

## Methodenworkshop „Metadaten in den historischen Wissenschaften“

**Veranstalter:** Leibniz-Institut für Europäische Geschichte Mainz; DARIAH-DE: Aufbau von Forschungsinfrastrukturen für die eHumanities

**Datum, Ort:** 24.09.2012-25.09.2012,

**Bericht von:** Julia Kleinschmidt, Leibniz-Institut für Europäische Geschichte, Mainz

Digital Humanities sind zurzeit „in“. Dennoch sind die vielfältigen, teils erst noch entstehenden Verknüpfungen von fachlicher Digitalisierung und geisteswissenschaftlichen Projekten nicht immer offensichtlich, werden sogar häufig von der Zunft der Geschichtsschreibern eher skeptisch betrachtet.<sup>1</sup> Mit den Herausforderungen und Möglichkeiten des Digitalen Zeitalters stellen sich neue Probleme, aber auch ungeahnte Lösungsansätze, wie mit den immer größeren Mengen an Informationen über Quellen, Archivierungsformen und Forschungsinfrastrukturen, also mit Metadaten, umgegangen werden kann. Am Leibniz-Institut für Europäische Geschichte (IEG) in Mainz fand am 24. und 25. September 2012 der Methodenworkshop „Metadaten in den historischen Wissenschaften“ statt, der den Spagat zwischen Erklärung der Funktion und Nutzen von Metadaten für (Einzel-)ForscherInnen und für große Institutionen probierte. Das IEG und DARIAH-DE<sup>2</sup> führten an diesen zwei Tagen in laufende Projekte von DARIAH-DE ein und ließen die Teilnehmenden aus allen Bereichen der historischen Wissenschaften diese erproben.

Der Direktor der Universalgeschichtlichen Abteilung am IEG, JOHANNES PAULMANN (Mainz), eröffnete den Workshop mit einleitenden Worten zu der vom IEG angestrebten dualen Publikationsstrategie und den bereits bestehenden digitalen Projekten wie EGO<sup>3</sup> (Europäische Geschichte Online), IEG-Maps<sup>4</sup> und recensio.net<sup>5</sup>. Er betonte die Bedeutung von Metadaten für die Verwirklichung von Forschungsprojekten und forderte innerhalb der Forschungspraxis, sich in Zukunft noch mehr auf Vielfalt zu stützen, statt in Einfalt zu verweilen.

Im Anschluss führten die Veranstalter STEFAN SCHMUNK (Mainz) und CHRISTOF

SCHÖCH (Würzburg) in das Tagungsthema ein. Schöch betonte, dass dies die erste Tagung innerhalb einer geplanten Veranstaltungsreihe sei, während der die verschiedenen Zielgruppen (NachwuchswissenschaftlerInnen, DozentInnen, ExpertInnen) angesprochen werden sollen, um Kompetenzen zu vermitteln. Gleich zu Beginn zeigte sich, dass es sich bei den Digital Humanists um eine ganz eigene Community handelt, die den Übergang von und die Zusammenarbeit mit den „reinen“ GeisteswissenschaftlerInnen erst noch verstärkt suchen muss und darum mit Workshops wie diesen einen begrüßenswerten Schritt an die historische Forschung heran trat.

Zunächst stellte Schöch das europaweite Projekt DARIAH und seine deutsche Abteilung vor. An DARIAH-DE beteiligt seien bisher sieben Universitäten, vier Rechenzentren, drei fachspezifische Einrichtungen, eine Bibliothek, eine Akademie der Wissenschaften und ein kommerzieller Anbieter. DARIAH helfe hier, digitale Sammlungen anzulegen, zu pflegen und zu bewahren. Diese (Nach-)Nutzungsmöglichkeiten eröffneten neue Wege, Daten zu analysieren und zu interpretieren. Das Netzwerk soll dazu beitragen, dass ICT-gestützte Methoden und Werkzeuge auch von Informatik-Laien verstanden und genutzt werden können, auch, um eine Zusammenarbeit über Disziplingrenzen hinaus verstärkt zu fördern. Angesprochen, so Schöch, seien beispielsweise PhilologInnen, HistorikerInnen und MusikwissenschaftlerInnen. Der Fokus von DARIAH-DE, den Bedürfnissen der jeweiligen Fachwissenschaften gerecht zu werden und einer übergeordneten interdisziplinären Ebene der

