⁴ Sahle, Patrick: Digitale Editionsformen: zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, Bd. 1-3, Norderstedt 2013 (Schriften des Instituts für Dokumentologie und Editorik 7-9), Bd. 2, S. 123.

sowie wirtschaftliche und rechtliche Bedenken im Hinblick auf Verlage⁹⁹⁵ auf. Solche Institutionen müssen sich den Vorwurf gefallen lassen, dass sie weniger Lösungen suchen, als lieber an Altbewährtem und teilweise Überholtem festhalten wollen, denn verschiedene digitale Darstellungsformen haben sich als sehr nützlich erwiesen und werden rege genutzt.⁹⁹⁶ Beispielsweise werden Kommunikationsmittel wie Webauftritte, Blogs, Facebook, Twitter, E-Mail oder Voice-over-IP Services wie Skype auch von Forschenden eingesetzt. Digitale Dokumentationsmittel wie elektronische Laborbücher ergänzen oder ersetzen handschriftliche Notizbücher der Forschenden. Insbesondere Blogs haben sich für die Veröffentlichung von Forschungsergebnissen oder wissenschaftlichen Beiträgen bei einem immer grösser werdenden Kreis von Wissenschaftlern als Publikationsmittel etablieren können. Von einer noch überwiegenden Anzahl jedoch werden sie leider noch nicht als wissenschaftliche Publikation im engeren Sinne⁹⁹⁷ akzeptiert. Die Mehrheit⁹⁹⁸ der Wissenschaftler wehrt sich gegen diese *Revolution im traditionellen Publikationsprozess*.⁹⁹⁹ Eine kritische Haltung ist durchaus berechtigt, denn "wenn es keine übergeordnete Erzählung mehr gibt, die den einzelnen Geschichten eine Richtung unterlegt, dann läuft Geschichte Gefahr, als rein subjektivistische Veranstaltung missverstanden zu werden."¹⁰⁰⁰ Dieser Gefahr können die Geschichtswissenschaftler mit entsprechend geschickter Auswahl des Digitalmediums und einer darauf angepassten Produktionsweise entgegen.

Es wäre für die Geschichtswissenschaft von Vorteil, wenn geeignete digitale Darstellungsformen den traditionellen Publikationsformen gleichgestellt würden, denn digitale Objekte können anders meistens gar nicht adäquat dargestellt werden. Wenn in einer wissenschaftlichen Arbeit hauptsächlich digitale Objekte untersucht werden und Verwendung finden, dann sollte für die Publikation der Forschungsergebnisse eine digitale Darstellungsform bevorzugt werden. Damit könnte auch ein unnötiger Medienbruch verhindert werden, denn die Überführung eines digitalen Objekts in eine analoge Form ist mit grossen Verlusten verbunden, weil es nicht im vom Autor intendierten Sinne präsentiert werden kann. Digitale Quellen mit

⁹⁹⁵ Arend, Sabine: Rezension zu: Sahle, Patrick: Digitale Editionsformen, Rezensionen, H / SOZ / KULT, 14.08.2014, URL: <http://hsozkult.geschichte.hu-berlin.de/rezensionen/2014-3-109>, Stand: 14.08.2014.

⁹⁹⁶ Die Nutzung von Darstellungsformen unterliegt einem ständigen Wandel. Was heute verwendet wird, kann morgen schon wieder 'veraltet' sein.

⁹⁹⁷ Unter 'im engeren Sinne' werden systematische, intersubjektive, angemessene und sinnvolle Beiträge zur Forschungsdiskussion verstanden (obwohl jedes Kriterium problembehaftet ist). Nach: Stock, Wolfgang G.: Was ist eine Publikation? Zum Problem der Einheitenbildung in der Wissenschaftsforschung, in: Gesellschaft für Wissenschaftsforschung e.V. (Hg.): Wissenschaft und Digitale Bibliothek: Wissenschaftsforschung Jahrbuch 1998, Berlin 2010², S. 252-256. URL: http://www.wissenschaftsforschung.de/JP98_239-282.pdf, Stand: 23.06.2015.

⁹⁹⁸ Ob sie tatsächlich die Mehrheit bilden, ist noch nicht untersucht worden.

⁹⁹⁹ Siehe: Collins u. a.: Going Digital, 2015, S. 37.

¹⁰⁰⁰ Epple, Angelika: Verlinkt, vernetzt, verführt – verloren? Innovative Kraft und Gefahren der Online-Historiographie, in: Haber, Peter; Epple, Angelika (Hg.): Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0, Zürich 2004 (Geschichte und Informatik 15), S. 27.

hypertextuellen Elementen können viel besser in die Publikation integriert werden, was zur Reduktion sonst nötiger Erklärungen beitragen kann. Eine der digitalen Quelle angemessene Publikationsform erfordert eine digitale Darstellung. Deshalb sollte auch bei akademischen Qualifikationsarbeiten die Möglichkeit bestehen, sie in einer digitalen Form einzureichen und anerkennen zu lassen. Unterstützt wird diese Forderung beispielsweise vom Digital Peer Publishing (DiPP), einer Initiative des Ministeriums für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen. DiPP will explizit neue Formen des digitalen Informationsmanagements in der Wissenschaft etablieren.¹⁰⁰¹

Vom *Einstellen* einer Publikation in ein Repository abgesehen, was nach Wolfgang Schmale nicht als digitale Geschichtswissenschaft zu bezeichnen sei¹⁰⁰², sind zurzeit Webauftritte¹⁰⁰³ und im Speziellen Blogs die bevorzugten Darstellungsformate für die digitale wissenschaftliche Publikation und Präsentation.

Vor allem wissenschaftliche Blogs¹⁰⁰⁴ sind "Wissensorte der Forschung"¹⁰⁰⁵, wie die Historikerin Mareike König, Abteilungsleiterin Digital Humanities des Deutschen Historischen Institutes in Paris, festgestellt hat, die "tradierte Formate und Standards von Wissenschaftlichkeit und damit unsere Forschungskultur insgesamt infrage stellen."¹⁰⁰⁶ Sie werden als Werkzeug der (Selbst-)Publikation, der Kommunikation und der Vernetzung genutzt und dienen dem Informations-, Identitäts- und Beziehungsmanagement. Blogs sind interaktiv und erlauben einen schnellen, direkten und unvermittelten fachlichen Austausch über Kommentare und Verlinkungen und geben zusätzlich Einblick in den Forschungsprozess. Sie sind zudem sowohl für die akademische als auch für die breite Öffentlichkeit sichtbar respektive führen beide Gruppen an einem Veröffentlichungsort zusammen.¹⁰⁰⁷ König sieht Blogs als "multimediales, schriftliches Gespräch"¹⁰⁰⁸ zwischen dem Blogger und dem Fachpublikum einerseits sowie zwischen diesem und der interessierten Öffentlichkeit

¹⁰⁰¹ URL: <http://www.dipp.nrw.de/>.

¹⁰⁰² Siehe: Schmale: Digitale Geschichtswissenschaft, 2010, S. 123.

¹⁰⁰³ Bspw: L.I.S.A. (URL: <http://www.lisa.gerda-henkel-stiftung.de/>), historypin (URL: <https://www.historypin.org/>), histoGraph (URL: http://eipcm.org/histogram_experiment/). Zu Webauftritten für geschichtswissenschaftliche Inhalte siehe noch immer: Cohen; Rosenzweig: Digital history, 2006.

¹⁰⁰⁴ Eine Einführung bieten bspw.: Haber; Pfanzer (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, 2013.

¹⁰⁰⁵ Hier und im Folgenden nach: König: Blogs als Wissensorte der Forschung, 2015.

¹⁰⁰⁶ König, Mareike: Herausforderung für unsere Wissenschaftskultur: Weblogs in den Geisteswissenschaften, in: Schmale, Wolfgang (Hg.): Digital Humanities: Praktiken der Digitalisierung, der Dissemination und der Selbstreflexivität, Stuttgart 2015 (Geschichte), S. 57. Der Beitrag gibt einen Einblick in wissenschaftliche Blogs.

¹⁰⁰⁷ Nach: Mounier, Pierre: Die Werkstatt öffnen: Geschichtsschreibung in Blogs und Sozialen Medien, in: Haber, Peter; Pfanzer, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013, S. 52.

¹⁰⁰⁸ König: Blogs als Wissensorte der Forschung, 2015, Fol. 10.

andererseits. Für die Kommunikationswissenschaftlerin Lisa Bolz und den Historiker und Blogspezialisten Sascha Foerster ist ein Blog eine Übungsanlage, eine Kommunikationsform und eine Gedankenstütze, die hilft, Netzwerke zu bilden.¹⁰⁰⁹ Einen Blog zu schreiben macht – aus eigener Erfahrung – auch viel Spass.

Eine grosse Vielfalt an Blogtypen zeugt von den verschiedenen Formen der Wissens- und Wissenschaftskommunikation im Zeichen der gemeinsamen Nutzung, auch unter dem Begriff *Sharing* bekannt.¹⁰¹⁰ So gibt es beispielsweise themenübergreifende oder themenzentrierte Blogs, Einzel- oder Gruppenprojekte, veranstaltungs-, publikations- oder projektbegleitende Diskussionsforen. Die Beiträge können unter anderem Hypothesen, Forschungsfragen oder -ergebnisse beinhalten, Veranstaltungen ankündigen, Publikationen rezensieren, den aktuellen Forschungsstand aufzeigen oder eine Diskussion mit den Lesern anregen. Das (Experten-)Wissen sichtbar zu machen und damit den wissenschaftlichen Diskurs mitzugestalten ist das gemeinsame Ziel der Blogs. Zusätzlich können mit hypertextuellen und taxonometrischen Verknüpfungen unterschiedliche Informationskomplexe verbunden werden, die sonst kaum Berührungspunkte haben. Das führt dazu, dass Blogs viel besser vernetzt sind, als dies bei physischen Publikationen je möglich sein wird.¹⁰¹¹

Sehr erfolgreich ist die Plattform *hypotheses.org*. Beim Start wurde im französischen Sprachraum anstelle von Blog der Begriff *carnets de recherche* (Forschungsjournale oder Forschungshefte) etabliert, um die Vorstellung von Blog als Austauschmedium für Belanglosigkeiten zu vermeiden.¹⁰¹² Diese Plattform befriedigt das Bedürfnis "nach direkter, vernetzter und schneller Kommunikation im Bereich der Geisteswissenschaften".¹⁰¹³ Eine Qualitätskontrolle findet zwar nur beschränkt statt, aber das Projekt konnte sich dennoch eine sehr gute Reputation für diese Publikationsform in den Geistes- und Sozialwissenschaften erarbeiten.¹⁰¹⁴

¹⁰⁰⁹ Nach: Bolz, Lisa; Foerster, Sascha: Die Blogplattform der deutschsprachigen Geistes- und Sozialwissenschaften, Workshop, DHd2015, Von Daten zu Erkenntnissen: Workshop II, Graz 23.02.2015, Fol. 2. URL: <http://gams.uni-graz.at/o:dhd2015.w.02>, Stand: 30.07.2015.

¹⁰¹⁰ Nach: Hecker-Stampehl: Bloggen in der Geschichtswissenschaft als Form des Wissenstransfers, 2013, S. 41. Weitere Einblicke in die Blogosphäre bietet bspw.: Mounier: Die Werkstatt öffnen: Geschichtsschreibung in Blogs und Sozialen Medien, 2013.

¹⁰¹¹ Siehe: Hecker-Stampehl: Bloggen in der Geschichtswissenschaft als Form des Wissenstransfers, 2013, S. 38-40.

¹⁰¹² Nach: König, Mareike: Die Entdeckung der Vielfalt: Geschichtsblogs der europäischen Plattform *hypotheses.org*, in: Haber, Peter; Pfanzer, Eva (Hg.): *historyblogosphere*. Bloggen in den Geschichtswissenschaften, München 2013, S. 182.

¹⁰¹³ Ebd., S. 184.

¹⁰¹⁴ Es werden nur wissenschaftliche Blogs aus den Geistes- und Sozialwissenschaften, die in ein wissenschaftliches Projekt eingegliedert sind, akzeptiert und aufgeschaltet.

Jedoch wird auch Kritik an der Verwendung von Blogs für die Wissenschaftskommunikation geübt.¹⁰¹⁵ Blogs würden dem traditionellen Peer Review entgehen, so der wahrscheinlich meistgenannte Vorwurf, und seien deshalb wenig glaub- und vertrauenswürdig.¹⁰¹⁶ Der Neuzeithistoriker Alex Cummings und der Mediävist Jonathan Jarrett sind der Meinung, dass jegliches Einbinden von Review-Verfahren zwei wichtige Eigenschaften von Blogs untergraben: Geschwindigkeit und Verbreitung.¹⁰¹⁷ Blogs können nach Jarrett nur die Wissensgewinnung unterstützen, aber nicht deren Validation ermöglichen. Nach Cummings dienen Blogs als zusätzliches Mittel zum Schreiben, zum Austausch und zur Diskussion, das die bisherigen Publikationsverfahren jedoch nicht ersetzen kann.¹⁰¹⁸ Jarrett und Cummings ist solange beizupflichten, bis verlässliche Repositories aufgebaut und Sicherheitsmerkmale eingeführt sein werden, mit denen auch Blogbeiträge für wissenschaftliche Zwecke verwendet werden können.

Valentin Groebner fragt nach dem Nutzen von Blogs. Er sieht in der *grauen Literatur*¹⁰¹⁹, die er als "schnelles, ironisches, grossmäuliges Zeug [bezeichnet], in dem oft Ideen zum ersten (und häufig auch zum letzten) Mal formuliert wurden"¹⁰²⁰, den direkten Vorläufer digitaler Kommunikationskanäle. Der Kern des Problems sei der Informationsüberfluss, der gefiltert werden müsse, denn es stehe nur eine begrenzte Zeit zum Lesen und Verarbeiten zur Verfügung. Groebner ist zu entgegnen, dass nicht all das 'coole Zeug' zu lesen ist, wie er behauptet¹⁰²¹, sondern es sind die informativen und qualitativ hochstehenden Blogs zu lesen, die vom Nutzer selbst ausgewählt werden können und die von der Fachgemeinschaft empfohlen werden. Recht hat er damit, dass alles, was nicht gelesen wird, wieder in der grauen Unsichtbarkeit des Web verschwindet.

Die Soziologin Sherry Turkle übt Kritik am Schreibstil von Blogs und ähnlichen Publikationsformaten, die den Nutzer mit Informationshäppchen füttern und ihn nicht zu weiterem Nachdenken anregen würden:

¹⁰¹⁵ Eine Replik auf diverse Kritik wie Wissenschaftlichkeit, Akzeptanz, Qualitätssicherung oder Mehrwert gibt: König: Herausforderung für unsere Wissenschaftskultur: Weblogs in den Geisteswissenschaften, 2015.

¹⁰¹⁶ Laut David Nicholas sind junge Forschende eher geneigt, Forschungsressourcen ohne Peer Review (wie bspw. Blogs) zu verwenden, als ältere Kollegen. Siehe: Nicholas, David; Watkinson, Anthony; Jamali, Hamid R. u. a.: Peer review: still king in the digital age, in: *Learned Publishing* 28 (1), 2015, S. 18. DOI: 10.1087/20150104.

¹⁰¹⁷ Siehe: Cummings, Alex Sayf; Jarrett, Jonathan: Only Typing? Informal Writing, Blogging, and the Academy, in: Dougherty, Jack; Nawrotzki, Kristen (Hg.): *Writing history in the digital age*, Ann Arbor 2013 (*Digital humanities*), S. 251.

¹⁰¹⁸ Nach: Ebd., S. 252 f.

¹⁰¹⁹ Graue Literatur: nicht durch kommerzielle Verlage vertrieben und nicht im Buchhandel erhältlich.

¹⁰²⁰ Groebner: *Wissenschaftssprache digital. Die Zukunft von gestern*, 2014, S. 92. Ergänzend: Föhr, Pascal: Rezension zu: Groebner, Valentin: *Wissenschaftssprache digital. Die Zukunft von gestern*, in: *Schweizerische Zeitschrift für Geschichte* 65 (1), 2015, S. 144-145.

¹⁰²¹ Siehe: Groebner: *Wissenschaftssprache digital. Die Zukunft von gestern*, 2014, S. 114.

*"I would feel bereft if because technology wants us to read short, simple stories, we bequeath to our children a world of short, simple stories. What technology makes easy is not always what nurtures the human spirit."*¹⁰²²

Turkles Technologiekritik ist gleichzeitig eine Gesellschaftskritik und sollte auch eine Warnung für die Wissenschaftskommunikation sein, denn mit ausschliesslich kurzen Beiträgen lässt sich keine Geschichte erzählen. Da Blogs aber die Möglichkeit bieten Beiträge zu kommentieren, kann eine rege Diskussion entstehen, in der die Nutzer trotzdem zum Nachdenken angeregt werden.

Blogs stellen eine Publikationsform dar, deren Stellenwert in der wissenschaftlichen Kommunikation stetig zunimmt und bieten viele Möglichkeiten des Informations- und Meinungsaustausches unter Wissenschaftlern. Sie sind bereits in einem kleinen Kreis von Digitalwissenschaftlern anerkannt und könnten, mit entsprechenden Sicherheitsmerkmalen versehen, die Anzahl der Nutzer schnell vergrössern. Man muss nicht bloggen, wie der Titel von David Perrys Post *Be Online Or Be Irrelevant*¹⁰²³ suggeriert, aber es hilft sichtbar(er) zu sein oder zu bleiben.

5.6.2. Schreiben

*"In der digitalen Kultur gehen wir vom Einzelkämpfertum zur Gruppenarbeit über."*¹⁰²⁴ (Clivaz, 2015)

Digitale Publikationsformen ermöglichen sowohl die Nachbearbeitung von bereits veröffentlichten Objekten als auch unmittelbare Dialoge zwischen Autor, Nutzer und anderen Autoren. Visualisierungen, Nichtlinearität und Hypertextualität sowie die technischen Möglichkeiten des kollaborativen Arbeitens verändern den Schreibprozess. Der Autor hat sein Werk zu vernetzen und zu modularisieren.

Bei digitalen Darstellungsformen ist das verwendete Informationssystem mit seinen Möglichkeiten und Limitierungen ausschlaggebend dafür, welche Visualisierungen umgesetzt werden können. Nicht nur für die Bearbeitung von Big Data-Beständen hat eine entsprechende Infrastruktur mit Supercomputern und speziellen Programmen zur Verfügung zu stehen, sondern ebenso auch für einen Fachartikel von wenigen Seiten.

Zudem sind die *charakteristischen Merkmale* eines Medienformats zu beachten und dem Verwendungszweck entsprechend auszuwählen. Beispielsweise ist in traditionellen Qualifi-

¹⁰²² FRONTLINE: Interview: Sherry Turkle, PBS, FRONTLINE, Digital Nation, 02.02.2010, URL: <http://www.pbs.org/wgbh/pages/frontline/digitalnation/interviews/turkle.html>, Stand: 06.10.2015.

¹⁰²³ Perry, Dave: Be Online or Be Irrelevant – AcademHack – Thoughts on Emerging Media and Higher Education, 11.01.2010, URL: <http://academhack.outsidethetext.com/home/2010/be-online-or-be-irrelevant/>, Stand: 07.02.2018.

¹⁰²⁴ Clivaz, Claire, in: Goubet, Fabien: Algorithmen treffen auf Papyrus, in: Horizonte (105), 2015, S. 13.

kationsarbeiten das Schreiben in der Ich-Form verpönt, in einem Blog hingegen ist sie unabdingbar, weil die Subjektivität zum Charakter des Formats gehört.¹⁰²⁵ In den Kommentaren treten als Partizipierende anschliessend verschiedene *Ichs* auf, die jeweils ihre subjektiven Meinungen ausdrücken. In einem anderen Format kann Schreiben Unmittelbarkeit bedeuten, beispielsweise wenn auf einer Plattform jede Eingabe sofort für alle Teilnehmer sichtbar ist.¹⁰²⁶

Bei der Verwendung von digitalen Darstellungsformen ist auch der *Schreibprozess* anzupassen, weil digitale Objekte anders gelesen werden.¹⁰²⁷ Die Vorgehensweise "Wir beginnen vorne und hören hinten auf, wir erzählen eines nach dem anderen"¹⁰²⁸ sollte durch einen hypertextuellen, multimedialen und trans- respektive interdisziplinären Ansatz ersetzt werden.¹⁰²⁹ Ein Digitalisat oder ein Objekt auf einer Webseite zu publizieren, reicht hierbei nicht aus, denn das Objekt hat darüber hinaus die Charakteristika der Hypertextmedien zu berücksichtigen, die sich durch Modularität der informationellen Einheiten, Nichtlinearität, Vernetztheit, Prozessualität, Kollaboration, Lesegewohnheiten, Metadaten und Aktualität definieren.¹⁰³⁰

Digitale informationelle Einheiten sind aufgrund der Erkenntnisse über die neuen Lesegewohnheiten relativ kurz zu halten und als Module zu konzipieren, die flexibel kombiniert werden können. Jedes Objekt ist gleichwertig, das heisst es hat alles zum Verständnis Notwendige zu beinhalten und alles Unnötige oder Weiterführende wird in weitere Module aufgenommen und verlinkt. Damit eine kohärente Abfolge entsteht, ist auf semantisch stringente und argumentativ sinnvolle Verbindungen zwischen den Modulen zu achten. Links sind bewusst zu setzen, um Inhalte und auch Akteure miteinander zu vernetzen und zu integrieren.¹⁰³¹ Dem Nutzer wird damit ein nichtlineares Lesen ermöglicht.

Niels-Oliver Walkowski wirft digitalen Publikationen vor, dass sie "mehr als kombinatorisches Vorgehen denn als Schreibprozess" zu verstehen sind, weil ihre Erstellung "sehr häufig im Zusammenfügen bereits bestehender und anderenorts hergestellter Ressourcen wie Daten, Multimedia-Objekte, Software und auch Texte"¹⁰³² bestehen würde. Tatsächlich ist es bei

¹⁰²⁵ Nach: Schreiner, Julia: Neue (Auf)Schreibsysteme. Verändern Weblogs die Konventionen des geschichtswissenschaftlichen Schreibens?, in: Haber, Peter; Pfanzelter, Eva (Hg.): *historyblogosphere. Bloggen in den Geschichtswissenschaften*, München 2013, S. 89.

¹⁰²⁶ Bspw. bei Etherpad, URL: <http://etherpad.org/>.

¹⁰²⁷ Eine fundierte Einführung bieten bspw.: Dougherty, Jack; Nawrotzki, Kristen (Hg.): *Writing history in the digital age*, Ann Arbor 2013 (Digital humanities). URL: <http://writinghistory.trincoll.edu/>. Vertiefend siehe auch: Haber: *digital past* (unv. Habil.), 2009, S. 202-231.

¹⁰²⁸ Epple: *Verlinkt, vernetzt, verführt – verloren? Innovative Kraft und Gefahren der Online-Historiographie*, 2005, S. 21.

¹⁰²⁹ Nach: Schmale: *Digitale Geschichtswissenschaft*, 2010, S. 123.

¹⁰³⁰ Siehe bspw.: Krameritsch: *Hypertext schreiben*, 2010; Haber: *Geschichte schreiben im Digitalen Zeitalter*, 2011; Thillosen: *Schreiben im Netz: Neue literale Praktiken im Kontext Hochschule*, 2008.

¹⁰³¹ Vgl.: Krameritsch: *Hypertext schreiben*, 2010, S. 84.

¹⁰³² Walkowski: *Text, Denken und E-Science*, 2013, S. 43.

digitalen Publikationen viel einfacher, verschiedene Ressourcen miteinander zu verknüpfen, als bei physischen Publikationen. Dennoch ist der Vorwurf nicht berechtigt. Eine grosse Anzahl an Forschungsblogs enthält meist wertvolle Fachbeiträge und regt mit neuen Ideen die internationale Forschungsdiskussion an.¹⁰³³ Deshalb dürfen sie nicht pauschal als 'Kombination bestehender Objekte' gebrandmarkt werden.

Unter *Schreiben* ist nicht nur das Verfassen von rezipierbaren Texten zu verstehen, sondern auch das *Programmieren*. Zwar können Programmanweisungen als eine hochstrukturalisierte Sprache angesehen werden, der Schreibprozess jedoch gestaltet sich andersartig. Mit einer Programmiersprache werden nur Anweisungen an ein Informationssystem geschrieben, das in einem Verarbeitungsprozess die im Code enthaltenen Befehle ausführt. Das Ergebnis ist anschliessend an einem peripheren Gerät wie einem Bildschirm oder einem Roboterarm sichtbar. Programmieren bedeutet deshalb eine aktive Einflussnahme auf das Ergebnis und setzt deshalb fundierte Kenntnisse der Funktion von elektronischen Informationssystemen und der verwendeten Programmiersprache voraus.

Ein wichtiger Teil des Schreibprozesses ist die Dokumentierung von digitalen Objekten mit *Metadaten*. Damit lassen sich digitale Objekte verlässlich (wieder-)auffinden und identifizieren, ebenso lässt sich deren Überlieferungsgeschichte nachvollziehen. Je mehr Metadaten zur Verfügung stehen, desto besser ist das Objekt für wissenschaftliche Zwecke verwendbar.

Die verwendeten Methoden, Techniken und die ergebnisrelevanten Einzelschritte wie Suchkriterien bei der Auswertung einer Datenbank oder Anpassungen an einen Code sind zu dokumentieren. Wie bei naturwissenschaftlichen Publikationen sind diese Informationen für das Nachvollziehen der Forschungsergebnisse relevant. Damit erleichtert der Wissenschaftler die Überprüfung seiner Erkenntnisse und steigert durch die Transparenz auch seine Reputation.

Das Schreiben kann bei digitalen Objekten auch nach der Publikation weitergeführt werden. Digitale Objekte sind Manuskripte, denn "werden wissenschaftliche Arbeiten digital publiziert, so muss [...] die Vorstellung eines fertigen, definitiven Textes aufgegeben werden."¹⁰³⁴ Ein Groebner'sches 'Stopp' kann zwar nach der Publikation mit einer Quellensicherung durchaus erfolgen, es kann aber im Gegensatz zum gedruckten Text jederzeit auch wieder aufgehoben werden. Inhaltliche Anpassungen, Aktualisieren von Links oder Ergänzungen durch kollaborative Elemente können durch den Autor oder einen Editor auch immer wieder vorgenommen werden. Gerade dieser Umstand macht digitale Publikationen wie Blogs so popu-

¹⁰³³ Siehe bspw. unter hypotheses.org (URL: <http://hypotheses.org/>).

¹⁰³⁴ Wampfler: «online fist». *Geisteswissenschaften als Social Media*, 2013, S. 86.

lär.¹⁰³⁵ Das Projekt *Writing History in the Digital Age* stellt anschaulich dar, wie ein Manuskript durch den Autor und die Mitwirkenden verändert wird, was ein 'Stopp' bewirkt und auch, was dieser nicht bewirkt.¹⁰³⁶ Die Kollaboration und Diskussion unter allen Mitwirkenden hat einerseits die Grenze zwischen Buch und Konferenz aufgehoben, andererseits trug der fruchtbare Dialog zur Schärfung der Thesen bei.¹⁰³⁷

Mit den Möglichkeiten des digital kollaborativen Schreibens, verstanden als gemeinschaftliche Co-Konstruktion eines digitalen Objekts durch mehrere Personen, wird einerseits die Autorschaft vergrößert, andererseits sind damit auch verschiedene neuartige Anforderungen verbunden.¹⁰³⁸ Das Ziel sollte stets sein, dass sich die Summe der Beiträge zu einem besseren Ganzen summieren. Projekte wie die von den Erziehungswissenschaftlern Kristen Nawrotzki und Jack Dougherty als 'Experiment' bezeichnete Publikation *Writing History in the Digital Age*¹⁰³⁹ oder die hybride Publikation *historyblogosphere*¹⁰⁴⁰ von Pfanzer und Haber zeigen auf, dass das Erarbeiten und Kritisieren mit digitalen Kollaborationsmitteln viel transparenter und kollaborativer gestaltet werden können als mit traditionellen Verfahren. Die Autoren hatten während des Erarbeitens ihrer Beiträge nicht nur die Möglichkeit, Anmerkungen und Vorschläge der anderen Autoren in ihren eigenen Beitrag einzuarbeiten, sondern können selbst bei den anderen Autoren mitarbeiten. Damit verweben sich die Beiträge viel besser, womit das Ganze wertvoller wird als die Summe der einzelnen Beiträge.

Wie die kollaborative Arbeit an der Publikation selbst, können externe kritische Beiträge öffentlich einsehbar publiziert werden. Der Autor kann in einem *Open Peer Review*-Verfahren¹⁰⁴¹ auf einer entsprechenden Plattform während allen Phasen des Erstellungsprozesses¹⁰⁴² einem bestimmten oder offenen Nutzerkreis erlauben, sein Manuskript zu

¹⁰³⁵ Nach: Fitzpatrick, Kathleen: The Digital Future of Authorship: Rethinking Originality, in: *Culture Machine* 12 (0), 18.02.2011, S. 11. URL: <http://www.culturemachine.net/>, The Digital Future of Authorship: Rethinking Originality, Stand: 17.04.2013.

¹⁰³⁶ Vgl.: Dougherty; Nawrotzki (Hg.): *Writing history in the digital age*, 2013, insb. S. 264-274. Zusätzlich URL: <http://writinghistory.trincoll.edu/>.

¹⁰³⁷ Nach Charlotte Rochez in URL: <http://writinghistory.trincoll.edu/conclusions-2012-spring/>.

¹⁰³⁸ Siehe dazu: Burckhardt, Daniel; Schiel, Juliane: *Kollaboratives Schreiben*, in: Gasteiner, Martin; Haber, Peter (Hg.): *Digitale Arbeitstechniken: für die Geistes- und Kulturwissenschaften*, Wien, Köln, Weimar 2010, S. 83-95.

¹⁰³⁹ Dougherty; Nawrotzki (Hg.): *Writing history in the digital age*, 2013.

¹⁰⁴⁰ Haber; Pfanzer (Hg.): *historyblogosphere*. *Bloggen in den Geschichtswissenschaften*, 2013.

¹⁰⁴¹ Peer: Fachkollege. *Review: Besprechung und Überprüfung*. Siehe weiterführend bspw.: Nentwich, Michael; König, René: *Peer Review 2.0: Herausforderungen und Chancen der wissenschaftlichen Qualitätskontrolle im Zeitalter der Cyber-Wissenschaft*, in: Gasteiner, Martin; Haber, Peter (Hg.): *Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften*, Wien, Köln, Weimar 2010, S. 143-163; Fahrmeir, Andreas: *Bloggen und Open Peer Review in der Geschichtswissenschaft: Chance oder Sackgasse?*, in: Haber, Peter; Pfanzer, Eva (Hg.): *historyblogosphere*. *Bloggen in den Geschichtswissenschaften*, München 2013, S. 23-35, insb. 30-34; Haber: *digital past* (unv. Habil.), 2009, S. 260-270.

¹⁰⁴² Nicht nur wenn eine Vorabfassung (preprint) publiziert wird, wie bspw. Lengwiler oder Nentwich suggerieren. Vgl.: Lengwiler: *Praxisbuch Geschichte*, 2011, S. 269; Nentwich; König: *Peer Review 2.0*, 2010.

lesen und mit Anmerkungen, Verweisen, Ergänzungen und Korrekturvorschlägen zu versehen.¹⁰⁴³ Die Peers können die Arbeit unmittelbar mitverfolgen und ihre Anmerkungen und Kommentare direkt einbringen – auch nach der definitiven Publikation. Das beim traditionellen Review oft intransparente Verfahren sowie dabei allenfalls entstehende "Befangenheiten und tendenziöse bzw. bewusst falsche Besprechungen wissenschaftlicher Literatur"¹⁰⁴⁴ werden zu Gunsten einer offenen, nicht anonymisierten *Kritikkultur* einer Fachgemeinschaft abgeschafft, was zu vielfältigeren Beiträgen und anregenden Diskussionen führen kann. Damit können auch nicht publizierbare Arbeiten besprochen werden, die thematisch durch keine Fachzeitschrift abgedeckt sind. Wichtig ist auch hier, dass Metadaten zur Verfügung gestellt werden, damit die Beiträge mittels Researcher Driven Archiving als Quelle verwendet werden können.

Auf solchen Kollaborationsplattformen wird die Art und Weise der Zusammenarbeit klar festgelegt. Durch die Vergabe von Lese- und Schreibrechten wird der Zugriff geregelt, weil sonst alle Autoren das Objekt gleichzeitig verwenden und manipulieren könnten. Änderungen werden protokolliert und sind somit nachvollziehbar. Mit einem Versionsmanagement kann zudem eine Vorversion jederzeit wiederhergestellt werden. Schliesslich muss die Autorschaft festgelegt, Metadaten eingefügt und das Objekt in einer adäquaten Form publiziert werden. Bei kollaborativem Schreiben ist ein klares Rechte- und Versionsmanagement nötig.

Andreas Fickers sieht das grosse Potenzial für die zukünftige Geschichtsschreibung in der umsichtigen Kombination von verschiedenen Erzählangeboten, die über verschiedene Medienformen verteilt sind. Dabei beruft er sich auf den Medienwissenschaftler Henry Jenkins, der das "transmedia storytelling"¹⁰⁴⁵ propagiert. Darunter versteht er das Erzählen einer Geschichte mit unterschiedlichen Medien, die aber unterschiedliche Teile der Geschichte erzählen. Idealerweise werden die Medienformen so eingesetzt, dass zur Geschichte (nur) das beitragen, wofür sie am Besten geeignet sind.¹⁰⁴⁶ Jenkins erklärt dies anhand der Geschichte der Filmtrilogie *The Matrix* (1999) / *The Matrix Reloaded* (2003) / *The Matrix Revolutions* (2003), die erst dann einen sinnvollen Handlungsstrang ergeben, wenn das Computerspiel *Enter the Matrix* gespielt wird. Adaptiert auf die Erzählung und Darstellung von historischen Werken, würde dies eine grundlegend neue Art des Schreibens über verschiedene Medienformen bedeuten, die sehr wahrscheinlich nur in Kollaboration mit Spezialisten erfolgen könnte. Transmedia Storytelling steckt für wissenschaftliche Arbeiten

¹⁰⁴³ Peer Review Plattformen bspw.: Publons (URL: <https://publons.com/>) oder PubPeer (URL: <https://pubpeer.com/>).

