

Geisteswissenschaftliches Forschungsdatenmanage- ment in der Lehre – Konzepte, Methoden, Erfahrungen

Blumtritt, Jonathan

jonathan.blumtritt@uni-koeln.de
Universität zu Köln, Cologne Center for eHumanities (CCeH); Universität zu Köln, Data Center for the Humanities (DCH)

Helling, Patrick

patrick.helling@uni-koeln.de
Universität zu Köln, Data Center for the Humanities (DCH); Universität zu Köln, Institut für Digital Humanities (IDH)

Mathiak, Brigitte

bmathiak@uni-koeln.de
Universität zu Köln, Data Center for the Humanities (DCH); Universität zu Köln, Institut für Digital Humanities (IDH)

Neuefeind, Claes

c.neuefeind@uni-koeln.de
Universität zu Köln, Cologne Center for eHumanities (CCeH); Universität zu Köln, Data Center for the Humanities (DCH)

Rau, Felix

f.rau@uni-koeln.de
Universität zu Köln, Data Center for the Humanities (DCH); Universität zu Köln, Institut für Linguistik (IfL)

Schildkamp, Philip

philip.schildkamp@uni-koeln.de
Universität zu Köln, Data Center for the Humanities (DCH)

Wieners, Jan Gerrit

jan.wieners@uni-koeln.de
Universität zu Köln, Institut für Digital Humanities (IDH)

Abstract

Mit diesem Beitrag möchten wir unsere Erfahrungen in der Vermittlung von Strategien und Ansätzen im Bereich des Forschungsdatenmanagements (FDM), die wir unter Einbeziehung unterschiedlicher Zielgruppen in der Lehre an der Universität zu Köln (UzK) gesammelt haben, vorstellen und mit der DH-Community diskutieren. Wir möchten dabei insbesondere die Frage aufwerfen, ob das Thema FDM als obligatorischer Lehrinhalt verstetigt werden sollte und wenn ja, inwieweit sich das Thema dauerhaft in die Lehre einbringen lässt, damit frühzeitig Sensibilität für Herausforderungen in der Erfassung, Bereitstellung und langfristigen Archivierung von Forschungsdaten aufgebaut werden kann.

Entwicklung des institutionellen Rahmens

Die Beschäftigung mit den Themen FDM und Langzeitarchivierung ist am *Cologne Center for eHumanities* (CCeH), dem *Kölner Data Center for the Humanities* (DCH) sowie dem *Institut für Digital Humanities* (IDH) (2017 hervorgegangen aus dem Institut für Historisch-Kulturwissenschaftliche Informationsverarbeitung und der Abteilung Sprachliche Informationsverarbeitung des Instituts für Linguistik) intensiv mit Forschung, Lehre und Projektpraxis verknüpft.¹ So flossen bereits früh Erfahrungen aus Projekten wie dem bis 2010 von der europäischen Union geförderten Projekt *PLANETS* oder dem *Digitalen Archiv NRW* in Bachelor- und Masterveranstaltungen ein.²

Ausgehend von der Annahme, dass es längerfristig zu einer Verbesserung der nachhaltigen Sicherung von wissenschaftlichen Inhalten führt, wenn Forschende bereits zu Beginn ihrer Ausbildung mit dem Problembereich und den verfügbaren Lösungsansätzen vertraut gemacht werden, haben wir am DCH mehrere Lehrveranstaltungen zum Themenbereich FDM konzipiert und im Rahmen der Studiengänge "Informationsverarbeitung" und "Medieninformatik" durchgeführt.

Die Lehrveranstaltungen thematisieren sowohl das Management von Forschungsprimärdaten, wie auch die Nutzbarkeit von erzeugten digitalen Artefakten (z.B. Präsentationssysteme, Dateneingabemasken und weitere). Zielgruppe der Lehrveranstaltungen sind vor allem Studierende der (digitalen) Geisteswissenschaften. Darüber hinaus veranstalten wir regelmäßig einen Workshop zum Thema Forschungsdatenmanagement für Promovierende an der a.r.t.e.s. Graduiertenschule der Philosophischen Fakultät, in dem insbesondere praktische FDM-Aspekte während der Promotion im Fokus stehen.³

Ziel ist es unter anderem direkte Rückmeldung von den Studierenden/Promovierenden als – wenn auch punktuelle – praktische Anwender von FDM-Angeboten zu erhalten.

