

Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften

Fabian Cremer, Max Weber Stiftung

Lisa Klaffki, DARIAH-DE / Herzog August Bibliothek Wolfenbüttel

Timo Steyer, Forschungsverbund Marbach Weimar Wolfenbüttel / Herzog August Bibliothek Wolfenbüttel

Zusammenfassung:

Forschungsdaten nehmen in den Geistes- und Kulturwissenschaften eine immer wichtigere Rolle im Forschungsprozess ein. Doch dieser Vorgang ist nicht frei von Widersprüchen und Missverständnissen – das aus den Naturwissenschaften stammende Konzept „Forschungsdaten“ kann nur bedingt auf traditionelle Arbeitsverfahren in den Geisteswissenschaften übertragen werden. Der vorliegende Beitrag thematisiert anhand der Spezifika geisteswissenschaftlicher Forschungsdaten das Spannungsverhältnis zwischen Relevanz und Wahrnehmung bei den Forschenden sowie zwischen Konzeption und Umsetzung im Forschungsalltag.

Summary:

Research data is getting more and more important in the humanities and cultural studies and now plays an important role in the research process. But contradictions and misunderstandings are also part of this development. The concept of research data comes from the sciences, and can only be transferred to the traditional scholarly methods of the humanities to a limited degree. Based on the specifics of research data in the humanities, the article deals with the tensions between relevance and familiarity among the researchers as well as between the conception and implementation in the research process.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/2018H2S142-162>

Autorenidentifikation: Cremer, Fabian: GND: 14061773, ORCID: <https://orcid.org/0000-0001-8251-9727>; Klaffki, Lisa: GND: 1041562500, ORCID: <https://orcid.org/0000-0002-8791-7422>;

Steyer, Timo: GND: 1053806175, ORCID: <http://orcid.org/0000-0003-0218-2269>

Schlagwörter: Forschungsdaten, Digital Humanities

1. Genese mit Widersprüchen

In einem Blogbeitrag stellte Thomas Stäcker die Frage, ob Forschungsdaten in den Geisteswissenschaften “[...] lediglich eine Chimäre, ein von den Technik- und Naturwissenschaften geborgter Begriff [sein]?”¹

1 Thomas Stäcker, „Noch einmal: Was sind geisteswissenschaftliche Forschungsdaten?“, *DHd-Blog*, 6. Dezember 2015, zuletzt geprüft am 16.01.2018, <http://dhd-blog.org/?p=5995>. Der Blogbeitrag war eine Replik auf ein von Peter Andorfer verfasstes DARIAH-DE Working Paper. Vgl. Peter Andorfer, *Forschungsdaten in den (digitalen) Geisteswissenschaften. Versuch einer Konkretisierung*, DARIAH-DE Working Papers 14 (Göttingen: DARIAH-DE, 2015), <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2>.

Gut zwei Jahre nach der Formulierung dieser Frage lässt sich eine zunehmende Relevanz von Forschungsdaten in geisteswissenschaftlichen Forschungsprozessen feststellen, die sich auch in einer Institutionalisierung in Form von Arbeitsstellen und Datenzentren ausdrückt.² Dabei spielen die Aspekte der Forschungsdatenpublikation und der Anreicherung von Forschungsdaten hinsichtlich der weiteren Nachnutzung eine zentrale Rolle. Das Chimärenhafte bleibt diesem Begriff jedoch weiterhin verhaftet, nicht länger als Trugbild, dafür als vielgestaltiges Mischwesen erscheinen die unterschiedlichen Praktiken und Perspektiven zu Forschungsdaten in den Geisteswissenschaften.³ In dieser noch nicht abgeschlossenen Begriffsgenese wird der Begriff Forschungsdaten auf unterschiedliche Qualitäts- und Aggregatsstufen von Daten, angefangen bei Grunddaten wie Digitalisaten oder maschinell erstellten Volltexten bis zu nach komplexen Richtlinien erschlossenen Datensammlungen oder Datenrepräsentation bzw. Datenbanken, angewendet und ist nicht frei von Widersprüchen und Akzeptanzproblemen.⁴

Während einige Wissenschaftler/innen vor allem im Bereich der digitalen Geisteswissenschaften (Digital Humanities) von einem Paradigmenwechsel hin zu maschinenlesbaren, strukturierten und frei nachnutzbaren Daten sprechen, hat der/die eher traditionell hermeneutisch arbeitende Geisteswissenschaftler/in oft Vorbehalte gegenüber der Publikation von Daten, die als Arbeitsmittel auf dem Weg zur Monographie betrachtet werden, wie z.B. Exceltabellen oder Filemakerdatenbanken sowie dem zeitintensiven Anreichern dieser Arbeitsdaten. Besonders in diesem Kontext sind im Gegensatz zu den in standardisierten Digitalisierungsworkflows generierten Forschungsdaten die skizzierten Widersprüche virulenter und die Positionen noch nicht konsensfähig.⁵ Auch kann generell hinterfragt

- 2 So widmet sich das Historische Datenzentrum Sachsen-Anhalt an der Universität Halle-Wittenberg dem geisteswissenschaftlichen Forschungsdatenmanagement. Vgl. <http://www.geschichte.uni-halle.de/struktur/hist-data/>, zuletzt geprüft am 16.01.2018. Auch wäre in diesem Kontext das Projekt „Geisteswissenschaftliche Infrastruktur für Nachhaltigkeit“ (gwin) zu nennen, welches im Rahmen des Zentrums für nachhaltiges Forschungsdatenmanagement eine Infrastruktur für Forschungsdaten und Forschungsanwendungen an der geisteswissenschaftlichen Fakultät der Universität Hamburg erstellt. Vgl. <https://www.fdm.uni-hamburg.de/ueber-uns/gwin.html>, zuletzt geprüft am 16.01.2018. Eine Bündelung der Initiativen von geisteswissenschaftlichen Datenzentren findet in der Arbeitsgruppe Datenzentren des DHD-Vereins statt, zuletzt geprüft am 16.01.2018, <https://dig-hum.de/ag-datenzentren>.
- 3 Vgl. Patrick Sahle und Simone Kronenwett, „Jenseits der Daten: Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner 'Data Center for the Humanities',“ *LIBREAS. Library Ideas*, 23 (2013), <http://nbn-resolving.de/urn:nbn:de:kobv:11-100212726>.
- 4 Michael Knoche konstatiert in diesem Kontext einen Wildwuchs an unterschiedlichen Initiativen und das Fehlen klarer Vorstellungen wie Wissenschaftler/innen beim Thema Forschungsdaten von bibliothekarischer Seite unterstützt werden sollen. Vgl. Michael Knoche, *Die Idee der Bibliothek und ihre Zukunft* (Göttingen: Wallstein, 2017), 22.
- 5 Die höhere Standardisierung bei institutionell erhobenen Forschungsdaten, wie z.B. Digitalisaten, wird durch die standardbildenden Förderrichtlinien gestärkt. Seit 2009 empfiehlt die DFG einen fachspezifischeren Umgang mit Forschungsprimärdaten und deren Organisation, eine standardisierte Speicherung mit Rechteverwaltung, den freien Zugang, die Auszeichnung mit Metadaten, sowie Verfahren zur Qualitätssicherung, vgl. Deutsche Forschungsgemeinschaft, Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme, Unterausschuss für Informationsmanagement, „Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten,“ 2009, zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901.pdf; DFG, „Antragstellung,“ zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf; DFG, „Leitlinien zu den Forschungsdaten,“ 2015, zuletzt geprüft am 19.02.2018, http://www.dfg.de/foerderung/info_wissenschaft/2015/info_wissenschaft_15_66/; Schwerpunktinitiative "Digitale Information" der Allianz der deutschen Wissenschaftsorganisationen, „Grundsätze zum Umgang mit Forschungsdaten,“ 2010, <http://dx.doi.org/10.2312/ALLIANZOA.019>.

werden, inwieweit ein Bewusstsein für die Wichtigkeit von Daten und die Aufwendung eigener Zeit- und Arbeitsressourcen für deren Verarbeitung in den Geisteswissenschaften etabliert ist.⁶

Hinzu kommt die unklare infrastrukturelle Verantwortlichkeit für die Erwerbung, Präsentation, Pflege und Archivierung von Forschungsdaten. Das Label Forschungsdaten wird dabei aufgrund seiner Aktualität auf bereits etablierte Konzepte, wie z.B. auf Meta- oder Erschließungsdaten, übertragen oder inflationsartig für jede Art von Daten verwendet. Cum grano salis: Während für die einen Forschungsdaten per se auch Forschungsergebnisse sind, sind Forschungsdaten für andere lediglich relevante, aber nicht publikationsfähige Nebenprodukte ihres Forschungsvorhabens – ihre Aufwertung zu Forschungsdaten erscheint als Trugbild des digitalen Wandels. Entscheidend ist dabei immer die Frage nach dem Erkenntnisgewinn, der abgesehen von einem häufig theoretischen Diskurs über Nachnutzungsszenarien noch nicht beantwortet ist.⁷ Der vorliegende Beitrag versteht sich dabei nicht als Versuch, eine allgemein gültige Definition für den Begriff der Forschungsdaten zu liefern, sondern möchte vielmehr die Spezifika geisteswissenschaftlicher Forschungsdaten herausarbeiten und darauf aufbauend der Frage nachgehen, warum trotz der Relevanz von Forschungsdaten ein Spannungsverhältnis zwischen Theorie (Anforderungen) und Praxis (Verwendung) existiert.⁸

2. Kreation statt Generierung

Mit der zunehmenden Etablierung der Digital Humanities in den Geistes- und Kulturwissenschaften bilden bewusst als solche wahrgenommene und gezielt erzeugte Forschungsdaten eine neue Facette innerhalb der Geistes- und Kulturwissenschaften. Die auf eine Forschungsfrage zielorientierte Datengenerierung wird zunehmend als notwendiger Teil für die Erlangung von neuen Erkenntnissen definiert und rückt somit auch immer stärker in das Bewusstsein der Forschungscommunity.⁹ Zwar bildet die massenhafte Digitalisierung von Objekten des in den Gedächtnisinstitutionen aufbewahrten kulturellen Erbes weiterhin einen signifikanten Anteil der generierten Daten in den Geistes- und Kulturwissenschaften, aber digital innerhalb eines Forschungsprozesses erzeugte Daten (digital-born) ergänzen zunehmend das Portfolio der Datenerzeugung.¹⁰ Während digital-born data also forschungs-

6 Vgl. Christoph Schöch, „Big? Smart? Clean? Messy? Data in the Humanities,“ *Journal of Digital Humanities* 2, Nr. 3 (2013), zuletzt geprüft am 24.02.2018, <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>.

