
Einleitung

Wir wollen unser Denken und Handeln auch in der virtuellen Arbeits- und Lebensweise verfestigen und schreiben es auf. So vergewissern wir uns, und teilen das Aufgeschriebene ändern mit. Der Brief, das Journal und der Bericht von gestern sind die digitalen sozialen Medien von heute. *I post, so I am*. Die digitale Kommunikation prägt mittlerweile auch das Berufsbild der Historikerinnen und Historiker. Nun gilt als Norm, was vor 20 Jahren noch die Ausnahme war: Das Arbeiten mit digitalen Dokumenten (und anderen Ressourcen). Zukünftig wird es darum gehen, historische Welten digital zu vermessen und analog zu interpretieren.

Drei Beispiele aus den USA zeigen, wie diese digitale, virtuelle, historische Welt zukünftig aussehen könnte: *Valley of the Shadow* macht Quellen zum amerikanischen Bürgerkrieg zugänglich – Zeitungsartikel, Bilder, offizielle Dokumente, Schlachtpläne, Briefe und Tagebücher von beteiligten Soldaten. Diese Quellen orientieren sich um einen Referenzraum, in dem eine *Timeline* einen übersichtlichen Zugang organisiert.¹ Das zweite Beispiel verweist auf eine moderne, digitale Inhaltsanalyse: *Echo Chamber or Public Sphere* identifiziert politische *Tweeds* und ordnet sie automatisch Demokraten beziehungsweise Republikanern zu – das Ergebnis einer Kombination von *Machine learning* und einer sozialen Netzwerk-Analyse.² Und das dritte Beispiel verweist auf neue Produktionsweisen von historischem Wissen: Shawn Graham, Ian Milligan und Scott Weingart haben gemeinsam mit der Hilfe von Leserinnen und Lesern im Internet ein Buch verfasst: *Exploring Big Historical Data: The Historian's Macroscope*.³ Die Autoren sind überzeugt, dass das Schreiben von Geschichte zukünftig offen und in Teams erfolgt.

Der digitale Wandel erfasst die Produktion und Vermittlung von historischem Wissen und wirkt auf die Voraussetzungen für das Schreiben von Geschichte zurück. Es ist eine »stille Revolution«⁴,

welche den Alltag von »Geistesarbeiterinnen« und »Geistesarbeitern« vollumfänglich verändert. Verantwortlich dafür sind nicht allein die leise summenden Rechenmaschinen, sondern vor allem die auf ihnen geräuschlos laufenden Programme, Anwendungen und Algorithmen. Sie ermöglichen die Verarbeitung von nie dagewesenen Mengen an Daten und schaffen mit uns Menschen eine neue, digitale Informationsgesellschaft. Bruno Latour, der Grenzgänger zwischen Philosophie und Soziologie, beschreibt sie als Netzwerk, das heißt als Interaktion beziehungsweise Konstellation von Menschen und Dingen.⁵

Die Folgen dieses digitalen Wandels sind enorm: So wie im 19. und 20. Jahrhundert dampfgetriebene und elektrische Maschinen die Arbeitsabläufe von Industrie-Arbeitern umwälzten, verändern im 21. Jahrhundert Computer und Algorithmen den beruflichen Alltag von Beamten, Archivaren, Bibliothekaren, Journalisten, Autoren, Herausgebern – einem überwiegenden Teil der mit Informationen arbeitenden Mittelschicht. Viele vergleichen diesen Wandel mit der Revolution zu Zeiten Gutenbergs, der Erfindung und Verbreitung des Buchdrucks mit beweglichen Lettern und der Gestaltung von neuen Arbeits- und Geschäftsmodellen in der Renaissance.

Dieses Buch beschreibt den gegenwärtigen Stand des digitalen Wandels für den Bereich der Geschichte als Teil der Geisteswissenschaften und diskutiert die Perspektiven über seinen weiteren Verlauf. Es enthält überdies einen Serviceteil, in dem Initiativen zum Aufbau von Infrastrukturen, Plattformen und Portale, Anwendungen, *Tools* und Standards sowie Zeitschriften und *Blogs* vorgestellt werden. Dabei bildet die Themen- und Inhaltsanalyse mit ihrem Potenzial für eine neue *Digital History* einen Schwerpunkt.

Diese Einleitung steckt zunächst den Untersuchungsgegenstand ab, das neue Praxisfeld *Digital Humanities*, in dem die Geschichte ihren Platz sucht. Sie umschreibt sodann die relevanten Begriffe und das in diesem Buch verwendete Verständnis historiographischer Produktion. Die Einleitung skizziert schließlich auch den Rahmen für die Beschreibung und die Analyse des Stands der Dinge und der Perspektiven, der beiden inhaltlichen Hauptbestandteile dieses Buches.

»Digitaler Wandel« bezeichnet eine wachsende gesellschaftliche, das heißt technische, wirtschaftliche und kulturelle Veränderung seit der zweiten Hälfte des 20. Jahrhunderts durch den Einsatz von Rechenmaschinen (Computer), durch die Datenverarbeitung und Automatisierung sowie durch die fortschreitende Digitalisierung und Vir-

tualisierung. Man nimmt an, dass um die Wende zum Jahr 2000 weltweit erstmals mehr Informationen digital als analog gespeichert wurden: Waren um das Jahr 1995 erst rund 5 % der Informationen digital verfügbar, so waren es um das Jahr 2010 schon rund 95 %. Als ein entscheidendes Merkmal des digitalen Wandels gilt denn auch das exponentielle Wachstum an verfügbaren Informationen. Es ist deshalb mittlerweile oft auch vom »Informationszeitalter« die Rede, einer Gesellschaft, in der Informationen, ihr Austausch und ihre Vernetzung einen grundlegenden wirtschaftlichen Produktionsfaktor bilden.

