

Geschichtsforschung im digitalen Raum. Über die Notwendigkeit der Digital Humanities als historische Grundwissenschaft

Malte Rehbein, Passau (Stand 16. April 2015, Preprint)

Einleitung¹

5 „The purpose of computing is insight, not numbers“, so ein berühmtes Zitat des US-amerikanischen Mathematikers Richard Hamming (1915-1998).² Die Nutzung von Computern sei nicht reiner Selbstzweck, sondern diene etwas Höherem, dem Gewinnen von Einsicht und Erkenntnis. Im Kontext eines „computational turns“ auch in den Geschichtswissenschaft, der häufig unter dem allgemeinen Dach der „Digital Humanities“ gefasst wird, können wir unter Abwandlung von Hamming's Zitat als
10 Postulat formulieren: The purpose of these Digital Humanities is insight, not computers. Nicht die Zahlen und die Zählbarkeit sind der Zweck einer Geschichtsforschung im digitalen Raum, sondern das Gewinnen und Mehren historischer Erkenntnis.

Die Abwandlung des berühmten Zitats von Hamming soll zu Beginn dieses Beitrages folgendes verdeutlichen: sprechen wir von digitaler Geschichtsforschung oder Geschichtsforschung im digitalen
15 Raum, dann meinen wir primär Geschichtsforschung, die von Historikerinnen und Historikern betrieben wird, deren Ziel es ist, zur historischen Kenntnis und Erkenntnis beizutragen, wobei das Digitale ein Mittel zum Zweck ist. Die gegenwärtigen und künftigen Möglichkeiten, die uns die Computertechnologie bietet, mögen dabei andere Fragestellungen hervorbringen als es bisherige Methoden vermochten oder neue Akzente setzen, Sichtweisen erweitern und Perspektiven
20 verschieben. Und es werden ganz sicher neue methodische Ansätze ermöglicht, von denen einige in diesem Band vorgestellt werden.³ Aber: egal ob digital oder nicht digital, es wird das gleiche Ziel

¹ Dieser Beitrag basiert auf dem Erlanger Vortrag vom 21. Februar 2015. Der Text wurde stark überarbeitet, wobei der Vortragsstil im Wesentlichen beibehalten wurde. Gegenüber dem Vortrag ergänzt wurde unter anderem die eigene statistische Auswertung von Daten der Neuen Deutschen Biographie.

² Hamming, Richard W. (1962): Numerical methods for scientists and engineers. New York, London: McGraw-Hill (International series in pure and applied mathematics), 276.

³ Es ist zu vermuten und teilweise bereits auch schon zu beobachten, dass die Methoden der Digital Humanities das historische Arbeiten verändern. Dies begründet aber keine neue *historische* Disziplin. Zwar sind die Digital Humanities ein eigenständiges Forschungs- und Lehrgebiet und stellen damit eine akademische Disziplin dar. Aus Sicht der Geschichtswissenschaft sind sie aber eine (wie ich noch zeigen möchte: elementare) Grund- oder Hilfswissenschaft. Der Begriff der „Digital Humanities“, der sich seit etwa Anfang des 21. Jahrhunderts international durchgesetzt hat, ist dabei leider etwas irreführend, impliziert er doch einen Gegensatz von *Digital Humanities* und *non-Digital* (traditionellen) Humanities bzw. schlägt die Existenz einer parallelen, eben digital betriebenen Geisteswissenschaft vor, die disjunkt von der sonstigen Geisteswissenschaft zu verstehen sei. Stattdessen bieten aber die Digital Humanities nicht ersetzende, sondern *zusätzliche* und komplementäre Methoden zur Gewinnung von Erkenntnis an, und dies tun sie nicht erst seit heute. Zum Ursprung des Begriffs „Digital Humanities“ siehe Humanist Discussion Group, Vol. 28, No. 840 (<http://www.dhhumanist.org/>).

verfolgt, nämlich das der historischen Erkenntnis: „Die Digital Humanities bereichern die traditionellen Geisteswissenschaften konzeptionell und methodisch - ihre Werkzeuge und Verfahren ergänzen das ‘Wie’ unserer Praxis um eine empirisch ausgerichtete Epistemologie“.⁴

Zu Stand und Perspektiven

5 Die Frage nach den Geschichtswissenschaften in einem *digitalen Raum*⁵ stellt sich dabei keineswegs erst heute. Wir wissen, dass Roberto Busa bereits im Jahr 1949 begonnen hat, IBM-Rechenmaschinen zu benutzen, um die Erstellung einer Wort-Konkordanz der Werke Thomas von Aquins, den Index Thomisticus, maschinell zu unterstützen.⁶ Zum Durchbruch verhalf der
10 Elektronischen Datenverarbeitung in den Geschichtswissenschaften dann das Auftreten des Personal Computers. Und nun stellte man fest, dass die Interessen, Computer für die Geschichtsforschung einzusetzen, doch recht breit gestreut waren; der digitale Raum füllte sich zunehmend, und man begann zu fragen, wie sich denn die Geschichtsforschung darin positionieren könne und solle.

15 1986 wurde die internationale Association for History and Computing gegründet,⁷ einige Jahre später folgte die deutsche „Arbeitsgemeinschaft Geschichte und EDV“, die jüngst ihr 20-jähriges Bestehen mit einem Sammelband feierte.⁸ Interessant hierbei ist, welche Themen vor 20 Jahren diskutiert

⁴ Thesenpapier „Digital Humanities 2020“ des Fachverbandes, vorgestellt im März 2014 (<http://www.dhd2014.uni-passau.de/programm/dh-2020/>).

⁵ Der ursprüngliche Vortragstitel lautete „Geschichtswissenschaft im digitalen Zeitalter“. Dass sich aus dem gegenwärtigen (gesellschaftlichen) „computational turn“ aber etwas entwickelt, das legitim als ein „Zeitalter“ bezeichnet werden kann, sollten wir der Retrospektive künftiger Generationen überlassen, die die Frage, ob es sich vielleicht nur um eine Übergangserscheinung oder um ein Phänomen, das von weiterreichenden gesellschaftlichen Entwicklungen überstrahlt wird, beantworten können werden. Dass wir uns in einer digitalen oder zunehmend digitalisierten Welt bewegen, steht wohl außer Zweifel, weswegen ich im Folgenden lieber vom „digitalen Raum“ spreche.

⁶ Busa, Roberto (1951): Sancti Thomae Aquinatis Hymnorum Ritualium Varia Specimina Concordantiarum: A First Example of Word Index Automatically Compiled and Printed by IBM Punched Card Machines. Milano: Fratelli Bocca, Editori. Busa gilt damit als der „Gründungsvater“ der Digital Humanities. Dabei dachte er bereits 1951 an „various interesting extensions of this system“ (ebd., 36) und skizziert in faszinierender Weitsicht die Zukunft der Technologie für die Textwissenschaften: “[...] it will be possible to make extremely quick comparative analyses of the composition and frequency of the vocabulary of various authors, useful for example in psychological research, criticism of texts or historical relationships; searching for all the words containing a given root; glottological comparisons of different languages and so forth” (ebd.).

⁷ Thaller, Manfred (2013): Die International Association for History and Computing. In: Jörn Kobes, Kai Ruffing und Wolfgang Spickermann (Hg.): 20 Jahre Arbeitsgemeinschaft Geschichte und EDV. Gutenberg: Computus-Dr. Satz & Verl (Abhandlungen der Arbeitsgemeinschaft Geschichte und EDV, 2), S. 25–41.

⁸ Kobes, Jörn; Ruffing, Kai; Spickermann, Wolfgang (Hg.) (2013): 20 Jahre Arbeitsgemeinschaft Geschichte und EDV. Arbeitsgemeinschaft Geschichte und EDV. Gutenberg: Computus-Dr. Satz & Verl (Abhandlungen der Arbeitsgemeinschaft Geschichte und EDV, 2).

- wurden. Die Tagung der Arbeitsgemeinschaft von 1994 hatte acht Themenkomplexe,⁹ wovon uns die meisten recht bekannt vorkommen, denn diese Art von Diskussionen, wie sie damals angestoßen wurden, führen wir noch immer. So erörterte man 1995 etwa „Chancen und Risiken des World Wide Web für Historiker und Soziologen“,¹⁰ während auf dem Göttinger Historikertag von 2014 eine
- 5 Sektion „Was gewinnt und was verliert die historische Forschung durch Science 2.0?“ auf dem Programm stand.¹¹ Oder man überlegte damals, wie ein Curriculum zu Geschichte und EDV gestaltet werden könne; jüngst ist dies das Thema einer Arbeitsgruppe innerhalb des 2012 in Hamburg gegründeten Fachverbandes Digital Humanities im deutschsprachigen Raum (DHD),¹² die erste Ergebnisse im März 2015 in Graz vorstellte.¹³
- 10 Es ist durchaus erstaunlich, dass Fragen nach der Verortung des Digitalen in Forschung und Lehre immer wieder und immer noch diskutiert werden. Pioniere wie zunächst Roberto Busa,¹⁴ dann mit dem Aufkommen der PCs und später des Internets in der Geschichtswissenschaft im deutschsprachigen Raum Manfred Thaller¹⁵ oder in den Niederlanden Onno Boonstra, Leen Breure

⁹ Spickermann, Wolfgang (2013): 20 Jahre AGE: Konzepte und Akteure. In: Jörn Kobes, Kai Ruffing und Wolfgang Spickermann (Hg.): 20 Jahre Arbeitsgemeinschaft Geschichte und EDV. Gutenberg: Computus-Dr. Satz & Verl (Abhandlungen der Arbeitsgemeinschaft Geschichte und EDV, 2), S. 11–23, hier: S. 14.

¹⁰ Ebd., 16.

¹¹ Programm „Gewinner und Verlierer“ des 50. Deutscher Historikertag, Göttingen, 23. - 26. September 2014, S. 40-41.