7 Vgl. Constanze Baum und Thomas Stäcker, „Die Digital Humanities im deutschsprachigen Raum. Methoden – Theorien – Projekte,“ in *Grenzen und Möglichkeiten der Digital Humanities*, hrsg. Constanze Baum und Thomas Stäcker, Sonderband der Zeitschrift für digitale Geisteswissenschaften 1 (Wolfenbüttel: Herzog August Bibliothek, 2015), http://dx.doi.org/10.17175/sb001_023.

8 Hubertus Kohle bezeichnet die derzeitige Publikationstätigkeit von Forschungsdaten in den Geisteswissenschaften als „sporadisch“. Vgl. Hubertus Kohle, „Digitales Publizieren,“ in *Digital Humanities: Eine Einführung*, hrsg. Fotis Jannidis, Hubertus Kohle und Malte Rehbein (Stuttgart: J. B. Metzler, 2017), 201.

9 In diesem Kontext kann auch der Titel „Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation“ der Jahrestagung 2015 des Vereins Digital Humanities im deutschsprachigen Raum gewertet werden. Die Tagung fand vom 23.02. bis zum 27.02.2015 an der Universität Graz statt. Vgl. Maik Fiedler und Andreas Weiß, „Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation. DHd-Jahrestagung 2015,“ *H-Soz-Kult* (Blog), 6. Juli 2015, zuletzt geprüft am 13.02.2018, <http://www.hsozkult.de/conferencereport/id/tagungsberichte-6059>.

10 Gemessen am Gesamtbestand ist die Anzahl an digitalisierten Quellen noch immer gering. Leider fehlt es an einer koordinierten nationalen Strategie für die Digitalisierung des gesamten kulturellen Erbes. Trotz übergreifender Projekte wie z. B. dem VD 17 stellen viele Digitalisierungsmaßnahmen Einzelvorhaben dar. Vgl. Lisa Klaffki, Stefan Schmunk und Thomas Stäcker, *Stand der Kulturgutdigitalisierung in Deutschland: Eine Analyse und Handlungsvorschläge*

getrieben (research-driven) erzeugt werden, sind vor allem die Surrogate von historischen Beständen von Bibliotheken und Archiven aus Gründen der nachhaltigen Erschließung produziert (curation-driven). Das aus den Naturwissenschaften bekannte Konzept der Kategorisierung der Daten nach Primär- und Sekundärdaten greift in den Geisteswissenschaften nicht, da auch innerhalb desselben Forschungsprozesses Sekundärdaten wieder zu Primärdaten werden; dies trifft vor allem auch für die Nachnutzung von Forschungsdaten zu.¹¹ Ebenfalls einen Unterschied zu den naturwissenschaftlichen Fächern bildet das Verhältnis von automatisierten Verfahren zu manuellen Erfassungs- und Erschließungsarbeiten. Trotz erheblicher Fortschritte in den Bereichen der Volltexterkennung, der automatisierten Erschließung und Annotation der Daten sind bei der Erhebung von hochqualitativen wissenschaftlichen Forschungsdaten in den Geisteswissenschaften nach wie vor größere manuelle Arbeitsanteile vonnöten. Zum einen ist dies durch die hochspezialisierten Fragestellungen bedingt, zum anderen durch die Unikalität und die Spezifika der historischen Überlieferung.¹² Die Mehrzahl der geisteswissenschaftlichen Projekte basiert nicht auf großen Datenmengen, sondern auf sehr kleinen und komplexen Datensammlungen, die die Historizität der zugrundeliegenden Quellen wiedergeben.¹³

So wurde im Projekt Frühneuzeitliche Gelehrtenbibliotheken des Forschungsverbunds Marbach Weimar Wolfenbüttel (MWW) die Erschließung der Bibliothek des Mathematikers und Chiliasten Benedikt Bahnsen über die Auswertung des Auktionskatalogs der Privatbibliothek durch eine weitestgehend manuelle Transkription der historischen Vorlage anhand eines definierten Datenmodells vorgenommen.¹⁴

des DARIAH-DE Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘, DARIAH-DE Working Papers Nr. 26 (Göttingen: DARIAH-DE, 2018), <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2018-1-3>; sowie das zugehörige Memorandum: Lisa Klaffki, „Memorandum: Digitalisierung von Kulturgut,“ *DHd-Blog*, 17. Mai 2018, zuletzt geprüft am 05.06.2018, <https://dhd-blog.org/?p=9849>.

- 11 Es wird daher empfohlen, in den Geisteswissenschaften keine Unterscheidung zwischen Primär- und Sekundärdaten vorzunehmen und allgemein von Forschungsdaten zu sprechen.
Vgl. Jens Klump, „Digitale Forschungsdaten,“ in *Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*, hrsg. Heike Neuroth et al., Version 2.3 (Göttingen: nestor, 2010), Kap. 17.10, 17:105, <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>; vgl. Johanna Puhl et al., *Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften*, DARIAH-DE Working Papers Nr. 11 (Göttingen: DARIAH-DE, 2015), 9, <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-4-4>.
- 12 Exemplarisch können die Probleme bei der Volltexterkennung von frühneuzeitlichen Leichenpredigten angeführt werden. Vgl. Maria Federbusch, „OCR für Drucke der Frühen Neuzeit? Erfahrungen und Perspektiven am Beispiel von Funeralschriften,“ *Bibliotheksdienst* 49, Nr. 7 (2015): 713–724; Thomas Stäcker, „Erfahrungsbericht Helmstedter Drucke Online an der Herzog August Bibliothek Wolfenbüttel,“ in *Volltext via OCR: Möglichkeiten und Grenzen. Test-szenarien zu den Funeralschriften der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz*, hrsg. Maria Federbusch und Christian Polzin, Beiträge aus der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz 43 (Berlin: Staatsbibliothek zu Berlin, 2013), 123–135.
- 13 Generell kann in den Geisteswissenschaften trotz der Massendigitalisierung nicht von Big Data gesprochen werden. Vgl. Schöch, „Data in the Humanities“.
- 14 Vgl. Hartmut Beyer et al., „Bibliotheken im Buch: Die Erschließung von privaten Büchersammlungen der Frühneuzeit über Auktionskataloge,“ in *Kodikologie und Paläographie im digitalen Zeitalter 4* (Codicology and Palaeography in the Digital Age), hrsg. Hannah Busch, Franz Fischer und Patrick Sahle (Norderstedt: Books on Demand, 2017), 43–70, <http://nbn-resolving.de/urn:nbn:de:hbz:38-77794>.

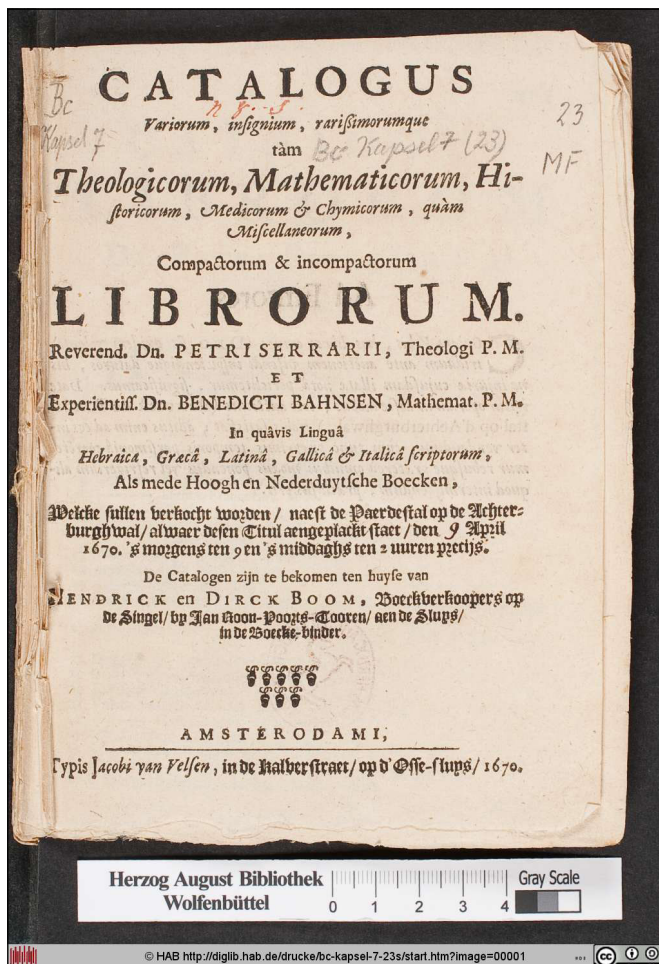


Abb. 1: Titelblatt des Auktionskatalogs der Bibliothek von Benedikt Bahnsen.

