Digitale Diplomatik - Die historische Arbeit mit Urkunden in der digitalen Welt

Veranstalter: Ludwig-Maximilians-Universtität Historisches Seminar Abt. Geschichtliche Hilfswissenschaften

mit Unterstützung der DFG, der Mission Historique Francaise en Allemagne und von Monasterium.net

unter der Schirmherrschaft der Commission Internationale de Diplomatique

Datum, Ort: 28.02.2007-02.03.2007, München

Bericht von: Georg Vogeler, Historisches Seminar, Abt. Geschichtliche Hilfswissenschaften, Ludwig-Maximilians-Universität München

In verschiedenen Publikationen und Tagungen der vergangenen acht Jahre ist deutlich geworden, dass die mittelalterlichen und frühneuzeitlichen Urkunden ein Paradebeispiel für die Digitalisierung von Quellenmaterial sein können.[1] Dass diese Entwicklung durchaus auch dem Interesse der Urkundenforscher entspricht, betonte Theo Kölzer in seinem Einführungsreferat unter dem Titel „Diplomatik, Edition und Computer“. Er stellte die Möglichkeiten der EDV in den Vordergrund, mit großen Mengen Text umzugehen und somit das Problem der Masse spätmittelalterlicher Überlieferung zu lösen. Er ordnete die Möglichkeiten textlinguistischer Analyse von großen Textcorpora in die Tradition des Diktatvergleichs ein und hob hervor, dass damit aus seiner Sicht keine neue Form der Diplomatik entstünde, sondern eine evolutionäre Weiterentwicklung der Traditionen der Urkundenforschung – einer Wissenschaft, die durch die neuen Techniken nicht abgelöst werde. Die digitale Diplomatik biete im Gegenteil eine weitere Chance, die institutionelle Krise der Diplomatik in Deutschland zu überwinden.

Auch wenn das Verhältnis zwischen gedruckten und digitalen Publikationen nicht explizit Thema der Sitzung „Edition“ war, kam es aber doch in mehreren Beiträgen zur Sprache: Die von Clemens Radl vorgestellten Urkundeneditionen innerhalb der digitalen Monumenta Germaniae Historica (dMGH) waren als genaues, zitierfähiges Abbild der gedruckten Edition konzipiert. Inzwischen liegen die Urkunden aber auch in einer beta-Version schon als Volltext vor. Radl zeigte Nutzungsperspektiven der Volltexte durch weitere Erschließung auf, wie z.B. durch kartographische Darstellung von Informationen. Was bei den dMGH mit der institutionellen Unterstützung eines DFG-Projektes möglich war, kann aber teilweise auch in Eigeninitiative verwirklicht werden: Tom Graber und André Thieme vom Institut für Sächsische Geschichte und Volkskunde haben die 25 Bände der seit dem späten 19. Jahrhundert erscheinenden „Codex diplomaticus Saxoniae Regiae“ ins Netz gestellt und damit auch nichtfachhistorische Nutzerkreise erreichen. Jürgen Sarnowsky analysierte die Erfahrungen mit den drei Hamburger Projekten (Hamburgisches Urkundenbuch, Preußisches Urkundenbuch und Hospitaller Sources) und zog ein eher ernüchtertes Fazit: Der gegenwärtige Stand der Dinge macht die gedruckter Edition als wichtigste Publikationsform noch nicht überflüssig. Um die digitalen Editionsformen nicht nur als Experimentierfeld zu verwenden, äußerte er Bedürfnisse nach einer Vernetzung der Projekte, nach einem Editionsstandard für digitale Urkundenbücher und nach „Editorentools“ als Entlastung für konventionelle Editoren. Auch für die Regesta Imperii sind die digitalen Formen Vorstufen zu den

