

Umfrage zu Forschungsdaten an der Philosophischen Fakultät der Universität zu Köln

Mathiak, Brigitte

bmathiak@uni-koeln.de

Universität zu Köln, Deutschland

Kronenwett, Simone

simone.kronenwett@uni-koeln.de

Universität zu Köln, Deutschland

Executive Summary

Um den aktuellen Bedarf an der Philosophischen Fakultät der Universität zu Köln im Umgang mit Forschungsdaten möglichst genau identifizieren zu können, wurde vom Data Center for the Humanities (DCH) in Kooperation mit dem Dekanat der Philosophischen Fakultät sowie der Universitäts- und Stadtbibliothek (USB) Köln 2016 eine Online-Umfrage unter dem akademischen Personal der Fakultät durchgeführt. Ziel der Erhebung ist es, sowohl die aktuellen Bestände zu charakterisieren, als auch Informationen zum Bedarf in den Bereichen Forschungsdatenmanagement (FDM) und Beratung zu erhalten. Im Vortrag werden die Ergebnisse der Umfrage präsentiert und diskutiert sowie mögliche Schlussfolgerungen erörtert.

Aktualität und Relevanz

Eines der wichtigsten neuen Handlungsfelder der Forschung, welche im Zuge der Digitalisierung von Information entstanden ist, betrifft das Management von Forschungsdaten. Die Hochschulen müssen sich darauf einstellen, ihren Wissenschaftlern und Forschern die notwendigen Infrastrukturen und Services zur Verfügung zu stellen. Auf diese Dringlichkeit verwies auch jüngst der Rat für Informationsinfrastrukturen in seinen Empfehlungen *Leistung aus Vielfalt* (Rat 2016). Denn noch immer gehen laut Schätzungen der DFG bis zu 90% der digital produzierten Daten und Ergebnisse nach kurzer Zeit verloren bzw. "verschwinden in der Schublade" (Kramer 2014) und stehen somit keiner weiteren Verwendung und Nachnutzung zur Verfügung (Winkler-Nees 2011). Auch deshalb verabschiedete die Hochschulrektorenkonferenz (HRK) gleich zwei Grundsatzpapiere, in denen das Management von Forschungsdaten als zentrale strategische Herausforderung für die Hochschulleitungen angesehen

wird (Hochschulschulrektorenkonferenz 2014 und 2015). Um einerseits die vielfältigen Aktivitäten und Akteure zu koordinieren und andererseits die Anschlussfähigkeit möglichst aller Hochschulen in den Scientific Communities auf nationaler und internationaler Ebene zu gewährleisten, erarbeitete die HRK einen 6-Punkte-Leitfaden, die sich aus ihrer Sicht beim Auf- oder Ausbau des institutionellen FDM ergeben und berücksichtigt werden sollen (Hochschulrektorenkonferenz 2015: 6-15). Im Rahmen dieses Maßnahmenkatalogs wird explizit empfohlen, zu Beginn eine Standortbestimmung an der jeweiligen Hochschule vorzunehmen, "z.B. mittels geeigneter interner Erhebungen zum Verhalten der Wissenschaftlerinnen und Wissenschaftler, aber auch zu deren Bedarfen." (Hochschulrektorenkonferenz 2015: 9)

Methodischer Ansatz

Der gewählte methodische Ansatz der Umfrage-basierten Studie orientiert sich an den sechs Leitlinien von (Müller et al. 2014) sowie an den einschlägigen Aufsätzen des Handbuchs *Methoden der Bibliotheks- und Informationswissenschaft* von Umlauf et al (Seadle 2013: 41-63; Fühles-Ubach 2013: 114-127; Fühles-Ubach 2013: 96-113; Fühles-Ubach, Umlauf 2013: 80-95).