Ebenso erfolgte die Anreicherung der erhobenen biobibliographischen Daten mit Normdaten in gemeinsamer Arbeit von Wissenschaftler/in und Bibliothekar/in. Erst im Zuge der Auswertung und Analyse der Daten wurden vermehrt automatisierte Verfahren angewandt. Ferner wurden die Druckorte der verzeichneten Bücher mittels des DARIAH-DE Geobrowser in Relation zum Druckjahr visualisiert sowie ausgewählte Inhalte des Kataloges, wie z.B. Autoren, inhaltliche Kategorien oder Hauptsprachen, mittels unterschiedlicher Visualisierungsvarianten von Wortwolken bis Kreisdiagrammen abgebildet.¹⁵

15 Vgl. die Rekonstruktion der Bibliothek von Benedict Bahnsen, zuletzt geprüft am 19.02.2018, <http://dev.hab.de/auktionskataloge/bahnsen/bahnsen-catSubjectFormat.html>.

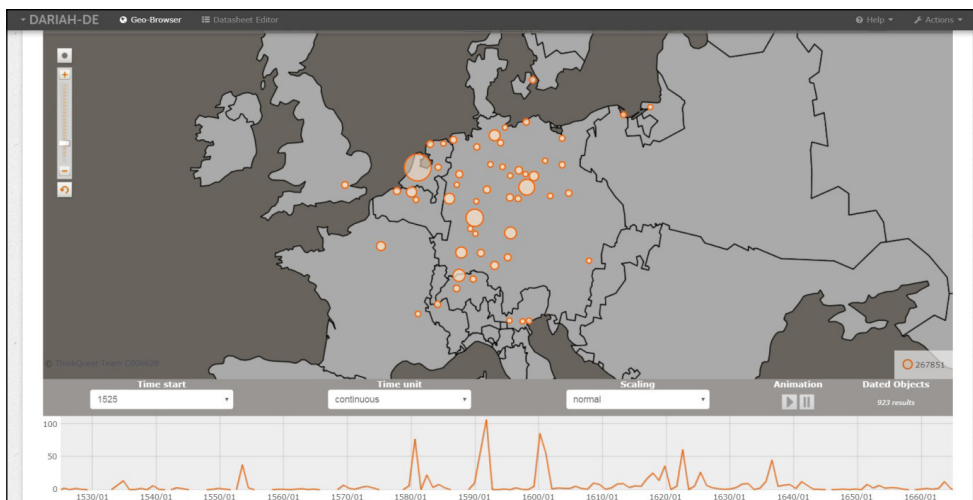


Abb. 2: Visualisierung der Druckorte / Druckjahre des Auktionskatalogs von Benedikt Bahsen im DARIAH-DE Geobrowser.

Das Beispiel zeigt, dass die Erstellung eines Forschungsdatensets eher eine domänenspezifischen Kreation denn ein automatisierter Datengenerierungsprozess ist. Gerade durch diesen Zuschnitt wird die Bedeutung von wissenschaftlichen Bibliotheken als Bereitsteller der Quellen, Kompetenzzentrum für deren Erschließung und als Anbieter für technische Lösungen – in summa die Bibliothek als traditionelle Infrastruktur für die Geisteswissenschaften – auch bzw. insbesondere im digitalen Forschen gestärkt.¹⁶

3. Unbehagen am Trend

Die Vielfalt und Heterogenität von geisteswissenschaftlichen Forschungsdaten korrespondiert zwangsläufig mit einer sehr weiten Definition, wie sie z.B. im DARIAH-DE Working Paper zum Research Data LifeCycle zu finden ist: „Unter digitalen geistes- und kulturwissenschaftlichen Forschungsdaten werden innerhalb von DARIAH-DE all jene Quellen/Materialien und Ergebnisse verstanden, die im Kontext einer geistes- und kulturwissenschaftlichen Forschungsfrage gesammelt, erzeugt, beschrieben und/oder ausgewertet werden und in maschinenlesbarer Form zum Zwecke der Archivierung, Zitierbarkeit und zur weiteren Verarbeitung aufbewahrt werden können.“¹⁷ Während die Funktionen des Sammelns, Beschreibens und Auswertens etablierte Vorgänge in den Geistes- und Kulturwissenschaften darstellen und keineswegs ein Spezifikum oder ein Resultat des digital turn sind, stellen die Aspekte der Zitierbarkeit, der Archivierung und des Teilens unter dem Paradigma der Maschinenlesbarkeit das Innovationspotenzial der Forschungsdaten heraus.¹⁸ Während die für den

16 Vgl. Petra Maier, *Digital Humanities und Bibliothek als Kooperationspartner*, DARIAH-DE Working Papers Nr. 19 (Göttingen: DARIAH-DE, 2016), 9, <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2016-5-6>.

17 Johanna Puhl et al., *Diskussion*, 14.

18 Vgl. Angelina Kraft, „Die FAIR Data Prinzipien für Forschungsdaten,“ TIB-Blog – Weblog der Technischen Informationsbibliothek (TIB), 12. September 2017, zuletzt geprüft am 24.01.2018, <https://blogs.tib.eu/wp/tib/2017/09/12/>

Erkenntnisgewinn notwendigen Zwischenprodukte, wie z.B. tabellarische Auswertungen, Arbeitsdatenbanken, Literaturlisten oder Exzerpte, nicht oder nur rudimentären Eingang in das schriftliche Ergebnis des Forschungsprozesses fanden, ist es nun technisch möglich, auch diese Zwischenprodukte zu publizieren und sie z.B. in Form einer enhanced publication mit dem Resultat zu verbinden.¹⁹ Die Publikation von Forschungsdaten würde den Forschungsweg transparenter und nachvollziehbarer machen, gleichzeitig das Anknüpfen an das Forschungsvorhaben erleichtern.

Trotz dieser Vorteile ist nicht abzusehen, ob sich dieser Trend abseits der Digital Humanities in den Geistes- und Kulturwissenschaften durchsetzt; scheint er für viele Geisteswissenschaftler/innen doch ihrem Verständnis von geisteswissenschaftlicher Forschung zu widersprechen: Angefangen bei dem ungewohnten Vorgang der zügigen Publikation von Zwischenergebnissen, der Transparenz des eigenen Forschungsprozesses sowie dem Unbehagen von technischen Abhängigkeiten bis zur Beurteilung der Datenqualität, der Reputation für die Publikation von Forschungsdaten bis hin zu der Bearbeitung durch Dritte im Zuge der Anreicherung und des Nachweises von Forschungsdaten. Dabei werden häufig die unterschiedlichen Qualitätsstufen von Datensets nicht ausreichend thematisiert bzw. gekennzeichnet, denn auch unstrukturierte und fehlerbehaftete Daten können für weitere Forschungsfragen von Relevanz sein bzw. als Grundlage für elaborierte Forschungsvorhaben dienen. Hier bestehen bereits Aktivitäten, wie z. B. die Kennzeichnung der Genauigkeit von OCR-Verfahren oder die Möglichkeit für Nutzer/innen, sich an der Optimierung von publizierten Texten zu beteiligen.²⁰ Gerade durch den Prozess des automatischen Indexierens werden Bibliotheken gefragt sein, sich in diesen Fragen zu positionieren.

4. Kosmos im Kontext

Die Ansprüche an geeignete Publikationsformate für die Repräsentation geisteswissenschaftlicher Daten sind hoch: Sie sollen die Heterogenität und Multimedialität, aber auch die Historizität sowie die Kontextualisierung der Untersuchungsobjekte adäquat darstellen können. In den Anfängen des „elektronischen Publizierens“ in den Geisteswissenschaften wurden vor allem die multimedialen und interaktiven Möglichkeiten der Webtechnologien als Verbesserung und Anreiz wahrgenommen.²¹ Als Resultat entstanden Webseiten, die Forschungsergebnisse und Untersuchungsgegenstände gleichzeitig repräsentieren und kontextualisieren können. Diese digitalen Forschungsdatensammlungen (*Thematic Research Collections*) ersetzen nicht klassische Publikationsformate wie Buch und Artikel,

[die-fair-data-prinzipien-fuer-forschungsdaten/](#).

- 19 Vgl. Andreas Degkwitz, „Enhanced Publications Exploit the Potential of Digital Media,“ in *Evolving Genes of ETDs for Knowledge Discovery. Proceedings of ETD 2015*, 18th International Symposium on Electronic Theses and Dissertations, 51-59, zuletzt geprüft am 19.02.2018, https://www.researchgate.net/profile/Andreas_Degkwitz/publication/283783359_Enhanced_Publications_Exploit_the_Potential_of_Digital_Media/links/5646fe2908ae9f9c13e8c4fa/Enhanced-Publications-Exploit-the-Potential-of-Digital-Media.pdf.
- 20 Vgl. das Tool DTAQ (Deutsches Textarchiv Qualitätssicherung); „Kollaborative Qualitätssicherung im Deutschen Textarchiv,“ zuletzt geprüft am 19.02.2018, <http://www.deutschestextarchiv.de/daq/about>.
- 21 Zu den Anfängen des digitalen Publizierens (dort noch elektronisches Publizieren genannt) in den Geisteswissenschaften ist das Positionspapier der DFG lesenswert: DFG, *Elektronisches Publizieren* (Bonn: DFG, März 2005), zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/programme/lis/pos_papier_elektron_publizieren_0504.pdf; hierzu auch die Empfehlungen des Unterausschusses für elektronisches Publizieren und des Ausschusses für wissenschaftliche Bibliotheken und Informationssysteme, Juni 2006, zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/programme/lis/elektr_publizieren.pdf.

sondern bilden eine eigenständige Publikationsform.²² Denn die Webtechnologien erlauben es, die Grundprinzipien geisteswissenschaftlicher Methodik – nach John Unsworth: Annotation, Vergleich, Referenz, Auswahl, Illustration und Repräsentation²³ – in unmittelbarer Form darzustellen, während dies in reinen Texten nur mittelbar durch eine sprachliche Beschreibung möglich ist.

