

Forschungsdaten in der Geschichtswissenschaft

Veranstalter: Peter E. Fäßler, Arbeitsbereich Zeitgeschichte, Historisches Institut, Universität Paderborn; Torsten Hiltmann, Center for Digital Humanities, Westfälische Wilhelms-Universität Münster; Katrin Moeller, Historisches Datenzentrum Sachsen-Anhalt (Hist-Data), Martin-Luther-Universität Sachsen-Anhalt; Martin Dröge, Arbeitsgemeinschaft ‚Digitale Geschichtswissenschaft‘ im Verband der Historiker und Historikerinnen Deutschlands (VHD) / Heinz-Nixdorf-MuseumsForum; Landesinitiative NFDI der Digitalen Hochschule NRW
Datum, Ort: 07.06.2018–08.06.2018, Paderborn

Bericht von: Sven Siemon, Historisches Institut, Universität Paderborn

Die voranschreitende Digitalisierung verändert unser Arbeits- und Privatleben unaufhaltsam. Auch die Wissenschaft befindet sich in einem grundlegenden Transformationsprozess. Dies gilt auch für die Geschichtswissenschaft, die hier jedoch noch eher am Anfang steht. So sind auch in der historischen Forschung Projektförderungen durch Drittmittelgeber immer häufiger mit der Maßgabe verbunden, den Umgang mit den Forschungsdaten zu erläutern und einen Datenmanagementplan zu Beginn eines Projekts zu erstellen. HistorikerInnen stehen hierbei jedoch vor nicht alltäglichen Fragen: Was sind eigentlich Forschungsdaten? Wie können diese Daten organisiert, gespeichert, veröffentlicht und damit schließlich auch nachhaltig nutzbar gemacht werden? Welche Standards sind hier zu verwenden, welche müssen noch entwickelt werden?

Die Konsequenzen des Digitalisierungsprozesses in der Geschichtswissenschaft und die sich daraus für das Fach ergebende Relevanz des Themas „Forschungsdaten“ unterstrich TORSTEN HILTMANN (Münster) in seiner Einführung zur Tagung. Dabei zeigte er auf, wie der fortschreitende Medienwandel schließlich zu einem Methodenwandel führen wird, wenn die digitalisierten Quellen nicht mehr nur als Substitut ihrer analogen Vorlage, sondern auch als das wahrgenommen wer-

den, was sie tatsächlich sind: nämlich als Daten. Als solche können sie nicht mehr nur mit hermeneutischen, sondern auch mit digitalen Methoden bearbeitet werden. Dafür aber, so hob er hervor, braucht es einen einheitlichen Umgang mit den Daten und eine Verständigung über gemeinsame Standards.

Die Bedeutung von Forschungsdaten als wesentlichen Grundpfeiler wissenschaftlicher Erkenntnis in der modernen Geschichtswissenschaft unterstrich auch SILKE SCHWANDT (Bielefeld) in ihrem Vortrag. Die Besonderheit digitaler Forschungsdaten bestehe neben ihrer Vielfältigkeit auch darin, dass sie während des Forschungsprozesses entstehen. Aus diesem Grund müsse ein professionalisiertes Forschungsdatenmanagement über den gesamten Daten-Lebenszyklus hinweg gedacht und in enger Kommunikation zwischen den einzelnen Kooperationspartnern abgestimmt werden. Wie dies in der Praxis aussehen kann, zeigte Schwandt anhand ausgewählter Beispiele.

Die von beiden Rednern bereits erwähnte Vielfältigkeit geschichtswissenschaftlicher Forschungsdaten war Thema des ersten Teils der Tagung. MATTI STÖHR (Dresden) nahm sich der Texte an und erörterte in seinem Vortrag beispielhaft, wie historische Textquellen in enger Kooperation zwischen Infrastruktureinrichtungen, Forschenden und Lehrenden effektiv erschlossen werden können. Er resümierte, dass die Verfahren, Tools und Infrastrukturen prinzipiell vorhanden sind, deren Kenntnis und die Fähigkeit zu ihrer Nutzung jedoch intensiviert werden müsste.