¹⁰⁴⁴ Haber, Peter; Pfanzer, Eva; Schreiner, Julia: Vorwort, in: Haber, Peter; Pfanzer, Eva (Hg.): *historyblogosphere. Bloggen in den Geschichtswissenschaften*, München 2013, S. 51-59.

¹⁰⁴⁵ Siehe: Jenkins, Henry: *Convergence Culture: Where Old and New Media Collide*, New York and London 2006, S. 93-130.

¹⁰⁴⁶ Vgl.: Ebd., S. 95f.

noch in den Kinderschuhen, ist aber ein höchst interessanter Ansatz, wie digitale Objekte in digitaler Form publiziert werden könnten.

Die Arbeit ist mit dem *digitalen Schreiben* nicht zwingend beendet. Beispielsweise ermöglicht eine Kommentarfunktion auch nachträglich Dialoge zwischen Autor(en) und Nutzern. Umfangreiche Nacharbeit ist nötig, wenn ein Transfer aus dem Digitalen ins Analoge vorgesehen ist. Einige Blogger haben zum Beispiel so viel Aufmerksamkeit und Interesse erzeugt, dass ein Verlag mit ihnen ein Buchprojekt durchgeführt hat. Der Blogger kann aber aufgrund der sehr unterschiedlichen Medienformen nicht einfach seine Blogbeiträge unredigiert ausdrucken, sondern muss sie umschreiben und sie aus der hypertextuellen Nichtlinearität in ein lineares, physisches Format überführen. Dies bedeutet die Rückkehr in die traditionelle Art von Schreiben.

Der Schreiber hatte sich schon bei traditionell physischen Publikationsformen mit den medialen Eigenschaften auseinanderzusetzen. Nachbearbeitungen am Objekt sind dort jedoch nicht möglich und auch der Dialog mit den Nutzern oder kollaboratives Arbeiten ist relativ umständlich sowie zeitintensiv. Digitale Publikationsformen verlangen vom Autor neue Schreibprozesse, die nichtlineares Lesen, Vernetzungen, Visualisierungen und Dokumentation der Auswertungsmethodik berücksichtigen. Dialoge mit Nutzern und kollaboratives Arbeiten mit anderen Autoren wie Nutzern sind nun unmittelbar und einfach möglich.

5.6.3. Belege

*"Print citation itself should not be regarded as a scholarly promised land."*¹⁰⁴⁷
(Blaney, 2012)

Wissenschaftliches Arbeiten bedeutet auch, dass das, was geschrieben wurde, auch zu belegen ist. Belege für physisch vorhandene Werke werden zwar nach anerkannten Konventionen mittels Zitation, Literaturverzeichnis und Fussnotenverweisen erstellt, sind aber nicht so weit standardisiert wie gerne gewünscht.¹⁰⁴⁸ Die Erweiterung dieser Konventionen auf digitale Quellen ist problematisch, weil deren Eigenschaften und die eines digitalen Informationssystems nicht beachtet werden.¹⁰⁴⁹ Informationen im Internet sind grundsätzlich flüchtig, weshalb der Zugriff unter einer angegebenen Netzwerkadresse schon nach einer kurzen Zeit verunmöglich sein kann. Die Informationen können auch aus dem Deep Web stammen und

¹⁰⁴⁷ Blaney, Jonathan: The Problem of Citation in the Digital Humanities, Proceedings of the Digital Humanities Congress 2012. Studies in the Digital Humanities, Sheffield; HRI Digital, 2014, URL: <http://www.hrionline.ac.uk/openbook/chapter/dhc2012-blaney>, Stand: 16.08.2015.

¹⁰⁴⁸ Ein einheitlicher Standard hat sich bisher nicht durchsetzen können. Das in dieser Arbeit verwendete Literaturverwaltungsprogramm Zotero (URL: <https://www.zotero.org/>) listet 7'824 verschiedene Zitationsstile auf, davon 62 für den Fachbereich History (URL: <https://www.zotero.org/styles>, Stand: 16.08.2015).

¹⁰⁴⁹ Vgl.: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 22-30.

damit nur einem bestimmten Nutzerkreis zugänglich sein. Identische Informationen können an verschiedenen Speicherorten gleichzeitig vorliegen, oder ein Objekt liegt an einem einzigen Speicherort vor und ist in verschiedene Publikationsplattformen integriert. Eine Änderung des Objekts wirkt sich automatisch auf alle damit verknüpften Plattformen aus.

Für eine Ressource im Web wird hauptsächlich folgende Zitierform als Belegkette empfohlen und verwendet:

Autor(schaft): Titel. URL/URI. Zeitpunkt des Zugriffs auf die Information.

Nach Margulies ist diese Struktur aus drei Gründen problematisch. Das erste Problem besteht in der Verwendung einer URL/URI, die auf einen *Speicherort* innerhalb des Web verweist. Wird die Ressource gelöscht oder verschoben, ist die Angabe nutzlos – obwohl sie unter Umständen im Internet an einem anderen Ort noch verfügbar sein könnte.

Es stehen noch wenige TDR von Institutionen mit gesetzlichem Archivierungsauftrag wie die Bibliothèque nationale de France, die die Adresspersistenz einer digitalen Forschungsressource garantieren, zur Verfügung. Selbst Zitationen aus physischen Quellen sind nicht immer persistent, insbesondere bei Einzelstücken in Archiven. Wenn beispielsweise das Archiv aufgelöst und in ein anderes integriert wird, ändern sich auch die Angaben.¹⁰⁵⁰

Das zweite Problem ist der *Zeitpunkt* des Zugriffs. Nahezu alle Zitationsvorschriften fordern die Angabe des Zeitpunkts, an dem das Dokument heruntergeladen worden ist. Die internationale Norm ISO 690:2010 empfiehlt sogar minutengenaue Angaben.¹⁰⁵¹ Auch in einem Ratgeber zum wissenschaftlich korrekten Arbeiten wird für Zitationsnachweise digitaler Quellen explizit die Angabe des Abfragedaums gefordert.¹⁰⁵²

Der Abfragezeitpunkt kann zwar tatsächlich einen Hinweis auf die verwendete Version geben. Aber wenn die Quelle verschoben, gelöscht oder manipuliert worden ist, sind die Angaben wertlos, denn die verwendete Version kann dadurch nicht mehr aufgerufen werden. Die Zeitangabe im Web ist mit der einer physischen Publikation nicht vergleichbar, denn sie stellt eine scheinbare Sicherheit dar und wird nur mangels Alternativen verwendet, worauf die Sprach- und Kommunikationswissenschaftler Jens Runkehl und Torsten Siever bereits früher hingewiesen haben.¹⁰⁵³ Einen Ausdruck als Zitatbeleg anzufertigen, ist zwar ein pragmatischer Ansatz, aber angesichts des Verlusts der Digitalität, der Möglichkeit der Manipulation

¹⁰⁵⁰ Bei professionell geführten Archiven wird auch dann noch der Zugriff bspw. durch Referenzlisten gewährleistet.

¹⁰⁵¹ Nach: ISO 690:2010, S. 19 f. Siehe URL: <http://www.iso.org/>.

¹⁰⁵² Wytrzens, Hans Karl; Schauppenlehner-Kloyber, Elisabeth; Sieghardt, Monika u. a.: *Wissenschaftliches Arbeiten: eine Einführung*, Wien 2014, S. 115. Ähnlich auch: Preissner, Andreas: *Wissenschaftliches Arbeiten: Internet nutzen – Text erstellen – Überblick behalten*, München 2012, S. 105.

¹⁰⁵³ Vgl.: Runkehl; Siever: *Zitieren und Belegen*, 2010, S. 96 f. und 138.

durch den Nutzer und die Aufbewahrung in einem für andere nicht zugänglichen, persönlichen Archiv nicht sinnvoll.

Die vielfältigen *Formen der Darstellung* von Inhalten in Dateien oder digitalen Informationssystemen bergen das dritte Problem. Wird beispielsweise ein Text von einer Webseite zitiert, können keine Seitenzahlen angegeben werden, und nur auf wenigen Webseiten sind die Absätze mit Nummern versehen, die in die bibliographischen Angaben übernommen werden könnten.¹⁰⁵⁴ Zudem wird derselbe Inhalt in verschiedenen Browsern unterschiedlich dargestellt. Mit einer Volltextsuche können zitierte Stellen innerhalb kurzer Zeit gefunden werden. Diese Möglichkeit besteht nur bei Digitalisaten oder bei e-Publikationen, die entsprechend dafür aufbereitet wurden.

Aufgrund der vielfältigen und heterogenen Formen der Darstellung und Kommunikation in digitalen Medien wie Twitter oder Mind-Maps sind konkrete Zitieranweisungen für jedes einzelne Programm nicht sinnvoll, wie sie beispielsweise der Betriebswirtschaftler Andreas Preissner für Twitter empfiehlt.¹⁰⁵⁵ Dem Zitierenden obliegt die Verantwortung, dass in der Belegkette alle nötigen Angaben zum Wiederauffinden der verwendeten digitalen Forschungsressource vorhanden sind. Style-Guides und Style-Vorlagen in Literaturverwaltungssystemen, Lehrmitteln und Studienrichtlinien akademischer Institutionen sowie Fachliteratur zum wissenschaftlichen Arbeiten geben dazu Hinweise.¹⁰⁵⁶

Lösungsansatz Belegkette

Als praktische Anwendung obiger Ausführungen schlage ich die Anpassung der Belegkette für digitale Objekte vor, mit folgenden Merkmalen:

- *Cite What You Used – CWYU*:¹⁰⁵⁷ Wenn digitale Forschungsressourcen verwendet werden, sind diese aus Transparenzgründen entsprechend auszuweisen.¹⁰⁵⁸ Insbesondere bei Digitalisaten ist die Angabe des physischen Äquivalents nur dann zulässig, wenn es tatsächlich konsultiert wurde. Wenn nur die digitale Version eingesehen wurde, ist dies offenzulegen. Lange oder komplexe (Hyperlink-)Angaben sind hierbei kein Hindernis, weshalb auch URL-Shortener aufzulösen sind.

¹⁰⁵⁴ Bspw. das Online-Journal 'zeitenblicke', URL: <http://www.zeitenblicke.de/>.

¹⁰⁵⁵ Twitter Zitatformat: [Nachname], [Vorname]; [@Twittername]: ["Text des Tweets"], Tweet vom [Datum und Uhrzeit des Tweets]. Siehe: Preissner: Wissenschaftliches Arbeiten, 2012, S. 108.

¹⁰⁵⁶ Siehe zu Spezialformen des Bibliographierens und Zitierens bspw.: Runkehl; Siever: Zitieren und Belegen, 2010, S. 93-108.

¹⁰⁵⁷ Nach: Blaney: The Problem of Citation in the Digital Humanities, 2014.

¹⁰⁵⁸ Bspw. ist zwischen Buch, Digitalisat des Buchs und Online-Publikation des Buchs zu unterscheiden. Noch nicht in allen Literaturverwaltungsprogrammen umgesetzt (bspw. Zotero, Version 4.0.28.1).

- *Autor(schaft)*: Es sind genaue Angaben zum Autor oder dessen Pseudonym anzugeben. Sind keine Informationen vorhanden, ist der Autor mit entsprechendem Hinweis als anonym zu bezeichnen.
- *Titel*: Der Titel der Quelle ist anzugeben. Bei nicht vorhandenem oder unbekanntem Titel ist ein als solcher kenntlich gemachter Titel vom Untersuchenden selbst zu setzen.
- *Persistenter Identifikator oder Speicherort*: Es sollte ein persistenter Identifikator verwendet werden. Solange dies nicht möglich ist, ist der Speicherort der Quelle möglichst genau anzugeben. Das umfasst den gesamten Pfad: Titel der 1. Einheit / Titel der 2. Einheit / ... / Titel der (n-1). Einheit / Titel der n. Einheit.¹⁰⁵⁹ Dieses Vorgehen hat den Vorteil, dass unabhängig von der Repräsentation des Objekts auf den Inhalt selbst verwiesen wird, dass auch in nicht hierarchischen Texten navigiert und auf physisch nicht publizierte Quellen verwiesen werden kann.
- *Link-Check-Element*: Mit diesem neuen Element zeigt der Zitierende auf, welche Hyperlinks des Objekts und wie tief diese überprüft wurden. Diese Analyse dient zur Feststellung von den Inhalt eventuell ergänzenden Informationen in externen, verlinkten Objekten.
- *Version und Datum*: Als zusätzliche Angaben können Version und Datum einer Quelle angegeben werden, wenn verlässliche Angaben vorhanden sind. Auf die Angabe mehrerer Daten oder eine minutengenaue Zeitangabe kann verzichtet werden.
- *Sammlung*: Wird eine Forschungsressource aus einer TDR-Sammlung verwendet, so ist neben dem Identifikator oder Speicherort auch die zugehörige Sammlung anzugeben, um die Ressource zu (re)kontextualisieren.
- *Seitenangaben*: Auf Seitenangaben kann in elektronisch durchsuchbaren Objekten wie HTML-Webseiten verzichtet werden. Sind Absätze nummeriert oder werden in bestimmten Formaten Seitenzahlen angegeben, können solche übernommen werden.
- *Kreativität*: Es gibt Formate (und es werden immer wieder solche entstehen), für die noch keine Konventionen bestehen. Der Zitierende hat nach bestem Wissen und Gewissen und im Sinne wissenschaftlicher Redlichkeit eine nachvollziehbare Belegkette nach den Richtlinien des wissenschaftlichen Arbeitens zu erstellen.

¹⁰⁵⁹ Vgl.: Margulies: Digitale Daten als Quelle der Geschichtswissenschaft, 2009, S. 27.

Daraus ergibt sich folgende Belegkette, die entsprechend der vorhandenen Informationen möglichst vollständig angegeben werden sollte:

Autor, Titel, Identifikator/Speicherort, Link-Check, Version, Datum, Sammlung, Seitenangaben.

Die Belegkette selbst unterscheidet sich nicht wesentlich von der traditionellen Zitierweise. Allerdings sind die zugehörigen Informationen teilweise mit anderen Methoden zu eruieren und zudem eingehender zu verifizieren. Die MIS sollte alle diese Angaben enthalten, um die Erstellung eines Anmerkungsapparats zu erleichtern.

5.6.4. Visualisieren

Mit einer visuellen Darstellung werden Aussagen reduziert und symbolisch dargestellt. Digitale Visualisierungen erlauben nicht nur die direkte Darstellung von Daten aus verschiedenen Quellen, sondern auch deren unmittelbare Aktualisierung und dynamische Betrachtung auf verschiedenen Ebenen.

Die Kernidee der Visualisierung ist das Abbilden von Eigenschaften eines Datensets in einer visuellen Darstellung.¹⁰⁶⁰ Visualisierungen sind sowohl Mittel des Forschens und Erkennens als auch Ausdrucksmittel zur Veranschaulichung von Erkenntnissen, die leicht verständlich und intuitiv nachvollziehbar sind.¹⁰⁶¹ Sie basieren nicht zwingend auf digitalen Daten und computerunterstützten Auswertungen, denn Visualisierungen sind bereits vor dem Computational Turn erfolgreich entwickelt und eingesetzt worden wie beispielsweise Harry Becks *Tube Map* der Londoner U-Bahn anschaulich beweist. Seit dem Aufkommen des Forschungsfelds Digital Humanities werden in der Geisteswissenschaft Visualisierungstechniken aber nicht mehr nur als Hilfsmittel, sondern als zentrales methodisches Element genutzt.¹⁰⁶² Beispielsweise sind Netzwerkanalysen einer grossen Anzahl beteiligter Personen ohne eine visuelle Darstellung kaum mehr erfassbar.¹⁰⁶³

¹⁰⁶⁰ Hier und folgend hauptsächlich nach: Manovich, Lev: What is Visualization?, Data Visualisation, 09.11.2010, URL: http://manovich.net/content/04-projects/064-what-is-visualization/61_article_2010.pdf, Stand: 06.05.2018.

¹⁰⁶¹ Nach: Burkhardt: Digitale Datenbanken, 2015, S. 317.

¹⁰⁶² Vgl.: Puschmann, Cornelius: Visualization in the Digital Humanities, 16.08.2010. URL: <http://de.slideshare.net/coffee001/visualization-in-the-digital-humanities>, Stand: 11.11.2013. Beispiel für Visualisierungsprojekte in den Geisteswissenschaften: Wissenschaftliche Kommunikations-Infrastruktur (WissKI) der Universität Erlangen. URL: <http://wiss-ki.eu/>.

¹⁰⁶³ Zur Visualisierung historischer Quellen siehe bspw.: Düring, Marten: From Hermeneutics to Data to Networks: Data Extraction and Network Visualization of Historical Sources, Programming Historian, 18.02.2015, URL: <http://programminghistorian.org/lessons/creating-network-diagrams-from-historical-sources.html>, Stand: 17.06.2015. Beispiele für Visualisierungen aufgrund historischer Daten: Repertorium Academicum Germanium (URL: <http://www.rag-online.org/>) oder Repertorium Vernice Time Machine (URL: <http://vtm.epfl.ch/>).

Bei einer statischen oder dynamischen Visualisierung¹⁰⁶⁴ werden die zwei Schlüsseltechniken *Reduktion* und *Raumverhältnisse* eingesetzt. Aussagen zu entdeckten Datenmustern oder -strukturen sowie Beziehungen einzelner Objekte untereinander werden mittels graphischer Grundformen wie Pfeilen oder Linien dargestellt und somit stark reduziert. Um wichtige Unterschiede zwischen gleichartigen Daten hervorzuheben, werden Raumverhältnisse wie Position, Abstände und Grösse der Datensymbole verwendet.

Der Medientheoretiker Lev Manovich ist der Ansicht, dass mit der Weiterentwicklung von Informationssystemen vermehrt auf Reduktion verzichtet werden kann. Mit "*direct visualization or visualization without reduction*"¹⁰⁶⁵ werden die vorhandenen Daten nur neu angeordnet und mit Funktionen wie dem Heraus- und Hineinzoomen sowie Neufokussieren können verschiedene Datenebenen in derselben Visualisierung erreicht werden. Manovich weist auch darauf hin, dass die *Direct Visualization* nicht neu ist, sondern dass beispielsweise die *Tag Cloud*¹⁰⁶⁶ mit einem Register in einem Buch auch als eine solche Visualisierung verstanden werden kann. Während hingegen bei physischen Visualisierungen normalerweise zuerst die vorhandenen Daten quantifiziert werden mussten, bevor sie visualisiert werden konnten, ist bei einer digitalen Visualisierung eine direkte Darstellung ohne Quantifizierung möglich.¹⁰⁶⁷ Die digitale Visualisierung unterscheidet sich aber von der physischen durch die Möglichkeit der Nutzung und Vernetzung externer Datenbestände sowie des kontrollierten Einsatzes der Parameter durch den Nutzer. Ich bin gleicher Meinung wie Manovich, weise aber darauf hin, dass auch er die Reduktion weiterhin verwendet. Es ist stark vom Datenbestand abhängig ob auf eine Reduktion verzichtet werden kann, denn insbesondere bei Big Data-Beständen können die zu visualisierenden Ergebnisse so komplex und umfangreich sein, dass eine Reduktion erfolgen muss.

Für eine erfolgreiche Visualisierung ist nicht nur die Wahl der Reduktion ausschlaggebend, sondern auch die Darstellung der Grössenverhältnisse, die sich nach den Prinzipien der *Nähe, Ähnlichkeit, Geschlossenheit, Kontinuität und Symmetrie* richten.¹⁰⁶⁸ Erst damit werden die Beziehungen der Daten zueinander und deren Abgrenzung voneinander klar. Damit der Leser¹⁰⁶⁹ einer graphischen Darstellung die Inhalte sofort erfassen kann, sollten die vom Informationswissenschaftler Edward Tufte formulierten drei Prinzipien für eine exzellente Visualisierung beachtet werden:

¹⁰⁶⁴ Beispiele für Visualisierungen siehe URL: <http://www.visualcomplexity.com/>.

¹⁰⁶⁵ Manovich: *What is Visualization?*, 2010, S. 12 f. Kursiver Schriftschnitt im Original fett.

¹⁰⁶⁶ *Tag Cloud*: Stichwörter werden flächig und in verschiedenen Grössen angeordnet, je öfter ein Wort in einem definierten Wortbestand vorkommt, desto grösser wird es dargestellt.

¹⁰⁶⁷ Vgl.: Manovich: *What is Visualization?*, 2010, S. 14.

¹⁰⁶⁸ Nach: Cristofoli, Pascal: *Principes et usages des dessins de réseaux en SHS*, in: Natale, Enrico; Sibille, Christiane; Chacherau, Nicholas u. a. (Hg.): *Visualisierung von Daten in der Geschichtswissenschaft / La Visualisation des données en histoire*, Zürich 2015 (Geschichte und Informatik), S. 28.

¹⁰⁶⁹ Der Leser ist von einem Betrachter zu unterscheiden, der nur die visuelle Gestaltung und nicht die inhaltliche Aussage der Visualisierung untersucht.

"Graphical excellence is the well-designed presentation of interesting data – a matter of substance, of statistics, and of design.

Graphical excellence consists of complex ideas communicated with clarity, precision, and efficiency.

Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space."¹⁰⁷⁰

Visualisierungen unterliegen allerdings gesellschaftlichen Rezeptionskonventionen, die bei der Erstellung zu berücksichtigen sind, um Missverständnissen vorzubeugen. Die verwendeten Parameter sind, im Idealfall mit dem zugrundeliegenden Datenbestand, zu dokumentieren und zu publizieren, um den Nachvollzug der visualisierten Aussagen zu gewährleisten und um diese eventuell nach einer sich veränderten, aktuelleren oder andersartigen Rezeptionskonvention neu erstellen zu können. Parameter sind für die Erstellung der Visualisierung zentral und können Aussagen unter Umständen verzerren oder verfälschen. Es ist deshalb eminent wichtig, dass auch Visualisierungen mit Metadaten dokumentiert werden.

5.6.5. Open Access und Open Data

*"Access to research data was and is a central issue in the humanities."*¹⁰⁷¹ (Boonstra, Breure und Doorn, 2004)

Open Access und Open Data sind zwei sehr ähnliche Konzepte: das erste will den freien Zugriff auf Publikationen, das zweite den freien Zugriff und die freie Verwendung von Datenbeständen sicherstellen. Für beide wird argumentiert, dass Publikationen oder Daten, die von der öffentlichen Hand finanziert wurden, nicht privatwirtschaftlich vermarktet werden dürfen. Open Access und Open Data haben sich mittlerweile etablieren können und werden von immer mehr Wissenschaftlern, Institutionen und Behörden unterstützt, nur von den Publikationsgesellschaften werden sie nach wie vor bekämpft. Für die freie Forschungsgesellschaft sind Open Access und Open Data jedoch zentral und werden hoffentlich noch mehr Anhänger gewinnen.

Im Jahre 2002 definierte die *Budapest Open Access Initiative*, ein Zusammenschluss internationaler Wissenschaftler, den Begriff *Open Access (OA)* und verband damit auch Forderungen an die Wissenschaft und Privatwirtschaft. Zehn Jahre später wurde die Definition bestätigt:¹⁰⁷²

¹⁰⁷⁰ Tufte, Edward Rolf: *The visual display of quantitative information*, Cheshire, Connecticut 1984, S. 51. URL: <http://www.colorado.edu/UCB/AcademicAffairs/ArtsSciences/geography/foote/maps/assign/reading/TufteCoversheet.pdf>, Stand: 06.10.2015. Schriftschnitt im Original invers.

¹⁰⁷¹ Boonstra; Breure; Doorn: *Past, present and future of historical information science*, 2004, S. 9.

¹⁰⁷² Chan, Leslie; Cuplinskas, Darius; Eisen, Michael u. a.: *Read the Budapest Open Access Initiative*, Budapest Open Access Initiative, 14.02.2002, URL: <http://www.budapestopenaccessinitiative.org/>,

"By 'open access' to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited."

"'Open Access' meint, dass [= Peer-Review-Fachliteratur] kostenfrei und öffentlich im Internet zugänglich sein sollte, sodass Interessenten die Volltexte lesen, herunterladen, kopieren, verteilen, drucken, in ihnen suchen, auf sie verweisen und sie auch sonst auf jede denkbare legale Weise benutzen können, ohne finanzielle, gesetzliche oder technische Barrieren jenseits von denen, die mit dem Internet-Zugang selbst verbunden sind. In allen Fragen des Wiederabdrucks und der Verteilung und in allen Fragen des Copyrights überhaupt sollte die einzige Einschränkung darin bestehen, den Autoren Kontrolle über ihre Arbeit zu belassen und deren Recht zu sichern, dass ihre Arbeit angemessen anerkannt und zitiert wird."

Unter Fachliteratur werden dabei Publikationen verstanden, die die traditionellen Regeln wissenschaftlicher Texte wie der Nachvollziehbarkeit durch einen Fussnotenapparat einhalten.¹⁰⁷³

Verschiedene Plattformen wie das bereits erwähnte *arXiv.org*¹⁰⁷⁴ bieten Publikationsdienstleistungen für Wissenschaftler. Bei arXiv besteht allerdings das Problem, dass die Authentizität der eingereichten Dokumente nicht nachgewiesen werden kann, denn die Registration verlangt nur eine gültige E-Mail-Adresse und keine verifizierte Identifikation.¹⁰⁷⁵ Die hohe Vertrauenswürdigkeit oder Reputation dieser Publikationsplattform gibt jedoch wichtige, das Objekt authentifizierende Hinweise bei der quellenkritischen Überprüfung. Der Bibliothekar Jeffrey Beall verfolgte seit 2008 die Entwicklungen von Online OA-Journalen auf seinem Blog *Scholarly Open Access*¹⁰⁷⁶ und bewertet diese auch nach ihrer Vertrauenswürdigkeit. Seine – allerdings von verschiedenen Seiten kritisierte und mittlerweile nicht mehr frei verfügbare – Liste von fragwürdigen Online-Verlegern weist darauf hin, welche Online-Publikationen wissenschaftlichen Standards eventuell nicht genügen. Eine kritische Haltung fordert auch der *Nature*-Journalist Declan Butler, der in seinem kurzen Artikel *The Dark Side Of Publishing* nicht nur die Probleme mit Online-Verlegern beschreibt,

Stand: 20.03.2015. Dt. Übersetzung: Budapest Open Access Initiative; German Translation, Budapest Open Access Initiative, 2012, URL: <http://www.budapestopenaccessinitiative.org/boai-10-translations/german-translation>, Stand: 06.07.2015.

¹⁰⁷³ Siehe: Hecker-Stampehl: Bloggen in der Geschichtswissenschaft als Form des Wissenstransfers, 2013, S. 40. Weiterführend zu Open Access bspw.: Mruck, Katja; Gersmann, Gudrun (Hg.): Neue Medien in den Sozial-, Geistes- und Kulturwissenschaften: Elektronische Publizieren und Open Access: Stand und Perspektiven – Sonderheft, in: *Historical Social Research* 29 (1), 2004.

¹⁰⁷⁴ URL: <http://arxiv.org/>.

¹⁰⁷⁵ Test am 06.07.2015 durch den Autor.

¹⁰⁷⁶ Beall schloss den Blog anfangs 2017. Veraltete URL: <http://scholarlyoa.com/>.

sondern auch in einer Checkliste zusammenfasst, wie vertrauenswürdige Verleger erkannt werden können.¹⁰⁷⁷ Beall und Butler bestätigen meine in den vorhergehenden Ausführungen mehrfach genannte Aussage, dass für die sichere, dauerhafte Aufbewahrung und Archivierung von digitalen Objekten, hier Publikationen, nur staatlich beauftragte Institutionen infrage kommen dürfen.¹⁰⁷⁸ Verlässlicher und qualitativ hochstehend sind die im *Directory of Open Access Journals*¹⁰⁷⁹ eingetragene Journale. Das Verzeichnis gilt als qualitätsgeprüftes Auskunftsinstrument.

Open Data (OD) verfolgt die gleichen Ziele wie Open Access, bezieht sich aber auf den freien Zugang und die freie Nutzung von (Forschungs-)Daten: "Daten öffentlich, frei verfügbar und nutzbar zu machen für mehr Transparenz, Innovation und Effizienz"¹⁰⁸⁰, so die Maxime. Anders ausgedrückt: "data that anyone can access, use and share."¹⁰⁸¹

Daten sind wertvoll geworden und sind entsprechend aufzubewahren, auszutauschen und zu verwalten, was ein entsprechendes Data Handling-Konzept erfordert. Diese Tätigkeiten werden laut dem Wissenschaftsforscher Christoph Hoffmann anscheinend noch immer als minderwertig angesehen, obwohl sie für die Weiternutzung von Daten eminent wichtig sind.¹⁰⁸² Die Veröffentlichung von Daten als Open Data bedarf guter Vorbereitung, vor allem im Bereich der Wahl des Datenformats und der Infrastruktur, damit auch die dauerhafte Verfügbarkeit sichergestellt werden kann. Zudem sollten die Daten so strukturiert sein, dass sie möglichst einfach weiterverwendet werden können. Das Ziel sollte sein, dass die Daten in einem Netzwerk untereinander verlinkt (sog. *Linked Open Data*) und in einem zukünftigen semantischen Web¹⁰⁸³ genutzt werden können. Der Physiker und Informatiker Tim Berners-Lee hat dazu das *5-Star Deployment Scheme* entwickelt, in dem die Vernetzung von Daten hierarchisch nach ihrer Weiternutzbarkeit aufgezeigt wird.¹⁰⁸⁴ Ergänzend ist der Vorschlag von Daniel Falster, Richard FitzJohn, Matthew Pennell und William Cornwell, die mit ihrem

¹⁰⁷⁷ Siehe: Butler, Declan: Investigating journals: The dark side of publishing, in: Nature 495 (7442), 27.03.2013, S. 435. DOI: 10.1038/495433a.

¹⁰⁷⁸ Zwielfichtige Journale (Beall nennt sie 'Predatory Journals'), die meist nach der Bezahlung einer Gebühr und ohne Qualitätssicherung fast jeden Artikel publizieren, stellen ein Problem für den Wissenschaftsbetrieb dar. Einerseits weisen viele Artikel Qualitätsmängel auf, andererseits wird bei Berufungen Wert auf eine möglichst lange Publikationsliste gelegt, aber kaum qualitativ überprüft. Siehe dazu bspw.: Amrein, Martin: So tricksen Schweizer Forscher die Hochschulen aus, NZZ am Sonntag, 03.02.2018, URL: <https://nzzas.nzz.ch/wissen/so-tricksen-schweizer-forscher-hochschulen-aus-Id.1353872>, Stand: 07.02.2018.

¹⁰⁷⁹ URL: <https://doaj.org/>

¹⁰⁸⁰ Definition nach der Organisation 'opendata.ch'. Siehe URL: <http://opendata.ch/>. opendata.ch ist eine Sektion der Open Knowledge Foundation (URL: <https://okfn.org/>).

¹⁰⁸¹ Definition nach dem 'open data institute'. Siehe URL: <http://theodi.org/what-is-open-data/>.

¹⁰⁸² Aussage von Christoph Hoffmann an der infoclio.ch-Tagung 'Daten und Geschichtswissenschaften' vom 16.10.2015 in Bern (pers. Notiz).

¹⁰⁸³ Das Ziel des Semantic Web ist die maschinelle Les- und Verknüpfbarkeit der im Internet publizierten Informationen nicht nur auf dem Niveau von Realität, sondern als Information über Realität. Nach: Fischer; Hofer: Lexikon der Informatik, 2011; Burkhardt: Digitale Datenbanken, 2015, S. 271.

¹⁰⁸⁴ Siehe URL: <http://5stardata.info/>; URL: <http://www.w3.org/DesignIssues/LinkedData.html>.

System der *Versioned Data Delivery*, mit dem Forschungsdaten von Individualforschenden und kleineren Forschungsgruppen einfach, schnell und kostengünstig der Forschungsgemeinschaft zur Verfügung gestellt werden können.¹⁰⁸⁵

Für die Wissenschaft bedeuten OA und OD, dass zumindest alle mit öffentlichen Geldern finanzierten Datensammlungen, Publikationen oder Forschungsergebnisse auch öffentlich zugänglich sein sollten.¹⁰⁸⁶ Dies kann bei traditionellen Publikationsmethoden schwierig oder gar unmöglich sein, weil beispielsweise der finanzielle Aufwand zu hoch ist, die Daten zu umfangreich sind oder das Publikationsmedium den allgemeinen Zugriff mittels Paywall oder hohen Abonnementspreisen stark einschränkt.¹⁰⁸⁷ Die Soziologin und Bibliothekarin Heike Andermann und der Informationswissenschaftler Andreas Degkwitz bezeichneten die hohen Abonnementspreise schon 2004 als 'Krise der wissenschaftlichen Informationsversorgung' aufgrund der "wachsende[n] Lücke zwischen den Preisen für die wissenschaftliche Information und deren Finanzierbarkeit durch die Bibliotheken"¹⁰⁸⁸ und plädierten für ein neues Geschäftsmodell im digitalen Kontext mit kostenlosem Zugang zur wissenschaftlichen Information.¹⁰⁸⁹ Dass sich mittlerweile auch grosse universitäre Institutionen gegen die hohen Kosten wehren und Verträge mit Verlagen aus Protest kündigen, unterstützt die Bestrebungen nach offenem Zugang zu wissenschaftlichen Publikationen.¹⁰⁹⁰ Bei OA-Publikationen sind im Vergleich zu physischen Publikationen die genannten Hürden massiv niedriger, was zu einem ungemein besseren Preis-Leistungs-Verhältnis führt.¹⁰⁹¹

¹⁰⁸⁵ Als Preprint zur Verfügung gestellt unter: <https://peerj.com/preprints/3401/>.

