

E-Editionen: Neue Publikationsmaßstäbe für historische Quellenprojekte

Veranstalter: Institut für Ost- und Südosteuropaforschung, Regensburg

Datum, Ort: 12.04.2013, Regensburg

Bericht von: Hans Christian Bauer, Institut für Ost- und Südosteuropaforschung, Regensburg

Dass die veränderte Publikationssituation im Digitalen und die Prognose eines *digital turn* in den Geisteswissenschaften zu Verunsicherung führen, verdeutlichen heftige, öffentlich geführte Kontroversen. Insbesondere die Konsequenzen für Autorschaft und Integrität eines Werkes, die Problematik von Renommee und Qualität sowie die vordergründige Dominanz technischer Komponenten beim elektronischen Publizieren provozieren bisweilen emotionale Kritik¹ oder zumindest, wie etwa zuletzt zum Thema Review-Verfahren in der Mailingliste InetBib², hitzige Diskussionen. So widersprüchlich der Einsatz computergestützter Verfahren in den Geisteswissenschaften bewertet werden mag, hat sich dieser in der Editorik doch soweit etabliert, dass sogar ursprünglich als Druckwerk geplante Editionsprojekte zusätzlich für das Internet aufbereitet werden.³ Daher erwies sich der Ansatz des von der Universitätsstiftung Hans Vielberth geförderten Workshops ‚E-Editionen‘, vorrangig die editorische Praxis zu fokussieren und sich an konkreten Projekten zu orientieren, als erfrischend sachlich. Mit der Suche nach gemeinsamen Editionskriterien und nach Lösungsvorschlägen für bisher bestehende Rezeptionshindernisse intendierte die Veranstaltung, greifbare Ziele zur Anwendung erarbeiten zu wollen, um sowohl Lernprozesse hinsichtlich technischer Kompetenzen anzuregen als auch den Blick des kritischen Autors zu schulen.

KONRAD CLEWING (Regensburg) und PETER MARIO KREUTER (Regensburg) wiesen unter dem einleitenden Diktum, das Medium habe noch nicht zu seiner Form gefunden⁴, auf die spezifische Problematik von Editionen in ihrer elektronischen Realisierung hin. Zwar seien Editionsprojekte allein aufgrund des auf diesem Wege erreichbaren Publikationsvolumens und der damit verbun-

denen, auch finanziellen Aufwendungen für eine online-Wiedergabe in Open Access prädestiniert, allerdings herrsche anders als in der klassischen Printausgabe eine „skandalöse“ Heterogenität von Standards und Regeln, die sich schließlich qualitätsmindernd niederschläge und die Nutzerfreundlichkeit beeinträchtigen. Die diversen Strukturen von online-Ausgaben historischer Dokumente forderten dem Nutzer immer neue Recherchestrategien ab, deren Ergebnisse oftmals nicht zufriedenstellend seien – angefangen beim Bibliographieren, bei dem ein klassischer OPAC verlässliche Angaben lieferte, während ein Überblick über die Landschaft elektronischer Quellen nur schwer zu gewinnen sei. Weitreichende Defizite machte Clewing bezüglich der Autorenfreundlichkeit aus. Bereits die komplexe Erstellung der Metadatenätze als Grundlage für die Langzeitarchivierung verlangte von Autorinnen und Autoren eine intensive Auseinandersetzung mit den formalen Voraussetzungen, für die ein einheitlicher Standard zudem fehle. Besonders ins Gewicht fielen Zweifel an der Sichtbarkeit der Autorschaft bei der Präsentation von E-Editionen, deren namentliche Zuordnung aber als Voraussetzung für die Evaluation der individuellen Forschungsleistung gelte. So bietet beispielsweise das Verzeichnis digitaler Editionen des Instituts für Dokumentologie und Editorik keine Suchfunktion über die Editio-

¹ Thomas Thiel mahnte in der F.A.Z. das Risiko der „Austreibung des Geistes aus den Geisteswissenschaften“ an. Thomas Thiel: Digital Humanities. Eine empirische Wende für die Geisteswissenschaften?, in: Frankfurter Allgemeine Zeitung, 24.07.2012, <<http://www.faz.net/aktuell/feuilleton/forschung-und-lehre/digital-humanities-eine-empirische-wende-fuer-die-geisteswissenschaften-11830514.html>> (02.05.2013).

