

HT 2018: Forschungsdaten: rechtliche Herausforderungen und wissenschaftliche Reputation. Forschungsdatenmanagement als Bestandteil einer neuen Wissenschaftskultur

Veranstalter: Verband der Historiker und Historikerinnen Deutschlands (VHD); Verband der Geschichtslehrer Deutschlands (VGD)

Datum, Ort: 25.09.2018–28.09.2018, Münster

Bericht von: Andreas Müller / Anne Pürschwitz, Institut für Geschichte, Martin-Luther-Universität Halle-Wittenberg

Die Digitalisierung stellt Historiker/innen nicht nur vor zahlreiche Herausforderungen, sondern wirkt sich auch auf die Wissenschaftskultur aus. Der Umgang mit Forschungsdaten im digitalen Zeitalter gehört momentan zu den besonders intensiv diskutierten wissenschaftlichen Arbeitsfeldern, und löst auch in der Geschichtswissenschaft zahlreiche Kontroversen aus.

Zur Einführung in das Thema skizzierte PATRICK SAHLE (Köln) in seinem Vortrag die Besonderheiten geisteswissenschaftlicher Forschungsdaten und sprach sich für einen auf den Forschungskontext bezogenen Forschungsdatenbegriff aus. Dabei plädierte er für einen erweiterten Datenbegriff, der letztlich über die gängige sachliche wie institutionelle Trennung von Quellen – Daten – Forschungsergebnissen hinausragt und skizzierte an einem Beispiel Aufgabengebiete und Methoden der Datensicherung. Sehr anschaulich demonstrierte der Referent wie das Verständnis von Forschungsdaten auch die Bewertung und vor allem Prinzipien der Qualitätssicherung beeinflusst. Mit seinen Schlagworten der Forschungsdaten wahlweise als „Abfallprodukt“, „Öl des Forschungsgetriebes“ bzw. „Gold zukünftiger Forschung“ gab er für die nachfolgende Podiumsdiskussion wichtige Impulse und zeigte auf, wie eng Forschungsdaten mit ihrem Entstehungskontext verbunden sind und welcher Aufwand gegebenenfalls notwendig ist, um diese Informationsschätze zu heben. Gleichzeitig verwies Sahle auf eine notwendige institutionelle Breite künftiger Einrichtungen des Forschungsdatenmanagements, um die verschiedenen,

komplexen Problemschichten sachgerecht zu adressieren. Sahle wünschte sich die Entwicklung eines Netzwerks von Institutionen, das ausgehend von den klassischen Gedächtnisinstitutionen wie Bibliotheken, Archive, Museen und Verlage auch fachbezogene, thematische und lokale Datenzentren integriert. Erst ein solches Netzwerk könne sowohl generische Daten-Dienste für zentrale Aufgaben wie der Ergebnissicherung, Langzeitspeicherung, Metadatenproduktion und allgemeinen Erschließung sicherstellen, als auch fachspezifische Angebote zur Unterstützung der digitalen Datenproduktion und -kuration, fachmethodischer Klassifikation, Analyse und Nachnutzung bereitstellen.

Während der geplante Vortrag zu den rechtlichen Herausforderungen leider ausfallen musste, diskutierte KATRIN MOELLER (Halle) Möglichkeiten zur Förderung des *data-sharings* durch stärkere Anreizbildungen innerhalb der Wissenschaftskulturen. In der digitalen Forschungslandschaft entstanden in den letzten Jahrzehnten Arbeitsfelder, die über die üblichen Verfahren der Qualitäts- und Reputationsmessung vor allem in der auf Monografien und Sammelbänden aufbauenden Ergebnissicherung der Geschichtswissenschaft kaum erfasst werden. Eine adäquate wissenschaftliche Würdigung der Aufwendungen für die Bereitstellung von Forschungsportalen, Datenbanken und digitalen Werkzeugen erfolgt hingegen bisher nicht. Letztlich muss aber gerade das Spannungsverhältnis von relativ hohen Aufwendungen für die Bereitstellung von Forschungsdaten, Urheberrecht und wissenschaftlicher Reputation so gelöst werden, dass ein optimales Verfahren für eine möglichst unkomplizierte, rechtlich tragfähige Datennachnutzung entsteht. Dazu sei es erforderlich, in gleichem Maß die Interessen von Datengebern und Datennehmern zu berücksichtigen und das Teilen von Daten vor allem auch für den Datengeber attraktiv zu machen. Über einen Vergleich verschiedener Möglichkeiten der Autorennennung bei der Verwendung gängiger CC-BY-Lizenzen spürte die Referentin den Vor- und Nachteilen der Co-Autorenschaft, des Zitats und der Namensnennung im Vorwort von Forschungsergebnissen über die Fachgrenzen der Geisteswissenschaften nach.

