

1. Jahrestagung der Digital Humanities im deutschsprachigen Raum (DHd 2014)

Universität Passau · 25.-28. März 2014

Thema 2: Digitale Infrastrukturen für die Geisteswissenschaften

Simone Kronenwett, Cologne Center for eHumanities, Universität zu Köln

Poster-Abstract

Bei fast allen Tagungen und Konferenzen zum Thema DH ist die Frage nach einer sinnvollen Weiterentwicklung der *digitalen Infrastrukturen für die Geisteswissenschaften* ein zentraler Diskussionspunkt.¹ Zwar besteht größtenteils Konsens darin, dass die Etablierung dedizierter digitaler Infrastrukturen für die Geisteswissenschaften im deutschsprachigen Raum derzeit führend ist. Auch wird der damit verknüpfte Professionalisierungsprozess der DH durch große laufende Infrastrukturprojekte wie DARIAH und CLARIN weiter vorangetrieben und ausgebaut.² Gleichzeitig eröffnen sich damit weiterführende Fragen, die es zu beantworten gilt: Wie können stabile und damit langfristig finanzierte Infrastrukturen in den DH geschaffen und gesichert werden?³ Und wie können bestehende Desiderata beispielsweise hinsichtlich geisteswissenschaftlicher Forschungsdaten und Ressourcen geschlossen werden? Denn die Sicherung und die langfristige Zugänglichkeit von Forschungsdaten und Projektergebnissen ist auch ein maßgebliches Bewilligungskriterium aller Förderinstitutionen bei Projektanträgen im Sinne guter wissenschaftlicher Praxis.⁴ Die Wissenschaftsorganisationen haben deshalb in diesem Zusammenhang in den Strategiepapieren der letzten Jahre den Aufbau von entsprechenden Datenzentren empfohlen.⁵

Data Center for the Humanities

Um der konkreten Problematik einer dauerhaften Bereitstellung geisteswissenschaftlicher Forschungsdaten aktiv zu begegnen und damit die oben genannten Desiderata zu schließen, wurde Ende 2012 von der Philosophischen Fakultät der Universität zu Köln das Data Center for the Humanities (DCH) gegründet, dessen aktueller Entwicklungsstatus im Rahmen der Posterpräsentation vorgestellt wird.⁶

Das vom Cologne Center for eHumanities (CCEH) organisierte Datenzentrum zielt nicht nur darauf ab, die digitale Langzeitarchivierung von Forschungsdaten zu gewährleisten, sondern wird auch Forschungsdatenmanagement betreiben, dauerhafte Präsentationssysteme hosten und betreuen,

¹ So jüngst auf der *Closing Panel Discussion* der Herrenhausen-Konferenz der VolkswagenStiftung „(Digital) Humanities Revisited – Challenges and Opportunities in the Digital Age“, 5.-7. Dezember 2013, Hannover, <http://www.volkswagenstiftung.de/digitalhumanities>.

² Vgl. DARIAH-DE (Digital Research Infrastructure for the Arts and Humanities), <https://portal-de.dariah.eu/>; CLARIN-D (Common Language Resources and Technology Infrastructure), <http://de.clarin.eu/de/home>; vgl. auch BMBF (Hrsg.): *Forschungsinfrastrukturen für die Geistes- und Sozialwissenschaften*, Mülheim an der Ruhr, Stand: Februar 2013, S. 27ff., http://www.bmbf.de/pub/forschungsinfrastrukturen_geistes_und_sozialwissenschaften.pdf.

³ Als Vorbild wird immer wieder das naturwissenschaftliche Forschungsinfrastrukturprojekt CERN (Conseil Européen pour la Recherche Nucléaire, <http://home.web.cern.ch/>) genannt, vgl. BMBF: *Forschungsinfrastrukturen*, S. 7.

⁴ Vgl. Deutsche Forschungsgemeinschaft (DFG): *Vorschläge zur Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“*, Denkschrift, Weinheim 1998, http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf.

⁵ Vgl. Wissenschaftsrat: *Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020*, 13.07.2012, Berlin, S. 11, <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>.

