

Nachhaltigkeit durch Zusammenschluss: Die DARIAH Data Re-Use Charter

Baillet, Anne

anne.baillet@gmail.com
Centre Marc Bloch, Deutschland

Busch, Anna

anna_busch@gmx.de
Forschungsverbund Weimar Marbach Wolfenbüttel

Puren, Marie

marie.puren@inria.fr
Inria Paris

Mertens, Mike

mike.mertens@dariah.eu
Goettingen Center for Digital Humanities

Romary, Laurent

laurent.romary@inria.fr
Inria Paris

Bedarf und Forschungsstand

Ausgangspunkt der Charter ist die Feststellung, dass es bis dato keinen klaren Rahmen gibt, der die wissenschaftlichen Weiternutzungsbedingungen von digitalen Daten regelt, die auf Beständen von Kulturerbeinstitutionen basieren. Einige allgemeine Richtlinien wurden entwickelt, die allerdings nicht spezifisch die Interaktion zwischen digitalen Kulturerbedaten und Forschung in den Blick nehmen, sondern breiter angelegt sind. So etwa die Data Re-Use Models DANS oder die UNESCO Guidelines for the preservation of digital heritage (die wohlgerne nicht ausschliesslich Kulturerbedaten in den Blick nehmen). Andere Initiativen haben die Evaluation der digitalen Best-Practices in den Mittelpunkt gestellt, so etwa der Data Seal of Approval oder die Richtlinien zur Certification and Assessment of Digital Repositories des Center for Research Libraries. Im Bereich des Forschungsökosystems selbst jedoch kann in diesem Sinne hauptsächlich auf den CERN Code of Conduct hingewiesen werden, der zwar wenig konkrete Verpflichtungen definiert, dafür aber den Schwerpunkt auf eine Ethik der wissenschaftlichen Zusammenarbeit legt, dessen Geist ebenfalls in der hier dargestellten Charter von wesentlicher Bedeutung ist.

Der Mangel eines klaren Rahmens für die Zusammenarbeit an und mit digitalen Kulturerbedaten macht sich für alle beteiligten Akteure im Alltag bemerkbar: Forscher und Forscherinnen, die vor einem Scan stehen, ohne zu wissen, wie sie diesen weiterverwenden und zitieren dürfen; Kulturerbeeinrichtungen, die ihre Metadatensätze mit den entstehenden Forschungsarbeiten verknüpfen möchten; Equipments, die Anfragen für die gleichen Artefakte immer wieder bekommen; Datenzentren, die der Schnittstelle zwischen Kulturerbeeinrichtungen und Forschung fernbleiben bzw. von der Verwaltung der unterschiedlichen Rechte der von ihnen gehosteten Daten herausgefordert sind; Forschungseinrichtungen, denen Strukturen fehlen, um ihren Mitarbeitern und Mitarbeiterinnen good-practice-Empfehlungen an die Hand zu geben. Der Mangel an definitorischer Schärfe des betroffenen Informationsaustauschs steht im Gegensatz sowohl zu dem damit für alle Beteiligten zusammenhängenden Bedarf als auch zu dem offensichtlichen Vorteil, den ein solcher für das ganze Ökosystem repräsentiert. Nachhaltig kann ein solcher Austausch nur dann werden, wenn neben den Datenproduzenten (Forscher/innen, Kulturerbeeinrichtungen, Equipments) auch Datenzentren daran beteiligt sind, wie es hier vorgesehen ist. Einige Einrichtungen übernehmen zwar mehrere dieser Funktionen, aber die Kommunikation zwischen den entsprechenden Abteilungen ist auch dort nicht immer optimal.