Das Internet, der Auf- und Ausbau von Informations- und Kommunikationsnetzen, ist gleichermaßen ein Produkt und Treiber dieses Wandels. Es steht im Grunde aber nur für seine überraschend schnelle Erfassung fast aller Lebensbereiche und für seine rasche weltweite Ausdehnung. Das eigentliche Fundament des digitalen Wandels sind die Verbreitung der digitalen Technik – integrierte Schaltkreise und Mikrochips in den Rechenmaschinen –, die schnelle Steigerung ihrer Leistung und die dadurch mögliche, immer weiter gehende Automation und Informatisierung von Produktion und Konsum.

Dieser Wandel lässt sich am Einsatz von Computern veranschaulichen: komplizierte und überaus teure Rechenmaschinen wurden zwischen den 1950er und den 1970er Jahren fast nur für die Verarbeitung von damals umfangreichen Mengen von Daten eingesetzt – in der Wissenschaft zum Beispiel für die Statistik, in der Verwaltung zum Beispiel für die Erfassung von Personendaten (Register)⁶, in Unternehmen zum Beispiel für die Buchhaltung, die Sitzplatzreservierung bei Fluggesellschaften oder die Berechnung von Kapazitäten bei Eisenbahngesellschaften und, schließlich, beim Militär zum Beispiel für die Luftüberwachung.⁷ Ab den 1980er Jahren finden sich sogenannte personalisierte Computer dann auch in immer mehr Büros kleinerer und mittlerer Firmen sowie in Haushalten.

Ausschlaggebend für den enormen *Impact* des digitalen Wandels ist der Umstand, dass digitale Objekte im Unterschied zu analogen Objekten beliebig oft kopiert und benutzt werden können. Sie lassen sich so kostengünstig produzieren und direkt verteilen. Die digitale Vervielfachung wirkt sich in allen Bereichen aus, in denen Informationen einen wesentlichen Anteil an der Produktion und Realisierung von wirtschaftlichem Mehrwert haben – auch in Wissenschaft und Kultur.

Einen wesentlichen Teil an diesem Wandel tragen die Konsumentinnen und Konsumenten. Sie sind bereit, in der Wertschöpfungskette

Arbeiten zu übernehmen, die zuvor die Produzenten erledigt haben. Entwickler verkaufen mittlerweile technische Grundlagen und die Rechte für die Beteiligung an Anwendungen wie Online-Spielen, welche die *Gamer* dann unter sich selbst (weiter) entwickeln. *Wikipedia* funktioniert nach einem ähnlichen Modell. Und dieses findet mittlerweile auch in der Historiographie Anwendung, wie wir etwa am Beispiel des Buchprojekts *Exploring Big Historical Data: The Historian's Macroscope* gesehen haben.⁸

Der Wandel ist noch lange nicht abgeschlossen. Die Entwicklung von selbstreflexiven Maschinen wird heute nicht mehr ausgeschlossen – Stichwort »künstliche Intelligenz«. Ein Beispiel dafür ist *Deep Dream*, ein Projekt von *Google*, das zeigt, wie eine Maschine lernt, Bilder zu berechnen und wieder zu erkennen. Es wird im Kapitel »Die Philosophie der *Digital History*« im Perspektiventeil dieses Buches näher beschrieben.

Auch in der Geschichtswissenschaft nimmt das Vertrauen in Computer zu. Wir befassen uns mit Analysen, die von Programmen erarbeiteten wurden, und nehmen Computer vermehrt als Erweiterung unserer eigenen kognitiven Fähigkeiten wahr. Dem Menschen vorbehalten bleibt die tiefe Interpretation von Texten und der Umgang mit Widersprüchlichkeit in Quellen (in welcher Form auch immer).

Ein wesentliches gesellschaftliches Merkmal des digitalen Wandels ist das Verdrängen von traditionellen Formen der Arbeit durch den Einsatz von billigen Computern und Programmen. Es stellen sich dabei viele grundlegende Fragen nach der Würde von Menschen, die sich über Arbeit definieren, und der Privatsphäre von Menschen. Die weltweite Überwachung von Aktivitäten im Internet verlangt nach transnationalen Lösungen für die Wahrung von Persönlichkeitsrechten und den Datenschutz, was aber nicht das Thema dieses Buches ist. Vieles von dem, was es kritisch zu debattieren gilt, findet sich in dem Buch *You are not a Gadget* zum Umgang mit Daten.⁹ Der Autor, der Internet-Pionier Jaron Lanier, wurde unter anderem dafür mit dem Friedenspreis des Deutschen Buchhandels ausgezeichnet.

Der digitale Wandel als eine umfassende gesellschaftliche Veränderung seit der zweiten Hälfte des 20. Jahrhunderts betrifft, wie schon mehrfach erwähnt, auch die Wissenschaft. Wenn wir den Einsatz von Computern, den Gebrauch von automatisierten Methoden und die Bearbeitung digitaler Objekte wie Daten, Dokumenten, Bildern oder

Tönen in den Geisteswissenschaften kombinieren, erhalten wir eine Umschreibung für den Begriff der *Digital Humanities*. Er bezeichnet also ein wissenschaftliches Praxisfeld, das sich gegenwärtig aus der Verbindung von traditionellen geisteswissenschaftlichen Methoden einerseits und digitalen, informationstechnischen Verfahren und Standards andererseits entwickelt.