[1] Vgl. die Publikationen von Ansani, Michele, Diplomatica (e diplomatisti) nell'arena digitale, http://dobc.unipv.it/scrineum/ansani.htm, in: Scrineum 1 (1999), S. 1-11 = Ansani, Michele: Diplomatica (e diplomatisti) nell'arena digitale, in: Archivio storico italiano 158 (2000), S. 349-379); Uhde, Karsten, Urkunden im Internet - Neue Präsentationsformen alter Archivalien, in: AfD 45 (1999), S. 441-464; Poupeau, Gautier, Comprendre les enjeux techniques de l'édition électronique, in: AfD 52 (2006), S. 467-478; Sahle, Patrick; Schaßan, Thorsten, Das Hansische Urkundenbuch in der digitalen Welt. Vom Druckwerk zum offenen Quellenrepertorium, in: Hansische Geschichtsblätter 118 (2000), S. 133-155; Sahle, Patrick; Vogeler, Georg, Urkundenforschung und Urkundenedition im digitalen Zeitalter, in: Burkhardt, Daniel; Hohls, Rüdiger; Ziegeldorf, Vera (Hrsg.), Geschichte und Neue Medien in Forschung, Archiven, Bibliotheken und Museen. Tagungsband .hist2003, Berlin 2005 (Historisches Forum. Schriftenreihe von Clio-online 7,1), S. 333-382; Urkunden-Editionen im Internet. Einführung und Überblick, in: AfD 52 (2006), S. 429-448; Vogeler, Georg u.a., Towards a standard of encoding medieval charters with XML, in: Literary and Linguistic Computing 20 (2005), S. 269-280.; Vom Nutz und Frommen digitaler Urkundeneditionen, in: AfD 52 (2006), S. 443-466.). Siehe auch die Tagungen bzw. Tagungssektionen München 2004: Vogeler, Georg, Ein Standard für die Digitalisierung mittelalterlicher Urkunden. Bericht zum Workshop (München 5./6. April 2004), auch in Historicum.net: http://www.forschung.historicum.net/tagungsberichte/muenchen-200404.html, in: AHF-Informationen 2004 Nr. 023 (#http://www.ahf-muenchen.de/Tagungsberichte/Berichte/pdf/2004/023-04.pdf#, 2004), Vogeler, Georg, Ein Standard für die Digitalisierung mittelalterlicher Urkunden mit XML. Bericht von einem internationalen Workshop in München 5./6. April 2004, in: AfD 50 (2004), S. 23-34), Wien/Göttweig 2005; Aigner, Thomas; Winter, Karin (Hrsg.), Alte Archive – Neue Technologien / Old Archives - New Technologies, St. Pölten 2006.

© H-Net, Clio-online, and the author, all rights reserved.

gedruckten. Doch Andreas Kuczera stellte zusammen mit den in einfachem TEI ausgezeichneten „Regesta Imperii Online“ auch gleich Ansätze für die von Sarnowsky eingeforderten Tools vor (FuD-System Trier), die insbesondere die Registererstellung erleichtern sollen.

Damit war der Beitrag von Kuczera auch Teil des zweiten Schwerpunkts der Beiträge der Sektion, der sich um die „Erschließung“ drehte. Der Eröffnungsbeitrag der Sektion von J.W.J. Burgers und F.G. Hoekstra stellte das Projekt einer digitalen Edition der Register der Grafen von Holland und Seeland aus der Zeit von 1316-1345 vor, das Bilddigitalisierungen mit XML-Daten verknüpft. Sie beschrieben die Wege, die das Projekt eingeschlagen hat, um Formalbeschreibung und Texte für ein Register zu erschließen. Eine wichtige Rolle spielte dabei der einfache Editor (ING Forms), der auf Grund der flachen Struktur der Auszeichnung eine datenbankartige Maske für die Erfassung zur Verfügung stellt. Dass die Erschließung einen sehr hohen Komplexitätsgrad erreichen kann, zeigte besonders der Beitrag von Arianna Ciula, die zwei Projekte am Centre for Computing in the Humanities des Kings College London präsentierte: Die Digitalisierung Fine Rolls (einer Art Supplikenregister) Heinrichs III. („Henry III Fine Rolls project“, http://www.frh3.org.uk) und das Pilotprojekt der „Anglo-Saxon charters“. Letzteres beruht auf den Vorarbeiten, die insbesondere in Cambridge in den letzten Jahren erstellt worden sind. Diese werden nun mit einer um diplomatische Eigenschaften erweiterte TEI-Syntax ausgezeichnet. Insbesondere versucht das Zentrum, die sachlichen Informationen der Urkunden, allen voran Personennamen, mit Hilfe von Ontologien abzubilden und so historischen Fragestellungen besser zugänglich zu machen. Valeria Leoni vertrat eines der Flagschiffe digitaler Urkundenpräsentation, den „Codice diplomatico della Lombardia Medievale“, der systematisch die älteren Urkundenbestände der Lombardei online in einem für das Projekt entwickelten XML-Format ediert und mit einer Suchmaschine erschlossen hat, um neben Volltexten auch Personen und Orte aber auch das diplomatische Formular gezielt durchsuchen zu können.