¹² <http://www.dig-hum.de>

¹³ Während es vor zwanzig Jahren vornehmlich um einzelne Lehrveranstaltungen unter dem Schlagwort „Geschichte und EDV“ ging, liegt der Schwerpunkt heute auf der Frage, wie Studiengänge in Digital Humanities aufgebaut und gestaltet werden können. Leider aber machen wir uns weiterhin viel zu wenig Gedanken darüber, wie die notwendige digitale Kompetenz innerhalb der historischen Proseminare und Übungen vermittelt werden kann. Zur Grazer Tagung siehe: <http://dhd2015.uni-graz.at>; zur Arbeitsgruppe: <http://www.dig-hum.de/arbeitsgruppe-referenzcurriculum-digital-humanities>. Siehe auch Rehbein, Malte; Sahle, Patrick (2013): Digital Humanities lehren und lernen. Modelle, Strategien, Erwartungen. In: Heike Neuroth, Norbert Lossau und Andrea Rapp (Hg.): Evolution der Informationsinfrastruktur. Kooperation zwischen Bibliothek und Wissenschaft. Glückstadt: vvh Hülsbusch, S. 209–228. Prinzipielle und pragmatische Fragen zur Lehre in den Digital Humanities zudem diskutiert Hirsch, Brett D. (Hg.) (2012): Digital humanities pedagogy. Practices, principles and politics. Cambridge: Open Book Publ.

¹⁴ Zu seiner 1949 begonnenen Arbeit am Index Thomisticus schreibt Busa: “I have in this way given an example [...] for the great and small facilities with which [...] accounting and statistical machines can bestow index material to men of letters. Other facilities can be excogitated” (Busa, Roberto (1951): Sancti Thomae Aquinatis Hymnorum Ritualium Varia Specimina Concordantiarum: A First Example of Word Index Automatically Compiled and Printed by IBM Punched Card Machines. Milano: Fratelli Bocca, Editori, 48).

¹⁵ Exemplarisch: Thaller, Manfred (1982): Numerische Datenverarbeitung für Historiker, Wien-Köln (=Materialien zur Historischen Sozialwissenschaft 1); Ders. (1991): The Historical Workstation Project, in: Historical Social Research / Historische Sozialforschung 16; Ders. (1993): Kleio. A Database System, St.

und Peter Doorn erkennen frühzeitig das Potential und die Perspektiven der digitalen Methoden. Letztere postulieren 1990: „The historian who refuses to use a computer as being unnecessary, ignores vast areas of historical research and will not be taken serious anymore”.¹⁶ 14 Jahre später schauen die gleichen Autoren mit ihrer Veröffentlichung „Past, Presence, and Future of Historical Information Science“ jedoch etwas enttäuscht zurück. In ihrer gründlichen Analyse von 2004, die leider recht wenig Beachtung gefunden hat, stellen sie fest: „in the late 1980s and early 1990s the debates on history and computing flourished. The internet stimulated many heritage institutes [...] to digitize their collections. But since the late 1990s and in the first years of the 21st century, the ‘history and computing movement’ seems to have lost momentum.”¹⁷ Sie erläutern weiter: “‘history’ [im Sinne der Geschichtswissenschaft; MR] has failed to acknowledge many of the tools ‘computing’ had come up with”, ein “computing”, das beabsichtige, die Historiker mit “computerized tools and methods” zu versorgen, die sie benutzen könnten, um das Potential und die Qualität ihrer Forschung zu verbessern.¹⁸

Boonstra, Breure und Doorn analysieren Vergangenheit und Gegenwart dessen, was sie noch recht präzise mit “Historical Information Science” bezeichnen und was wir heute gerne mit dem allgemeineren Sammelbegriff “Digital Humanities” versehen. Dabei definieren sie, ohne Anspruch auf Vollständigkeit zu erheben, sieben „areas of further research“ in diesem Gebiet: 1. „modelling sources and user behaviour“, 2. „supporting editorial processes“, 3. „discovering structures and patterns“, 4. „tuning statistical techniques to historical research“, 5. „tuning information retrieval to historical requirements“, 6. “multimedia, reconstruction and simulation”, 7. „publishing historical databases“.¹⁹ Die Prognose von 2004 erscheint zumindest auf den ersten Blick recht zutreffend, obwohl sich die Entwicklung der Technologien und ihrer Anwendungen in den Geschichtswissenschaften seitdem in vielseitige und teils auch schwer zu erahnende Richtungen

Katharinen (= Halbgraue Reihe zur Historischen Fachinformatik B 11); Ders. (2003): Historische Fachinformatik - Ein Kölner Modell. In: Daniel Burckhardt, Rüdiger Hohls und Vera Ziegeldorf (Hg.): Geschichte und Neue Medien in Forschung, Archiven, Bibliotheken und Museen. Tagungsband.hist 2003. Berlin: Humboldt-Univ (Veröffentlichungen von Clio-online, 7, Teilbd. 1), S. 45–72.

¹⁶ Boonstra, O. W. A.; Breure, Leendert; Doorn, Peter K. (1990): Historische informatiekunde. Inleiding tot het gebruik van de computer bij historische studies. Hilversum: Verloren.

¹⁷ Boonstra, Onno W.; Breure, Leendert; Doorn, Peter (2004): Past, present and future of historical information science. In: Historical social research, 7.

¹⁸ Ebd., 9.

¹⁹ Ebd., 94. Es wäre sicherlich wertvoll, an dieser Stelle aber nicht leistbar, für alle die von Boonstra und Kollegen prognostizierten sieben Aufgabenfelder die tatsächlichen Entwicklungen der letzten zehn Jahre genauer zu untersuchen.

weiterentwickelt. Die digitalen Geschichtswissenschaften haben viele der vorgeschlagenen Bereiche tatsächlich erreicht. In diesem Fortschreiten digitaler Verfahren für die Geschichtswissenschaften zeigt sich zugleich aber auch das Risiko einer Fehlentwicklung, das im weiteren Verlauf dieses Beitrages thematisiert wird.

5 Verweilen wir zunächst noch einen Moment lang bei den „areas of further research“, die ja zunächst recht technisch formuliert sind, aber einen ganz guten Ausblick in die Geschichtswissenschaften im digitalen Raum bieten. Modellieren wir die Arbeit der Historiker einmal salopp als den Dreiklang von „lesen“, „denken“ und „schreiben“, so finden wir bei den Überlegungen von Boonstra, Breure und Doorn zur Computer-Unterstützung der Geschichtswissenschaften genau diese Trias wieder:

10 „content creation“, „intelligence and selection“ und „presentation“.²⁰ Eine solche Dreiteilung erscheint günstig, um die doch große Breite an Computer-Unterstützung in den Geschichtswissenschaften durch Klassifizierung der Methoden und Werkzeuge in den Griff zu kriegen.²¹ Der erste Bereich, die Digitalisierung bzw. „content creation“ ist Grundlage für Vieles, vor allem für die weitere Verarbeitung von Date; er unterstützt jedoch vorrangig das Lesen. Wer will

15 heute noch ernsthaft bestreiten, dass die digitale Verfügbarkeit von Quellen und Sekundärliteratur erhebliche Vorteile bringt? Vorteile allerdings nur für diejenigen, die kritisch (im Sinne der Informationskritik) mit diesen Daten und der Datenflut umgehen können, eine Informationskritik, an der es leider häufig mangelt. Der dritte Bereich „presentation“ oder Kommunikation wird im Kontext von Forschungsinfrastrukturen und neuen Publikationsformen bereits weitreichend diskutiert.²² Ich

20 möchte mich daher auf den zweiten Bereich konzentrieren. „Intelligence and selection“ nennen es Boonstra, Breure und Doorn, ich würde im Englischen wohl den Begriff des „reasoning“ bevorzugen. Also: wie unterstützen uns digitale Methoden beim Denken und Schlussfolgern, und warum besteht auch hier die Gefahr einer mangelhaften Informationskritik? Was kann eine Konsequenz für die Geschichtswissenschaft im digitalen Raum sein, wenn die Potentiale der computergestützten

²⁰ Ebd., 95.

²¹ Ähnliche Klassifizierungen finden wir auch in Peter Habers „Digital Past“ oder für den Bereich der Kunstgeschichte in Hubertus Kohles „Digitale Bildwissenschaften“, zwei der wenigen aktuellen Publikationen, die brauchbare Ansätze zu einer Theoriebildung für das Digitale in den Geschichtswissenschaften beitragen. Siehe: Haber, Peter (2011): Digital past. Geschichtswissenschaft im digitalen Zeitalter. München: Oldenbourg und Kohle, Hubertus (2013): Digitale Bildwissenschaft. Glückstadt: Hülsbusch. Auch das Studienangebot der Universität Passau zu Digital Humanities greift diese Dreiteilung auf. Die Bereiche heißen dort „Digitalisierung des kulturellen Erbes“, „Computergestützte Informationsanalyse und –verarbeitung“ und „Wissenskommunikation“.

²² Einen ersten Überblick zu Forschungsinfrastrukturen verschafft Neuroth, Heike; Lossau, Norbert; Rapp, Andrea (Hg.) (2013): Evolution der Informationsinfrastruktur. Kooperation zwischen Bibliothek und Wissenschaft. Glückstadt: vwh Hülsbusch.

Methoden auch zum “reasoning” der Historiker weiter erschlossen und herangezogen werden?
Hierzu ein Fallbeispiel.

„Charting Culture“

Im Sommer 2014 legte der Kunsthistoriker Maximilian Schich mit seinen Mitarbeitern eine
5 diskussionswürdige Arbeit unter der Überschrift „A network framework of cultural history“ vor, die
in Science, einer der international wichtigsten Fachzeitschriften, allerdings für Naturwissenschaften,
publiziert wurde.²³ Schich et al. stellen hierin eine Methode vor, deren Grundidee es ist, Merkmale
von Kultur in Raum und Zeit zu kartieren und damit deren Entwicklung, also eine Art Kulturgeschichte
10 der Welt nachzuzeichnen. Annahmen, die dahinter liegen, umfassen: 1. die Menschheit bildete im
Laufe ihrer Geschichte kulturelle Zentren aus (Beispiele etwa: Athen, Rom); 2. diese kulturellen
Zentren seien durch die Anwesenheit von „notable individuals“ charakterisiert, also ein kulturelles
Zentrum sei ein Ort, an dem nachweislich viele bedeutende Personen wie Akademiker oder Künstler
leben; 3. diese Zentren verlagerten sich im Laufe der Zeit dadurch, dass die Personen wandern,
migrieren. Schichs et al. Hypothese ist nun, dass sich durch quantitative Analyse von biographischen
15 Daten einer großen Menge solcher „notable individuals“ Trends in der Kulturgeschichte messen und
visualisieren ließen, sich also etwa die Verlagerung von kulturellen Zentren im Zeitverlauf (sozusagen
die *longue duree* kulturell-räumlicher Entwicklung) durch statistische Auswertungen nachweisen und
sichtbar machen ließe.