Vorläufer für diese neue wissenschaftliche Praxis sind die Fachinformatik, die Informationswissenschaften und die Computerlinguistik. Soweit sie ähnliche digitale Methoden und Verfahren einsetzen, scheinen die verschiedenen geisteswissenschaftlichen Disziplinen sich einander anzunähern. Ein Konzept für diese Konvergenz zeichnet sich zwar in Umrissen ab, ein theoretisches Gerüst für die *Digital Humanities* fehlt aber noch weitgehend. Die beiden nachfolgenden Teile zum Stand der Dinge und zu den Perspektiven versuchen, diese Linien für das Auftreten einer digitalen, virtuellen, hybriden Geschichtsschreibung nachzuzeichnen.

Der Gebrauch digitaler Medien und computergestützter analytischer Verfahren für die Produktion und Vermittlung historischer Forschungsergebnisse umreißt ein Praxisfeld, das als *Digital History* bezeichnet werden soll. Es handelt sich zum einen um eine wissenschaftliche Praxis, die sich an Methoden der *Digital Humanities* anlehnt. Es handelt sich zum andern aber um ein Feld, das sich aufgrund von eigenen Erfahrungsbereichen konstituiert. Zu den Vorläufern der digitalen Geschichtswissenschaft zählen die historische Fachinformatik und die quantitative Geschichte.

Gegenwärtig prägt vor allem eine Anwendung die digital aufbereitete und vermittelte Geschichte: Die Ausweitung der Präsentation von Dokumentationen und Rechercheergebnissen zu einer »öffentlichen Geschichte« im *Web*. Sichtbar wird die *Public Digital History* insbesondere durch thematische Präsentationen, *Timelines*, interaktive Karten und virtuelle historische Räume im Internet. Insbesondere die Möglichkeiten der *Geographischen Informationssysteme* (GIS) werden recht häufig genutzt. Hingegen werden computergestützte quantitative Analysen in der Geschichtswissenschaft im deutschsprachigen Raum noch kaum eingesetzt. Das ist ein wesentlicher Unterschied zu den Sozialwissenschaften, wo diese Methoden und Verfahren mittlerweile sehr oft zum Einsatz kommen.

Fallbeispiele für die Analyse umfangreicher historischer Datenmengen und die dafür notwendige »unternehmerische« oder »kollaborati-

ve« Arbeitsweise werden im »Stand der Dinge« und die Umrisse weiterer Gestaltungsmöglichkeiten in den »Perspektiven« vorgestellt.

Das Schreiben von Geschichte lässt sich als Produktion von historischem Wissen umschreiben. Es handelt sich um einen Prozess, der mit dem Sammeln und Organisieren von Informationen beginnt. Diese Daten und Fakten müssen aufbereitet und analysiert werden. Aufbereitete, analysierte und synthetisierte Informationen, das Wissen also, wird sodann an ein Publikum vermittelt sowie von diesem angewendet und geteilt. Dieses Publikum kann ein Fachpublikum – die wissenschaftliche *Community* – oder ein breiteres Publikum sein – die Öffentlichkeit. Oft richtet sich solches Wissen auch an einen ganz bestimmten Teil einer Bevölkerung, an Gruppen, die von bestimmten historischen Ereignissen besonders betroffen sind – an die Opfer von fürsorglichen Zwangsmaßnahmen in der Schweiz, beispielsweise.¹⁰

Karl Mannheim hat sich als Pionier intensiv mit dieser »Soziologie des Wissens« beschäftigt¹¹ und Peter Burke, der Historiker des Wissens, hat sich jüngst mit der Beschleunigung dieses Prozesses auseinandergesetzt.¹² Dieses Buch beleuchtet die Stationen des historiographischen Produktionsprozesses unter der Voraussetzung des digitalen Wandels. Die Leserin und den Leser erwartet folgendes:

»Der Stand der Dinge« beschäftigt sich zunächst mit der »stillen Revolution«, den Möglichkeiten von Algorithmen, deren Bedeutung wir noch nicht vollumfänglich verstehen. Algorithmen, Daten und Netze tragen zur Verflüssigung der Information über die Welt bei, welche die Produktion und Vermittlung von Wissen auf eine ganz neue Basis stellt. Als Zweites werden grundlegende Bedingungen zur Produktikon von Geschichte besprochen: Die Anstrengungen von Archiven und Bibliotheken, sich den technischen Veränderungen und den neuen Erwartungen der Kunden anzupassen. Gefragt wird dabei auch nach der Überlieferung digitaler Unterlagen – »Wer archiviert die SMS von Angela Merkel?«¹³. Es geht sodann, drittens, um den Nutzen großer Zahlen, das heißt um *Big Data* sowie um die historische Quantifizierung und Analyse. Es folgt, viertens, eine Erörterung über einige neue Praktiken in den Geisteswissenschaften, die durch den digitalen Wandel induziert wurden. Der Aufbau der dafür notwendigen Infrastruktur und ihre neuen Anbieter sind das fünfte Thema. Es folgt, sechstens, die Präsentation verschiedener Fallbeispiele von *Digital History* und ihrem Überschneidungsbereich der *Digital Humanities*. »Der Stand der Dinge« schliesst mit einer Zwischenbi-

lanz zum »Nachzügler digitale Geschichte« und bereitet so den Teil zu den »Perspektiven« vor.

In den »Perspektiven« wird zunächst die Öffnung der Werkstatt, das Arbeitsumfeld des Historikers, der Historikerin besprochen. Es folgt sodann eine Auseinandersetzung mit virtuellen Räumen und Zeiten. An dritter Stelle geht es um die *Pro-Amateur*-Revolution, das heißt um die Ausweitung der Beteiligung an der Produktion von historischem Wissen, sowie um ihren zivilgesellschaftlichen Aspekt, das heißt, letztlich, um *Public Digital History*. In *Close reading* geht es, viertens, um die Vorteile, welche die Aufmerksamkeit Texten gegenüber und die vertiefte Interpretation von textlichen Informationen mit sich bringen, um zwei wichtige Alleinstellungsmerkmale der Historiographie, also. Dabei darf eine Reminiszenz an einen *Bestseller* der Renaissance nicht fehlen: ein kurzer Ausblick auf die Zukunft des historischen Buches. An fünfter Stelle wird Hybridität diskutiert, eine Gemengelage, die nicht nur die Medien (analog-digital) sondern auch die Kulturen und damit die Prämissen für die Produktion von Geschichte zunehmend prägt. Es folgt, sechstens, ein Ausblick auf die Philosophie der *Digital History*, bevor ein Fazit das Ausloten der Perspektiven für das zukünftige Vermessen und Interpretieren historischer Welten diesen Teil des Buches schließt.