---

<sup>1</sup> Vgl. hierzu Tagungsbericht HT 2010: Virtuelle Grenzen der Geschichtswissenschaft. Stand und Perspektiven der Digitalen Geschichtsforschung. 28.09.2010-01.01.2010, Berlin, in: H-Soz-u-Kult, 16.06.2011, <<http://hsozkult.geschichte.hu-berlin.de/tagungsberichte/id=3692>> (07.12.2012).

<sup>2</sup> Deutsches Teilprojekt des DARIAH-EU-Konsortiums, d.h. der „Digital Research Infrastructure for the Arts and Humanities“, <<http://de.dariah.eu/>> (07.12.2012).

<sup>3</sup> <[http://www.ieg-ego.eu/?set\\_language=de](http://www.ieg-ego.eu/?set_language=de)> (07.12.2012).

<sup>4</sup> <<http://www.ieg-maps.uni-mainz.de/info.htm>> (07.12.2012).

<sup>5</sup> <<http://www.recensio.net/>> (07.12.2012).

---

institutionalisierten Nachhaltigkeit, Langzeitverfügbarkeit und Interoperabilität von Daten(Sätzen) zu gewährleisten, scheint noch ein großes Experimentierfeld zu sein. Darum betonte Schmunk auch, dass DARIAH als ein „offener Prozess“ zu verstehen sei, der als sich stetig erneuernder „Arbeitsfluss“ (workflow) charakterisiert werden könne, an dem GeisteswissenschaftlicherInnen aller Disziplinen sich beteiligen und die von DARIAH entwickelten Dienste nutzen können. Besonders interessant war der Ausblick, dass alle ForscherInnen auch unabhängig ihrer institutionellen Bindung diesen Verbund für sich in absehbarer Zeit nutzen können.

Der erste Tag war den theoretischen Einführungen zu Metadaten und den damit verbundenen technischen Möglichkeiten zur Informationsspeicherung gewidmet, um zunächst für alle Teilnehmenden einen Zugang zu dem Komplex von Digitalisierung von Quellen und (Sekundär-) Literatur in den historisch arbeitenden Geisteswissenschaften aufzuzeigen. Die Konzentration auf „Metadaten“ erschien sinnvoll, da diese die Grundwerkzeuge bilden, mit denen erst die Informationen von Quellen und Literatur digital beschrieben und erschlossen werden, um sie anschließend anhand von Datenbanken stetig abrufbar machen zu können.

JENS LUDWIG (Göttingen) gab einen ersten Einblick in die methodische Vielfalt, die sich (auch) GeisteswissenschaftlerInnen mit den digitalen Technologien eröffnet. Das Problem, das sich dabei stelle, seien neben der benötigten Organisationsstruktur (Angestellte, Räume, technisches Gerät), dass mit dem Instrument „Metadaten“ nicht überall dasselbe gemeint sei. Metadaten seien aber unumgänglich, sie seien die strukturierten Daten über Daten – d.h. nicht einfache Beschreibungen, sondern Informationen, die einer Systematik unterliegen und selbst Inhalte beschreiben. Der Klassiker von Metadaten ist der Eintrag im Bibliothekskatalog. Weiterhin können sie aber noch vielmehr, wie Ludwig an den inzwischen populär gewordenen Musikverwaltungsprogrammen wie iTunes und NokiaMusik anschaulich machte.

Anschließend sprach GREGOR HORSTKEMPER (München) über „Nutzungsmöglichkeiten bibliographischer Metadaten und

Erschließungssysteme“. Bekannt sind hier unter ForscherInnen die Möglichkeiten, bibliographische Daten aus den Bibliothekskatalogen in die eigenen Literaturdatenbanken wie Citavi, Zotero und Endnote zu kopieren. Weniger bekannt sind allerdings die E-Maildienste, Neuerwerbungsdienste und RSS-Feeds, die nach Ort, Epoche und Raum klassifiziert werden könnten. Über diese erweiterten Dienste können ForscherInnen spezifisch und ständig aktualisiert in den bekannten großen Bibliothekskatalogen suchen bzw. sich automatisch über Neuerscheinungen und die Veränderung ihres Forschungsgebiets informieren lassen.