Definition der Forschungsziele

Der Anlass für diese Studie ist die HRK-Empfehlung zur Durchführung einer Umfrage zu Forschungsdaten an Hochschulen als Grundlage für eine institutionelle FDM-Strategieentwicklung. Denn im Gegensatz zu einigen anderen deutschen Hochschulen fehlt für die drittgrößte Universität in Deutschland sowohl eine entsprechende Erhebung als auch weiterführend eine universitätsweite FDM-Policy. Der Fokus liegt allerdings *nicht* auf einer quantitativen Totalerhebung zu Forschungsdaten an der Kölner Volluniversität. Da professionelles FDM fachbereichsspezifisch erfolgen sollte (Sahle et al. 2013), richtet sich der Blick gezielt auf die Forschungsdaten an der Philosophischen Fakultät der Universität zu Köln, einer der größten geisteswissenschaftlichen Fakultäten Europas (UzK 2016). In diesem Kontext werden die Ergebnisse der Umfrage helfen, zur konzeptionellen Weiterentwicklung und Optimierung des DCH-Beratungs- und Serviceangebots, ein zentrales Dienstleistungsangebot der Fakultät, beizutragen.

Bestimmung der Zielgruppe

Die Teilnehmergruppe ist beschränkt auf das akademische Personal der Philosophischen Fakultät der Universität zu Köln. Die Umfrage zielte dabei besonders auf Wissenschaftler und Forscher, die direkt für datengestützte Forschungsprojekte verantwortlich sind.

Spezifizierung des Fragebogendesigns

Für die inhaltliche, konzeptionelle und methodische Gestaltung der Umfrage wurden die bisher verfügbaren Erhebungen zu Forschungsdaten an nationalen und internationalen wissenschaftlichen Institutionen und fachspezifischen Forschungseinrichtungen analysiert (forschungsdaten.org 2016; Burger et al. 2013) und auf die besonderen Gegebenheiten an der Philosophischen Fakultät zugeschnitten (Andorfer 2015; Stäcker 2015; CCeH 2016, DCH 2016a). Es wurden auch die Ergebnisse mehrerer Experteninterviews des DCH mit Wissenschaftlern der Philosophischen Fakultät berücksichtigt, die im Vorfeld der Erhebung im Rahmen von FDM-Beratungen durch das DCH geführt wurden. Der Fragebogen wurde insgesamt in fünf Teilbereiche untergliedert: 1) Forschungsdaten 2) Nutzung von Datenarchiven 3) Unterstützung beim Umgang mit Forschungsdaten 4) Fachbereich und Position 5) Interesse.

Überprüfung und Pretests

Eine Word-Version des Fragebogens wurde zunächst an alle Kooperationspartner des Projektes verschickt. Nach Einarbeitung aller Rückmeldungen wurde der Fragebogen online programmiert und der Testlink an Wissenschaftler aller acht Fächergruppen der Fakultät sowie an externe Experten mit soziologischem Hintergrund für Pretests versendet. In mehreren Iterationen wurde der Fragebogen immer weiter adaptiert. Dies betraf u.a. die Reihenfolge der einzelnen Frageblöcke und die Auswahl der verwendeten Definitionen sowie die Präzision der Fragestellungen. Alle Schritte erfolgten in enger Absprache mit dem Datenschutzbeauftragten der Universität zu Köln.

Umsetzung und Einführung

Der Fragebogen wurde mit Hilfe des Online-Befragungstools von (Kronenwett&Adolphs 2016) erstellt und war vom 30.05.2016 bis 12.06.2016 aktiv. Die Auswertung der Ergebnisse erfolgte mit R.

Weiterführende Informationen zur Online-Umfrage sowie zu dem umfassenden Ergebnisbericht können auf der DCH-Webseite zur Umfrage eingesehen werden: <http://dch.phil-fak.uni-koeln.de/umfrage-2016.html> (DCH 2016b).