Die Spezifika von „digital born sources“ zeigte OLIVER KIECHLE (Düsseldorf) am Beispiel von „Usenet“-Daten auf. Im Gegensatz zu analogen Quellen sei die Unterscheidung zwischen Original und Kopie bei den digital entstandenen Daten technisch kaum möglich. Zudem würden die „digital born sources“ einer extrem hohen Dynamik unterliegen, die nur schwer abgebildet werden könne. Die existierende Software zur Sammlung und Auswertung der Datensätze sei hierzu nur bedingt geeignet. Kiechle forderte damit eine engere Kooperation zwischen der Informatik und der Geschichtswissenschaft ein.

Sodann zeigte CORD PAGENSTECHER (Berlin) anhand von Zeitzeugen-Interviews die rechtlichen Herausforderungen auf, die mit jüngeren Forschungsdaten verbunden sind, und demonstrierte die Möglichkeiten der digitalen Verarbeitung und Präsentation audiovisueller Quellen.

CHRISTIAN LOTZ (Marburg) schließlich besprach die Frage der Forschungsdaten im Bereich der historischen Kartographie und wies hier insbesondere auf die mit der Digitalisierung verbundenen Möglichkeiten hin. So könne im Digitalen auf Vereinheitlichungen und Vereinfachungen in der historischen Geografie und Kartografie zugunsten von Mehrsprachigkeit und Multiperspektivität verzichtet werden: Die Reduzierung von Komplexität bei der kartografischen Darstellung von historischen Geografie-Forschungsdaten bei analogen Karten stelle eine Art „Altlast“ dar, die durch die Möglichkeiten der Digitalisierung, die flexible und dynamische Darstellungsformen bereithalte, überholt sei.

In Anbetracht der Vielfalt an Forschungsdaten und dem Umstand, dass im Forschungsprozess weitere Forschungsdaten entstehen, ist ein langfristiges und nachhaltiges Forschungsdatenmanagement von zentraler Bedeutung. Entsprechend war der zweite Teil der Tagung dem Forschungsdatenmanagement und den Datenzentren selbst gewidmet.

MARINA LEMAIRE (Trier) stellte anhand der virtuellen Forschungsumgebung für die Geistes- und Sozialwissenschaften (FUD) des Servicezentrum eSciences der Universität Trier einen beispielhaften Workflow vor. Sie unterstrich, dass die Quellen und Methoden des jeweiligen Forschungsprojekts die Grundlage für die Entwicklung eines geeigneten Datenmodells bilden sollten. Obwohl HistorikerInnen bereits das passende Handwerkszeug wie die kritische Auseinandersetzung mit Quellen oder das Rekonstruieren komplexer System mitbrächten, sei die Erstellung und Planung eines Datenmanagementplans jedoch noch immer eine Hürde für viele Forschende.

MORITZ KURZWEIL (Leipzig) machte mit seinem Einblick in die außeruniversitäre Forschung Mut, denn das Bewusstsein für Forschungsdatenmanagement steige bei den Forschenden. Forschungsdatenmanagement sei

nicht unilateral, sondern finde stets zwischen mehreren Akteuren wie zum Beispiel Forschenden und Infrastruktureinrichtungen statt und unterstrich somit die kooperative Dimension des Forschungsdatenmanagements.

LARS MÜLLER (Berlin) zeigte mögliche Probleme durch divergierende Sichten und Anforderungen von Forschung und Archiven oder Bibliotheken bei der Verarbeitung und Archivierung von Forschungsdaten auf und wies darauf hin, dass die Herausforderung nur in einem ständigen Kommunikationsprozess gemeistert werden kann.

Die gewonnenen Forschungsdaten sollten zur nachhaltigen Sicherung und Nutzbarkeit in einem Forschungsdatenrepositorium abgelegt werden. FABIAN CREMER (Bonn) stellte in seinem Vortrag das DARIAH-DE Repositorium vor. JULIAN SCHULZ (München) gab einen Einblick in das Modellvorhaben „Forschungsdatenmanagement Bayern“, das den uneinheitlichen Umgang mit Forschungsdaten anhand von „best practice“-Modellen vereinheitlichen will. Ziel des Projekts sei ein schnittstellenkompatibles System mit zahlreichen Vernetzungsmöglichkeiten. Damit spricht das bayerische Projekt eine Herausforderung an die Datenrepositorien an: Dezentrale Datenbanken sind oftmals nicht interoperabel. So war denn auch der dritte Abschnitt der Tagung den Standards und der Standardentwicklung gewidmet.