CHRISTIANE FRITZE (Göttingen) vertiefte in ihrem Beitrag die Frage, woher die einzelnen Metadatenysteme stammen und wie so hier eine Interoperabilität überhaupt gewünscht bzw. notwendig werde. Mit dem Fokus auf die aktuellen Standards (zum Beispiel CEI, TEI, CDWA, COinS) erklärte Fritze, wie Daten zwischen einzelnen Systemen (nicht) austauschbar sind. Während dieses Vortrags wurde deutlich, dass Laien die Denkform bzw. Sprache von Digital Humanists erst lernen müssen, um mit Begriffen wie Vokabular, Modell und Kodierungssyntax (bspw. XML) einen souveränen Umgang entwickeln zu können. Fritze gelang es trotz dieses fachspezifischen Inhalts – das zwar Basiswissen jeglichen Programmierens darstellt, auf GeisteswissenschaftlerInnen jedoch erst abschreckend wirken kann – die Problematiken der historisch gewachsenen Wissensstrukturen herauszustellen. Aufgrund unterschiedlicher „Schulen“, die sich Metadaten widmen, gibt es heute verschiedene Communities und damit auch die etablierten Standards, die erst durch eine gemeinsame Syntax interoperabel gemacht werden können.

TOBIAS GRADL (Bamberg) und MATTEO ROMANELLO (Berlin) widmeten sich im letzten Beitrag des ersten Tages einer, wie Gradl es bezeichnete, „Vision vom Austausch von Wissen“ durch interoperable Sammlungen und Metadaten. Sie legten vor allem die immer stärker gewünschte Interdisziplinarität von Forschungsfeldern, Projekten und Fragestellungen als Ausgangsanforderungen an Projekte wie DARIAH-DE dar. Dabei erschienen zwei Probleme prägnant, die zukünftig

gelöst werden müssten. Zum einen können Sammlungen nur mit Kontextwissen benutzt werden. Datenbanken, die eine Nachnutzung oder auch parallele Anwendung finden wollen, müssen also Wege zur Explizierung des Wissens über diese Datenbanken aufzeigen. Zum anderen wurde hervorgehoben, dass Archive bisher meist autonom organisiert sind und von ihrer Ausrichtung vor der digitalen Revolution, vor allem aber vor der Etablierung des World Wide Web als Kommunikationsraum, nicht unbedingt mit anderen Einrichtungen integrierbar sein wollten/sollten. Der Gedanke an einen umfassenden Open Access scheint noch nicht in allen Bereichen geklärt, was in der Wissenschaft aufgrund des Konzepts vom geistigen Eigentum als immanenten Bestandteil seriöser Forschung noch zu klären gilt. Wichtigste Bausteine der angestrebten Interoperabilität seien, so Gradl und Romanello, neben der Bewahrung eigener origineller Forschungsergebnisse ein Minimieren von Informationsverlust und das Aufstellen von Transformationsregeln oder sogenannten Crosswalks zwischen Archiven in einer zentralen Registry.

Den Abschluss des ersten Tages bildete der öffentliche Abendvortrag von WOLFGANG SCHMALE (Wien) mit dem Titel „Kulturwissenschaften im ‚Digital Turn‘“. Schmale thematisierte in seinem spannenden und zugleich kurzweiligen Vortrag das „Phänomen der Digitalität“ in den Geistes- und Kulturwissenschaften und problematisierte anhand der Aspekte Rationalisierung, Verflüssigung, Entgrenzung, Dekontextualisierung und Personalisierung das Veränderungspotential, die Stärken und die Schwächen von digitalen Methoden und Verfahren für geisteswissenschaftliche Forscher. Gerade die Akzeptanz von digital publizierten Forschungsergebnissen und Qualifikationsarbeiten schätze er zurzeit in den Geistes- und Kulturwissenschaften noch gering ein und attestierte hierbei einen ungeheuren Entwicklungsbedarf. Darüber hinaus vertrat Schmale die These, dass ein zentraler Mehrwert der „Digital Humanities“ bei der Erstellung von Daten- und Quellensammlungen läge. Denn nur durch digitale Verfahren und die Verwendung von Metadaten schemata und Normdatenklassifikationen könnten „Meta-

