

Forschungsdatenmanagement

von *Katrin Moeller*

Datenautorschaft als Voraussetzung genderübergreifender Nachwuchsförderung

Belastbare aktuelle Analysen zur Geschlechterverteilung im Bereich der digitalen Wissenschaften – vor allem in den neu entstehenden Arbeitsfeldern des Forschungsdatenmanagements – fehlen noch. Wer sich allerdings sowohl in den fachspezifischen wie überfachlichen Veranstaltungen der Communities zum Forschungsdatenmanagement umsieht, kann eine interessante Beobachtung machen: Bei den Zusammenkünften der MINT-Fächer lässt sich ein deutlich überproportionaler Anteil von jungen Frauen ausmachen, der sich dem Forschungsdatenmanagement zuwendet. In den Fällen, in denen die Geistes- und Kulturwissenschaften eher unter sich bleiben, ist der Frauenanteil hingegen deutlich geringer, ja scheint manchmal kaum existent zu sein. In der Zusammenkunft der Arbeitsgemeinschaft Datenzentren 2016 innerhalb der Digital Humanities[1] lag der Frauenanteil bei gerade einmal 26 % – eine angesichts des traditionell eher hohen Anteils von Studentinnen, Wissenschaftlerinnen und Professorinnen in den Sprach- und Kulturwissenschaften doch überraschende Beobachtung. Denn eigentlich müsste der Frauenanteil im Forschungsdatenmanagement von MINT- und Geistes- sowie Sozialwissenschaften genau umgekehrt zueinander verteilt sein. Immerhin stellen sie mit 42,0 % Professorinnen in den Sprach- und Kulturwissenschaften beziehungsweise mit 21,0 % in den MINT-Fächern jeweils das Maximum und das Minimum des Professorinnenanteils im Vergleich der verschiedenen Fachdisziplinen der Hochschulen Deutschlands, der bekanntlich mit durchschnittlich 22 % in Deutschland trotz aller Förderungen immer noch gering ausfällt.[2] Woran liegt es also, dass sich gerade in den Disziplinen mit einem geringeren Frauenanteil eher die weiblichen Wissenschaftlerinnen dem Forschungsdatenmanagement zuwenden, während das umgekehrt nicht der Fall ist?

Die Überlegungen verweisen auf zwei Antworten, die generell für die Zukunft der Digital Humanities von entscheidender Bedeutung sein dürften und deren Diskussion die Ausgestaltung und Akzeptanz des Datenmanagements in der Wissenschaftscommunity auch über den Genderaspekt hinaus positiv beeinflussen könnte.

1. Geschlechtsspezifische Präferenzen oder wohin gehören

Forschungsdatenzentren?

Eine naheliegende Antwort könnte darin bestehen, auf die geschlechtsspezifischen Präferenzen im Umgang mit digitalen Arbeitstechniken und Technologien zu verweisen. Frauen – so das Klischee – trauen sich aufgrund von Rollen- und Geschlechterzuschreibungen tendenziell eher weniger Kompetenzen im Bereich der digitalen Welt zu als Männer. Dieser Effekt zeige sich gerade in den Geistes- und Sozialwissenschaften, weil bereits die Fächerwahl schlechthin Prämissen für den Umgang mit digitalen Arbeitstechniken umfasse.

Gleichermaßen ältere[3] wie aktuellere Umfragen und Studien haben jedoch nachgewiesen, dass die Wirklichkeit wesentlich vielschichtiger ist und das Stereotyp der mangelnden Technikkompetenz eben auch nur ein Konstrukt unserer lang tradierten Geschlechterbilder repräsentiert. Bei alltäglichen computergestützten Arbeitsprozessen etwa – so die Ergebnisse einer Studie[4] an der Wissenschaftlichen Hochschule Lahr – zeigten Frauen und Männer keine wesentlichen Unterschiede. Den täglichen Umgang mit dem Internet oder mit Textverarbeitungsprogrammen beherrschten beide Geschlechter gleichermaßen souverän. Größere Unterschiede offenbarten sich hingegen dort, wo es sich um spezifische Arbeitstechniken und Kompetenzen handelte, deren Gebrauch die Überwindung gewisser Barrieren beinhaltet. Sich digitaler Methoden wie Netzwerkanalyse, XML-Skripting oder Geoinformation zu bedienen oder selbst kleine Programme zu schreiben gehörte offenbar nicht mehr zu den alltäglichen Fertigkeiten. Hier ließen sich nicht nur deutliche Unterschiede in den Handlungsmustern zwischen den Geschlechtern nachweisen, frühzeitig wurden zudem auch Unterschiede zwischen den eher kunst- und geisteswissenschaftlich ausgerichteten Disziplinen einerseits und den MINT-orientierten Fächern andererseits deutlich. Die Kompetenzunterschiede zwischen Informatikern und Geisteswissenschaftlern etwa fielen sehr viel größer aus als die zwischen den Geschlechtern, auch wenn die digitale Technik nicht zum Spezialgebiet des Informatikers gehört.[5]

