

Neue Erkenntnisse durch digitalisierte Geschichtswissenschaften? Zur hermeneutischen Reichweite aktueller digitaler Methoden in informationszentrierten Fächern.

Während mit computerlinguistischen und auf Statistik beruhenden Methoden in der Verarbeitung natürlicher Sprache interessante neue Ergebnisse erzielt und neue Forschungsfragen aufgeworfen worden sind (Autorenschaftszuweisungen, Plagiatserkennung etc.), scheint die digitale Geschichtswissenschaft jenseits der Computerlinguistik bislang recht konventionell daher zu kommen. Die Mehrzahl der Angebote konzentriert sich auf Information Retrieval sowie digitale Editionen und soll mithin den klassischen Forschungsprozess erleichtern. Diese Entwicklung lässt sich direkt auf die Forschungsfragen und die Methoden dieser Fächer zurückführen: der historischen (Re-)Konstruktion historischen Geschehens auf der Grundlage von Einzelinformationen, die oft mühsam aus verschiedenen Quellen zusammengetragen werden. Welchen Mehrwert bietet die kritische digitale Edition in diesem Zusammenhang über die größere Freiheit der Darstellung der Textgenese hinaus, wenn die gewonnen Einzelinformationen am Ende erneut im Kopf des Forschers neu zusammengesetzt werden? Und wenn eine computerbasierte, (teil)autonome Form der Informationsverknüpfung schon möglich wäre, würde man diese überhaupt wünschen?

Peter Haber hat in seinen Arbeiten die Entwicklung, aber auch zukünftige Potentiale der digitalen Geschichtswissenschaft mit Weitblick beschrieben. Bislang sind digitale Werkzeuge für Historiker oft nur die moderne Form des analogen Zettelkastens, mit dessen Hilfe schon im 19. Jahrhundert erstaunliches geleistet wurde. Welche Bereiche der digitalen Geschichtswissenschaft kann man also heute von wirklich neuen Methoden sprechen, die auch grundsätzlich neue Erkenntnisse erwarten lassen? Drei Bereiche erscheinen mir dabei besonders vielversprechend:

In den aktuellen Methoden der **Netzwerkanalyse** lassen sich in der Tat neue Einsichten in soziale Beziehungsstrukturen gewinnen, die aufgrund ihrer Komplexität früheren Forschergenerationen grundsätzlich verschlossen blieben. Nicht zuletzt aus diesem Grund einer qualitativen Neubewertung von seriellen prosopographischen Quellen findet diese Methode immer breitere Akzeptanz in den Geschichtswissenschaften.

Noch mutet OCR für Manuskripte ein wenig wie Science Fiction an, aber spätestens seitdem Frederic Kaplan den Plan einer Digitalisierung der Venezianischen Handschriften im dortigen Staatsarchiv (Venice Time Machine) verkündete, scheinen Optical Character and Structure Recognition von Manuskripten greifbarer geworden zu sein. Hier geht es um die Beziehung von **Digitalen Methoden** zu den **historischen Hilfswissenschaften** – ein bislang nicht ausreichend beleuchteter Bereich der digitalen Geschichtswissenschaft, der, ähnlich wie bei google books, aufgrund der reinen Quantität des so verfügbaren Materials einen qualitativen Durchbruch ermöglichen könnte. Dies gilt übrigens grundsätzlich auch für die born digital Quellen der Zeitgeschichte, die gerade erst entstehen und in Zukunft ganz neue Formen digitaler Hilfswissenschaft mit Blick auf die langfristige Lesbarkeit von Datenformaten notwendig machen werden.

Im Rahmen des Konzepts eines **Semantic Web** auf Grundlage von maschinenlesbaren Beschreibungen von Ressourcen rücken die Möglichkeit der künstlichen Intelligenz und des automatisierten „reasoning“ zumindest theoretisch in erreichbare Nähe. In der Realität wird die Extraktion impliziten Wissens aber in den Geschichtswissenschaften bislang kaum genutzt. Facettierungen bieten Hilfe bei der Beurteilung von Suchergebnissen, aber wirkliche „künstliche Intelligenz“ ist dies noch nicht. Offensichtliche Fehler in Schemata können erkannt werden, aber können Computer schon kreativ Informationen verknüpfen und daraus einen neuen „Gedanken“

generieren, der Ausgangs- oder Endpunkt einer Forschungsfrage sein könnte? Immerhin bringt das gemeinsame Datenformat die Möglichkeit mit sich, Informationen in neuer Weise miteinander automatisiert zu verknüpfen und so neue Zusammenhänge aufzudecken.