Der Serviceteil stellt eine Auswahl wichtiger Infrastrukturen, Projekte, Plattformen, Anwendungen, Standards, Zeitschriften und *Blogs* zur digitalen Geschichte beziehungsweise zur digitalen Geschichtswissenschaft im deutsch- und englischsprachigen Raum vor. Einen Schwerpunkt bilden dabei das *Topic Modeling* und die Inhaltsanalyse mit ihren Potenzialen für eine Geschichte, die digital vermessen und analog interpretiert werden wird.

Der Anhang besteht aus einem Abkürzungsverzeichnis sowie einem Verzeichnis der verwendeten Literatur und Webseiten.

Teil I

Der Stand der Dinge

Die Geisteswissenschaften unterliegen einem steten, in jüngster Zeit aber erheblich beschleunigten Wandel. Dieser hat mittlerweile einen ziemlich prägnanten Namen: *Digital Humanities*. Und dieser Name ist Programm – die Geisteswissenschaften sollen fit gemacht werden für das digitale Zeitalter. Emmanuel Le Roy Ladurie, der berühmte Autor von *Montaillou*¹⁴, prophezeite schon 1973, dass zukünftige Historiker entweder Programmierer oder dann gar nicht mehr sein würden.¹⁵

Viele vergleichen die gegenwärtige Umwälzung mit der Erfindung des Buchdrucks in der Renaissance. Der Wandel betrifft nicht nur die Ressourcen, Instrumente und Methoden auf allen Gebieten, in den Informationen generiert, aufbereitet und ausgewertet werden. Dieser Wandel bringt auch die Entwicklung neuer und den Verlust alter Geschäftsfelder mit sich. Ein grundlegender Wandel also, den zwei Protagonisten der digitalen Geisteswissenschaften in Europa, Frédéric Clavert und Serge Noiret, als *Digital turn* bezeichnen.¹⁶ Sie plädieren für eine Zusammenarbeit zwischen Computerspezialisten und Historikern, um diesem Wandel zu entsprechen. Le Roy Laduries Prophezeiung scheint sich also tatsächlich zu bewahrheiten.

Eine besondere Chance und Herausforderung ist das Internet. Seine Einsatzmöglichkeiten destabilisieren die Produktion und Vermittlung von Wissen und demontieren professionelle, zum Teil festgefahrene Autoritäten. Das traditionelle Modell des Autors, der Autorin löst sich auf eine Art und Weise auf, die an die Theorien der Dekonstruktion von Jacques Derrida und Michel Foucault erinnern.¹⁷ Es wird gegenwärtig neu erfunden, wie wir am Beispiel *The Historian's Mac-*

roscope sehen werden. Clavert und Noiret fordern für die Geschichtswissenschaft, ihre Methoden zu erneuern und sich den *Digital Humanities* anzuschliessen. Sie skizzieren damit eine Entwicklung, die eine Konvergenz der Methodik in den verschiedenen Disziplinen der Geisteswissenschaft impliziert.

Serge Noiret macht aus diesem Konzept *Digital History 2.0*. Er orientiert sich kritisch an *Wikipedia*, einem Informationsangebot, das sehr viele *User* anzieht, dessen Inhalte unter Wissenschaftlern aber nach wie vor umstritten sind. *Wikipedia* und ähnliche Wissensangebote arbeiten mit einer je logischen, sichtbaren und aktiven Ebene, einer digitalen Architektur, deren Möglichkeiten für die Geschichte es sorgfältig zu analysieren gilt. Das *Web* bietet eine Chance für *Public History*, in der sich das Verhältnis zwischen historisch Arbeitenden und Publikum neu formiert. Mittlerweile zeigen viele Beispiele, wie Menschen sich beteiligen und ganze »Geschichtslandschaften« bauen, die zuvor nicht denkbar waren.

Der digitale Wandel hat aber auch Auswirkungen auf die Rohstoffe für die Produktion von Geschichte, die Quellen in den Archiven, Druckschriften in den Bibliotheken und Daten in den Repositorien. Und auch die Art und Weise der Produktion von historischem Wissen ändert sich.

Die stille Revolution: Algorithmen, Daten und Netze

Wir sind uns oft gar nicht bewusst, wie rasch digitale Innovationen die Welt verändern. *Google* zum Beispiel hat als ein von der amerikanischen Forschungsbehörde subventioniertes Projekt begonnen und nur kurze Zeit später unsere Medien- und Wissenswelt mit seiner mächtigen Suchmaschine (und anderen Angeboten) komplett verändert. Das gilt auch, um ein zweites Beispiel zu nennen, für *Stats Monkey*, ein von Studenten in den USA entwickeltes Schreib-Programm, das nun selbstständig zum Beispiel einfache Sportberichte erstellt. Die aus Agenturmeldungen halbautomatisch generierten Inhalte für Medien sollen es ermöglichen, durch Preissenkungen Leserinnen und Leser zurückzugewinnen, welche Zeitungen im Zuge der Digitalisierung an das Internet verloren haben. Kurz: Algorithmen erlauben es, Kulturtechniken zu automatisieren. Damit verliert der Mensch eine bis dahin ihm vorbehaltene Domäne an die Maschine.

