

## Einleitung Teil III: Fortschreitende Digitalisierung im ABD- Bereich: neue Möglichkeiten und neue Herausforderungen

Ulrich Reimer

Die Digitalisierung schreitet auch im ABD-Bereich voran. Sie ermöglicht einerseits neue Lösungen, andererseits stellt sie uns vor neue Herausforderungen. Das Potenzial neuer Ansätze und deren Grenzen zeigt der Beitrag von Lynn Zimmermann auf. Sie beschreibt ihre Erfahrungen aus einem Pilotprojekt zur automatischen Erschliessung grosser Dokumentenbestände mit Hilfe von auf Data Science basierenden Ansätzen. Der Beitrag von Ursula Kähler wiederum erläutert, worum es beim Aufbau eines Repositoriums von nur noch digital vorliegender grauer Literatur geht. Zwei weitere Beiträge widmen sich Problemen, die aufgrund der Digitalisierung entstanden sind bzw. sich dadurch verschärft haben, und stellen Lösungen dafür vor, die ohne Digitalisierung wiederum nicht möglich wären: Angela Gastl-Hartmann befasst sich mit der Archivierung von Forschungsdaten, die im Rahmen von Open Science zunehmend wichtiger wird, während Janett Seewer sich mit der Archivierung von CAD-Plänen auseinandersetzt und u. a. damit konfrontiert ist, dass kein herstellerunabhängiges Standardformat existiert. Eine ganz grundlegende Frage behandelt Marcel Kuchler in seinem Beitrag. Man würde annehmen, es sei zumindest in der Gesetzgebung klar definiert, was unter Daten zu verstehen ist, doch zeigt sich, dass dies nicht der Fall ist. Da Daten die Basis für jegliche Digitalisierung sind und die Gesetze das legale Fundament dafür bilden, besteht hier Klärungsbedarf. Der Autor formuliert schliesslich eigene Gedanken für eine rechtlich sinnvolle Fassung des Datenbegriffs.

Die einzelnen Beiträge werden im Folgenden näher vorgestellt.

### Ansätze zur computergestützten Erschliessung von gleichförmigen Massenakten von Lynn Zimmermann

Die Erschliessung grosser Mengen analog vorliegender Akten kann Archive vor erhebliche Ressourcenprobleme stellen. Eine Unterstützung durch automatisierte Verfahren wäre eine willkommene Erleichterung. Basierend auf den Erfahrungen mit der Digitalisierung und Erschliessung einer Kartei über die Grossratsmitglieder des Kantons Thurgau untersucht der Beitrag von LYNN ZIMMERMANN, heute wissenschaftliche Archivarin im Staatsarchiv des Kantons Zürich, die Praxistauglichkeit bestehender computergestützter Verfahren. Spezielle Herausforderungen waren dabei

das nicht einheitliche Layout der Karteikarten, wodurch bestimmte Informationen nicht immer an derselben Stelle aufzufinden waren, sowie die Beschriftung der Karteikarten sowohl durch Schreibmaschine als auch mit handschriftlichen Bleistiftanmerkungen.

Die im Rahmen des Pilotprojekts automatisch zu erschliessende Information soll pro Karteikarte neben dem Namen des betreffenden Grossratsmitglieds auch Angaben zu seiner Herkunft und zur Zeitspanne, in welcher die Person im Grossen Rat sass, umfassen. Diese Angaben sollen grammatikalisch korrekte Phrasen für den Dossiertitel bilden und nicht einfach nur Wortaufzählungen sein.

Die für die Erschliessung eingesetzte Software basiert auf Verfahren des maschinellen Lernens, lernt somit ein Erkennungsmodell anhand einer Menge manuell transkribierter und erschlossener Karteikarten. Dieses Modell ist somit spezifisch an die zu leistende Aufgabe angepasst. Nach Durchlauf aller Prozessschritte wurden die Dossiertitel in ein Archivinformationssystem importiert. Es zeigte sich, dass einige Titel praktisch fehlerfrei erzeugt wurden, die Mehrzahl jedoch eine hohe Fehlerrate aufweist. Die Autorin analysiert die Gründe dafür und leitet aus den gewonnenen Erkenntnissen Vorschläge ab, wie man den Erschliessungsprozess – im konkret betrachteten Fall sowie ganz allgemein – besser angeht und dabei auch die zur Verfügung stehenden personellen und finanziellen Ressourcen berücksichtigt.