Ihr Fazit bestand im Vorschlag einer neuen Form der Datenautorenschaft, die eine funktionale Trennung zwischen Text- bzw. Ergebnisproduzenten auf der einen und allen anderen Beiträgern zu einem Forschungsergebnis auf der anderen Seite (Software, nachgenutzte Daten, Editionen etc.) vorsieht, sofern die Benutzung der entsprechenden Daten sich nicht mit einem üblichen Kleinzitat vereinbaren lasse. Beiden Formen der „Autorenschaft“ sollten dabei nur die funktionellen Zuschreibungen zufallen, für die sie jeweils verantwortlich sein können. Während Ergebnis- und Textproduzenten für das neu entstehende Forschungsergebnis Verantwortung übernehmen, bliebe die Zuständigkeit der Beiträger auf ihre eigene Leistung (Software, Daten, Edition) begrenzt, womit sich eine in den Geisteswissenschaften anerkannte Form der wissenschaftlichen Reputation verbinden ließe (Co-Autorenschaft). Damit entfalle die Hürde der Kontaktaufnahme und eine einfache Weiterverwendung von Daten analog zur Nachweisung in einem Zitat könnte etabliert werden.

Im Anschluss an diesen Vortrag entspann sich eine Debatte um die bis heute ungeklärte Frage, ob geisteswissenschaftliche Forschungsdaten über die erforderliche Schöpfungshöhe¹ verfügen, um Urhebererschaft geltend machen zu können. Hier fordern Geisteswissenschaftler/innen mehr Rechtssicherheit, weil nur so die Anreiz- und Folgewirkungen für die Bereitstellung von Forschungsdaten Berücksichtigung erfahren würden. Diese Diskussion zeigte, wie nachteilig sich ein wissenschaftliches Reputationssystem auswirke, dass letztlich weitgehend auf urheberrechtlich geschützter Autorenschaft in Kombination mit Rezensionen aufbaut.

Daran anschließende Diskussionsbeiträge thematisierten die notwendige innerfachliche Akzeptanz von alternativen Formen der Reputationswürdigung (etwa die zunehmende Bedeutung von Zitationsindizes) bzw. bekräftigten den Bedarf neuer „Rollen“ von „Autorenschaften“ – durchaus auch abseits des Urheberrechts. Als Fazit dieses Diskussionsstranges kann festgehalten werden, dass (nicht nur) die Geschichtswissenschaft aufgefördert ist, Strategien einer digitalen Ergebnis- und Qualitätssicherung zu entwickeln und

zu diskutieren. Diese Strategien sollen nicht einfach nur die analoge Handhabung (Urheberschaft und Rezensionswesen) ins Digitale übertragen, sondern die vielfältige digitale Ergebnisproduktion über neue Formen der Qualitätssicherung in den Blick nehmen.