² Diskussion unter dem thread, erster Listeneintrag: Walther Umstätter: Alternde Türsteher der Wissenschaft, in: InetBib, <<http://www.ub.uni-dortmund.de/listen/inetbib/msg50106.html>> (02.05.2013).

³ Siehe: Thomas Stäcker: Wie schreibt man Digital Humanities richtig? - Überlegungen zum wissenschaftlichen Publizieren im digitalen Zeitalter, in: Bibliotheksdienst 47 (2013) 1, S. 24-50, hier S. 25.

⁴ zitiert nach Patrick Sahle: Vom editorischen Fachwissen zur digitalen Edition: Der Editionsprozess zwischen Quellenbeschreibung und Benutzeroberfläche, in: Fundus – Forum für Geschichte und ihre Quellen, Heft 2, S. 75-102, hier S. 75, <<http://webdoc.sub.gwdg.de/edoc/p/fundus/2/sahle.pdf>> (02.05.2013).

ren bzw. Herausgeber.⁵ Dies wiederum wirke sich auf die Rezensionsexemplarpraxis aus, welche elektronische Ressourcen derzeit nur am Rande zur Kenntnis nähme, was natürlich auch dem Umstand geschuldet sei, dass Rezensionsexemplare hier nicht mit einem Rezensionsexemplar rechnen könnten. Alternative Modelle hingegen, wie etwa das Angebot der Präsentation der Publikation durch den Autor selbst⁶, gerieten leicht in den Ruch der Eigenwerbung und widersprächen damit guter wissenschaftlicher Praxis. Mit eben geschilderten Strukturproblemen hätten aktuelle Editionsprojekte am IOS umzugehen. Peter M. Kreuter plant die Erschließung des Zensus der Kleinen Walachei, der einzigen systematisch durchgeführten Erhebung zu Bevölkerung und Infrastruktur im Donaufürstentum Walachei vor 1800. Bisher nur unzureichend ausgewertet, bestehe der Informationswert des Zensus nicht zuletzt in seinen noch weitgehend unbearbeiteten qualitativen Angaben etwa zum Wegenetz in der Region (hier könnten mit Visualisierung, Georeferenzierung, Pop-ups die Mehrwertfunktionen digitaler Editionen gut zu Geltung kommen). Konrad Clewing bearbeitet zum einen den an Kaiser Franz I. von Österreich gerichteten Bericht des Gouverneurs Lilienberg von seiner Bereisung Dalmatiens, 1832-1834 – ein Bericht, in dem nicht nur umfassende Angaben zu Gesellschaft und Administration beinhaltet sind, sondern in denen Lilienberg auch gegen das zentralistische habsburgische Gefüge opponiere. Zum anderen beleuchtet ein Forschungsprojekt zu den „Kaiserreisen“ von Joseph II. und wiederum Franz I. (II.) die Herrschaftspraxis im Habsburgerreich. Die äußerst umfangreichen Dokumente zu den Reisen sollen in Auswahl hybrid in Buchform auf Kroatisch sowie digital auf Deutsch ediert werden.

THOMAS STÄCKER (Wolfenbüttel) wies bezüglich der Editionsprojekte am IOS darauf hin, dass die Wiedergabe historischer Dokumente als Digitalisat allein noch keine digitale Edition ausmache. PDF, das bei der Digitalisierung meistgenutzte Format, bezeichnete er in diesem Zusammenhang als „Inkunabel des digitalen Zeitalters“. Konstitutiv für eine digitale Edition sei vielmehr, dass Daten nicht nur gelesen, sondern gerechnet werden könnten.⁷

