

Forschungsinfrastrukturen in den digitalen Geisteswissenschaften.

Wie verändern digitale Infrastrukturen die Praxis
der Geisteswissenschaften?

Martin Huber, Sybille Krämer, Claus Pias
Symposienreihe „Digitalität in den Geisteswissenschaften“

Gefördert durch

DFG Deutsche
Forschungsgemeinschaft

IMPRESSUM

HERAUSGEBER

Martin Huber, Sybille Krämer, Claus Pias

KONTAKT

Julia Menzel

Digitalität in den Geisteswissenschaften

DFG-geförderte Symposienreihe

Universität Bayreuth

Universitätsstr. 30

95447 Bayreuth

www.digitalitaet.dfg@uni-bayreuth.de

1. Auflage November 2019

Wir danken der Deutschen Forschungsgemeinschaft (DFG Projekt Projektnummer 287972711) für die Förderung.

Ist da eine Veränderung in den Geisteswissenschaften? Infrastrukturen und ihre Folgen für die Praxis der Geisteswissenschaften

Gerhard Lauer (Basel)

Im zweiten Buch seiner Abhandlung *The Advancement of Learning* von 1605 hat Francis Bacon die Infrastrukturen für die zu gründenden modernen Wissenschaften umrissen. Für das, was wir heute näherungsweise Geisteswissenschaften nennen, nennt er zwei wesentliche Voraussetzungen gelingender Wissenschaften. Bacon schreibt:

The works touching books are two — first, libraries, which are as the shrines where all the relics of the ancient saints, full of true virtue, and that without delusion or imposture, are preserved and reposed; secondly, new editions of authors, with more correct impressions, more faithful translations, more profitable glosses, more diligent annotations, and the like.¹

Bacon zählt mit guten Gründen unter die Gründungsväter der modernen Wissenschaften im konfessionellen Zeitalter.² Infrastrukturen für die neuen Wissenschaften, das waren für Bacon Bibliotheken und Editionen. Sie ermöglichen, was wir heute zumeist Geisteswissenschaften nennen und das umso mehr, wenn man unter Bibliotheken auch Archive und Museen, unter Editionen auch Sammlungen und die vielen Formen des Weiterschreibens und Übersetzens in Büchern, Zeitschriften bis zum Feuilleton versteht, die geisteswissenschaftliches Arbeiten vielfach ausmachen. Man sieht unschwer, dass die geisteswissenschaftlichen Infrastrukturen zunächst buchzentriert waren. Ohne Bibliotheken und Editionen sind Geisteswissenschaften kaum eine Praxis und keine Disziplinen und das bis heute. Dass den buchzentrierten Fächern der Theologie und der Klassischen Philologie lange eine historische Vorrangstellung zukam, ist daher kein historischer Zufall. Sie sind die Buchwissenschaften schlechthin.

Doch Infrastrukturen waren in der Frühen Neuzeit bald schon mehr als Bibliotheken und Editionen. Hinzu kamen Bibliographien und Rezensionsorganen wie die *Göttingischen Gelehrten Anzeigen*, Akademien für die Forschung, da in den Universitäten nicht geforscht, sondern gelehrt wurde. Dazu kamen auch gelehrte Korrespondenznetzwerke und schließlich auch gegen Ende des 19. Jahrhunderts größere, lang angelegte Forschungsprojekte wie Boeckhs *Corpus Inscriptionum Graecarum*, Mommsens epigraphische und prosopographische Projekte, die *Monumenta Germaniae Historica*, Kirchenväterkommission, *Prosopographie Imperii Romani* oder das *Corpus Inscriptionum Latinarum*.³ Diese Projekte wurden in enger Nachbarschaft und zugleich Konkurrenz zu der entstehenden Großforschung in den Naturwissenschaften konzipiert. Mommsen und Harnack, Geistes- und Naturwissenschaften teilten die identischen forschungspolitischen Strukturen des Kaiserreichs.