Wie schon in der Editionssektion immer wieder speziell für die Anwendungszwecke entwickelte Software vorgestellt worden war, präsentierte auch die Sektion „Techniken“ solche Instrumente. Das von Ben Burkard vorgestellte Editionstool „EditMOM“ ist vielleicht die am weitesten entwickelte Antwort auf Jürgen Sarnowskys Forderung nach „Editorentools“. Es ist eine Arbeitsumgebung zur Regestierung, Transkription und Kommentierung der im Monasterium-Projekt enthaltenen Urkundenbilder, die technische Kenntnisse über das dabei erzeugte XML überflüssig macht, ohne die Mächtigkeit der Technik allzu sehr einzuschränken. Das Tool ist zusätzlich kollaborativ konzipiert und war damit ein Anwendungsfall für Manfred Thallers theoretische Analyse der Konzeption eines kollaborativen Editionstools. In Thallers Vorstellung soll das Redaktionssystem die Dokumente in seiner visuellen, zeichenhaften und interpretativen Ebene darstellen und die Beiträge der Bearbeiten auf den jeweiligen Ebenen sichtbar machen. Mit einer sauberen Trennung der interpretativen Ebene vom Objekt würde die Forschungsdiskussion auch in der virtuellen Welt den Ort bekommen, den sie von der Geschichtswissenschaft zugewiesen hat: Die Diskussion über die Echtheit einer Urkunde wären so direkt an die Edition angefügt und direkt am Urkundenbild überprüfbar.

Welche technischen Änderungen nötig waren, um ein zunächst als Arbeitshilfsmittel konzipiertes System zur Dokumentation von Urkunden in ein aktuellen Ansprüchen genügendes Onlinesystem umzuwandeln, konnte man am Vortrag von Christian Domenig studieren, der die Datenbank zum Editionsprojekt „Die Urkunden und Briefe der Grafen von Cilli (1341-1456)“ vorstellte.

Der Beitrag von Michael Gervers und Michael Margolin stellte die für das DEEDS-Projekt entwickelte noch komplexere Lösung vor, die einerseits wie schon der Beitrag von Arianna Ciula aufzeigte, mit modernen Ontologie-Technologien (RDF, OWL) arbeitet und andererseits den Text und seine Metadaten als mehrere Ebenen konzipiert, das heißt zu ein und demselben Text verschiedene Arten der Auszeichnung parallel verwaltet.

Bernhard Assmann brachte geläufige Nutzungskonstruktionen in die elektronische Präsentation ein: Er forderte, „Systeme kanonischer Referenz“ wie z.B. die eingeführten Zitierweisen der MGH, auch in den elektronischen Systemen abzubilden, um den Zugriff aus der Sache heraus dauerhaft sicherzustellen. Seinen Vorschlag kann man auch als Gegenargument gegen die für Buch und digitales Faksimile immer wieder ins Feld geführte Zitierfähigkeit werten, denn eine sachliche Referenz auf z.B. eine Urkundennummer besitzt von der jeweiligen Präsentationsform unabhängige Dauerhaftigkeit.

© H-Net, Clio-online, and the author, all rights reserved.