Die Datenbasis, die er verwendet, ist dabei verhältnismäßig groß (im Falle von Freebase ca. 120.000
20 Datensätze), aber die Daten sind nach einer Vorverarbeitung einfach strukturiert: ausgewertet wird,
wer wann und wo geboren wurde und wann und wo stirbt.²⁴ Die Daten stammen aus verschiedenen
Datenbanken, die wichtigste hierbei ist freebase.com. Freebase ist eine sich selbst als “community-
curated database of well-known people, places, and things” bezeichnete offen zugängliche
Datenbank,²⁵ die sich aus zwei Quellen speist: zum einen durch Zuarbeiten der “crowd”, das ist in

²³ Schich, M.; Song, C.; Ahn, Y.-Y.; Mirsky, A.; Martino, M.; Barabasi, A.-L.; Helbing, D. (2014): A network framework of cultural history. In: Science 345 (6196), S. 558–562. DOI: 10.1126/science.1240064.

²⁴ Auch die Berufsklassifizierung und das Geschlecht wurden herangezogen. Der Analyse voran ging eine Vorverarbeitung der Daten, wie die Autoren in den Supplementary Materials (<http://www.sciencemag.org/content/345/6196/558/suppl/DC1>) beschreiben. Auch stehen dort die auf diese Weise entstandenen Daten zur Überprüfung der formulierten Ergebnisse und für eigene Analysen zur Verfügung. Beispiele für solche Datensätze sind in den jeweiligen Datenbanken zu finden, für eine ausführliche Vorstellung und Erläuterung fehlt hier der Platz.

²⁵ <http://www.freebase.com/>. Die anderen herangezogenen Datensätze sind das *Allgemeine Lexikon der Bildenden Künstler von der Antike bis zur Gegenwart* von Thieme und Becker (AKL), die *Getty Union List of Artists*

etwa das Prinzip, das auch hinter Wikipedia steht, zum ändern durch ein sogenanntes “harvesting”, durch welches automatisiert Daten aus anderen Datenquellen zusammengetragen werden. Im Gegensatz zu biographischen Lexika findet man bei Freebase zwar biographische Daten zu Personen, aber keine beschreibenden Texte. Kern der Datenbank sind Metadaten in einer stark formalisierten

5 Aufbereitung. Das macht sie für eine quantitative Auswertung reizvoll und mit dem entsprechenden Know-how auch recht einfach zugänglich. In ihrer Arbeit analysieren Schich et al. nun solche biographischen Metadaten zu “notable individuals”, vor allem Künstler und Akademiker, und sie verwenden Geburts- und Sterbedaten (Zeitpunkt und Ort), die sie im Zeitverlauf visualisieren. Herausgekommen ist dabei unter anderem ein sehr ansprechendes, vor allem ästhetisch gut

10 gemachtes Video, publiziert als „nature video“ im Juli 2014 auf Youtube unter dem Titel “Charting Culture”.²⁶ Die Visualisierungen im Film wie im Science-Artikel zeigen dann etwa einen deutschen Föderalismus (“many subcenters compete with each other in federal clusters”) gegenüber einem französischen Zentralismus (“a winner-takes-all regime”) auf,²⁷ Migrationsströme von Europa nach Nordamerika²⁸ oder eine Ost-Westwanderung in Nordamerika.²⁹

15 Die Arbeit ist aus zwei Gründen bemerkenswert. Zum einen ist sie ästhetisch sehr ansprechend, die Visualisierungen im Film sind aufwendig gemacht. Sebastian Herrmann schreibt in der Süddeutschen Zeitung zu diesem Projekt: “Nie sah Geisteswissenschaft cooler aus”³⁰, und eine solche Arbeit wird in der Öffentlichkeit, nicht nur unter Fachleuten, sichtbar.³¹ Zum anderen aber, und das ist das Wichtige, ist die Arbeit bemerkenswert, weil sie ein ungeheures Potential (Chancen wie Risiken)

20 computerbasierter Verfahren gerade für die Geschichtswissenschaft aufzeigt. Lassen wir bei der folgenden Diskussion außer Acht, dass *Charting Culture* uns zunächst kaum etwas über die Entwicklung von Kultur(en) sagt, das wir nicht schon wüssten. Es geht bei diesem Ansatz primär nicht um die Gewinnung neuer historischer Erkenntnis, sondern darum, wie die Autoren selbst betonen,³²

Names (ULAN) sowie das Winckelmann *Corpus der antiken Denkmäler*. Die Daten von Freebase ragen deshalb ein heraus, weil sie Grundlage für das “nature video” (siehe unten) und Aussagen zur Mobilität bildeten.

²⁶ <https://www.youtube.com/watch?v=4glhRkCcD4U>.

²⁷ Schich et al., S. 562. Vgl. dort Abb. 1E und 3A sowie im genannten Video ab 1:56 min.

²⁸ Im Video ab 2:30 min.

²⁹ Im Video ab 3:30 min.

³⁰ SZ, 1.12.2014, online: <http://www.sueddeutsche.de/wissen/datenvisualisierung-europas-leben-und-sterben-im-zeitraffer-1.2071567>.

³¹ Ende März 2015 verzeichnet YouTube über 984.000 Besucher des Videos.

³² “As such, these results are consistent with Ravenstein’s laws of migration [...] and other empirical observations of human mobility [...] Our findings are nevertheless relevant, as (i) we can determine these

methodisch nachzuweisen, datenbasiert zu solcher Erkenntnis zu gelangen, ohne die Datensätze dabei einzeln lesen und interpretieren zu müssen. Quantitative Analyse, Data-driven Research, Distant Reading fallen hier als Stichworte ein, mit denen solche methodischen Ansätze bezeichnet werden.³³ Verankert sind diese Ansätze schon bei Boonstras, Breures und Doorns *areas of further research*, die sie mit “discovering structures and patterns” und “tuning statistical techniques to historical research” umschreiben. In diesem Verständnis zeigen Schich et al. Strukturen, Muster und Trends auf und verwenden Techniken der Statistik als Grundlage. Mit der folgenden kritischen Betrachtung der *Charting Culture* soll gezeigt werden, wie groß die Gefahr ist, das postulierte „tuning statistical techniques to historical research“ umzukehren, so dass eben nicht die statistischen Methoden für den Einsatz in der historischen Forschung maßgeschneidert werden, sondern die historische Forschung an die Statistik angepasst wird.



Abbildung 1: Screenshot des Nature Videos "Charting Culture" nach 0:29 Min.

Visualisierung eines Weltbilds

patterns from a relatively small fraction of birth and death location pairs, and (ii) we demonstrate that the patterns hold for more than eight centuries [...]" (Schich et al., S. 560).

³³ Zu "Distant Reading" vgl. Moretti, Franco (2005): *Graphs, maps, trees. Abstract models for a literary history*. London: Verso. Zu "Data-Driven Research" siehe unten. Auch der Begriff "Big Data" in den Geisteswissenschaften wird hier gerne genannt. Siehe etwa Schöch, Christof (2013): *Big? Smart? Clean? Messy? Data in the Humanities*. In: *Journal of Digital Humanities* 2 (3). Online verfügbar unter <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities>.

Abbildung 1, entnommen aus dem Video,³⁴ visualisiert "Migrationsströme" um das Jahr 1250. Die Grafik zeigt Konzentrationen, starke Knoten etwa in Mittelitalien, Rom wahrscheinlich, aber auch in West- und Mitteleuropa und England. Es handelt sich um den Nachweis von „notable individuals“ und ihrer Quantität zu diesem Zeitpunkt an diesen Orten, gewonnen aus den biographischen Daten der oben genannten Datenbanken. Im Verständnis der Informationsvisualisierung wird also ein visueller Eindruck von Strukturen der Verortung kultureller Zentren am Übergang vom hohen ins späte Mittelalter geboten.

Wenden wir nun aber in der gleichen Grafik unseren Blick nach Osten, etwa nach China. Hier offenbart die Visualisierung augenscheinlich, dass die Anzahl der "notable individuals" in China im Jahr 1250 deutlich geringer sei als die in Europa. Die offensichtliche Schlussfolgerung wäre die, dass die chinesische Kultur ebenso deutlich hinter der europäischen zurückläge, wie es der visuelle Eindruck hinterlässt, dass kulturelle Zentren (man ist geneigt zu sagen: Kultur) in diesem Raum kaum ausgeprägt gewesen wären. Doch aber war China zu Zeiten des europäischen Mittelalters kulturell hoch entwickelt. Frederick Mote charakterisiert die damals in China herrschende Sung-Dynastie (960-1279) als "bedeutendes Zeitalter der Pflege der Gelehrsamkeit, der höchsten künstlerischen und technischen Maßstäbe im Buchdruck und in der Förderung gewaltiger Projekte enzyklopädischer Sammlungen von Monumenten der Wissenschaft"; Gelehrsamkeit, Reichtum und materieller Wohlstand hätten "auf weit höherer Stufe [ge]standen, als sie in Europa oder wohl in jedem anderen Teil der Welt damals bekannt waren."³⁵ In der gleichen Visualisierung scheint etwa auch die mittelamerikanische Maya-Kultur, zu jener Zeit immerhin eine Schriftkultur, nicht zu bestehen; Afrika scheint kulturell gänzlich zu schlafen -- insgesamt ein visueller Widerspruch zu etabliertem Wissen über die Vergangenheit.