Deskriptive Datenanalyse: Ergebnisauswahl

Der Fragebogen wurde von 191 Personen begonnen und von 136 Teilnehmern vollständig beantwortet. 71.20% der Teilnehmer, die den Fragebogen begonnen haben, haben die Befragung auch beendet, d.h. sie haben alle Fragen vollständig beantwortet und sich bis zur Abschlusseite durchgeklickt. Die folgende Auswahl der Datenauswertung

berücksichtigt nur diese Teilnehmer (n=136). Unser Ziel bei der Erstellung des Fragebogens war folgende Fragen zu beantworten:

- Welche Forschungsdaten gibt es?
- Welchen Bedarf gibt es bezüglich Forschungsdaten?
- Welche Unterstützung wünschen sich die Mitglieder der Fakultät von uns?

Zur ersten Frage war uns Nachhaltigkeit und Volumen wichtig. Zur Nachhaltigkeit konnten wir feststellen, dass die Mehrzahl der Befragten die Daten auf ihren lokalen Rechnern speichert: 70% auf dienstlichen Rechnern, 70% auf privaten Rechnern, Mehrfachantworten waren möglich (vgl. Abb. 1). Nur 14% speichern ihre Daten in einem Datenarchiv, eine Zahl, die sich auch in anderen Fragen reflektiert wird, etwa wie viele sich vorstellen können ihre Daten in einem Datenarchiv abzulegen.

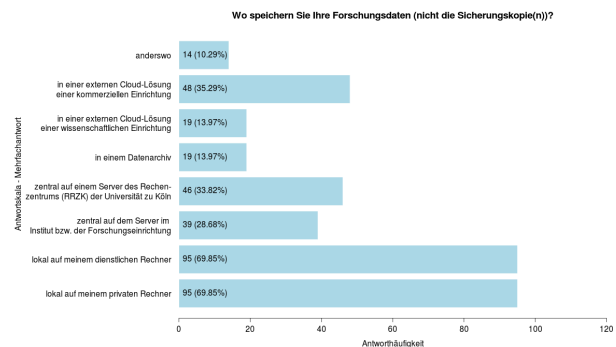


Abb. 1: Speicherort der Forschungsdaten (n=136)

Dies ist für die Nachhaltigkeit fatal, da nur in einem Datenarchiv ein strukturierter Zugriff und insbesondere auch Auffindbarkeit gewährleistet sind. Cloud-Lösungen, die auch weit verbreitet sind (35% Nutzung von kommerziellen Anbietern und 14% von wissenschaftlichen Anbietern), stellen zwar sicher, dass der Nutzer immerfort und von überall auf die Daten zugreifen kann und diese auch teilen kann. Aber für die Nachvollziehbarkeit von Forschungsergebnissen und die langfristige Sicherung sind diese denkbar ungeeignet.

Wir gehen davon aus, dass dies daran liegt, dass die Forscher ihre Handlungsweise nicht bezüglich Nachhaltigkeit und Nachvollziehbarkeit reflektieren. Die Selbsteinschätzung zu den eigenen Kenntnissen im Bereich FDM (vgl. Abb. 2) zeigt, dass die Kenntnisse größtenteils als durchschnittlich oder noch geringer (71%) eingeschätzt werden.

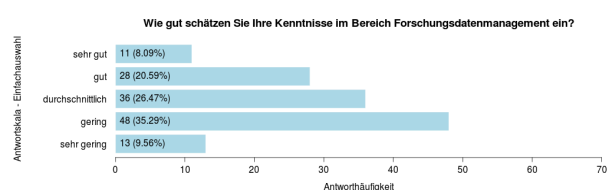


Abb. 2: Selbsteinschätzung der Kenntnisse im Forschungsdatenmanagement (n=136)

Es kann aber auch ein Faktor sein, dass selbst bei hohen Kenntnissen schlicht die Möglichkeiten fehlen, die Daten zu publizieren, oder es keine Motivation bzw. Ressourcen gibt, dies auch tatsächlich zu tun.