Der Fokus liegt hierbei besonders auf der Nutzererfahrung. So wird in den Kursen unter anderem diskutiert, wie aufwendig FDM in der Forschungsarbeit ist, und ob es in manchen Fällen als hinderlich oder gar belastend empfunden wird. Im Zuge der Veranstaltungen wird auch der Bekanntheitsgrad typischer FDM-Werkzeuge evaluiert. Diese direkte Rückmeldung aus der Erfahrungswelt angehender Wissenschaftler*innen fließt unmittelbar in die Beratungs- und Schulungskonzepte des DCH ein.

Im Folgenden beschreiben wir zunächst die Lehrinhalte des Workshops und der beiden Übungen, bevor wir abschließend ein kurzes Fazit ziehen, das als Grundlage für eine Diskussion mit der Community dienen soll.

Workshop „Forschungsdatenmanagement“ an der a.r.t.e.s. Graduiertenschule der Philosophischen Fakultät

Um Promovierende frühzeitig mit dem Thema Forschungsdatenmanagement vertraut zu machen, veranstalten wir seit dem Sommersemester 2018 einmal im Jahr einen zweitägigen Workshop an der a.r.t.e.s. Graduiertenschule der Philosophischen Fakultät der UzK. Der Workshop findet aufgeteilt in vier Sessions à vier Stunden/Woche statt, ist auf maximal 15 Promovierende ausgelegt und gehört zum Wahlpflichtbereich.

Im ersten Workshop im Sommersemester 2018 lag der Fokus auf der theoretischen Betrachtung von Forschungsdatenmanagement innerhalb der Geisteswissenschaften. Als Struktur dienten hierzu die acht Säulen des Forschungsdatenmanagements der AG Datenzentren des Verbandes “Digital Humanities im deutschsprachigen Raum” (DHd) [Helling, Moeller und Mathiak 2018].⁴

In der Evaluation des Workshops wurde deutlich, dass sich die Informations- und Servicebedarfe der Promovierenden deutlich von denen von BA-/MA-Studierenden oder Forschenden im universitären Regelbetrieb unterscheiden: Neben Informationen über Servicestrukturen an der UzK wurden auch insbesondere praktische Lösungen für FDM in Bezug auf das eigene Promotionsprojekt nachgefragt.

Entsprechend wurde für den zweiten Workshop im Sommersemester 2019 gemeinsam mit Vertreter*innen des Regionalen Rechenzentrums Köln (RRZK) ein Dienstportfolio für Software- und Hardware-services an der UzK zusammengestellt, das in einer der Sessions vorgestellt und diskutiert wurde.⁵ Der gesamte Workshop wurde deutlich praxisorientierter geplant und durchgeführt: Nach einem verkürzten Theorie- und Einführungsblock wird in einer praktischen Übung der Forschungsdatenlebenszyklus aus einer datengebenden und datennehmenden Perspektive nachvollzogen. Mit einer Rechercheübung wird anhand eines Fragenkatalogs das

passende Repositorium für den eigenen Fachbereich gesucht, eigene Forschungsdaten werden mit Metadaten beschrieben und in einer Diskussionsrunde werden die Themen Urheberrecht und Persönlichkeitsrecht und deren Auswirkungen auf das Forschungsdesign behandelt. Als Kursabschluss formulieren die Teilnehmenden einen Datenmanagementplan für ihr Promotionsprojekt nach dem Datenmanagementplan-Template der UzK.⁶

FDM-Übung für Masterstudierende an der Philosophischen Fakultät

Für Studierende im Masterstudiengang “Informationsverarbeitung” bieten wir ebenfalls Veranstaltungen an. Die im Sommersemester 2017 von Simone Kronenwett und Jan G. Wieners durchgeführte Veranstaltung “Forschungsdatenmanagement und Langzeitarchivierung” intendierte, Teilnehmende an verschiedenste Aspekte des FDM und der Langzeitarchivierung heranzuführen und animierte dazu, eigene Forschungsfragen zu entwickeln.⁷ So gossen Gastvorträge von domänenrelevanten Fachwissenschaftler*innen aus Einrichtungen wie dem Forschungsdatenzentrum Archäologie & Altertumswissenschaften IANUS oder dem Rheinisch-Westfälischen Wirtschaftsarchiv das Fundament für Ausarbeitungen der Studierenden, die sich beispielsweise mit Fragen nach guten Speicherformaten für Rastergraphiken oder die Sicherheit physischer Speichermedien im Hinblick auf Langzeitverfügbarkeit beschäftigten.⁸