Für die Historiographie stellt sich die Frage, ob die neuen digitalen Medien »geschichtsmächtig« sind. Sie verwandeln Raum und Zeit in fließende Kategorien, deren Grenzen zu verschwimmen scheinen. Digitale Medien besitzen eine Tiefendimension, die wir noch gar nicht vollständig ausgelotet haben. Arbeiten vernetzen sich, werden im Internet öffentlich zugänglich, der Zugang zu den Quellen beschleunigt sich und die Vermittlung von Forschungsergebnissen erfolgt immer häufiger im Internet. Es findet eine Öffnung statt, welche die Grenzen zwischen den Produzenten und den Konsumenten von historischem Wissen verwischt. Immer mehr Nutzerinnen und Nutzer sind bereit, in Wertschöpfungsketten Arbeiten zu übernehmen, die zuvor Produzenten erledigt haben. Diese Verlagerung führt zu einer massiven Erweiterung des Kreises von Personen, die sich beteiligen. Kurz: Das Potenzial des digitalen Wandels für geisteswissenschaftliche Innovationen ist ohne Zweifel groß.

Für die Geschichtswissenschaft ist essentiell, dass sie der Offenheit des *Web* Standards für die Qualität von Inhalten gegenüber stellt. Das Netz verändert die Beziehungen zwischen Individuum, Kollektiv und Wissen: Dem Verlust von individueller Autorität in Bezug auf Wissen steht der Gewinn von Bedeutung in der kollektiven historischen Sinnbildung gegenüber. Um das organisatorisch zu bewältigen, braucht es neue Forschungsinfrastrukturen. Die Summe all dieser Faktoren ergibt dann eine neue »digitale Geschichtswissenschaft«, sagt Wolfgang Schmale.¹⁸ Auch das *Massachusetts Institute of Technology* (MIT) schätzt das geisteswissenschaftliche Potenzial des digitalen Wandels als hoch ein:

»The Digital Humanities will revitalize the liberal arts tradition in the electronically inflected, design-driven, multimedia language of the twenty-first century«,

fassen die Autoren und Autorinnen diesen Wandel einleitend zusammen.¹⁹ Es geht dabei nicht allein um den Aufbau von digitalen Speichern und das Schaffen von Konventionen für den Zugang, die Nutzung und Publikation digitaler Verzeichnisse, Quellen und Textprodukte, sondern um viel mehr: Der Wechsel von analogen zu digitalen Medien und moderne Web-Technologien ziehen zahlreiche Möglichkeiten wie auch Zwänge nach sich. Dies erfordert ein experimentelles Denken und ein Wirken über die digitale beziehungsweise analoge Spaltung hinweg.

Die Geschichtswissenschaft benötigt Innovationen, einen Weg, um Wissen mit neuen Kompetenz-Sets wie Data-Mining, Inhaltsmodellierung und Design zu produzieren. Benötigt werden eigene Programmiersprachen und neue, nicht ausschließlich quantitative Techniken, wie wir sie heute schon kennen: Interpretatives Mapping, Visualisierung und andere Methoden könnten dann unseren Analysen historischer Welten Formen verleihen, die kulturell und medial unseren hybriden Lebenswelten besser entsprechen. Mehr davon im Kapitel »Hybridität« im Perspektiventeil dieses Buches (vgl. S. 84).

Eine eindruckliche Sammlung von Beispielen darüber, was insbesondere die Visualisierung für die Vermittlung von Wissen bewirken kann, ist der *Atlas of Science*. Der Herausgeber, das MIT, hat den Atlas in Anlehnung an Modelle der Kartographie gestaltet. Er schreibt dazu in der Einleitung:

»Cartographic maps have guided our explorations for centuries, allowing us to navigate the world. Science maps have the potential to guide our search for knowledge in the same way, helping us navigate, understand and communicate the dynamic and changing structure of science and technology. Allowing us to visualize scientific results, science maps help us make sense of the avalanche of data generated by scientific research today.«²⁰

Für historisch Arbeitende ist wichtig zu erkennen, dass *Tools* nicht einfach nur Werkzeuge, sondern kognitive Schnittstellen sind, die Wissen produzieren und vermitteln und deshalb organisiert werden müssen.²¹ Das Vertrauen in Computer für die Auswahl, das Aggregieren und für die Synthese von Daten wächst. Wir befassen uns schon jetzt mit Texten, Zusammenfassungen und Analysen, die von Maschinen erarbeitet wurden. Programme liefern Filter, die den Zugang zu nach bestimmten Kriterien bearbeiteten Aufzeichnungen vereinfachen: Jeder Filter stellt auf unterschiedlichen Ebenen aggregierte Daten für differenzierte Synthesen bereit. Kurz: Begriffe wie *Distant reading*, Inhaltsmodellierung oder Wissensrepräsentation werden bald ebenso vertraut sein, wie Netzwerk und soziale Medien es heute schon sind. Das wiederum bedeutet, dass wir Computer konsequenterweise als Erweiterung unserer eigenen kognitiven Fähigkeiten wahrnehmen sollten, wie Anne Burdick, Johanna Drucker und ihre Kollegen vom MIT betonen.²²

Aber: Wenn das Potenzial des digitalen Wandels so groß ist, weshalb setzt dann nicht auch die Historiographie schon längst auf den Einsatz von Algorithmen? Diese Frage lässt sich nur über den Umweg über das Selbstverständnis der Geschichtswissenschaft und über ihre produktive Voraussetzung beantworten: Eines ihrer Alleinstellungsmerkmale ist traditionell die Interpretation, das Erarbeiten eines tiefen Verstehens von einzelnen Quellen. Sie setzt also auf die qualitative Analyse primär in Einzelfällen. Demgegenüber steht die quantitative Analyse, die in der Regel nur dann Sinn macht, wenn ein umfangreicher Quellenkorpus verarbeitet wird. Und genau diese Voraussetzung muss für die Geschichte erst noch geschaffen werden: Quellen müssen digitalisiert und datafiziert werden. Dies ist eine überaus aufwendige und teure Aufgabe als Grundlage für quantitative historische Analysen. Überdies, das sei an dieser Stelle auch erwähnt, ist Sprache derart komplex, dass automatische inhaltliche Analysen das sorgfältige Le-

sen von Texten kaum je vollständig wird ersetzen können. Mehr dazu im Kapitel *Close reading*, im Perspektiventeil dieses Buches.