Eine weit fortgeschrittene Lösung stellte ALBAN FREI (Bern) anhand des schweizerischen histHub vor, das Datensätze aus unterschiedlichen Repositorien mit Hilfe von „Linked Data“ und dem „Semantic Web“ vernetzt. Der Referent zeigte auf, wie das Projekt histHub vorgeht, um über verschiedene Textrepositorien zu einer Normierung und Vernetzung der jeweils enthaltenen Daten zu gelangen. Dabei werden mit Hilfe von Ontologien unterschiedliche Datenmodelle und Datenbanken in Bezug gesetzt, während die Normierung dezentral auf Ebene der Verbindung zwischen den einzelnen Belegen in den jeweiligen Repositorien erfolgt.

Deutlich wurde, dass die Interoperabilität von Forschungsdaten aus unterschiedlichen Repositorien nur dann gewährleistet werden

kann, wenn die Forschungsdaten ein Mindestmaß an Einheitlichkeit beziehungsweise Normierung bieten. Hierbei macht der Forschende einen Spagat zwischen Anschlussfähigkeit, Interoperabilität und den eigenen Projektanforderungen, wie BIANCA PRAMANN (Braunschweig) anhand von Metadaten für historische Schulbücher demonstrierte.

Welchen Einfluss wiederum unterschiedliche Strategien von Normvokabularen zur Beschreibung von Forschungsdaten auf die Trefferzahl und Präzision von Suchanfragen in Datenbanken haben, zeigte JOHANNES BRACHT (Marburg) in seinem Beitrag am Beispiel eines Projekts zur „Strategischen Weiterentwicklung des Forschungsdatenmanagements am Herder-Institut für historische Ostmitteleuropa-Forschung“ auf.

Ohne Zweifel werden zukünftig mehr und mehr digitale Forschungsdatensätze veröffentlicht. Doch wie müssen mögliche Rahmenbedingungen der Datenautorenschaft und des Datasharing geschaffen sein? ANNA NEOVESKY (Mainz) erläuterte, dass die Anerkennung von Datenpublikationen als Forschungsleistung Anreize liefern kann, die erhobenen Daten zu teilen. Die Impulse hierfür müssten aus den einzelnen Fachcommunities kommen.

Die verschiedenen Möglichkeiten der Unterscheidung zwischen Text- und Datenautorenschaft diskutierte KATRIN MOELLER (Halle / Saale) in ihrem Vortrag. Sie zeigte auf, dass eine funktionale Trennung zwischen Textautoren und Datenautoren von Nöten sei. Aufgrund ihrer fluiden Eigenschaft als Original oder Kopie könnten dabei bei genuin digitalen Quellen andere Modelle der Datenautorenschaft notwendig sein.

Der abschließende Beitrag von ANIA LÓPEZ (Duisburg / Essen) von der nordrhein-westfälischen Landesinitiative NFDI schließlich gab Gelegenheiten, die Themen und Diskussionen der Tagung in die allgemeinen Entwicklungen und vor dem Hintergrund des geplanten Aufbaus einer „Nationalen Forschungsdateninfrastruktur“ (NFDI) einzuordnen. Eine verlässliche Forschungsdateninfrastruktur sei der Schlüssel für die Verbreitung digitaler Methoden, denn sonst sei stets die Sorge der Nachhaltigkeit

und Nutzbarkeit ein Hindernis für die Forschung.

Torsten Hiltmann leitete mit einem strukturierten Input die Abschlussdiskussion ein, was zu einem intensiven, fast einstündigen Austausch führte. Die Fragen vom Anfang der Tagung vor dem Hintergrund der verschiedenen Beiträge wieder aufnehmend, wurde darüber diskutiert, was nun tatsächlich unter historischen Forschungsdaten zu greifen ist, welche Kategorien hier unterscheidbar sind und woran das spezifisch historische dieser Daten liegt. Dabei wurde ebenso über die engen Verknüpfungen zu den anderen Fächern gesprochen wie auch über die Rolle der die Quellen bewahrenden und diese damit auch als Daten aufbereitenden Institutionen.