¹ vgl. Bacon, Francis: *The Advancement of Learning*. In: Francis Bacon. *The Major Works*. Edited with an Introduction and Notes by Brian Vickers. Oxford: Oxford University Press 1996 [1605], S. 170.

² vgl. Gaukroger, Stephen: *Francis Bacon and the Transformation of Early-Modern Philosophy*. Cambridge: Cambridge University Press 2010.

³ vgl. Rebenich, Stefan: *Theodor Mommsen und Adolf Harnack. Wissenschaft und Politik des ausgehenden 19. Jahrhunderts*. Berlin: de Gruyter 2012.

Forschungsinfrastrukturen waren daher schon zu Beginn des 20. Jahrhunderts mehr als nur Editionen und Bibliotheken, waren bald schon nicht nur buchzentriert. Sie sind es noch weniger als mit den Massenuniversitäten des 20. Jahrhunderts auch die Geisteswissenschaften weit mehr als nur Editionen und Bibliotheken brauchen, um zu funktionieren. Der Humboldt-Kult der 20er Jahre des 20. Jahrhunderts war schon eine Reaktion darauf, dass auch die Geisteswissenschaften kaum den idealistischen Idealen der Weimarer Klassik entsprachen als sie zwischen den Kriegen ihren Status ausbauen konnten. Das Pathos von Einsamkeit und Freiheit, das damals so gerne beschworen wurde und sich bis heute in den bildungspolitischen Debatten gehalten hat, war eine ideologische Konstruktion vor hundert Jahren und die Humboldt-Edition von Eduard Spranger u.a. ist ein Zeugnis für diese Selbstverklärung der Geisteswissenschaften angesichts von Großforschung und – wie man damals schon es empfand – Massenuniversität der 20er Jahre.⁴ Wenn schon im 19. und erst recht im 20. Jahrhundert Infrastrukturen komplexe Institutionen waren, um wieviel weniger passen Sentimentalitäten vom guten Buch und den einfachen Infrastrukturen zu den Geisteswissenschaften des 21. Jahrhunderts.

1. Transformation der Geisteswissenschaften

Damit bin ich bei meiner ersten These: Die Digitalisierung verstanden als Transformation einer Wissenschaft durch digitale Ressourcen und veränderte Methoden, ihrer anderen Darstellung und Vernetzung in der Gesellschaft, diese Digitalisierung ist nur ein, wenn auch gewichtiger Faktor in der Umgestaltung der Wissenschaften. Nicht alles ist neu, was unter Digitalisierung ausgeflaggt wird, aber vieles schon und das hat längerfristige Folgen, von denen hier zu reden sein wird. Dennoch wäre es eine grobe Vereinfachung, Faktoren wie die veränderten Bildungszugänge, die Globalisierung der Wissenschaften, ihre interne Dynamik der Ausdifferenzierung, die massiven ökonomischen Interessen und andere Faktoren außen vor zu lassen oder diese unter Digitalisierung zu subsumieren. Moderne Gesellschaften sind multifaktoriell bestimmt. Wenn daher von der Digitalisierung, hier den digitalen Infrastrukturen, die Rede sein muss, ist darauf zu achten, den Entdifferenzierungen und Simplifikationen aus dem Weg zu gehen, die sich nur zu leicht einstellen, wenn von Digitalisierung die Rede ist. Das Schlagwort ‚Digitalisierung‘ tendiert zur Simplifikation. Es spricht viel dafür, die damit implizierten ideologisierten Debatten zu vermeiden.

Was also heißt und zu welchem Ende reden wir von digitalen Infrastrukturen? Ich schichte das Problem ab, indem ich versuche differenzierter zu fragen: Welche Veränderungen im Gegenstandsfeld der Geisteswissenschaften gibt es und was zeitigt Folgen für die Infrastrukturen der Geisteswissenschaften? Welche Veränderungen der Methoden ziehen Veränderungen in den Infrastrukturen nach sich? Wie verändern sich Rollen und Akteure im geisteswissenschaftlichen Forschungs- und Lehrbetrieb, wie ändern sich die Umwelten, so dass Infrastrukturen für die Geisteswissenschaften des 21. Jahrhunderts andere sind als die für das 20. Jahrhundert.