Dass sich der Aufwand für die Bereitstellung von digitalen Texten aber durchaus lohnen kann, zeigen Beispiele aus den Politischen Wissenschaften: Mosteller und Wallace haben schon 1963 automatisch die Autoren von nicht verzeichneten *Federalist Papers* erschlossen, in dem sie die Texte dieser Publikation als Daten repräsentierten. Sie haben dazu die funktionalen Wörter eines Textes ausgezählt und das Ergebnis einem Autor zugeordnet.²³ Ein späteres, ausgefeiltes Beispiel für eine Inhaltsanalyse ist das *Expressed Agenda Model* von Justin Grimmer. Der amerikanische Politologe hat 2010 gemessen, welche Bedeutung US-amerikanische Senatoren bestimmten Themen geben. Als Quellenkorpus benutzte er Pressemitteilungen.²⁴

Diese Entwicklung könnte zu einer Zweiteilung innerhalb der Geschichtswissenschaft führen. Das Szenario sieht so aus: Die Protagonisten für Datenanalysen übernehmen Aufgaben, die den quantitativ und empirisch arbeitenden Sozialwissenschaften nahe kommen. Sie untersuchen zum Beispiel die historische Ausbreitung von bestimmten kulturellen Mustern. Demgegenüber verteidigen die Vertreterinnen und Vertreter traditioneller heuristischer Ansätze die Bedeutung der Interpretation für die Forschung und den Zuwachs an historischem Wissen. Diese geisteswissenschaftliche Kompetenz sei wichtig, um Subjektivität, Mehrdeutigkeit und Komplexität als Aspekte der Forschung nicht aus dem Blick zu verlieren, argumentieren sie. Wohl zurecht: Das Wissen um die menschliche Dimension maschineller Verfahren ist gerade im digitalen Zeitalter besonders gefragt.

Diesen Punkt unterstreicht auch der Kulturwissenschaftler Gerhard Lauer. Für ihn ist die Digitalisierung zwar eine Revolution, welche die Hierarchie der geisteswissenschaftlichen Werte erschüttert. Die Aufmerksamkeit gilt neu dem digitalen *Service* am Text, das heißt seiner Standardisierung in Form von Daten und Metadaten, seiner algorithmischen Berechenbarkeit. Hermeneutisches Wissen ist aber deswegen noch lange nicht überflüssig. Denn Daten sind noch keine Informationen. Wer Daten zum Sprechen bringen will, muss sie kontextualisieren.²⁵ Wir kennen diese Diskussion aus der Statistik: Zahlen allein erzählen keine Geschichte.

Es geht also um die Zukunft der Geschichtswissenschaft unter digitalen Vorzeichen. Vieles von dem, was Historikerinnen und Historiker erwartet, zeichnet sich bereits in den *Digital Humanities* ab. Die

digitalen Geisteswissenschaften sind ein interdisziplinär ausgerichtetes Praxisfeld, das sich, ausgehend von Universitäten, Forschungsprojekten und Gedächtnisinstitutionen, zurzeit rasch entwickelt. Es umfasst insbesondere den Aufbau digitaler Forschungsinfrastrukturen sowie die Entwicklung von computergestützten Verfahren und Anwendungen. Dabei zeichnet sich ab, dass die Methoden verschiedener geisteswissenschaftlichen Disziplinen unter digitalen Vorzeichen bis zu einem gewissen Grad konvergieren.

Der Begriff der digitalen Geisteswissenschaften ist eine Übersetzung der angloamerikanischen Bezeichnung *Digital Humanities*. Die digitalen Geisteswissenschaften haben im deutschsprachigen Raum aber auch eigene Vorläufer wie die historische Fachinformatik, Computerlinguistik oder Computerphilologie. Die *Digital Humanities* können in ihrem Verhältnis zu den schon länger bestehenden Fächern auf unterschiedliche Weise beschrieben werden. Manche sehen in ihnen eine Hilfswissenschaft zu den geisteswissenschaftlichen Fächern, deren Fragestellungen sie aufgreifen. Manche betrachten sie als eine angewandte Informatik, die für einen speziellen Gegenstandsbereich Informatik-Lösungen entwickelt. Wieder andere sehen in ihr eine allumfassende Geisteswissenschaft auf der Ebene der Methoden und digitalen Praktiken. Erst mit eigenen Forschungsfragen und eigenen Methoden würden die *Digital Humanities* tatsächlich eine eigenständige wissenschaftliche Disziplin abgeben. Vielleicht müsste man dann aber die Bezeichnung ändern, weil das digitale Moment allein nicht hinreichend ist, um einen wissenschaftlichen Bereich abzugrenzen.