Die Webseite mit zugehöriger Datenbank als wissenschaftliches Publikationsformat und komplexes Informationsobjekt zugleich impliziert, dass nicht nur die hinterlegten und präsentierten Daten, sondern die Präsentationsschicht und Softwarearchitektur selbst Informationsträger wissenschaftlicher Erkenntnis sind. Dies erschwert wesentlich die für eine langfristige Erhaltung notwendige Trennung zwischen der webbasierten Präsentations- und einer dateibasierten Archivierungsschicht.²⁴ Die komplexen Webseiten haben sich damit zu echten Sorgenkindern für die Informationsinfrastruktureinrichtungen entwickelt.²⁵ Diesen Herausforderungen lässt sich durch Investitionen in Forschungsdatenmanagement bei Forschenden, Förderern und Informationsinfrastruktureinrichtungen begegnen und hier sind auf allen Ebenen deutliche Anzeichen der Bewusstseinsbildung und Handlungswille zu konstatieren.²⁶

Die zahlreichen Anforderungen des modernen Forschungsdatenmanagement wurden bei der Online-Edition der "Teutschen Academie" von Joachim von Sandrart (sandrart.net) nahezu vorbildlich erfüllt: standardisierte Datenmodelle (XML/TEI, CIDOC-CRM), persistente Identifikatoren, Normdaten (GND, TGN, etc.), offene Schnittstellen und Webservices (REST-API, LOD/RDF, Beacon), offene Lizenzen (ODbL, CC-BY-NC-SA).²⁷ Die Textdaten liegen zudem in einem Repositorium.²⁸ Dennoch vermag die Datenschicht nicht, die komplexen Funktionalitäten des nutzerfreundlich und effizient gestalteten Webinterface zu erhalten: Register, facettierte Suche, Links und interaktive Verknüpfungen sowie die einzelnen Seiten zu Personen, Orten, Kunstwerken, bibliographische Nachweise und den Quellen, die als aufbereitete Dossiers das komplexe Beziehungsgeflecht in den Daten erst rezipierbar

22 Vgl. Carole L. Palmer, „Thematic Research Collections,“ in *A Companion to Digital Humanities*, hrsg. Susan Schreibman, Ray Siemens und John Unsworth (Oxford: Blackwell 2004), Kap. 24, zuletzt geprüft am 19.02.2018, <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-4-5&toc.depth=1&toc.id=ss1-4-5&brand=default>.

23 John Unsworth, *Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?* Vortrag, Humanities Computing: Formal Methods, Experimental Practice, King's College London, 13. Mai 2000, zuletzt geprüft am 19.02.2018, <http://www.iath.virginia.edu/~jmu2m/Kings.5-00/primitives.html>.

24 Wolfgang Pemppe, „Geisteswissenschaften,“ in *Langzeitarchivierung von Forschungsdaten: Eine Bestandsaufnahme*, hrsg. Heike Neuroth et al. (Göttingen: Universitätsverlag Göttingen / Boizenburg: Verlag Werner Hülsbusch, 2012), 141, <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>.

25 Vgl. hierzu Stefan Buddenbohm, Claudia Engelhardt und Ulrike Wuttke, „Angebotsgenese für ein geisteswissenschaftliches Forschungsdatenzentrum,“ *Zeitschrift für digitale Geisteswissenschaften* (2017), http://dx.doi.org/10.17175/2016_003.

26 Zu Maßnahmen der Infrastruktureinrichtungen siehe Anm. 2, aber auch aus der Wissenschaft werden Standards definiert. So hat etwa das Fachkollegium Literaturwissenschaft der DFG Förderkriterien aufgestellt, zuletzt geprüft am 19.02.2018, http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/foerderkriterien_editionen_literaturwissenschaft.pdf.

27 Joachim von Sandrart, *Teutsche Academie der Bau-, Bild- und Mahlerey-Künste, Nürnberg 1675-1680, wissenschaftlich kommentierte Online-Edition*, hrsg. Thomas Kirchner et al., 2008-2012, zuletzt geprüft am 19.02.2018, <http://ta.sandrart.net>.

28 Joachim von Sandrart, *L'Academia Todeasca. della Architectura, Scultura & Pittura: Oder Teutsche Academie der Edlen Bau- Bild- und Mahlerey-Künste. Bd. 1,1.* Nürnberg: 1675. Deutsches Textarchiv, <http://nbn-resolving.de/urn:nbn:de:kobv:b4-30378-3>.



Startseite

Die Teutsche Academie im Volltext
Gliederung der Teutschen Academie
Übersetzungen

Suche
Personen
Orte
Kunstwerke
Bibliographie

Editionsinformationen
Projektinformationen
Herausgeber
Kontakt/Impressum



Permanente URL:
<http://ta.sandrart.net/artwork-4842>

Mischwesen (TA 1680, Tafel CC, 5)

Basis-Daten

Datierung
1680

Material/Technik
Radierung

Signatur und Bezeichnung
ed Sandr:

loh Meyer fecit.

Abbildung(en) in der »Teutschen Academie«
[TA 1680, Iconologia Deorum, Tafel CC. \(nach S. 196\)](#)

Erwähnungen in der Teutschen Academie

»5. Chimaera, oder Amuletum.«
[TA 1680, Iconologia Deorum, S. 196](#)

»Das nachfolgende Stück soll ein Chimaera seyn: das ich aber viel eigentlicher ein Amuletum wider Zauberey und alles Böse/...«
[TA 1680, Iconologia Deorum, S. 196](#)

»... einige unversehene Schrecken Panici genennet worden/ so von diesem Gott sollen hergerühret haben. Diese Abbildung aber ist aus einem alten Niccolo-Stein abgesehen worden.«
[TA 1680, Iconologia Deorum, S. 197](#)

Kommentare

Als graphische Vorlage konnte ein Blatt in [Agostini, Gemme antiche](#) identifiziert werden, welches das Mischwesen gegenseitig abbildet.
► [Carolin Ott, 14.06.2011](#)

Wurde radiert von
– [Meyer, Johannes <der Jüngere>](#)

Kunstwerk-Bezüge [Symbolansicht](#)

Ist Teil von
– Meyer: [Tafel mit sechs Medaillons \(TA 1680, Tafel CC\)](#) *Graphik*

Steht in Zusammenhang mit
– Unbekannter antiker Künstler: [Gemme mit Darstellung eines Mischwesens](#) *Kunstgewerbe*

Geschaffen nach
– Galestruzzi: [Mischwesen 1 \(Agostini, Gemme antiche\)](#) *Graphik*

Abb. 3: Entitätsseite eines Kunstwerks in sandrart.net

machen (Abb. 3). Die Webseite funktioniert so als Benutzerschnittstelle, die nicht nur präsentiert, sondern als „virtuelle Forschungsplattform“ die Edition unmittelbar nutzbar macht. Die Webseite mit Oberfläche und Interface wird glücklicherweise von einer wissenschaftlichen Bibliothek betreut und gepflegt.²⁹ Die Projektverantwortlichen haben auf die Herausforderung der langfristigen Nutzbarkeit vorbildliche Antworten gefunden,³⁰ jedoch muss betont werden, dass das ehemalige DFG-Projekt

29 Die Herzog August Bibliothek Wolfenbüttel hat die Digitalisierung der Bände der „Teutschen Academie“ übernommen und sorgt für die Langzeitverfügbarkeit der Edition. „Kooperationen,“ zuletzt geprüft am 19.02.2018, <http://www.sandrart.net/de/kooperationen/>.

30 Patrick Sahle, „Review von ‚Sandrart.net,‘“ *RIDE* 1 (2014): Abs. 11, <https://dx.doi.org/10.18716/ride.a.1.5>.

technologische und informationswissenschaftliche Verfahren einsetzen konnte, die erst heute als zeitgemäße Standards gelten. In der Regel müssen sich Forschungsprojekte aber auf die verfügbaren Best Practices beschränken und können selten Ressourcen in technologische oder informationswissenschaftliche Innovationen investieren.

Sandrat.net demonstriert, dass die Verknüpfung von Texten und Objekten in der Präsentationsebene der Webseiten keineswegs das Resultat fehlender technischer und informationswissenschaftlicher Kompetenz sind, sondern vielmehr Ausdruck geisteswissenschaftlicher Methodik bei der durch Kontextualisierung komplexe Zusammenhänge analysiert werden. Weder lässt sich also eine einfache Trennung zwischen Präsentations- und Datenschicht etablieren; auch der Informations- und Nachnutzungsgehalt einer Datenschicht steht damit in Frage. Auswahl und Aufbereitung der Daten sind in den Geisteswissenschaften an spezifische Fragestellungen, bestimmte disziplinabhängige Methoden und theoretische Perspektiven gebunden. Eine unmittelbare Nachnutzung ist daher nicht ohne Weiteres möglich.³¹ Für nachnutzbare, maschinenlesbare Forschungsdaten bedarf es in den Geisteswissenschaften daher auch geeigneter Repräsentations- und Publikationsformate, die die notwendige Dokumentation und Kontextualisierung der Daten transportieren können. Die webbasierten *Thematic Research Collections* weisen hier den Weg: Sie sind nicht als Nachweis, sondern als nutzbare Ressource konzipiert und gestaltet.