¹⁰⁸⁶ Nach: Haber: digital past (unv. Habil.), 2009, S. 261. Vgl. auch: Cope, Bill; Phillips, Angus: Introduction, in: Cope, Bill; Phillips, Angus (Hg.): The future of the academic journal, Oxford 2014 (Chandos information professional series), S. 3-6.

¹⁰⁸⁷ Siehe dazu auch bspw.: Hirschi, Caspar: Das geisteswissenschaftliche Buch im digitalen Zeitalter. Der Schweizerische Nationalfonds und seine Open-Access-Strategie., in: NZZ-Online, Zürich 19.05.2014. URL: <http://www.nzz.ch/feuilleton/der-schweizerische-nationalfonds-und-seine-open-access-strategie-1.18304812>. Mit dem Tool *unpaywall* kann nach Artikeln gesucht werden, die sich hinter einer bestimmten Paywall befinden, aber unter einer anderen URL frei erhältlich sind. Siehe: <http://unpaywall.org/>.

¹⁰⁸⁸ Andermann, Heike; Degkwitz, Andreas: Neue Ansätze in der wissenschaftlichen Informationsversorgung: ein Überblick über Initiativen und Unternehmungen auf dem Gebiet des elektronischen Publizierens, in: Historical Social Research, 2004, S. 10.

¹⁰⁸⁹ Siehe: Ebd., S. 18.

¹⁰⁹⁰ Kwon, Diana: Major German Universities Cancel Elsevier Contracts, News, The Scientist, 17.07.2017, URL: <https://www.the-scientist.com/?articles.view/articleNo/49906/title/Major-German-Universities-Cancel-Elsevier-Contracts/>, Stand: 07.02.2018.

¹⁰⁹¹ Siehe dazu: Van Noorden, Richard: Open access: The true cost of science publishing, in: Nature 495 (7442), 27.03.2013, S. 426-429. DOI: 10.1038/495426a; Harnad, Stevan: The post-gutenberg open access journal, in: Cope, Bill; Phillips, Angus (Hg.): The future of the academic journal, Oxford 2014 (Chandos information professional series), S. 186 f. Weiterführend zu Open Access siehe bspw.: Björk, Bo-Christer; Solomon, David: Article processing charges in OA journals: relationship between price and quality, in: Scientometrics 103 (2), 13.03.2015, S. 373-385. DOI: 10.1007/s11192-015-1556-z.

Kritik gegenüber (Linked) OD und OA wird vor allem von privatwirtschaftlichen Unternehmen geäußert, die dadurch ihr Geschäftsmodell bedroht sehen.¹⁰⁹² Verhandlungen zwischen den Interessensgruppen Staat, Privatwirtschaft, Wissenschaft und Private über Zugang, Zugriff und Nutzung von solchen Daten sind noch im Gange, weshalb der öffentlich zugängliche Datenbestand bislang nur langsam zunimmt. Gesetzliche Bestimmungen wie Urheberrecht oder Datenschutz, die die Veröffentlichung von Teilen oder eines gesamten Datenbestandes verhindern können, wirken sich zusätzlich einschränkend auf die Umsetzung von OA- und OD-Projekten aus.

Es ist zu konstatieren, dass die Wissenschaftsgemeinschaft erst langsam zu erkennen scheint, dass mit Open Peer Review, Open Access und Open Data sehr gute Kollaborations-Diskussions- und Publikationsmöglichkeiten zur Verfügung stehen würden, mit denen der wissenschaftliche Informationsaustausch gefördert werden könnte.¹⁰⁹³

5.6.6. Zusammenfassung

Neue Publikations- und Darstellungsformen wie Webseiten, Blogs oder kollaborativ erarbeitete, wissenschaftliche Texte brechen mit der bisherigen stringenten, bindenden Darstellungsordnung, indem sie diese durch eine hypertextuelle und multimediale sowie trans- und interdisziplinäre Darstellung ersetzen. Die neuen Darstellungsformen sollten insbesondere deshalb verwendet werden, weil nur dadurch eine digitale Quelle adäquat und ohne Medienbruch eingebunden und publiziert werden kann. Trotz der noch zu bewältigenden Hürden und der von verschiedenen Seiten wie den Verlagen oder Wissenschaftsgremien geäußerten Vorbehalte, werden Webseiten oder Blogs bereits von einem immer grösser werdenden Kreis von Forschenden als wissenschaftlich relevante Publikationsformen akzeptiert. Hypertextmedien, die sich durch Modularität der informationellen Einheiten, veränderte Lesegewohnheiten, Nichtlinearität, Vernetztheit, Prozessualität, Kollaboration, über Metadaten und Aktualität charakterisieren, verlangen nach neuen Schreibprozessen. Multimediale Formen oder Transmedia Storytelling bieten interessante Ansätze, wie digitale Objekte in digitaler Form publiziert werden könnten. Programmierkenntnisse werden im Schreibprozess zu einer immer wichtigeren Kompetenz, insbesondere um Forschungsergebnisse aus Rohdaten korrekt darstellen und visualisieren zu können. Bei Visualisierungen sind vor allem die Schlüsseltechniken Reduktion und Raumverhältnisse korrekt zu beherrschen sowie die gesellschaftlichen Rezeptionskonventionen zu beachten und zu dokumentieren. Weiterhin sind die

¹⁰⁹² Insbesondere durch die Verlage, bspw. den Beitrag von: Wiedmer, Hans-Rudolf: Publizieren im Zeitalter von Open Access. Die Verlagsperspektive, in: *Traverse: Zeitschrift für Geschichte* 22 (1), 2015, S. 147-156.

¹⁰⁹³ Nach: Eisen, Michael, in: Van Noorden: *Open access*, 2013, S. 429.

verwendeten Quellen, nun mit angepasster Belegkette (Autor, Titel, Identifikator/Speicherort, Link-Check, Version, Datum, Sammlung, Seitenangaben), nachzuweisen.

Die Publikation von Forschungsergebnissen erfolgt weiterhin in traditionellen Formaten und in vertrauenswürdigen Verfahren, aber Konzepte wie Open Peer Review und Open Access / Open Data gewinnen immer mehr Vertrauen. Open Access / Open Data fordern den freien Zugang auf und die freie Nutzung von allen mit öffentlichen Geldern finanzierten Datensammlungen, Publikationen oder Forschungsergebnissen. Digitale Daten sollten als Linked Open Data möglichst strukturiert zur Verfügung stehen, um deren einfachen und schnellen Austausch zu ermöglichen.

5.7. Digital Humanities Cycle



Abbildung 30: Digital Humanities Cycle nach Verhoef und Wever (Darstellung: Föhr)

Die Geisteswissenschaftler Jesper Verhoef und Melvin Wever der Universität Utrecht haben als Arbeitsmethode für den Umgang mit digitalen Objekten den *Digital Humanities Cycle* (DHC) vorgestellt.¹⁰⁹⁴ Der Zyklus beinhaltet die Schritte 'Heuristics, Tool Criticism, Hermeneutics, Source Criticism und Corpus Selection', die iterativ durchlaufen werden. Der Fokus wird dabei auf die für die Beantwortung einer Forschungsfrage zu verwendenden

¹⁰⁹⁴ Teilweise von ihnen auch als 'Digital History Cycle' bezeichnet. Siehe: Verhoef; Wevers: The Digital Humanities cycle, 2015, S. 1.

Informatiktechniken gelegt, was die Autoren als "Question-Driven Use of Tools"¹⁰⁹⁵ bezeichnen.

Da der DHC erst kürzlich vorgestellt und von Verhoef und Wever nur exemplarisch begründet wurde, wird er hier kurz diskutiert.

Ausgangslage:

Wenn sich Historiker mit digitalen Objekten beschäftigen und elektronische Informationssysteme nutzen, werden sie mit technischen, epistemologischen und methodologischen Fragen konfrontiert, die aus der traditionellen historisch-kritischen Methode nicht zureichend sind.

Ziel:

Die Entwicklung eines Modells für die methodologische Vorgehensweise bei der Nutzung von informationstechnischen Werkzeugen, den Tools, unter Einbezug der klassischen respektive traditionellen Methoden.

Diskussion:

Ein iteratives, hybrides Modell, das digitale und klassische Methoden vereint, ist grundsätzlich sinnvoll und wird von mir in dieser Arbeit ebenfalls befürwortet. Im Gegensatz zur vorgestellten, angepassten historisch-kritischen Methode werden im DHC die Informationssystemkritik und die Auswahl des Korpus als eigenständige, zusätzliche Elemente aufgeführt, wobei Quellenkunde, Quellensicherung und die Darstellung im Modell fehlen.

Die Trennung der Korpusauswahl von der Quellenkritik kann gerechtfertigt werden, denn einerseits ist die Auswahl der Quelle das Resultat der Quellenkritik und andererseits kann die Trennung zur Übersichtlichkeit beitragen. Insbesondere wird damit die Arbeit mit grossen Datenbeständen oder solchen, die zu einem grösseren Korpus wie für die TEI zusammengefasst werden, erleichtert. Quellen eines zusammengefassten Korpus unterscheiden sich von bisher bekannten Objekten dadurch, dass sie für die Interpretation zuerst durch ein Informationssystem aufbereitet und allenfalls visualisiert werden müssen.

Die Informationssystemkritik hingegen sollte nicht als eigenständiger Schritt definiert werden, weil bei digitalen Quellen stets ein elektronisches Informationssystem vorausgesetzt wird. Jedes der in den einzelnen Prozessschritten eingesetzten Systeme ist kritisch zu beurteilen, weshalb die Informationssystemkritik eine grundlegend anzuwendende Methode darstellt.

¹⁰⁹⁵ Verhoef, Jesper; Wevers, Melvin: The Digital Humanities Cycle: Doing Digital History Iteratively (Präsentation), Digital Humanities Benelux Conference, Antwerpen 08.06.2015, Fol. 3. URL: <http://www.slideshare.net/MelvinWevers/the-digital-humanities-cycle-doing-digital-history-iteratively>, Stand: 04.08.2015.

Der Schritt der Quellenkunde fehlt im DHC oder wurde unter dem Schritt Heuristik subsumiert. Dies könnte mit entsprechender Begründung eventuell akzeptiert werden, ist aber in den vorhandenen Unterlagen nicht formuliert. Ebenso fehlen die wichtigen Schritte Quellensicherung und Darstellung, die laut meinen Ausführungen zur (erweiterten) historisch-kritischen Methode nicht übergangen werden dürfen. Die Quellensicherung wird bei digitalen Objekten als zwingender Schritt angesehen, weil nur dadurch Persistenz und Nachvollziehbarkeit gewährleistet werden können. Auch die Darstellung respektive Publikation der Forschungsergebnisse darf nicht vernachlässigt werden, da ohne diese die Forschung nicht abgeschlossen werden kann und keine Partizipation am wissenschaftlichen Austausch möglich ist.¹⁰⁹⁶ Vielleicht wurde auf diesen Schritt verzichtet, um einer Diskussion um Publikationsformate wie Blogs auszuweichen. Der Entscheid kann aber aufgrund der vorhandenen Unterlagen nicht nachvollzogen werden.

Fazit:

Das Modell stellt einen interessanten Ansatz einer neuen Methode dar, mit der digitale Quellen kritisch beurteilt werden können. Allerdings wird der Fokus der Kritik mehr auf die digitalen Tools gelegt als auf eine vollständige historisch-kritische Analyse digitaler oder hybrider Objekte. Durch die fehlende Dokumentation sind auch einige Punkte ungeklärt, unter anderem warum die wichtigen Schritte "Quellensicherung" und "Darstellung" fehlen. Deshalb sollte dieses Modell (noch) nicht angewendet werden.

5.8. Zusammenfassung (Fragen und Thesen)

Historiker haben sich mit der neuen Quellengattung auseinanderzusetzen und entsprechende wissenschaftliche Arbeitsmethoden und -techniken anzuwenden. Solche wurden in den einzelnen Prozessschritten Heuristik – Quellenkunde – Quellensicherung – Quellenkritik – Interpretation – Darstellung exemplarisch diskutiert.¹⁰⁹⁷ Anschliessend wurde dargelegt, dass der Digital Humanities Cycle (noch) nicht verwendet werden sollte, weil er zu stark auf die digitalen Tools fokussiert und keine vollständige historisch-kritische Analyse bietet.

Bereits für die Suche nach potentiellen digitalen Forschungsressourcen hat der Recherchierende eine hohe (Digital-)Medienkompetenz über digitale Objekte sowie digitale Informations- und Kommunikationsmedien vorzuweisen: kritisches Beurteilen von Resultatlisten einer Suchmaschine, überprüfen auf Vollständigkeit, Rekontextualisierung oder Kennen der rechtlichen Rahmenbedingungen sind wichtige Merkmale von hybriden Suchtechniken.

¹⁰⁹⁶ Siehe Kap. 3.1.5. Darstellung.

¹⁰⁹⁷ Siehe auch die Unterkapitel 'Zusammenfassung' in den jeweiligen Kapiteln der Prozessschritte.

Digitale Objekte können die Funktion von Quartär- und Quintärquellen übernehmen und sind der Kategorie Elektronisch der Kategorie Immateriell / Virtuell in einem angepassten Quellengattungsschema zuzuordnen.

Weil digitale Objekte grundsätzlich volatil sind und manipuliert werden können, ist die Quellsicherung als neuer Prozessschritt etabliert worden, um den Zugriff auf eine digitale Forschungsressource sowie deren Verfügbarkeit dauerhaft zu gewährleisten. Es ist festzulegen welches Objekt in welcher Form in welchem Format und an welchem Speicherort mit welchen standardisierten Metadaten gesichert wird. Zudem sind allenfalls Hard- und Softwarekomponenten zu archivieren. Die digitalen Objekte sollten mit dem *Researcher Driven Archiving (RDA)* individuelle in einer extern-institutionellen Forschungsinfrastruktur, einem sogenannten *Trusted Digital Repository (TDR)* gespeichert werden. Falls kein solches zur Verfügung steht, könnte ein lokales *Personal Trusted Archive (PTA)* eingesetzt werden. Die Sicherung erfolgt mittels *Trusted Saving Procedure (TSP)*. Die *Metadata Information Section (MIS)* könnte zum Austausch verlässlicher Metadaten beitragen.

Die Konzepte der *Authentizität*, der *Integrität* und der *Persistenz* sind für die kritische Untersuchung eines digitalen Objekts grundlegend, bedeuten aber in der Geschichtswissenschaft und in der Informatik Unterschiedliches. Der Begriff Authentizität kann geschichtswissenschaftlich Original bedeuten, aber auch originalgetreu; das Untersuchungsobjekt kann beschädigt oder unbeschädigt sein, von bekannter oder unbekannter Herkunft stammen. Für die nachvollziehbare Beurteilung der Authentizität eines digitalen oder physischen Objekts kann die *Authentizitätsapproximation* verwendet werden. In der Informatik ist Authentizität an den eindeutigen Vergleich von Identifikationsmerkmalen gebunden. Informationstechnische Integrität bedeutet die Unversehrtheit des Bitcodes. Die historische Integrität eines Objekts ist gewahrt, wenn es materiell und/oder inhaltlich unbeschädigt sowie unversehrt und unverändert ist. Es können Informations-, Darstellungs- und Nutzungsintegrität unterschieden und die Integrität der verwendeten elektronischen Informationssysteme sollte überprüft werden. Physische Objekte haben durch ihre Materialität eine massiv höhere Persistenz als digitale, bei denen zwischen Objekt- und Adresspersistenz unterschieden wird.

Die informationstechnische *Objektbeschreibung* bildet die nötige Voraussetzung für die Darstellung eines digitalen Objekts durch ein elektronisches Informationssystem. Die *Echtheit* eines singulären digitalen Objekts kann nur anhand der im Objekt enthaltenen Überlieferungsgeschichte und/oder unter Zuhilfenahme von dessen Metadaten überprüft werden. Die *Datierung* eines digitalen Objekts sollte sich auf Sicherheitsmerkmale wie die qualifizierte elektronische Signatur mit Zeitstempel und/oder Informationen aus TDR abstützen. Anstelle des *Erscheinungsorts* sollte ein *persistenter Identifikator*, der um die URL eines gesetzlich beauftragten TDR ergänzt wurde, angegeben werden. Auch bei der Identifikation der Autor-

schaft ist eine fortgeschrittene digitale Signatur hilfreich. Anonymität, Pseudonyme und Softbots erschweren die Arbeit. *Autor und Adressat* können je nach verwendeter Kommunikationstechnik eine manuskriptive Kollaboration eingehen. Der Adressat erhält dadurch einen argumentativen Machtzuwachs. Die Darstellung ist daraufhin zu untersuchen, ob sie so intendiert und korrekt ist. Um den *Inhalt* zu beurteilen, sind alle Ebenen des digitalen Objekts zu untersuchen. Inhalte im Web können domain-, orts- oder browserspezifisch generiert werden. Hyperlinks können das digitale Objekt inhaltlich ergänzen, indem sie qualitative *Beziehungen* zu anderen Objekten herstellen. Der Link-Check weist auf die erfolgte Analysetiefe hin.

Für die Überprüfung eines digitalen Objekts ist dessen Authentizität massgebend. Ein Objekt kann in unterschiedlichem Masse authentisch sein, abhängig von der Fragestellung, von den zur Verfügung stehenden Informationen, von Sicherheitsmerkmalen und auch von den medien- und informationstechnischen Kompetenzen des Nutzers. Deshalb ist der Näherungswert für jede potenzielle Quelle individuell zu erarbeiten.

T5: Die quellenkritische Überprüfung eines digitalen Objekts bedeutet die Erarbeitung eines Näherungswerts für die Wahrscheinlichkeit der Authentizität des Objekts.

Mit obigen Ausführungen wurde die sechste Frage 'Welche Methoden können für die Quellenkritik digitaler Objekte in der Geschichtswissenschaft angewendet werden?' beantwortet:

F6: Es sind für die Quellenkritik digitaler Objekte sowohl die traditionelle Methode wie auch Methoden aus der Informationstechnik anzuwenden. Letztere kommen immer dann zum Einsatz, wenn Fragen zu den Eigenschaften des digitalen Objekts, dessen Verwendung innerhalb eines digitalen Informationssystems (inkl. Netzwerken wie das Internet) oder zum System selbst beantwortet werden sollen.

Auch die siebte Frage 'Verändert sich die Arbeitsweise der historischen Zunft durch die veränderten quellenkritischen Anforderungen?' ist beantwortet:

F7: Der Einbezug informationstechnischer Arbeitsmethoden und -techniken in die historisch-kritische Methode und die neuen quellenkritischen Anforderungen verändern auch die Arbeitsweise der Historiker, weil ohne neue Methoden neue Quellen nicht bearbeitet und kritisch untersucht werden können.

Die sechste These ist bestätigt:

T6: Die Arbeitsweise der historischen Zunft verändert sich durch Digitalität.

6. Digitalhistoriker

*"Going digital is not a choice but a given"*¹⁰⁹⁸ (Zaagsma, 2013)

Das Bewusstsein, dass das Digitale nicht ignoriert werden kann, ist noch nicht überall vorhanden. Wer sich ausschliesslich mit physischen Quellen beschäftigt, meint oft, er verwende keine digitalen Objekte und habe sich deshalb auch nicht damit auseinanderzusetzen. Viele Nutzer eines Computersystems sind sich nicht bewusst, dass wenn sie damit Texte schreiben, ein Dokument im Archiv mit einer Digitalkamera fotografieren, ein Bild bearbeiten, in einen Online-Bibliothekskatalog recherchieren oder ein E-Mail schreiben bereits digital arbeiten. Viele elektronischen Informationssysteme unterstützen oder ermöglichen erst die Arbeit, indem mit diesen der Fernverkehr disponiert oder die Nahrungsmittelproduktion gesteuert wird. Die Durchdringung der Gesellschaft und der Wissenschaft mit Computertechnologie ist so gross, dass die Anwendung informationstechnischer Methoden und Techniken unausweichlich ist. Die Gesellschaft ist *hybrid* geworden – *"hybridity is the new normal."*¹⁰⁹⁹

Alle Wissenschaftler und damit auch die Historiker haben sich mit der neuen Quellengattung auseinanderzusetzen und entsprechende wissenschaftliche Arbeitsmethoden und -techniken anzuwenden. Sie müssen sich mit den Eigenschaften digitaler Objekte auskennen und digitale oder digitalisierte Daten bearbeiten können. Damit werden sie zu Digitalhistorikern – vorerst noch, denn (hoffentlich) bald wird das 'Digitale' nicht mehr besonders sein, sondern selbstverständlicher Bestandteil der Methodenkompetenz eines Historikers sein.

Die Erarbeitung von digitalen Arbeitsmethoden und -techniken für die Geschichtswissenschaften ist allerdings noch wenig fortgeschritten, weil viele Historiker gegenüber Veränderungen ihrer Arbeitsmethoden wenig aufgeschlossen sind oder meinen, dass sie das Thema nicht betreffe. "[N]ot nearly enough attention is being paid to how doing history in the digital age is changing"¹¹⁰⁰ meinte Gerben Zaagsma bereits vor zwei Jahren. Diese Aussage ist heute noch gültig, denn es sind zwar Bestrebungen zur Erarbeitung eines Curriculums solcher Kompetenzen und Arbeitsmethoden erkennbar¹¹⁰¹, aber bisher ist noch kein Konsens für deren Integration in ein geschichts- resp. geisteswissenschaftliches Studium in Sicht. Dies ist aber auch dem Umstand geschuldet, dass sich Informationssysteme schnell weiter-

¹⁰⁹⁸ Zaagsma: On Digital History, 2013, S. 17.

¹⁰⁹⁹ Ebd. Zur Abhängigkeit der Gesellschaft von Computersystemen siehe auch bspw. Sherry Turkle: Jenkins: «Does this Technology serve Human Purposes?» A «Necessary Conversation» with Sherry Turkle, 2015; Turkle, Sherry: Verloren unter 100 Freunden: wie wir in der digitalen Welt seelisch verkümmern, München 2012.

¹¹⁰⁰ Zaagsma: On Digital History, 2013, S. 5.

¹¹⁰¹ Vgl. bspw.: Gasteiner; Haber (Hg.): Digitale Arbeitstechniken, 2010; Galgano; Arndt; Hyser: Doing history research and writing in the digital age, 2008; Sahle: Digitale Editionsformen, 2013; American Historical Association: Guidelines for the Professional Evaluation of Digital Scholarship in History – Draft, 04.2015. URL: <http://historians.org/Documents/Teaching and Learning/Current Projects/Digital Scholarship Evaluation/Guidelines for the Professional Evaluation of Digital Scholarship in History.pdf>, Stand: 27.07.2015.

entwickeln. Bis sich die Historiker, die eher als "tool conservative"¹¹⁰² gelten, mit einer Informationstechnik beschäftigen, ist diese oft schon obsolet geworden. Eine Ausnahme sind die Archivwissenschaften, in denen die Problematik der dauerhaften Speicherung von digitalen Objekten bereits früh erkannt wurde. Seit längerem wird diese nachhaltig erforscht und es werden Lösungen entwickelt.¹¹⁰³

In diesem Kapitel wird kurz auf die curriculare Einbindung der wichtigsten Kompetenzen in geistes- und vor allem geschichtswissenschaftliche Studiengänge sowie auf die dafür notwendige Infrastruktur eingegangen.

6.1. Ausbildung

*"It is here that much more education is needed in order for historians to be able to make an informed choice as to which tools to employ in their research."*¹¹⁰⁴ (Zaagsma, 2013)

In den bisherigen Ausführungen habe ich aufgezeigt, dass für eine adäquate historisch-kritische Beurteilung eines digitalen Objekts nicht nur die traditionell vermittelten Kompetenzen wie interdisziplinäres Denken, wissenschaftliches Arbeiten, Methodenkompetenz oder Fachwissen vorhanden sein müssen, sondern dass auch neue Kompetenzen in der Informationstechnik und in der Kommunikation unter Forschenden und von Forschungsergebnissen verlangt werden. Fred Gibbs und Trevor Owens fordern, dass Historiker nicht nur lernen traditionelle Quellen zu finden, zu sammeln, zu organisieren und daraus Wissen zu generieren, sondern auch Daten zu erfassen, zu manipulieren, zu analysieren und zu repräsentieren.¹¹⁰⁵ Digital Natives haben dabei einen ebenso hohen Bedarf an fundierter Ausbildung wie Digital Immigrants. Deren Anzahl wird in nächster Zeit mit der Anbindung bisher von der elektronischen Informationsinfrastruktur Ausgeschlossener, beispielsweise aus afrikanischen Ländern, zunehmen.

Die Anstrengungen *Referenzcurricula* für Lehrgänge in den Digital Humanities zu erarbeiten, in denen die Ausbildung der geforderten Kompetenzen einen hohen Stellenwert eingeräumt wird, sind begrüßenswert.¹¹⁰⁶ Allerdings sollten solche nicht nur für Studiengänge in den

¹¹⁰² Burckhardt; Schiel: Kollaboratives Schreiben, 2010, S. 108.

¹¹⁰³ Vgl. bspw.: Arms, Caroline R.: Keeping Memory Alive: Practices for Preserving Digital Content at the National Digital Library Program of the Library of Congress, in: RLG DigiNews Volume 4 (Number 3), 15.06.2000. URL: <http://webdoc.gwdg.de/edoc/aw/rlgdn/preserv/diginews/diginews4-3.html#feature1>, Stand: 23.07.2014; Cullen, Charles T.; Hirtle, Peter B.; Levy, David u. a.: Authenticity in a Digital Environment, Washington, D.C 2000; Gladney: Preserving Digital Information, 2007; Masanès: Web Archiving, 2006; Brügger: Archiving Websites, 2005; Brügger: Web Archiving, 2012.

¹¹⁰⁴ Zaagsma: On Digital History, 2013, S. 18.

¹¹⁰⁵ Gibbs; Owens: The Hermeneutics of Data, 2013, S. 169.

¹¹⁰⁶ Siehe bspw.: Sahle, Patrick: DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum – Milestone 2.3.3, 19.03.2013. URL: <https://dev2.dariah.eu/wiki/download/attachments/14651583/>

Digital Humanities, sondern für alle wissenschaftlichen Lehrgänge erarbeitet werden. Kein Wissenschaftszweig kann behaupten, dass er von den Veränderungen des Digitalen Zeitalters nicht beeinflusst wird. Bereits bestehende Forderungskataloge wie die *Guidelines for the Professional Evaluation of Digital Scholarship in History* der American Historical Association¹¹⁰⁷ sind zwar vielversprechend, aber erst als Rohkonzept verfügbar und noch zu diskutieren.

Die Betonung des *digital* wird mit der immer tieferen Durchdringung der (Wissens-)Gesellschaft durch elektronische Informationssysteme in naher Zukunft nicht mehr sinnvoll sein. 'Digitale Arbeitsmethoden' werden nicht mehr 'besonders' digital sein, sondern werden ganz selbstverständlich schlicht zu 'Arbeitsmethoden', wie Peter Haber mehrfach betonte und Gerben Zaagsma prognostiziert:

*"I would hope that within a decade or so there will be no more talk of 'digital history' as all history is somehow 'digital' in terms of incorporation of new types of sources, methods and ways of dissemination (just as all humanities will be inherently 'digital')."*¹¹⁰⁸

Allerdings sind bis dahin die dafür nötigen *Schlüsselkompetenzen* allen Wissenschaftlern zu vermitteln und aufgrund der sich immer weiterentwickelnden Informationstechnik auch ständig zu aktualisieren und zu erweitern.

Die einzelnen Kompetenzen können nicht klar voneinander getrennt oder hierarchisch geordnet werden, weil einige Merkmale oder Qualifikationen meist mehreren Kompetenzbereichen zugeordnet werden können.¹¹⁰⁹ Allgemein sollten sich Historiker Kompetenzen aneignen, mit denen die Grundlagen digitaler Technologien und die Interaktion mit den Menschen verstanden werden können.¹¹¹⁰ Peter Haber hat bereits ausführlich dargelegt, welcher Kanon an Kompetenzen vermittelt werden sollte, und bezeichnet die Informations- und die Medienkompetenz als besonders wichtig.¹¹¹¹

Informationskompetenz wird als die Fähigkeit verstanden, zu einer Frage den Informationsbedarf zu ermitteln, eine Suchstrategie zu entwickeln, die geeigneten Informationsquellen zu kennen und benutzen zu können, die erhaltenen Daten zu bewerten und zu selektieren und

DARIAH-M2-3-3_DH-programs_1_1.pdf?version=2&modificationDate=1366904376117&api=v2, Stand: 04.04.2015. Allerdings stellt Sahle im Dokument nur ein Projekt zur Erarbeitung vor, wenige inhaltliche Angaben zum Curriculum sind enthalten.

¹¹⁰⁷ Siehe: American Historical Association: *Guidelines for the Professional Evaluation of Digital Scholarship in History – Draft*, 2015, URL: <http://historians.org/Documents/Teaching and Learning/Current Projects/Digital Scholarship Evaluation/Guidelines for the Professional Evaluation of Digital Scholarship in History.pdf>.

¹¹⁰⁸ Zaagsma: *On Digital History*, 2013, S. 16.

¹¹⁰⁹ Die vielfältigen Definitionsansätze einer einzelnen Kompetenz verunmöglichen eine Ordnung. Es ist hier nicht möglich, eine Differenzierung und Definierung vorzunehmen.

¹¹¹⁰ Vgl.: Capurro, Rafael: *Digital Hermeneutics – An Outline*, Rafael Capurro, 09.11.2017, URL: <http://www.capurro.de/digitalhermeneutics.html>, Stand: 07.02.2018.

¹¹¹¹ Vgl.: Haber: *digital past* (unv. Habil.), 2009, S. 188-201.

die Informationen so zu verarbeiten, dass die Fragestellung effektiv beantwortet werden kann.¹¹¹² Der in dieser Definition beschriebene Prozess entspricht dem der historisch-kritischen Methode. Da zudem für Historiker der Umgang mit analogen und digitalen Informationen eine der wichtigsten Kompetenzen darstellt, wird die Informationskompetenz als grundlegend erachtet.

Für die Nutzung geeigneter Informationsquellen ist eine fundierte *Medienkompetenz* elementar. Wie ich in Kapitel 5.1.1. Suchstrategie bereits ausgeführt habe, wird darunter ein Bündel verschiedener technischer und intellektueller Teilkompetenzen verstanden, die für eine fundierte Medien- und Informationskritik elementar sind.

*"Medienkompetenz' meint also grundlegend nichts anderes als die Fähigkeit, in die Welt aktiv aneignender Weise auch alle Arten von Medien für das Kommunikations- und Handlungsrepertoire von Menschen einzusetzen."*¹¹¹³

Für Haber ist Medienkompetenz eine konzeptionelle Ausweitung der Lesekompetenz.¹¹¹⁴ Der Historiker hat für die Recherche nicht nur die geeigneten analogen und digitalen Medien auszuwählen, sondern auch deren technische und kommunikative Eigenschaften sowie den Kontext der durch diese Medien erstellten Informationen zu kennen.¹¹¹⁵ Jan Hodel verwendet für Kompetenzen im Umgang mit Digitalmedien den Begriff der *historischen Online-Kompetenz*. Er suggeriert aber mit *Online* eine Einschränkung auf Forschungsressourcen im Internet. Medienkompetenz umfasst jedoch alle Medien, unabhängig von ihrer Materialität. Zum Grundlagenwissen gehören beispielsweise die Entwicklung einer Suchstrategie oder das 'Lesenkönnen' eines Mediums.

*"Any robust digital research methodology must allow the scholar to move easily between distant and close reading, between the bird's eye view and the ground level of the texts themselves."*¹¹¹⁶

Es sind Massnahmen zu treffen, um Digital Natives und betroffene Digital Immigrants durch entsprechende Ausbildung zu befähigen, physische und digitale Objekte genau, vollständig und kritisch lesen zu können. Dies kann innerhalb des akademischen Curriculums erfolgen

¹¹¹² Definition nach dem Projekt 'Informationskompetenz' der Universitätsbibliothek Konstanz, Abschluss 2005. Siehe: Bibliothek der Universität Konstanz: Informationskompetenz an der Bibliothek der Universität Konstanz, Bibliothek der Universität Konstanz, 04.06.2008, URL: <http://www.ub.uni-konstanz.de/bibliothek/projekte/informationskompetenz/>, Stand: 18.08.2015.

¹¹¹³ Baacke, Dieter: Medienkompetenz – Begrifflichkeit und sozialer Wandel, in: Rein, Antje von (Hg.): Medienkompetenz als Schlüsselbegriff, Bad Heilbrunn 1996 (Theorie und Praxis der Erwachsenenbildung), S. 119. Kursiv im Original. Vgl. zur Konzeption des Begriffs auch: Groeben, Norbert; Hurrelmann, Bettina (Hg.): Medienkompetenz: Voraussetzungen, Dimensionen, Funktionen, Weinheim 2002 (Lesesozialisation und Medien).

¹¹¹⁴ Nach: Haber: digital past (unv. Habil.), 2009, S. 188.

¹¹¹⁵ Siehe auch: Hodel, Jan: Digital lesen, digital schreiben, digital denken? Über den kompetenten Umgang mit Geschichte im Zeitalter des digitalen Medienwandels, in: Eggs, Cindy; Jorio, Marco (Hg.): Wort. Am Anfang ist das: Lexika in der Schweiz, Baden 2008, S. 116 f.