⁶ Vgl. Data Center for the Humanities (DCH), www.dch.uni-koeln.de.

Services zur Datenkommunikation aufbauen und Werkzeuge zur Arbeit mit den Daten vorhalten.⁷ Die Zusammenarbeit und Kooperation mit verschiedenen Akteuren (wie z.B. der Universitäts- und Stadtbibliothek Köln oder dem Regionalen Rechenzentrum der Universität zu Köln) mit ihren jeweiligen Kompetenzen ist dabei für das Datenzentrum grundlegend. Im Mittelpunkt stehen allerdings die FachwissenschaftlerInnen und ihre Forschungsprojekte, deren Inhalte nachhaltig gesichert und dauerhaft zugänglich gemacht werden müssen.

Geisteswissenschaftliche Forschungsdaten

Unter Forschungsdaten werden gemeinhin sämtliche Daten verstanden, die als Grundlage für die Forschung dienen.⁸ Aufgrund der großen Methodenvielfalt in den geisteswissenschaftlichen Disziplinen zeichnen sich diese Daten besonders durch ihre Heterogenität aus.⁹ Um den Mehrwert der Daten in ihrer Kontextualität und Diskursivität zu wahren und nachnutzen zu können, kann häufig nicht einfach zwischen „Primärdaten“, die bloß zu archivieren wären, und den oft komplexen und individuellen Systemen der Präsentation und Benutzung von „Ergebnisdaten“ unterschieden werden. Neben der Langzeitarchivierung von Daten stehen deshalb die Pflege und Betreuung dieser komplexen Systeme im Fokus des Datenzentrums. Damit Forschungsdaten auch weiterhin vernetzt und kontextualisiert dargestellt und werden können, bedarf es eines Forschungsdatenmanagements und einer Begleitung der Forschung, die den Spezifika geisteswissenschaftlicher Fragestellungen und Methoden sowie der Präsentation ihrer Ergebnisse gerecht werden.

Geisteswissenschaftlicher Forschungsprozess

Geisteswissenschaftliche Forschung vollzieht sich in zyklischen Prozessen. In der Operationalisierung und Beantwortung von Forschungsfragen entstehen Daten, die selbst wieder das Ausgangsmaterial für neue Fragestellungen sein können. Damit Informationen dauerhaft nutzbar bleiben und Wissen nicht verloren geht, bedarf es eines professionellen Datenmanagements, das die Forschung begleitet und ihre Ergebnisse dokumentiert, mit Metadaten anreichert, archiviert, an Schnittstellen bereitstellt, präsentiert und für den andauernden Zugriff pflegt.

DCH-Schichtenmodell

Im Zentrum der DCH-Präsentation steht deshalb neben dem Lebenszyklus geisteswissenschaftlicher Forschungsdaten das Schichtenmodell des Datenzentrums. Archivierung, Bereitstellung, Adressierbarkeit, Präsentation und Nutzung von Diensten und Werkzeugen bauen hier aufeinander auf. Die modulare Struktur des Datenzentrums entspricht dabei den unterschiedlichen Anforderungen aus den verschiedenen Forschungsprojekten und gewährleistet insgesamt (1.) eine langfristige Sicherung und dauerhafte Bereitstellung, (2.) die allgemeine Zugänglichkeit und Nutzbarkeit und (3.) eine erhöhte Sichtbarkeit der Forschungsdaten und Ressourcen, die zugleich eine bessere Vernetzung der Projekte und Daten bedeutet und die Grundlage für die zukünftige Nutzung in Forschung und Lehre bildet.

⁷ Vgl. Cologne Center for eHumanities (CCEH), www.cceh.uni-koeln.de.

⁸ Vgl. z.B. Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen: Grundsätze zum Umgang mit Forschungsdaten, 04. Juni 2010, <http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsaeetze/>.

⁹ Vgl. Jasmin Hügi, René Schneider: Digitale Forschungsinfrastrukturen in den Geistes- und Geschichtswissenschaften, Genf 2013, S. i, http://www.infoclio.ch/sites/default/files/standard_page/studie_forschungsinfrastrukturen_small.pdf.