Die Nachhaltigkeit an sich wird schon als Problem gesehen. 66% der Befragten geben an, dass sie befürchten die Daten zu verlieren, wenn sich nach Projektende niemand mehr für die dazugehörigen Webseiten zuständig fühlt. 60% fürchten Datenkonversionsprobleme. Aber auch für Probleme mit der Auffindbarkeit (45%) und der Dokumentation (41%) besteht eine prinzipielle Sensibilität. In Abbildung 3 finden Sie noch weitere Probleme, die von den Befragten genannt wurden. Interessant ist in diesem Zusammenhang auch, dass nur 11% der Befragten den Datenschutz bzw. die Datensicherheit als Problem sehen. Dies könnte aber auch im Zusammenhang damit stehen, dass wir explizit nach Problemen mit Forschungsdaten aus der Nutzerperspektive und nicht aus der Datengeberperspektive gefragt haben.

In unserer Beratungspraxis und auch in der Frage welche Serviceleistungen von einem Datenzentrum gewünscht werden (Abb. 4) spielen rechtliche Aspekte und Zugriffseinschränkungen eine sehr große Rolle. 74% wünschen sich diesbezüglich eine Beratung. Es ist damit das meistgewünschte Thema. Ebenfalls in der Spitzengruppe sind Beratungen zu technischen Themen (73%) und allgemeiner Natur (66%), sowie die konkrete Bereitstellung von Speicherplatz zur Archivierung und Publikation von Forschungsdaten (72%). Im Mittelfeld wird Unterstützung beim Erstellen eines Datenmanagementplans, z.B. für Drittmittelanträge gewünscht (54%), Beratung für Archivierung und Zitation (50%) und der Betrieb von laufenden Anwendungen (46%). Letzteres gestaltet sich für uns äußerst schwierig umzusetzen. Auch in der Beratungspraxis werden wir immer wieder mit diesem Wunsch konfrontiert. Die technischen Hürden und notwendigen Ressourcen sind jedoch zum Teil beträchtlich.

Abb. 3: Allgemeine Probleme mit Forschungsdaten (n=136)

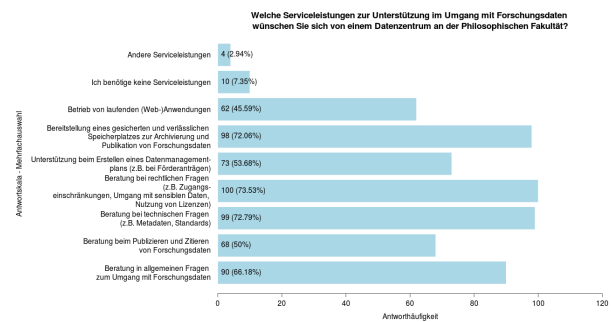


Abb. 4: Gewünschte Serviceleistungen (n=136)

Schlussfolgerungen und Ausblick

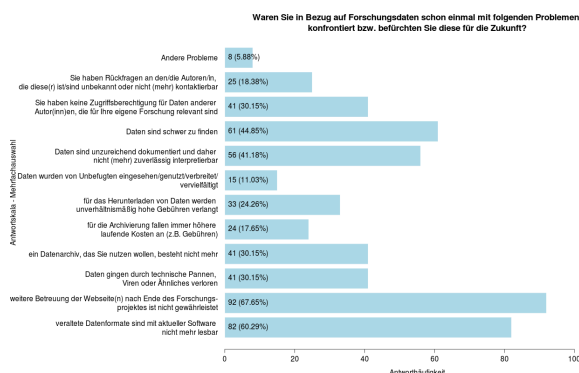
Im Vortrag werden wir noch genauer auf unsere Ergebnisse eingehen und diese auch mit anderen Studien vergleichen, die bereits an anderen Hochschulen sowohl in Deutschland (forschungsdaten.org 2016) als auch international (Kuipers et al. 2009; Bauer et al. 2016) durchgeführt wurden. Umfragen, wie die von uns durchgeführte, sind ein wichtiges Mittel für die strategische Positionierung von Institutionen, die sich mit Forschungsdaten beschäftigen. Wir werden daher im Vortrag auch kurz darauf eingehen, wie die von uns erhobenen Ergebnisse die Strategie des DCH (Data Center for the Humanities) an der Philosophischen Fakultät zu Köln beeinflusst hat (Kronenwett 2017).