¹¹¹⁶ Gibbs, Frederick W.; Cohen, Daniel J.: A Conversation with Data: Prospecting Victorian Words and Ideas, in: Victorian Studies 54 (1), 01.10.2011, S. 76. DOI: 10.2979/victorianstudies.54.1.69.

oder durch zusätzliche Kurse, den die Teilnehmer mit einem Fähigkeitszeugnis abschließen. Veranstaltungen wie *#Lesen – Transformationen traditioneller Rezeptionskonzepte im digitalen Zeitalter*¹¹¹⁷ des DFG Graduiertenkollegs 1787, die sich Fragen zu Veränderungen im Leseverhalten und in der/den Lesekultur(en) sowie deren Auswirkungen auf die Produktions- und Vermittlungsseite widmen, sollten weiterführende Erkenntnisse über notwendige, auszubildende Qualifikationen hervorbringen.

Mit der Medienkompetenz, insbesondere der Mediennutzung, sind auch neue *Kommunikationskompetenzen* verbunden. Wenn der Nutzer Digitalmedien für die wissenschaftliche Forschung und zur Publikation seiner Ergebnisse verwendet, sollte er die Grundsätze der Kommunikation kennen, insbesondere die Art und Weise und der Zeitpunkt der Veröffentlichung, die Wahl des Mediums und rechtliche Aspekte. Beispielsweise sind in einem Blog der Zeitpunkt der Publikation, die sprachliche Formulierung oder auch die Art der Visualisierung zu beachten, wenn Informationen bestimmte Adressaten erreichen sollen. Besonders Hypertexte stellen neue Anforderungen an das historische Schreiben, denn mit den kurzen und in sich geschlossenen Informationseinheiten brechen sie mit den bewährten Traditionen und Konventionen des Fachs¹¹¹⁸ Kommunikation bedeutet nicht nur Austausch mit Fachkollegen oder interessierten Kreisen, mit denen eventuell anders zu kommunizieren ist, sondern auch das Filtern und (Re-)Organisieren aller verfügbaren Informationen, Anregungen und Hinweise.

Dank seiner *Informatikkompetenz* oder Informationstechnikkompetenz ist der Nutzer befähigt, digitale Informationssysteme und deren Produkte, die digitalen Objekte, kritisch zu untersuchen und zu beurteilen. Dazu ist ein grundlegendes Verständnis über Hard- und Software nötig. Zum Grundlagenwissen gehören beispielsweise Kenntnisse über den Aufbau von Computersystemen mit Zentraleinheit und Peripheriegeräten, Zahlensysteme, Logik, Boolesche Algebra, Algorithmen, Datenstrukturen, Speicher- und Kompilierungsprozesse, Betriebssysteme, Software Engineering, Datenbanksysteme, Datenkompression, Kryptographie, Vernetzung, Internet und Web sowie Programmierung. In welcher Detailtiefe die einzelnen Qualifikationen von einem Geisteswissenschaftler zu beherrschen sind, ist Gegenstand aktueller Diskussionen.

Insbesondere die letzte Forderung, dass Historiker das *Programmieren* lernen sollten, ist sehr umstritten.¹¹¹⁹ Auf der einen Seite wird argumentiert, dass nur mit

¹¹¹⁷ URL: <https://www.uni-goettingen.de/de/tagungen/524561.html>.

¹¹¹⁸ Nach: Hodel: Digital lesen, digital schreiben, digital denken? Über den kompetenten Umgang mit Geschichte im Zeitalter des digitalen Medienwandels, 2008, S. 119.

¹¹¹⁹ Eine kurze Darstellung der Auseinandersetzung bietet: Messner, Daniel: Coding History – Software als kulturwissenschaftliches Forschungsobjekt, in: Schmale, Wolfgang (Hg.): Digital Humanities: Praktiken der Digitalisierung, der Dissemination und der Selbstreflexivität, Stuttgart 2015 (Geschichte), S. 157-174.

Programmierkenntnissen die Interpretation von digitalen Forschungsressourcen und das kritische Lesen der digitalen Objekte möglich seien. Zudem könnten Forscher erfolgreicher mit Programmierern zusammenarbeiten und wären auch von diesen unabhängiger.¹¹²⁰ Auf der anderen Seite werden Programmierkenntnisse beispielsweise von Peter Haber¹¹²¹ oder Toni Weller¹¹²² als nicht notwendig bezeichnet, und es werden technische Systeme gefordert, die es einem Historiker auch ohne Programmierkenntnisse ermöglichen sollen, eine kritische Analyse eines digitalen Objekts vorzunehmen.¹¹²³ Der Nutzer hat aber auch solche Systeme zu verstehen, was wieder die Ausbildung von Hard- und Softwarekenntnissen erfordert. Meines Erachtens müssen Historiker nicht eine Ausbildung zum Informatiker abschliessen, aber sie sollten Grundkenntnisse des Programmierens erlernen. Damit können sie Informationssysteme und deren Algorithmen als kognitive Schnittstellen erst erkennen.¹¹²⁴ Softwarekenntnisse werden unter Umständen auch für das Publizieren, Kommunizieren und Visualisieren der Forschungsergebnisse benötigt. Obwohl viele Programme und Tools die Arbeit erleichtern, sollte der Nutzer sich vergewissern können, dass die eingesetzte Software die von ihm intendierte Verarbeitung auch genau so erledigt wie intendiert. Der Nutzer sollte zudem bei einfachen und offensichtlichen Fehlern selbst eingreifen können.

Es ist jedoch nicht sinnvoll für die Geschichtswissenschaft eigene Programmiersprachen zu entwickeln, wie es beispielsweise Guido Koller fordert.¹¹²⁵ Solche spezifischen Programmiersprachen stellen Insellösungen dar und müssten sehr gut dokumentiert sein, um deren Funktionalität auch nach Jahren noch nachvollziehen zu können. Deshalb sollten eher populäre und standardisierte Programmiersprachen verwendet werden, die von vielen Hardware- und Software-Plattformen unterstützt werden.

Um eine algorithmische Auswertung verstehen und nachvollziehen zu können, deren Voraussetzungen, Annahmen und Einschränkungen zu kennen und sie schliesslich kritisieren zu können, sollten sich Wissenschaftler grundlegende Kenntnisse im Programmieren aneignen. Durch das Erlernen der Syntax wichtiger Programmiersprachen wie HTML, Java oder C++ können Algorithmen auch dann untersucht werden, wenn eine unbekannte Sprache

¹¹²⁰ Vgl. bspw.: Rieder; Röhle: Digital Methods, 2012, S. 81; Berry: The Computational Turn, 2011, S. 15.

¹¹²¹ Haber, Peter: Wir werden uns die Technik noch stärker zu nutze machen, Audiointerview, 00:01:46, 12.09.2011. URL: http://www.lisa.gerda-henkel-stiftung.de/historiker_muessen_keine_informatiker_sein?nav_id=1807, Stand: 06.12.2015.

¹¹²² Weller, Toni: Introduction, in: Weller, Toni (Hg.): History in the digital age, London 2013, S. 1; Weller: Conclusion. A changing field, 2013, S. 195.

¹¹²³ Siehe bspw. Kommentar von Eva P. (Eva Pfanzer) vom 27.04.2012 auf: Föhr, Pascal: Historiker müssen HTML interpretieren können, Billet, 21.03.2012, URL: <http://hsc.hypotheses.org/38>, Stand: 19.08.2015.

¹¹²⁴ Vgl. unter Berufung auf Bruno Latour: Koller: Geschichte digital, 2016, S. 67 und 84.

¹¹²⁵ Ebd., S. 18.

verwendet wurde, weil deren Aufbau relativ schnell ersichtlich ist und nachvollzogen werden kann. Die Kenntnis über das Funktionieren des Algorithmus bedeutet Macht, weil damit dessen Ergebnisse kritisch untersucht und selbständig interpretiert werden können und weil der Nutzer den Algorithmus nach seinen Bedürfnissen ändern kann.

Historiker sollten zukünftig auch im Umgang mit grossen Datenbeständen ausgebildet werden, was ich hier als *Datenkompetenz* bezeichne. Sie beschreibt die Fähigkeit, grosse digitale Datenbestände (Big Data) mit geeigneten Mitteln auf ihre Herkunft, Qualität und den Verwendungszweck hin adäquat zu beurteilen, zu beurteilen und visualisieren respektive darstellen zu können. Datenkompetenz ermöglicht es dem Historiker, visualisierte Datenbestände zu kritisieren, zu (re-)kontextualisieren und zu interpretieren. Ohne eine fundierte Datenkompetenz ist der Wissenschaftler auf Spezialisten angewiesen und gezwungen, den digitalen Informationssystemen zu vertrauen.

Informatik- und Datenkompetenzen haben auch Archivare und wissenschaftliche Mitarbeiter in Archiven zukünftig vorzuweisen. Beispielsweise müssen Staatsarchive, die sicherzustellen haben, dass die Verwaltungstätigkeit einer behördlichen Einheit nachvollziehbar ist, immer mehr digitale Objekte sichern, bewerten, übernehmen, dauerhaft speichern sowie diese den Nutzern wieder zur Verfügung stellen. Hier werden Personen benötigt, welche sowohl fundierte Kenntnisse über archivische Tätigkeiten, den historischen Wert der vorhandenen physischen und digitalen Informationen als auch informationstechnische Kenntnisse vorweisen können, um die digitale Langzeitarchivierung sicherzustellen.¹¹²⁶

Historiker haben sich neue Kompetenzen anzueignen und diese ständig aufzufrischen, denn "innovative Technik braucht kompetente Menschen, die sie beherrschen, statt von ihr beherrscht zu werden."¹¹²⁷ Welche spezifischen Qualifikationen in ein Curriculum aufgenommen werden sollen, ist noch Gegenstand von Diskussionen und Verhandlungen, vor allem weil die meisten Kompetenzbereiche in ihrer inhaltlichen Konzeption nicht einheitlich aufgefasst werden.¹¹²⁸ Trotzdem sind zwingend Grundlagen in allen oben genannten Kompetenzbereichen zu vermitteln, damit die Historiker mit einer reflektierten Methodenkompetenz für die zukünftige Entwicklung des Wissenschaftsgebiets gerüstet sind.

¹¹²⁶ Erfahrungen aus eigener Tätigkeit als wiss. Mitarbeiter im Staatsarchiv Solothurn.

¹¹²⁷ Gigerenzer, Gerd: Technik braucht Menschen, die sie beherrschen, in: Das Digital-Manifest, 2015, S. 31. Siehe auch die Forderung des VHD Verband der Historiker und Historikerinnen Deutschlands: Quellenkritik im digitalen Zeitalter. Die Historischen Grundwissenschaften als zentrale Kompetenz der Geschichtswissenschaft und benachbarter Fächer, Stellungnahmen, VHD, 10.2015, URL: <http://www.historikerverband.de/verband/stellungnahmen/quellenkritik.html>, Stand: 29.11.2015. Das Positionspapier erhält positive Resonanz mit teilweise weitergehenden Forderungen. Siehe bspw.: Diskussionsforum: Historische Grundwissenschaften und die digitale Herausforderung, in: H-Soz-Kult, 15.11.2015, URL: <http://www.hsozkult.de/text/id/texte-2890>.

¹¹²⁸ Siehe dazu bspw. die Kritik von Baacke am Begriff der 'Medienkompetenz', in: Baacke: Medienkompetenz – Begrifflichkeit und sozialer Wandel, 1996, S. 119-121.

"Eine Methodenausbildung vermittelt nicht nur praktische Kenntnisse für den Umgang mit Daten, sondern versetzt Studierende und Forschende auch in die Lage, gehaltvolle Fragen zu stellen und sich bewusst zu machen, wie ihre Entscheidungen an einer beliebigen Stelle im Forschungsprozess sich auf alle folgenden Phasen auswirken."¹¹²⁹

Die Ausbildung solcher Kompetenzen würde auch der oft geäußerten Forderung nach wissenschaftlicher Interdisziplinarität entsprechen, den Historikern den Blick über die Fachgrenzen hinaus ermöglichen und helfen, neue Forschungsfragen zu bearbeiten.

6.2. Infrastruktur

Mit der Ausbildung neuer Kompetenzen sind auch neue digitale Infrastrukturen aufzubauen, die die Erstellung, Verarbeitung, Nutzung und Archivierung von Forschungsdaten ermöglichen. Alle Wissenschaftszweige sollten den Austausch mit den Archivwissenschaften suchen und institutionalisieren, damit die dauerhafte Archivierung und Verfügbarkeit der Daten gewährleistet bleibt. Archive, die bereits eine Infrastruktur zum Speichern von digitalen Objekten betreiben,¹¹³⁰ suchen händeringend nach Digitalhistoriker, die sie darin unterstützen.

Drei wichtige Punkte sind bei der Umsetzung solcher Vorhaben zu beachten: Erstens sollten die dafür verantwortlichen Institutionen mit einem gesetzlichen Auftrag zur Langzeitarchivierung beauftragt werden, damit die Rechte und Pflichten der beteiligten Institutionen klar geregelt sind und damit auch die Finanzierung langfristig gesichert ist. Zweitens sollten Infrastrukturen nicht als Projekte geplant und nur mit Anschubfinanzierungen unterstützt werden, denn es geht um die Langzeitarchivierung von digitalen Objekten für die nächsten Tausend Jahre.¹¹³¹ Drittens sollten jegliche digitalen Objekte in einer solchen Infrastruktur gesichert werden können, damit sie als Forschungsressource dauerhaft zur Verfügung stehen. Dies setzt eine Archivierungsstrategie voraus, die klar definiert, wie ein solches Objekt in das Archiv aufgenommen und wie es im Archiv selbst behandelt wird. Im Unterschied zu physischen Objekten sind digitale Objekte in regelmässigen Abständen zu überprüfen und auf andere Datenträger oder in ein anderes Datenformat zu migrieren, um ihre Lesbarkeit zu gewährleisten.

Damit Wissenschaftler unabhängig von einem Forschungsprojekt ihre digitalen Quellen sichern können, wird das *Researcher Driven Archiving (RDA)* vorgeschlagen, das die individuelle Sicherung digitaler Objekte in *Trusted Digital Repositories (TDR)* ermöglicht. Dafür

¹¹²⁹ Mahrt; Scharnow: Wert von Big Data, 2014, S. 232.

¹¹³⁰ Bundes-, Kantons- und Gemeindebehörden müssen in den nächsten Jahren entsprechende Infrastrukturen aufbauen und anschliessend dauerhaft betreiben. Bspw. betreibt das Staatsarchiv Basel-Landschaft bereits eine entsprechende Infrastruktur, das Staatsarchiv Solothurn baut die nötige Infrastruktur erst auf.

¹¹³¹ Die genannte Zahl soll nur den langen Zeitraum verdeutlichen.

sollten benutzerfreundliche Softwares entwickelt werden, mit der sich jedes Objekt durch eine *Trusted Saving Procedure (TSP)* in einem TDR eines gesetzlich beauftragten Providers sichern lässt. Wird die digitale Forschungsressource als Quelle verwendet, kann sie unter einem *persistenten Identifier (PI)* oder einer/einem eindeutigen URI/URL veröffentlicht werden.¹¹³² Jede nachfolgende Änderung oder Ergänzung wird protokolliert und ist von Drittpersonen einsehbar. Alternativ oder zusätzlich sollte eine Software entwickelt werden, die eine vertrauenswürdige Sicherung mittels TSP in einem *Personal Trusted Archive (PTA)* erlaubt, das wiederum in einem TDR veröffentlicht werden kann.

Jede Sicherung ist mit standardisierten Metadaten zu versehen, die während des Sicherungsprozesses entstehen und vom System automatisch hinzugefügt und durch den Nutzer ergänzt werden können. Es wird die Einführung einer *Metadata Information Section (MIS)* vorgeschlagen, die in alle wissenschaftlichen oder wissenschaftsrelevanten digitalen Objekte integriert wird, alle relevanten bibliographischen und identifizierenden Metadaten enthält und automatisch ausgelesen werden kann. Bestimmte Metadaten wie Zeitstempel oder Provenienz sollten dabei vom Nutzer nicht verändert werden können. Die MIS sollte in alle wissenschaftlichen, analogen und digitalen Publikationen implementiert werden, um den wissenschaftlichen Austausch mit verlässlichen Metadaten zu fördern.

6.3. Zusammenfassung

Da die Informatik in allen Bereichen der Wissenschaft Einzug gehalten hat, haben sich auch die Historiker mit der neuen Quellengattung auseinanderzusetzen. Dazu wäre es vorteilhaft, wenn sie bereits während des Studiums entsprechende Kompetenzen vermittelt erhielten:

- *Informationskompetenz* als Fähigkeit, zu einer Frage den Informationsbedarf zu ermitteln, eine Suchstrategie zu entwickeln, die geeigneten Informationsquellen zu kennen und benutzen zu können, die erhaltenen Daten zu bewerten und zu selektieren und die Informationen so zu verarbeiten, dass die Fragestellung effektiv beantwortet werden kann.
- *Medienkompetenz* als Fähigkeit, in die Welt aktiv aneignender Weise auch alle Arten von Medien für das Kommunikations- und Handlungsrepertoire von Menschen einzusetzen.
- *Kommunikationskompetenz* als Fähigkeit, Kommunikationskanäle und -medien zum Austausch mit Fachkollegen und der Öffentlichkeit zu nutzen und alle verfügbaren Informationen, Anregungen und Hinweise zu filtern und zu (re-)organisieren.

¹¹³² Rechtliche Einschränkungen werden hier nicht beachtet.

- *Informatikkompetenz* oder Informationstechnikkompetenz (inkl. Programmieren) als Fähigkeit, digitale Informationssysteme und deren Produkte, die digitalen Objekte, kritisch zu untersuchen und zu beurteilen.
- *Datenkompetenz* als Fähigkeit, grosse digitale Datenbestände (Big Data) mit geeigneten Mitteln auf ihre Herkunft, Qualität und den Verwendungszweck hin adäquat zu beurteilen und zu visualisieren respektive darzustellen.

Mit diesen Kompetenzen können sie die Herausforderungen, die sich insbesondere in der Langzeitarchivierung von digitalen Objekten stellen, annehmen. Archive haben bereits entsprechende Infrastrukturen in Betrieb oder bauen diese in nächster Zeit auf. Wichtig ist, dass die Institutionen mit einem gesetzlichen Auftrag beauftragt werden, dass die Finanzierung langfristig sichergestellt ist dass jegliche digitalen Objekte in einer solchen Infrastruktur gesichert werden können.

7. Schlussbetrachtungen

*"In short, historical understanding answers the questions basic to all humanistic endeavors: how and why? [...] These basic questions should be the guiding force behind all activities in digital history, from the building of new tools to interpretative projects."*¹¹³³ (Sternfeld, 2014)

Wie? und Warum? sind auch im Digitalen Zeitalter zentrale Fragen für den inhaltlichen Erkenntnisgewinn in der Geschichtswissenschaft. Ebenso zentral bleiben die bereits seit Johann Gustav Droysen etablierten Grundsätze und Methoden auf denen die Quellenkritik des Digitalen aufbaut.¹¹³⁴ Mit der vorliegenden Untersuchung wurden Fragen zu den methodologischen Veränderungen bei der quellenkritischen Analyse eines digitalen Objekts, und wie diese innerhalb des Fachbereichs Geschichte zu konzeptualisieren sind, beantwortet.¹¹³⁵

In einem ersten Schritt wurden die Eigenschaften eines digitalen Objekts definiert und herausgearbeitet, welche Eigenschaften es von bisher bekannten Objekten unterscheidet (kursiv): es ist *genuin digital*, multimodal und -dimensional, *nicht nachvollziehbar manipulierbar*, volatil, *datenträgerunabhängig*, *verlustfrei vervielfältig- und (gleichzeitig) wiederverwendbar* und *von einem spezifischen, digitalen Informationssystem abhängig*. Damit ist ein digitales Objekt eine neue Art von Quelle, die sich durch ihre Eigenschaften und inhaltlichen Beziehungen zu anderen Objekten von bisher bekannten Objekten abgrenzt.

Drei neue Quellentypen wurden in einem zweiten Schritt identifiziert. Erstens den *integrierenden Quellentypus*, der verschiedene Objekte miteinander nicht nachvollziehbar verbindet. Es ist beispielsweise auf einer Webseite durch hypertextuelle Elemente möglich, externe Objekte in die Darstellung zu integrieren, ohne dass dies erkennbar ist. Zweitens den *einobjektreduzierenden Quellentypus*, der sich durch die Eigenschaft der verlustfreien Vervielfältigbarkeit in Klone und die Datenträgerunabhängigkeit eines digitalen Objekts auszeichnet. Mehrere Klone können immer auf einen einzelnen Klon reduziert werden. Damit ist auch die Unterscheidung zwischen Original und Kopie obsolet. Der dritte *Quellentypus beinhaltet unsichtbar codierte Informationen*. Einerseits liegen digitale Objekte sowohl als Darstellung als auch als Code in mehreren Ebenen vor. Im Programmcode können Informationen wie Kommentare hinterlegt sein, die nur dort zu finden sind, oder nur nach einer bestimmten Aktion dargestellt werden. Andererseits können digitale Objekte wie Datenbanken auch Daten oder Informationen enthalten, die nicht in ihrer Gesamtheit, sondern nur durch bestimmte Abfragen mit einem bestimmten Informationssystem ausschnittsweise dargestellt werden.

¹¹³³ Sternfeld, Joshua: Historical Understanding in the Quantum Age, Bd. 3, Nr. 2 (Sommer), Journal of Digital Humanities, 25.08.2014, URL: <http://journalofdigitalhumanities.org/3-2/historical-understanding-in-the-quantum-age/>, Stand: 04.08.2015.

¹¹³⁴ Vgl. bspw.: Haber: Digital Past, 2011, S. 106.

¹¹³⁵ Vgl.: Ebd.

Digitale Objekte werden als *virtuelle, elektronische Überreste* in einem angepassten Quellengattungsschema verortet. Die aufgezeigten grossen Unterschiede zwischen den bisher bekannten traditionell analogen und den andersartigen digitalen Quellen rechtfertigen die Neuordnung des Quellengattungsschemas mit der Unterscheidung von materiell / physischen und immateriell / digitalen Quellen. Des Weiteren können digitale Objekte nicht nur die Funktion von Primär-, Sekundär-, oder Tertiärquellen übernehmen, sondern auch die von *Quartär-* (von einem Algorithmus generierte Objekte) und *Quintärquellen* (Objekte, die durch einen von einem Algorithmus geschaffenen Algorithmus entstanden sind). Diese zwei neuen Funktionen unterstreichen, dass digitale Objekte von bisher bekannten klar abzugrenzen sind.

Die tl;dr-Generation¹¹³⁶ von Wissenschaftlern nutzt im Digitalen Zeitalter zwar hauptsächlich Digitalmedien, aber eher oberflächlich und unkritisch. Wenn Geschichtswissenschaftler aber nicht mehr fähig sind, ihre Quellen intensiv zu studieren sowie hermeneutisch und epistemologisch zu erfassen, wird die Qualität wissenschaftlicher Arbeiten massiv leiden. Deshalb wäre es vorteilhaft, wenn im Studium Informations-, Medien-, Kommunikations-, Informatik- und Datenkompetenzen vermittelt würden. Nur mit diesen Schlüsselkompetenzen ist es möglich, digitale Forschungsressourcen kritisch zu untersuchen, denn ihre genuine Digitalität, ihre Volatilität und dass sie nicht nachvollziehbar manipuliert werden können, kann mit bisherigen Methoden nur ungenügend untersucht werden. Zwar ist keine neue historisch-kritische Methode zu erarbeiten, aber die traditionelle Methode ist mit dem Prozessschritt der *Quellensicherung* zur Gewährleistung der Nachvollziehbarkeit einer digitalen Quelle sowie mit *neuen Fragen und Arbeitsmethoden* aus der Informationstechnik zu ergänzen. Letztere kommen vor allem dann zum Einsatz, wenn *Fragen zu den Eigenschaften des digitalen Objekts, dessen Verwendung innerhalb eines digitalen Informationssystems (inkl. Netzwerken wie das Internet) oder zum System selbst* zu beantworten sind.

Im neuen Prozessschritt der *Quellensicherung* speichert der Forschende potenzielle Quellen meist als *Surrogat* in einem *Trusted Digital Repository (TDR)*, in dem das Objekt mittels eines *persistenten Identifikators (PI)*¹¹³⁷ *dauerhaft*, unter einer *eindeutigen Adresse*, *integer*, *authentisch* und *manipulationssicher* archiviert wird. Mit einem speziellen Verfahren, der in der vorliegenden Arbeit von mir postulierten *Trusted Saving Procedure (TSP)*, wird sichergestellt, dass die Sicherung direkt zwischen dem Anbieter der Quelle und dem Repository erfolgt und nicht durch den Nutzer beeinflusst werden kann. Für Individualforschende könnten alternativ oder ergänzend sogenannte *Trusted Personal Archives (TPA)* entwickelt werden, die die Sicherung von Objekten in ein verlässliches, lokales Softwarearchiv mit

¹¹³⁶ tl;dr: too long; didn't read.

¹¹³⁷ Zusammengesetzt aus: URL des TDR, ISIL-Nummer des Repository, Nutzernummer, Projekt- / Sammlungsnummer, Laufnummer. Fiktive URL: <https://www.tdr.bar.ch/CH-101111-1.00120.AOY4.0218>.

einem dafür angepassten Sicherungsverfahren ermöglichen. Mit den bereits bestehenden oder in Entwicklung begriffenen Infrastrukturen wie dem Data and Service Center for the Humanities der Schweizerischen Akademie der Geistes- und Sozialwissenschaften¹¹³⁸ wurden bereits sehr gute Grundlagen geschaffen. Es wäre aber wünschenswert, wenn nicht nur Forschungsdaten aus institutionell-wissenschaftlichen Projekten, sondern im Sinne eines *Researcher Driven Archiving (RDA)* auch solche von Individualforschenden gesichert werden könnten, die Sicherung automatisch und manipulationsresistent erfolgen könnte und die Finanzierung auf gesetzlicher Basis dauerhaft gewährleistet wäre.

Da die Authentizität eines digitalen Objekts zu einem grossen Teil von den zugehörigen Metadaten abhängt, sollte eine *Metadata Information Section (MIS)* eingeführt werden. Sie enthält in einem separaten, von Menschen und Maschinen lesbaren Abschnitt im Objekt selbst die wichtigsten bibliographischen Informationen. Zusätzlich werden diese, zusammen mit sicherheitstechnischen Angaben, auf der publizierenden Plattform maschinell auslesbar hinterlegt. Trotz Vorbehalten bezüglich der Langzeitarchivierung von Sicherheitsmerkmalen ist die *qualifizierte digitale Signatur mit Zeitstempel* die derzeit beste und einfachste informationstechnische Möglichkeit zur eindeutigen Feststellung verschiedener Metadaten eines Objekts.

Die mit der erweiterten historischen Methode erfolgten Untersuchungen eines digitalen Objekts und seiner Umwelt haben ergeben, dass – im Unterschied zu den analogen Sekundärquellen – seinen Metadaten und Paratexten grundsätzlich nicht zu vertrauen ist und diese deshalb zu verifizieren sind. Die Verifikation des Inhalts, die vorhandenen Sicherheitsmerkmale, die Reputation von Informationsübermittlern und andere vertrauenswürdige Zusatzinformationen geben Hinweise auf die Echtheit eines digitalen Objekts und steigern die Wahrscheinlichkeit, dass es authentisch ist. *In der quellenkritischen Überprüfung eines digitalen Objekts wird durch Verifikation ein Näherungswert erarbeitet, der die Wahrscheinlichkeit der Authentizität des Objekts ausdrückt.* Für die strukturierte und nachvollziehbare Beurteilung der Authentizität eines digitalen – und physischen – Objekts kann die *Authentizitätsapproximation* berechnet werden. Sieben Faktoren wurden entwickelt, die bewertet, gewichtet und miteinander in Beziehung gesetzt werden sollen. Das Resultat ist ein subjektiver Prozentwert, der ausdrückt, für wie authentisch der Nutzer das untersuchte Objekt hält. Dieser Näherungswert ist abhängig von der Fragestellung, von den zur Verfügung stehenden Informationen, von Sicherheitsmerkmalen und auch von den Medien- und informationstechnischen Kompetenzen des Nutzers.

Diese Kompetenzen sind auch bei der Interpretation der Quellen und der Darstellung der Forschungsergebnisse notwendig. Neue Darstellungs- und Visualisierungsformen sowie

¹¹³⁸ Siehe URL: <http://www.sagw.ch/de/sagw/die-akademie/unternehmen/DaSCH.html>.

neue Diskussionsräume bedingen Veränderungen im Schreibprozess und erlauben neue Formen der Interaktion zwischen Autorschaft und Leser, die miteinander manuskriptive Kollaborationen eingehen können. Eine wichtige Veränderung ist auch in der Formulierung der *Belegkette* von digitalen Objekten vorzunehmen, in der Seitenangaben und Zeitangaben entfallen und dafür zusätzlich der Speicherort, ein *Link-Check-Element* und allenfalls die Sammlung, aus der das digitale Objekt stammt, angegeben werden sollte.

Die *Digitalität verändert die Arbeitsweise der historischen Zunft* aufgrund der Integration von *neuen Arbeitsmethoden* und -techniken aus der Informatik in die historisch-kritische Methode. Neue Quellentypen, neue Quellenarten, der neue Prozessschritt der Quellensicherung, neue Paratexte und Hypertextualität fordern neue Arbeitsweisen und Kompetenzen. Gleichzeitig ist zu betonen, dass die traditionellen Arbeitsmethoden und -techniken nicht vernachlässigt werden dürfen.

Peter Habers Frage, ob für die Ausbildung von Geschichtswissenschaftlern an den Universitäten eine 'digitale Hilfswissenschaft' etabliert werden sollte, oder ob digitale Methoden zum Kernbestandteil der Geschichtswissenschaften gehören sollten¹¹³⁹, kann ich basierend auf den obigen Ausführungen nun klar beantworten: Die für die Untersuchung eines digitalen Objekts nötigen Methoden aus der Informationstechnik sowie den Medien-, Kommunikations-, Archiv- und Bibliothekswissenschaften gehören zwingend zum Curriculum eines Historikers und somit zum Kernbestandteil der Geschichtswissenschaften. Daraus folgt, dass *nicht von einer Quellenkritik des Digitalen gesprochen werden sollte, sondern weiterhin von der historisch-kritischen Methode*, die um den Schritt der Quellensicherung erweitert wird sowie Arbeitsmethoden und Techniken aus der Informatik enthält. Die interpretative und hermeneutische Arbeit wird darin mit den informationstechnischen Methoden verschmolzen.¹¹⁴⁰ Da sich die Forschenden immer mehr in digitalen und seltener in ausschliesslich analogen Räumen bewegen, haben sie die dafür nötigen Kompetenzen zu erarbeiten.

Das Digitale Zeitalter ist Realität, weshalb sowohl Wissenschaft wie Politik die Verantwortung für das digitale Erbe zu übernehmen haben. Aus- und Weiterbildungsprogramme sollten geschaffen werden, damit die Akteure der Wissenschaft mit den mittlerweile wichtigsten Forschungsressourcen umzugehen wissen. Auch Geschichtswissenschaftler haben sich neue Kompetenzen anzueignen und den vertretbaren Mehraufwand bei der Anwendung neuer Methoden zu leisten. Die Informationstechnik wird sich laufend weiterentwickeln, insbesondere entstehen ständig neue Datenformate und -formen, oder sie werden überholt und obsolet. Dieses Phänomen beeinflusst insbesondere die digitale Langzeitarchivierung,

¹¹³⁹ Vgl. Haber: *Digital Past*, 2011, S. 106.

¹¹⁴⁰ Nach: Zaagsma: *On Digital History*, 2013, S. 24.

die mit entsprechenden Massnahmen die gesicherten digitalen Objekte dauerhaft authentisch zur Verfügung zu stellen hat. Archivare, Informatiker und Historiker sollten sich deshalb ständig austauschen, damit sowohl der Aufbau und die Weiterentwicklung der Infrastruktursysteme, als auch der Datenimport und das Data Management nach den Anforderungen der Anspruchsgruppen und den informationstechnischen Möglichkeiten der Archive erfolgen können. Nur mit gegenseitigem Wissens- und Informationsaustausch kann gewährleistet werden, dass die digitalen Objekte, im Besonderen die Forschungsdaten, dauerhaft nutzbar bleiben.

Da mittlerweile analoge Suchräume immer häufiger ignoriert und damit auch nicht mehr alle zur Verfügung stehenden Informationen in eine Untersuchung miteinbezogen werden, sollte es auch ein Anliegen der Geschichtswissenschaft sein, dass online zur Verfügung stehende Quellen das breitestmögliche Spektrum an Themengebieten abbilden.¹¹⁴¹ Die Weigerung, physische Originale zu digitalisieren, ist nicht zielführend, denn was im Web nicht gefunden wird, existiert für eine stark zunehmende Personengruppe auch nicht. Zudem trägt eine solche Haltung nicht zur Weiterentwicklung des Wissensstandes bei und ist vor allem dann gefährlich, wenn die Auswahl an digital vorhandenen Objekten ein falsches und unausgewogenes Bild vermittelt. Es besteht folglich dringender Handlungsbedarf, denn es sollten so viele historische Quellen wie möglich digital und vertrauenswürdig zur Verfügung gestellt werden. Für wissenschaftliche Forschungen ist es aber unabdingbar, dass beide Formen uneingeschränkt zugänglich sind, denn je nach Forschungsfrage ist auch das physische Original zwingend zu konsultieren. Deshalb muss der Zugriff auf das physische Original, auch wenn es digitalisiert wurde, weiterhin gewährleistet sein. Die Gedächtnisinstitutionen werden deshalb auch in Zukunft eine sehr wichtige Rolle spielen.