Auch die Übung “Forschungsdatenmanagement”, die wir seit 2019 anbieten, richtet sich an Studierende in den Masterstudiengängen “Informationsverarbeitung” und “Medieninformatik”, hat dabei aber einen stärkeren praktischen Fokus.⁹ Anders als bei Kursen, die zukünftige Wissenschaftler*innen ausschließlich als potentielle Datenproduzent*innen adressieren, legen wir den Fokus nicht nur auf FDM-Grundlagen, Awareness und Fortbildung, sondern beschäftigen uns aktiv mit den Problemen aus der Sicht von Forschungsdatenmanager*innen und -kurator*innen. Es wird das Datenmanagement und die -kuration als mögliches Berufsfeld vorgestellt.

Bei der inhaltlichen Gestaltung haben wir uns an Kursen orientiert, die bereits andernorts erfolgreich durchgeführt wurden. Insbesondere der Kurs von Iris Vogel von der Universität Hamburg hat uns in vielen Punkten inspiriert, wie auch schon die bereits erwähnte Veranstaltung von 2017.¹⁰ Ähnlich wie bei der Veranstaltung für Doktorand*innen setzen wir auf eine Mischung von Theorie und Praxis. Etwa die Hälfte der Sitzungen schließt mit einer praktischen Hausaufgabe, für die andere Hälfte soll Literatur vorbereitet werden, die in einem Quizformat diskutiert wird.

Während das Format bei den Studierenden grundsätzlich gut ankam, waren die Themen zum Teil doch sehr weit von ihrem Arbeitsalltag entfernt. Dazu kam, dass die Gruppe sehr heterogen war. Während die linguistisch orientierten Studierenden bereits Forschungsdaten für ihre Projekte genutzt hatten und gerne über die Details sprechen wollten, waren die eher medien- oder literaturwissenschaftlich ausgerichteten Studierenden zum Teil noch gar nicht mit dem Thema Forschungsdaten in Berührung gekommen und dementsprechend skeptisch. In der Zukunft planen wir dies zu adressieren, indem wir beide Gruppen noch stärker fachspezifisch abholen und beispielsweise digitale Editionen und die dazugehörige Datenmodellierung stärker in den Fokus nehmen.

Seminar „Virtualisierungstechnologien für Forschungssoftware“

In einem weiteren Seminar, das wir seit dem Wintersemester 2018/2019 anbieten, fokussieren wir das spezifische Problem der nachhaltigen Sicherung von Forschungssoftware. Im Gegensatz zu datenbezogenem FDM ist die Frage nach dem Umgang mit Forschungssoftware bislang noch weitgehend unbeantwortet und stellt deshalb eine besondere Herausforderung im Hinblick auf die langfristige Sicherung von Forschungsergebnissen dar. Kontext der Seminare ist das von der DFG geförderte Projekt “SustainLife”, welches die Erprobung eines Konzepts zur nachhaltigen Sicherung lebender Systeme auf Basis des OASIS-Standards TOSCA (*Topology and Orchestration Specification for Cloud Applications*) [OASIS 2013, 2016] zum Gegenstand hat [Barzen et. al 2018, Neuefeind et al. 2018, 2019].¹¹ Die im Projekt erarbeiteten Methoden und Erfahrungen sollen dazu beitragen zukunftssichere Standards und Nachhaltigkeitsstrategien für den Umgang mit Forschungssoftware zu etablieren. Dies soll unter anderem durch die Modellierung von Applikationstopologien mittels der Open-Source Implementierung OpenTOSCA erreicht werden, die durch den Projektpartner IAAS Stuttgart bereitgestellt wird.¹²

Im Rahmen des Seminars werden zum einen verschiedene Nachhaltigkeitsstrategien beleuchtet, zum anderen wollen wir den im Projekt entwickelten Lösungsansatz auf Grundlage des TOSCA-Standards von Studierenden erproben lassen. Im Mittelpunkt steht dabei die Frage, ob die TOSCA-konforme Modellierung heterogener Softwarelösungen in einer Weise vermittelt werden kann, dass Studierende der (digitalen) Geisteswissenschaften in der Lage sind, eigene Use Cases mit akzeptablem Aufwand zu bearbeiten.