Die beobachteten Phänomene haben viel mit dem Selbstverständnis der einzelnen Fachdisziplinen zu tun. Kenntnisse der Informatik und der Datenverarbeitung gehören zwar zu den Kernkompetenzen der Informatik, aber eben nicht unbedingt auch der Geisteswissenschaften. Dabei dürfte unter anderem auch die Angst, durch eine stärkere Hinwendung zu digitalen Arbeitstechniken und damit verwandten Themen eventuell den Anschluss an wichtige Entwicklungen im Rahmen des eigenen Faches zu verlieren, eine Rolle spielen. Und tatsächlich ist eine solche Angst nicht völlig unbegründet. Der Aufwand, den es braucht, um entsprechende Zusatzqualifikationen zu erlangen, dürfte in den Geistes- und Kulturwissenschaften heute noch deutlich höher zu veranschlagen sein als in den MINT-Fächern.

Eine zentrale Konsequenz dieser Erkenntnisse für das Forschungsdatenmanagement dürfte daher wohl in der Forderung nach einer

Veralltäglicung sowohl des fachspezifischen als auch des fächerübergreifenden Forschungsdatenmanagements bestehen. Der geschlechts- und fachneutrale Umgang mit Technologien des Datenmanagements sollte langfristig über Lehr- und Lernformen verstetigt werden, die geeignet sind, vorhandene Barrieren abzubauen und so letztlich auch die Akzeptanz in den weniger technikaffinen Fachdisziplinen zu erhöhen.

Genau an dieser Schnittstelle gibt es allerdings aufgrund der momentanen Förderpolitik massive Probleme. Gefördert wird derzeit nämlich vor allem ein institutioneller Aufbau von Forschungsdatenrepositorien an bestehenden, institutionell abgesicherten Einrichtungen wie Bibliotheken und Archiven. So plausibel ein solches Vorgehen in Hinblick auf die Verbesserung der zur Erschließung und Verfügbarmachung von Forschungsdaten erforderlichen Infrastrukturen und Kompetenzen ist, so schwerfällig nimmt es sich in der Forschungspraxis aus. Bibliotheken und Archive repräsentieren Institutionen, denen jeweils nur am Ende des Entstehungsprozesses von Forschungsdaten Aufgaben zufallen. Wirkungsvolles Datenmanagement greift jedoch vor allem, wenn es bereits bei der Konzeption eines Forschungsvorhabens oder Projektes mitbedacht wird. Da Archive und Bibliotheken bisher lediglich mit den Endergebnissen der Forschung betraut sind, fällt es diesen Einrichtungen häufig schwer, sich nun bereits zu Beginn von Forschungsprozessen aktiv einzubringen. Für den Aufbau nachhaltiger Forschungsdatenbestände reichen die bestehenden Maßnahmen daher meist nicht aus, weil erhebliche Potenziale bei der Kreierung von langfristig nutzbaren Datenkonvoluten verschenkt werden.