Nicht zuletzt für die Weiterentwicklung der Digital Humanities handelt es sich bei dem hier geschilderten Zusammenhang um ein zentrales Anliegen, bilden diese Daten ja die Grundlage für eine beachtliche Reihe geisteswissenschaftlicher Forschungen etwa in den Bereichen der Archäologie, der Kunstgeschichte, der Musikwissenschaft, aber auch u.a. der Literaturgeschichte und der Editionswissenschaft. Selbst vom Standpunkt nicht-historischer Fächer ist es auf Dauer von Vorteil, wenn für digitale Daten aller Art hinsichtlich der Formate und Standards Rücksprache gehalten wird, wenn Equipments die bereits durchgeführten Scanarbeiten transparent machen, wenn auf einen Blick klar werden kann, zu welchen Kulturerbebeständen in welchen Datenzentren Arbeiten vorhanden sind. Die DARIAH Data Re-Use Charter hat zum Ziel, die Transaktionen zwischen all den Akteuren, die an der wissenschaftlichen Arbeit mit den digitalen Daten, die auf Kulturerbedaten basieren, Interesse haben, einfacher, transparenter und nachhaltiger zu machen, sodass alle daraus einen Nutzen für ihre eigene Arbeit ziehen können.

Grundprinzipien

Die DARIAH Data Re-Use Charter ist eine Interaktionsplattform für alle an der wissenschaftlichen Nutzung digitaler Daten von Kulturerbeeinrichtungen interessierten Akteure. Die Charter stellt diesen

Textbausteine zur Verfügung, damit sie die Bedingungen ihrer Zusammenarbeit gemeinsam definieren können. Auf diesem Weg kann eine Kulturerbeinstitution die Nutzungsbedingungen ihres Gesamtbestands, aber auch von Sondersammlungen unterschiedlich beschreiben. Eine Universität kann ihre Befürwortung der Open Access-Prinzipien deklarieren und ihren Mitarbeiter/innen empfehlen, ihre Kooperationen mit Kulturerbeeinrichtungen in diesem Sinne zu konzipieren. Wie in diesen zwei Beispielen liefert die Charter ausformulierte Bausteine für die Zusammensetzung einer Kooperationsvereinbarung, an der ebenfalls Infrastruktureinrichtungen beteiligt sind, die mit der nachhaltigen Sicherung der Primär- bzw. Sekundärdaten betraut sind. Auf diese Art und Weise werden Empfehlungen in folgenden Bereichen formuliert (wobei mehrere Formulierungen pro Bereich zur Verfügung stehen sowie jeweils ein freies Textfeld):

Zugang zu Metadaten, Texten, Bildern (beispielsweise im Fall einer zu edierenden Handschrift: Archivmetadaten, Transkription und Annotation, Scan, die dann auch explizit miteinander verlinkt werden)

- Lizenzierung der Inhalte (mit Verweis auf weiterführende, informierende Dokumentation)
- Formate und Standards (ebenso)
- Anreicherungen; Verknüpfung der wissenschaftlichen Anreicherungen der Kulturerbedaten mit den Metadaten
- Streuung sowohl der Kulturerbedaten als auch der Anreicherungen
- Qualitätssicherung bei allen beteiligten Akteuren

Darüber hinaus nennt jede Einrichtung den/die einschlägige(n) Ansprechpartner/in, damit eine Kommunikation erleichtert wird.

Neben den technischen Aspekten, bei denen mit Sicherheit Aufklärungsbedarf besteht, geht es auch, wenn nicht vorrangig, darum, eine digitale Kooperationsethik zu fördern, die wie im CERN Code of Conduct auf dem Respekt vor dem Werk anderer, der guten Zusammenarbeit, der Förderung von Kooperation und der Offenheit gegenüber der Öffentlichkeit beruht.

Umsetzung

Die Charter wird die Form eines Webinterfaces annehmen, auf dem sich die Akteure in ihrer jeweiligen Funktion (Forscher/in, Kulturerbeinstitution, Datenzentrum, Equipment, Forschungseinrichtung) registrieren lassen können und von dort ausgehend unter den ihnen zur Verfügung stehenden Optionen diejenigen aussuchen können, zu denen sie sich bekennen möchten. Im Vortrag wird nach einer Einführung zum Grundgedanken der Charter spezieller auf den Teil des Interfaces eingegangen, das dem Forscher/der Forscherin gewidmet ist.