Die Übungen wurden trotz der stark spezialisierten Ausrichtung sehr gut angenommen. In der ersten Übung zeigte sich dabei eine sehr deutliche

Lücke zwischen geisteswissenschaftlich-konzeptueller und informationstechnologischer Kompetenz. Während das grundlegende Konzept einer TOSCA-konformen Modellierung und dessen Mehrwert von den Studierenden schnell erfasst und gut nachvollzogen wurde, stellte sich die praktische Umsetzung als deutlich größere Herausforderung heraus. Dies konnte durch Gruppenarbeiten zwar zu Teilen aufgefangen werden, dennoch wurde deutlich, dass die konkrete Umsetzung von Anwendungsmodellen ein erhebliches softwaretechnologisches Vorwissen voraussetzt. In der Folge adressierten wir dies dadurch, dass wir die reine Modellierung von der technischen Umsetzung stärker trennen, so dass die Gruppen sich nach technischen Kompetenzen aufteilen können.

Fazit und Ausblick

Die hier vorgestellten Lehrveranstaltungen fokussieren auf unterschiedliche Weise verschiedene Aspekte des Forschungsdatenmanagements und sprechen unterschiedliche Zielgruppen an. Über die genannten Veranstaltungen hinaus wird das Thema FDM an der UzK auch in weiteren Veranstaltungen thematisiert, wenn auch nicht so ausschließlich. Insgesamt haben wir gute Erfahrungen mit den Veranstaltungen gemacht. Das Feedback zeigt, dass eine starke Verankerung in der Praxis bei den Studierenden gut ankommt.

In unserem Poster werden wir den Aufbau der Kurse genauer beleuchten und den Benefit für die Studierenden diskutieren. Wir laden dazu ein, Erfahrungen auszutauschen, um zu eruieren welche Kompetenzen bereits im Studium zu Forschungsdatenmanagement erworben werden sollten und welche Aspekte für besonders dringlich gehalten werden.

Fußnoten

1. Cologne Center for eHumanities (CCEH) an der Universität zu Köln, Online: <https://cceh.uni-koeln.de/> (letzter Zugriff: 12.09.2019); Data Center for the Humanities (DCH) an der Universität zu Köln, Online: <https://dch.phil-fak.uni-koeln.de/> (letzter Zugriff: 12.09.2019); Institut für Digital Humanities (IDH) an der Universität zu Köln, Online: <https://dh.uni-koeln.de/> (letzter Zugriff: 12.09.2019).
2. Preservation and Long-term Access through Networked Services, Online: (letzter Zugriff: 12.09.2019); Digitales Archiv NRW, Online: (letzter Zugriff: 12.09.2019).
3. a.r.t.e.s. Graduate School der Philosophischen Fakultät der Universität zu Köln, Online: <http://artes.phil-fak.uni-koeln.de/> (letzter Zugriff: 12.09.2019); Workshop Forschungsdatenmanagement a.r.t.e.s., Online: <https://artes.phil-fak.uni-koeln.de/40285.html> (letzter Zugriff: 12.09.2019).