Die Komponenten ‚echter‘ digitaler Editionen erläuterte Thomas Stäcker im Anschluss anhand der Digitalisierungspraxis an der Herzog August Bibliothek Wolfenbüttel (HAB). Auf Basis faksimilierter Anzeige des Quelloriginals, das dabei auch extern über Verlinkung eingebracht werden könne, wobei jedoch sowohl Rechtsfragen als auch die Zuverlässigkeit der Verlinkung zu beachten seien, werde die Transkription dargestellt. Die Transkription sei das erklärte Ziel jeder Edition, aber in analogen Ausgaben schlicht nicht in vollem Umfang zu realisieren. Via *double keying* oder OCR ließe sich das Quantum bewältigen⁸, wenn auch mangelnde Genauigkeit weiterhin ein großes Problem darstelle. Die eigentliche Generierung der digitalen Edition beruhe auf der XML-basierten Kodierung der Transkription, welche die semantischen Informationen des linguistischen Basistextes fachlich-thematisch nach den Standards der Text Encoding Initiative (TEI) aufschlüssele und maschinenlesbar mache. An dieser Stelle sei die Herausforderung für den Editor zu verorten, nicht nur die Auszeichnung des Textes durch XML zu bewältigen, sondern zugleich die vielseitigen TEI-Standards durch eigene Regeln eindeutig auf die Edition zuzuschneiden: Ersteres böte dabei ein neues Feld zur Kooperation mit Bibliothekaren, letzteres entspräche den genuinen Aufgaben des Wissenschaftlers, der in der Einleitung zu seiner Edition klare Richtlinien für seine Forschung formuliere. Allgemein kann festgehalten werden, dass für die Arbeit an der Kodierung wie für das klassische Verlagswesen gilt: das Produkt ist nicht das Buch oder die online-Oberfläche, sondern die abstrakte En-

⁵ <<http://www.digitale-edition.de/>> (02.05.2013).

⁶ Vgl. <<http://www.recensio.net/presentationen/>> (02.05.2013).

⁷ zur terminologischen Abgrenzung vgl. Patrick Sahle: Digitales Archiv – Digitale Edition. Anmerkungen zur Begriffsklärung, in: Michael Stolz (Hrsg.), Literatur und Literaturwissenschaft auf dem Weg zu den neuen Medien, Zürich 2007, S. 64-87.

⁸ Die digitale Edition „negates the importance of selection as a formal principle.“, Sören Steding zitiert nach: Eva Maria Glaser, Digitale Edition als Gegenstand bibliothekarischer Arbeit: Probleme, Umsetzung und Chancen am Beispiel der Wolfenbütteler Digitalen Bibliothek (WDB), Berlin 2013 (Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft, 339), S. 23.

tität Text. Mit Übersetzung, Einleitung und Kommentar, Register und Bibliographie werde die Edition – analog zu ihrem gedruckten Pendant – vervollständigt. Die digitale TEI-kodierte Variante überträte allerdings auf der Funktions- und Strukturebene die Druckvariante bei Weitem, würde wie an der HAB beispielsweise von Zitaten im Quelltext auf deren Volltext gesprungen, das Register mit der Gemeinsamen Normdatei (GND) verknüpft (Erhöhung der Sichtbarkeit), Kontexte angezeigt, Vergleichs- und Suchfunktion angelegt. Insofern bewege sich das Konzept „Edition“ von einem reinen Wissenscontainer in Richtung einer neu zu bestimmende Kategorie (knowledge site, Forschungs Umgebung).

Einen spezifischen Typus der Edition beschrieb SIEGFRIED GRUBER (Rostock) mit dem Datenbankprojekt „Mosaic“ am Max-Planck-Institut für demografische Forschung.⁹ Verfügten die USA¹⁰ und Nordwesteuropa¹¹ mit belastbaren Datenbeständen zu historischen Bevölkerungen über wirksame Forschungsinstrumente, bestehe diesbezüglich in Kontinentaleuropa eine große Lücke, die im Verlust von Quellenbeständen, aber auch mangelnder Zusammenarbeit von Historikern und Genealogen begründet liege. Mosaic steuere diesem Umstand entgegen, indem zensusrelevante Quellenbestände vor 1900 gesichtet, stichprobenweise geprüft und schließlich in einer Datenbank harmonisiert werden. Mosaic greife dabei nicht nur auf Volkszählungsdaten, sondern auch auf kirchliche Quellen, Kopfsteuerbeschreibungen und Einwohnerzählungen zurück. Besonderen Wert lege man auf die qualitative Substanz der Daten, die, um Vergleichbarkeit zu garantieren, feste Merkmale besitzen müssten: Ort / Jahr der Zählung, Angabe von Namen und Alter, Unterscheidbarkeit von Einzelhaushalten etc. Die Daten würden in der Regel extern kodiert und unter Anwendung von 30 Variablen dokumentiert, die nicht nur als Identifikatoren der erfassten Personen bzw. Haushalte fungieren, sondern zusätzlich qualitative Zusammenhänge (etwa Verwandtschaftsverhältnisse) abbildeten. Mosaic stelle derzeit in acht Länderverzeichnissen mit 17 standardisierten und sieben unstandardisierten Dateien Informationen zu beinahe 380.000 Personen zur Verfügung, die