Mit dem Punkt des Mehrwerts und der Qualitätssicherung startete anschließend die sehr lebhaft Podiums- und Publikumsdiskussion. RÜDIGER HOHLS (Berlin) reklamierte die fehlende Infrastruktur sowohl der digitalen Ergebnissicherung unmittelbar in der Forschung, wie der mangelnden Reputation anderer Arbeitsbereiche des Digitalen – abseits von Forschungsdaten. Der hohe Arbeitsaufwand für die nachhaltige Sicherung von Forschungsdaten könne nicht allein über fach- und raumübergreifende Dienste und Werkzeuge des Forschungsdatenmanagements aufgefangen werden, sondern müsse beim Forschenden ansetzen. Dies erfordere ein „Re-Design“ des Forschungsprozesses. Dabei sei in der Abwägung von Zeit- und Personalressourcen auch eine strikte Konzentration auf den Forschungsprozess notwendig. Ähnliche Positionen wurden aus dem Publikum mit mehrfachen Beiträgen unterstützt. Besonders starkes Interesse erfuhr die Frage nach dem Mehrwert der Bereitstellung nachnutzbarer Forschungsdaten verbunden mit Überlegungen, ob damit letztlich nicht eher eine Bürokratisierung und Wegführung von der Forschung verbunden sei.

THOMAS STÄCKER (Darmstadt) betonte in seinem Podiumsbeitrag die Chancen von FAIR(en)-Forschungsdaten und die Vorteile eines nachhaltigen Arbeitens mit diesen Daten. Der Mehrwert speise sich nach seinem Dafürhalten nicht nur aus der kreativen Quer- und Fremdnutzung von Daten, die ursprünglich vielleicht einem ganz anderen Entstehungskontext entstammen, sondern ebenso aus einer multiplen breiten methodischen Neuinterpretation. Für ihn war nicht nachvollziehbar, warum gerade Historiker/innen als Nutznießende von Forschungsdaten aller Couleur eigentlich über den Mehrwert von Forschungsdatenmanagement dis-

¹Jan Kröger, Urheberrecht, in: forschungsdaten.info, Konstanz 2018, URL: <https://www.forschungsdaten.info/themen/rechte-und-pflichten/urheberrecht/>.

kutieren müssten, schließlich seien doch gerade sie auf die Nachnutzung von Daten angewiesen. Er betonte nachdrücklich den hohen Wert digitaler Daten und den großen Nutzen von strukturierten Forschungsmethoden und Datenbeständen.

Aus dem Publikum angeregt wurde daran anschließend eine alternative Diskussion darüber, was Forschungsdaten sind. Letztlich gehe es aus wissenschaftlicher Perspektive gar nicht so sehr darum über einzelne Datentypen und ihre Besonderheiten zu verhandeln, sondern vielmehr um Qualitätsanforderungen von Daten. Hier machte die Diskussion eine Vielzahl von Problemschichten sichtbar: Welche Qualitätsstufen von Daten sollen vornehmlich gesichert werden? Wer bewertet und entscheidet, was aufgehoben wird? Gelten für alle – gleich ob es sich um ein Dissertationsprojekt, ein großes Drittmittel oder gar ein Verbundprojekt handelt – die gleichen Qualitätsanforderungen? Können oder sollen von Seiten der Wissenschaft verbindliche Normen und Standards festgelegt werden oder grenzt dies die Handlungsspielräume zu stark ein?

Dabei gab es starke Positionen und Argumente, einerseits für ein „Quick & Dirty-Prinzip“, da auch „dreckige“ Daten einen erheblichen Mehrwert für die Erschließung von Quellen oder die Nachnutzung weitgehend unerschlossener Datenbestände bieten. Ein Archivar befürwortete nachdrücklich die gute Anbindungsmöglichkeit solcher – auch hermeneutischer Forschungsdaten – für Archive bzw. die Notwendigkeit, die Annotation und Erschließung an die einschlägigen Quellen zurückzubinden. Zudem sah er die Archive, als klassische Institutionen der Konservierungsbewertung, auch für Forschungsdaten einschlägig als zuständig an. Ähnliches wünschten sich Wissenschaftler/innen von Forschungsdatenzentren anhand von Forschungsmethoden und Dokumentationen. Google & Co. würden letztlich ihren Erfolg gerade aus prozessualen Daten ziehen, weil diese schnell und massenhaft verfügbar seien. Letztlich müsse die komplementär dazu sich entwickelnde Automatisierung Datenqualität erzeugen. Katrin Moeller betonte, dass hierfür wiederum ein Wandel der Wissenschaftskultur erfolgen müsse. Ein