4. Digital Humanities im deutschsprachigen Raum (DHd), Online: <https://dig-hum.de/> (letzter Zugriff: 12.09.2019); AG Datenzentren im DHd-Verband, Online: <https://dhd-ag-datenzentren.github.io/> (letzter Zugriff: 12.09.2019).
5. Regionales Rechenzentrum (RRZK) der Universität zu Köln, Online: <https://rrzk.uni-koeln.de/> (letzter Zugriff: 12.09.2019); Dienstportfolio des RRZK, Online: <https://rrzk.uni-koeln.de/datenmanagement.html> (letzter Zugriff: 12.09.2019).
6. Datenmanagementplan-Template der Universität zu Köln, Online: <https://fdm.uni-koeln.de/27602.html> (letzter Zugriff: 12.09.2019).
7. Forschungsdatenmanagement und Langzeitarchivierung - Übung Sommersemester 2017, Online: <http://lehre.idh.uni-koeln.de/lehveranstaltungen/sosem17/forschungsdatenmanagement-und-langzeitarchivierung/> (letzter Zugriff: 12.09.2019).
8. Forschungsdatenzentrum Archäologie & Altertumswissenschaften IANUS, Online: <https://www.ianus-fdz.de/> (letzter Zugriff: 12.09.2019); Rheinisch-Westfälische Wirtschaftsarchiv, Online: <https://www.rwwa.de/> (letzter Zugriff: 12.09.2019).
9. Forschungsdatenmanagement - Übung Sommersemester 2019, Online: <http://lehre.idh.uni-koeln.de/lehveranstaltungen/sosem19/forschungsdatenmanagement/> (letzter Zugriff: 12.09.2019).
10. Stefan Thiemann, Iris Vogel und Juliane Jacob: Materialien zum Kurs "Forschungsdatenmanagement" an der Universität Hamburg, Online: <https://www.hoou.de/materials/forschungsdatenmanagement> (letzter Zugriff: 12.09.2019).
11. SustainLife - Erhalt lebender, digitaler Systeme für die Geisteswissenschaften; Projektnummer 379522012. Online: <https://gepris.dfg.de/gepris/projekt/379522012> (letzter Zugriff: 12.09.2019).
12. Institut für Architektur von Anwendungssystemen, Universität Stuttgart; Online: <https://www.iaas.uni-stuttgart.de> (letzter Zugriff: 12.09.2019).

ISSN (Print) 0720-6763, DOI: <https://doi.org/10.1515/abitech-2018-3006>.

Neuefeind, Claes / Schildkamp, Philip / Mathiak, Brigitte / Mar#i#, Alexander / Hentschel, Frank / Harzenetter, Lukas / Breitenbücher, Uwe / Barzen, Johanna / Leymann, Frank (2019): "Sustaining the Musical Competitions Database: a TOSCA-based Approach to Application Preservation in the Digital Humanities". To Appear in: Book of Abstracts of the Digital Humanities Conference 2019 (DH2019). Utrecht, Netherlands. 09.07-12-07.2019. Web: <https://dev.clariah.nl/files/dh2019/boa/0574.html> [letzter Zugriff 5.9.2019]

Neuefeind, Claes / Harzenetter, Lukas / Schildkamp, Philip / Breitenbücher, Uwe / Mathiak, Brigitte / Barzen, Johanna / Leymann, Frank (2018): "The SustainLife Project – Living Systems in Digital Humanities" in: Proceedings of the 12th Advanced Summer School on Service-Oriented Computing, 2018 (IBM Research Report RC25681), S.101-112.

OASIS (2013): „Topology and Orchestration Specification for Cloud Applications Version 1.0“. 25 November 2013. OASIS Standard. <http://docs.oasis-open.org/tosca/TOSCA/v1.0/os/TOSCA-v1.0-os.html> . [letzter Zugriff 5.9.2019]

OASIS (2016): „TOSCA Simple Profile in YAML, Version 1.0“. Edited by Derek Palma, Matt Rutkowski, and Thomas Spatzier. 21 December 2016. OASIS Standard. <http://docs.oasis-open.org/tosca/TOSCA-Simple-Profile-YAML/v1.0/os/TOSCA-Simple-Profile-YAML-v1.0-os.html> . [letzter Zugriff 5.9.2019]

Bibliographie

Barzen, Johanna / Blumtritt, Jonathan / Breitenbücher, Uwe / Kronenwett, Simone / Leymann, Frank / Mathiak, Brigitte / Neuefeind, Claes (2018): "SustainLife – Erhalt lebender, digitaler Systeme für die Geisteswissenschaften." in: Book of Abstracts der 5. Jahrestagung der Digital Humanities im deutschsprachigen Raum (DHd 2018), Köln 26.02-02.03.2018, S. 471-474, DOI: <https://doi.org/10.18716/KUPS.8085> .

Helling, Patrick / Moeller, Katrin / Mathiak, Brigitte (2018): "Forschungsdatenmanagement in den Geisteswissenschaften - der Dienstekatalog der AG Datenzentren des Verbands 'Digital Humanities im deutschsprachigen Raum' (DHd)" in: ABI Technik, Band 38, Heft 3, Seiten 251–261, ISSN (Online) 2191-4664,