Digital Humanities, in denen sich die Methoden verschiedener Disziplinen unter digitalen Vorzeichen einander nähern, schafft neue Voraussetzungen für interdisziplinäre Forschungen und Projekte. So gesehen ist die Vorstellung von einer zukünftigen Zweiteilung der *Community* relativ. Die Grenzen sind nicht immer klar. Auf der *Digital-Humanities-Konferenz* 2012 in Hamburg befürchteten einige Wissenschaftler den Import eines reduktiven, nicht auf konkrete Inhalte bezogenen Denkens in die Geisteswissenschaften, während andere die *Digital Humanities* entweder als Unterstützung für oder sogar als Speerspitze einer vollständigen Überführung der Geisteswissenschaften begrüßten.²⁶

Sicher ist, dass sich der Status des historischen Dokuments und mit ihm die Vorstellung von Autorenschaft gerade massiv verändern. Texte und andere digitale Objekte werden »flüssiger«. Immer öfter arbei-

ten verschiedene Autoren an einem Textkorpus, Schreiben und Auswerten wird immer häufiger als ein kooperativer Prozess verstanden. Unsere Vorstellung vom Dokument wird sich entsprechend anpassen müssen. Es wird, ganz im Sinne von Michel Foucault, sich in ein diskursives Objekt verwandeln, in ein Produkt additiver Interventionen verschiedener Autoren.²⁷

Der Geschichte kommt dabei deshalb eine besondere Sprengkraft zu, weil ihre Quellen in ihrer Materialhaftigkeit eine besondere Wirklichkeit aufweisen, wie insbesondere *Wikileaks* eindrücklich zeigt. Die Geschichtswissenschaft produziert ihre Daten nicht selbst, im Gegensatz z.B. zu den Sozialwissenschaften, sondern findet sie vor. Quellen sind Zeugnisse einer vergangenen Wirklichkeit, die Spur einer verstummen Stimme. Sie zehren von der Kraft ihrer Authentizität. Denn, wenn etwas wirklich ist, bedeutet das noch lange nicht, dass es auch wahr sein muss. Diese »Lücke des Wirklichen« wird anhand von Wissen überprüft.²⁸ Bei Quellen zeugt die Echtheit des Materials von der Wirklichkeit. Sie macht die Quelle zum beständigen Zeugen. All dies rückt das Quellenmaterial in die Nachbarschaft einer vermeintlichen Wahrheit.

Es geht also zukünftig auch in der Geschichte um die Interaktion zwischen Mensch und Maschine, um die Gestaltungskraft der technischen Infrastruktur im Verhältnis zu Symbolsystemen wie Quellen, Texten, Bildern, Filmen, Videos, Büchern, sozialen Medien oder dem Internet. Menschen sind mit ihrer Fähigkeit zur (Selbst-)Darstellung auf das Schaffen von Wirklichkeit hin angelegt (Ian Hacking). Die Konstitution von historischen Erfahrungsräumen ist dabei besonders spannend, weil sie anhand von verschiedenen Zeugnissen überprüft werden können. Der Historiographie bietet sich die große Chance einer historischen Sinnstiftung für ein breites Publikum, das in der virtuellen, digitalen Welt nach stabilem Wissen und Orientierung verlangt.

Der Wechsel von analogen zu digitalen Medien und moderne Web-Technologien ziehen also viele Möglichkeiten wie auch Zwänge mit und nach sich. Dies erfordert ein experimentelles Denken über die digitale beziehungsweise analoge Spaltung hinweg. Dabei stellen aber die Sicherung digitaler Quellen und der Umgang mit ihnen eine besondere Herausforderung dar, wie das folgende Kapitel zeigt.

Zeitmaschinen: Archive, Bibliotheken

Wer wird die SMS von Angela Merkel archivieren? Kiran Klaus Patel hat diese Frage schon 2011 gestellt.²⁹ Sie ist bis heute nicht gelöst. Nicht nur die Geschichte, sondern auch ihr Rohstoff, die Quellen, deren Überlieferung und deren Verfügbarkeit unterliegen einem tiefgreifendem Wandel. Es ist deshalb notwendig, dass Archiv- und Geschichtswissenschaft die materiellen und konzeptionellen Grundlagen insbesondere der Zeitgeschichte vertieft diskutieren.

Patel fühlte sich vor allem durch die Menge der von *Wikileaks* veröffentlichten Dokumente aus dem US-Außenministerium überwältigt und herausgefordert. Er fragte sich, wer all diese Unterlagen auf ihre Authentizität hin prüfen und dann auswerten solle. Die Gründer von *Wikileaks* beantworteten diese Frage teilweise, in dem sie die Unterlagen verschiedenen Zeitungen insbesondere in Deutschland, Großbritannien und den USA zur Verfügung stellten.³⁰ Dieses spezielle *Crowdsourcing* löst die grundlegende Frage aber nicht: Die Frage nach der Überlieferung und Bewertung von Informationen aus dem Internet und den sozialen Medien wie SMS, *Twitter*, *Facebook* oder *WhatsApp*.

Patel verlangt, dass solche Informationen von Regierungen sicher gestellt und der Forschung möglichst rasch zur Verfügung gestellt werden. Die in öffentlichen Archiven übliche Schutzfrist von 30 Jahren müsse abgeschafft werden. Das hätte einen massiven Wandel in der Praxis der Zeitgeschichtsforschung zur Folge – sie erfolgte dann sehr zeitnah zu den Ereignissen. Während Gesetze zum Öffentlichkeitsprinzip der Verwaltungen in den europäischen Ländern diesem Anliegen entgegen kommen, wird es umgekehrt durch Datenschutz und Geheimhaltung schwieriger, diesem Verlangen nach Transparenz zum Durchbruch zu verhelfen.