### Ein digitales Repositorium für graue Bernensia. Umfeldanalyse und praktische Empfehlungen für digitale Publikationen in der Bernensia-Sammlung an der Universitätsbibliothek Bern von Ursula Kähler

Der Anteil rein online erscheinender Publikationen steigt stetig, so auch im Bereich der grauen Literatur. Dies stellt Bibliotheken mit dem Auftrag, das dokumentarische Erbe eines geographisch festgelegten Raumes zu sichern, vor besondere Herausforderungen. Dazu zählen der langfristige Erhalt des ursprünglichen Erscheinungsbildes einer Publikation sowie die lückenlose Sammlung auch bei fehlender Pflichtabgabe. Der Artikel von URSULA KÄHLER, Filmwissenschaftlerin und frühere Mitarbeiterin der Fachstelle Historische Bestände ZHB der Universitätsbibliothek Bern, beleuchtet die Problematik und stellt die Erfahrungen vor, die im Rahmen eines Pilotprojekts für den Aufbau und Betrieb eines digitalen Repositoriums für graue Literatur an der Universitätsbibliothek Bern gesammelt wurden.

Die Autorin erörtert zunächst das aktuelle Bernensia Sammelkonzept der Universitätsbibliothek Bern, das zurzeit noch eine klare Ausrichtung auf Print-Medien hat, aber auf digitale Medien erweitert werden soll. Die Darstellung wird ergänzt um eine Beschreibung der bestehenden bzw. geplanten Sammelpraxis für digitale Regionalia und ihres Zugriffs über entsprechende Portale an einigen anderen ausgewählten Bibliotheken.

Schliesslich stellt der Beitrag ein Pilotprojekt für den Aufbau und Betrieb eines digitalen Repositoriums für graue Bernensia an der Universitätsbibliothek Bern vor. Die Autorin beschreibt, wie graue Bernensia überhaupt begrifflich zu fassen sind, welche Medien im Pilotprojekt berücksichtigt wurden, die Nutzung bereits vorhandener technischer Infrastruktur sowie das Management des Sammlungsprozesses.

### Bewertung und Archivierung von Forschungsdaten als neues Arbeitsgebiet für Hochschularchive von Angela Gastl-Hartmann

Das professionelle Management von Daten ist auch in Forschungsprojekten heutzutage eine Notwendigkeit. Dies betrifft alle Daten, die während eines Forschungsprojekts gesammelt, beobachtet oder erstellt wurden. Ein Teil dieser Forschungsdaten bleibt über das Projektende hinaus relevant und sollte der Wissenschaftsgemeinschaft zur Verfügung stehen, um die Forschungsergebnisse überprüfen und nachvollziehen, aber auch um für weitere Forschungsaktivitäten darauf aufbauen zu können. Damit gewinnt das Thema der Forschungsdatenarchivierung zunehmend an Wichtigkeit.

Der Beitrag von ANGELA GASTL-HARTMANN, wissenschaftliche Archivarin im Hochschularchiv der ETH Zürich, untersucht zunächst, inwieweit Hochschulbibliotheken und -archive diese Aufgabe schon übernehmen. Es zeigt sich, dass Bibliotheken vor allem im Rahmen von Open Access auch die Publikation von Forschungsdaten als Ergänzungen zum eigentlichen Artikel unterstützen. Der Übergang von Open Access zu Open Research Data bzw. Open Science ist hier fließend. Dagegen übernehmen die Hochschularchive zurzeit nur in wenigen Ausnahmefällen eine aktive Rolle in der Archivierung von Forschungsdaten.

Ausgehend vom beschriebenen Status Quo diskutiert die Autorin verschiedene Kooperationszenarien für die Forschungsdatenarchivierung und leitet daraus anschliessend konkrete Handlungsempfehlungen ab. Neben den Hochschulbibliotheken und -archiven kommen hier als Partner auch die Hochschulrechenzentren sowie schon etablierte, externe Fachrepositorien in Frage.

### Ein PDF, und fertig? Überlegungen zur digitalen Archivierung von CAD-Plänen von Janett Seewer

Die Planung von Bauprojekten erfolgt seit längerem mit Hilfe von Computer Aided Design (CAD), welches mannigfaltige Vorteile bietet. Eine Archivierung solcher CAD-Pläne kann sehr nützlich sein, wenn Gebäude zu einem späteren Zeitpunkt verändert werden sollen. Hierfür ist eine digitale Archivierung hilfreicher als das Papierformat, da auf einer digitalen Version idealerweise weitergearbeitet werden kann. Daneben kann eine Archivierung von Plänen bestimmter Gebäude auch zur Dokumentation von Technikgeschichte sinnvoll sein. Vor diesem Hintergrund geht der Beitrag von JANETT SEEWER, heute Leiterin des Archivs der Universität St. Gallen, der

Frage nach, welche digitalen Formate sich für die Archivierung von CAD-Plänen eignen.