5. Raum oder Wegweiser

Die thematische Vielfalt der Forschungsprojekte bedingt auch stets eine technische Vielfalt der erstellten Ressourcen, aus denen heterogene Datensammlungen entstehen.³² Momentan werden diese, wenn überhaupt, häufig separat von der „eigentlichen Publikation“ in verschiedene Repositorien hochgeladen. Als Repositorien werden derzeit verstärkt domänenspezifische³³ Lösungen umgesetzt, Publikationen zu verwandten Themen und mit einem ähnlichen Publikum befinden sich dann in unmittelbarer Nachbarschaft und können entdeckt werden – ein erster Schritt „vom Datencontainer zum Diskursraum“.³⁴ Für Forschungsdatenrepositorien lassen sich die Infrastrukturen

31 Vgl. Sahle und Kronenwett, *Jenseits der Daten*, 82.

32 Vgl. Tobias Gradl, Andreas Henrich und Christoph Plutte, „Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen“ in *Grenzen und Möglichkeiten der Digital Humanities*, hrsg. Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1 (Wolfenbüttel: Herzog August Bibliothek und Forschungsverbund Marbach Weimar Wolfenbüttel, 2015), http://nbn-resolving.de/10.17175/sb001_020.

33 Als eines der ältesten Publikationsplattformen und Volltextserver „Art Dok,“ Repositorium für die Kunstgeschichte (zuletzt geprüft am 13.02.2018, <http://archiv.ub.uni-heidelberg.de/artdok/>), seit wenigen Wochen online das „GenderOpen-Repository“ (zuletzt geprüft am 13.02.2018, <https://www.genderopen.de/>), am 1. Oktober 2018 startet „Archivum Medii Aevi Digitale (AMAD),“ ein Repositorium für Open Access-Publikationen aus der Mediävistik, Karoline Döring, „DFG-Projekt „Archivum Medii Aevi Digitale – Mediävistisches Fachrepositorium und Wissenschaftsblog für die Mittelalterforschung (AMAD);“ *Mittelalter. Interdisziplinäre Forschung und Rezeptionsgeschichte* (Blog), 6. Februar 2018, zuletzt geprüft am 19.02.2018, <http://mittelalter.hypotheses.org/12026>.

34 Hubertus Kohle, „Repositorien: vom Datencontainer zum Diskursraum,“ *arthistoricum.net* (Blog), 9. Dezember 2017, zuletzt geprüft am 13.02.2018, <https://blog.arthistoricum.net/beitrag/2017/12/09/repositorien-vom-datencontainer-zum-diskursraum/>.

grob in institutionsweite³⁵ oder übergreifende³⁶ Angebote einordnen. Im ersteren Falle erfolgt konsequenterweise auch der Nachweis über institutionsgebundene Verzeichnisse, etwa an der Universität Heidelberg als Teilkatalog der Hochschulbibliographie, in der Daten zu allen Fachgebieten verzeichnet sind.³⁷

Wie die Forschenden aber für die eigene Arbeit relevante Datensammlungen suchen und finden können, ohne jedes Repositoryum einzeln abzufragen, ist noch offen, denn es fehlt an zentralen Nachweisinstrumenten – es fehlt der Wegweiser in Form domänenspezifischer Kataloge, aber Räume wie die Publikationsrepositorien gibt es auch nicht. Existierende Ansätze sind etwa aggregierende Metalösungen wie re3data³⁸ als Register von Forschungsdatenrepositorien oder OpenAIRE³⁹ als Zugang zu den Ergebnissen aus EU-Projekten. Viele Vorhaben legen (Zwischen-)Ergebnisse und Dokumentation in GitHub⁴⁰ ab, ein (kommerzielles) Repositoryum, das hauptsächlich auf Code ausgerichtet ist, in dem aber auch Textdateien oder Bilder gespeichert werden können – die Auffindbarkeit ist hier ohne Wegweiser auch nicht gegeben. Ein noch im Aufbau befindliches Konzept zur Beschreibung und Verzeichnung ist die DARIAH-DE Collection Registry⁴¹, ein Katalog für Forschungsdatensammlungen. Diese sind mit DCDDM, dem an Dublin Core angelehnten DARIAH Collection Data Description Modell,⁴² standardisiert beschrieben und damit auch maschinell nutzbar.

Die Forschenden wählen das Repositoryum für die Datenpublikation selbst, insofern muss die Verzeichnung und damit die Voraussetzung für die Auffindbarkeit andernorts umgesetzt werden – die Kompetenz und Erfahrung hierfür ist in den Bibliotheken traditionell vorhanden.⁴³ Denn auch Datenpublikationen (z.B. in einem Repository publiziert, mit DOI und ORCID versehen) können katalogisiert werden, analog zu gedruckten Publikationen. Allerdings ist dies bei Daten komplexer, erfordert unter Umständen domänenspezifisches Wissen und ist (noch) ohne institutionelle Strukturen und ohne etablierte Workflows. Auffindbar würden diese Publikationen dann über die einschlägigen Kataloge oder Projekte wie die an der Bielefelder Universitätsbibliothek angesiedelte Suchmaschine BASE.⁴⁴ Doch die Sichtbarkeit und Zugänglichkeit von Forschungsdaten kann auch aktiv von den Forschenden

35 Wie z.B. das „Servicezentrum Forschungsdaten und -publikationen“ an der TU Berlin (Kooperation von UB, Rechenzentrum und Forschungsabteilung), zuletzt geprüft am 13.02.2018, www.szf.tu-berlin.de oder „heiDATA“, das Repositoryum des Competence Centre for Research Data der Universität Heidelberg, zuletzt geprüft am 13.02.2018, <https://heidata.uni-heidelberg.de/>.

36 Für die Geisteswissenschaften „DARIAH-DE Repository“, zuletzt geprüft am 18.02.2018, <https://de.dariah.eu/repository>; vgl. auch „User Guide, Das DARIAH-DE Repository und das TextGrid Repository“, vom 14.02.2018, zuletzt geprüft am 18.02.2018, <https://wiki.de.dariah.eu/display/publicde/Das+DARIAH-DE+Repository+und+das+TextGrid+Repository>; „CLARIN-D Repository“, zuletzt geprüft am 19.02.2018, <http://www.sfs.uni-tuebingen.de/ascl/clarin-center/repository.html>.

37 „Heidelberger Forschungsdatenverzeichnis“, zuletzt geprüft am 19.02.2018, <http://data.uni-heidelberg.de/datenverzeichnis.html>.

38 „Registry of Research Data Repositories“, zuletzt geprüft am 19.02.2018, <https://www.re3data.org/>.

39 „OpenAIRE“, zuletzt geprüft am 19.02.2018, <https://www.openaire.eu/>.

40 „GitHub“, zuletzt geprüft am 19.02.2018, <https://github.com/>.

41 „Informationen zur Collection Registry“, zuletzt geprüft am 19.02.2018, <https://de.dariah.eu/collection-registry>; „Collection Registry“, zuletzt geprüft am 19.02.2018, <https://colreg.de.dariah.eu/colreg-ui/>.

42 „DCDDM“, zuletzt geprüft am 19.02.2018, <https://github.com/DARIAH-DE/DCDDM>.

43 Roland Bertelmann und Hans Pfeiffenberger, „Forschungsdaten und Bibliotheken“, in *Praxishandbuch Bibliotheksmanagement*, hrsg. Rolf Griebel, Hildegard Schäffler und Konstanze Söllner (Berlin, Boston: De Gruyter Saur, 2014), 651-664.

44 „Bielefeld Academic Search Engine“, zuletzt geprüft am 19.02.2018, <https://www.base-search.net/>.

den selbst erhöht werden: Anders als in den Naturwissenschaften werden die Daten in der Literatur bisher häufig nicht zitiert, weshalb auch keine Notwendigkeit zu bestehen scheint, überhaupt welche zu publizieren.

6. Kulturwandel vor Technologie

Die technologischen Voraussetzungen wie Repositorien, Schnittstellen, Standardisierung sind weitgehend entwickelt. Die Technologie stellt damit nicht mehr die wesentliche Barriere dar.⁴⁵ Ob die Publikation von Forschungsdaten zu einer gelebten wissenschaftlichen Praxis wird, ist eng an Veränderungen des Wissenschaftssystems geknüpft. Die Grundvoraussetzung für das Teilen von Forschungsdaten ist die Bereitschaft der Forschenden. Dies erfordert jedoch einen Kulturwandel der zentralen Akteure des Wissenschaftsbetriebes. Während Technologien und Methoden weit fortgeschritten sind, fehlen Abläufe, Strukturen, Tradition, Kreditierung und nicht selten ein grundsätzliches Bewusstsein für die Relevanz von Forschungsdaten.⁴⁶ Diese Faktoren gelten fast ausnahmslos für alle Wissenschaften, aber für einen Kulturwandel müssen die gerade die disziplinspezifischen Eigenheiten berücksichtigt werden:

1. Wie hier ausgeführt wird, hat sich in den Fächern noch keine gemeinsame Begriffsdefinition für geisteswissenschaftliche Forschungsdaten herausgebildet und die datengestützte Forschung in Geisteswissenschaften ist selbst noch in der Entwicklungsphase. Das hat Auswirkungen auf das Reputationssystem, bei Evaluationen und Berufungen spielen Forschungsdaten keine Rolle.
2. Das traditionelle Verständnis wissenschaftlicher Nachvollziehbarkeit der Forschungsergebnisse basiert auf der Identifizierung der Quellen durch Verzeichnisse (Fußnotenapparat, Quellen- und Literaturverzeichnis), nicht durch die unmittelbare Bereitstellung der Quellen. Eine Nachnutzbarkeit ist (noch) nicht intendiert.⁴⁷
3. Anders als in naturwissenschaftlichen Feldern erhalten viele Forschungsdaten ihren Wert als Untersuchungsgegenstand über längere Zeit. Nicht die erste Entdeckung und Analyse, sondern die beste Interpretation wird als höchstes gewichtet.