⁴ vgl. Gerhard (Hg.): Wilhelm von Humboldt. Schriften zur Bildung. Stuttgart: Reclam 2017.

2. Infrastrukturen und die Gegenstände der Geisteswissenschaften

Das Gegenstandsfeld der Geisteswissenschaft unterlag und unterliegt einer Erweiterung längst vor der Digitalisierung. Die Ausdehnung von den klassischen und biblischen Gegenständen auf mittelalterliche Überlieferungen im 19. Jahrhundert und bald schon auch auf außer-europäische Gegenstandsbereiche hat viele neue Fächer wie die Orientalistik, Keltologie oder Germanistik erfunden. Fächer wie die Amerikanistik oder Neugermanistik kamen im 20. Jahrhundert dazu, weil sich die Relevanz von Fragestellungen gewandelt hatte. Die Geisteswissenschaften haben schließlich das Gegenstandsfeld geisteswissenschaftlicher Fächer soweit ausgedehnt, dass in vielen Fächern nur noch wenig Konsens darüber herrscht, was Gegenstand des jeweiligen Fachs ist. Ist in der Anglistik der Fahrplan der Londoner Tube, in der Archäologie auch die Wissenschaftspopularisierung in der Museumskultur der Gegenwart ein disziplinär akzeptierter Gegenstand? Die Antworten auf solche und ähnliche Fragen fallen unterschiedlich aus.

Die Reihe lässt sich leicht fortsetzen, wenn es etwa Debatten darüber gibt, ob *Latino studies* und *Latina studies* zwei verschiedene Gegenstandsfelder und evtl. auch Disziplinen meinen könnten. Dem teilweise bizarren Erfindungsreichtum der Geisteswissenschaften für neue Gegenstandsfelder lassen sich leicht reichlich ironische Wendungen abgewinnen. Dennoch lohnt es sich genauer hinzusehen. Gegenstandsfelder sind disziplinäre Konstruktionen, kein Zweifel, und sie folgen binnenwissenschaftlichen wie außerwissenschaftlichen Faktoren.⁵ Digitalisierung ist daher kein notwendiger Gegenstand der Geisteswissenschaften und nicht wenige Kolleginnen und Kollegen sehen das genau so. Dass es mehr Kunst, Musik und Literatur außerhalb des Kanons gibt, führt eher zu einem Unbehagen in der Kultur. Die Reaktionen sind dann mehr oder weniger gut begründete Varianten der Kulturkritik vom Niedergang der Bildung, dem moralischen Verfall der Jugend und dem Ende der Kultur. Das kaum noch jemand lesen, ist die gängigste Variante dieser Kulturkritik. Und sie führt zu einer stillen Selbstabschaffung der Geisteswissenschaften, der die Studenten wegbleiben.

Es ist also eine in jedem Fach, aber auch in den Fakultäten und den Universitäten jeweils zu treffende Entscheidung, ob die objektiven Veränderungen in der Kultur disziplinäre Veränderungen in den Geisteswissenschaften und ihren Infrastrukturen nach sich ziehen. Mit objektiven Veränderungen im Gegenstandsfeld der Geisteswissenschaften meine ich zunächst die digitale Massenerschließung des kulturellen Erbes von der Keilschrift-Überlieferung in Portalen wie [CDLI](#) bis zu [Bildersammlungen der Getty-Foundation](#), der Maya-Inschriften bis zur [digitalen Mozart-Ausgabe](#). Das dehnt das mögliche Gegenstandsfeld der Fächer erheblich aus und vertieft deren Forschungsmöglichkeiten zugleich. Die Breite der erschlossenen Überlieferung wie auch deren Tiefe, genauer die Granularität, mit der Details etwa von Keilschrift-Inschriften erfasst werden können, so dass anhand hochauflösender Photographien Zusammenhänge zwischen Tontafeln sichtbar werden, die für das bloße Auge nicht erkennbar sind, das alles ist eine neue Qualität im Gegenstandsfeld.