Bibliographie

Andorfer, Peter (2015): *Forschung und Forschungsdaten in den Geisteswissenschaften: Zwischenbericht einer Interviewreihe*. DARIAH-DE working papers 10. Göttingen: GEODOC, Dokumenten- und Publikationsserver der Georg-August-Universität Göttingen <http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2015-10.pdf> [letzter Zugriff 30. August 2016].

Bauer, Bruno / Ferus, Andreas / Gorraiz, Juan et al. (2015): *Forschende und ihre Daten. Ergebnisse einer österreichweiten Befragung*. Report 2015, Version 1.2 <https://phaidra.univie.ac.at/view/o:407513> [letzter Zugriff 30. August 2016].

Burger, Marleen / Kindling, Maxi / Liebenau, Lisa et al. (2013): *Forschungsdatenmanagement an Hochschulen. Internationaler Überblick und Aspekte eines Konzepts für die Humboldt-Universität zu Berlin*. Version 1.1 vom 3. März 2013 <http://edoc.hu-berlin.de/oa/reports/reZ8xHXx2cLyc/>



PDF/28q8QGIHKwrRw.pdf [letzter Zugriff 30. August 2016].

Cologne Center for eHumanities (eds.) (2016): *Digital Humanities. Strukturen - Lehre - Forschung*. Universität zu Köln <http://ceeh.uni-koeln.de/broschue-digital-humanities-2016/> [letzter Zugriff 30. November 2016].

Data Center for the Humanities (2016a): Homepage, <http://dch.phil-fak.uni-koeln.de/> [letzter Zugriff 12. August 2016].

Data Center for the Humanities (2016b): Homepage, Unterseite "Umfrage Forschungsdaten 2016", <http://dch.phil-fak.uni-koeln.de/umfrage-2016.html> [letzter Zugriff 12. August 2016].

Drees, Bastian (2016): „Zukunft der Informationsinfrastrukturen: Das deutsche Bibliothekswesen im digitalen Zeitalter“, in: *Perspektive Bibliothek* 5.1: 25–48 urn:nbn:de:bsz:16-pb-313858.

DV-ISA (2016): *Umgang mit digitalen Daten in der Wissenschaft: Forschungsdatenmanagement in NRW. Eine erste Bestandsaufnahme*, 14. April 2016, Version 0.7 [Final] https://www.dh-nrw.de/fileadmin/dh-nrw/PDF/Veroeffentlichungen/DV-ISA-Bestandsaufnahme_FDM.pdf [letzter Zugriff 10. August 2016].

Fühles-Ubach, Simone (2013a): „Quantitative Befragungen“, in: Umlauf, Konrad / Fühles-Ubach, Simone / Seadle, Michael (eds.): *Handbuch Methoden der Bibliotheks- und Informationswissenschaft. Bibliotheks-, Benutzerforschung, Informationsanalyse*. Berlin: De Gruyter 96–113.

Fühles-Ubach, Simone (2013b): „Online-Befragungen“, in: Umlauf, Konrad / Fühles-Ubach, Simone / Seadle, Michael (eds.): *Handbuch Methoden der Bibliotheks- und Informationswissenschaft. Bibliotheks-, Benutzerforschung, Informationsanalyse*. Berlin: De Gruyter 114–127.

Fühles-Ubach, Simone / Umlauf, Konrad (2013): „Quantitative Methoden“, in: Umlauf, Konrad / Fühles-Ubach, Simone / Seadle, Michael (eds.): *Handbuch Methoden der Bibliotheks- und Informationswissenschaft. Bibliotheks-, Benutzerforschung, Informationsanalyse*. Berlin: De Gruyter 80–95.

forschungsdaten.org (2016): Homepage, Unterseite "Umfragen zum Umgang mit Forschungsdaten an wissenschaftlichen Institutionen", http://www.forschungsdaten.org/index.php/Umfragen_zum_Umgang_mit_Forschungsdaten_an_wissenschaftlichen_Institutionen [letzter Zugriff 29. August 2016].