Die bisherigen Erfahrungen zeigen: Datenmanagement beginnt bereits bei der Antragsstellung und der ideellen Konzeption von Projekten. Fachliche Kompetenzen, die in Studiengänge einmünden könnten, entstehen nur durch die eigene Praxis. Forschungsdatenmanagement umfasst zudem auch immer Elemente der Operationalisierung und Strukturierung von Daten, die sich originär aus wissenschaftlichen Belangen speisen. Die Produktion von Forschungsdaten, in die Forschende heute mitunter bis zu 80% ihrer Arbeitszeit investieren, bildet ein Herzstück wissenschaftlicher Forschung - und kann eben deshalb auch nicht allein durch zwar wissenschaftsnahe, aber Forschungsdaten in erster Linie verwaltende und nicht produzierende Institutionen geleistet werden. Nicht zufällig entstehen Datenzentren momentan besonders häufig von unten, und zwar dort, wo wissenschaftlich der Bedarf an Datenspeicherung entsteht.[6] Insbesondere die Archive, aber auch viele Bibliotheken tun sich hingegen eher schwer damit, die ihnen zugeschriebene neue Rolle des Kompetenzvermittlers und Wissenschaftsorganisations auszufüllen. Erforderlich sind letztlich Zwitterinstitutionen, die sowohl Personal aus der Forschungspraxis als auch Fachkompetenzen zur sicheren Archivierung und Erschließung umfassen und die klassische Arbeitsteilung zwischen den Bereichen Forschung, Archiv und Bibliothek unterlaufen.

2. Berufschancen und die Frage der Nachwuchsförderung

Eine zweite Antwort auf die eingangs formulierte Frage könnte im Verweis auf ein Phänomen bestehen, das als Konsequenz der ungleichen Chancenverteilung zu beobachten ist: Absolvent_innen der MINT-Fächer besitzen nicht nur in puncto Berufseinstieg die besseren Chancen, sondern ihnen stehen auch außerhalb der Wissenschaft zahlreiche und attraktive Karrierewege offen.[7] Das legt die Vermutung nahe, das Forschungsdatenmanagement als Daueraufgabe nicht zuletzt deshalb gerade auf junge Frauen und Mütter attraktiv wirkt, weil sie sich davon langfristige Anstellungsverhältnisse und eine bessere Vereinbarkeit von Familie und Beruf erhoffen. In den Geistes- und Sozialwissenschaften, in denen der Berufseinstieg tendenziell schwerer fällt und Karrierewege außerhalb der Wissenschaft häufiger auf Positionen führen, die nicht dem Qualifikationsniveau der Absolvent_innen entsprechen, bietet das wissenschaftsnahe Forschungsdatenmanagement hingegen attraktive Berufsoptionen für beide Geschlechter. Darüber hinaus könnte auch die schärfere berufliche Konkurrenzsituation Relevanz für die eingangs festgestellte Genderverteilung besitzen.

Seziert man die Problematik weiter, so geben die genannten Beobachtungen Anlass, die Frage der wissenschaftlichen Reputation des Forschungsdatenmanagements zu diskutieren. Der jeweilige Stellenwert, den die langfristige Archivierung und vor allem die Publikation von Daten besitzt, ist in den einzelnen Fachdisziplinen sehr unterschiedlich ausgestaltet und kann hier nicht in seiner ganzen Vielfalt gewürdigt werden. So viel aber lässt sich sagen: Soll die Nachnutzung von kostspielig erhobenen Daten auf eine breitere Basis gestellt werden, sind intensive Überlegungen darüber notwendig, wie sich die Attraktivität des arbeitsintensiven Berufsfelds Forschungsdatenmanagement steigern lässt.

Überdies lassen sich zwei sehr verschiedene wissenschaftliche Bedürfnisse ermitteln, die letztlich auch mit den verschiedenen Rollen von Datengeber und Datennutzer verknüpft sind: Einerseits besteht ein hoher Bedarf an anonymen Daten, die möglichst ohne größere rechtliche und wissenschaftliche Barrieren frei genutzt und verschnitten werden können. Big-Data-Visionen leben nun einmal davon, große Datenmengen ungehindert nutzen und miteinander kombinieren zu können. Überall dort, wo Daten maschinell erzeugt werden, ist dieser unkomplizierte Zugang zu und Umgang mit Daten relativ unproblematisch. Die Hoffnungen, die sich seitens der Wissenschaft mit der Nutzung solcher Datenkonvolute verbinden, sind dabei nicht nur weitreichend, sondern häufig auch durchaus plausibel. Diese Art des Datenmanagements kollidiert allerdings mit den Bestimmungen des Urheberrechts und der Sichtbarmachung von Forschungsleistungen.