Keine dieser Erschließungen und Editionen ist möglich ohne eine mehr als proprietäre Infrastruktur. [FuD](#) in Trier und das [DaSCH](#) in Basel sind virtuelle Forschungsumwelten für Editionen,

⁵ vgl. Diemer, Alwin: Art. Geisteswissenschaften. In: Historisches Wörterbuch der Philosophie 3 (1974), S. 211-215, sowie Hollinger, David. A.: *The Humanities and the Dynamics of Inclusion since World War II*. Baltimore: Johns Hopkins University Press 2006.

in die auch ältere Pioniereditionen wie etwa Walter Morgenthals Gottfried Keller-Ausgabe gehoben werden. Standards wie TEI und jetzt in der Musikwissenschaft MEI haben sich aus dieser Ausdehnung des Gegenstandsfelds ergeben und weltweit geteilte Standards haben sich dank der Beharrlichkeit einer kleinen Gruppe von Kolleginnen und Kollegen entwickelt. Das wird durch höherstufige Infrastrukturen wie [DARIAH](#) und [CLARIN](#) unterstützt, die helfen, digitale Ressourcen nach gemeinsamen Standards zu archivieren, zu teilen und Werkzeuge zu deren Erschließung einzubinden. Hat man diesen Pfad einmal eingeschlagen, dann hat das wiederum Rückwirkungen auf das Gegenstandsfeld. Denn statt nur Editionen in der Tradition der Werkausgaben zu betreiben, werden Gegenstandsfelder unter Portalen wie etwa [Perseus](#) zusammengeführt, welches die gesamte schriftliche Überlieferung der Antike und ihre Rezeptionsgeschichte vereint. Genauer: Aus Werken werden Korpora. In der Linguistik hat dies ein ganzes Fach, eben die Korpuslinguistik, hervorgebracht. In kleinen Fächern wie der Koptologie, die nur selten die Werkästhetik der großen Fächer kennen, führt dies zu einer Neuvermessung ihrer Gegenstände. Statt nur die wenigen, traditionell hochbewerteten Gegenstände in den Blick zu nehmen, werden historisch lange Reihen vielfach nur höchst fragmentarischer Überlieferungen zum Gegenstand solcher Fächer. Der Gegenstandsbereich dieser Fächer verändert sich damit nachhaltig. Gerade für die kleinen Fächer ist daher die Digitalisierung ihrer Gegenstände die Chance für eine fachpolitische Erneuerung in Richtung auf eine Vertiefung ihrer Gegenstände und deren Zusammenhänge. Mehr Gegenstände in einer höheren Auflösung und in einer historischen Vernetzung, die noch vor kurzem unmöglich schienen, beginnt die Fächer und deren Gewichte zu verschieben, vorausgesetzt, ihnen gelingt der Aufbau von fachadäquaten Infrastrukturen.

Das ist im Detail dann aufwändig und nicht alle Fächer verfügen dafür über ausreichende Ressourcen, solange die Fakultäten in ihren Strukturen des 20. Jahrhunderts aufgestellt bleiben, wie das gegenwärtig noch überwiegend der Fall ist. Wir haben versucht, ein Korpus zur Klassischen Moderne aufzubauen (<https://kolimo.uni-goettingen.de/index.html>). Es umfasst etwa 40.000 Texte von mehr als 2.000 Autoren der deutschen Literatur um 1900. Da eine Korpusliteraturwissenschaft (noch) nicht existiert, gibt es auch keine Infrastrukturen aus Datenbanken-Standards, Webspaces, Urheberrecht und überhaupt einem Geist des Teilens von Ressourcen, um ein repräsentatives und balanciertes Korpus der Literatur um 1900 aufzubauen, also ein Korpus, das die verschiedenen Register, Domänen, Gattungen und Genre, Autorinnen und Autoren, Schulen und Richtungen historisch möglichst genau abbildet.⁶ Vom Rechenzentrum bis zu den Metadatenstandards müssen jeweils erst einmal Lösungen formuliert, vorgeschlagen und erprobt werden, das aber in einem Fach, das selbst noch gar nicht weiß, ob Infrastrukturen für den Bau und Betrieb von Korpora überhaupt zu diesem Fach, hier der Neugermanistik, gehören. Texte von Kafka oder Thomas Mann sind zu einem erheblichen Teil für eine nicht nur buchzentrierte Forschung gar nicht zugänglich, jedenfalls nicht legal. Speicherformate und Metadaten sind nicht vereinheitlicht, wenn man etwa die Standards des [Deutschen Textarchivs](#) mit denen von [Gutenberg](#) und [TextGrid](#) vergleicht. Zitierfähige URLs für wissenschaftliches Bibliographieren müssen erst entwickelt und Datenzentren für geisteswissenschaftliche Daten erst etabliert werden. Das alles ist aufwändig und es genießt im Fach kein Ansehen.

