



Guido Koller

Geschichte digital

Historische Welten
neu vermessen

Kohlhammer

150 Jahre
Kohlhammer

Guido Koller

Geschichte digital

Historische Welten neu vermessen

Verlag W. Kohlhammer

1. Auflage 2016

Alle Rechte vorbehalten

© W. Kohlhammer GmbH, Stuttgart

Gesamtherstellung: W. Kohlhammer GmbH, Stuttgart

Print:

ISBN: 978-3-17-028929-1

E-Book-Formate:

pdf: ISBN 978-3-17-028930-7

epub: ISBN 978-3-17-028931-4

mobi: ISBN 978-3-17-028932-1

Für den Inhalt abgedruckter oder verlinkter Websites ist ausschließlich der jeweilige Betreiber verantwortlich. Die W. Kohlhammer GmbH hat keinen Einfluss auf die verknüpften Seiten und übernimmt hierfür keinerlei Haftung.

Inhaltsverzeichnis

Einleitung	_____	7
Teil I	Der Stand der Dinge _____	15
1	Die stille Revolution: Algorithmen, Daten und Netze	17
2	Zeitmaschinen: Archive, Bibliotheken	23
3	Vom Nutzen großer Zahlen: Big Data und historische Analyse	29
4	Neue Praktiken in der Geschichtswissenschaft	36
5	Neue Anbieter: Digitale Infrastrukturen für die Forschung	43
6	Digital History: Fallbeispiele	49
7	Zwischenbilanz: Digitale Geschichte, der Nachzügler	64
Teil II	Perspektiven _____	67
1	Die Werkstatt der Historiker öffnen	70
2	Virtuelle Räume und Zeiten	73
3	Everyman His Own Historian: Die Pro-Amateur-Revolution	77
4	Close reading: Vom Vorteil der Aufmerksamkeit	80
5	Hybridität (Medien, Kultur, Geschichte)	84
6	Die Philosophie der Digital History	87

7	Fazit: Das Vermessen und Interpretieren historischer Welten	91
Teil III	Services _____	99
1	Infrastruktur	100
2	Portale / Plattformen	103
3	Verbände und Zeitschriften	110
4	Blogs	115
5	Text Mining / Topic Modeling / Inhaltsanalyse	122
6	Tools / Maschinenlesbare Sprachen / Standards	127
7	Editionen	136
	Abkürzungen _____	139
	Literatur, Artikel _____	142
	Webseiten und soziale Medien _____	147
	Endnoten _____	150

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 Umriss- weite- terer Gestaltungsmöglichkeiten in den »Perspektiven« vorgestellt.

Das Schreiben von Geschichte lässt sich als Produktion von histori- schem 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. Aufbe- reitete, analysierte und synthetisierte Informationen, das Wissen also, wird sodann an ein Publikum vermittelt sowie von diesem angewen- det und geteilt. Dieses Publikum kann ein Fachpublikum – die wissen- schaftliche *Community* – oder ein breiteres Publikum sein – die Öff- entlichkeit. 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 fürsorgerischen 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 Wis- sens, hat sich jüngst mit der Beschleunigung dieses Prozesses ausein- andergesetzt.¹² Dieses Buch beleuchtet die Stationen des historiogra- phischen 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 Produktik 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 histori- sche 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 not- wendigen Infrastruktur und ihre neuen Anbieter sind das fünfte The- ma. 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-*