Die Autoren der *Charting Culture* sind sich dieser Problematik durchaus bewusst und sprechen in Bezug auf die "spatial coverage" von einem "bias", also einer Abweichung der Daten von ihrem statistischen Erwartungswert.³⁶ Die Raumabdeckung der Daten sei voreingenommen, unausgewogen, sie bevorzuge bestimmte geographische Räume gegenüber anderen. Dies sei kurz belegt.

³⁴ Nach 0:29 Min.

³⁵ Frederick W. Mote (1964): China von der Sung-Dynastie bis zur Ch'ing-Dynastie. In: Propyläen Weltgeschichte, hg. von Golo Mann und August Nitschke, 6. Auflage, 6. Band, S. 265-342, hier S. 281.

³⁶ Schich et al., Supplementary Materials, S. 7-8.

	Datensätze mit Todesjahr 20. Jh.	Einwohnerzahl 1950 (in Mio)	Statistischer Erwartungswert	Abweichung absolut	relative Abweichung
Deutschland	10,600	70.1	3,886	6,714	0.63
Frankreich	5,177	41.8	2,317	2,860	0.55
Italien	3,424	46.4	2,572	852	0.25
China	506	573.8	31,809	-31,303	-61.86
Japan	1,856	82.2	4,557	-2,701	-1.46
India	2,180	376.3	20,860	-18,680	-8.57
Kanada	5,674	13.7	759	4,915	0.87
USA	47,954	157.8	8,748	39,206	0.82
Ägypten	221	21.54	1,194	-973	-4.40
Mexiko	679	28.3	1,569	-890	-1.31
Summe	78,271	1,411.9	78,271	0	
Freebase gesamt	84,129	2,525.8			

Tabelle 1: Verzeichnete Personen aus ausgewählten Ländern in der Freebase-Datenbank (Stand: 31.3.2015).³⁷

Tabelle 1 zeigt eine Zählung der Datensätze zu den in der Freebase-Datenbank verzeichneten

5 Personen (mit Todesjahr im 20. Jahrhundert).³⁸ Fünf der hier gelisteten Länder (Deutschland,

³⁷ Erwartungswert und relative Abweichung beziehen sich auf die Gesamtzahl der hier herangezogenen Daten ausgewählter Länder, nicht auf alle Freebase-Daten. Als Referenzpunkt werden die Einwohnerzahlen des Jahres 1950 nach der UN-Weltbevölkerungsstatistik (<http://esa.un.org/unpd/wpp/Excel-Data/population.htm>, 7. April 2015) genommen. Dies berücksichtigt nicht die Veränderung der Einwohnerzahlen im Zeitverlauf, insbesondere nicht das starke Wachstum Chinas und Indiens im 20. Jahrhundert, wodurch die Abweichung vom Erwartungswert für diese Länder noch höher ausfallen würde.

³⁸ Die Auswertung wurde von mir durch Queries in der *Metaweb Query Language MQL* räumlich nach der Zuordnung „/people/person/nationality“ auf der Freebase-Datenbank unter <http://www.freebase.com/queries> am 6. April 2015 vorgenommen und spiegelt den Stand der Datenbank von diesem Tag wieder. Seit dem 31. März 2015 wurde die Datenbank auf „read-only“ gesetzt. Für die nationale Zuordnung wurden die Werte „Germany“, „France“, „Italy“, „China“, „India“, „Japan“, „Canada“, „United States of America“, „Egypt“ und „Mexico“ angesetzt. Dass der Begriff der „Nation“ vor allem im historischen Kontext durchaus problematisch ist, kann hier hingenommen werden, da es in der Auswertung nicht um absolute Werte geht, sondern lediglich darum, ihr ungefähres Verhältnis untereinander zu illustrieren. Ebenso hingenommen wird die Inkonsistenz der Freebase-Daten. So wird beispielsweise Thomas Manns Nationalität mit „German Empire“ (neben der seiner späteren tschechoslowakischen und US-amerikanischen) angegeben, die seiner Frau Katia jedoch mit

Frankreich, Italien, USA und Kanada) machen dabei lediglich 13,06% der Weltbevölkerung aus, werden aber durch 93,05% aller Datensätze der Freebase-Datenbank repräsentiert. Bei den in Tabelle 1 ausgewählten Ländern ist im Einzelnen zunächst die deutliche Abweichung vor allem Chinas gegenüber dem statistischen Erwartungswert auffällig. Bei einer Zufallsverteilung auf Basis der Einwohnerzahl wären über 31.000 Einträge zu erwarten, tatsächlich sind es nur 506. Anders etwa Deutschland (10.600 statt 3.886 Einträge) und die anderen europäischen und nordamerikanischen Länder. Für das zuvor diskutierte 13. Jahrhundert ist die Unausgewogenheit der Daten ähnlich: Mit den zentralen europäischen Ländern Frankreich, Deutschland, Italien, Spanien und England sind 375 Personen klassifiziert, mit China lediglich 11.

10 Was die Visualisierung der *Charting Culture* uns demnach zeigt, ist keineswegs die unterstellte Ausprägung kultureller Zentren und ihre Entwicklung. Was vielmehr visualisiert wird, und dies ist ein elementarer Unterschied, ist das Zustandekommen der zugrunde liegenden Daten und ihre Qualität.³⁹ Die Autoren führen uns vor Augen, wie wir (in Europa und Nordamerika) die Welt sehen und welche Daten wir zu dieser Welt erheben, nicht wie die Welt war oder ist. Die Daten in Freebase werden von Menschen ausgewählt und produziert, die ein bestimmtes Weltbild haben, ein Weltbild, in dem eine Idee von europäischen und nordamerikanischen kulturellen Zentren bereits fest verankert ist. Aus der Visualisierung solcher Daten nun wiederum die Herausbildung dieser Zentren begründen zu wollen, ist ein Zirkelschluss.

20 Um die geographische Schiefverteilung der Daten zu berücksichtigen, beschränken sich die Autoren der *Charting Culture* nach der anfänglichen globalen Perspektive auf den am besten repräsentierten Raum: "We focus on Europe and North America as we continue", heißt es im Video folglich.⁴⁰ Damit wird dann zwar nicht mehr der Eindruck erweckt, mit der statistischen Analyse dieser Daten eine Weltgeschichte zu schreiben, aber letztendlich bleibt die eigentliche Problematik der Datenzusammensetzung und -qualität nur verschleiert. Auch bleibt unberücksichtigt, dass die Daten auch innerhalb Europas schief verteilt sind. Während Deutschland und Frankreich ähnlich stark

25

"Germany". Im Datensatz zum Sohn Klaus fehlt die Historie der Staatsbürgerschaft gänzlich, er ist lediglich mit "United States of America" gelistet. Für tiefgründigere Analysen wären dies und die Qualität der Daten im Allgemeinen aber unbedingt zu berücksichtigen.

³⁹ An dieser Stelle sei notiert, dass die Analysen und Visualisierungen von Schich nicht auf primären *Quellen* beruhen, sondern auf Sekundär- bzw. aggregierten Tertiärdaten.

⁴⁰ Nach 1:10min.

repräsentiert sind, liegt Italien im Vergleich zu diesen Ländern deutlich hinten an. Zudem erscheinen die nordamerikanischen Länder gegenüber Europa überrepräsentiert.⁴¹

Datenmodelle

Den *Charting Culture* liegt eine recht einfache Definition von "Migration" zu Grunde: Migration im Kontext dieser Arbeit bedeutet eine Abweichung von Sterbeort und Geburtsort ("birth-to-death migration")⁴². Eine Person, die an Ort A geboren wurde und an Ort B gestorben ist, sei von A nach B migriert, so die vereinfachende Annahme. Schich et al. folgen mit diesem Ansatz damit den grundlegenden Migrationsstudien von Ravenstein aus dem späten 19. Jahrhundert.⁴³ Zugleich könne ein wachsendes kulturelles Zentrum damit definiert werden, dass an diesem Ort ein Überhang von Todesfällen ("death attractors") gegenüber Geburten ("birth sources") nachweisbar ist.⁴⁴ Der Definition von Migration folgend unterliegen das Datenmodell und auch die Berechnungen Schichs et al. folgerichtig dieser Vereinfachung. Oft können wir nur mit Modellen rechnen, da die Realität so komplex ist, dass sie nicht vollständig, sondern nur abstrahierend beschrieben werden kann -- eine Einschränkung, die nicht nur für die Naturwissenschaften, sondern auch für die Betrachtung kulturhistorischer Phänomene mit größeren, vor allem aber komplexen Datenbeständen gilt. Modelle unterliegen dabei immer einer zweckgebundenen Verkürzungsparadigmatik, die Stachowiak vor 40 Jahren beschrieben hat.⁴⁵

Aus der Perspektive der zu Grunde liegenden Fragestellung verkürzen die Autoren der Arbeit zu den *Charting Culture* nun das komplexe Phänomen der Migration durch Reduktion auf den Vergleich von

⁴¹ Die Vermutung liegt nahe, dass vor allem bei Datensammlungen durch die "crowd" die eigene Nation besondere Berücksichtigung findet, und dass etwa Nordamerikaner deutlich mehr zur Freebase-Datenbank beigetragen haben als Zulieferer anderer Nationen. Für die Interpretation der Visualisierung ist dies aber bedeutsam. Denn sie zeigen genau dies an, nämlich wie viele Daten zu den jeweiligen Ländern in der Datenbank gesammelt wurden, nicht aber wie viele "notable individuals" in den Ländern gelebt haben. Sie können damit zur Forschungsfrage nach der quantitativen Ausprägung kultureller Zentren nicht beitragen.

⁴² Schich et al., 558.

⁴³ Ernst Georg Ravenstein (1885): "The Laws of Migration." *Journal of the Statistical Society of London*, Vol. 48, No. 2, S. 167-235. Ravensteins erster Aufsatz basiert auf Census-Daten des Vereinigten Königreiches und beschränkt sich damit auf Bewegungen von Personen mit einem Ziel Sterbeort auf den Britischen Inseln, bietet also keine globale Perspektive.