Die Geschichtswissenschaft erhält von verschiedenen Forschungsfeldern, insbesondere den Digital Humanities, vielfältige Impulse. Sehr interessant ist die methodische Annäherung an die Naturwissenschaften und die Informatik, die der Geschichtswissenschaft zwar nicht empirischere, präzisere und zeitgemässere Antworten ermöglichen, wie der Wissenschaftsjournalist Urs Hafner spöttisch bemerkte¹¹⁴², sondern neue Fragestellungen und umfassendere Erkenntnisse erlauben. Wie historische Erkenntnisse mit und aus digitalen Objekten und den zugehörigen Informationssystemen gewonnen werden können, sollte weiter vertiefend und interdisziplinär erforscht werden. Mit neuen digitalen Methoden und Techniken können Historiker jetzt schon ihre eigenen Forschungsfelder ausweiten und tiefer ausloten. Das ist äusserst spannend!

¹¹⁴¹ Gerben Zaagsma spricht von 'das Marginale nicht weiter marginalisieren'. Siehe: Ebd., S. 23.

¹¹⁴² Siehe: Hafner, Urs: Digital History: Der Irrtum der Zeitmaschinen, in: Neue Zürcher Zeitung, Zürich 27.05.2016. URL: <http://www.nzz.ch/feuilleton/zeitgeschehen/digital-history-historiografie-des-zeitfeils-id.85000>, Stand: 27.05.2016.

8. Literaturverzeichnis

- Ackermann, Judith: Avatars: Identitäten und digitale (Ab-)Bilder im Netz, in: Stiegler, Christian; Breitenbach, Patrick; Zorbach, Thomas (Hg.): New Media Culture: Mediale Phänomene der Netzkultur, Bielefeld 2015 (Digitale Gesellschaft), S. 51-65.
- Adobe Systems Incorporated: Digital Advertising Report. Adobe Digital Index Q2 2015. URL: http://www.cmo.com/content/dam/CMO_Other/ADI/ADI_Mobilegeddon/Q2-ADI-Digital-Ad-Report.pdf, Stand: 24.07.2015.
- Akademien der Wissenschaften Schweiz (Hg.): Wissenschaftliche Integrität. Grundsätze und Verfahrensregeln, 2008. URL: http://www.snf.ch/de/derSnf/forschungspolitische_positionen/wissenschaftliche_integritaet/Seiten/default.aspx.
- Alvarado, Rafael C.: The Digital Humanities Situation, in: Gold, Matthew K. (Hg.): Debates in the digital humanities, Minneapolis 2012, S. 50-55. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- American Historical Association: Guidelines for the Professional Evaluation of Digital Scholarship in History – Draft, 04.2015. URL: <http://historians.org/Documents/Teaching and Learning/Current Projects/Digital Scholarship Evaluation/Guidelines for the Professional Evaluation of Digital Scholarship in History.pdf>, Stand: 27.07.2015.
- Amrein, Marcel: Archivzugang mit Hindernissen, in: Neue Zürcher Zeitung, Zürich 28.07.2014, S. 7.
- Amrein, Martin: So tricksen Schweizer Forscher die Hochschulen aus, NZZ am Sonntag, 03.02.2018, URL: <https://nzzas.nzz.ch/wissen/so-tricksen-schweizer-forscher-hochschulen-aus-ld.1353872>, Stand: 07.02.2018.
- Andermann, Heike; Degkwitz, Andreas: Neue Ansätze in der wissenschaftlichen Informationsversorgung: ein Überblick über Initiativen und Unternehmungen auf dem Gebiet des elektronischen Publizierens, in: Historical Social Research, 2004.
- Anderson, Chris: The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, WIRED MAGAZINE, 27.06.2008, URL: http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory, Stand: 07.06.2014.
- Andress, Jason (Hg.): The basics of information security: understanding the fundamentals of InfoSec in theory and practice, Amsterdam 2011.
- Ankerson, Megan: Writing web histories with an eye on the analog past, in: New Media & Society (May 2012), 12.04.2012, S. 384-400. DOI: 10.1177/1461444811414834.
- Arbeitsgruppe zum Urheberrecht: Schlussbericht AGUR12, Eidgenössisches Institut für Geistiges Eigentum, Bern 28.11.2013. URL: https://www.ige.ch/fileadmin/user_upload/recht/national/d/urheberrecht/Schlussbericht_der_AGUR12_vom_28_11_2013_DE.pdf, Stand: 06.05.2018.
- Arend, Sabine: Rezension zu: Sahle, Patrick: Digitale Editionsformen, Rezensionen, H / SOZ / KULT, 14.08.2014, URL: <http://hsozkult.geschichte.hu-berlin.de/rezensionen/2014-3-109>, Stand: 14.08.2014.
- Arms, Caroline R.: Keeping Memory Alive: Practices for Preserving Digital Content at the National Digital Library Program of the Library of Congress, in: RLG DigiNews Volume 4 (Number 3), 15.06.2000. URL: <http://webdoc.gwdg.de/edoc/aw/rlgdn/preserv/diginews/diginews4-3.html#feature1>, Stand: 23.07.2014.
- Arnold, Klaus: Quellen, in: Jordan, Stefan (Hg.): Lexikon Geschichtswissenschaft: Hundert Grundbegriffe, Stuttgart 2007, S. 251-255.
- Ashenfelder, Mike: The Personal Digital Archiving 2015 Conference, The Signal, 03.08.2015, URL: <https://blogs.loc.gov/thesignal/2015/08/report-on-the-personal-digital-archiving-2015-conference/>, Stand: 06.05.2018.
- Australian Government Productivity Commission: Intellectual Property Arrangements, Report 78, Canberra 23.09.2016. URL: <http://www.pc.gov.au/inquiries/completed/intellectual-property/report>, Stand: 26.06.2016.
- Baacke, Dieter: Medienkompetenz – Begrifflichkeit und sozialer Wandel, in: Rein, Antje von (Hg.): Medienkompetenz als Schlüsselbegriff, Bad Heilbrunn 1996 (Theorie und Praxis der Erwachsenenbildung), S. 112-124.

- Baacke, Dieter: Medienpädagogik, Tübingen 1997 (Grundlagen der Medienkommunikation).
- Bachfeld, Daniel: 2,2 Millionen URLs bei URL-Verkürzerdienst manipuliert, Newsticker, heise Security, 17.06.2009, URL: <http://www.heise.de/security/meldung/2-2-Millionen-URLs-bei-URL-Verkuerzerdienst-manipuliert-182569.html>, Stand: 30.04.2015.
- Baecker, Dirk: Metadaten. Eine Annäherung an Big Data, in: Geiselberger, Heinrich; Moorstedt, Tobias (Hg.): Big Data: das neue Versprechen der Allwissenheit, Berlin 2013 (edition unseld), S. 156-186.
- Bargheer, Margo; Bellem, Saskia; Schmidt, Birgit: Open Access und Institutional Repositories – Rechtliche Rahmenbedingungen, in: Spindler, Gerald (Hg.): Rechtliche Rahmenbedingungen von Open Access-Publikationen, Göttingen 2006 (Göttinger Schriften zur Internetforschung), S. 1-20.
- Baron, Naomi S.: The Plague of tl;dr, in: The Chronicle of Higher Education, 09.02.2015. URL: <http://chronicle.com/article/The-Plague-of-tl-dr/151635/>, Stand: 07.04.2015.
- Barrenechea, Mark: Enterprise Information Management (EIM): The Next Generation of Enterprise Software, OpenText White Paper Series, 12.2012. URL: <http://www.legaltechnology.com/wp-content/uploads/2013/07/OpenText-EIM-Summary.pdf>, Stand: 02.11.2015.
- Bassler, Moritz: Die kulturpoetische Funktion und das Archiv: eine literaturwissenschaftliche Text-Kontext-Theorie, Tübingen 2005 (Studien und Texte zur Kulturgeschichte der deutschsprachigen Literatur).
- Baumann, Jan; Hügi, Jasmin: compas: Strukturiertes Forschen im Web, infoclio, 03.2015, URL: <http://www.compas.infoclio.ch/de>, Stand: 24.03.2015.
- Baumgärtel, Tilman: Per Tweet ins Museum: Internetkunst, in: Neue Zürcher Zeitung, 06.09.2013. URL: http://www.nzz.ch/aktuell/feuilleton/kunst_architektur/per-tweet-ins-museum-1.18145506?extcid=Newsletter_06092013_Top-News_am_Morgen, Stand: 06.03.2014.
- Beck, Klaus: Medien, in: Lexikon Kommunikations- und Medienwissenschaft, Wiesbaden 2013 (Studienbücher zur Kommunikations- und Medienwissenschaft), S. 201-202.
- Becker, Rainer; Castilho, Richard Eckart de: Archive-in-a-box (Internal Demonstrator) – Service-Concept M 1.4.1.1, 31.08.2012. URL: https://dev2.dariah.eu/wiki/download/attachments/14651583/M1411_Archive_in_a_box.pdf?version=2&modificationDate=1366904206713&api=v2.
- Beinecke, Julius: Zahlen, bitte! 3.333.360 Punkte bis zum «Kill-Screen» von Pac-Man, News, heise online, 28.06.2016, URL: <http://www.heise.de/newsticker/meldung/Zahlen-bitte-3-333-360-Punkte-bis-zum-Kill-Screen-von-Pac-Man-3227685.html>, Stand: 17.07.2016.
- Benjamin: L'œuvre d'art à l'époque de sa reproduction mécanisée, in: Zeitschrift für Sozialforschung (5), 1936, S. 40-68.
- Bentele, Günter; Brosius, Hans-Bernd; Jarren, Otfried (Hg.): Lexikon Kommunikations- und Medienwissenschaft, Wiesbaden 2013 (Studienbücher zur Kommunikations- und Medienwissenschaft).
- Bergman, Michael K.: White Paper: The Deep Web: Surfacing Hidden Value, in: The Journal of Electronic Publishing 7 (1), 01.08.2001. DOI: 10.3998/3336451.0007.104.
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, Berliner Erklärung, Open Access – Max-Planck-Gesellschaft, 22.10.2003, URL: <http://openaccess.mpg.de/Berliner-Erklaerung>, Stand: 18.11.2014.
- Bernheim, Ernst: Lehrbuch der historischen Methode und der Geschichtsphilosophie mit Nachweis der wichtigsten Quellen und Hilfsmittel zum Studium der Geschichte, Leipzig 1903.
- Berry, David M.: Understanding digital humanities, Houndmills, Basingstoke, Hampshire - New York 2012.
- Berry, David M.: The Computational Turn: Thinking About the Digital Humanities, in: Culture Machine 12 (0), 18.02.2011. ew/440, Stand: 14.01.2014.
- Bibliothek der Universität Konstanz: Informationskompetenz an der Bibliothek der Universität Konstanz, Bibliothek der Universität Konstanz, 04.06.2008, URL: <http://www.ub.uni-konstanz.de/bibliothek/projekte/informationskompetenz/>, Stand: 18.08.2015.

- Bidder, Benjamin: Interview zu angeblichen MH17-Manipulationen: «Bellingcat betreibt Kaffeesatzleserei», in: Spiegel Online, 03.06.2015. URL: <http://www.spiegel.de/politik/ausland/mh17-satellitenbilder-bellingcat-betreibt-kaffeesatzleserei-a-1036874.html>, Stand: 01.07.2015.
- Birkerts, Sven: *The Gutenberg elegies the fate of reading in an electronic culture*, Boston (Mass.) [etc.] 1994.
- Bischoff, Frank M.: Bewertung elektronischer Unterlagen und die Auswirkungen archivarischer Eingriffe auf die Typologie zukünftiger Quellen, in: *Archivar, Zeitschrift für Archivwesen* 1, 2014, S. 40-52.
- Björk, Bo-Christer; Solomon, David: Article processing charges in OA journals: relationship between price and quality, in: *Scientometrics* 103 (2), 13.03.2015, S. 373-385. DOI: 10.1007/s11192-015-1556-z.
- Blaney, Jonathan: *The Problem of Citation in the Digital Humanities*, Proceedings of the Digital Humanities Congress 2012. Studies in the Digital Humanities, Sheffield; HRI Digital, 2014, URL: <http://www.hrionline.ac.uk/openbook/chapter/dhc2012-blaney>, Stand: 16.08.2015.
- Bloch, Marc: *Aus der Werkstatt des Historikers. Zur Theorie und Praxis der Geschichtswissenschaft*, Frankfurt a.M. 2000.
- Böhler, Fritz; Engler, Balz; Glauser, Jürg u. a.: *Positionspapier – Für eine Erneuerung der Geisteswissenschaften. Empfehlungen der SAGW zuhanden der Leitungsorgane der Hochschulen, der Lehrenden, der Förderorganisationen und des Staatssekretariats für Bildung und Forschung*, 05.12.2012. URL: www.sagw.ch/dms/sagw/schwerpunkte/wissenschaft_und_technikgeschichte/publikationen/Positionspapier_mit_Ruecken, Stand: 06.05.2018.
- Böhnke, Alexander; Schröter, Jens (Hg.): *Analog / Digital – Opposition oder Kontinuum? Zur Theorie und Geschichte einer Unterscheidung*, Bielefeld 2004 (Medienumbrüche). URL: <http://www.fk615.uni-siegen.de/de/publikationeinzeln.php?band=1>, Stand: 19.04.2013.
- Bolz, Lisa; Foerster, Sascha: *Die Blogplattform der deutschsprachigen Geistes- und Sozialwissenschaften*, Workshop, DHd2015, Von Daten zu Erkenntnissen: Workshop II, Graz 23.02.2015. URL: <http://gams.uni-graz.at/o:dhd2015.w.02>, Stand: 30.07.2015.
- Boonstra, Onno; Breure, Leen; Doorn, Peter: Past, present and future of historical information science, article, *Historical Social Research*, 2004, URL: <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-30846>, Stand: 06.05.2018.
- Broadband Commission for Digital Development: *The State of Broadband 2015: Broadband as a Foundation for Sustainable Development*, ITU & UNESCO, 21.09.2015. URL: <http://www.broadbandcommission.org/publications/Pages/SOB-2015.aspx>, Stand: 06.05.2018.
- Borowsky, Peter; Vogel, Barbara; Wunder, Heide: *Einführung in die Geschichtswissenschaft I: Grundprobleme, Arbeitsorganisation, Hilfsmittel*, Opladen 1989 (Studienbücher moderne Geschichte 1).
- Bortnikov, Vyacheslav: *Netzneutralität und Bedingungen kommunikativer Selbstbestimmung: Pflichten des freiheitlichen Verfassungsstaates zur Gewährleistung der Neutralität des Internets im Lichte der grundrechtlichen Schutzpflichtenlehre*, München 2013 (Schriftenreihe des Instituts für Rundfunkrecht an der Universität zu Köln).
- boyd, danah; Crawford, Kate: *Six Provocations for Big Data*, SSRN Scholarly Paper ID 1926431, Social Science Research Network, Rochester, NY 21.09.2011, S. 17. URL: <http://papers.ssrn.com/abstract=1926431>, Stand: 24.02.2015.
- Brabazon, Tara: *The university of google. Education in the (post) information age*, Aldershot 2008.
- Brey, Mariano Rajoy: *Modifica el texto refundido de la Ley de Propiedad Intelectual (11404)*, in: *Boletín Oficial Del Estado* (268), 05.11.2014, S. 90404-90439.
- Briegleb, Volker: *EU-Länder einigen sich auf Eckpunkte der Datenschutzreform*, Newsticker, heise online, 13.03.2015, URL: <http://www.heise.de/newsticker/meldung/EU-Laender-einigen-sich-auf-Eckpunkte-der-Datenschutzreform-2574894.html>, Stand: 01.05.2015.
- Brown, Adrian: *Practical Digital Preservation: a How-to Guide for Organizations of any Size*, London 2013.

- Brügger, Niels: Web Archiving, Web History, and the Web in Digital Humanities, Präsentation, Digital Humanities Luxembourg – Symposium, Luxembourg 21.03.2012.
- Brügger, Niels: Web History and the Web as a Historical Source, in: Zeithistorische Forschungen/Studies in Contemporary History, Online-Ausgabe 9 (2), 2012. URL: <http://www.zeithistorische-forschungen.de/16126041-Bruegger-2-2012>, Stand: 06.05.2018.
- Brügger, Niels: When the Present Web is Later the Past: Web Historiography, Digital History, and Internet Studies, in: Historical Social Research 37 (4), 2012, S. 102-117.
- Brügger, Niels: Web history, New York 2010.
- Brügger, Niels: Website Analysis. Elements of a conceptual architecture, Aarhus 2010 (CFI monograph series 12). URL: <http://cfi.au.dk/publications/monographseries/#a12>, Stand: 24.09.2015.
- Brügger, Niels: Archiving Websites. General Considerations and Strategies, Aarhus 2005.
- Bruns, Axel: Blogs, Wikipedia, Second Life, and Beyond from Production to Prousage, New York 2008 (Digital formations).
- Buchmann, Bertrand Michael: Einführung in die Geschichte, Wien 2002 (Facultas-Manuals).
- Budapest Open Access Initiative; German Translation, Budapest Open Access Initiative, 2012, URL: <http://www.budapestopenaccessinitiative.org/boai-10-translations/german-translation>, Stand: 06.07.2015.
- Bundesamt für Kommunikation: Netzneutralität, Bericht zur Arbeitsgruppe, Schweizerische Eidgenossenschaft, Bern 23.10.2014. URL: <http://www.bakom.admin.ch/themen/internet/04810/index.html?lang=de>, Stand: 30.06.2015.
- Bundesamt für Sicherheit in der Informationstechnik: Sicherheitsanalyse TrueCrypt, Publikationen, Bundesamt für Sicherheit in der Informationstechnik, 16.11.2015, URL: <https://www.bsi.bund.de/DE/Publikationen/Studien/TrueCrypt/truecrypt.html>, Stand: 18.04.2016.
- Bundesministerium der Justiz und für Verbraucherschutz: Gesetz über Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz), 31.08.2015. URL: <http://www.gesetze-im-internet.de/urhg/>, Stand: 09.10.2015.
- Bundesministerium für Justiz: Bundesgesetz über das Urheberrecht an Werken der Literatur und der Kunst und über verwandte Schutzrechte (Urheberrechtsgesetz), 13.08.2015. URL: <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10001848>, Stand: 09.10.2015.
- Bundesversammlung der Schweizerischen Eidgenossenschaft: Bundesgesetz über die Erfindungspatente (Patentgesetz), 01.01.2012. URL: <https://www.admin.ch/opc/de/classified-compilation/19540108/index.html>, Stand: 06.10.2015.
- Bundesversammlung der Schweizerischen Eidgenossenschaft: Bundesgesetz über das Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz), 01.01.2011. URL: <https://www.admin.ch/opc/de/classified-compilation/19920251/index.html>, Stand: 09.10.2015.
- Bundesversammlung der Schweizerischen Eidgenossenschaft: Bundesgesetz über Zertifizierungsdienste im Bereich der elektronischen Signatur (Bundesgesetz über die elektronische Signatur, ZertES), 01.08.2008. URL: <http://www.admin.ch/opc/de/classified-compilation/20011277/index.html>, Stand: 24.04.2015.
- Burchert, Heiko; Sohr, Sven: Praxis des wissenschaftlichen Arbeitens: Eine anwendungsorientierte Einführung, München 2008 (Studien- und Übungsbücher der Wirtschafts- und Sozialwissenschaften).
- Burckhardt, Daniel; Schiel, Juliane: Kollaboratives Schreiben, in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken: für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 83-95.
- Burkhardt, Marcus: Digitale Datenbanken: eine Medientheorie im Zeitalter von Big Data, Bielefeld 2015 (Digitale Gesellschaft).
- Bush, Vannevar: Wie wir denken sollten, in: Wirth, Uwe (Hg.): Kulturwissenschaft, Frankfurt a.M. 2008, S. 392-416.

- Butler, Declan: Investigating journals: The dark side of publishing, in: *Nature* 495 (7442), 27.03.2013, S. 433-435. DOI: 10.1038/495433a.
- Büttner, Sabine: Einführung: Arbeiten mit Quellen, *historicum-estudies.net*, 12.02.2007, URL: <http://www.historicum-estudies.net/etutorials/tutorium-quellenarbeit/?L=962>, Stand: 06.05.2018.
- Capurro, Rafael: Digital Hermeneutics – An Outline, Rafael Capurro, 09.11.2017, URL: <http://www.capurro.de/digitalhermeneutics.html>, Stand: 07.02.2018.
- Carl, Janusch; Rutz, Andreas: Bits and Bytes statt Pergament und Papier? Das digitale historische Archiv Köln und die Zukunft des Kölner Stadtarchivs im Web 2.0, in: *Archivar, Zeitschrift für Archivwesen* 2, 2012, S. 143-153.
- Carpenter, Julie; Wetheridge, Louise; Tanner, Sophie: Researchers of Tomorrow: The research behaviour of Generation Y doctoral students, JISC, Bristol, London 28.06.2012. URL: <http://www.jisc.ac.uk/publications/reports/2012/researchers-of-tomorrow>, Stand: 04.03.2014.
- Carr, Nicholas: Is Google Making Us Stupid?, in: *The Atlantic*, 08.2008. URL: <http://www.theatlantic.com/magazine/archive/2008/07/is-google-making-us-stupid/306868/>, Stand: 06.06.2014.
- Carvajal, Doreen: Anne Frank's Diary Gains 'Co-Author' in Copyright Move, in: *The New York Times*, New York (USA) 13.11.2015. URL: <http://www.nytimes.com/2015/11/14/books/anne-frank-has-a-co-as-diary-gains-co-author-in-legal-move.html>, Stand: 04.12.2015.
- Case, Donald O.: The Concept of Information (2012), in: *Information Science*, Bd. I / IV, London 2014 (Critical concepts in media and cultural studies), S. 15-42.
- Cayley, John: The Code is not the Text (unless it is the Text), electronic book review, 10.09.2002, URL: <http://www.electronicbookreview.com/thread/electropoetics/literal>, Stand: 25.08.2014.
- Centre for Educational Research and Innovation of the OECD: Are the New Millennium Learners Making the Grade? Technology Use And Educational Performance in PISA, Paris 2010. URL: <http://www.oecd-ilibrary.org/content/book/9789264076044-en>, Stand: 08.05.2015.
- Chan, Leslie; Cuplinskas, Darius; Eisen, Michael u. a.: Read the Budapest Open Access Initiative, Budapest Open Access Initiative, 14.02.2002, URL: <http://www.budapestopenaccessinitiative.org/>, Stand: 20.03.2015.
- Christiansen, Per; Gehlen, Dirk von; Hofmann, Jeanette u. a.: Ein Gedankenexperiment zur Neuordnung des Urheberrechts, *irights.info*, 27.08.2015. URL: <http://irights.info/2015/08/29/ein-gedankenexperiment-zur-neuordnung-des-urheberrechts/25980>, Stand: 01.09.2015.
- Clerwall, Christer: Enter the Robot Journalist, in: *Journalism Practice* 8 (5), 03.09.2014, S. 519-531. DOI: 10.1080/17512786.2014.883116.
- CNN: Web newspaper registration stirs debate, *CNN.com*, 14.06.2004, URL: http://chnm.gmu.edu/digitalhistory/links/cached/chapter6/6_28a_registration.htm, Stand: 17.04.2013.
- Cohen, Daniel J.: Collecting History Online, in: *Clio wired – The future of the Past in the Digital Age*, 2011, S. 124-151.
- Cohen, Daniel J.; Frisch, Michael; Gallagher, Patrick u. a.: Interchange: The Promise of Digital History, in: *The Journal of American History* 95 (2), 01.09.2008, S. 452-491. DOI: 10.2307/25095630.
- Cohen, Daniel J.; Rosenzweig, Roy: *Digital history: A Guide to Gathering, Preserving, and Presenting the Past on the Web*, Philadelphia 2006.
- Collins, Ellen; Jubb, Michael: How do Researchers in the Humanities Use Information Resources?, in: *Liber Quarterly* 21 (2), 2012, S. 176-187.
- Collins, Sandra; Harrower, Natalie; Haug, Dag Trygve Truslew u. a.: Going Digital: Creating Change in the Humanities, ALLEA E-Humanities Working Group Report, ALLEA – All European Academies, Berlin 07.05.2015. URL: http://www.allea.org/wp-content/uploads/2015/07/Going-Digital_digital-version.pdf, Stand: 06.05.2018.
- Cope, Bill; Phillips, Angus (Hg.): *The future of the academic journal*, Oxford 2014 (Chandos information professional series).

- Cope, Bill; Phillips, Angus: Introduction, in: Cope, Bill; Phillips, Angus (Hg.): The future of the academic journal, Oxford 2014 (Chandos information professional series), S. 1-8.
- Cornelissen, Christoph: Geschichtswissenschaften: Eine Einführung, Frankfurt a.M. 2004.
- Corrado, Greg: Computer, respond to this email., Research Blog, 03.11.2015, URL: <http://googleresearch.blogspot.com/2015/11/computer-respond-to-this-email.html>, Stand: 03.11.2015.
- Cowburn, Pam: ORG's Blocked project finds almost 1 in 5 sites are blocked by filters, Blog, Open Rights Group, 02.07.2014, URL: <https://www.openrightsgroup.org/blog/2014/blockedproject>, Stand: 08.05.2015.
- Cristofoli, Pascal: Principes et usages des dessins de réseaux en SHS, in: Natale, Enrico; Sibille, Christiane; Chacherau, Nicholas u. a. (Hg.): Visualisierung von Daten in der Geschichtswissenschaft / La Visualisation des données en histoire, Zürich 2015 (Geschichte und Informatik), S. 23-57.
- Cullen, Charles T.; Hirtle, Peter B.; Levy, David u. a.: Authenticity in a Digital Environment, Washington, D.C 2000.
- Cummings, Alex Sayf; Jarrett, Jonathan: Only Typing? Informal Writing, Blogging, and the Academy, in: Dougherty, Jack; Nawrotzki, Kristen (Hg.): Writing history in the digital age, Ann Arbor 2013 (Digital humanities), S. 246-258.
- DARIAH-DE: DARIAH-DE – Aufbau von Forschungsinfrastrukturen für die eHumanities, in, Göttingen 2012. DOI: 10.3249/webdoc-3589.
- Darnton, Robert: Die Bibliothek im Informationszeitalter. 6000 Jahre Schrift, in: Becker, Konrad (Hg.): Deep Search: Politik des Suchens jenseits von Google, Innsbruck 2009, S. 37-52.
- Data; Protection Working Party (DAPIX); Data: Proposal from Presidency to Council: General Data Protection Regulation, 09.03.2015. URL: <http://data.consilium.europa.eu/doc/document/ST-6834-2015-INIT/en/pdf>.
- Dauer, Antje: Pressemitteilung der Herzog August Bibliothek Wolfenbüttel: Freie Nutzung von Digitalisaten dank neuer Lizenzen, 23.05.2013. URL: <http://www.hab.de/files/2013-05-22-creativecommonslizenz.pdf>, Stand: 20.03.2015.
- Davidson, Cathy N.: Humanities 2.0: Promise, Perils, Predictions, in: Gold, Matthew K. (Hg.): Debates in the digital humanities, Minneapolis 2012, S. 476-489. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- Day, Andreas: The Long-Term Preservation of Web Content, in: Masanès, Julien (Hg.): Web Archiving, Berlin 2006, S. 177-199.
- De Cristofaro, Emiliano; Friedman, Arik; Jourjon, Guillaume u. a.: Paying for Likes? Understanding Facebook Like Fraud Using Honeypots, in: ACM Internet Measurement Conference 2014, arXiv:1409.2097 [cs.SI], Vancouver 2014. URL: <http://arxiv.org/abs/1409.2097>, Stand: 30.06.2015.
- Dean, Brian: Google Ranking Factors: The Complete List, Backlinko, 21.03.2016, URL: <http://backlinko.com/google-ranking-factors>, Stand: 28.03.2016.
- Die Grünen - Europäische Freie Allianz: Kreatives Schaffen und Urheberrecht im digitalen Zeitalter – Positionspapier, 21.11.2011. URL: <http://www.greens-efa.eu/de/kreatives-schaffen-und-urheberrecht-im-digitalen-zeitalter-4539.html>, Stand: 31.08.2015.
- Diederich, Jörg; Müller-Rowold, Boris: Kriterien für wissenschaftliche Arbeiten, Text, Stiftung Universität Hildesheim, 16.09.2010, URL: <http://www.uni-hildesheim.de/index.php?id=2239>, Stand: 30.11.2012.
- Dommann, Monika: Autoren und Apparate. Die Geschichte des Copyrights im Medienwandel, Frankfurt a.M. 2014 (Fischer Wissenschaft).
- Dommann, Monika: Schlüssel für Clio: Digitale Daten und die Suche nach einer neuen Quellenkritik, in: Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Digital Humanities und Web 2.0, 2012.
- Dommann, Monika; Gugerli, David: Geschichtswissenschaft in Begutachtung: acht Kommentare zur historischen Methode der Gegenwart, in: Traverse, 2011.

- Dong, Xin Luna; Gabrilovich, Evgeniy; Murphy, Kevin u. a.: Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources, arXiv:1502.03519 [cs.DB], 11.02.2015. URL: <http://arxiv.org/abs/1502.03519>, Stand: 03.03.2015.
- Dougherty, Jack; Nawrotzki, Kristen (Hg.): Writing history in the digital age, Ann Arbor 2013 (Digital humanities).
- Dougherty, Jack; Nawrotzki, Kristen; Rochez, Charlotte D. u. a.: Conclusions. What We Learned from Writing History in the Digital Age, in: Dougherty, Jack; Nawrotzki, Kristen (Hg.): Writing history in the digital age, Ann Arbor 2013 (Digital humanities), S. 259-277.
- Droysen, Johann Gustav: Grundriss der Historik, Leipzig 1875².
- Düring, Marten: From Hermeneutics to Data to Networks: Data Extraction and Network Visualization of Historical Sources, Programming Historian, 18.02.2015, URL: <http://programminghistorian.org/lessons/creating-network-diagrams-from-historical-sources.html>, Stand: 17.06.2015.
- Eden, Bradford Lee (Hg.): Twenty-first century metadata operations challenges, opportunities, directions, London 2012.
- Eder, Franz X.; Berger, Heinrich; Casutt-Schneeberger, Julia u. a.: Geschichte Online: Einführung in das wissenschaftliche Arbeiten, Wien, Köln, Weimar 2006.
- Editorial Board: A prestigious research publisher gives in to China's censorship, The Washington Post, 04.11.2017, URL: https://www.washingtonpost.com/opinions/a-prestigious-research-publisher-gives-into-chinas-censorship/2017/11/04/4e4ff9de-c0c6-11e7-959c-fe2b598d8c00_story.html, Stand: 07.02.2018.
- Egloff, Wili: Der Zugang zu Fachinformationen – zwischen gesellschaftlicher Notwendigkeit und privatem Verwertungsinteresse, in: Gartska, Hansjürgen; Coy, Wolfgang (Hg.): Wovon – für wen – wozu. Systemdenken wider die Diktatur der Daten. Wilhelm Steinmüller zum Gedächtnis, Berlin 2014, S. 349-373. URL: <http://edoc.hu-berlin.de/docviews/abstract.php?lang=ger&id=40657>, Stand: 19.03.2015.
- Eh, Doris; Schütte, Simone: Literatur finden, in: Franck, Norbert; Sary, Joachim (Hg.): Die Technik wissenschaftlichen Arbeitens eine praktische Anleitung, Paderborn 2013 (UTB 724), S. 33-64.
- Eickels, Klaus van; Weichselbaumer, Ruth; Bennewitz, Ingrid (Hg.): Mediaevistik und Neue Medien, Ostfildern 2004.
- Eidgenössisches Finanzdepartement EFD: Verordnung des EFD über elektronische Daten und Informationen (EIDI-V), 01.01.2010. URL: <http://www.admin.ch/opc/de/classified-compilation/20092054/index.html>.
- Eidgenössisches Institut für Geistiges Eigentum: Bundesgesetz über das Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz, URG) – Entwurf, 11.12.2015. URL: <https://www.ige.ch/urheberrecht/modernisierung-des-urheberrechts-2015.html>, Stand: 03.04.2016.
- Enderle, Wilfried: Fakten und Informationen im digitalen Raum, in: Busse, Laura; Enderle, Wilfried; Hohls, Rüdiger u. a. (Hg.): Clio Guide. Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften, Berlin 2016 (Historisches Forum 19). URL: <http://guides.clio-online.de/guides/arbeitsformen-und-techniken/fakten-und-informationen/2016>, Stand: 10.07.2016.
- Epple, Angelika: Verlinkt, vernetzt, verführt – verloren? Innovative Kraft und Gefahren der Online-Historiographie, in: Haber, Peter; Epple, Angelika (Hg.): Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0 Digital Humanities, Zürich 2004 (Geschichte und Informatik 15-Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0), S. 15-32.
- EuGH: Urteil des Europäischen Gerichtshofes; Rechtssache C-131/12, InfoCuria – Rechtsprechung des Gerichtshofs, 13.05.2014, URL: <http://curia.europa.eu/>.
- EU-Ratspräsidentschaft (Lettland): Proposal from Presidency to Permanent Representatives Committee: General Data Protection Regulation – Chapter III, 27.04.2015. URL: <http://data.consilium.europa.eu/doc/document/ST-6834-2015-INIT/en/pdf>.
- European Commission: Inquiry on Electronic Mass Surveillance of EU Citizens – Exchange of views on the Russian communications interception practices (SORM), Audiovisual Services, 22.01.2014, URL: <http://ec.europa.eu/avservices/video/shotlist.cfm?ref=95402>, Stand: 13.10.2015.