Als registrierter Forscher/registrierte Forscherin soll man sich in drei Bereichen positionieren, die jeweils entweder im öffentlichen Profil oder im privaten Bereich (nur für die eigens ausgewählten Kooperationspartner - Kulturerbeinstitution, Datenzentrum, Equipment - zugänglich) erscheinen.

Der erste Bereich betrifft den Zugang zu den Daten. Dort verpflichtet sich der Forscher dazu, den Anforderungen der Kulturerbeinstitutionen zu folgen, was die Verwendung (insbes. Zitierweise) der digitalen Kulturerbedaten angeht. Hier kann der Forscher ebenfalls die Bestände anklicken, für die er sich interessiert und diese Information auch publik machen oder nur den betroffenen Institutionen zugänglich machen.

Der zweite Bereich betrifft die Streuung der vom Forscher auf der Grundlage der Kulturerbedaten produzierten Sekundärdaten oder angereicherten Metadaten. An dieser Stelle kann der Forscher die Lizenzen nennen, die er bevorzugt (weiterführende Dokumentation zu diesem Thema wird anklickbar gemacht). Eine Identifikation mittels einer ORCID-Nummer, die Referenzierung einer Id-HAL oder Vergleichbares können ebenfalls an dieser Stelle angegeben werden.

Im dritten Bereich geht es um den Umgang mit den anderen Charter-Partnern: Best-Practices wie die systematische Nennung der Kooperationspartner oder die explizite Nennung der Art und Weise, wie man selbst zitiert werden möchte, werden dort deklariert.

Diese drei Bereiche, die die Nutzung von Primärdaten, die Produktion- und Streuung von Sekundärdaten und die allgemeine Kooperationsethik bedingen und definieren, machen die Grundlage des Forscherprofils aus. Von dort ausgehend kann er dann die Einrichtungen, Bestände und Dienstleistungen recherchieren, mit denen er zusammenarbeiten möchte.

Die Charter hat somit eine doppelte Dimension: die einer Information- und Selbstpositionierungsplattform und die eines sozialen Netzwerkes, wobei auf die Balance zwischen Transparenz und Respekt der Privatsphäre stets geachtet wird ;

Perspektiven

Die DARIAH Data Re-Use Charter wird zwischen Herbst 2016 und Frühjahr 2017 soweit entwickelt sein, dass ein Soft Launch im Frühjahr 2017 stattfinden kann. Das Kernteam arbeitet sowohl an der Entwicklung der Webseite als auch an der Einholung von Feedback interessierter Akteure, um dieses in der Entwicklung des Interfaces im Allgemeinen und der relevanten Bausteine im Besonderen zu integrieren. Sie wurde bereits in ihren Ansätzen auf Konferenzen dargestellt und wird im Herbst in ähnlichen Kontexten vorgestellt, damit weiteres Feedback eingeholt werden kann. Darüber hinaus findet im November 2016 in Berlin eine dedizierte Arbeitssitzung statt sowie eine Woche später eine entsprechende in Paris und im Januar eine in Rom - womöglich können über den Winter

ebenfalls in anderen EU-Ländern entsprechende Sitzungen stattfinden. Bis zur DHd wird das Interface als das Ergebnis dieser breitgefächerten Konsultation demonstrierbar sein und kurz vor dem Launch stehen. Das bereits signalisierte Interesse zahlreicher Einrichtungen und Forscher/innen lässt vermuten, dass die Vernetzungsfunktion der Plattform rasch Konturen gewinnen wird.

Ausgerechnet dieser Aspekt gilt es mit Blick auf Nachhaltigkeit zu unterstreichen: Wenn Equipments systematischer zitiert werden, wenn Datenzentren expliziter an der Schnittstelle zwischen Kulturerbeinstitutionen und Forschung arbeiten können, wenn Forscher/innen ohne akademische Anbindung - und es gibt immer mehr prekäre Wissenschaftler/innen, die nichtsdestotrotz forschend tätig sind - einen Kooperationsraum finden können, dann ist für das gesamte Forschungsökosystem viel gewonnen und ein entscheidender Beitrag zur Nachhaltigkeit von digitalen Primär- und Sekundärkulturerbedaten geleistet.