⁴⁴ In Ravensteins Terminologie wäre ein solcher Ort als ein Ort des "absorption" charakterisiert: "a population more or less in excess of the number of its natives enumerated" (ebd.).

⁴⁵ Stachowiak, Herbert (1973): *Allgemeine Modelltheorie*. Wien: Springer. Die Datenmodellierung ist heute ein zentraler Bereich in den Digital Humanities, weil nur über sie die Verbindung von der fachwissenschaftlichen Fragestellung zu den Quellen und Daten erfolgen kann.

Sterbe- mit Geburtsort auf einen einmaligen Wanderungsvorgang von permanenter Dauer. Diese Art der Migration trifft allerdings nur einen Teil zu, viele Männer und Frauen seien “für eine Reihe von Jahren ins Ausland gegangen und dann zurückgekehrt; sie sind vielfach hin und her (zirkulär) gewandert; sie sind in Etappen gewandert”, so der Migrationsforscher Dirk Hoerder.⁴⁶ Das in den *Charting Culture* zu Grunde liegende Migrationsmodell trifft folglich ebenfalls nur auf einen Teil der Personen zu, und das Datenmodell erlaubt keine Differenzierung, wer zur ersten Gruppe gehört und wer nicht. Zudem erscheint insbesondere problematisch, dass sich die in der Arbeit untersuchte Personengruppe der Intellektuellen häufig durch komplexere Biographien charakterisieren lässt als lediglich durch die zwei Metadaten Geburt- und Sterbeort. Für die Untersuchung sozialer Wanderungsprozesse und Ortsveränderungen scheint ein differenzierteres Konzept von “Ort” in Bezug auf den Untersuchungsgegenstand und die Fragestellung erforderlich zu sein, wobei Wirkorte und Lebensmittelpunkte gerade bei den untersuchten “notable individuals” geeignetere Daten bieten als die Orte, an denen eine Person vielleicht zufällig geboren und möglicherweise ebenso zufällig gestorben ist. Albert Einstein etwa wäre nach dem Modell der *Charting Culture* von Ulm nach Princeton migriert. Nur hat er in Ulm nie gewirkt, er war ein Jahr alt, als die Eltern nach München umzogen. Für Einstein als Wissenschaftler sind Bern, Zürich, Prag und Berlin entscheidend, bevor er endgültig nach Amerika ging.⁴⁷ Nach gleichem Modell wäre Thomas Mann von Lübeck (Geburtsort) nach Zürich (Sterbeort) migriert. Die letztendliche Wanderung nach Zürich ist damit zwar korrekt erfasst, aber seine Migrationen in die USA und wieder zurück in die Schweiz bleiben ebenso unberücksichtigt wie etwa sein Wirkort München.⁴⁸

Digitale Heuristik

Die Verkürzungsparadigmatik kann damit zu einem weiteren Problem werden, nämlich dann, wenn man sich zu absoluten Aussagen verleiten lässt. So antwortet Schich in einem Interview im SPIEGEL (35/2014) auf die Frage “Sind die Intellektuellen über die Jahrhunderte mobiler geworden?” unter Berücksichtigung seiner Analysen: “Nicht wirklich. Wir müssen unser Bild von der Globalisierung

⁴⁶ Hoerder, Dirk (2010): *Geschichte der deutschen Migration. Vom Mittelalter bis heute*. München: Beck, S. 11.

⁴⁷ Laue, Max von, "Einstein, Albert" in: *Neue Deutsche Biographie* 4 (1959), S. 404-408 [Onlinefassung]; URL: <http://www.deutsche-biographie.de/ppn118529579.html> (9. April 2015).

⁴⁸ Schröter, Klaus, "Mann, Thomas" in: *Neue Deutsche Biographie* 16 (1990), S. 43-50 [Onlinefassung]; URL: <http://www.deutsche-biographie.de/ppn118577166.html> (9. April 2015). Kurioserweise fällt auch der für die Migrationsforschung und Demographie grundlegend geltende Ernst Georg Ravenstein (siehe oben) in diese Kategorie: Ravenstein (1834-1913) verbrachte zwar die größte Zeit seines Lebens in England und wirkte dort, wurde aber in Frankfurt / Main geboren und starb in Hofheim am Taunus, fast am Stadtrand des heutigen Frankfurt. Vgl. Grigg, D.B (1977): E. G. Ravenstein and the 'laws of migration'. In: *Journal of Historical Geography* 3 (1), S. 41–54.

überdenken. Im 14. Jahrhundert lag die Distanz zwischen Geburts- und Sterbeort im Mittel bei 214 Kilometern. Im 21. Jahrhundert sind es 382 Kilometer: noch nicht einmal doppelt so viele.“⁴⁹ Das ist eine interessante aber gewagt erscheinende These, die an Hand von zwei Aspekten hinterfragt werden soll.

- 5 Zum ersten überrascht die Absolutheit der Aussage mit ihren konkreten Zahlen, und es stellt sich die Frage, wie sichergestellt werden kann, dass die für die Untersuchung herangezogene Datengrundlage repräsentativ und in ihrer Qualität hoch genug ist, um ein belastbares Ergebnis zu erlauben. Für eine Überprüfung der Aussage sei im Folgenden die Neue Deutsche Biographie (NDB) herangezogen, ein für den deutschen Sprach- und Kulturraum führendes biographisches Lexikon, dessen umfangreiche Daten digital zur Verfügung stehen.⁵⁰ Ebenso wie die von den Autoren der *Charting Culture* zu Grunde gelegten Datenbasen verzeichnet die NDB “notable individuals” und ihre Biographien. Im Gegensatz etwa zur Freebase-Datenbank ist hier die Personenauswahl von vornherein um ihre Bedeutung für den deutschen Sprach- und Kulturraum eingeschränkt, so dass die Ergebnisse -- wenn überhaupt -- lediglich für diesen Raum Gültigkeit besitzen können.

Jahrhundert	Arith. Mittel	Median	Max	n
14. Jh.	182.72	95.00	1,309	96
15. Jh.	193.41	110.50	2,000	502
16. Jh.	297.70	126.00	12,272	765
17. Jh.	330.06	125.50	12,198	780
18. Jh.	285.07	121.50	15,459	2,492
19. Jh.	622.93	231.00	16,326	8,302
20. Jh.	898.28	316.00	15,993	899

15 Tabelle 2: Durchschnittliche Distanz zwischen Geburts- und Sterbeort auf Basis der NDB.

⁴⁹ Zur Herleitung der Daten vgl. Schich et al., S. 560. Für den Durchschnittswert setzen Schich et al. den Medianwert an.

⁵⁰ Selbstverständlich unterliegen auch die Einträge der NDB einem “bias”, denn es wird bei der Auswahl der aufzunehmenden Personen zwar versucht, einen repräsentativen Querschnitt durch die Gesellschaft zu erreichen, die Auswahl selbst wird aber dem jeweils vorherrschenden Zeitgeist oder einer gewissen Subjektivität unterworfen sein. Zu Organisation und Geschichte der NDB siehe Hockerts, Hans Günter (2008): Vom nationalen Denkmal zum biographischen Portal. Die Geschichte von ADB und NDB 1858–2008. In: Lothar Gall (Hg.): “...für deutsche Geschichts- und Quellenforschung”. 150 Jahre Historische Kommission bei der Bayerischen Akademie der Wissenschaften. München: Oldenbourg, S. 229–269.

Tabelle 2⁵¹ zeigt die Berechnung der Distanzen zwischen Geburts- und Sterbeort analog zur Methodik Schichs et al. auf Basis von 13.905 Datensätzen der NDB. Abzulesen ist hier ein Anstieg der durchschnittlichen Distanz zwischen Geburts- und Sterbeort vom 14./15. Jahrhundert zum 20. Jahrhundert im Durchschnitt etwa um den Faktor drei, wenn man den Median ansetzt; beim gegenüber Ausreißern störanfälligeren arithmetischen Mittel wächst die als Indiz für die Mobilität herangezogene Distanz fast um den Faktor fünf. Auch die absoluten Werte weichen ab (nach unten); ein von Schich et al. beobachtetes Minimum im 17. Jahrhundert⁵² kann nicht bestätigt werden.

Insgesamt kann die Aussage Schichs damit in ihrer Absolutheit nicht bestätigt werden, dafür sind die Abweichungen der Zahlen zu deutlich. Diese legen nahe, dass noch weitere Grundlagenforschung zur Validität solcher Methoden geleistet werden muss, die letztendlich zu einer Theoriebildung in den Digital Humanities,⁵³ an der es bislang noch mangelt, beiträgt. Auch bei Schichs Aussage zur Veränderung der Mobilität vom späten Mittelalter zur Gegenwart bleiben Zweifel, und auch hier besteht die Notwendigkeit, die Daten, die für solche Analysen zu Grunde gelegt werden, zu hinterfragen und weitere Studien durchzuführen.

Dies führt zum zweiten Aspekt. Was sagt die Distanz zwischen Geburts- und Sterbeort wirklich über die Mobilität der Menschen aus, wenn man Biographien wie die von Thomas Mann (siehe oben) betrachtet, die dann aber in die Berechnung statt mit etlichen Tausend nur mit einigen Hundert Kilometern Einzug gehalten hat? In Bezug auf die Untersuchung der menschlichen Mobilität ist dies eine heuristische Frage und zeigt eine Schwierigkeit im Umgang mit Daten auf. Die Autoren von *Charting Culture* vermerken hierzu: "Moreover, birth and death locations are usually the two most frequent locations associated with historical individuals"⁵⁴. Dies ist eine fragliche Aussage. Zwar verzeichnet ULAN in der Tat lediglich Geburts- und Sterbeort der erfassten Künstler, in den biographischen Artikeln der NDB und anderer biographischer Werke wie dem herangezogenen AKL aber wird hingegen versucht, das ganze Leben der Personen darzustellen (gleiches gilt übrigens auch

⁵¹ Angaben in Kilometern. Für die zeitliche Klassifizierung ist das Geburtsjahr angesetzt. Die Analyse basiert auf einer Ortsnamen-Identifikation und Geo-Referenzierung, die durch das Projekt "Deutsche Biographie" (<http://www.deutsche-biographie.de>) vorgenommen wird, die Differenzberechnung basiert auf dem Haversine-Algorithmus. Werte vor dem 15. Jahrhundert sind auf Grund ihrer geringen Quantität statistisch nicht belastbar. Der für die Analyse verwendete Zugriff auf die Daten über einen SparQL-Endpoint der Deutschen Biographie befindet sich derzeit (April 2015) noch in der Entwicklung. Ich danke Herrn Sebastian Gassner für die Vorbereitung der Daten.