Die Schwierigkeiten nehmen noch zu, wenn man das digital born-Material in den Blick nimmt. Denn nicht nur historisch in die Tiefe und kulturell in die Breite erlaubt die Digitalisierung die

⁶ Herrmann, Berenike / Lauer, Gerhard: Korpusliteraturwissenschaft. In: Osnabrücker Beiträge zur Sprachtheorie 92 (2018), S. 127-156.

Transformation der geisteswissenschaftlichen Gegenstandsfelder. Erstmals können wir heute in Echtzeit verfolgen, wer, was, wo schreibt, teilt, kommentiert, und weiterschreibt. Remix ist das Stichwort, das die historisch einmaligen Möglichkeiten des Fortschreibens im digitalen Zeitalter umschreibt, das die Werkästhetik nur als einen Fall neben vielen anderen Verdichtungen und Auflösungen von Ausdrucksformen kennt. Nicht dass es E-Books und Games gibt, verändert das Gegenstandsfeld der Geisteswissenschaften so sehr, sondern die Möglichkeiten des Fort- und Umschreibens in fast alle Richtungen. Dafür verfügen die meisten geisteswissenschaftlichen Fächer noch nicht über ausreichende Infrastrukturen. Und die Ablehnung der digitalen Kulturwelt bzw. der Rückzug auf die Werkästhetik dominiert viele Fächer.

Ein Beispiel mag das Gesagte illustrieren: [Wattpad](#). Hier schrieben bislang mehr als 40 Millionen Menschen in sehr vielen Sprachen dieser Welt Literatur, vor allem Romanzen und Fantasy. [Wattpad](#) wirbt damit, dass eine von drei jungen Frauen einen [Wattpad](#)-Account habe. Das dürfte eine Übertreibung sein, macht aber den Anspruch deutlich, dass es eine digitale Kultur gibt, von der kein Feuilleton Notiz nimmt und die doch für so viele Menschen ihre Kultur ist. Entscheidend ist hierbei, dass diese digitale Kultur beständig Daten erzeugt. Wir verfügen nicht nur über die Texte, sondern auch über die Daten der Nutzer. Wir wissen, welche Geschichten und Plots mehr Aufmerksamkeit auf sich ziehen als andere, was die nationalen und regionalen Präferenzen etwa von Hörbüchern sind, welche Altersgruppen welche Musik hören und welche nicht, wie sich Musik durch Streamingdienste verändert oder ob Frauen und Männer immer noch anderes lesen – fast. Denn diese Daten gehören zumeist den Unternehmen wie [Wattpad](#). Die Extraktion von Daten über die ästhetischen Präferenzen der Gegenwart ist nur äußerst aufwändig zu erheben. Die Rechtslage ist, vorsichtig gesagt, unklar. Unklar damit, wie Daten geteilt werden und Ergebnisse überprüfbar gemacht werden können. Der Umfang der Daten passt nicht mehr in die etablierten Gefäße der philologischen Fächer. Von Tausenden, ja Millionen von Datenpunkten zu sprechen, ist keine Übertreibung, sondern das Problem. Für unsere Forschergruppe zum Lesen und Schreiben auf den sozialen Plattformen wie [Wattpad](#), [Goodreads](#), [Amazon Reviews](#), müssen wir Infrastrukturen erst aufbauen, um zu erforschen, welche Literatur die heutige Gesellschaft schreibt, teilt und liest.