Wie geht es in Zukunft in einer mehr und mehr digitalisierten Geschichtswissenschaft weiter? Die Tagung zeigte auf, dass sich die Geschichtswissenschaft in einem Transformationsprozess befindet und klären muss, wie mit digitalen Forschungsdaten umzugehen ist. Zahlreiche Initiativen und Projekte haben hier bereits Meilensteine gesetzt. Neben der Vorstellung des „status quo“ wurden bei der Tagung weitere Schritte zum Umgang mit digitalen, historischen Forschungsdaten zur Etablierung digitaler Forschungsumgebungen in den Geschichtswissenschaften diskutiert. Die Entwicklung geeigneter Methoden, Metadatenschemata und Workflows kann nur durch Sensibilisierung, Kommunikation und Austausch der einzelnen Fachbereiche, Fachcommunities und Infrastruktureinrichtungen vorangetrieben werden. Erst dann kann das Ziel erreicht werden, Forschungsdaten nachhaltig, interoperabel, maschinenlesbar und universell auffindbar abzuspeichern und bereitzustellen.

Die Präsentationen zu den einzelnen Beiträgen sind auf dem Blog der AG „Digitale Geschichtswissenschaften“ unter der Adresse <https://digigw.hypotheses.org/2265> einsehbar.

Konferenzübersicht:

Martin Dröge (Paderborn) / Torsten Hiltmann (Münster): Begrüßung und Einführung
Silke Schwandt (Bielefeld): Quellen, Daten, Interpretationen. Heterogene Forschungsda-

ten und ihre Publikation als Herausforderung in der Geschichtswissenschaft

Sektion 1: Forschungsdaten in der Geschichtswissenschaft

Matti Stöhr (Dresden): Texte als Forschungsdaten

Oliver Kiechle (Düsseldorf): Der Umgang mit historischen Daten aus Webarchiven. Das Beispiel Usenet

Cord Pagenstecher (Berlin): Zeitzeugen-Interviews als Forschungsdaten

Christian Lotz (Marburg): Forschungsdaten in der historischen Geografie und Kartografie

Sektion 2: Forschungsdatenmanagement und Datenzentren

Marina Lemaire (Trier): Von Primärquellen zu Forschungsdaten. Entwicklung digitaler Konzepte zur Integration geschichtswissenschaftlicher Arbeits- und Datenmanagementprozesse

Moritz Kurzweil (Leipzig): Nur nicht allein auf die Insel - Forschungsdatenmanagement in der außeruniversitären Forschung

Lars Müller / Stefan Cramme / Sabine Reh (Berlin): Forschungsdatenmanagement als Kommunikationsprozess: neue Ansätze in einem Projekt zur Erforschung historischer Abituraufsätze

Fabian Cremer (Bonn): Nutzungserfahrungen des DARIAH-DE Repository

Julian Schulz (München): Das Modellvorhaben „Forschungsdatenmanagement Bayern“ aus Sicht der Geisteswissenschaften: Erfahrungen - Probleme – Perspektiven

Sektion 3: Datenstandards

Alban Frei (Bern): Historische Forschungsdaten vernetzen und normieren. histHub - ein Beispiel aus der Schweiz

Bianca Pramann (Braunschweig): Datenstandards für historische Schulbuchquellen. Wege der Integration in CLARIN-D

Johannes Bracht / Stephanie Palek (Marburg): So einfach wie möglich, so kompliziert wie nötig? Fachliche Normvokabulare in der Beschreibung von Forschungsdaten aus der

Sicht der Forschenden am Beispiel der historischen Ostmitteleuropa-Forschung

Sektion 4: Datenautorenschaft

Anna Neovesky (Mainz): Autorschaft von Daten und Forschungssoftware - Möglichkeiten der Sichtbarmachung und Etablierung in der Geschichtswissenschaft

Katrin Moeller (Halle / Saale): Datenautorenschaft und Forschungsdatenmanagement

Impulsreferat zur Abschlussdiskussion

Ania López (Duisburg / Essen): Forschungsdaten in der Geschichtswissenschaft. Diskussionsstand und Herausforderungen vor dem Hintergrund der Initiative zum Aufbau einer Nationalen Forschungsdateninfrastruktur (NFDI)

Abschlussdiskussion

Tagungsbericht *Forschungsdaten in der Geschichtswissenschaft*. 07.06.2018–08.06.2018, Paderborn, in: H-Soz-Kult 15.09.2018.