Auch Datenproduzent_innen zeigen das Bedürfnis, die eigenen wissenschaftlichen Leistungen auszuweisen und zu dokumentieren – insbesondere wenn die

betreffenden Daten anschließend anderen zur Nachnutzung überlassen werden. Noch stärker ausgeprägt ist das Bedürfnis, wenn die Datenproduktion mit einem hohen personellen und finanziellen Aufwand sowie mit der Erledigung originär wissenschaftlicher Aufgaben einhergeht, wie es beispielsweise in den Geschichtswissenschaften der Fall ist. Daten müssen hier nicht nur aus alten Handschriften transkribiert, sondern zudem auch inhaltlich erschlossen und wissenschaftlich annotiert werden. Eine umfassende Dokumentation und eine genaue Beschreibung des Prozesses der Datenerfassung erfordert zudem häufig erhebliche zeitliche Kapazitäten. Oft sind es genau diese Punkte, die einer effektiven Nachnutzung von Daten entgegenstehen.[8]

Halten wir fest: Die Datenerfassung ist häufig ein sehr aufwändiger und nicht selten auch kognitiv anspruchsvoller Prozess, der ungeachtet der Tatsache, dass in ihn bereits Kernelemente wissenschaftlicher Arbeit einfließen, gerade in den Geisteswissenschaften bislang kaum die Reputation genießt, die er verdient. Um das zu ändern braucht es – nicht zuletzt unter dem Blickwinkel von Chancengleichheit und Nachwuchsförderung – Mechanismen, die geeignet sind, die wissenschaftliche Reputation von Datenproduktionen zu erhöhen, wie dies unlängst etwa auch in den Leitlinien der DFG zum Forschungsdatenmanagement festgehalten wurde.[9] Einen wichtigen Schritt in diese Richtung könnte die gezielte Verbesserung der Möglichkeiten zur Datenpublikation im Rahmen von Forschungsdatenrepositorien darstellen. Hier bedarf es gezielter Anreize, um die Akzeptanz solcher digitaler Veröffentlichungen in der Wissenschaft zu erhöhen und die positiven Effekte kooperativer Strategien tatsächlich wirksam werden zu lassen.

Einstweilen jedenfalls genießen digitale Publikationen und Datenveröffentlichungen noch nicht den gleichen Stellenwert wie analog publizierte Forschungsergebnisse. So gaben bei einer Befragung von wissenschaftlichen Mitarbeiter_innen und Student_innen der Geistes- und Sozialwissenschaften der Martin-Luther-Universität Halle-Wittenberg mehr als 50 % an, digitale Veröffentlichungen besäßen nicht oder nur unter Vorbehalt die gleiche Reputation wie analoge Publikationen.[10] Ebenfalls fraglich bleibt, ob die vollständige Verwendung von Forschungsdaten durch nachnutzende Projekte durch den Akt der Zitierung eine ausreichende Würdigung erfährt. Unabhängig davon, dass sich diese Praxis unter juristischen Gesichtspunkten nur schwer mit dem bestehenden Urheberrecht vereinbaren lässt, bleiben die formalen Vorgaben für das Zitieren deutlich hinter den entsprechenden Maßstäben der Fachwissenschaft im analogen Bereich zurück. Man stelle sich vor, ein Wissenschaftler würde Artikel oder Monografien im Volltext zitieren, und auf deren Autor lediglich mit einer Fußnote verweisen! Für Datenpublikationen werden solche Modelle jedoch projiziert.

Fasst man diese widerstreitenden Bedürfnisse der neuen kooperativen Strategien des Data-Sharings zusammen, liegt der Schluss nahe, dass hier neue Strategien der wissenschaftlichen Reputation erforderlich sind, um den Visionen von Big-Data näher zu kommen. In einigen Wissenschaftsbereichen hat sich daher ein dezidiertes System der multiplen Autorenschaft entwickelt, um den oben beschriebenen

Bedürfnissen gerecht zu werden. Datenproduzenten tauchen dabei in den betreffenden wissenschaftlichen Veröffentlichungen als Zweit- beziehungsweise Coautoren auf. Das macht allerdings den Einsatz von Big-Data auch komplizierter, denn da nur derjenige Autor im Sinne des bestehenden Urheberrechts sein kann, der dieser Autorschaft auch prinzipiell zugestimmt hat, setzt das Prinzip der Zweitbeziehungsweise Coauthorschaft die Erreichbarkeit eines Datenautors voraus. Strukturell wird das Problem der Autor- beziehungsweise Urheberschaft bisher noch nicht von den in Entstehung begriffenen Forschungsdatenrepositorien erkannt und technisch umgesetzt. In Forschungslandschaften, in denen bis heute Monografien der vorherrschende Publikationstypus sind, taugen solche Lösungen auch nur bedingt.