Anders gesagt, bedeuten die Veränderungen im Gegenstandsfeld der Geisteswissenschaften, dass sie nur sehr bedingt abgebildet werden können, einfach weil keine Infrastrukturen dafür vorhanden sind und der Aufbau solcher Infrastrukturen kein oder nur sehr wenig Anerkennung in den Fächern erbringt. Zitierfähige DOI sind ein wichtiger Schritt, um solche Forschung zu erlauben. Aber warum sich die Mühe machen, wenn es so viel einfacher ist, noch einen Walter Benjamin-Aufsatz zu schreiben. Anerkennungssysteme sind träge, so auch hier.

Zusammenfassend lässt sich zunächst nur der vorsichtige Schluss ziehen, dass die Transformation der geisteswissenschaftlichen Gegenstandsfelder in den nächsten Jahren nur sehr begrenzte Veränderungen der Infrastrukturen nach sich ziehen wird, eher in den kleinen Fächern, eher mit Blick auf die Erschließungsmöglichkeiten des kulturellen Erbes und eher in Fächern, die große Daten handhaben müssen wie die Linguistik oder Archäologie. Für sie macht die Digitalisierung des Gegenstandsfeldes einen Unterschied. Die anderen werden an der eher kanonischen Konstitution ihres Gegenstandes festhalten.

Für diejenigen, die genügend Neugierde und Forscherdrang spüren, ist eine Veränderung jedoch schon jetzt vielfach zu benennen. Sie wird zumeist und eher missverständlich unter Big

Data verhandelt. Genauer handelt es sich um eine Verschiebung weg vom Kanon, noch genauer um die Annäherung an die Grundgesamtheit, statistisch gesprochen. Während die Geisteswissenschaften – nicht alle, aber erhebliche Teile – stark am Kanon der wenigen, besonderen Werke ausgerichtet sind und von dort aus generalisieren, rückt vor allem mit dem digital born-Material nicht nur Kanon, nicht nur Stichproben, sondern die Grundgesamtheit dessen, was an Musik gehört oder an Literatur geschrieben wird, in Reichweite der Forschung. Damit ändert sich, was und wie generalisiert werden kann. Was der Roman des 19. Jahrhundert ist, das kann jetzt auch nahe an tatsächlich allen im 19. Jahrhundert gedruckten Romanen untersucht werden. Das ändert die Aussagenreichweite, die Exaktheit, Reliabilität, Replikation und ändert deren Validierung. Infrastrukturen müssen sich an diese veränderten Forschungsgegenstände anpassen, – das wäre zumindest die Forderung der Avantgarde der nicht etablierten Geisteswissenschaften.

3. Infrastrukturen und die Methoden der Geisteswissenschaften

Damit ist schon die zweite Veränderung, die der Methoden und der Methodik, implizit angesprochen. Der Kanon kann von Auge und Hand gelesen werden, andere Forschungsgegenstände des 21. Jahrhunderts brauchen Mal das Mikroskop oder mal das Teleskop, also ‚scalable reading‘⁷, Methoden, die den Abstand zum Text bzw. Objekt je nach Fragestellung wählen und dann entsprechende Instrumente nutzen. Solche Transformationen der Methodik lassen sich in einigen Fächern fraglos beobachten. Die Archäologie ist so selbstverständlich digitale Archäologie geworden, dass das Wort ‚digital‘ verschwunden ist, so wie sich das Fach Bioinformatik auflöst, weil es selbstverständlicher Teil der Biologie geworden ist. Computergestützte Explorationsverfahren wie LIDAR-Radartechnik, GIS-Verfahren für Ausgrabungen oder Wegeaufwandsberechnungen haben im Fach Archäologie so eminente Vorteile, weil sie eine Forschung ermöglichen, die ohne computergestützte Verfahren nicht möglich wäre, so dass diese digitalen Methoden zum Selbstverständnis eines Fachs dazu gehören. Ähnliches gilt für die Linguistik und Editorik. Auch hier lässt man den Zusatz ‚digital‘ weg, so selbstverständlich ist die Transformation der Methoden geworden. Dass die Linguistik mit *CLARIN* über eine eigene, disziplinäre Infrastruktur verfügt, ergibt sich aus der langen Tradition quantitativer und korpusbasierter Methoden, für die die Nutzung computergestützter Methoden nur die Fortsetzung ihrer bisherigen Forschungsmethodik mit anderen Mitteln ist. Dass das *Deutsche Archäologische Institut* über eigene digitale Infrastrukturen zur Handhabung seiner Daten aufgebaut hat, ist schiere Notwendigkeit für das Funktionieren dieser Forschung. Und Editionen etwa im Rahmen der *Union der deutschen Akademien* sind innerhalb einer Dekade ganz selbstverständlich digitale Editionen geworden.