- Evens, Aden: Web 2.0 and the Ontology of the Digital, in: DHQ Digital Humanities 6 (Futures of Digital Studies: 2) (2), 2012. URL: <http://www.digitalhumanities.org/dhq/vol/6/2/000120/000120.html>, Stand: 17.08.2012.
- Fahrmeir, Andreas: Bloggen und Open Peer Review in der Geschichtswissenschaft: Chance oder Sackgasse?, in: Haber, Peter; Pfanzelter, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013, S. 23-35.
- Falkvinge, Rick: The Entire Copyright Monopoly Idea is Based on a Colossal Lie, Opinion, TorrentFreak, 21.06.2015, URL: <https://torrentfreak.com/the-entire-copyright-monopoly-idea-is-based-on-a-colossal-lie-150621/>, Stand: 22.06.2015.
- Fanger, Reto: Digitale Dokumente als Beweis im Zivilprozess, Basel 2005 (Basler Studien zur Rechtswissenschaft).
- Fehrmann, Gisela; Linz, Erika; Schumacher, Eckhard u. a.: Originalkopie: Praktiken des Sekundären – Einleitung, in: Originalkopie: Praktiken des Sekundären, Bd. 11, Köln 2004 (Mediologie), S. 7-17.
- Fehrmann, Gisela; Linz, Erika; Schumacher, Eckhard u. a.: Originalkopie: Praktiken des Sekundären, Bd. 11, Köln 2004 (Mediologie).
- Fickers, Andreas: Towards A New Digital Historicism? Doing History In The Age Of Abundance, in: Journal Of European Television History & Culture (1), 2012, S. 19-26.
- Finanzdepartement der Schweizerischen Eidgenossenschaft: Expertengruppe "Zukunft der Datenbearbeitung und Datensicherheit" eingesetzt, Medienmitteilungen, 27.08.2015, URL: <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-58475.html>, Stand: 28.08.2015.
- Fischer, Martin: Bitte 10 Bit! Eingeschränkte Farbwiedergabe beim Mac Pro, in: c't: Magazin für Computertechnik (10), 2015, S. 36.
- Fischer, Peter; Hofer, Peter: Lexikon der Informatik, Berlin Heidelberg 2011.
- Fitzpatrick, Kathleen: The Humanities, Done Digitally, in: Gold, Matthew K. (Hg.): Debates in the digital humanities, Minneapolis 2012, S. 12-15. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- Fitzpatrick, Kathleen: The Digital Future of Authorship: Rethinking Originality, in: Culture Machine 12 (0), 18.02.2011. URL: <http://www.culturemachine.net/index.php/view/fulltext/article/111>, Stand: 17.04.2013.
- Föhr, Pascal: Anpassung Urheberrechtsgesetz, Historical Source Criticism, 01.03.2016, URL: <http://hsc.hypotheses.org/382>, Stand: 26.03.2016.
- Föhr, Pascal: Rezension zu: Groebner, Valentin: Wissenschaftssprache digital. Die Zukunft von gestern, in: Schweizerische Zeitschrift für Geschichte 65 (1), 2015, S. 144-145.
- Föhr, Pascal: 'just a historian', Historical Source Criticism, 09.08.2013, URL: <http://hsc.hypotheses.org/241>, Stand: 06.11.2014.
- Föhr, Pascal: Historiker müssen HTML interpretieren können, Billet, 21.03.2012, URL: <http://hsc.hypotheses.org/38>, Stand: 19.08.2015.
- Foster, Ian: How Computation Changes Research, in: Bartscherer, Thomas; Coover, Roderick (Hg.): Switching codes: Thinking through digital technology in the humanities and the arts, Chicago 2012, S. 15-37.
- Franck, Norbert; Stary, Joachim (Hg.): Die Technik wissenschaftlichen Arbeitens: Eine praktische Anleitung, Paderborn 2013 (UTB 724).
- Freeman, Elisabeth: HTML mit CSS und XHTML von Kopf bis Fuss, Beijing 2006 (Ein Buch zum Mitmachen und Verstehen).
- Freytag, Nils: Kursbuch Geschichte. Tipps und Regeln für wissenschaftliches Arbeiten, Paderborn 2011 (UTB 2569, Ed. 5).
- Friedman, Bruce: How Google is Changing Our Information-Seeking Behavior, Lab Soft News, 06.02.2008, URL: http://labsoftnews.typepad.com/lab_soft_news/2008/02/how-google-is-c.html, Stand: 20.03.2015.

- Fritch, John W.; Cromwell, Robert L.: Evaluating Internet Resources: Identity, Affiliation, and Cognitive Authority in a Networked World, in: *Journal of the American Society for Information Science and Technology* 52 (6), 2001, S. 499-507. DOI: 10.1002/asi.1081.
- FRONTLINE: Interview: Sherry Turkle, PBS, FRONTLINE, Digital Nation, 02.02.2010, URL: <http://www.pbs.org/wgbh/pages/frontline/digitalnation/interviews/turkle.html>, Stand: 06.10.2015.
- Fuchs, Konrad; Raab, Heribert: Quellenkritik, in: *Wörterbuch zur Geschichte*, München 2002.
- Fuchs, Konrad; Raab, Heribert: Zeitalter, Weltalter, Äon, in: *Wörterbuch zur Geschichte*, München 2002.
- Fuchs, Konrad; Raab, Heribert: *Wörterbuch zur Geschichte*, München 2002.
- Furlough, Mike: What We Talk About When We Talk About Repositories, in: *Reference & User Services Quarterly* 49 (1), 01.10.2009, S. 18-32.
- Furness, Hannah: Hay Festival 2013: Teenagers' mistakes will stay with them forever, warns Google chief Eric Schmidt, in: *The Telegraph*, 25.05.2013. URL: <http://www.telegraph.co.uk/technology/eric-schmidt/10080596/Hay-Festival-2013-Teenagers-mistakes-will-stay-with-them-forever-warns-Google-chief-Eric-Schmidt.html>, Stand: 01.08.2014.
- Galgano, Michael J.; Arndt, J. Chris; Hyser, Raymond M.: *Doing history research and writing in the digital age*, Boston, Mass 2008.
- Galton, Francis: Vox Populi, in: *Nature* (75), 07.03.1907, S. 450-451. DOI: 10.1038/075450a0.
- Gänser, Georg: Archive, Nachvollziehbarkeit und Vertrauen, in: *scrinium* 71-2017, 2017, S. 102-113.
- Gantert, Klaus: *Elektronische Informationsressourcen für Historiker*, Berlin 2011 (Bibliotheks- und Informationspraxis 43).
- Gasteiner, Martin; Haber, Peter (Hg.): *Digitale Arbeitstechniken: für die Geistes- und Kulturwissenschaften*, Wien, Köln, Weimar 2010.
- Geiselberger, Heinrich; Moorstedt, Tobias (Hg.): *Big Data: das neue Versprechen der Allwissenheit*, Berlin 2013 (edition unseld).
- Geschonneck, Alexander: *Computer-Forensik: Computerstraftaten erkennen, ermitteln, aufklären*, Heidelberg 2011.
- Giaretta, David: *Advanced Digital Preservation*, Berlin 2011.
- Gibbs, Fred; Owens, Trevor: The Hermeneutics of Data and Historical Writing, in: Dougherty, Jack; Nawrotzki, Kristen (Hg.): *Writing history in the digital age*, Ann Arbor 2013 (Digital humanities), S. 159-170. URL: <https://muse.jhu.edu/book/27633>, Stand: 05.02.2018.
- Gibbs, Frederick W.; Cohen, Daniel J.: A Conversation with Data: Prospecting Victorian Words and Ideas, in: *Victorian Studies* 54 (1), 01.10.2011, S. 69-77. DOI: 10.2979/victorianstudies.54.1.69.
- Giessmann, Sebastian; Burkhardt, Marcus: Was ist Datenkritik? Zur Einführung. URL: <http://www.medialekontrolle.de/wp-content/uploads/2014/09/Giessmann-Sebastian-Burkhardt-Marcus-2014-03-01.pdf>, Stand: 28.04.2016.
- Gigerenzer, Gerd: Technik braucht Menschen, die sie beherrschen, in: *Das Digital-Manifest*, 2015, S. 30-32.
- Gladney, Henry M.: *Preserving Digital Information*, Berlin 2007.
- Goertz, Hans-Jürgen: *Geschichte. Ein Grundkurs*, Reinbek bei Hamburg 1998.
- Gold, Matthew K. (Hg.): *Debates in the digital humanities*, Minneapolis 2012. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- Gold, Matthew K.: Day of DH: Defining the Digital Humanities – Blog Posts, in: Gold, Matthew K. (Hg.): *Debates in the digital humanities*, Minneapolis 2012, S. 67-71. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- Gomez, Jeff: *Beside Myself: An Interactive novel for the iPad*, *Beside Myself*, 2016, URL: <http://www.besidemysself.com/>, Stand: 24.07.2016.

- Gomez, Jeff: Die erzählerische Singularität: Geschichten erzählen im digitalen Zeitalter, Aus Politik und Zeitgeschichte, 02.10.2012, URL: <http://www.bpb.de/apuz/145376/geschichten-erzaehlen-im-digitalen-zeitalter?p=all>, Stand: 11.02.2014.
- Google: Algorithmen – Alles über die Suche, Google, URL: <http://www.google.com/intl/de/insidesearch/howsearchworks/algorithms.html>, Stand: 30.12.2013.
- Google, Unternehmen, URL: <http://www.google.com/intl/de/about/company/>, Stand: 24.12.2013.
- Götzer, Klaus (Hg.): Dokumenten-Management Informationen im Unternehmen effizient nutzen, Heidelberg 2014.
- Goubet, Fabien: Algorithmen treffen auf Papyrus, in: Horizonte (105), 2015, S. 10-16.
- Grabs, Anne; Bannour, Karim-Patrick: Follow me!: Social Media Marketing mit Facebook, Twitter, XING, YouTube und Co. Inkl. Empfehlungsmarketing, Crowdsourcing und Social Commerce, 2011.
- Graefe, Andreas; Haim, Mario; Haarmann, Bastian u. a.: Readers' perception of computer-generated news: Credibility, expertise, and readability, in: Journalism, 17.04.2016. DOI: 10.1177/1464884916641269.
- Grass, Robert N.; Heckel, Reinhard; Puddu, Michela u. a.: Robust Chemical Preservation of Digital Information on DNA in Silica with Error-Correcting Codes, in: Angewandte Chemie International Edition 54 (8), 16.02.2015, S. 2552-2555. DOI: 10.1002/anie.201411378.
- Grassegger, Hannes: Ein zerteiltes Internet wäre nutzlos – Interview mit Vinton Cerf, in: NZZ am Sonntag, Zürich 09.02.2014, S. 26.
- Grassegger, Hannes: Das Ende des Internets, in: NZZ am Sonntag, Zürich 09.02.2014, S. 24-26.
- Graves, Michael W.: Digital archaeology: the art and science of digital forensics, London 2014.
- Green, Muriel: Harry Stanley Green's letters and medals, URL: https://www.europeana.eu/portal/en/record/2020601/contributions_17034.html#&gid=1&pid=21 Stand: 06.05.2018
- Grey, Anthony J. G.: Jim Gray on eScience: A Transformed Scientific Method, in: Hey, Anthony J. G.; Tansley, Stewart; Tolle, Christin (Hg.): The fourth paradigm data-intensive scientific discovery, Redmond, Wash 2009, S. XVII–XXXI. URL: http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf.
- Groeben, Norbert; Hurrelmann, Bettina (Hg.): Medienkompetenz: Voraussetzungen, Dimensionen, Funktionen, Weinheim 2002 (Lesesozialisation und Medien).
- Groebner, Valentin: Wissenschaftssprache digital. Die Zukunft von gestern, Paderborn 2014.
- Groebner, Valentin: Welches Thema? Was für eine Art Text? – Vorschläge zum wissenschaftlichen Schreiben 2009 ff., in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 15-23.
- Gunter, Barrie; Rowlands, Ian; Nicholas, David: The Google Generation. Are ICT innovations changing information-seeking behaviour?, Oxford 2009 (Chandos information professional series).
- Haber, Peter: Zeitgeschichte und Digital Humanities, Version: 1.0, Archive, Docupedia, 24.09.2012, URL: http://docupedia.de/zg/Digital_Humanities?oldid=84594, Stand: 17.09.2014.
- Haber, Peter: Wir werden uns die Technik noch stärker zu nutze machen, Audiointerview, 00:01:46, 12.09.2011. URL: http://www.lisa.gerda-henkel-stiftung.de/historiker_muessen_keine_informatiker_sein?nav_id=1807, Stand: 06.12.2015.
- Haber, Peter: Digital Past: Geschichtswissenschaft im digitalen Zeitalter, München 2011.
- Haber, Peter: Geschichte schreiben im Digitalen Zeitalter, in: Digitale Wissenschaft: Stand und Entwicklung digital vernetzter Forschung in Deutschland ; 20./21. September 2010, Köln ; Beiträge der Tagung, Köln 2011, S. 21-26.
- Haber, Peter: Sprung in eine andere Welt? Mediengeschichte im Zeichen von Digitalität und Remediation, in: Schweizerische Zeitschrift für Geschichte 60 (1), 2010, S. 121-132. DOI: <http://dx.doi.org/10.5169/seals-109698>.
- Haber, Peter: digital past – Geschichtswissenschaft im digitalen Zeitalter, unveröffentlichte Habilitationsschrift, Universität Basel, Basel 2009.

- Haber, Peter: «Google-Syndrom». Phantasmagorien des historischen Allwissens im World Wide Web, in: Haber, Peter; Epple, Angelika (Hg.): Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0, Zürich 2004 (Geschichte und Informatik 15), S. 73-89.
- Haber, Peter; Gasteiner, Martin: Digitale Arbeitstechniken. Eine Einführung, in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 11-14.
- Haber, Peter; Pfanzer, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013.
- Haber, Peter; Pfanzer, Eva; Schreiner, Julia: Vorwort, in: Haber, Peter; Pfanzer, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013, S. 51-59.
- Hafner, Urs: Digital History: Der Irrtum der Zeitmaschinen, in: Neue Zürcher Zeitung, Zürich 27.05.2016. URL: <http://www.nzz.ch/feuilleton/zeitgeschehen/digital-history-historiografie-des-zeitpeils-ld.85000>, Stand: 27.05.2016.
- Hagner, Michael: Vor und nach den Digital Humanities. Eine Übersicht, SAGW-Tagung 13, Bern 28.11.2013.
- Hall, Gary: There Are No Digital Humanities, in: Gold, Matthew K. (Hg.): Debates in the digital humanities, Minneapolis 2012, S. 133-138. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- Hammond, Kris: The End of Big Data: AI and the Rise of the Narrative, Blog, NarrativeScience, 06.03.2015, URL: <http://www.narrativescience.com/blog/big-data-ai-rise-narrative/>, Stand: 23.05.2015.
- Harnad, Stevan: The post-gutenberg open access journal, in: Cope, Bill; Phillips, Angus (Hg.): The future of the academic journal, Oxford 2014 (Chandos information professional series), S. 179-193.
- Haugeland, John: analog und analog, in: Böhnke, Alexander; Schröter, Jens (Hg.): Analog/Digital – Opposition oder Kontinuum? Zur Theorie und Geschichte einer Unterscheidung, Bielefeld 2004, S. 33-48.
- Hayles, Katherine N.: Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis, in: Poetics Today 25 (1), 2004, S. 67-90. DOI: 10.1215/03335372-25-1-67.
- Haynes, David: Metadata for information management and retrieval, London 2004.
- Hecker-Stampehl, Jan: Bloggen in der Geschichtswissenschaft als Form des Wissenstransfers, in: Haber, Peter; Pfanzer, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013, S. 37-49.
- Helbing, Dirk: »BIG NUDGING« – zur Problemlösung wenig geeignet, in: Das Digital-Manifest, 2015, S. 37-39.
- Helbing, Dirk; Frey, Bruno; Zicari, Roberto u. a.: Digitale Demokratie statt Datendiktatur, in: Das Digital-Manifest, 2015, S. 5-19.
- Hermann, Felix: Rezension zu: Gold, Matthew K. (Hrsg.): Debates in the Digital Humanities, Minneapolis 2012 / Berry, David M. (Hrsg.): Understanding Digital Humanities, New York 2012, Buchrezensionen, H-Soz-Kult, 09.01.2015, URL: <http://www.hsozkult.de/publicationreview/id/rezbuecher-21673>, Stand: 09.01.2015.
- Herold, Helmut; Lurz, Bruno; Wohlrab, Jürgen: Grundlagen der Informatik: praktisch – technisch – theoretisch, München, Boston 2007.
- Hess, Demian: DAM and the Need for Flexible Metadata Models, Digital Asset Management, Avalon Consulting, LLC – Blogs, 17.11.2014, URL: <http://blogs.avalonconsult.com/blog/generic/dam-and-the-need-for-flexible-metadata-models/>, Stand: 21.08.2015.
- Hirschi, Caspar: Das geisteswissenschaftliche Buch im digitalen Zeitalter. Der Schweizerische Nationalfonds und seine Open-Access-Strategie., in: NZZ-Online, Zürich 19.05.2014. URL: <http://www.nzz.ch/feuilleton/der-schweizerische-nationalfonds-und-seine-open-access-strategie-1.18304812>.
- Hochschulrektorenkonferenz Deutschland: Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien, 10.11.2015. URL: <https://www.hrk.de/positionen/beschluss/detail/wie->

- hochschulleitungen-die-entwicklung-des-forschungsdatenmanagements-steuern-koennen-orientierungsp/, Stand: 06.05.2018.
- Hodel, Jan: Verkürzen und Verknüpfen: Geschichte als Netz narrativer Fragmente. Wie Jugendliche digitale Netzmedien für die Erstellung von Referaten im Geschichtsunterricht verwenden, Bern 2013 (Geschichtsdidaktik heute).
- Hodel, Jan: Recherche: Google – and Far Beyond, in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 25-37.
- Hodel, Jan: Digital lesen, digital schreiben, digital denken? Über den kompetenten Umgang mit Geschichte im Zeitalter des digitalen Medienwandels, in: Eggs, Cindy; Jorio, Marco (Hg.): Wort. Am Anfang ist das: Lexika in der Schweiz, Baden 2008, S. 113-125.
- Hodel, Jan: Historische Online-Kompetenz. Überlegungen zu einem hybriden Kompetenzmodell, in: Haber, Peter; Eppe, Angelika (Hg.): Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0, Zürich 2004 (Geschichte und Informatik 15), S. 139-161.
- Holdren, John P.: Increasing Access to the Results of Federally Funded Scientific Research. Memorandum For The Heads Of Executive Departments And Agencies, Executive Office of the President. Office of Science and Technology Policy, 22.02.2013.
- Holiday, Ryan: Trust me, I'm lying. Confessions of a media manipulator, New York 2013.
- Howell, Martha; Prevenier, Walter: Werkstatt des Historikers: Eine Einführung in die historischen Methoden, Köln 2004.
- Hügi, Jasmin; Schneider, René: Digitale Forschungsinfrastrukturen in den Geistes- und Sozialwissenschaften, Haute école de gestion den Genève, Genf 24.01.2013.
- Hühnlein, Detlef; Korte, Ulrike: Grundlagen der elektronischen Signatur: Recht, Technik, Anwendung, Ingelheim 2006.
- Hui, Yuk: The computational turn, or, a new Weltbild, in: Junctures – The Journal for Thematic Dialogue (13), 2010, S. 41-51.
- Hunger, Francis: SETUN eine Recherche über den sowjetischen Ternärcomputer, Leipzig 2007.
- Hyun-ju, Ock: 200 professors to be indicted over brazen copyright breach, Social Affairs, The Korea Herald, 24.11.2015, URL: <http://www.koreaherald.com/view.php?ud=20151124001042>, Stand: 26.06.2016.
- Ilves, Toomas Hendrik: The greatest cyber threat comes from violation of data integrity, Steering Board of the European Cloud Partnership, Berlin 14.11.2013. URL: <http://www.president.ee/en/media/press-releases/9604-president-ilves-in-berlin-the-greatest-cyber-threat-comes-from-violation-of-data-integrity/index.html>, Stand: 21.07.2014.
- Immenhauser, Beat: SAGW goes digital, in: Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Digital Humanities und Web 2.0, 2012, S. 29-31.
- Information Commissioner's Office: Data Protection Act 1998 – Superprovisory Powers of the Information Commissioner – Enforcement Notice, 18.08.2015. URL: <https://ico.org.uk/media/action-weve-taken/enforcement-notice/1432380/google-inc-enforcement-notice-18082015.pdf>, Stand: 21.08.2015.
- Innis, Harold A: Bias of Communication, Toronto 1951.
- Innis, Harold A: Empire and Communications, London 1950.
- Institut für Geschichtswissenschaft der Universität Bremen: Studienleitfaden Geschichtswissenschaft. Einführung in die Techniken und Methoden des wissenschaftlichen Arbeitens, 2012.
- Jacso, Peter: Google Scholar's Ghost Authors, Library Journal, URL: <http://lj.libraryjournal.com/2009/11/industry-news/google-scholars-ghost-authors/>, Stand: 30.12.2013.
- Jaeger, Friedrich: Geschichtstheorie, in: Goertz, Hans-Jürgen (Hg.): Geschichte. Ein Grundkurs, Reinbek bei Hamburg 1998, S. 724-756.
- Jenkins, Henry: «Does this Technology serve Human Purposes?» A «Necessary Conversation» with Sherry Turkle, in: Svensson, Patrik; Goldberg, David Theo (Hg.): Between humanities and the digital, Cambridge 2015, S. 69-81.

- Jenkins, Henry: *Convergence Culture: Where Old and New Media Collide*, New York and London 2006.
- Jenks, Stuart; Marra, Stephanie: Über die Verlässlichkeit von Informationen im Internet, in: Jenks, Stuart; Marra, Stephanie (Hg.): *Internet-Handbuch Geschichte*, Köln 2001, S. 265-271.
- Jobin, Anna: *La mémoire kaléidoscopique : l'histoire au prisme des algorithmes d'autocomplétion*, Panel: *Le pouvoir des algorithmes. Technologies numériques et rapports de force au XXe-XXIe siècles*, Lausanne 09.06.2016. URL: <https://geschichtstage.ch/referat/77/la-memoire-kaleidoscopique-l-histoire-au-prisme-des-algorithmes-d-autocompletion>, Stand: 26.06.2016.
- Jonas, Jeff; Sokol, Lisa: *Data Finds Data*, in: Segaran, Toby; Hammerbacher, Jeff (Hg.): *Beautiful data: [the stories behind elegant data solutions]*, Sebastopol, Calif 2009 (Data modeling & design), S. 105-118.
- Jordan, Stefan: *Einführung in das Geschichtsstudium*, Stuttgart 2013 (Reclams Universal-Bibliothek).
- Jordan, Stefan: *Theorien und Methoden der Geschichtswissenschaft. Orientierung Geschichte*, Paderborn 2008.
- Jordan, Stefan: *Lexikon Geschichtswissenschaft: Hundert Grundbegriffe*, Stuttgart 2007.
- Jung, Matthias: *Hermeneutik zur Einführung*, Hamburg 2012⁴.
- Kannenbergh, Axel: *Bürgerrechtler: Britischer Pornofilter blockt fast jede fünfte Website*, Newsticker, heise online, 03.07.2014, URL: <http://www.heise.de/newsticker/meldung/Buergerrechtler-Britischer-Pornofilter-blockt-fast-jede-fuenfte-Website-2248673.html>, Stand: 08.05.2015.
- Karpathy, Andrej; Fei-Fei, Li: *Deep Visual-Semantic Alignments for Generating Image Descriptions*, Stanford University, Stanford 11.2014, S. 10. URL: <http://cs.stanford.edu/people/karpathy/deepimagesent/devisagen.pdf>, Stand: 19.11.2014.
- Keller-Marxer, Peter: *Archivierung elektronischer digitaler Daten und Akten (ARELDA) im Bundesarchiv, VSA/AAS-Arbeitstagung «Elektronische Akten – eine Strategie. Und dann?»*, Bern 22.11.2002.
- Kelly, Matthew; Bielby, Jared (Hg.): *Information Cultures in the Digital Age: A Festschrift in Honor of Rafael Capurro*, Wiesbaden 2016.
- Kelly, T. Mills: *Teaching history in the digital age*, Ann Arbor 2013 (Digital humanities).
- Kemman, Max; Kleppe, Martijn; Scagliola, Stef: *Just Google It – Digital Research Practices of Humanities Scholars*, in: Mills, Clare; Pidd, Michael; Ward, Esther (Hg.): *Proceedings of the Digital Humanities Congress 2012. Studies in the Digital Humanities – HRI Online Publications*, arXiv:1309.2434 [cs.DL], Sheffield 2014. URL: <http://arxiv.org/abs/1309.2434>, Stand: 06.08.2015.
- Kesselring, Thomas: *Mens sana in corpore sano. Integrität aus ethischer Sicht*, in: *Tagung «Wieviel Körper darf es sein?»*, Bern 2007. URL: <http://docplayer.org/32570566-Thomas-kesselring-mens-sana-in-corpore-sano-integrtaet-aus-ethischer-sicht.html>, Stand: 06.05.2018.
- King, Samuel T.; Tucek, Joseph; Cozzie, Anthony u. a.: *Designing and implementing malicious hardware*, in: *Firt Usenix Workshop on Large-Scale Exploits and emergent Threats*, San Francisco, CA, USA 2008. URL: https://www.usenix.org/legacy/events/leet08/tech/full_papers/king/king.pdf, Stand: 17.09.2015.
- Kirn, Paul: *Einführung in die Geschichtswissenschaft*, Berlin [-West] 1968 (Sammlung Göschen).
- Kittler, Friedrich A.: *Grammophon, Film, Typewriter*, Berlin 1986.
- Klein, Torsten; Windeck, Christof: *WikiCon: Konfliktfeld Wikipedia*, Newsticker, heise online, 05.10.2014, URL: <http://www.heise.de/newsticker/meldung/WikiCon-Konfliktfeld-Wikipedia-2411752.html>, Stand: 27.05.2015.
- Klinke, Harald; Surkemper, Liska (Hg.): *What is Digital Art History?*, München 2015 (International Journal for Digital Art History). URL: <http://www.dah-journal.org/>, Stand: 27.02.2016.
- Kohle, Hubertus: *Digitale Bildwissenschaft*, Glückstadt 2013 (E-Humanities).
- Kohler, Josef: *Das Autorenrecht – Eine zivilistische Abhandlung zugleich ein Beitrag zur Lehre vom Eigentum, vom Miteigentum, vom Rechtsgeschäft und vom Individualrecht*, Jena 1880.
- Koller, Guido: *Geschichte digital: Historische Welten neu vermessen*, Stuttgart 2016.

- König, Mareike: Jenseits der Metaphorik: Experimente in den Digital Humanities #dhiha6, Billet, digitale:geschichte, 07.06.2015, URL: <https://dguw.hypotheses.org/257>, Stand: 19.02.2018.
- König, Mareike: Blogs als Wissensorte der Forschung, Präsentation, Die Zukunft der Wissenspeicher: Forschen, Sammeln und Vermitteln im 21. Jahrhundert, Konstanz 06.03.2015. URL: <http://cms.uni-konstanz.de/wissenschaftsforum/veranstaltungen/veranstaltungsarchiv/veranstaltungen-2015/die-zukunft-der-wissenspeicher/>, Stand: 17.06.2015.
- König, Mareike: Herausforderung für unsere Wissenschaftskultur: Weblogs in den Geisteswissenschaften, in: Schmale, Wolfgang (Hg.): Digital Humanities: Praktiken der Digitalisierung, der Dissemination und der Selbstreflexivität, Stuttgart 2015 (Geschichte), S. 57-74.
- König, Mareike: Die Entdeckung der Vielfalt: Geschichtsblogs der europäischen Plattform hypotheses.org, in: Haber, Peter; Pfanzelter, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013, S. 181-197.
- König, Mathias; König, Wolfgang: Twitter kontrolliert Auffindbarkeit von Tweets, Praxis, politik & kommunikation, 10.03.2016, URL: <http://www.politik-kommunikation.de/ressorts/artikel/twitter-kontrolliert-auffindbarkeit-von-tweets-1486134101>, Stand: 26.06.2016.
- König, René; Nentwich, Michael: Cyberscience 2.0: Wissenschaftskommunikation in der Beta-Gesellschaft, in: kommunikation @ gesellschaft 15 (Sonderausgabe), 2014, S. 26.
- Koosel, Stacey M.: Exploring Digital Identity: Beyond the Private Public Paradox, in: Runnel, Pille; u.a. (Hg.): The digital turn – user's practices and cultural transformations, Frankfurt a.M. 2013, S. 161-173.
- Koskimaa, Raine: From the Gutenberg Galaxy to the Internet Galaxy. Digital Textuality and the Change of the Cultural Landscape, in: Runnel, Pille; u.a. (Hg.): The digital turn – user's practices and cultural transformations, Frankfurt a.M. 2013, S. 271-283.
- KOST: Katalog archivischer Dateiformate, Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen, 06.2015, URL: <http://kost-ceco.ch/wiki/whelp/KaD/index.php>, Stand: 14.04.2016.
- Krameritsch, Jakob: Hypertext schreiben, in: Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, 08.04.2010, S. 83-95.
- Krameritsch, Jakob: Geschichte(n) im Netzwerk. Hypertext und dessen Potenziale für die Produktion, Repräsentation und Rezeption der historischen Erzählung, Münster 2007 (Medien in der Wissenschaft).
- Krawetz, Neal: A picture's worth: Digital image analysis, in: Black Hat DC, Washington (USA) 2008. URL: <https://www.blackhat.com/html/bh-dc-08/bh-dc-08-archives.html>, Stand: 17.06.2015.
- Kretzschmar, Robert: Archive als digitale Informationsinfrastrukturen. Stand und Perspektiven, in: Archivar, Zeitschrift für Archivwesen 2, 2013, S. 146-153.
- Krikorian, Gaëlle; Kapczynski, Amy (Hg.): Access To Knowledge In The Age Of Intellectual Property, New York 2010.
- Krüger, Stefanie: Die Erschließung digitaler und analoger Suchräume. Anforderungen an heuristische Verfahren, in: Haber, Peter; Epple, Angelika (Hg.): Vom Nutzen und Nachteil des Internets für die historische Erkenntnis: Version 1.0, Zürich 2004 (Geschichte und Informatik 15), S. 91-105.
- Krüger, Uwe: Meinungsmacht: der Einfluss von Eliten auf Leitmedien und Alpha-Journalisten – eine kritische Netzwerkanalyse, Köln 2013 (Reihe des Instituts für Praktische Journalismus- und Kommunikationsforschung 9).
- Kühl, Eike: Künstliche Intelligenz: Auf Fake News folgt Fake Porn, Zeit Online, 26.01.2018, URL: <http://www.zeit.de/digital/internet/2018-01/kuenstliche-intelligenz-deepfakes-porno-face-swap/komplettansicht?print>, Stand: 07.02.2018.
- Kuhlee, Lorenz; Völzow, Victor: Computer-Forensik Hacks, Köln 2012 (Hacks series).
- Kuhlen, Rainer: Hypertext: Ein nicht-lineares Medium zwischen Buch und Wissensbank, Berlin 1991 (Edition SEL-Stiftung).
- Kühmstedt, Estella: Klug recherchiert: für Historiker, Göttingen 2013 (UTB 3940).