Linguistische Probleme im Umgang mit den Quellen sprach Markus Heller an, der approximative Verfahren zur Suche in frühneuhochdeutschen Texten vorstellte und dabei Hoffnung machte, dass die geläufigen graphischen Variationen für Suchanfragen zusammengefasst werden können.
Paolo Buonora analysierte zutreffend, wie ein auf einer Matrix aufbauendes Kompressions- und Speicherverfahren wie jpg2000 die Langlebigkeit von digitalen Bildern erhöht, in dem Fehler auf einzelne, zu isolierende Teile der Bilddatei beschränkt werden können.
In die Praxis des digitalen Edierens führte Gautier Poupeaus Beitrag zurück, der die Möglichkeiten ausforschte, einen weit verbreiteten Standard zur Auszeichnung von geisteswissenschaftlichen Texten (TEI P5) an individuelle und lokale Gegebenheiten anzubinden und die Veränderungen zu dokumentieren. Als Ergebnis der Sektion zur Technik kann also festgehalten werden, dass die Forderung nach im Editionsalltag nutzbaren Tools gehört wird und in verschiedenen Feldern an Lösungen gearbeitet wird, die den Diplomatiker und Editor in den Vordergrund stellen und ihm angemessene Bearbeitungs- und Präsentationsformen zu entwickeln versuchen.

Dass der Historikerarchivar, der sich auch an diplomatischer Arbeit beteiligt noch lange nicht von den Managerarchivaren verdängt ist, zeigte die Sektion „Archive“ am Freitag. Der Beitrag von Claes Gejrot verwies darauf, dass die Archive auch aktiv an digitalen Editionsprojekten beteiligt sind: Die Digitalisierung des „Diplomatarium Suecanum“ wird als Erweiterung der Urkundenkartei des Stockholmer Riksarikiv verstanden und aktiv vom Archiv betrieben. Auch das monasterium-Projekt, dessen organisatorische Erfahrungen Karl Heinz vorstellte, verbindet Archivmaterial mit wissenschaftlichen Editionen. Heinz stellte insbesondere die Wege vor, mit denen die Kooperation einer Vielzahl von Institutionen überhaupt erst möglich gemacht werden kann. Auch das Landesarchiv Württemberg beschränkt sich im Projekt des Württembergische Urkundenbuch Online, das Maria Magdalena Rückert und Nicola Wurthmann vorstellten, nicht auf die Wiedergabe der vorhandenen Archivinformationen oder die Digitalisierung eines vorhandenen Druckes, sondern verbindet beide Arbeiten und beteiligt sich an der weiteren editorischen Verarbeitung in der online Präsentation. Die Grenze zur Edition zog dagegen deutlich der Beitrag von Raffaele Santoro, der die Digitalisierung des Urkundencorpus des Archivio di Stato di Venezia in die archivische Arbeit einordnete und das Projekt in die Tradition der Codice diplomatici seit dem Mittelalter stellte.

Der Überblick über das Verhältnis der Archive in Tschechien, Slowakei und Ungarn zur Präsentation von Archivmaterial im Netz von Juraj Sedivý zeigte die ganze Bandbreite von reinen „Visitenkarten“ über die Recherche in den Findmitteln bis zur umfassenden Regestenerschließung und Angebote von Bildern der Archivalien auf. Die von ihm vorgestellte Archivregion war für ihn auch ein Beispiel für die Notwendigkeit von mehrsprachigen Konzepten bei der Präsentation von Archivalien im Netz. Eher an der Archivpraxis orientiert waren auch die Beiträge von Giovanni Tartaglione und Joachim Kemper: Giovanni Tartaglione stellte sein Modell der zentralen baumartigen Erfassung der Daten bei der Erschließung und Digitalisierung von Archivmaterial am Beispiel des Staatsarchivs Lucca vor und forderte, dass die Archivare sich intensiv mit den auf XML basierenden Technologien auseinanderzusetzen müssen, die der Struktur der Archivmaterialien und archivischer Arbeitsweise am besten angepasst sind. Joachim Kemper listete die Positionen des Bayerischen Hauptstaatsarchivs zur Digitalisierung von Urkunden auf, das sich der aktuellen Entwicklung öffnet und so das exceptionelle Material des Archivs in überregionale Projekte einbringt.

Die Beiträge der Sektion zeigten einmal mehr, dass für die Archive die Übertragung ihres Materials in eine digitale Repräsentation nicht als Bedeutungsverlust zu werten ist, sondern ihre Stellung im Forschungsprozess und in der Öffentlichkeit stärken. Nicht nur schon die Erstbenutzung am Bildschirm die Originale, sondern auch die überregionale Vernetzung und die größere Öffentlichkeit des im Archiv bewahrten Kulturerbes sind Argumente für die groß angelegte Digitalisierung von Archivgut, deren Durchführbarkeit immer wieder deutlich wurde.