Das gilt aber eben nur für einige Fächer, für andere nicht. In unserem Band *Die Rückkehr des Autors* von 1999 haben wir einen Beitrag des Pioniers der Autorenstylistik, John Burrows, abgedruckt. Unter allen Beiträgen blieb dieser Beitrag gänzlich unbeachtet, obwohl Burrows‘ Forschung gezeigt hat, dass Autorschaft nicht nur eine nachgetragene, soziale Konstruktion ist, sondern sich Autoren, Autorengruppen und Epochen anhand der Sprache unterscheiden lassen.⁸ Während in den literaturwissenschaftlichen Fächern noch die Meinung dominiert,

⁷ vgl. Mueller, Martin: <https://scalablereading.northwestern.edu/>.

⁸ vgl. Burrows, John. F.: *Computation into Criticism. A Study of Jane Austen’s Novels and an Experiment in Method*. Oxford: Clarendon Press 1987.

dass sich fiktionale und faktuale Texte nicht objektiv unterscheiden ließen, sondern auch dies nur eine soziale eingeübte gesellschaftliche Konvention sei, hat die computergestützte Forschung von Andrew Piper⁹ beispielhaft gezeigt, dass es dazu mindestens gegenläufige Befunde gibt. Solche Befunde können Fächer wie die Anglistik problemlos ignorieren und tun das auch. Die Infrastrukturen, die hinter dieser Forschung stehen, – das sind vor allem die Entwicklung von Skriptsprachen Python und R und die Software-Plattform [Github](#) –, werden der Informatik zugeschlagen und spielen für die Fächer der philosophischen Fakultät keine größere Rolle. Die Entwicklungen zu einer Semantisierung der kulturellen Daten durch rdf und [IIIF](#) sind ohne Bedeutung für die Kunstgeschichte. Auch hier gilt, dass es erst dann einen Druck auf die Fächer gibt, sich digitalen Infrastrukturen als Aufgabe zu stellen, wenn die Ergebnisse nicht mehr ignoriert und die Ergebnisse nur auf den computergestützten Wegen zu erreichen sind.

Das ist gegenwärtig besonders deutlich im Feld des Machine Learning. Während das in den meisten Fächern vor allem kulturkritische Reflexe auslöst, hat Machine Learning einen rasant anwachsenden Einfluss auf ganz unterschiedliche Fächer. In Fraktur gedruckte Quellen sind auf einmal maschinenlesbar, damit erhebliche Teile der deutschsprachigen Kulturtraditionen. Muster in den Farben und Formen in großen Bildersammlungen lassen sich ermitteln und Zusammenhänge aufzeigen, die mit dem bloßen Auge nicht zu erkennen sind. Handschriften-Bestände wie etwa das globalhistorisch so spannende *Archiv der Basler Mission und Schweizerischen Handelsgesellschaft* können jetzt erschlossen werden und eröffnen Forschungsmöglichkeiten für globalhistorische Fragestellungen. Doch verlangt Machine Learning den Philologen und Kunsthistorikern eine steile Lernkurve ab, die Fachvertreter mit zunehmendem Etablierungsgrad nicht auf sich nehmen, sondern bestenfalls etwa an Digitalisierungszentren in Bibliotheken zu delegieren versuchen. Noch, denn es ist nicht auszuschließen, dass die computationellen Lernverfahren bald schon über einzelne Anwendungen hinaus etwa die Zusammenhänge der antiken Mittelmeerkultur oder die Wanderung buddhistischer Ideenwelt in einer Geschichte der *longue durée* aufzeigen können, die mit herkömmlichen Methoden nicht aufgezeigt werden können. Dann dürften auch die etablierten Fachvertreter ihr Methoden-Portfolio überdenken.