kostenlos zugänglich und in Einzeldatensätzen zitierfähig sind. Da Mosaic keine weiteren Services integriert, die Datenauswertung also allein beim Nutzer liegt, illustriert es mit der Verfügbarmachung von Forschungsdaten eine Möglichkeit innovativer elektronischer Publikationsformen.

Digitale Editionen nicht als starres kanonisiertes Endprodukt, sondern als eine offenstehende Ressource zur Weiterverarbeitung von Texten und Metadaten im Hinblick auf „neue Publikationsmaßstäbe“ zu begreifen, dafür plädierte MICHAEL PIOTROWSKI (Mainz). Bei online edierten Texten den semantischen Träger ‚Sprache‘ zu berücksichtigen, ein Bewusstsein für die Spezifika historischer Sprachen zu entwickeln, sei unabdingbar, um mit der Weiterentwicklung von Techniken automatischer Spracherkennung (Natural Language Processing, NLP) die Tiefe historischer Texte voll zu durchdringen. NLP ermögliche durch Verfahren wie Termextraktion, automatisierte Themenerkennung (Topic Modelling), Lemmatisierung, speech tagging und Analysefunktionen die Chance zur Analyse und Durchsuchbarkeit großer Textmengen und damit zu einer Automatisierung des Erkenntnisgewinns. Um dieses Potential ausschöpfen zu können, müsse sich die community von exklusiv-projektbezogenem Denken lösen und auf eine stärkere Formalisierung hinarbeiten. Piotrowski umriss erste einfache Schritte auf diesem Weg und forderte eine möglichst standardisierte Dokumentation von Transkriptionsregeln und deren eindeutige Identifizierung über IDs (als PURL), so dass die Austauschbarkeit von Regeln gewährleistet sei. Er schlug des Weiteren ein Abrücken von der klassischen Editionsstruktur in Bänden vor, so dass aus für sich stehenden Einzeltexten verschiedener Editionsprojekte zeitlich und / oder räumlich kombinierte „virtuelle Textkorpora“ gebildet werden könnten – eine Voraussetzung für den Entwurf neuer kontrollierter Vokabulare. Piotrowski berührte damit die eingangs geäußerte Kritik an der „skandalösen“ Uneinheitlichkeit der Standards und legte die Verantwortung hierfür zurück in die Hände der eigent-

⁹ <<http://www.censusmosaic.org>> (02.05.2013).

¹⁰ <<https://usa.ipums.org/usa/>> (02.05.2013).

¹¹ <<https://www.nappdata.org/napp/>> (02.05.2013).

lichen Akteure, der Wissenschaftler, die ihre Forschung nachhaltiger dokumentieren und verknüpfen sollten.