Unter dem Stichwort „Forschung“ versammelten sich Beiträge, die die Nutzung der Digitalisierungsprojekte aus verschiedenen Perspektiven beleuchteten: Das von Andreas Gniffke vorgestellte Corpus der altdeutschen Originalurkunden bis zum Jahr 1300 im Internet ist entstanden als Instrument germanistischer Forschung, bietet aber durch die Digitalisierung Forschungsmaterial für viele weitere Wissenschaften. Die Tiefenerschließung – insbesondere Orte- und Personen, aber auch der

© H-Net, Clio-online, and the author, all rights reserved.

Umgang mit zahlreichen Sonderzeichen – ist aber noch eine größere Herausforderung.
Die Aktivitäten der Bayerischen Staatsbibliothek auf dem Bereich der Massendigitalisierung, stellten Alessandra Sorbello Staub und Gregor Horstkemper vor. Die Bayerische Staatsbibliothek sieht sich dabei besonders in der systematischen Digitalisierung von bibliographisch nachgewiesenen Beständen (VD 16) und Handschriften engagiert. Sorbello Staub betonte, dass die vorgestellten Strategien, diese digitalen Faksimiles auch tiefer zu erschließen, die Kooperation mit der Fachwissenschaft brauchen. Horstkemper konzentrierte sich auf die von der Bayerischen Staatsbibliothek eingesetzten Tools zur Langzeitarchivierung. Die Bibliothek stellte sich so als Dienstleister für die Wissenschaft dar, massenhaft Digitalisate zu erstellen und dauerhaft verfügbar zu machen.

Christian-Emil Ore präsentierte Überlegungen, wie die Vielzahl der von den einzelnen Urkundendigitalisierungsprojekten erstellten Ergebnisse zusammengeführt werden können. Er analysierte vorhandene Modelle und hob besonders das für Museen und andere Institutionen des Kulturerbes entwickelte CIDOC CRM (Conceptual Reference model) hervor. Das Modell schlägt ein abstraktes Schema zur Organisation von Beschreibungsmodellen vor, das einen Austausch zwischen den lokalen und objektspezifischen Beschreibungen ermöglichen soll. Er stellte das CIDOC CRM neben die Weiterentwicklung des bibliothekarischen FRBR-Modells (Functional requirements for bibliographical records), das als FRBRoo (object oriented) auch auf „Werke“ übertragbar ist, die keine gedruckten Bücher sind. In Anwendung des CIDOC-CRM auf die Urkundenerschließung bekam für Ore das „Ereignis“ besonderes Gewicht: Die Urkundenausstellung ist für ihn der logische Ort, an dem Personen, Orte, Rechtssachverhalte usw. zusammenlaufen.

Den Umgang mit der Sprache der Urkunden stellten die drei folgenden Beiträge in den Vordergrund. Herbert Stoyan skizzierte, wie man die Konzepte des „Textmining“ auf Bestände digitalisierter Urkunden übertragen und dabei auf Regel- und Regelmäßigkeiten stoßen kann, die von historischem Interesse sind. Benoît Tock untersuchte an Beispielen aus ARTEM und dem Chartae-Galliae-Projekt die Möglichkeiten, Bezüge zwischen Urkunden an der Verwendung auffälliger Wörter und Formeln herzustellen. Margarete Springeth stellte die Mittelhochdeutsche Begriffsdatenbank vor, die seit neuestem auch Urkundentexte (Urkunden des Klosters Herzogenburg) enthält und damit die tiefe semantische und formale Durchdringung des Datenbankprojektes auch auf Urkundensprache anwendet. An den Referaten von Stoyan, Heller und Springeth wird deutlich, dass auch einige Unschärfen der Sprache in Regelsysteme übertragen werden können, mit deren Hilfe der Computer historisches Sprachmaterial jenseits des reinen Zeichenkettenvergleichs auswerten kann. Bedauerlich war nur, dass keine vergleichbaren Initiativen aus mittellateinischer Sicht auf der Tagung präsentiert wurden.