⁵² Schich et al., S. 560.

⁵³ Siehe auch: Hall, Gary (2012): Has Critical Theory Run Out of Time for Data-Driven Scholarship? In: Matthew K. Gold (Hg.): *Debates in the digital humanities*. Minneapolis: Univ. of Minnesota Press.

⁵⁴ Ebd., Supplementary Materials, S. 7.

für die Wikipedia) und nicht nur Geburts- und Sterbeorte zu verzeichnen.⁵⁵ So liest sich ein typischer Eintrag im Österreichischen Musiklexikon (ÖML) etwa zum Organisten Georg Motz wie folgt:

5 “* 24.12.1653 Augsburg/D, † 25.9.1733 Tilsit/Ostproußen (Sovetsk/RUS). Organist, Kantor,
Theoretiker. Erhielt seine musikalische Ausbildung bei Georg Schmetzer in Augsburg und kam
mit 16 Jahren nach Worms/D. Hier beendete er den cursus scholasticus und übersiedelte
anschließend nach Wien. In der Folge in Diensten des Fürsten J. S. v. Eggenberg auf dessen
Schloss Eggenberg bei Graz (heute Graz XIV), unternahm M. 1679 eine Italienreise. 1680 verließ
er die Steiermark und übersiedelte nach Krumau (Český Krumlov/CZ), wo er Organist bei J. Chr.
v. Eggenberg wurde. Ab 1681 auf Wanderschaft, besuchte er u. a. Prag, Dresden/D,
10 Wittenberg/D, Berlin, Hamburg/D, Danzig (Gdansk/PL) und Königsberg (Kaliningrad/RUS).
1682–1720 war er Kantor an der königlichen Provinzialschule in Tilsit”.⁵⁶

In einem solchen Beitrag sind die Wirkorte des Protagonisten (Augsburg, Worms, Wien usw.) sehr wohl verzeichnet, müssten allerdings für eine computerbasierte und quantitative Auswertung zunächst durch spezielle Verfahren extrahiert werden, da sie, wie in vielen biographischen Lexika in
15 natürlicher Sprache formuliert sind und in ihrer ganzen Komplexität durch automatisierte Verfahren korrekt erfasst werden müssten.⁵⁷

Es ist zuzustimmen, dass eine solche Extraktion und die Überführung der Daten (d.h. die wesentlichen Stationen eines Lebens mit räumlicher und zeitlicher Verortung) in ein standardisiertes, strukturiertes Format noch eine Herausforderung darstellt (“building full life-paths for a large
20 number of individuals remains a persistent challenge [...] which is why we focus exclusively on birth and death locations”).⁵⁸ Nur ist die möglichst vollständige Berücksichtigung der Lebenspfade aber entscheidend für die gestellten Fragen nach Mobilität, wie die Beispiele von Albert Einstein, Georg Motz und Thomas Mann gezeigt haben. Dass die hier vorgenommene Reduktion auf Geburts- und Sterbeorte nicht von der fachlichen Fragestellung her begründet wird, ist zu kritisieren.

⁵⁵ Daher betrachten wir diese Personen als “notable”: sie sind es uns wert, ihre Biographien im allgemeinen Gedächtnis zu bewahren. Auch die Freebase-Datenbank enthält -- in höchst unterschiedlicher Tiefe und Qualität sowie uneinheitlich -- solche Daten bereit, zu Albert Einstein etwa als “employment history”; zu Thomas Mann sind jedoch keine Informationen zu finden.

⁵⁶ <http://www.musiklexikon.ac.at>.

⁵⁷ Dies geschieht gerade in einem von der DFG geförderten Projekt für die Deutsche Biographie, vgl. <http://www.deutsche-biographie.de/ueber>.

⁵⁸ Schich et al., Supplementary Materials, S. 7.

Man unterliegt leicht einer Vereinfachungsmetaphorik, "zu der das Digitale verleitet", wie der Publizist und Internetkritiker Evgeny Morozov im Allgemeinen argumentiert.⁵⁹ Der Reiz, die Verfügbarkeit von Daten und die Methodik der Quantifizierung auszunutzen, ist dabei groß. Aber um daraus historische Erkenntnis zu gewinnen, ist der Weg doch noch weit. Die Möglichkeiten, die uns digitale Methoden bieten, sind in Verbindung mit fachwissenschaftlichen Fragestellungen und zweckgerichteten Datenmodellen noch längst nicht ausgeschöpft und die fachlichen Notwendigkeiten nicht hinreichend spezifiziert. Denn auch im Digitalen brauchen die Geschichtswissenschaften eine Heuristik, auch sie benötigen fundierte, valide Methoden, auch ihre Quellen müssen der Fragestellung entsprechend ausgewählt und Daten zweckmäßig modelliert sein, und auch ihre Ergebnisse müssen falsifiziert werden können, sonst verliert sie ihre Wissenschaftlichkeit.

Potentiale des Digitalen

Das hier Gezeigte veranschaulicht, dass das Problem nicht in den Methoden an sich liegt. Das Problem ist, dass die hergeleiteten Ergebnisse bei der vorhandenen Datenauswahl und -qualität Gefahr laufen, unglaublich zu werden. Es veranschaulicht aber auch eine unmittelbare Folgeproblematik: die Ergebnisse, die durch digitale Methodik erzielt werden, können nur mit Hilfe digitaler Methodik oder wenigstens mit derer guten Kenntnis überprüft werden. Für Harald Welzer bedeutet Aufklärung heute die "Gewinnung von Unterscheidungsvermögen"⁶⁰, für Klaus Mainzer geht es um die "Stärkung unserer Urteilskraft".⁶¹ Dem ist zuzustimmen, aber es ist gleichzeitig zu fragen, wie denn ein Unterscheidungsvermögen bei digitalen Methoden gewonnen werden kann, ohne die dahinterliegenden Algorithmen zu verstehen. Das ist einer der Gründe, warum Digital Humanities nicht nur als Forschungs- und Lehrgebiet zu etablieren sind, sondern sie auch Einzug in die historischen Proseminare und Übungen, in die Grundausbildung einer jeden Historikerin und eines jeden Historikers halten sollten. Denn die Digital Humanities vereinen genau die benötigten Kompetenzen an der Schnittstelle zwischen Geisteswissenschaften und Informatik / Technologie und liefern damit die Grundlage für das kritische Beurteilen und Überprüfen neuer methodischer Ansätze und für das erforderliche Unterscheidungsvermögen.

⁵⁹ SZ vom 17.9.2014, S.11.

⁶⁰ Welzer, Harald (2014): Selbst denken. Eine Anleitung zum Widerstand. 4. Aufl. Frankfurt am Main: Fischer-Taschenbuch, S. 16.

⁶¹ Mainzer, Klaus (2014): Die Berechnung der Welt. Können Big-Data-Ergebnisse Theorie und Beweis ersetzen? In: Deutscher Hochschulverband (Hg.): Glanzlichter der Wissenschaft 2014. Ein Almanach. Stuttgart: Lucius & Lucius (Glanzlichter der Wissenschaft), S. 117–121.

Das Potential, große Mengen historischer Daten auswerten zu können, ist enorm, und die Perspektiven der Geschichtsforschung im digitalen Raum sind damit sehr vielversprechend – dies führen uns die Autoren der *Charting Culture* durch ihre methodischen Ansätze, die hier nur teilweise diskutiert werden konnten, vor Augen: was etwa, wenn wir in der Lage sind, Hunderttausende noch unerschlossener Handschriften des lateinischen Mittelalters automatisiert zu transkribieren oder wenigstens klassifizieren. Wenn wir große Datenmengen und heterogene Quellen (zum Beispiel Urkunden und Briefe) zur Erforschung der Papstgeschichte in Relation setzen und Zusammenhänge aufzeigen können, die bislang bestenfalls Zufallsfunde waren? Wenn Millionen Seiten historischer Zeitungen nicht nur zur Volltextrecherche zur Verfügung stehen, sondern etwa zum Finden und Analysieren von Co-Relationen oder zur Untersuchung von Trends im Verlauf von Zeit und Raum? Schichs et al. Arbeit wie auch die von anderen sind hierfür zugleich wegweisend, wie sie uns auch die gegenwärtig noch bestehenden Grenzen aufzeigen.

Zur Deutungshoheit über die Geschichte im digitalen Raum

Der Ansatz der *Charting Cultures* fällt in die Kategorie eines data-driven-research,⁶² datenmotivierter Forschung also, die auch in den Geschichtswissenschaften in der Zukunft eine wichtige Rolle spielen wird. Und es wird deutlich: mit digitalen Daten allein ist es nicht getan, wir benötigen digitale Methoden, und wir brauchen eine digitale Quellenkritik bzw. eine Quellenkritik des Digitalen.⁶³ Hierzu grundlegend fehlen derzeit mindestens zwei Dinge: nämlich eine Theorie quantitativer wie qualitativer digitaler Methoden für die Geschichtswissenschaften, wie sie die Digital Humanities hervorbringen sollte und zweitens die Verankerung der digitalen Methoden in den Fachwissenschaften. Denn es müssen die Historikerinnen und Historiker sein, die die Fragestellungen entwickeln, sie müssen die Quellen und Daten auswählen und die Daten modellieren. Natürlich muss dabei verkürzt werden, die Daten müssen auf das für den Zweck Erforderliche reduziert werden, wobei allerdings gleichzeitig das für den Zweck also für die Forschungsfrage Notwendige erhalten bleiben muss. Schließlich sind es auch nur die Historikerinnen und Historiker, die die Daten kritisch interpretieren können.