Die Diskussion um die Publikation von Forschungsdaten wird vor allem von den Naturwissenschaften vorangetrieben. Spezifisch geisteswissenschaftliche Herausforderungen bleiben so häufig unberücksichtigt. Hier lassen sich Parallelen zur Open-Access-Bewegung identifizieren und darin liegt eine große Gefahr. In den Geisteswissenschaften finden sich viele Stimmen, die Open Access als

45 Obwohl es noch Bedarf an technologischer Entwicklung gibt, konstatiert daher eine Expertengruppe für die European Open Science Cloud: „The majority of the challenges to reach a functional EOSC are social rather than technical.“ The Commission High Level Expert Group on the European Open Science Cloud (EOSC), *Realising the European Open Science Cloud* (Luxembourg: Publications Office of the European Union, 2016), 6, <http://dx.doi.org/10.2777/940154>.

46 Eine gute Übersicht über die Barrieren in allen Wissenschaften liefert Christine Borgman, *Scholarship in the digital age: information, infrastructure, and the Internet* (Cambridge, London: MIT Press, 2007).

47 Siehe hierzu Andorfer, *Forschungsdaten*, 19.

übergestülptes und nicht passfähiges Konzept aus den Naturwissenschaften zurückweisen.⁴⁸ Die geisteswissenschaftlichen Fächer müssen daher eigene Konzepte für eine Publikationspraxis von Forschungsdaten entwickeln. Diese Konzeption steht vor der Herausforderung, die fachspezifischen Bedürfnisse insoweit zu berücksichtigen, dass sie mit den etablierten technologischen und informationswissenschaftlichen Standards umzusetzen sind. Idealerweise erfolgt dies gemeinsam mit den Informationsinfrastruktureinrichtungen, die die Forschungsdaten nach der Publikation bereitstellen und betreuen.

7. Kuration statt Friedhof

Die DFG empfiehlt als gute wissenschaftliche Praxis für die einer Publikation zugrunde liegenden Daten eine „Langzeitarchivierung“ von zehn Jahren.⁴⁹ Geisteswissenschaftliche Forschungsdaten sind tendenziell aber viel länger interessant oder bekommen, vom momentanen Standpunkt aus vielleicht gar nicht vorhersehbar, erst viel später eine erneute Relevanz. Dabei steht weniger die Reproduzierbarkeit der Ergebnisse im Vordergrund, sondern vielmehr die Nachnutzung der gesammelten (und bearbeiteten) Quellen durch andere Forschende sowie die langfristige Sicherung von hochwertigen Digitalisaten, Volltexten etc.

Bei der Erhaltung der Daten und dem Zugang zu ihnen sind verschiedene Aspekte zu unterscheiden, die alle ihrem jeweils spezifischen Alterungsprozess unterliegen und die Daten zu unterschiedlichen Zeitpunkten „sterben“ lassen können: Die (technische) Aktualität der Inhalte, der Erschließung und schließlich des technischen Systems. Um einem Datenfriedhof rechtzeitig entgegenzuwirken, ist die Kuration von Forschungsdaten zur Sicherstellung ihrer Unversehrtheit, Lesbarkeit und Interpretierbarkeit unabdingbar.⁵⁰

Da in den Geisteswissenschaften häufig projektbezogen, also mit begrenzter Zeit und limitierten Ressourcen geforscht wird, die Daten aber über das Projektende hinaus aufbewahrt und gepflegt werden müssen, entstehen sog. Humanities Data Center, die die Archivierung, Bereitstellung und Präsentation dieser Daten langfristig begleiten und übernehmen, wobei die Organisation und Zuständigkeit dieser Datenzentren sehr divers ausgestaltet ist.⁵¹

48 Übersichtlich referiert im Debattenbeitrag von Ben Kaden, „Publikationsfreiheit.de, Open Access und Geisteswissenschaften,“ *LIBREAS* (Blog), 31. Juli.2017, zuletzt geprüft am 13.02.2018, https://libreas.wordpress.com/2017/07/31/open-access_publikationsfreiheit/.

49 Deutsche Forschungsgemeinschaft (DFG), *Sicherung Guter Wissenschaftlicher Praxis*, Weinheim: Wiley-VCH, 2013, <http://dx.doi.org/10.1002/9783527679188.oth1>.

50 Michael Lautenschlager, „Institutionalisierte ‚Data Curation Services‘,“ in *Handbuch Forschungsdatenmanagement*, hrsg. Stephan Büttner, Hans-Christoph Hobohm und Lars Müller (Bad Honnef: BOCK + HERCHEN, 2011), 149, <http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2412>.

51 DHd AG Datenzentren, *Geisteswissenschaftliche Datenzentren im deutschsprachigen Raum. Grundsatzpapier zur Sicherung der langfristigen Verfügbarkeit von Forschungsdaten* (Hamburg: DHd e.V., 2017), <http://dx.doi.org/10.5281/zenodo.1134759>.

8. Lizenzen für Gesetzeshürden

Eine restriktive und komplexe Rechtslage kann für geisteswissenschaftliche Forschungsdaten als charakteristisch bezeichnet werden. Eine verlässliche Rechtslage ist für die Forschenden und publizierenden Einrichtungen jedoch ein entscheidender Faktor, der damit die Zugänglichkeit und die Verwertbarkeit der Forschungsdaten unmittelbar beeinflusst. Zum einen stehen Untersuchungsgegenstände wie zeitgenössische Werke unter urheberrechtlichem Schutz. Aber auch wenn ein Kulturgut selbst nicht mehr dem Urheberrecht unterliegt, kann dessen digitale Reproduktion geschützt und damit die Publikation dieser Forschungsdaten ohne Einigung mit dem Rechteinhaber nicht möglich sein. Der Gegenstand der Forschung kann auch weitere Rechtsbereiche berühren, etwa Persönlichkeitsrechte oder Datenschutz.⁵² Noch komplexer wird die Situation, wenn es sich – wie häufig in den Geisteswissenschaften – um eine Sammlung handelt, sei es eine Zusammenstellung ähnlicher Inhalte, beispielsweise mehrerer Handschriftendigitalisate, oder eine Sammlung im weiteren Sinne, die unterschiedliche Daten wie Digitalisate, Metadaten, eine Datenbank und Auswertungen enthält.⁵³

Die Lösung gegen unklare Rechtslage und Unsicherheit wäre, eine – wo möglich – freie Lizenz wie eine der Creative Commons-Lizenztypen⁵⁴ zu vergeben.⁵⁵ Somit wissen Forschende auf einen Blick, was sie unter welchen Bedingungen mit den Daten machen können und ob diese veröffentlicht werden dürfen. Dabei kommt es darauf an, was im Einzelnen lizenziert wird, ob die Lizenz die Daten selbst, die Datenbank an sich oder sogar die Metadaten und Beschreibung der Sammlung betrifft.

Ein bislang noch ungelöstes Problem sind Sperrfristen, die nach einer bestimmten Zeit ablaufen oder durch neue rechtliche Bedingungen verkürzt werden könnten.⁵⁶ Theoretisch könnte dann die Lizenz angepasst, der Zugang ermöglicht werden, doch hierfür gibt es keinen Automatismus, auch hier muss eine Form von Data Curation betrieben werden.

- 52 Auf die rechtlichen Bestimmungen kann an dieser Stelle nicht im Detail eingegangen werden. Einstiegspunkte für eine weitere Beschäftigung im Bereich der Geisteswissenschaften wären beispielsweise: Paul Klimpel und John H. Weitzmann, *Forschen in der digitalen Welt: Juristische Handreichung für die Geisteswissenschaften*. DARIAH-DE Working Papers Nr. 12 (Göttingen: DARIAH-DE, 2015), <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-5-0>; Nikolaos Beer et al., *Datenlizenzen für geisteswissenschaftliche Forschungsdaten. Rechtliche Bedingungen und Handlungsbedarf*. DARIAH-DE Working Papers Nr. 6 (Göttingen: DARIAH-DE, 2014), <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2014-4-8>; „Portal zu Forschungslizenzen,“ zuletzt geprüft am 16.02.2018, <http://forschungslizenzen.de/>.
- 53 Vgl. zu einer inklusiven Definition von „Forschungsdaten“ und „wissenschaftlicher Sammlung“ Jenny Oltersdorf und Stefan Schmunk, „Von Forschungsdaten und wissenschaftlichen Sammlungen. Zur Arbeit des Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘ in DARIAH-DE,“ *Bibliothek Forschung und Praxis* 40, Nr. 2 (2016): 181f., <http://dx.doi.org/10.1515/bfp-2016-0036>.
- 54 „Erläuterung zu Creative Commons,“ zuletzt geprüft am 16.02.2018, <https://de.creativecommons.org/index.php/was-ist-cc/>.
- 55 „Übersicht über freie und restriktive Lizenzmodelle,“ zuletzt geprüft am 16.02.2018, <http://www.ifross.org/lizenz-center/>.
- 56 Eric Steinhauer, „Recht als Risiko für das kulturelle Erbe,“ in *Mit gutem Recht erinnern*, hrsg. Paul Klimpel (Hamburg: Hamburg University Press, 2018), 134-136, <http://dx.doi.org/10.15460/HUP.178>.