Hochschulrektorenkonferenz (2014): *Management von Forschungsdaten als strategische Aufgabe der Hochschulleitungen*. Empfehlung der 16. HRK-Mitgliederversammlung am 13. Mai 2014 in Frankfurt am Main, https://www.hrk.de/uploads/tx_szconvention/HRK_Empfehlung_Forschungsdaten_13052014_01.pdf [letzter Zugriff 29. August 2016].

Hochschulrektorenkonferenz (2015): *Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien*. Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel, https://www.hrk.de/uploads/tx_szconvention/Empfehlung_Forschungsdatenmanagement_final_Stand_11.11.2015.pdf [letzter Zugriff 29. August 2016].

Kramer, Bernd (2014): „Datenflut an Unis: Forscher müssen teilen lernen“, in: *Spiegel Online*, 26. Februar 2014, <http://www.spiegel.de/unispiegel/jobundberuf/umgang-mit-daten-der-glaeserne-forscher-a-954958.html> [letzter Zugriff 01. August 2016].

Kronenwett, Simone (2017): *Forschungsdaten an der Philosophischen Fakultät der Universität zu Köln* (= Kölner Arbeitspapiere zur Bibliotheks- und Informationswissenschaft 78). Technische Hochschule Köln.

Kronenwett&Adolphs (2016): Homepage, <http://www.kronenwett-adolphs.com/de> [letzter Zugriff 12. August 2016].

Müller, Hendrik / Sedley, Aaron / Ferrall-Nunge, Elizabeth (2014): „Survey research in HCI“, in: Olson, Judith S. / Kellogg, Wendy A. (eds.): *Ways of Knowing in HCI*. New York: Springer 229–266.

Philosophische Fakultät (2016): Homepage, Unterseite "Fächergruppen", <http://phil-fak.uni-koeln.de/9785.html> [letzter Zugriff 12. August 2016].

Rat für Informationsinfrastrukturen (2016): *Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*. Göttingen, <http://www.rfii.de/?wpdmdl=1998> [letzter Zugriff 30. August 2016].

Sahle, Patrick / Kronenwett, Simone (2013): „Jenseits der Daten: Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner ‚Data Center for the Humanities‘“, in: *LIBREAS. Library Ideas* 23: 76–96 urn:nbn:de:kobv:11-100212726.

Schöpfel, Joachim / Prost, Hélène (2016): „Research data management in social sciences and humanities: A survey at the University of Lille (France)“, in: *LIBREAS. Library Ideas* 29: 98–112 urn:nbn:de:kobv:11-100238193.

Seadle, Michael (2013): „Entwicklung von Forschungsdesigns“, in: Umlauf, Konrad / Fühles-Ubach, Simone / Seadle, Michael (eds.): *Handbuch Methoden der Bibliotheks- und Informationswissenschaft. Bibliotheks-, Benutzerforschung, Informationsanalyse*. Berlin: De Gruyter 41–63.

Stäcker, Thomas (2015): „Noch einmal: Was sind geisteswissenschaftliche Forschungsdaten?“, in: *DHdBlog*, 06.12.2015, <http://dhd-blog.org/?p=5995> [letzter Zugriff 30. Juli 2016].

Universität zu Köln (2016): Homepage, Unterseite "Philosophische Fakultät", <http://phil-fak.uni-koeln.de/studieninteressierte.html?&L=0> [letzter Zugriff 30. Juli 2016].

Winkler-Nees, Stefan (2011): „Vorwort“, in: Büttner, Stephan / Hobohm, Hans-Christoph / Müller, Lars (eds.): *Handbuch Forschungsdatenmanagement*. Bad Honnef: Bock+Herchen Verlag 5–6.