Was nämlich passieren kann, wenn dies nicht geschieht, zeigt ein anderer, sehr bekannter, beachteter und kontrovers diskutierter Ansatz, der der *Culturomics*, basierend auf den Volltexten der

⁶² Für die Naturwissenschaften prägte hier Jim Gray mit den Begriff des "fourth paradigm", das den älteren (natur-)wissenschaftlichen Forschungsparadigmen des Experiments, der Theorie und der Simulation folge. Vgl. Hey, Tony; Tansley, Stewart; Tolle, Kristin: Jim Gray on eScience: a transformed scientific method. In: The fourth paradigm. Data-intensive scientific discovery, S. xvii–xxxii..

⁶³ Vgl. Haber, Digital Past, S. 104–112.

Digitalisierung im Rahmen des Google-Books-Projektes.⁶⁴ Die *Culturomics* mit ihrem nGram-Viewer zeigen uns eben nicht nur exemplarisch auf, wie solche Verfahren etwa zum Studium kultureller Entwicklung genutzt werden können, also was machbar ist. Sie unterstellen gleichzeitig, dass das, was sie aufzeigen, Geschichtsforschung oder die Kulturwissenschaft *sei*, und leider wird das in der öffentlichen Wahrnehmung häufig so gesehen.⁶⁵ Eine Heuristik im geisteswissenschaftlichen Sinne, also die Auswahl von geeigneten Quellen und Methoden in Bezug auf eine Forschungsfrage, scheint nicht mehr notwendig zu sein, ebenso wenig die kritische Hinterfragung der Zusammensetzung, des Zustandekommens und der Qualität der Daten und Metadaten (in diesem Falle aus dem Google-Books-Projekt), die Verfahren selbst (hier vor allem quantitativ-statistisch) sind relativ agnostisch gegenüber dem Untersuchungsgegenstand, die Ergebnisse dieser Analyse sprechen scheinbar für sich, und das kritische Deutungsmuster, das "Warum?" tritt hinter der Statistik und ästhetisch ansprechenden visuellen Aufbereitungen in den Hintergrund. Unterschwellig kann hier der Eindruck entstehen, dass man eigentlich die Kulturwissenschaftler gar nicht mehr brauche. Und das ist sehr gefährlich.⁶⁶ In einem denkbaren Szenario wäre damit nicht nur die Deutungshoheit über die Digital Humanities, die derzeit ausgiebig zwischen Informatikern und Spezialisten der Digital Humanities ausdiskutiert wird,⁶⁷ in Frage gestellt, sondern die Deutungshoheit über die Geisteswissenschaften überhaupt. Geschichte würde nicht mehr von Historikern interpretiert, sondern von Datenanalysten.

⁶⁴ Michel, Jean-Baptiste; Shen, Yuan Kui; Aiden, Aviva Presser; Gray, Matthew K.; The Google Books Team; Pickett, Joseph P. et al.: Quantitative Analysis of Culture Using Millions of Digitized Books. In: Science and Engineering Ethics (14. Januar 2011), S. 176–182.

⁶⁵ Obwohl die Protagonisten dieses Projektes, Jean-Baptiste Michel und Erez Lieberman Aiden keine Geisteswissenschaftler sind (weder im traditionellen Sinne noch in der Ausprägung der Digital Humanities) sondern angewandte Mathematiker, wurde ihr Projekt lange synonym mit den Digital Humanities gesetzt. Denn die Ergebnisse in ihrem Projekt behandeln Fragen der Erinnerungskultur, der Medienzensur, der Linguistik und vieler weiterer, typisch geisteswissenschaftlicher Gebiete. "Culturomics", so argumentieren sie, "provides quantitative *evidence* [Hervorhebung MR] for scholars in many fields" (Michel et al., S. 181).

⁶⁶ Das Projekt wurde kontrovers diskutiert. Kritisch äußern sich u.a. Greetham, David (2012): The Resistance to Digital Humanities. In: Matthew K. Gold (Hg.): Debates in the digital humanities. Minneapolis: Univ. of Minnesota Press, Hitchcock, Tim (2011): Culturomics, Big Data, Code Breakers and the Casaubon Delusion, Historyonics, <<http://historyonics.blogspot.de/2011/06/culturomics-big-data-code-breakers-and.html>>; 9. April 2015> und Zimmer, Ben (2013): 'When Physicists Do Linguistics - Ideas - The Boston Globe', The Boston Globe, 10.2.2013 <<http://www.bostonglobe.com/ideas/2013/02/10/when-physicists-linguistics/ZoHNxhE6uunmM7976nWsRP/story.html>>; 9. April 2015>; optimistischer z.B. Cohen, Daniel (2010): 'Initial Thoughts on the Google Books Ngram Viewer and Datasets' <<http://www.dancohen.org/2010/12/19/initial-thoughts-on-the-google-books-ngram-viewer-and-datasets>>; 9. April 2015>.

⁶⁷ Vgl. das Motto der ersten Jahrestagung des Fachverbandes Digital Humanities im deutschsprachigen Raum DHd im März 2014 in Passau: "Digital Humanities - methodischer Brückenschlag oder ‚feindliche Übernahme‘? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik" (<http://www.dhd2014.uni-passau.de/>).

Dies könnte ein mögliches, bedenkliches Szenario für die Geschichtswissenschaften im 21. Jahrhundert sein. Nicht die statistischen Methoden werden an die Bedürfnisse der Historiker angepasst, wie Boonstra, Breure und Doorn es 2004 formulierten, sondern die Geschichte an die Daten.

5 Digital Humanities als historische Grundwissenschaft

In diesem Beitrag soll es aber nicht darum gehen, digitale Methoden abzulehnen. Ganz im Gegenteil, die Bereicherung der historischen Methode durch digitale Verfahren bietet enormes Potential. Aber wir müssen sie kritisch befragen und ihrer Herr werden. Dazu ist es wichtig, die Digital Humanities in den Geisteswissenschaften zu verorten, wozu folgende Analogie bemüht werden soll.⁶⁸ Die Paläographie lehrt uns (unter anderem) alte Texte zu lesen und zu datieren. Nur wächst aus dem Lesenkönnen der Texte allein noch keine historische Erkenntnis, deren Gewinn ja das Ziel der Geschichtswissenschaft ist, sondern erst aus der Kontextualisierung und Interpretation dieser Texte durch die Historiker. Dennoch verbessert die Paläographie stetig ihre Methodik, hat eigene Fragestellungen und Forschungsgegenstände, was sie als eigenständige, leider zu oft vernachlässigte Disziplin auszeichnet. Für die Geschichte und andere Disziplinen aber ist sie primär eine Hilfs- und Grundwissenschaft. Sie ist sogar eine *elementare* Hilfe, denn ohne das Lesenkönnen alter Texte könnten wir alte Texte nicht lesen, so einfach das klingt, und sie somit nicht als historische Quellen nutzen, aus denen wir Erkenntnis über unsere Vergangenheit gewinnen. Wir wären quasi blind.

Analog verhält es sich mit der *Digital History*. Die Entwicklung von Datenmodellen, Kodierungsstandards, computerbasierten Analyseverfahren, Visualisierungstechniken und Big-Data-Verfahren für historische Daten, inzwischen längst standardisierte Digitalisierungs- und Erschließungstechniken wie auch experimentelle, fortgeschrittene Digitalisierungs- und Analyseverfahren wie etwa die Multispektralfotographie -- sie alle liefern losgelöst und isoliert keine neue historische Erkenntnis. Aber sie können grundlegend sein für ihr Gewinnen: die Digitalisierung schafft Zugang zu Quellen, der ansonsten nicht oder nur aufwendig herzustellen wäre; das *Information Retrieval* filtert aus der Flut von Daten und Information die benötigten heraus; die *Visual*

⁶⁸ Neben dem hier bemühten Vergleich und der Einordnung der Digital Humanities als historische Grundwissenschaften, wird auch überlegt, wie digitale Methoden für die historischen Grundwissenschaften nutzbar gemacht werden können. Vgl. hierzu Patrick Sahle, "Historische Hilfswissenschaften" in Bärbel Biste, Rüdiger Hohls (Hg.), Fachinformation und EDV-Arbeitstechniken für Historiker: Einführung und Arbeitsbuch (=Historical Social Research Supplement, 12), 2000, S. 103-112 sowie aktuell: Hassner, Tal, Malte Rehbein, Peter A. Stokes, and Lior Wolf. "Computation and Palaeography: Potentials and Limits". Dagstuhl Manifesto, Vol. 2, Issue 1, 14-35. Weiteren Überblick verschaffen: Rehbein, Malte, Patrick Sahle, Torsten Schaßan (Hg.) (2009): Kodikologie und Paläographie im digitalen Zeitalter / Codicology and Palaeography in the Digital Age. Norderstedt: BoD und Franz Fischer, Christiane Fritze, Georg Vogeler (2010): Kodikologie und Paläographie im digitalen Zeitalter 2 / Codicology and Palaeography in the Digital Age 2. Norderstedt: BoD.

Analytics helfen uns bei der Erkennung von Mustern und Strukturen, die ansonsten im Verborgenen blieben. Analog zur Paläographie hat dann die Auswahl, Kontextualisierung und Interpretation dieser digitalen Daten und Informationen durch die Historiker zu erfolgen. Als Grundwissenschaft aber erwirkt die Digital Humanities das so elementar wichtige *Nutzenkönnen* digitaler Methoden und Daten.

Ein weiterer Gedanke betrachtet die gegenwärtigen digitalen Technologien in den Geschichtswissenschaften als Hype, und betonen wir dabei einen Paradigmenwechsel innerhalb der Wissenschaft, können wir von einem „Turn“ sprechen, dem „computational turn“. Solche Turns, aus der Vergangenheit kennen wir etwa den linguistic turn, den spatial turn oder den iconic turn, scheinen immer einem Muster zu folgen. Es gibt einen gewissen Hype, gepaart mit einer gleichzeitigen harschen Kritik, die die Befürworter eines neuen Paradigmas als Anhänger einer obskuren Sekte erscheinen lässt, dann aber bleibt doch so einiges von diesem Turn übrig, wird als selbstverständlich erachtet und erhält kanonisiert Einzug in das Methodenrepertoire.