Andreas Wirsching, der Direktor des Münchner *Instituts für Zeitgeschichte* (IfZ), hat gezeigt, wie Zeitgeschichtsschreibung unter hy-

per-modernen Bedingungen von digitalen Medien und »Informationsflut« funktionieren könnte. Sein Buch »Der Preis der Freiheit« von 2012 über die »Geschichte Europas in unserer Zeit«³¹ hielt sein Kollege, der Historiker Ulrich Herbert, für »ein mutiges Unterfangen«³². Wirsching hat nicht das Ablaufen von Sperrfristen in den Archiven abgewartet, sondern aus einer Menge von öffentlichen amtlichen Dokumenten, Presseartikeln und wissenschaftlichen Analysen geschöpft. Die Risiken eines solchen Unterfangens minimierte der Autor, in dem er nicht chronologisch, sondern thematisch vorging. Offenbar hat sich das Risiko gelohnt: Die *Neue Zürcher Zeitung* (NZZ) fand, dass Wirsching schon jetzt »ein Standardwerk zur Geschichte Europas seit 1989« vorgelegt habe.³³

Aber die Überfülle an Informationen ist nicht die einzige Herausforderung, die es zu bewältigen gilt: Im Informationszeitalter schlägt das quantitative Problem – die Menge an Informationen – in ein qualitatives Problem um. Viktor Mayer-Schönberger verlangt deshalb eine »Informationsökologie«³⁴, das heißt eine verbindliche Regelung, welche Informationen gesammelt, gespeichert und daher erinnert werden sollen – etwas das Archive schon lange tun: Das Bewerten der Überlieferungswürdigkeit von Dokumenten und Daten. Öffentliche Archive wie zum Beispiel das *Schweizerische Bundesarchiv* (BAR) übernehmen rund 10 % der von Parlament, Regierung und Verwaltung produzierten Unterlagen. Das entsprach im Papierzeitalter durchschnittlich rund 1600 Laufmeter jährlich. 2015 hielt das BAR 62 500 Laufmeter an Akten in seinen Magazinen. Dies alles zu digitalisieren stellt eine enorme Herausforderung dar.

Die Bewertungskriterien in den deutschsprachigen, öffentlichen Archiven sind zumeist der Evidenz- und der Informationswert: Welche Unterlagen sind wichtig für die Nachvollziehbarkeit eines Geschäftes? Welche haben einen hohen Informationswert für zukünftige Forschungen? Mit diesen auch an der Archivschule in Marburg entwickelten Prinzipien soll die demokratische Kontrolle im Nachhinein, eine zentrale Aufgabe behördlicher Archive, gewährleistet werden, ohne dass die Nutzer in der Flut an Informationen versinken.

Aber: Ist die Rede vom *Information Overload* überhaupt richtig? Im Zeitalter von *Big Data* – dazu mehr im nächsten Kapitel – und mächtigen Suchmaschinen stellt sich die Frage, ob Unterlagen überhaupt noch bewertet werden sollen. Mit den neuen digitalen Speichertechniken erleben wir eine völlig neue Dimension im Kampf gegen

das Vergessen. Mittlerweile scheint es weniger Zeit und Geld zu kosten, Daten zu speichern, als die Frage des Speicherns oder Löschens zu prüfen und zu entscheiden. Ob dies allerdings auch gilt, wenn die Sicherung eine sehr langfristige sein soll, ist noch offen. Die Lebensdauer von Datenträgern ist begrenzt, so dass Daten immer wieder migriert werden müssen. Professionelle Archive minimieren das Risiko von Datenverlusten mit redundanten Speichersystemen, so dass auch dadurch die Kosten steigen. Und ob die digitale Archivierung tatsächlich funktioniert, wird sich erst dann zeigen, wenn die in digitalen Magazinen »eingefrorenen«, das heißt von ihren Anwendungen getrennten Daten hervorgeholt und für eine erneue Nutzung wieder aufbereitet werden.

Jedenfalls: Die Digitalisierung hat vielfältige Auswirkungen auf die zukünftige Praxis sowohl der Anbieter wie der Nutzer von Archiv-beziehungsweise Bibliotheks- und Sammlungsmaterial. Dieser Schritt vom *Gate Keeper* zum *Data Broker* der Archive, Bibliotheken und anderen Dokumentationsstellen lässt sich anhand eines modellhaften Nutzungsvorgangs wie folgt veranschaulichen:

Jemand, der eine Information aus der Vergangenheit benötigte, suchte bis vor kurzem das Archiv, die Gedächtnisinstitution, physisch auf. Er meldete sich an, suchte in den Findbüchern, ließ sich einschlägige Akten in den Lesesaal bringen, blätterte darin, revidierte seine Suchstrategie, manchmal auch seine Fragestellung, suchte wieder in den Findbüchern, ließ sich neue Findbücher und Dokumente bringen und so weiter. Tage konnte er so verbringen, bis er fand, wonach er suchte. Kurz: Das Archiv ist eine Zeitmaschine. Die Vereinbarung Nazi-Deutschlands und der Schweiz von 1938, um die Pässe deutscher Juden separat zu markieren, zum Beispiel, wurde zusammen mit anderen Flüchtlingsakten in den 1950er Jahren an das BAR abgegeben. Sie können nun, 50 Jahre später, mit einigem Aufwand aus der Vergangenheit in die Gegenwart geholt werden.

Was bedeutet es für die Nutzung und die Forschung, wenn solche historische Quellen digital beziehungsweise in digitalisierter Form *online* zugänglich gemacht werden? Sie sind dann mit einem Klick auf dem Schirm, als Bild virtuell also ständig präsent. Werden sie so Teil der Gegenwart? Das digitalisierte Archivgut scheint aus der Vergangenheit heraus in die Gegenwart hinein zu fallen. Die Zeitmaschinen werden in ihrer traditionellen Form bald nicht mehr gebraucht.³⁵ Wer sich für die archivwissenschaftliche Perspektive dieser Fragestellung