Ein wie von Piotrowski gefordertes best practice-Beispiel lieferte SASCHA HINKEL (Münster) mit dem Langzeitprojekt der Edition der Nuntiaturreporte Eugenio Pacellis (1917-1929), des späteren Papstes Pius XII.¹² Während das Pontifikat von Pius XII. nicht erst seit Rolf Hochhuths Schauspiel „Der Stellvertreter“ wegen der Haltung des Heiligen Stuhls im Zweiten Weltkrieg umstritten bewertet würde, sei die Laufbahn Pacellis als Nuntius in München und Berlin weitgehend unbeachtet geblieben. Seit der Öffnung der Akten aus dem Pontifikat Pius' XI. können nun alle Berichte, Dokumentenentwürfe und Weisungen, die Pacelli zum Teil täglich nach Rom schickte, eingesehen werden. Insgesamt lägen mehr als 20.000 Dokumente vor, die ein neues Licht auf die besondere Beziehung Pacellis zu Deutschland werfen würden, aber auch die Außenpolitik der katholischen Kirche in der von politischen ‚Totalitarismen‘ geprägten Zwischenkriegszeit widerspiegeln könnten. Man habe sich, so Hinkel, angesichts der umfangreichen Quellengrundlage bei der Edition von pragmatischen Ansätzen mit dem Ziel schneller Veröffentlichung bei gleichzeitiger maximaler Verwertbarkeit leiten lassen. Als technische Basis konnte man auf ein von den Deutschen Historischen Instituten Rom und London entwickeltes System zurückgreifen, das den Editor in allen Phasen von der Aufnahme der Daten bis zur Veröffentlichung unterstütze. Das System Digitale Editionen Neuzeitlicher Quellen (DENQ)¹³ kombiniere eine XML-Datenbank, deren Eingabeoberfläche keine tiefgehenden IT-Kenntnisse voraussetze, mit einem Web-Interface zur Recherche. Die Datenbank eröffne Suche und Abruf von Personen-, Orts- und Schlagwortdatensätzen, die nicht mehr als klassische Fußnote, sondern in einem eigenen Fenster erschienen. Ein bemerkenswerter Vorteil entstehe durch Verknüpfungen der Texte – mit biographischen Daten via GND-Beacon, vor allem aber intern durch Parallelansicht von Ausfertigung eines Nuntiaturreports und den dazugehörigen Entwürfen, so dass Korrekturen und Entstehungsphasen nachvollziehbar würden, sowie Verweisen auf andere Berichte. Um der

Masse des Materials Herr zu werden, habe man sich entschieden, nur für die ausgefertigten Berichte Regesten auf Deutsch zu erstellen, die quasi als Metatext dienen, während der Bericht selbst und die Entwürfe im Original auf Italienisch verblieben. Zudem wären einige Aufgaben (Transkription) mit gutem Erfolg in die Öffentlichkeit (Crowd Sourcing) ausgelagert worden, was den Workflow beschleunige. Als mutiger Schritt des Projektes lässt sich bezeichnen, dass die Edition als Work in Progress entsteht, also transkribierte Berichte abrufbar sind, zu denen die Regesten und Kommentare später nachgereicht werden. Äußerst positiv reagierten die Teilnehmenden auf die sauber aufgearbeiteten Texte, die Funktionen der Suchmaske und die Nutzbarkeit durch feine Zitierbarkeit auf PURL-Ebene. Bedauert wurde allerdings, dass die Originale nicht gescannt veröffentlicht werden durften (das Vatikanische Geheimarchiv habe schließlich einen Ruf zu verteidigen), und dass solcher Komfort bei der Arbeit an Editionsprojekten mit einem technischen, personellen und damit finanziellen Rahmen verbunden sei, der vielen Einrichtungen bzw. Projekten nicht zur Verfügung stünde.

In seinem abschließenden Resümee kam PATRICK SAHLE (Köln) zum Ausgangspunkt zurück, „Wildwuchs im Editionsbe- reich zu reduzieren“. Er erinnerte daran, dass die digitale Editorik bereits auf bestehende Standards zurückgreifen könne (XML, TEI), dass zudem die Diskussion methodologisch- standardisierender Untersuchungen auf Beteiligung warte.¹⁴ Das klassische Editionswe- sen sei schließlich auch in einem Prozess stän- digen Austauschs entstanden. Bei allen in- dividualisierbaren Varianten der Wiederga- be oder Funktionalitäten dürften die „edito- rischen Basics“ nicht aus dem Blick geraten: Klare Regeln, klare Aufgabenstellung, klare Identifikation der Autoren in einem Vorwort seien das Mindestmaß, um Nutzbarkeit und Kredit digitaler Arbeiten zu garantieren. Sah-

¹² <<http://www.pacelli-edition.de>> (02.05.2013).