Der computational turn wird aber, so meine Vermutung, durchdringender sein und die Geschichtsforschung nachhaltiger beeinflussen und verändern als andere Turns der Vergangenheit. Zwar wird der gegenwärtige Hype abflachen, aber er wird konsolidiert vieles Gewinnbringende übrig lassen. Hiervon wird in der Zukunft allerdings das Wenigste als *Digital Humanities* aufgefasst werden, sondern es wird in den selbstverständlichen Methodenkanon der Historiker Einzug erhalten haben, so dass das Attribut „digital“ wegfallen kann.⁶⁹

Eigentlich ist die Unterscheidung schon heute hinfällig. Man stelle sich etwa einen Historiker vor, der zum ersten Weltkrieg forscht und als Hauptquelle regionale Tageszeitungen auswertet. Liegen diese Zeitungen digital vor, und liest er sie am Bildschirm, in einer digitalen oder virtuellen Arbeitsumgebung und wendet er etwa Volltextsuche an, um gezielt in den Quellen zu suchen, würden wir dies wohl als herkömmlichen Arbeitsprozess des Historikers bezeichnen und kaum weiter zur Kenntnis nehmen. Was aber, wenn er Verfahren des Distant Reading, Topic Modelling oder quantitative Analysen zur Untersuchung von Trends in der Berichterstattung nutzt? Er würde zum Beispiel die Zeitungskorpora in großer Menge analysieren: was und über welche Themen wurde berichtet, wie verändert sich das im Zeitverlauf (etwa von der Kriegsführung hin zur Heimatfront), und wie wird berichtet (die Verfahren der Sentimentanalyse könnten hier greifen). Das sind genau

⁶⁹ Wir beobachten das schon heute. So setzte etwa ein im Januar 2014 veröffentlichter Aufruf des Schweizer Nationalfonds für Projektvorschläge im Bereich der Edition fast selbstverständlich die digitale Publikation dieser Editionen voraus: „Es werden ein digitales Bearbeitungs- und Publikationskonzept sowie Überlegungen zur Langzeitarchivierung erwartet“ (http://www.snf.ch/SiteCollectionDocuments/call_editionen_d.pdf; 6. April 2015).

die Fragestellungen des Historikers. Dennoch würde man ihn nach der gegenwärtigen Lesart wohl als einen "Digital Humanist" einordnen.

Woran liegt es also, dass in einem Falle (Volltextsuche) niemand von Digital Humanities redet, im anderen (Sentimentanalyse) aber schon, und wo ist die Grenze etwa zwischen Volltextsuche und Sentimentanalyse? Es liegt vermutlich daran, dass die zuletzt genannten Verfahren noch nicht ohne spezielle Kenntnisse oder Kooperationen mit Informatikern nutzbar sind. Ich vermute, dies wird sich ändern, und die genannten Verfahren werden in den Methodenkanon der historischen Forschung Einzug halten, sobald sie in der Breite benutzbar werden -- auch dies eine Aufgabe einer grundwissenschaftlichen Disziplin *Digital Humanities*. Die "Digital Humanists" sind dann diejenigen, die die Methoden und Verfahren weiterentwickeln.

Digitale Quellen und digitale Methoden

Noch aber mangelt es den digitalen Geschichtswissenschaften an Theorie und Verankerung mit den Fachwissenschaften. Vieles wird ausprobiert, und weil die technischen Möglichkeiten bestehen, beherrscht uns ein Pragmatismus: die Daten sind da, also nutzen wir sie mittels maschineller Unterstützung und schauen, was herauskommt. Explorative Analyse⁷⁰ würden wir das nennen, *nGram-Culturomics* ist ein Beispiel dafür. Nun wissen wir aber, dass nicht alles was möglich ist, auch gut ist. Das ist durchaus eine ethische, primär aber eine fachwissenschaftliche Frage. Was ist also fachwissenschaftlich gut, und was ist richtig? Insbesondere im Bereich der modernen, datenzentrierten Wissenschaft, die zunehmend auch die Geschichtsforschung erreicht, sind diese Fragen berechtigt. Ramón Reichert stellt hierzu fest, dass „die der Big-Data-Research zugrundeliegenden fächerübergreifenden Praxisorientierungen eine nuancierte Genealogie, Datenkritik und Medienreflexion der daten-intensiven Formen der Wissensproduktion erfordern“.⁷¹ Auch wenn das Interesse an historischer Erkenntnis das Gleiche ist, ist der Umgang mit digitalen Daten, seien sie digitalisiert oder wie hier digital born, doch ein anderer. Und der Vergleich digitaler *Methoden* mit traditionellen Methoden wird nicht funktionieren, traditionelle Methoden lassen sich nur begrenzt auf diese Daten anwenden.

Charting Culture ist ein schönes Beispiel für eine „datengesteuerte Wissenschaft“, von der Reichert spricht. Die Wissensproduktion, so argumentiert er, sei von der Verfügbarkeit der technologischen Infrastruktur und von digitalen Methoden abhängig geworden. Bei den Wissenschaften im digitalen

⁷⁰ Im Gegensatz zur bestätigenden Analyse setzt die explorative Analyse keine Forschungshypothese voraus.

⁷¹ Interview von Georgios Chatzoudis vom 3.2.2015 (http://www.lisa.gerda-henkel-stiftung.de/big_data).

Zeitalter unterscheidet Reichert „digitalisierte Methoden“ von „digitalen Methoden“.⁷² Ersteres sei die Übertragung traditioneller Methoden (er nennt als Beispiel Fragebögen) in das digitale Medium (online-Fragebögen); letzteres seien Methoden auf großen Mengen digitaler Daten (wie Schich et al. es zeigen), um diese mit computergestützten Verfahren zu filtern, analysieren, aufzubereiten und darzustellen. Und wie es die Aufgabe der Digital Humanities bislang gewesen ist, Verfahren zur Digitalisierung von Quellen zu entwickeln, ist es die *künftige* Aufgabe dieser Digital Humanities, nun gemeinsam mit den Fachwissenschaften digitale Methoden zu entwickeln und diese theoretisch zu unterfüttern.

Reicherts Unterscheidung ist damit analog zur Ausdifferenzierung der Daten bzw. Quellen, der Unterscheidung von digitalisierten Quellen und *digital born* Quellen zu verstehen. Im Bereich der Geschichtsforschung müssen wir eigentlich sogar eine Dreiteilung zu Grunde legen: unsere Quellen in Schriftform und anderer, nicht digitaler Gestalt erheben und erschließen wir entweder 1. analog/traditionell oder 2. direkt digital, oder wir haben es 3. mit *digital born* Daten zu tun, die ohne Transformation direkt im digitalen Medium bearbeitet werden. Im Bereich der Papstforschung fiele in die erste Kategorie etwa die Transkription oder Edition einer Papsturkunde in Buchform, in die zweite Kategorie die digitale Edition dieser Urkunde, in die dritte Kategorie aber die Verarbeitung von online-Kommunikation des aktuellen Papstes. Damit ergeben sich nämlich weitere unmittelbare Konsequenzen für die Geschichtswissenschaft: spätestens, wenn digitale Quellen ihre Grundlage werden, greifen alte Methoden nicht mehr vollständig, denn die Twitter-Microblogs des Papstes Franziskus (@pontifex) können wir nicht aus ihrem digitalen Kontext lösen.

Es gibt also viel zu tun für die Geschichtswissenschaften im digitalen Raum. Um das digitale Potential aber zu wecken und für ernsthafte, gesellschaftlich als relevant angesehene und förderungswürdige Geschichtsforschung nutzbar zu machen, bedarf es im Kern Dreierlei: (1) einer *Integration* der Digital Humanities als Geisteswissenschaft innerhalb der Forschungslandschaft, die sich (2) einem “Ende der Theorie”⁷³ und einem drohenden Verlust der Deutungshoheit über die Geschichte entgegenstellt,

⁷² Zu “digital methods” siehe auch Rieder, Bernhard; Röhle, Theo (2012). In: David M. Berry (Hg.): Understanding digital humanities. Basingstoke: Palgrave Macmillan, S. 67–84.

⁷³ Wie von Chris Anderson postuliert: “massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves” (Anderson, Chris, ‘The End of Theory: The Data Deluge Makes the Scientific Method Obsolete’, Wired Magazine, 16.07 (2008) <http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory> (19 Juni 2014). Hingegen plädiert Mainzer für “die Besinnung auf die Grundlagen, Theorien, Gesetze und Geschichte, die zu der Welt führen, in der wir heute leben [...] Am Ende geht es um eine Stärkung der Urteilskraft, d.h. die Fähigkeit, Zusammenhänge zu erkennen, das 'Besondere', wie es bei Kant heißt, mit dem

wofür (3) die Verankerung digitaler Methoden und einer digitalen Quellenkritik in den Fachwissenschaften erfolgen muss.

5 Willard McCarty, einer der führenden Denker und Theoretiker der Digital Humanities und Herausgeber des 1987 von ihm ins Leben gerufenen HUMANIST⁷⁴ fordert “a digital practice that is *of* as well as *in* the humanities” als nächsten Schritt der Entwicklungen an der Schnittstelle zwischen Geisteswissenschaften und Computertechnologie: “This next step is the learned practitioner’s open-eyed, technologically informed, imaginative, critical, hands-on *questioning* of what happens at the cross-roads of actual work, where computing, scholar-practitioners and the humanities meet”.⁷⁵ Und dies muss Aufgabe der Digital Humanities als historische Grundwissenschaft sein.

'Allgemeinen' zu verbinden, in diesem Fall die Datenflut mit Reflexion, Theorie und Gesetzen, damit eine immer komplexer werdende und von Automatisierung beherrschte Welt uns nicht aus dem Ruder läuft.” (Mainzer, 121).

⁷⁴ <http://dhhumanist.org>

⁷⁵ McCarty, Willard. “Getting there from here. Remembering the future of digital humanities”. *Literary and Linguistic Computing* 29, Nr. 3 (2014): S. 283–306, hier: S. 295.