9. Dilemma mit Potenzial

Sind die geschilderten Faktoren wie die Integration der Forschungsdaten in ein vertrauenswürdiges Repositorium, der Eintrag in einem wissenschaftlich relevanten Nachweissystem, die Vergabe einer freien Lizenz und die Anlage einer Dokumentation für eine Forschungsdatensammlung erfüllt, so stehen sie der Wissenschaft für die Nachnutzung in einer adäquaten Form zur Verfügung. Im idealen Fall werden die Forschungsdaten direkt mit der Publikation verbunden: So sind in einem bildanalytischen Aufsatz von Waltraud von Pippich die Messwerte und Metadaten, die der Untersuchung zugrunde liegen, über Links in das Forschungsdatenrepositorium der Universität München nachgewiesen.⁵⁷ Das Beispiel demonstriert auch die Wichtigkeit der Persistent Identifier für die wissenschaftliche Verwendung von Forschungsdaten und die Relevanz von nachhaltiger Infrastruktur.⁵⁸

Es ist offensichtlich, dass die Publikation von Forschungsdaten an inhaltliche und technische Kriterien geknüpft ist, die zu einem erhöhten Arbeitsaufwand für die Forschenden führen. Der Mehraufwand kann nur bedingt durch andere technische oder manuelle Unterstützung gemindert werden, da verschiedene Prozesse an das Fachwissen des/der Domänenspezialist/inn/en gebunden sind. Nichtsdestotrotz sollten die Forschenden sich nicht noch in neue technische Kontexte einarbeiten müssen, die für die Publikation der Forschungsdaten relevant sind. Gerade in diesem Kontext sollten wissenschaftliche Bibliotheken ein Angebot für ihre Community bereitstellen, welches zumindest die Beratung umfasst, falls keine eigene Infrastruktur zur Verfügung steht. Es ist auch vorstellbar, dass die Fachinformationsdienste hier verstärkt aktiv werden. Ein *do-ut-des*-Szenario würde in der Unterstützung der Aufbereitung, Publikation und langfristigen Verfügbarkeit der Forschungsdaten auf der einen und in dem Einbringen von Zeit und Fachwissen auf der anderen Seite liegen.

Bevor sich ein solches Konzept etablieren kann, müssen „neue Blicke durch die alten Löcher“⁵⁹ geworfen werden, um es mit Georg Christoph Lichtenberg zu sagen. Denn trotz der Heterogenität der Forschungsdaten in den Geisteswissenschaften handelt es sich bei vielen der in diesem Beitrag skizzierten Fragen und Aspekte nicht um unüberwindbare Hindernisse, die außerhalb eines paradigmatischen Forschungsprozesses liegen. Das Dilemma besteht in der Frage nach dem Verhältnis von Relevanz, Investition und Ertrag – sprich dem Bewusstsein für die Bedeutung von Forschungsdaten, dem Mehraufwand der Publikation und deren wissenschaftlicher Reputation. Diese Misere kann nicht einseitig gelöst werden, aber an den Lehrstühlen können Praktiken vorgelebt und vermittelt werden, an den Universitäten können Leistungen anerkannt und honoriert werden und bei den Fördergebern die Bereitstellung und Nachnutzung von Forschungsdaten in den Geisteswissenschaften gefördert

57 Vgl. Waltraud von Pippich, „Rot rechnen,“ in *Grenzen und Möglichkeiten der Digital Humanities*, hrsg. Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1 (Wolfenbüttel: Herzog August Bibliothek und Forschungsverbund Marbach Weimar Wolfenbüttel, 2015), http://dx.doi.org/10.17175/sb001_016. Zu den Forschungsdaten vgl. Waltraud von Pippich, *Frequenzen und statistische Dispersion der Farben in 50 Bildern von Adolph Menzel* (München: Open Data LMU, 2014), <http://dx.doi.org/10.5282/ubm/data.79>.

58 Vgl. Kathrin Schroeder, „Persistent Identifier (PI) – ein Überblick,“ in *Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*, hrsg. Heike Neuroth et al., Version 2.3 (Göttingen: nestor, 2010), Kap. 9.4, 9:22-45, <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>. Zu den infrastrukturellen Kriterien vgl. das DINI-Zertifikat, zuletzt geprüft am 23.01.2018, <https://dini.de/dini-zertifikat/>.

59 Georg Christoph Lichtenberg, *Sudelbuch F, 1776-1779, F 879*, zit. nach „Projekt Gutenberg-DE,“ zuletzt geprüft am 19.02.2018, <http://gutenberg.spiegel.de/buch/-6445/7>.

und gefordert werden. Geisteswissenschaftliche Infrastrukturen schließlich können die Voraussetzungen dafür schaffen, dass Forschende nicht noch eine technische *terra incognita* betreten müssen, wenn sie ihre Daten publizieren wollen.

Das Unterstützungsangebot der Informationsinfrastruktureinrichtungen nimmt dabei eine Sonderrolle ein, es muss bereits etabliert sein, wenn sich die anderen Akteure bewegen. Erst dann kann auch die Datenautorschaft die Rolle einnehmen, die ihr in einer zunehmend digital geprägten Forschungskultur gebührt. Denn Forschungsdaten sind eben kein von den Naturwissenschaften geborgter Begriff, sondern vielmehr essentieller Bestandteil einer auf digitalen Formaten basierenden geisteswissenschaftlichen Forschungslandschaft.

Abbildungsnachweis:

- Abb. 1: „Titelblatt des Auktionskatalog von Benedikt Bahnsen,“ Digitalisat der Herzog August Bibliothek Wolfenbüttel, zuletzt geprüft am 23.01.2018, <http://diglib.hab.de/drucke/bc-kapsel-7-23s/start.htm?image=00001>.
- Abb 2: „Visualisierung der Druckorte / Druckjahre des Auktionskataloges von Benedikt Bahnsen im DARIAH-DE Geobrowser,“ zuletzt geprüft am 01.02.2018, <https://geobrowser.de.dariah.eu/?csv1=http://geobrowser.de.dariah.eu./storage/267851¤tStatus=mapChanged=Historical+map+of+1650>.
- Abb. 3: „Mischwesen. Datenbankeintrag zu Mischwesen,“ (TA 1680, Tafel CC, 6), aus: Joachim von Sandrart. *Teutsche Academie der Bau-, Bild- und Mahlerey-Künste, Nürnberg 1675-1680, wissenschaftlich kommentierte Online-Edition*, hrsg. von Th. Kirchner et al., 2008-012, zuletzt geprüft am 19.02.2018, <http://ta.sandrart.net>.

Literaturverzeichnis

- Andorfer, Peter. *Forschungsdaten in den (digitalen) Geisteswissenschaften: Versuch einer Konkretisierung*. DARIAH-DE Working Papers 14. Göttingen: DARIAH-DE, 2015. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2>.
- Baum, Constanze und Thomas Stäcker. „Die Digital Humanities im deutschsprachigen Raum. Methoden – Theorien – Projekte.“ In *Grenzen und Möglichkeiten der Digital Humanities*, herausgegeben von Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. Wolfenbüttel: Herzog August Bibliothek und Forschungsverbund Marbach Weimar Wolfenbüttel, 2015. http://dx.doi.org/10.17175/sb001_023.
- Beer, Nikolaos, Kristin Herold, Maurice Heinrich, Wibke Kolbmann, Thomas Kollatz, Matteo Romanello, Sebastian Rose, Felix Falco Schäfer und Niels-Oliver Walkowski. *Datenlizenzen für geisteswissenschaftliche Forschungsdaten. Rechtliche Bedingungen und*

Handlungsbedarf. DARIAH-DE Working Papers Nr. 6. Göttingen: DARIAH-DE, 2014. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2014-4-8>.

- Bertelmann, Roland und Hans Pfeiffenberger. „Forschungsdaten und Bibliotheken.“ In *Praxishandbuch Bibliotheksmanagement*, herausgegeben von Rolf Griebel, Hildegard Schäffler und Konstanze Söllner, 651-664. Berlin, Boston: De Gruyter Saur, 2014.
- Beyer, Hartmut, Jörn Münkner, Katrin Schmidt und Timo Steyer. „Bibliotheken im Buch: Die Erschließung von privaten Büchersammlungen der Frühneuzeit über Auktionskataloge.“ In *Kodikologie und Paläographie im digitalen Zeitalter 4 (Codicology and Palaeography in the Digital Age)*, herausgegeben von Hannah Busch, Franz Fischer und Patrick Sahle, 43-70. Norderstedt: Books on Demand, 2017. <http://nbn-resolving.de/urn:nbn:de:hbz:38-77794>.
- Borgman, Christine. *Scholarship in the digital age: information, infrastructure, and the Internet*. Cambridge, London: MIT Press, 2007.
- Buddenbohm, Stefan, Claudia Engelhardt und Ulrike Wuttke. „Angebotsgenese für ein geisteswissenschaftliches Forschungsdatenzentrum.“ *Zeitschrift für digitale Geisteswissenschaften* (2017). http://dx.doi.org/10.17175/2016_003.
- EOSC, The Commission High Level Expert Group on the European Open Science Cloud. *Realising the European Open Science Cloud*. Luxembourg: Publications Office of the European Union, 2016. <http://dx.doi.org/10.2777/940154>.
- Degkwitz, Andreas. „Enhanced Publications Exploit the Potential of Digital Media.“ In *Evolving Genres of ETDs for Knowledge Discovery. Proceedings of ETD 2015*, 18th International Symposium on Electronic Theses and Dissertations. 51-59. Zuletzt geprüft am 19.02.2018, https://www.researchgate.net/profile/Andreas_Degkwitz/publication/283783359_Enhanced_Publications_Exploit_the_Potential_of_Digital_Media/links/5646fe2908ae9f9c13e8c4fa/Enhanced-Publications-Exploit-the-Potential-of-Digital-Media.pdf.
- DFG, Deutsche Forschungsgemeinschaft. *Sicherung Guter Wissenschaftlicher Praxis*. Weinheim: Wiley-VCH, 2013. <http://dx.doi.org/10.1002/9783527679188.oth1>.
- DHd AG Datenzentren. *Geisteswissenschaftliche Datenzentren im deutschsprachigen Raum. Grundsatzpapier zur Sicherung der langfristigen Verfügbarkeit von Forschungsdaten*. Hamburg: DHd e.V., 2017. <http://dx.doi.org/10.5281/zenodo.1134759>.
- Döring, Karoline. „DFG-Projekt ‚Archivum Medii Aevi Digitale – Mediävistisches Fachrepositorium und Wissenschaftsblog für die Mittelalterforschung (AMAD)‘.“ *Mittelalter. Interdisziplinäre Forschung und Rezeptionsgeschichte* (Blog). 6. Februar 2018. Zuletzt geprüft am 19.02.2018. <http://mittelalter.hypotheses.org/12026>.

