

“Kann man da eben mal was eintragen und visualisieren?” Digitaler Praxistest für die DARIAH-DE-Infrastruktur

Klaffki, Lisa

klaffki@hab.de

Herzog August Bibliothek Wolfenbüttel, Deutschland

Steyer, Timo

steyer@hab.de

Herzog August Bibliothek Wolfenbüttel, Deutschland;
Forschungsverbund Marbach Weimar Wolfenbüttel,
Deutschland

Die Ausgangslage: Diskrepanz zwischen Anspruch und Wirklichkeit

In den letzten Jahren wurden vermehrt digitale Services und Ressourcen entwickelt, die EinzelforscherInnen, aber auch Projektgruppen in der Arbeit mit digitalen Methoden, der Generierung, dem Forschungsdatenmanagement und in toto bei der Transformation von Arbeitsprozessen in das digitale Medium unterstützen sollen. Signifikante Bestandteile des Research Data Lifecycles können bereits über digitale Angebote abgedeckt werden (Puhl et. al. 2015). Ein wesentliches Ziel, welches mit dem Aufbau der digitalen Infrastrukturen verfolgt wird, ist, dass auch diejenigen Forschungsvorhaben davon profitieren sollen, in denen keine oder nur geringe technischen Kenntnisse oder eine DH-Unterstützung vorhanden sind. Aktuell besteht zwischen Anspruch und Wirklichkeit aber noch eine Diskrepanz: Auf der einen Seite scheint mittlerweile die Mehrheit der WissenschaftlerInnen den Nutzen von Normdaten, Integration ihrer Forschungsdaten in Suchmaschinen oder die Visualisierungen ihrer Daten in Forschungsumgebungen zu erkennen. Auf der anderen Seite steht aber die Kritik, dass die oftmals propagierte einfache und intuitive Nutzung der digitalen Angebote nicht vorhanden ist. Der Mehrwert, den die Services den eher auf ein analoges Projektergebnis ausgerichteten Projekten bieten, steht in keinem effizienten Verhältnis von Aufwand und Ertrag. Auch Schulungsvideos oder Workshops sind häufig zu unspezifisch oder nicht auf die vorhandenen Daten anwendbar.

Der Usecase: ein frühneuzeitlicher Auktionskatalog

Als Usecase dient das Forschungsprojekt “Erschließung frühneuzeitlicher Auktionskataloge” des Forschungsverbundes Marbach Weimar Wolfenbüttel (MWW) (Autorenbibliotheken). Die aus einem Auktionskatalog von 1670 rekonstruierte Gelehrtenbibliothek gehörte dem Chiliasten Benedikt Bahnsen (gest. 1669). Der aus Norddeutschland stammende und nach Amsterdam emigrierte Bahnsen war als Verleger, Buchhändler und Bücheragent, Mathematiker und Rechenmeister tätig. In diesem Projekt wurden in einer Excel-Tabelle alle Losnummern, d.h. Titel des Katalogs, erfasst. Die Tiefenerschließung umfasst die bibliographischen Angaben der verzeichneten Bücher, Normdaten zu Personen und Werken, Geodaten für die Druckorte sowie Links zu Digitalisaten und in Nachweissysteme. Mit der Erschließung wird dem Alleinstellungsmerkmal dieser Bibliothek Rechnung getragen, bei der es sich um das umfangreichste Einzelrepositorium für nonkonformes und heterodoxes Schrifttum handeln dürfte (Beyer et. al. 2017, 43–70).

Zwar verfügt das Projekt durch die Zusammenarbeit mit den DH-MitarbeiterInnen des Forschungsverbundes über eine eigene Datenbank und eine Präsentation im WWW (Auktionskataloge). Aber begonnen hatte das Projekt ohne deren Unterstützung und genau die in dieser Phase in einer Exceltabelle erhobenen Daten sollen für den Usecase mit DARIAH-DE verwendet werden.

Die Werkzeuge: DARIAH-DE- Dienste

Innerhalb von DARIAH-DE, einem digitalen Forschungsinfrastrukturprojekt für die Geistes- und Kulturwissenschaften, wurde eine *Data Federation Architecture* (DFA) entwickelt (Gradl / Henrich / Plutte 2015; Gradl / Henrich 2016). Unter diesem Begriff sind mehrere modulare Komponenten gebündelt, die für sich alleine oder im Zusammenspiel genutzt werden können. Davon nutzt der hier skizzierte Workflow für den Usecase das jüngst in die Produktivphase übergegangene *DARIAH-DE-Repository* zum Speichern und persistenten Adressieren der Forschungsdaten, die *Collection Registry* zum Verzeichnen der Datensammlung mitsamt Schnittstelle, um die Daten in die *Generic Search* zu integrieren und wiederauffindbar und somit nachnutzbar zu machen.