- Kuhn, Thomas S.: The structure of scientific revolutions, Bd. 2, Chicago 1962 (International encyclopedia of unified science 2).
- Kunder, Maurice de: Geschatte grootte van het geïndexeerde World Wide Web, 29.03.2007. URL: <http://www.dekunder.nl/Media/Scriptie%20Maurice%20de%20Kunder%20-%20Grootte%20geïndexeerde%20web.pdf>, Stand: 12.03.2015.
- Kurzidim, Michael: Crowd-Testing: pro Bug 25 Franken, Businesspraxis, Computerworld.ch, 01.12.2014, URL: <http://www.computerworld.ch/businesspraxis/artikel/weg-mit-den-software-bugs-66855/>, Stand: 02.12.2014.
- Kwon, Diana: Major German Universities Cancel Elsevier Contracts, News, The Scientist, 17.07.2017, URL: <https://www.the-scientist.com/?articles.view/articleNo/49906/title/Major-German-Universities-Cancel-Elsevier-Contracts/>, Stand: 07.02.2018.
- Landwehr, Achim: Die Kunst, sich nicht allzu sicher zu sein: Möglichkeiten kritischer Geschichtsschreibung, in: geschichte und kritik (61), 2013, S. 7-15.
- Landwehr, Achim: Geschichte des Sagbaren. Einführung in die historische Diskursanalyse, Tübingen 2001 (Historische Einführungen).
- Langer, Marie-Astrid: Zu Besuch in der Zukunft: Fortschrittliches Estland, Europa, Neue Zürcher Zeitung, 27.02.2015, URL: http://www.nzz.ch/international/europa/zu-besuch-in-der-zukunft-1.18491451?extcid=Newsletter_27022015_Top-News_am_Morgen, Stand: 27.02.2015.
- Lanier, Jaron: Wem gehört die Zukunft? Du bist nicht der Kunde der Internet-Konzerne, du bist ihr Produkt, Hamburg 2014.
- Lauer, Gerhard: Die digitale Vermessung der Kultur. Geisteswissenschaften als Digital Humanities, in: Geiselberger, Heinrich; Moorstedt, Tobias (Hg.): Big Data: das neue Versprechen der Allwissenheit, Berlin 2013 (edition unseld), S. 99-116.
- Lawrence, Steve; Pennock, David M.; Flake, David William u. a.: Persistence of Web References in Scientific Research, in: Computer (34), 02.2001, S. 26-31.
- Lecher, Hanno E.: Small Scale Academic Web Archiving: DACHS, in: Web Archiving, Berlin 2006, S. 213-225.
- Lengwiler, Martin: Praxisbuch Geschichte: Einführung in die historischen Methoden, Zürich 2011.
- Leon, Sam: The Revenge of the Yellow Milkmaid: Cultural Heritage Institutions open up dataset of 20m+ items, Open Knowledge Foundation, 17.09.2012, URL: <http://blog.okfn.org/2012/09/17/the-revenge-of-the-yellow-milkmaid-cultural-heritage-institutions-open-up-dataset-of-20m-items/>, Stand: 14.08.2014.
- Leonelli, Sabina: Data Interpretation in the Digital Age, in: Perspectives on Science 22 (3), 06.03.2014, S. 397-417. DOI: 10.1162/POSC_a_00140.
- Lepore, Jill: The Cobweb. Can the Internet be archived?, The New Yorker, 19.01.2015, URL: <http://www.newyorker.com/magazine/2015/01/26/cobweb>, Stand: 28.01.2015.
- Levy, Steven: Can an Algorithm Write a Better News Story Than a Human Reporter?, Blog, WIRED, 24.04.2012, URL: <http://www.wired.com/2012/04/can-an-algorithm-write-a-better-news-story-than-a-human-reporter/>, Stand: 23.05.2015.
- Lewandowski, Dirk; Mayr, Philipp: Exploring the Academic Invisible Web, in: Library Hi Tech 24 (4), 2006, S. 529-539.
- Lindner, Jim: Toward a MediaLESS Archive, in: Aubert, Michelle; Billeaud, Richard (Hg.): Archiver et communiquer l'image et le son: les enjeux du 3ème millénaire: actes du Symposium technique mixte – JTS Paris 2000, Paris 2000.
- Lindstädt, Birte; Schmidt, Christoph: «Digital Preservation Summit 2011», Globales Gipfeltreffen zur digitalen Langzeitarchivierung; Tagungsbericht, in: Archivar, Zeitschrift für Archivwesen 2, 2012, S. 175-178.
- Lingelbach, Gabriele: Ein Motor der Geschichtswissenschaft? Zusammenhänge zwischen technologischer Entwicklung, Veränderungen des Arbeitsalltags von Historikern und fachlichem Wandel, zeitenblicke, 09.08.2011, URL: http://www.zeitenblicke.de/2011/1/Lingelbach/index_html, Stand: 30.07.2015.

- Lischka, Konrad: Studie zur Web-Haltbarkeit: Das Netz vergisst schnell -, SPIEGEL ONLINE – Nachrichten – Netzwelt, URL: <http://www.spiegel.de/netzwelt/web/web-archiv-studie-zur-haltbarkeit-von-online-quellen-a-856936-druck.html>, Stand: 25.09.2012.
- Liu, Peter J.; Saleh, Mohammad; Pot, Etienne u. a.: Generating Wikipedia by Summarizing Long Sequences, arXiv:1801.10198 [cs.CL], 30.01.2018. URL: <http://arxiv.org/abs/1801.10198>, Stand: 19.02.2018.
- Lorenz, Chris: Heuristik, in: Jordan, Stefan (Hg.): Lexikon Geschichtswissenschaft: Hundert Grundbegriffe, Stuttgart 2007, S. 139-141.
- Lorenz, Chris: Konstruktion der Vergangenheit, Köln 1997.
- Lorenz, Maren: Wikipedia – Zum Verhältnis von Struktur und Wirkungsmacht eines heimlichen Leitmediums, Essen 2006 (Werkstatt Geschichte 43), S. 84-95.
- Lovink, Geert: Zero comments Elemente einer kritischen Internetkultur, Bielefeld 2008 (Kultur- und Medientheorie).
- Luhmann, Niklas: Die Gesellschaft der Gesellschaft, Frankfurt a.M. 1998 (Suhrkamp Taschenbuch Wissenschaft).
- Lunt, Barry M.; Davis, Robert; Hansen, Douglas u. a.: Permanent Digital Data Storage: A Materials Approach, in: iPRES-2013 – 10th International Conference on Preservation of Digital Objects, Lissabon 2013. URL: http://www.purl.pt/24107/1/iPres2013_PDF/Permanent%20Digital%20Data%20Storage%20A%20Materials%20Approach.pdf, Stand: 05.12.2015.
- Lurk, Tabea (Hg.): ComputerKunstGeschichte Schweiz, in: Computergeschichte Schweiz: eine Bestandesaufnahme = Histoire de l'ordinateur en Suisse: un état des lieux, Zürich 2009 (Geschichte und Informatik 17).
- Lynch, Clifford A.: Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust, in: Cullen et al. (Hg.): Authenticity in a Digital Environment, Washington, D.C. 2000, S. 32-50.
- Lyon, David: Surveillance after Snowden, Cambridge 2015.
- Mahrt, Merja; Scharow, Michael: Der Wert von Big Data für die Erforschung digitaler Medien, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 221-237.
- Makarow, Marja; Moulin, Claudine; Žic Fuchs, Milena: Research Infrastructures in the Digital Humanities, Strasbourg 2011 (Science Policy Briefing 42). URL: http://resaw.eu/wp-content/uploads/2013/06/ESF_Research-Infrastructures-in-the-Digital-Humanities.pdf Stand: 06.05.2018.
- Makino, Takaki; Jung, Chaesang; Phan, Doantam: Finding more mobile-friendly search results, Official news on crawling and indexing sites for the Google index, Official Google Webmaster Central Blog, 26.02.2015, URL: <http://googlewebmastercentral.blogspot.com/2015/02/finding-more-mobile-friendly-search.html>, Stand: 17.06.2015.
- Makino, Takaki; Phan, Doantam: Rolling out the mobile-friendly update, Official Google Webmaster Central Blog, 21.04.2015, URL: <http://googlewebmastercentral.blogspot.com/2015/04/rolling-out-mobile-friendly-update.html>, Stand: 13.10.2015.
- Maltzahn, Christoph Frhr von: Niebuhr, Barthold Georg, in: Vom Bruch, Rüdiger; Müller, Rainer A. (Hg.): Historikerlexikon: von der Antike bis zur Gegenwart, München 2002, S. 236.
- Maltzahn, Christoph Frhr von: Droysen, Johann Gustav, in: Vom Bruch, Rüdiger; Müller, Rainer A (Hg.): Historikerlexikon: von der Antike bis zur Gegenwart, München 2002, S. 267.
- Manovich, Lev: The language of new media, Cambridge (USA) 2001 (Leonardo). URL: http://dss-edit.com/plu/Manovich-Lev_The_Language_of_the_New_Media.pdf, Stand: 29.04.2016.
- Manovich, Lev: What is Visualization?, Data Visualisation, 09.11.2010, URL: http://manovich.net/content/04-projects/064-what-is-visualization/61_article_2010.pdf, Stand: 06.05.2018.
- Margrethe II., Königin von Dänemark: Lov om pligtaflevering af offentliggjort materiale (Lov nr. 1439 af 22. december 2004), Pligtaflevering, 22.12.2004, URL: <http://www.pligtaflevering.dk/loven/index.htm>, Stand: 31.03.2015.

- Margulies, Simon B.: Digitale Daten als Quelle der Geschichtswissenschaft: Eine Einführung, Hamburg 2009 (Kölner Beiträge zu einer geisteswissenschaftlichen Fachinformatik 2).
- Masanès, Julien: Web Archiving, Berlin 2006.
- McCarty, Willard: A Telescope for the Mind?, in: Gold, Matthew K. (Hg.): Debates in the digital humanities, Minneapolis 2012, S. 113-123. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- McGuinness, Phillipa (Hg.): Copyfight, Sydney 2015.
- McLuhan, Marshall (Hg.): Understanding Media: The Extensions of Man, New York 1964.
- McLuhan, Marshall: The Gutenberg Galaxy: The Making of Typographic Man, London 1962.
- McLuhan, Marshall; Fiore, Quentin; Agel, Jerome: The medium is the message: An inventory of effects, New York 1967 (Bantam Books).
- McNamee, Joe: Open Letter to the President of the European Commission Jean-Claude Juncker concerning Data Protection, E-Mail, 21.04.2015. URL: https://edri.org/files/DP_letter_Juncker_20150421.pdf, Stand: 01.05.2015.
- Mersch, Dieter: Medientheorien zur Einführung, Hamburg 2013 (Zur Einführung 318, Ed. 3).
- Merten, Klaus: Zur Ausdifferenzierung der Mediengesellschaft. Wirklichkeitsmanagement als Suche nach Wahrheit, in: Arnold, Klaus; Tonnemacher, Jan (Hg.): Alte Medien – neue Medien: Theorienperspektiven, Medienprofile, Einsatzfelder. Festschrift für Jan Tonnemacher, Wiesbaden 2005 (Public relations), S. 21-39.
- Messner, Daniel: Coding History – Software als kulturwissenschaftliches Forschungsobjekt, in: Schmale, Wolfgang (Hg.): Digital Humanities: Praktiken der Digitalisierung, der Dissemination und der Selbstreflexivität, Stuttgart 2015 (Geschichte), S. 157-174.
- Metzger, Franziska: Geschichtsschreibung und Geschichtsdenken im 19. und 20. Jahrhundert, Bern 2011.
- Meyer, Steven: The Effect of Computers on Understanding Truth, in: Hagengruber, Ruth; Ess, Charles (Hg.): The Computational Turn: Past, Presents, Futures? Proceedings IACAP 2011, Münster 2011, S. 49-52. URL: <http://homepages.uc.edu/~freemago/papers/IACAP.pdf>, Stand: 06.05.2018.
- Miller, George A.: The Magical Number Seven, Cogprints, o. J., URL: <http://cogprints.org/730/1/miller.html>, Stand: 23.08.2015.
- Miller, Hannah: Investigating the Potential for Miscommunication Using Emoji, GroupLens, 04.05.2016, URL: <http://grouplens.org/blog/investigating-the-potential-for-miscommunication-using-emoji/>, Stand: 14.04.2016.
- Miller, Hannah; Thebault-Spieker, Jacob; Chang, Shuo u. a.: "Blissfully happy" or "ready to fight": Varying Interpretations of Emoji, in, Köln 2016. URL: http://grouplens.org/site-content/uploads/Emoji_Interpretation.pdf, Stand: 14.04.2016.
- Miller, Stephen J.: Metadata for Digital Collections: A How-to-do-it Manual, London 2011.
- Missomelius, Petra: Digitale Medienkultur Wahrnehmung, Konfiguration, Transformation, Bielefeld 2006 (Kultur- und Medientheorie).
- Moraes, Claude: Report on the US NSA surveillance programme, surveillance bodies in various Member States and their impact on EU citizens' fundamental rights and on transatlantic cooperation in Justice and Home Affairs, Report A7-0139/2014, Committee on Civil Liberties, Justice and Home Affairs, Bruxelles 21.02.2014. URL: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+REPORT+A7-2014-0139+0+DOC+PDF+V0//EN>, Stand: 04.05.2018.
- Mordvintsev, Alexander; Olah, Christopher; Tyka, Mike: Inceptionism: Going Deeper into Neural Networks, Google Research Blog, 17.06.2015, URL: <http://googleresearch.blogspot.com/2015/06/inceptionism-going-deeper-into-neural.html>, Stand: 24.07.2015.
- Moretti, Franco: Distant reading, London 2013.

- Morse, Andrew; Sherr, Ian: Senate Website Gets Hacked, in: Wall Street Journal, 14.06.2011. URL: <http://online.wsj.com/article/SB10001424052702303848104576383970053018848>, Stand: 13.11.2012.
- Mounier, Pierre: Die Werkstatt öffnen: Geschichtsschreibung in Blogs und Sozialen Medien, in: Haber, Peter; Pfanzelter, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013, S. 51-59.
- Mruck, Katja; Gersmann, Gudrun (Hg.): Neue Medien in den Sozial-, Geistes- und Kulturwissenschaften: Elektronische Publizieren und Open Access: Stand und Perspektiven – Sonderheft, in: Historical Social Research 29 (1), 2004.
- Muhlack, Ulrich (Hg.): Leopold von Ranke und die Begründung der quellenkritischen Geschichtsforschung, in: Historische Debatten und Kontroversen im 19. und 20. Jahrhundert: Jubiläumstagung der Ranke-Gesellschaft in Essen 2001, Wiesbaden 2003 (Historische Mitteilungen. Beiheft), S. 23-33.
- Muhlack, Ulrich: Verstehen, in: Goertz, Hans-Jürgen (Hg.): Geschichte. Ein Grundkurs, Reinbek bei Hamburg 1998, S. 99-131.
- Murali, Vijayaraghavan; Qi, Letao; Chaudhuri, Swarat; Jermaine, Chris: Neural Sketch Learning for Conditional Program Generation, arXiv:1703.05698 [cs.PL], 16.03.2017. URL: <http://arxiv.org/abs/1703.05698>, Stand: 30.04.2018.
- Münker, Stefan; Roesler, Alexander (Hg.): Was ist ein Medium?, Frankfurt a.M. 2008 (Suhrkamp Taschenbuch Wissenschaft).
- National Information Standards Organization: Permanence of Paper for Publications and Documents in Libraries and Archives, Standard ANSI/NISO Z39.48-1992 (R2009), Baltimore (USA) 2009. URL: <https://www.niso.org/publications/ansiniso-z3948-1992-r2009-permanence-paper-publications-and-documents-libraries-and>, Stand: 06.05.2018.
- Nationalrat der Republik Österreich: Bundesgesetz über die Presse und andere publizistische Medien (Mediengesetz – MedienG), 19.04.2015. URL: <http://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=1000719>, Stand: 09.09.2015.
- Nelson, Theodor Holm: Literary machines 93.1: the report on, and of, project Xanadu concerning word processing, electronic publishing, hypertext, Sausalito 1993.
- Nentwich, Michael: Cyberscience 2.0 research in the age of digital social networks, Frankfurt a.M. 2012 (Interaktiva).
- Nentwich, Michael: Cyberscience 2.0 oder 1.2? Das Web 2.0 und die Wissenschaft, in: ITA manu:script (ITA-ms-09-02), 11.2009. DOI: 10.1553.
- Nentwich, Michael; König, René: Peer Review 2.0: Herausforderungen und Chancen der wissenschaftlichen Qualitätskontrolle im Zeitalter der Cyber-Wissenschaft, in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 143-163.
- Neuroseptic: Predatory Journals Hit By «Star Wars» Sting, Blog, Neuroseptic, 22.07.2017, URL: <http://blogs.discovermagazine.com/neuroseptic/2017/07/22/predatory-journals-star-wars-sting/>, Stand: 20.02.2018.
- Neuroth, Heike; Strathmann, Stefan; Osswald, Achim u. a. (Hg.): Nestor Handbuch: eine kleine Enzyklopädie der digitalen Langzeitarchivierung, Göttingen 2010. URL: <http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/index.php>.
- New Zealand Foreign Affairs & Trade: Text of the Trans-Pacific Partnership, Trans-Pacific Partnership, 2016, URL: <https://www.tpp.mfat.govt.nz/text>, Stand: 03.04.2016.
- Nicholas, David; Herman, Eti (Hg.): Information Science, London 2014 (Critical concepts in media and cultural studies).
- Nicholas, David; Rowlands, Ian; Clark, David u. a.: Google Generation II: web behaviour experiments with the BBC, in: Aslib Proceedings 63 (1), 18.01.2011, S. 28-45. DOI: 10.1108/00012531111103768.
- Nicholas, David; Watkinson, Anthony; Jamali, Hamid R. u. a.: Peer review: still king in the digital age, in: Learned Publishing 28 (1), 2015, S. 15.21. DOI: 10.1087/20150104.

- Nickl, Roger; Gull, Thomas: Wer herrscht im Internet? Interview mit Hans Geser und Michael Latzer, in: magazin – Universität Zürich (4), 2011, S. 36-39.
- Nicolussi, Ronny: Arbeitsgruppe schlägt Zensur und Überwachung vor: Urheberrecht im digitalen Zeitalter, Schweiz, Neue Zürcher Zeitung, 06.12.2013, URL: <http://www.nzz.ch/aktuell/schweiz/arbeitsgruppe-schlaegt-zensur-und-ueberwachung-vor-1.18199613>, Stand: 04.05.2018.
- Nielsen, Jakob: F-Shaped Pattern For Reading Web Content, Articles, NN/g Nielsen Norman Group, 17.04.2006, URL: <http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/>, Stand: 07.04.2015.
- Nunberg, Geoffrey: Google's Book Search: A Disaster for Scholars, in: The Chronicle of Higher Education, 31.08.2009. URL: <http://chronicle.com/article/Googles-Book-Search-A/48245/>, Stand: 01.10.2015.
- Nygren, Thomas; Vikström, Lotta: A Bastard in the History Classroom: Benefits and Obstacles in Using Digital Databases in History Teaching, dhlu - Panel: Challenges and potentialities of online archives, Luxembourg 20.03.2012.
- O'Donnell, Daniel Paul: The credit line, 16.08.2014, URL: <http://people.uleth.ca/~daniel.odonnell/Blog/firs-thing-we-do-let-s-kill-all-the-authors>, Stand: 19.11.2014.
- O'Reilly, Tim: What is Web 2.0?, 30.09.2005, URL: <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>, Stand: 06.05.2018.
- Ostendorf-Rupp, Sonja: Kurz nachgefragt: Akzeptanz digitaler wissenschaftlicher Publikationen, Dresden Summer School 2012, 02.10.2012, URL: <http://dss.hypotheses.org/455>, Stand: 04.03.2014.
- Pachali, David: Foto-Reproduktionen: Wikimedia verliert Bilderstreit mit Mannheimer Museum (Update), Museen, iRights – Kreativität und Urheberrecht in der digitalen Welt, 23.06.2016, URL: <https://irights.info/artikel/foto-reproduktionen-wikimedia-verliert-bilderstreit-mit-mannheimer-museum/27582>, Stand: 16.07.2016.
- Padayachy, Meshendri: Copyright Amendment Bill 2015 – Draft, 27.07.2015. URL: <http://www.gov.za/documents/copyright-amendment-bill-comments-invited-27-jul-2015-0000>, Stand: 16.08.2015.
- Panzer-Steindel, Bernd: Data integrity, CERN, 08.04.2007. URL: http://indico.cern.ch/event/13797/contributions/1362288/attachments/115080/163419/Data_integrity_v3.pdf, Stand: 19.04.2016.
- Pariser, Eli: Filter Bubble: Wie wir im Internet entmündigt werden, München 2012.
- Pawliczek, Aleksandra: Geschichtsforschung und Geschichtsschreibung zwischen Digital Humanities und Archivquellen, in: geschichte und kritik (61), 2013, S. 37-42.
- Perri, Pascal (Hg.): Google un ami qui ne vous veut pas que du bien, Paris 2013.
- Perry, Dave: Be Online or Be Irrelevant – AcademHack – Thoughts on Emerging Media and Higher Education, 11.01.2010, URL: <http://academhack.outsidethetext.com/home/2010/be-online-or-be-irrelevant/>, Stand: 07.02.2018.
- Petrescu, Philip; Ghita, Mircea; Loiz, Dana: Google Organic CTR Study, Studie, Caphyon / Advanced Web Ranking, 2014, S. 56. URL: <https://www.advancedwebranking.com/google-ctr-study-2014.html>, Stand: 28.03.2016.
- Pfanzelter, Eva: Rezensieren, kommentieren, bloggen: Wissenschaftliche Onlinekommunikation, in: Bernsen, Daniel; Kerber, Ulf (Hg.): Praxishandbuch Historisches Lernen und Medienbildung im digitalen Zeitalter, Opladen Berlin Toronto 2017, S. 158-166.
- Pfanzelter, Eva: Quellenkritik vor der Zerreißprobe? Vom kritischen Umgang mit digitalen Ressourcen. Video/Abstract, Dokumentation, .hist2011, 15.09.2011, URL: [http://www2.hu-berlin.de/historisches-forschungsnetz/tagung/index.php?conference=hist2011&schedConf=index&page=pages&op=view&path\[\]=documentation%2Fpfanzelter](http://www2.hu-berlin.de/historisches-forschungsnetz/tagung/index.php?conference=hist2011&schedConf=index&page=pages&op=view&path[]=documentation%2Fpfanzelter), Stand: 19.12.2014.
- Pfanzelter, Eva: Von der Quellenkritik zum kritischen Umgang mit digitalen Ressourcen, in: Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, 2010, S. 39-49.
- Pfanzelter, Eva: Von der Quellenkritik zum kritischen Umgang mit digitalen Ressourcen, in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken: für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 39-49.

- Pfeifer, Karl-Nikolaus; Frey, Dieter; Rudolph, Matthias: Urhebervertragsrecht in der Reform: Der "Kölnener Entwurf", in: Gewerblicher Rechtsschutz und Urheberrecht, Praxis im Immaterialgüter und Wettbewerbsrecht (1), 2015, S. 1-25.
- Pirker, Eva Ulrike; Rüdiger, Mark: Authentizitätsfiktionen in populären Geschichtskulturen: Annäherungen, in: Echte Geschichte: Authentizitätsfiktionen in populären Geschichtskulturen, Bielefeld 2010 (Historische Lebenswelten in populären Wissenskulturen), S. 11-30.
- Poguntke, Werner: Basiswissen IT-Sicherheit: Das Wichtigste für den Schutz von Systemen und Daten, Dortmund 2013 (Informatik).
- Poster, Mark: Whats the Matter with the Internet?, Minneapolis 2001.
- Postman, Neil: Technopoly: The surrender of culture to technology, New York 1992.
- Potthoff, Jan; Rieger, Sebastian; Johannes, Paul C. u. a.: Elektronisches Laborbuch: Beweiswerterhaltung und Langzeitarchivierung in der Forschung, in: Digitale Wissenschaft: Stand und Entwicklung digital vernetzter Forschung in Deutschland ; 20./21. September 2010, Köln ; Beiträge der Tagung, Köln 2011, S. 149-156.
- Pratschke, Margarete: Die Kunstgeschichte, SAGW-Tagung 13, Bern 28.11.2013.
- Preissner, Andreas: Wissenschaftliches Arbeiten: Internet nutzen – Text erstellen – Überblick behalten, München 2012.
- Preuss, Roland: Guttenbergs Doktorarbeit: Summa cum laude? – «Mehr als schmeichelhaft», in: sueddeutsche.de, 16.02.2011. URL: <http://www.sueddeutsche.de/politik/guttenbergs-doktorarbeit-summa-cum-laude-mehr-als-schmeichelhaft-1.1060779>, Stand: 03.10.2014.
- Prevezanos, Christoph: Computer Lexikon 2013, München 2012 (Always learning).
- Prinz, Claudia: Kommunikation im digitalen Raum, in: Busse, Laura; Enderle, Wilfried; Hohls, Rüdiger u. a. (Hg.): Clio Guide. Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften, Berlin 2016 (Historisches Forum 19). URL: <http://guides.clio-online.de/guides/arbeitsformen-und-techniken/kommunikation-im-digitalen-raum/2016>, Stand: 10.07.2016.
- Purcell, Kristen; Rainie, Lee; Heaps, Alan u. a.: How Teens Do Research in the Digital World. A survey of Advanced Placement and National Writing Project teachers finds that teens' research habits are changing in the digital age, Pew Internet & American Life Project, 01.11.2012. URL: <http://www.pewinternet.org/Reports/2012/Student-Research.aspx>, Stand: 30.12.2013.
- Puschmann, Cornelius: Visualization in the Digital Humanities, 16.08.2010. URL: <http://de.slideshare.net/coffee001/visualization-in-the-digital-humanities>, Stand: 11.11.2013.
- Ranke, Leopold von: Sämtliche Werke, Bd. 33/34, Leipzig 1885.
- Rechenberg, Peter; Pomberger, Gustav (Hg.): Informatik-Handbuch, München 1999.
- Recht, Christian: Webarchivierung und Webharvesting der ÖNB, in: ipCompetence (2), 2009, S. 42-53.
- Reda, Julia: Implementation of Directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society, 2014/2256(INI)-P8_TA-PROV(2015)0273, European Parliament, 09.07.2015. URL: <http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2015-0273&language=EN>, Stand: 18.09.2015.
- Reich, Siegfried (Hg.): Human Computer Interaction, Heidelberg 2013 (HMD - Praxis der Wirtschaftsinformatik).
- Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014.
- Reichert, Ramón: Facebooks Big Data. Die Medien- und Wissenstechniken kollektiver Verdattung, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 437-452.
- Reichert, Ramón: Im Kino der Humanwissenschaften: Studien zur Medialisierung wissenschaftlichen Wissens, Bielefeld 2007 (Kultur- und Medientheorie).

- Reisinger, Gunther: Digital Source Criticism: Net Art as a Methodological Case Study, in: Daniels, Dieter; Reisinger, Gunther (Hg.): Net pioneers 1.0 contextualizing early net-based art, Berlin 2009, S. 123-142. URL: <http://www.netzpioniere.at/node/52>, Stand: 06.05.2018.
- Reppen, Konrad: Über Rankes Diktum von 1824: «Bloss sagen, wie es eigentlich gewesen», in: Historisches Jahrbuch 102, 1982, S. 439-449.
- Richterich, Annika: Infodemiologie – von «Supply» zu «Demand». Google Flu Trends und transaktionale Big Data in der epidemiologischen Surveillance, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 333-363.
- Rieder, Bernhard; Röhle, Theo: Digital Methods: Five Challenges, in: Understanding Digital Humanities, Houndmills, Basingstoke, Hampshire; New York 2012, S. 67-84.
- Rogers, Richard: Digital methods, Cambridge (USA) 2013.
- Röhle, Theo: Big Data – Big Humanities? Eine historische Perspektive, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 157-172.
- Romary, Laurent; Neuroth, Heike; Fritze, Christiane u. a.: DARIAH-EU – Digital Research Infrastructure for the Arts and Humanities, in: Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Digital Humanities und Web 2.0, 2012, S. 36-39.
- Romeyk, Horst: EDV [Elektronische Datenverarbeitung] und Archive, Siegburg 1981².
- Rosenberg, Daniel: Data before the Fact, in: Gitelman, Lisa (Hg.): «Raw data» is an oxymoron, Cambridge (USA) 2013 (Infrastructures series), S. 15-40.
- Rosenthaler, Lukas: Gespräch über digitale Langzeitarchivierung, 04.05.2016.
- Rosenthaler, Lukas: Archivierung im digitalen Zeitalter. Historische Entwicklung und Wege in eine digitale Zukunft, unveröffentlichte Habilitationsschrift, Universität Basel, Basel 2007.
- Rosenthaler, Lukas; Fornaro, Peter; Clivaz, Claire: DASCH: Data and Service Center for the Humanities, in: Digital Scholarship in the Humanities. The Journal of the Alliance of Digital Humanities Organizations 30 (Supplement 1), 2015, S. 43-49.
- Rosenzweig, Roy: Scarcity or Abundance? Preserving the Past, in: Clio wired – The future of the Past in the Digital Age, 2011, S. 1-27.
- Rosner, Bernd: Telematik. Vilém Flusser, in: Kloock, Daniela; Spahr, Angela (Hg.): Medientheorien: eine Einführung, Paderborn 2012 (UTB. Medienwissenschaft, Kommunikationswissenschaft 1986), S. 77-98.
- Rothenberg, Jeff: Digital Preservation in Perspective: How far have we come, and what's next?, Technologie, Future Perfect 2012: Digital Preservation by Design, Wellington 26.03.2012. URL: http://de.slideshare.net/FuturePerfect_/jeff-rothenberg-digital-preservation-perspective#, Stand: 22.07.2014.
- Rothenberg, Jeff: Avoiding technological quicksand: finding a viable technical foundation for digital preservation a report to the Council on Library and Information Resources, Washington (D.C.) Amsterdam 1999.
- Runkehl, Jens; Siever, Torsten: Zitieren und Belegen, in: Gasteiner, Martin; Haber, Peter (Hg.): Digitale Arbeitstechniken für die Geistes- und Kulturwissenschaften, Wien, Köln, Weimar 2010, S. 131-142.
- Sahle, Patrick: Forschungsdaten in den Geisteswissenschaften, in: Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Geisteswissenschaftliche Forschung und ihre Daten, Bern 2015 (SAGW Bulletin), S. 43-45. URL: http://www.sagw.ch/de/dms/sagw/bulletins_sagw/bulletins_2015/Bulletin_4_15/Dossier-SAGW_Bulletin_4_15.pdf.
- Sahle, Patrick: DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum – Milestone 2.3.3, 19.03.2013. URL: https://dev2.dariah.eu/wiki/download/attachments/14651583/DARIAH-M2-3-3_DH-programs_1_1.pdf?version=2&modificationDate=1366904376117&api=v2, Stand: 04.04.2015.

- Sahle, Patrick: Digitale Editionsformen: zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, Bd. 1-3, Norderstedt 2013 (Schriften des Instituts für Dokumentologie und Editorik 7-9).
- Sahle, Patrick; Kronenwett, Simone; Blumtritt, Jonathan: Sustainability?! Four Paradigms for Humanities Data Centers (Poster), Poster, digital humanities Lausanne - Switzerland '14, 10.07.2014, URL: <http://dharchive.org/paper/DH2014/Poster-285.xml>, Stand: 01.04.2015.
- SalahEldeen, Hany M.; Nelson, Michael L.: Losing My Revolution: How Many Resources Shared on Social Media Have Been Lost?, arXiv:1209.3026 [cs.DL], 13.12.2012. URL: <http://arxiv.org/abs/1209.3026>, Stand: 06.05.2018.
- Sandle, Mark: Studying the past in the digital age. From tourist to explorer, in: Weller, Toni (Hg.): History in the digital age, London 2013, S. 129-148.
- Sarasin, Philipp: Schlaue Maschinen. Peter Habers kritische Medienwissenschaft und unsere Lage im Netz heute, in: Gugerli et al. (Hg.): Digital Humanities, Zürich 2013 (Nach Feierabend 9), S. 191-200.
- Saunders, Joss: The future of copyright: what are the pressures on the present system?, in: Cope, Bill; Phillips, Angus (Hg.): The future of the academic journal, Oxford 2014 (Chandos information professional series), S. 249-257.
- Schanze, Helmut: Gibt es ein digitales Apriori?, in: Böhnke, Alexander; Schröter, Jens (Hg.): Analog/Digital – Opposition oder Kontinuum? Zur Theorie und Geschichte einer Unterscheidung, Bielefeld 2004, S. 67-79.
- Schassan, Thorsten: Digitale Quellen: Datei- und Datenformate, in: Busse, Laura; Enderle, Wilfried; Hohls, Rüdiger u. a. (Hg.): Clio Guide. Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften, Berlin 2016 (Historisches Forum 19). URL: <http://guides.clio-online.de/guides/arbeitsformen-und-techniken/digitale-quellen-dateiformate/2016>, Stand: 10.07.2016.
- Scheuermann, Leif: Die Abgrenzung der digitalen Geisteswissenschaften, in: Digital Classics Online 2 (1), 18.02.2016, S. 58-67.
- Scheuermann, Leif: Zu virtuellen Forschungsumgebungen, einer genuin digitalen Hermeneutik sowie deren Visualisierung, Abstract, DHd2016, 12-03.2016, URL: <http://www.dhd2016.de/abstracts/votr%C3%A4ge-024.html>, Stand: 19.02.2018.
- Schmale, Wolfgang: Digital Humanities: Praktiken der Digitalisierung, der Dissemination und der Selbstreflexivität, Stuttgart 2015 (Geschichte).
- Schmale, Wolfgang: Digital Humanities – Einleitung: Begriff, Definition, Probleme, in: Historische Mitteilungen. HMRG 26, 2013, S. 86-93.
- Schmale, Wolfgang: Digitale Geschichtswissenschaft, Wien, Köln 2010.
- Schneider, Jan Georg; Stöckl, Hartmut: Medientheorien und Multimodalität: Zur Einführung, in: Schneider, Jan Georg; Stöckl, Hartmut (Hg.): Medientheorien und Multimodalität: ein TV-Werbespot – sieben methodische Beschreibungsansätze, Köln 2011, S. 10-38.
- Schrage, Dominik: Flüssige Technokratie, in: Merkur deutsche Zeitschrift für europäisches Denken (Jg. 66, H. 9/10 (H. 760/761)), 2012, S. 817-825.
- Schreiber, Catherina: Genuine Internetdaten als historische Quellen – Entwurf einer korrealistischen Quellentheorie, in: Zeitschrift für digitale Geschichtswissenschaften 1, 16.08.2012. URL: <http://universaar.uni-saarland.de/journals/index.php/zdg/article/view/292>, Stand: 11.12.2012.
- Schreibman, Susan; Siemens, Raymond George; Unsworth, John: A companion to digital humanities, Malden 2004 (Blackwell companions to literature and culture 26).
- Schreiner, Julia: Neue (Auf)Schreibsysteme. Verändern Weblogs die Konventionen des geschichtswissenschaftlichen Schreibens?, in: Haber, Peter; Pfanzelter, Eva (Hg.): historyblogosphere. Bloggen in den Geschichtswissenschaften, München 2013, S. 89-99.
- Schüler, Hans-Peter: Urheberrecht: Spanien besteuert Web-Links, Newsticker, heise online, 26.07.2014, URL: <http://www.heise.de/newsticker/meldung/Urheberrecht-Spanien-besteuert-Web-Links-2268747.html>, Stand: 19.09.2014.
- Schulze, Winfried: Einführung in die neuere Geschichte, Stuttgart 2010.