Der Beitrag diskutiert anhand der drei oben skizzierten Beispiele Chancen und Grenzen sowie die hermeneutische Reichweite digitaler Methoden in den Geschichtswissenschaften vor einem erkenntnistheoretischen Hintergrund. Besondere Aufmerksamkeit liegt dabei auf den Erkenntnismöglichkeiten, die in der Zusammenschau und automatischen Verknüpfung von verteilten Daten bzw. Informationen unter einer gemeinsamen Oberfläche liegen. Linked Data kann heute schon aktiv genutzt werden, um digitale Ressourcen miteinander zu verknüpfen und damit Zusammenhänge evident zu machen, die bislang erst nach serieller Rezeption im Kopf des oder der Forschenden entstanden. Mit sog. „**Mashups**“ können so qualitativ neue Ergebnisse aus schon in standardisierten Datenformaten digital vorliegenden Informationen gewonnen werden, die das Potential besitzen, bislang unbekannte Zusammenhänge sichtbar(er) werden zu lassen. In diesem Kontext spielen Verknüpfungspunkte wie die Normdaten zu Personen und Orten aber auch die dbpedia oder Ontologien als (vereinfachte) Konzepte des Weltwissens eine entscheidende Rolle.

Literatur:

Alves, Daniel (Hg.) (2014): Digital Methods and Tools for Historical Research: A Special Issue. International Journal of Humanities and Arts Computing, Bd. 8, Heft 1 (April).

Düring, Marten & Ulrich Eumann (2013): Historische Netzwerkforschung: Ein neuer Ansatz in den Geschichtswissenschaften, in: Geschichte und Gesellschaft, Bd. 39, Heft 3, S. 369-390.

Dougherty, Jack & Kristen Nawrotzki (Hg.) (2013): Writing history in the digital age / Ann Arbor: University of Michigan Press.

Endres-Niggemeyer, Brigitte (Hg.) (2013): Semantic Mashups. Intelligent Reuse of Web Resources, New-York: Springer.

Haber, Peter (2011): Digital Past. Geschichtswissenschaft im digitalen Zeitalter, München.

Knap, Tomáš, Michelfeit J., Nečaský M. (2012): Linked Open Data Aggregation: Conflict Resolution and Aggregate Quality, in 36th Annual IEEE Computer Software and Applications Conference Workshops, Izmir, Turkey, IEEE Computer Society, S. 106-111.

Meroño-Peñuela, Albert, Schlobach, Stefan, van Harmelen, Frank (2013): "Semantic Web for the Humanities", In: Proceedings of the 10th Extended Semantic Web Conference, ESWC 2013, Montpellier, France, May 28-30, 2013. Philipp Cimiano et al. (Eds.), Lecture Notes in Computer Science 7882, Berlin Heidelberg: Springer, S. 645-649.

Reichert, Ramón (2014): Einführung, in: Big Data, hg. von Ramón Reichert, Bielefeld, S. 9-31.

Rosenthaler, Lukas, Fornaro, Peter, Clivaz, Claire (2014): National Data Curation and Service Center for Digital Research Data in the Humanities, Conference Proceedings DH 2014, Lausanne, S. 338-340.

Sacramento, Eveline R.; Casanova, Marco A.; Breitman, Karin K.; Furtado, Antonio L.; Macêdo, José Antonio F. de; Vidal, Vânia M. P. (2012): Dealing with inconsistencies in linked data mashups, Proceedings of the 16th International Database Engineering & Applications Symposium [http://delivery.acm.org/10.1145/2360000/2351496/p175-sacramento.pdf].

Sahle, Patrick (2013): Digitale Editionsformen, Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, 3 Bände, Norderstedt: Books on Demand. (Schriften des Instituts für Dokumentologie und Editorik 7-9).

Sequeda, Juan , Marcelo Arenas, Daniel P. Miranker (2012): On directly mapping relational databases to RDF and OWL. WWW 2012, S. 649-658.

Sequeda, Juan, Miranker, Daniel P. (2013): Ultrawrap: SPARQL execution on relational data. Journal of Web Semantics, Bd. 22, S. 19-39.

Wettlaufer, Jörg & Sina Westphal (2014): "Digital Humanities", Der Archivar, Heft 3, S. 270-277.

Zaagsma, Gerben (2013): On Digital History, Low Countries Historical Review, Bd. 128, Heft 4, S. 3-29.

Dr. Jörg Wettlaufer

Göttingen Centre for Digital Humanities (GCDH)

Akademie der Wissenschaften zu Göttingen (ADWG)

Papendiek 16

37073 Göttingen

Germany

Tel. +49 551 39 20477

email: jwettla@gwdg.de

<http://www.gcdh.de/en/people/researchers-project-staff/joerg-wettlaufer/>