Dass auch ein einfaches Nachweisinstrument, als das das Repertorium der Chartulare Frankreichs ursprünglich geplant war, durch die Digitalisierung zu einem veritablen Forschungsinstrument werden kann, bewies der Beitrag von Paul Bertrand, der aufzeigte, dass die komplexe Überlieferungssituation der mittelalterlichen Urkundenhandschriften in der digitalen Form besser dargestellt, ja eventuell überhaupt erst publizierbar gemacht werden kann.

Als eine Summe der Beiträge konnte man Patrick Sahles abschließende „Vorüberlegungen zur Portalbildung in der Urkundenforschung“ verstehen, die einerseits eine Community von interessierten Forschern einforderten und andererseits einen Vorschlag für die Datenstruktur eines Portals lieferten, in dem die Urkunden der anwachsenden Vielfalt von Urkundenpräsentationen im Netz nachgewiesen werden können. Wie schon im einleitenden Vortrag Theo Kölzers angeklungen war, ist dieses Portal eine wichtige Voraussetzung dafür, dass die digitalen Urkundenpräsentationen für Geschichtswissenschaft, Philologie und Diplomatik fruchtbar gemacht werden können, denn ein zentraler Vorzug der Online-Präsentation des europäischen Urkundenerbes ist die Möglichkeit, nach neuen Kriterien Corpora zu bilden und damit traditionelle Grenzen der Forschung zu überschreiten.

Dass die internationale wissenschaftliche Gemeinschaft, die zu einem Urkundenportal beitragen und mit seiner Hilfe unterstützt werden kann, schon Realität ist, zeigt die hohe Teilnehmerzahl der Tagung: Über 140 Teilnehmer aus 15 Ländern nutzten die Gelegenheit, sich über den Stand der Forschung zu informieren und Kontakte zu den Kollegen zu knüpfen.
Am Samstag 3. März versammelten sich die Forscher, die aktiv an der Gestaltung der von Patrick Sahle eingeforderten Community mitwirken wol-

© H-Net, Clio-online, and the author, all rights reserved.

len, bei der Sitzung der Charters Encoding Initiative. Das Treffen bestärkte noch einmal die gemeinsame Absicht, eine solche Community aufzubauen. Das Meeting arbeitete weiter an einem Vorschlag für XML-Elemente, die bei der Digitalisierung von Urkunden notwendig sind. Die Arbeit wird über die Mailing-Liste CEI-L fortgeführt. Wissenschaftler, die an dieser Arbeit interessiert sind oder ihre eigenen Erfahrungen bei der Digitalisierung von Urkunden einbringen wollen, werden aufgefordert, sich unter http://www.cei.lmu.de /cei-l.html in die Liste einzuschreiben. Die Ergebnisse der Arbeit der CEI werden auf der Seite der Gruppe http://www.cei.lmu.de kontinuierlich veröffentlicht und stehen der wissenschaftlichen Gemeinschaft zur Nutzung bei der Weiterentwicklung existierender Projekte und bei der Neuentwicklung von Projekten zur Verfügung.

Die Tagungsakten sollen als Beiheft des „Archiv für Diplomatik“ erscheinen. Dabei wird nicht nur die Vielfalt der Projekte und Lösungen sichtbar, mit denen die Urkunden von Archiven, Bibliotheken und Diplomatikern online präsentiert werden, sondern insbesondere der Trend, mit den technischen Hilfsmittel das Material in einer größeren Tiefe zugänglich zu machen als in der traditionellen Edition. Wie die so in vielen Einzelprojekten codierten Ergebnisse, die von prosopographischen Informationen über linguistische Instrumente wie der mittelhochdeutschen Begriffsdatenbank bis zur diplomatischen Tiefenerschließung reichen, jedoch in anderen Projekten weitergenutzt werden kann, ist eine Aufgabe, an der die Community noch intensiv arbeiten muss.

Tagungsbericht *Digitale Diplomatik - Die historische Arbeit mit Urkunden in der digitalen Welt*. 28.02.2007-02.03.2007, München. In: H-Soz-u-Kult 20.05.2007.

© H-Net, Clio-online, and the author, all rights reserved.