¹³ Vgl. Jörg Hörnschemeyer: DENQ – Software für die Online-Edition, in: *digiversity*, <<http://digiversity.net/2010/denq/>> (02.05.2013).

¹⁴ Vgl. Paper zur Rezeptionspraxis, Patrick Sahle u.a., Kriterien für die Besprechung digitaler Editionen, Version 1.0, in: I.D.E., <<http://www.i-d-e.de/aktivitaeten/reviews/kriterien-version-1>> (12.05.2013).

le prognostizierte den mittelfristigen Bedarf an technischer Kompetenz bei den wissenschaftlichen Institutionen selbst, wollten sie der Komplexität und Abstraktheit digitaler Editionen, die „Print plus x“ böten, gerecht werden. Vorerst regte er an, gute Projekte zu dokumentieren, sie nachzunutzen und aus ihnen generische Modelle zu entwickeln, wozu der modulare Aufbau digitaler Strukturen (Baukasten) einläde. Ein Weg zu diesen Modellen könne in Richtung von Portalbildungen und Serialisierungen, wie an der HAB, weisen. Portalbildungen förderten denn auch Verstetigung, Entlastung bei Nachbesserungen und nicht zuletzt die sichtbare Profilbildung der Beteiligten.¹⁵

Projekte wie Mosaic oder die Pacelli-Edition ermutigen zur Arbeit an digitalen Editionen. Allein die Aussicht, im WWW genug Raum zu finden, um nicht mehr nach dem *Entweder - Oder* fragen zu müssen, scheint eine Defragmentierung des Wissens zu versprechen. Bei allen Problemen bezüglich der Sichtbarkeit der Autorschaft, bei allen Vorbehalten gegenüber Versionierungen und kollaborativen Forschungsmodellen wäre es wichtig, Wissenschaftlerinnen und Wissenschaftler davon zu überzeugen, dass nicht nur der eigene Name auf einem Buchtitel, sondern gerade auch die aktive Mitgestaltung der Entwicklung hoch spezialisierter Standards, wie in der E-Editionswissenschaft notwendig, Reputation erzeugt. Zudem ließen sich dadurch in den Geisteswissenschaften ungewohnte Erfahrungen gewinnen, schließlich sei, so Patrick Sahle, „die digitale Edition ein Labor.“

Konferenzübersicht

Teil 1

Ulf Brunnbauer (Regensburg): Grußwort

Konrad Clewing / Peter Mario Kreuter (beide Regensburg): Neue Leitlinien für neuartiges Publizieren? Digitale Quellenvorhaben am Institut für Ost- und Südosteuropaforschung

Thomas Kuster (Innsbruck): Das italienische Reisetagebuch Kaiser Franz I. von Österreich aus dem Jahre 1819. Druck und digitale Ausgabe im Rezeptionsvergleich (Vortrag fiel aus)

Thomas Stäcker (Wolfenbüttel): Digitale

Quelleneditionen an der Herzog August Bibliothek Wolfenbüttel. Editorisch-inhaltliche Aspekte und technische Lösungen

Siegfried Gruber (Rostock): Digitale Quellaufbereitung am Beispiel des Mosaic-Projekts: Der Aufbau einer Datenbank zur Erforschung historischer europäischer Bevölkerungen

Teil 2

Michael Piotrowski (Mainz): Digitale Editionen: Technische Aspekte und aktuelle Entwicklungen

Sascha Hinkel (Münster): Die Nuntiaturreportagen Eugenio Pacellis 1917-1929 als Beispiel für eine tiefe Texterschließung

Patrick Sahle (Köln) / Moderation: Abschlussdiskussion. Gemeinsame Grundlinien digitaler historischer Editorik?

Tagungsbericht *E-Editionen: Neue Publikationsmaßstäbe für historische Quellenprojekte*. 12.04.2013, Regensburg, in: H-Soz-Kult 30.07.2013.

¹⁵ Als best practice nannte Sahle die *Éditions en ligne de l'École des chartes*, <<http://elec.enc.sorbonne.fr/>> (12.05.2013).