Quellensammlungen“ erstellt werden, die nicht nur Kontextualisierungen liefern, sondern darüber hinaus auch inhaltliche und semantische Vernetzungen generieren könnten. Gerade aus diesen Gründen müsste verstärkt ein wissenschaftlicher Diskurs um Digitalität und digitale Methoden in den Geistes- und Kulturwissenschaften angestoßen werden, der auch methodische und Theorie reflektierende Aspekte aufgreifen müsste.

Am zweiten Tag wurde es praktisch und Anhand von Beispielen und Übungseinheiten konnte noch einmal verdeutlicht werden, wozu Metadaten in der Forschung benötigt werden und was diese leisten können. Während am ersten Tag durch die theoretischen Überlegungen noch einige Fragezeichen stehen geblieben waren, konnten nun die TeilnehmerInnen nachvollziehen und später auch selbst ausprobieren, wie mit Metadaten Informationen über eine Quelle gesammelt, sortiert und abrufbar geschrieben werden müssen.

DANIEL KURZAWA (Göttingen), THOMAS KOLLATZ (Essen) und Christof Schöch führten in ihrem Panel anhand von historischen Fotos und Briefen vor, was deskriptive, administrative und strukturelle Metadaten voneinander unterscheidet und welche graduellen Informationsschritte je nach Metadatenstandart für die jeweilige Datenbank von Belang sind. Besonders die gut gegliederte Übung, in der ein Brief von Ludwig van Beethoven in eine Datenbank eingespeist werden sollte, zeigte die Unterschiede zwischen den Möglichkeiten der Metadaten-Standards Dublin Core (DC) und TEI. Hierbei erleichterte es den Teilnehmenden, wenn sie über Vorwissen über die syntaktischen Eigenheiten von HTML oder XML verfügten, um die Beispiele sofort anwenden zu können. Anhand der Übung war es nicht nur möglich, das unterschiedliche Potential der Standards zu erkennen, sondern auch sich darüber bewusst zu werden, welche Informationen über eine Quelle in der (digitalen) Registrierung bzw. Sammlung von Bedeutung sind.

Tobias Gradl und Matteo Romanello führten anschließend erste für Metadaten relevante Testversionen von Komponenten der DARIAH-DE-Infrastruktur vor. Gradl versuchte anhand einer Klicker-Umfrage mit den TeilnehmerInnen zu klären, was Suchmaschi-

---

nen leisten müssen. Da die Internet-oder Datenbanksuche eine tägliche Aktivität sei und DARIAH-DE Anfragen nicht nur an Text, sondern auch an Inhalte anderer Formate wie MP3 stellen können soll, muss die Suchmaschine filtern und klassifizieren können. Ziel sei also die generische Suche innerhalb der Forschungsdaten. In einer Demoversion zeigte Romanello, wie die Suchoberfläche von DARIAH-DE schließlich in etwa aussehen werde und welche Suchoptionen die BenutzerInnen auswählen können.

Der Workshop schloss inhaltlich mit dem Panel von PHILIPP VANSCHIEDT (Trier) und THOMAS KOLLATZ (Essen). Kollatz führte in virtuellen Karten das Anzeigen historischer Orte vor, die unabhängig der historischen Bezeichnung erkannt werden. Mit der Geolocation bekommen dann historische Ortsnamen wie Mainz oder Magenta die identischen Koordinaten und eine feste ID zugewiesen. Vanscheidt präsentierte ein Projekt, in dem bis 2013 der mittelalterliche Bestand der Benediktinerabtei St. Matthias in Trier digitalisiert werden soll. Etwa 500 Codices aus der Zeit vom 8. bis 16. Jahrhundert, die auf 25 Standorte weltweit zerstreut sind – liturgische, theologische und philosophische Schriften – sollen hier findbar werden. Probleme, wie die unterschiedliche Schreibweise des Namens des Verfassers und bisherige nicht vereinheitlichte Autoreninformationen sind die Herausforderungen, mit denen sich die Umsetzung zurzeit konfrontiert sieht und die für den Start im Jahr 2013 entscheidend sind.