Für viele der in den verschiedenen Schritten eines Data Research Lifecycles anfallenden Aufgaben können also Komponenten der DFA zum Einsatz kommen (vgl. Abb. 1). Die DFA deckt aber nicht alle Schritte eines Forschungsprozesses und damit letztlich auch des Research Data Lifecycles ab. So bleibt der Abschnitt der Analyse

oder Visualisierung offen, allerdings gibt in den digitalen Geisteswissenschaften hinreichend andere Tools, die diese Lücke zielgerichtet füllen können, zumal die häufig projektspezifischen Anforderungen hier kaum von einer generischen Lösung erfüllt werden können.

Für den Praxistest wird zur Visualisierung der räumlich-zeitlichen Daten ein ebenfalls aus dem DARIAH-DE-Projekt stammendes Tool verwendet, der *Geo-Browser* (Kollatz 2016).

Der Test: Von der grauen Theorie in die blaue DARIAH-DE-Praxis

Der Diskrepanz zwischen Wunsch und Wirklichkeit, zwischen Sinnhaftigkeit und Aufwand möchte sich diese Postereinreichung annehmen und überprüfen, wie exemplarische Forschungsdaten aus einem „analogen“, nicht von vornherein mit einer DH-Komponente geplanten Projekt in verschiedene Komponenten der Infrastruktur von DARIAH-DE integriert werden können.

Ziel ist es, dabei den Aufwand und die erforderlichen Kenntnisse zu evaluieren, die eine erfolgreiche Integration der Forschungsdaten bedingen. Die DARIAH-DE-Dienste eignen sich besonders für diesen Test, da sich die DARIAH-DE-Angebote ausdrücklich an EinzelforscherInnen richten (DARIAH-DE in Kürze).

Zwar handelt es sich nicht um eine belastbare empirische Studie, doch der mit den Projektbeteiligten durchgeführte Usecase hat dennoch aufgrund seines prototypischen Charakters richtungsweisende Bedeutung.

Gradl, Tobias / Henrich, Andreas (2016): “Die DARIAH-DE Föderationsarchitektur. Datenintegration im Spannungsfeld forschungsspezifischer und domänenübergreifender Anforderungen”, in: Neuroth, Heike et al. (Hrsg.): *Bibliothek – Forschung und Praxis* 40 (2): 222–228. DOI: <https://doi.org/10.1515/bfp-2016-0027>.

Kollatz, Thomas (2016): “Raum-Zeit-Analysen mit Geo-Browser und Datasheet-Editor”, in: Neuroth, Heike et al. (Hrsg.): *Bibliothek – Forschung und Praxis* 40 (2): 229–233. <https://doi.org/10.1515/bfp-2016-0032>.

Puhl, Johanna et. al. (2015): Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften (= DARIAH-DE Working Papers Nr. 11). Göttingen: DARIAH-DE. URN: <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-4-4>.

Bibliographie

Auktionskataloge: <http://dev.hab.de/auktionskataloge> [letzter Zugriff 11.09.2017].

Autorenbibliotheken: <http://www.mww-forschung.de/forschungsprojekte/autorenbibliotheken> [letzter Zugriff 11.09.2017].

Beyer, Hartmut et.al. (2017): “Bibliotheken im Buch: Die Erschließung von privaten Büchersammlungen der Frühneuzeit über Auktionskataloge”, in: Busch, Hannah / Fischer, Franz und Sahle, Patrick (Hrsg.): *Kodikologie und Paläographie im digitalen Zeitalter 4 (Codicology and Palaeography in the Digital Age)*. Norderstedt: Books on Demand 43–70.

DARIAH-DE in Kürze: <https://de.dariah.eu/dariah-de-in-kurze> [letzter Zugriff 11.09.2017].

Gradl, Tobias / Henrich, Andreas / Plutte, Christoph (2015): “Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen”, in: Baum, Constanze / Stäcker, Thomas (Hrsg.): *Grenzen und Möglichkeiten der Digital Humanities* (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). text/html Format. DOI: http://dx.doi.org/10.17175/sb001_020.