- Schumacher-Gebler, Christian: PIPER ebooks mit digitalem Wasserzeichen, News, Piper Verlag, 23.06.2015, URL: <http://www.piper.de/aktuelles/news/piper-e-books-ohne-kopierschutz>, Stand: 24.07.2015.
- Schwabe, Astrid: Historisches Lernen im World Wide Web: Suchen, flanieren oder forschen? Fachdidaktisch-mediale Konzeption, praktische Umsetzung und empirische Evaluation der regionalhistorischen Website Vimu.info, Bd. 4, Göttingen 2012 (Beihefte zur Zeitschrift für Geschichtsdidaktik).
- Schweizerische Akademie der Geistes- und Sozialwissenschaften: Final report for the pilot project «Data and Service Center for the Humanities» (DaSCH), Schlussbericht, Schweizerische Akademie der Geistes- und Sozialwissenschaften, Bern 30.03.2015. URL: www.akademien-schweiz.ch/fr/dms/publikationen/10/report1001.pdf, Stand: 06.05.2018.
- Schweizerische Akademie der Geistes- und Sozialwissenschaften (Hg.): Dossier Digital Humanities und Web 2.0, Bern 2012 (SAGW Bulletin). URL: http://www.sagw.ch/dms/sagw/bulletins_sagw/bulletins_2012/SAGWbulletin1-12.
- Schweizerische Gesellschaft für Geschichte: Le pouvoir des algorithmes. Technologies numériques et rapports de force au XXe-XXIe siècles, Panel Detail, 4. Schweizerische Geschichtstage, 09.06.2016, URL: <https://www.geschichtstage.ch/panel/52/le-pouvoir-des-algorithmes-technologies-numeriques-et-rapports-de-force-au-xxe-xxie-siecles>, Stand: 10.04.2016.
- Schweizerische Nationalbibliothek: e-Helvetica, URN-Handbuch, 04.2014. URL: https://www.nb.admin.ch/nb_professionnel/01693/01695/01706/index.html?lang=de, Stand: 20.04.2016.
- Schweizerischer Bundesrat: Verordnung über die Erfindungspatente (Patentverordnung), 01.09.2014. URL: <https://www.admin.ch/opc/de/classified-compilation/19770250/index.html>, Stand: 06.10.2015.
- Schweizerischer Bundesrat: Verordnung über die Führung und Aufbewahrung der Geschäftsbücher (Geschäftsbücherverordnung; GeBüV), 01.01.2013. URL: <http://www.admin.ch/opc/de/classified-compilation/20001467/index.html>.
- Schweizerisches Bundesarchiv: Archivtaugliche Dateiformate. Standards für die Archivierung digitaler Unterlagen, 01.2014. URL: <http://www.bar.admin.ch/dienstleistungen/01893/index.html?lang=de>, Stand: 09.09.2015.
- Senate and House of Representatives of the United States of America: Copyright Term Extension Act, 27.10.1998. URL: <http://www.copyright.gov/legislation/pl105-298.pdf>, Stand: 05.04.2015.
- Shannon, Claude Elwood: A Mathematical Theory of Communication, in: The Bell System Technical Journal (27), 1948, S. 379-423.
- Siemens, Raymond George: Understanding and Responding to (Inter)Disciplinary Change, unpubl. Vortrag, DH Summer School Switzerland, University of Berne 26.03.2013.
- Smith, Abby: Introduction, in: Cullen et al. (Hg.): Authenticity in a Digital Environment, Washington, D.C 2000, S. VI–VIII.
- Smith, Abby: Why digitize?, Washington (D.C.) 1999.
- soft Xpansion GmbH & Co. KG: Sichere PDF-Dokumente durch digitale Signaturen, 2011. URL: <http://soft-xpansion.de/utills/csp/sichere%20pdf-dokumente%20durch%20digitale%20signaturen.pdf>, Stand: 05.11.2015.
- Spahr, Angela: Magische Kanäle. Marshall McLuhan, in: Kloock, Daniela; Spahr, Angela (Hg.): Medientheorien: eine Einführung, Paderborn 2012 (UTB. Medienwissenschaft, Kommunikationswissenschaft 1986), S. 39-76.
- Spiro, Lisa: «This Is Why We Fight»: Defining the Values of the Digital Humanities, in: Gold, Matthew K. (Hg.): Debates in the digital humanities, Minneapolis 2012, S. 16-35. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- Stark, Birgit; Dörr, Dieter; Aufenanger, Stefan: Die Googleisierung der Informationssuche – Suchmaschinen im Spannungsfeld zwischen Nutzung und Regulierung, Universität Mainz, Mainz 16.12.2014.
- Stern, Fritz; Osterhammel, Jürgen (Hg.): Moderne Historiker: Klassische Texte von Voltaire bis zur Gegenwart, München 2011.

- Sternfeld, Joshua: Historical Understanding in the Quantum Age, Bd. 3, Nr. 2 (Sommer), Journal of Digital Humanities, 25.08.2014, URL: <http://journalofdigitalhumanities.org/3-2/historical-understanding-in-the-quantum-age/>, Stand: 04.08.2015.
- Stiegler, Bernard: Licht und Schatten im digitalen Zeitalter. Programmatische Vorlesung auf dem Digital Inquiry Symposium am Berkeley Center for New Media, in: Reichert, Ramón (Hg.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, Bielefeld 2014, S. 35-46.
- Stiller, Andreas: 20 Jahre FDIV-Bug: Ein Prozessor-Rechenfehler macht Geschichte, Newsticker, heise online, 30.10.2014, URL: <http://www.heise.de/newsticker/meldung/20-Jahre-FDIV-Bug-Ein-Prozessor-Rechenfehler-macht-Geschichte-2438283.html>, Stand: 18.11.2014.
- Stiller, Burkhard; Thouvenin, Florent: Gutachten: Netzsperrern, 16.09.2016. URL: <http://www.itsl.uzh.ch/dam/jcr:9607879d-3e04-4de2-b25a-bf46fa58d5ac/Gutachten%20Netzsperrern-final.pdf>, Stand: 26.04.2018.
- Stock, Wolfgang G.: Was ist eine Publikation? Zum Problem der Einheitenbildung in der Wissenschaftsforschung, in: Gesellschaft für Wissenschaftsforschung e.V. (Hg.): Wissenschaft und Digitale Bibliothek: Wissenschaftsforschung Jahrbuch 1998, Berlin 2010². URL: http://www.wissenschaftsforschung.de/JP98_239-282.pdf, Stand: 23.06.2015.
- Stone, Brad: Amazon Erases Orwell Books From Kindle, in: The New York Times, 18.07.2009. URL: <http://www.nytimes.com/2009/07/18/technology/companies/18amazon.html>, Stand: 19.03.2015.
- Strasser, Bruno J.; Edwards, Paul N.: Big data is the answer... but what is the question?, unveröffentlichtes Paper [zugestellt per E-Mail am 18.10.2015], infoclio.ch-Tagung: Daten und Geschichtswissenschaften, Bern 16.10.2015.
- Strohm, Klauspeter: Kulturwissenschaften und Medienkompetenz. Einige Überlegungen, in: Historische Anthropologie Kultur, Gesellschaft, Alltag (1), 2001.
- Stroud, Ronald S.: «Wie es eigentlich gewesen» and Thucydides 2.48.3, in: Hermes 115 (H. 3), 1987, S. 379-382.
- Surowiecki, James: The wisdom of crowds : why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations, London 2004.
- Svensson, Patrik: Beyond the Big Tent, in: Debates in the digital humanities, Minneapolis 2012, S. 36-49. URL: <http://www.upress.umn.edu/book-division/books/debates-in-the-digital-humanities>.
- Talbot, David; Schwan, Ben: Hongkong: Mesh-Netzwerk ermöglicht Demonstranten die Kommunikation, Infotech, Technology Review, 06.10.2014, URL: <http://www.heise.de/tr/artikel/Hongkong-Mesh-Netzwerk-ermoeglicht-Demonstranten-die-Kommunikation-2411653.html>, Stand: 07.10.2014.
- Tantner, Anton: Die ersten Suchmaschinen : Adressbüros, Fragämter, Intelligenz-Comptoirs, Berlin 2015.
- Taylor, Arlene G.; Joudrey, Daniel N.: The organization of information, Westport (USA) 2009 (Library and information science text series).
- Team #hwzmobile: Das Internet; kein rechtsfreier Raum, CAS Mobile Busines, Hochschule für Wirtschaft Zürich – Center for Digital Busines, 23.01.2015, URL: <http://hwzdigital.ch/das-internet-kein-rechtsfreier-raum/>, Stand: 03.03.2015.
- Terras, Melissa; Nyhan, Julianne; Vanhoutte, Edward (Hg.): Defining digital humanities: a reader, Farnham 2013.
- Text Encoding Initiative: TEI: Customization, Guidelines, TEI, 07.04.2015, URL: <http://www.tei-c.org/Guidelines/Customization/>, Stand: 09.04.2015.
- Thelwall, Michael: Introduction to Webometrics: Quantitative Web Research for the Social Sciences, in: Synthesis Lectures on Information Concepts, Retrieval, and Services 1 (1), 01.01.2009, S. 1-116. DOI: 10.2200/S00176ED1V01Y200903ICR004.
- Thies, Justus; Zöllhöfer, Michael; Stamminger, Marc u. a.: Face2Face: Real-time Face Capture and Reenactment of RGB Videos. URL: <http://www.graphics.stanford.edu/~niessner/thies2016face.html>, Stand: 26.06.2016.
- Thillosen, Anne: Schreiben im Netz: Neue literale Praktiken im Kontext Hochschule, Münster 2008 (Medien in der Wissenschaft).

- Thirlwell, Adam: Kapow!, London 2012.
- Thomas, Hartwig: Erster Entwurf eines Versuchs über den Zusammenstoss des Urheberrechts mit dem Internet, Zürich 2014. URL: <https://buchundnetz.com/online-buch/erster-entwurf-eines-versuchs-ueber-den-zusammenstoss-des-urheberrechts-mit-dem-internet/inhaltsverzeichnis/>.
- Thomas, William G., III: The Future of Digital History, #rrchnm20, William G. Thomas III, 16.11.2014, URL: <http://railroads.unl.edu/blog/?p=1146>, Stand: 19.11.2014.
- Tkacz, Nathaniel: Wikipedia and the politics of openness, Chicago 2015.
- Toebak, Peter M.: Records Management ein Handbuch, Baden 2007.
- Tripathi, Arun Kumar: The Significance of Digital Hermeneutics for the Philosophy of Technology, in: Kelly, Matthew; Bielby, Jared (Hg.): Information Cultures in the Digital Age: A Festschrift in Honor of Rafael Capurro, Wiesbaden 2016, S. 143-160.
- Tufte, Edward Rolf: The visual display of quantitative information, Cheshire, Connecticut 1984. URL: <http://www.colorado.edu/UCB/AcademicAffairs/ArtsSciences/geography/foote/maps/assign/reading/TufteCoversheet.pdf>, Stand: 06.10.2015.
- Turkel, William J.: Intervention. Hacking history, from analogue to digital and back again, in: De Groot, Jerome (Hg.): Public and popular history, London 2012, S. 139-148.
- Turkle, Sherry: Verloren unter 100 Freunden: wie wir in der digitalen Welt seelisch verkümmern, München 2012.
- Turner, Vernon; Gantz, John F.; Reinsel, David u. a.: White Paper: The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things, 04.2014. White Paper unter URL: <http://idcdocserv.com/1678>, Stand: 02.10.2014; Präsentation unter URL: <https://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>, Stand: 06.05.2018.
- Twitter Inc.: Quarterly Report, Form 10-Q, United States Securities and Exchange Commission, Washington (D.C.) 11.08.2014, S. 65. URL: http://www.sec.gov/Archives/edgar/data/1418091/000156459014003474/twtr-10q_20140630.htm, Stand: 30.06.2015.
- Uhl, Heidemarie; Rössner, Michael: Renaissance der Authentizität? : Über die neue Sehnsucht nach dem Ursprünglichen, Bielefeld 2012 (Kultur- und Medientheorie).
- Universität Zürich: Chaosseminar: AGUR12 und Urheberrechtsentwicklungen, Agenda, Universität Zürich, 11.12.2014, URL: <http://www.agenda.uzh.ch/record.php?id=25153>, Stand: 04.04.2015.
- Vaidhyanathan, Siva: The Googlization of everything (and why we should worry), Berkeley 2011.
- Valentino, Maura L.: Integrating Metadata Creation into Catalog Workflow, in: Eden, Bradford Lee (Hg.): Twenty-first century metadata operations challenges, opportunities, directions, London 2012, S. 40-49.
- Van Alstyne, Marshall; Brynjolfsson, Erik: Global village or cyber-balkans? Modeling and measuring the integration of electronic communities, in: Management Science 51 (6), 2005, S. 851-868.
- Van Noorden, Richard: Open access: The true cost of science publishing, in: Nature 495 (7442), 27.03.2013, S. 426-429. DOI: 10.1038/495426a.
- «Vanellus»: Wikipedia, in: Wikipedia, 04.05.2015. URL: <https://de.wikipedia.org/w/index.php?title=Wikipedia&oldid=141752818>, Stand: 08.05.2015.
- Verband der Historiker und Historikerinnen Deutschlands: Quellenkritik im digitalen Zeitalter. Die Historischen Grundwissenschaften als zentrale Kompetenz der Geschichtswissenschaft und benachbarter Fächer, Stellungnahmen, VHD, 10.2015, URL: <http://www.historikerverband.de/verband/stellungnahmen/quellenkritik.html>, Stand: 29.11.2015.
- Verein Schweizerischer Archivarinnen und Archivare (VSA): Kodex ethischer Grundsätze, VSA – Verein Schweizerischer Archivarinnen und Archivare, 1997, URL: <http://vsa-aas.ch/beruf/ethikkodex/>, Stand: 06.05.2018.
- Verhoef, Jesper; Wevers, Melvin: The Digital Humanities Cycle: Hermeneutics, Heuristics, and Source Criticism in a Digital Age, in: Digital Humanities Benelux Conference, Antwerpen 2015. URL: <http://www.dhbenelux.org/wp-content/uploads/2015/04/20.pdf>, Stand: 06.05.2018.

- Verhoef, Jesper; Wevers, Melvin: The Digital Humanities Cycle: Doing Digital History Iteratively (Presentation), Digital Humanities Benelux Conference, Antwerpen 08.06.2015. URL: <http://www.slideshare.net/MelvinWevers/the-digital-humanities-cycle-doing-digital-history-iteratively>, Stand: 04.08.2015.
- Verwayen, Harry; Arnoldus, Martijn; Kaufman, Peter B.: The Problem of the Yellow Milkmaid, White Paper 2, Europeana, 2011. URL: <http://pro.europeana.eu/publications>, Stand: 17.03.2014.
- Vinyals, Oriol; Toshev, Alexander; Bengio, Samy u. a.: Show and Tell: A Neural Image Caption Generator, arXiv – Cornell University Library, arXiv:1411.4555 [cs.CV], 17.11.2014. URL: <http://arxiv.org/abs/1411.4555>, Stand: 19.11.2014.
- Volz, Stephanie: Trennungsgebot und Internet: Ein medienrechtliches Prinzip in Zeiten der Medienkonvergenz, Bd. 58, Zürich 2014 (Publikationen aus dem Zentrum für Informations- und Kommunikationsrecht der Universität Zürich).
- Vompras, Johanna; Schirrwagen, Jochen; Horstmann, Wolfram: Die Bibliothek als Dienstleister für den Umgang mit Forschungsdaten, in: Digitale Wissenschaft: Stand und Entwicklung digital vernetzter Forschung in Deutschland ; 20./21. September 2010, Köln ; Beiträge der Tagung, Köln 2011, S. 101-106. URL: <http://pub.uni-bielefeld.de/publication/2422868>.
- W3C/IETF URI Planning Interest Group: URIs, URLs, and URNs: Clarifications and Recommendations 1.0, World Wide Web Consortium (W3C), 21.09.2001, URL: <http://www.w3.org/TR/uri-clarification/>, Stand: 30.07.2014.
- Walkowski, Niels-Oliver: Text, Denken und E-Science, Eine intermediale Annäherung an eine Konstellation, in: Gugerli et al. (Hg.): Digital Humanities, Zürich 2013 (Nach Feierabend 9), S. 37-54.
- Wampfler, Philippe: «online fist». Geisteswissenschaften als Social Media, in: Gugerli et al. (Hg.): Digital Humanities, Zürich 2013 (Nach Feierabend 9), S. 79-102.
- Wang, Gang; Wilson, Christo; Zhao, Xiaohan u. a.: Serf and Turf: Crowdturfing for Fun and Profit, arXiv:1111.5654 [cs.SI], 23.11.2011. URL: <http://arxiv.org/abs/1111.5654>, Stand: 03.09.2014.
- Warwick, Claire; Terras, Melissa; Nyhan, Julianne (Hg.): Digital humanities in practice, London 2012.
- Wathelet, Melchior: Opinion of Advocate General Wathelet. Case C-160/15 (GS Media BV v Sanoma Media Netherlands BV, Playboy Enterprises International Inc., Britt Geertuida Dekker), ECLI:EU:C:2016:221, Gerichtshof der Europäischen Union, Luxembourg 07.04.2016. URL: <http://curia.europa.eu/>.
- Weber, Rolf H.: E-Commerce und Recht: rechtliche Rahmenbedingungen elektronischer Geschäftsformen, Zürich 2010.
- Weber, Volker: Verse: Die Neuerfindung der E-Mail, according to IBM, News, heise online, 18.11.2014, URL: <http://www.heise.de/newsticker/meldung/Verse-Die-Neuerfindung-der-E-Mail-according-to-IBM-2459147.html>, Stand: 19.11.2014.
- Weigert, Martin: Onlinepräsenzen gehen verloren: Die mobile Revolution und das vergessliche Web, netzwertig, 15.01.2014, URL: <http://netzwertig.com/2014/01/15/onlinepraesenzen-gehen-verloren-die-mobile-revolution-und-das-vergessliche-web/>, Stand: 06.03.2014.
- Weinberger, David: Too Big To Know; Rethinking Knowledge. Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room, New York 2011.
- Weingart, Scott: A Working Definition of Digital Humanites, the scottbot irregular, 08.12.2013, URL: <http://www.scottbot.net/HIAL/?p=39749>, Stand: 25.02.2014.
- Weller, Katrin; Dröge, Evelyn; Puschmann, Cornelius: Citation Analysis in Twitter: Approaches for Defining and Measuring Information Flows within Tweets during Scientific Conferences, in: #MSM2011, 1st Workshop on Making Sense of Microposts, Heraklion, Kreta 2011, S. 12. URL: http://ceur-ws.org/Vol-718/paper_04.pdf, Stand: 06.05.2018.
- Weller, Toni (Hg.): History in the digital age, London 2013.
- Weller, Toni: Conclusion. A changing field, in: Weller, Toni (Hg.): History in the digital age, London 2013, S. 195-205.
- Weller, Toni: Introduction, in: Weller, Toni (Hg.): History in the digital age, London 2013, S. 1-19.

- Wellerstein, Alex: The year of the disappearing websites, Restricted Data: The Nuclear Secrecy Blog, 27.12.2013, URL: <http://blog.nuclearsecrecy.com/2013/12/27/year-disappearing-websites/>, Stand: 01.08.2014.
- Welskopp, Thomas: Erklären, in: Goertz, Hans-Jürgen (Hg.): Geschichte. Ein Grundkurs, Reinbek bei Hamburg 1998², S. 132-168.
- Wermke, Matthias; Klosa, Annette; Kunkel-Razum, Kathrin u. a. (Hg.): Duden - das Fremdwörterbuch, Bd. 5, Mannheim, Leipzig, Wien, Zürich 2001 (Der Duden in zwölf Bänden).
- Wiedmer, Hans-Rudolf: Publizieren im Zeitalter von Open Access. Die Verlagsperspektive, in: Traverse : Zeitschrift für Geschichte 22 (1), 2015, S. 147-156.
- Wiegand, Dorothee: Wikipedia: Zu gut, um nicht zitiert zu werden?, Newsticker, heise online, 04.10.2014, URL: <http://www.heise.de/newsticker/meldung/Wikipedia-Zu-gut-um-nicht-zitiert-zu-werden-2411605.html>, Stand: 07.10.2014.
- Wilbanks, John: Licence restrictions: A fool's errand, in: Nature 495 (7442), 28.03.2013, S. 440-441. DOI: 10.1038/495440a.
- WIPO: WIPO-Administered Treaties: Convention Establishing the World Intellectual Property Organization (28.09.1979), WIPO, URL: http://www.wipo.int/wipolex/en/treaties/text.jsp?file_id=283854#P50_1504, Stand: 31.08.2015.
- Wirth, Uwe: Hypertext, in: Grundbegriffe der Medientheorie, Paderborn 2005, S. 86-94.
- Wissenschaftsrat der Bundesrepublik Deutschland: Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020, 2359-12, Köln 13.06.2012. URL: <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>, Stand: 24.01.2014.
- Wittenhorst, Tilman: Zensur befürchtet: Türkei will Websites erst nach Lizenzvergabe zulassen, Newsticker, heise online, 06.02.2018, URL: <https://www.heise.de/newsticker/meldung/Zensur-befuerchtet-Tuerkei-will-Websites-erst-nach-Lizenzvergabe-zulassen-3961490.html>, Stand: 07.02.2018.
- Worm, Hans-Jürgen: Neue Informationstechnologien und Archive, in: Reimann, Norbert (Hg.): Praktische Archivkunde : ein Leitfaden für Fachangestellte für Medien- und Informationsdienste : Fachrichtung Archiv, Münster 2008, S. 219-237.
- Wrana, Daniel: DiskursNetz: Wörterbuch der interdisziplinären Diskursforschung, Berlin 2014 (Suhrkamp Taschenbuch Wissenschaft).
- Wulf, Tobias; Foerster, Sascha: Wissenschaftliches Bloggen: Digitale Miszelle oder kommunikativer "Turn"? Vortrag von Tobias Wulf und Sascha Foerster #digigw14, Digitale Geschichtswissenschaft, 16.09.2014, URL: <http://digigw.hypotheses.org/838>, Stand: 18.11.2014.
- Wytrzens, Hans Karl; Schauppenlehner-Kloyber, Elisabeth; Sieghardt, Monika u. a.: Wissenschaftliches Arbeiten: eine Einführung, Wien 2014.
- Zaagsma, Gerben: On Digital History, in: BMGN – Low Countries Historical Review 128 (4), 16.12.2013, S. 3-29.
- Zhang, Jingyu; Gecevičius, Mindaugas; Beresna, Martynas u. a.: 5D Data Storage by Ultrafast Laser Nanostructuring in Glass, University of Southampton, 2013. URL: <https://eprints.soton.ac.uk/364916/1/5973.pdf>, Stand: 06.05.2018.

9. Anhang

9.1. Zusammenfassung Lösungsansätze

Nachfolgend werden die von mir vorgeschlagenen Lösungsansätze noch einmal aufgeführt und abschliessend kurz erklärt:

- Quellengattungsschema anpassen:¹¹⁴³ Digitale Objekte werden im angepassten Quellengattungsschema als virtuell-elektronische Quellenart verortet.

Überreste						
materiell / physisch				immateriell / virtuell		
Überreste im engeren Sinne	Denkmäler	Bildlich	Schriftlich	Audio/ audiovisuell	Zustände	Elektronisch
<ul style="list-style-type: none"> - körperliche Reste - Werkzeuge - Wohnen - Produkte 	<ul style="list-style-type: none"> - Inschriften - Monumente - Urkunden 	<ul style="list-style-type: none"> - historische Gemälde - topographische Darstellungen - historische Skulpturen 	<ul style="list-style-type: none"> - historische Inschriften - Genealogien - Kalender - Annalen - Chroniken - Biographien - Memoiren - Urkunden und Akten (Vorgänge rechtlicher Natur) - Literatur - Geschäftsakten 	<ul style="list-style-type: none"> - Erzählung - Sage - Anekdoten - Sprichwörter - historische Lieder - Sprache - Ton, Musik 	<ul style="list-style-type: none"> - Sitten - Prozesse - Institutionen 	<ul style="list-style-type: none"> - Darstellung - Graphiken (dynamisch / statisch) - Texte - Webseiten - Musik - Film - Versteckt - Programm bibliotheken - Code - Datenbank-elemente

Tabelle 11: Quellengattungsschema nach Föhr (siehe Kap. 5.2.2.)

- Quellensicherung als eigenständigen Prozessschritt einführen:¹¹⁴⁴ Da digitale Objekte volatil sind, manipuliert werden können und nicht persistent sind, sollte der Forschende alle potenziellen digitalen Quellen in verlässlichen, manipulationssicheren Forschungsinfrastrukturen sichern. Vom Forschenden sind fundierte Medien- und Informatikkenntnisse gefordert, damit alle nötigen Elemente eines digitalen Objekts in angemessener Form und in einem langzeitarchivfähigen Format als Surrogat gesichert werden und anschliessend Drittpersonen zur kritischen Beurteilung zur Verfügung stehen.

¹¹⁴³ Siehe Kap. 5.2. Quellenkunde.

¹¹⁴⁴ Siehe Kap. 3.3. Objekt, Forschungsressource, Quelle und Kap. 5.3. Quellensicherung.

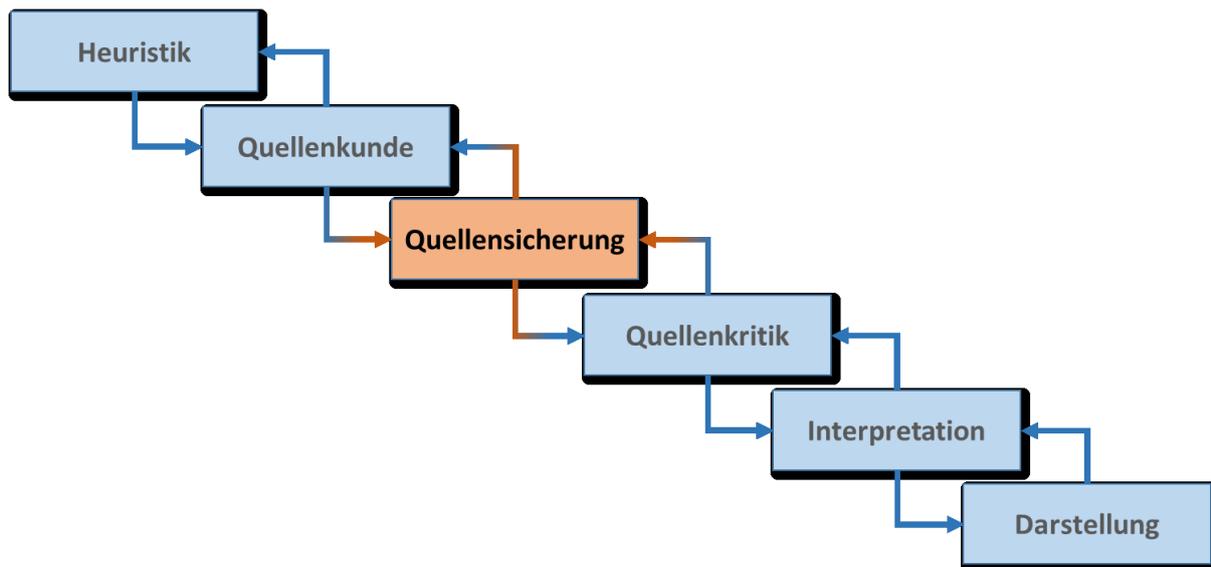


Abbildung 31: Historisch-kritischer Prozess nach Föhr (Kap. 5.)

- Researcher Driven Archiving (RDA) ermöglichen:¹¹⁴⁵ Forschende sollen selbständig die für ihre Arbeit verwendeten digitalen Quellen sichern können. Damit werden die tatsächlich forschungsrelevanten Objekte gesichert und der utopische Versuch der Speicherung aller digitalen Daten, insbesondere des Webs, kann abgebrochen werden.
Für die Speicherung sollten Trusted Digital Repositories (TDR) aufgebaut werden. Da mit dem Aufbau und Betrieb dieser Infrastrukturen hohe finanzielle Kosten verbunden sind und von ihnen die Nachvollziehbarkeit der verwendeten Quellen abhängt, sollten die dauerhafte Finanzierung und der Betrieb von Institutionen wie Bundes-, Staats- oder Universitätsarchiven mit einem gesetzlichen Auftrag gesichert werden.
- Trusted Saving Procedure (TSP) entwickeln:¹¹⁴⁶ Mit diesem softwaretechnischen Verfahren können digitale Objekte manipulationssicher in die Speicherinfrastrukturen überführt werden. Die Procedure und die genannten Repositories sollten gemeinsam (weiter-)entwickelt werden und die dafür nötigen Mittel sollten dauerhaft zur Verfügung gestellt werden.
- Personal Trusted Archive (PTA) entwickeln:¹¹⁴⁷ Mit dieser Software, die eine angepasste TSP enthält, können Forschende und andere Nutzer digitale Objekte in vertrauenswürdigen und manipulationsresistenten, aber persönlichen und lokalen Archiven speichern. Für die Benutzerfreundlichkeit sollten solche Archive in Literaturverwaltungsprogramme oder Data Management Systeme implementiert werden.

¹¹⁴⁵ Siehe Kap. 5.3.6. Sicherungsort.

¹¹⁴⁶ Siehe Kap. 5.3.6. Sicherungsort.

¹¹⁴⁷ Ebd.

- Metadata Information Section (MIS) einführen:¹¹⁴⁸ Die für eine wissenschaftliche oder wissenschaftsnaher Publikation relevanten bibliographischen Metadaten sollten in einem separaten Abschnitt des digitalen Objekts im Klartext enthalten sein. Dieselben Metadaten, ergänzt mit Sicherheitsmerkmalen wie dem zugehörigen Hashwert, sollten auf der Publikationsplattform ebenfalls im Klartext und maschinenlesbar aufgeführt sein. Um den Autoren diese Arbeit zu erleichtern, könnten entsprechende Module für bestehende Softwares wie Microsoft Word oder Wordpress entwickelt werden, die die Erstellung der MIS automatisieren. Ob sich dieser Vorschlag durchsetzen wird, hängt von der Benutzerfreundlichkeit wie auch davon ab, ob wissenschaftliche Institutionen diesen unterstützen und umsetzen. In der MIS sehe ich grosses Potenzial zur Verbesserung der Vertrauenswürdigkeit von digitalen Objekten und zur Sicherstellung der Vollständigkeit bibliographischer Angaben.
- Digitale Signatur verwenden:¹¹⁴⁹ Die Digitale Signatur stellt zurzeit die technisch beste und einfachste Lösung dar, um die Integrität und Authentizität eines digitalen Objekts zu garantieren. Die bei Wissenschaftlern fast unbekannt Technik könnte mit einer Informationskampagne, der kostenlosen Abgabe von Signaturschlüsseln und einem gewissen Druck durch die wissenschaftlichen Institutionen eingeführt werden.
- Persistenter Identifikator (PI) als Permalinks:¹¹⁵⁰ Mit der Vergabe von persistenten Identifikatoren an digitale Objekte durch ein TDR können tatsächlich persistente Adressen vergeben werden. Das Objekt wäre damit eindeutig identifizierbar und direkt adressierbar. Bei der Umsetzung dieses Systems ist allerdings mit Widerstand von Seiten der dadurch nicht mehr benötigten Resolver-Dienstleister zu rechnen.
- Eindeutige Datierung per Zeitstempel verwenden:¹¹⁵¹ Anstelle der sich in digitalen Objekten oft widersprechenden Datumsangaben generiert eine qualifizierte elektronische Signatur mit Zeitstempel eindeutige Zeit- und Datumsangaben, die als bibliographische Informationen verwendet werden können.
- Erscheinungsort durch persistenten Identifikator ersetzen:¹¹⁵² Da bei digitalen Objekten der Erscheinungsort keine Rolle spielt und der Speicherort nicht persistent ist, sollte ein von einem TDR vergebener, eindeutiger und persistenter Identifikator verwendet werden.
- Link-Check-Element verwenden:¹¹⁵³ Mit dem Link-Check untersucht ein Autor die in einem digitalen Objekt enthaltenen Hyperlinks und weist mit dem Link-Check-Element der Belegkette nach, welche Links er wie tief nachverfolgt hat. Der Leser kann

¹¹⁴⁸ Siehe Kap. 5.3.7. Metadaten.

¹¹⁴⁹ Siehe Kap. 5.4.2. Integrität.

¹¹⁵⁰ Siehe Kap. 5.4.3. Persistenz.

¹¹⁵¹ Siehe Kap. 5.4.6. Datierung.

¹¹⁵² Siehe Kap. 5.4.7. Entstehungs- / Erscheinungsort.

¹¹⁵³ Siehe Kap. 5.4.12. Beziehungen zu Objekten.

anhand der entsprechenden Link-Check-Ziffer nachvollziehen, ob zusätzliche und damit vom Autor nicht explizit erwähnte Informationen für die Interpretation des digitalen Objekts verwendet wurden. Das Link-Check-Element ist ein neues Element in der Belegkette und wäre beispielsweise in Softwares für Literaturverwaltung zu implementieren.

- Authentizitätsapproximation (AP) berechnen.¹¹⁵⁴ Der Nutzer errechnet bei der Authentizitätsapproximation einen subjektiven Prozentwert anhand von sieben gewichteten Kriterien, der seine Einschätzung der Authentizität eines Objekts ausdrückt. Diese Methode verhilft zu einer strukturierten und nachvollziehbaren Beurteilung eines Objekts. Es wäre benutzerfreundlich, wenn diese Berechnung in bestehende Literaturverwaltungssoftwares integriert werden würde. Darin könnte sie mit den Metadaten des digitalen Objekts und allenfalls auch mit seinem Surrogat verknüpft werden. Die AP sollte auch in der Aus- und Weiterbildung von Historikern als Teil der wissenschaftlichen Arbeitsmethodik vermittelt werden.

¹¹⁵⁴ Siehe Kap. 5.4.12. Beziehungen zu Objekten, Lösungsansatz Authentizitätsapproximation.