Es würde nahe liegen, ein offenes, herstellerunabhängiges Format für die Archivierung zu verwenden. Da ein solches Standardformat bislang nicht existiert, müssen andere, möglichst gängige Formate gewählt werden. Die Autorin untersucht die aktuelle Praxis in verschiedenen Archiven und betrachtet die ganz unterschiedlichen Bedürfnisse einzelner Benutzergruppen, z.B. die interessierte Öffentlichkeit oder Mitarbeitende eines staatlichen Hochbauamtes.

Schliesslich analysiert die Autorin auf der Basis signifikanter Eigenschaften von CAD-Modellen verschiedene in Frage kommende Dateiformate im Hinblick auf ihre jeweilige Eignung und diskutiert ihre Vor- und Nachteile. Signifikante Eigenschaften sind z.B. die Speicherung geometrischer Volumen- und Flächenmodelle im Vektorformat oder die Darstellung von Oberflächeneigenschaften. Es bietet sich eine Kombination mehrerer Dateiformate zusätzlich zum Originalformat an, das man natürlich so lange wie möglich aufbewahren sollte.

### Daten als Gegenstand des Rechts von Marcel Küchler

Im Rahmen der Digitalisierung, die mittlerweile praktisch alle Gesellschaftsbereiche erreicht hat, spielen Daten die zentrale Rolle – einerseits als Rohstoff und Treiber der Digitalisierung, andererseits als Ergebnis entsprechender Prozesse. Als wichtige wirtschaftliche Ressource sind Daten naturgemäss auch Gegenstand rechtlicher Fragen. Der Beitrag von MARCEL KÜCHLER, heute wissenschaftlicher Bibliothekar am Sozialversicherungsgericht des Kantons Zürich, geht deshalb der Frage nach, ob es überhaupt einen rechtlich verbindlichen Datenbegriff gibt. Dazu betrachtet der Autor, welcher Datenbegriff bestehenden Gesetzen zugrunde liegt, und fahndet in den Rechtswissenschaften, der Informatik und der Informationswissenschaft nach einem einheitlichen Verständnis des Begriffs.

Es zeigt sich, dass der zunächst als klar fassbar erscheinende Begriff von Daten viele Facetten und Bedeutungsvarianten hat, die es zu berücksichtigen gilt. Beispielsweise verwendet das Datenschutzgesetz der Schweiz einerseits einen Datenbegriff, der ihre Speicherung auf einem Datenträger voraussetzt, andererseits wird aber auch von Daten ganz allgemein gesprochen, unabhängig davon, ob sie in einer gespeicherten Form vorliegen. Im letzteren Fall sind im Grunde eher Fakten oder Tatsachen gemeint. Die zum Teil gleichzeitige Verwendung des Begriffs Daten in unterschiedlichen Bedeutungsvarianten macht es schwer, ihn für den juristischen Kontext genauer zu fassen.

Auch ein Blick in die Rechtswissenschaft hilft bei der Klärung des Datenbegriffs nicht weiter. In neuerer Zeit entstandene Definitionen zeigen ein recht unterschiedliches Verständnis, z.B. darüber, ob Daten maschinenlesbar auf Datenträgern

vorliegen müssen und ob sie nur in ihrer digitalisierten Form als Daten anzusehen sind. Wenig zielführend sind Definitionen, die andere, ebenso ungeklärte Begriffe heranziehen, z.B. indem Daten als bestimmte Arten von Information definiert werden. Schliesslich untersucht der Autor, inwieweit die Disziplinen der Informatik und Informationswissenschaft einen einheitlichen Datenbegriff kennen, wird jedoch auch dort nicht fündig, wenngleich sich vereinzelt klare Abgrenzungen der Begriffe Daten, Information und Wissen finden.

Aufbauend auf den zusammengetragenen Erkenntnissen erweitert und vertieft Marcel Küchler die Diskussion und formuliert schliesslich Aspekte, die aus seiner Sicht für eine rechtlich sinnvolle Fassung des Datenbegriffs zu berücksichtigen sind.