anderer Umgang mit Fehlern und ein Abrücken von Standards wie Double Keying, Peer Review-Verfahren und „Buchqualität“ erfordere eben auch eine Akzeptanz in der breiten Forschungslandschaft. MARINA LEMAIRE (Trier) und Katrin Moeller führten die Idee aus, den bereitgestellten Forschungsdaten unterschiedliche Qualitätssiegel hinzuzufügen, die die verschiedenen Qualitäten des Datenmanagements auf einen Blick erfassbar und sichtbar machen würden. Lemaire schlug vor, einfach mal den Versuch zur Publikation von prozessualen Daten durch etablierte Forschende anzustoßen, um Reaktionen aus dem Umfeld der Forschung auszutesten.

Thomas Stäcker und Patrick Sahle plädierten auf der anderen Seite für eine Konzentration auf hohe Qualitätsstandards, da erst dadurch wissenschaftliche Forschung vom freien Markt internetbasierter Daten abgrenzbar würde. Stäcker verwies in diesem Zusammenhang auf die Unterscheidung von Publikation (mit bisheriger Verfügbarmachung durch Bibliotheken) und Nachlass (Bewertung und Erschließung durch Archiv) mit unterschiedlichen Kompetenzen. Ein anderer Diskussionsteilnehmer betonte, dass nur publizierte und qualitätsgesicherte Daten auch Rückschlüsse auf den Urheber und Datenproduzenten ermöglichen und so wichtige Anhaltspunkte zur Quellenkritik leisten würden. Grundlegend wurde in diesem Kontext noch einmal auf eine fehlende Digitalisierung von Studium und Lehre verwiesen, die heute vor allem in der Breite der geschichtswissenschaftlichen Angebote nicht durchgesetzt werde, sondern zunehmend in Sonderstudiengänge externalisiert. Marina Lemaire bekräftigte jedoch nochmals einerseits die Ausweisung von fachspezifischen Curricula und die Grundlagen zur Datenmodellierung als fachinterne Kompetenz, andererseits aber auch die Notwendigkeit kooperativer Lehre, da Forschungsleistungen heute bereits oft nicht mehr als Leistung einer einzelnen Disziplin erbracht würden.

Insgesamt wurde auf die Notwendigkeit von fachbezogenen Standards und verlässlichen Werkzeugen und Methoden verwiesen, die verknüpft mit Automatisierung wesentlich zur Qualitätsverbesserung und Veralltäglichen des Datenhandlings beitragen kön-

nen. Der Wandel – dies betonte Torsten Hiltmann aus dem Publikum – ist vor allem ein innerfachlicher Wandel, der zu einem methodenbewussteren Forschen und Handeln in der Geschichtswissenschaft führen sollte und könnte.

Sektionsleitung: Katrin Moeller (Halle)

Patrick Sahle (Köln): Was sind Forschungsdaten in der Geschichtswissenschaft?

Katrin Moeller (Halle): Ausweisung der Datenautorenschaft als neues Verfahren zur Messung wissenschaftlicher Leistung

Podiumsdiskussion

Marina Lemaire (Trier): Moderation

Rüdiger Hohls (Berlin) / Katrin Moeller (Halle) / Patrick Sahle (Köln) / Thomas Stäcker (Darmstadt)

Tagungsbericht *HT 2018: Forschungsdaten: rechtliche Herausforderungen und wissenschaftliche Reputation. Forschungsdatenmanagement als Bestandteil einer neuen Wissenschaftskultur*. 25.09.2018–28.09.2018, Münster, in: *H-Soz-Kult* 30.11.2018.