In der Abschlussdiskussion riefen Schmunk und Schöch noch einmal die Fragen nach den Anforderungen von historischer Forschung an Metadaten(standards) vom Beginn des Workshops in Erinnerung. Unter den Teilnehmenden kamen zudem Fragen auf, wie man mit Nicht-Texten, zum Beispiel Objekten, umgehen könne und welches Funktionspotential sich mit großen visuellen Formaten wie den digitalen Karten für Forschungsfragen ergebe. Auch wurde die Überlegung diskutiert, wie realistisch die Umsetzung einer neuen umfassenden Datenbank sei, die sich neben anderen Projekten wie Europeana<sup>6</sup> und der Deutschen Digitalen Bibliothek<sup>7</sup> behaupten muss. Für

große Institutionen etwa müsste der Nutzen von DARIAH-DE als Angebotsplattform der eigenen Sammlungen und Forschungsdaten als attraktiver Standard in der Digitalisierungsbranche erkannt werden.

Die im Workshop aufgezeigten beeindruckenden Möglichkeiten, die sich innerhalb der (historischen) Geisteswissenschaften durch eine breite digitale Infrastruktur an Forschungsdaten ergeben und interessante (Forschungs-)Fragen aufwerfen, müssen in Zukunft trotz der gelungenen Bemühungen noch besser auf die einzelnen Zielgruppen eingehen. Dies betrifft sowohl die technischen Angebote als auch die Vermittlung an ForscherInnen. Nur so kann das Feld der Digital Humanities ein interdisziplinäres sein, das Projekte flächendeckend verknüpft und den Austausch einzelner Forschungsvorhaben fördert. Daher ist es wünschenswert, dass dieser Auftaktveranstaltung noch viele weitere folgen werden, die GeisteswissenschaftlerInnen helfen können, im Digitalen Zeitalter Fuß zu fassen.

#### **Workshopübersicht:**

##### *Begrüßung*

Johannes Paulmann (Mainz)

##### *Thematische Einführung*

Christof Schöch (Würzburg) / Stefan Schmunk (Mainz)

Jens Ludwig (Göttingen): Metadaten für Forschungsdatenmanagement und Langzeitarchivierung

Gregor Horstkemper (München): Nutzungsmöglichkeiten bibliographischer Metadaten und Erschließungsmöglichkeiten

Christiane Fritze (Göttingen): Metadaten als Vorbedingung einer interoperablen Nachnutzung

Tobias Gradl (Bamberg) / Matteo Romanello (Berlin): Interoperabilität aus technischer Perspektive

##### *Öffentlicher Abendvortrag*

---

<sup>6</sup> <<http://www.europeana.eu/portal/>> (07.12.2012).

<sup>7</sup> <<http://www.deutsche-digitale-bibliothek.de/>> (07.12.2012).

Wolfgang Schmale (Wien): Kulturwissenschaften im „Digital Turn“

Daniel Kurzawe (Göttingen) / Thomas Kollatz (Essen) / Christof Schöch (Würzburg): Grundlagen von Metadaten und Datenstandards – Erfassungs- und Vergabeprobleme von Metadaten

Tobias Gradl (Bamberg) / Matteo Romanello (Berlin): Using the interoperability components of DARIAH-DE

Philipp Vanscheidt (Trier) / Thomas Kollatz (Essen): Metadaten in Anwendungsbeispielen

*Abschlussdiskussion*

Christof Schöch (Würzburg) / Stefan Schmunk (Mainz)

Tagungsbericht *Methodenworkshop „Metadaten in den historischen Wissenschaften“*. 24.09.2012-25.09.2012, , in: H-Soz-u-Kult 19.12.2012.