Ebenfalls klärungsbedürftig ist die Frage, ob der Datengeber den Ergebnissen einer auf der Nachnutzung seiner Daten basierenden wissenschaftlichen Studie widersprechen darf oder nicht. Ein Schritt zur Lösung dieses Problems könnte in der Entwicklung eines neuen Typs von Urheberschaft liegen, der textbasierte Forschungsergebnisse von datenbasierten Forschungsleistungen entkoppelt. Mittels einer separat von der Textautorschaft anzugebenden Datenautorschaft ließen sich die Leistungen und Deutungen der beteiligten Autor_innen voneinander abgrenzen und damit auch Urheberschaften und Verantwortlichkeiten besser in Einklang bringen. Forschungsleistungen wären auf diese Weise klar nachweisbar und letztlich angemessener darstellbar als in dem bisherigen System der Coauthorschaft. Letztlich dürfte das der richtige Weg sein, um Forschungsdatenmanagement in der universitären Praxis attraktiver zu machen. Er führte zudem in kleinen Schritten auch zu einer größeren Attraktivität des Forschungsdatenmanagements für Frauen und Männer. So ließe sich vermutlich auch verhindern, dass die im Entstehen befindliche Branche des Forschungsdatenmanagements Sogkräfte vor allem auf ein Geschlecht entfaltet. Denn letztlich sagen die beobachtbaren Geschlechterproportionen im Forschungsdatenmanagement sehr viel mehr über Arbeitsmarktchancen und Berufsplatzierungsstrategien von weiblichen und männlichen Wissenschaftlerinnen aus, als sie über deren jeweilige Kompetenzen und Fähigkeiten verraten.

Fußnoten

[1] Es handelt sich hier um den Dachverband Digital Humanities im deutschsprachigen Raum, auf dessen Jahrestagung im März 2016 in Leipzig auch ein Treffen der Arbeitsgruppen stattfand.

[2] Statistisches Bundesamt (Hrsg.), Frauenanteil in der Professorenschaft in Deutschland von 1999 bis 2014, in: Statista - Das Statistik-Portal.

[3] Uwe Rutenfranz, Wissenschaft im Informationszeitalter. Zur Bedeutung des Mediums Computer für das Kommunikationssystem Wissenschaft, Opladen 1997, S. 214f.

[4] Bernd Remmele, Degendering von E-Learning. Das Forschungsprojekt „Das aufwendige Geschlecht“ an der WHL, in: AKAD. Das Hochschulmagazin 23 (2012), S.

10-12, hier S. 12.

[5] Rutenfranz, Wissenschaft im Informationszeitalter, S. 227.

[6] Beispielhaft genannt seien hier das Historische Datenzentrum Sachsen-Anhalt oder das Kölner Datenzentrum für die Geisteswissenschaften.

[7] Bundesagentur für Arbeit, Gute Bildung - gute Chancen. Der Arbeitsmarkt für Akademikerinnen und Akademiker, Nürnberg 2016, S. 6 und S. 101ff.; Bundesagentur für Arbeit, Der Arbeitsmarkt für IT-Fachleute in Deutschland, Nürnberg 2016, S. 9 und S. 14.

[8] Diane Opitz/Reiner Maier, Erfahrungen mit der Sekundärnutzung von qualitativem Datenmaterial. Erste Ergebnisse einer schriftlichen Befragung im Rahmen der Machbarkeitsstudie zur Archivierung und Sekundärnutzung qualitativer Interviewdaten, in: FQS Forum: Qualitative Social Research 6 (2005), S. 11.

[9] DFG, Leitlinien zum Umgang mit Forschungsdaten, Bonn 2015, S. 2.

[10] Katrin Moeller/Kerstin Völkl, Genderspezifische Aspekte des Forschungsdatenmanagements? Ergebnisse einer Umfrage zum Forschungsdatenmanagement, Projektbericht, Halle 2015.