- Federbusch, Maria. „OCR für Drucke der Frühen Neuzeit? Erfahrungen und Perspektiven am Beispiel von Funeralschriften.“ *Bibliotheksdienst* 49, Nr. 7 (2015): 713-724.
- Fiedler, Maik und Andreas Weiß. „Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation. DHd-Jahrestagung 2015.“ *H-Soz-Kult* (Blog). 6. Juli 2015. Zuletzt geprüft am 13.02.2018. <http://www.hsozkult.de/conferencereport/id/tagungsberichte-6059>.
- Gradl, Tobias, Andreas Henrich und Christoph Plutte. „Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen.“ In *Grenzen und Möglichkeiten der Digital Humanities*, herausgegeben von Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. Wolfenbüttel: Herzog August Bibliothek, 2015. http://nbn-resolving.de/10.17175/sb001_020.
- Kaden, Ben. „Publikationsfreiheit.de, Open Access und Geisteswissenschaften.“ *LIBREAS* (Blog), 31. Juli.2017. Zuletzt geprüft am 13.02.2018, https://libreas.wordpress.com/2017/07/31/open-access_publikationsfreiheit/.
- Klaffki, Lisa, Stefan Schmunk und Thomas Stäcker. *Stand der Kulturgutdigitalisierung in Deutschland: Eine Analyse und Handlungsvorschläge des DARIAH-DE Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘*. DARIAH-DE Working Papers Nr. 26. Göttingen: DARIAH-DE, 2018. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2018-1-3>.
- Klaffki, Lisa. „Memorandum: Digitalisierung von Kulturgut.“ *DHd-Blog*, 17. Mai 2018. Zuletzt geprüft am 05.06.2018, <https://dhd-blog.org/?p=9849>.
- Klimpel, Paul und John H. Weitzmann. *Forschen in der digitalen Welt. Juristische Handreichung für die Geisteswissenschaften*. DARIAH-DE Working Papers Nr. 12. Göttingen: DARIAH-DE, 2015. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-5-0>.
- Klump, Jens. „Digitale Forschungsdaten.“ In *Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*, herausgegeben von Heike Neuroth, Karsten Huth, Achim Oßwald, Regine Scheffel und Stefan Strathmann. Version 2.3, Kap. 17.10, 17:104-115. Göttingen: nestor, 2010. <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>.
- Kohle, Hubertus. „Digitales Publizieren.“ In *Digital Humanities: Eine Einführung*, herausgegeben von Fotis Jannidis, Hubertus Kohle und Malte Rehbein, 199-205. Stuttgart: J. B. Metzler, 2017.
- Kohle, Hubertus. „Repositorien: vom Datencontainer zum Diskursraum.“ *arthistoricum.net* (Blog), 9. Dezember 2017. Zuletzt geprüft am 13.2.2018. <https://blog.arthistoricum.net/beitrag/2017/12/09/repositorien-vom-datencontainer-zum-diskursraum/>.

- Knoche, Michael. *Die Idee der Bibliothek und ihre Zukunft*. Göttingen: Wallstein, 2017.
- Kraft, Angelina. „Die FAIR Data Prinzipien für Forschungsdaten.“ *TIB-Blog – Weblog der Technischen Informationsbibliothek (TIB)*, 12. September 2017. Zuletzt geprüft am 24.01.2018. <https://blogs.tib.eu/wp/tib/2017/09/12/die-fair-data-prinzipien-fuer-forschungsdaten/>.
- Lautenschlager, Michael. „Institutionalisierte ‚Data Curation Services‘.“ In *Handbuch Forschungsdatenmanagement*, herausgegeben von Stephan Büttner, Hans-Christoph Hobohm und Lars Müller, 149-156. Bad Honnef: BOCK + HERCHEN, 2011. <http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2412>.
- Lichtenberg, Georg Christoph. *Sudelbuch F, 1776–1779. F 879*. Zitiert nach „Projekt Gutenberg-DE.“ Zuletzt geprüft am 19.02.2018. <http://gutenberg.spiegel.de/buch/-6445/7>.
- Maier, Petra. *Digital Humanities und Bibliothek als Kooperationspartner*. DARIAH-DE Working Papers Nr. 19. Göttingen: DARIAH-DE, 2016. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2016-5-6>.
- Oltersdorf, Jenny und Stefan Schmunk. „Von Forschungsdaten und wissenschaftlichen Sammlungen: Zur Arbeit des Stakeholdergremiums ‚Wissenschaftliche Sammlungen‘ in DARIAH-DE.“ *Bibliothek Forschung und Praxis* 40, Nr. 2 (2016): 179-185. <http://dx.doi.org/10.1515/bfp-2016-0036>.
- Palmer, Carole L. „Thematic Research Collections.“ In *A Companion to Digital Humanities*, herausgegeben von Susan Schreibman, Ray Siemens und John Unsworth. Kap. 24. Oxford: Blackwell 2004. Zuletzt geprüft am 19.02.2018, <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-4-5&toc.depth=1&toc.id=ss1-4-5&brand=default>.
- Pompe, Wolfgang. „Geisteswissenschaften.“ In *Langzeitarchivierung von Forschungsdaten: Eine Bestandsaufnahme*, herausgegeben von Heike Neuroth, Stefan Strathmann, Achim Oßwald, Regine Scheffel, Jens Klump, Jens Ludwig, 137-159. Göttingen: Universitätsverlag Göttingen / Boizenburg: Verlag Werner Hülsbusch, 2012. <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>.
- Pippich, Waltraud von. *Frequenzen und statistische Dispersion der Farben in 50 Bildern von Adolph Menzel*. München: Open Data LMU, 2014. <http://dx.doi.org/10.5282/ubm/data.79>.
- Pippich, Waltraud von. „Rot rechnen.“ In *Grenzen und Möglichkeiten der Digital Humanities*, herausgegeben von Constanze Baum und Thomas Stäcker. Sonderband der Zeitschrift für digitale Geisteswissenschaften 1. Wolfenbüttel: Herzog August Bibliothek und

- Forschungsverbund Marbach Weimar Wolfenbüttel, 2015. http://dx.doi.org/10.17175/sb001_016.
- Puhl, Johanna, Peter Andorfer, Mareike Höckendorff, Stefan Schmunk, Juliane Stiller und Klaus Thoden. *Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften*. DARIAH-DE Working Papers Nr. 11. Göttingen: DARIAH-DE, 2015. <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-4-4>.
 - Sahle, Patrick. „Review von ‚Sandrart.net‘.“ *RIDE* 1 (2014): Abs. 11. <https://dx.doi.org/10.18716/ride.a.1.5>.
 - Sahle, Patrick und Simone Kronenwett. „Jenseits der Daten: Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner ‚Data Center for the Humanities‘.“ *LIBREAS. Library Ideas* 23 (2013): 76-96. <http://nbn-resolving.de/urn:nbn:de:kobv:11-100212726>.
 - Sandrart, Joachim von. *L'Academia Todesca. della Architectura, Scultura & Pittura: Oder Teutsche Academie der Edlen Bau- Bild- und Mahlerey-Künste. Bd. 1,1*. Nürnberg: 1675. Deutsches Textarchiv, <http://nbn-resolving.de/urn:nbn:de:kobv:b4-30378-3>.
 - Sandrart, Joachim von. *Teutsche Academie der Bau-, Bild- und Mahlerey-Künste, Nürnberg 1675–1680, wissenschaftlich kommentierte Online-Edition*, herausgegeben von Thomas Kirchner, Alessandro Nova, Carsten Blüm, Anna Schreurs und Thorsten Wübbena, 2008-2012. Zuletzt geprüft am 19.02.2018, <http://ta.sandrart.net>.
 - Schöch, Christoph. „Big? Smart? Clean? Messy? Data in the Humanities.“ *Journal of Digital Humanities* 2, Nr. 3 (2013). Zuletzt geprüft am 24.02.2018. <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>.
 - Schroeder, Kathrin. „Persistent Identifier (PI) – ein Überblick.“ In *Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*, herausgegeben von Heike Neuroth, Karsten Huth, Achim Oßwald, Regine Scheffel und Stefan Strathmann. Version 2.3, Kap. 9.4, 9:22-45. Göttingen: nestor, 2010. <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>.
 - Stäcker, Thomas. „Erfahrungsbericht Helmstedter Drucke Online an der Herzog August Bibliothek Wolfenbüttel.“ In *Volltext via OCR. Möglichkeiten und Grenzen. Testszenarien zu den Fundamentalschriften der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz*, herausgegeben von Maria Federbusch und Christian Polzin. Beiträge aus der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz 43, 123-135. Berlin: Staatsbibliothek zu Berlin, 2013).
 - Stäcker, Thomas. „Noch einmal: Was sind geisteswissenschaftliche Forschungsdaten?“ *DHd-Blog*, 6. Dezember 2015. Zuletzt geprüft am 16.01.2018. <http://dhd-blog.org/?p=5995>.

- Steinhauer, Eric. „Recht als Risiko für das kulturelle Erbe.“ In *Mit gutem Recht erinnern*. herausgegeben von Paul Klimpel, 125-136. Hamburg: Hamburg University Press, 2018. <http://dx.doi.org/10.15460/HUP.178>.
- Unsworth, John. *Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?* Vortrag, Humanities Computing: Formal Methods, Experimental Practice. King's College London, 13. Mai 2000. Zuletzt geprüft am 19.02.2018. <http://www.iath.virginia.edu/~jmu2m/Kings.5-00/primitives.html>.