Die Verknüpfung von linguistischen Datenreihen mit DNA-Daten ist gegenwärtig ein Beispiel, wie Innovationen mit Datenbanken und computergestützten Modellierungen die Trennung in Natur- und Geisteswissenschaften aufheben und radikal andere Methoden in die Geisteswissenschaften einbringen. Infrastrukturen bis hin zur Gründung eines eigenen *Max-Planck-Instituts für Menschheitsgeschichte* ermöglichen eine Forschung, die bis vor kurzem für unmöglich gehalten wurde, eben Menschheitsgeschichte zu betreiben, also die Zusammenhänge von paläoanthropologischen DNA-Daten mit der Verbreitung von Sprachen und deren Strukturen zu verstehen. Für Fächer wie die Geschichte oder die Sprachwissenschaft sind die Ergebnisse von Svante Pääbo, Simon Greenhill oder Patrick Geary mindestens eine Irritation und können nur soweit als Randphänomen außen vorgehalten werden, wie sie eher die Ur- und Frühgeschichte betreffen und nicht die Frage klären können, wie es zum Aufstieg des Nationalsozialismus gekommen ist. Ansonsten zeigen sie, dass ganz andere, hier naturwissenschaftliche Ansätze zum Methodenset der Geisteswissenschaften zählen können und eine solche Forschung völlig neue Infrastrukturen benötigt, eben auch DNA-Labore.

⁹ vgl. Piper, Andrew: Fictionality. In: Journal of Cultural Analytics (2016).

Anders gewendet gilt auch im Feld der Methodenentwicklung, dass aus neuen methodischen Möglichkeiten nicht zwingend Veränderungen der Fächer und ihrer Infrastrukturen folgen. Während die Linguistik über aufwändige Datenbanken zur Kartographierung der Sprachen der Welt verfügt, kennen so etwas die literaturwissenschaftlichen Fächer nicht, trotz deren emphatischer Rede von der Weltliteratur. Die Rekonstruktion der Namen der 6 Millionen ermordeten Juden braucht eine aufwändige Infrastruktur, [EHRI- European Holocaust Research Infrastructure](#), um Name Entities aus Deportationslisten und Steuerakten zusammenzuführen, andere Teile der Geschichtswissenschaft aber nicht. Die Wanderungen Ideen des Aristotelismus durch die verschiedenen Orden, Klöster und entstehenden Universitäten des Mittelalters nachzuverfolgen, ist jetzt möglich, auch die Korrelationen zu entdecken, die vielleicht mit der Geographie, Siedlungsdichte oder dem römischen Straßennetz zusammenhängen könnten. Für andere Bereiche der Philosophie- und Ideengeschichte werden solche Fragen nicht gestellt. Kurz: Neue Formen der Historiographie sind möglich, aber das sind bewusste Entscheidungen, keine Notwendigkeiten, die sich aus der Digitalisierung ergeben würden.

Ob man sich für solche, infrastrukturell aufwändigen Ansätze entscheidet, hängt an zwei Abwägungen. Das eine ist die Bereitschaft, neben den historischen-hermeneutischen Ansätzen auch solche der Formalisierung von Problemen zuzulassen. Praktisch alle computergestützten Ansätze verlangen, dass Probleme in möglichst exakte Kategorien modelliert werden müssen, um von Maschinen verarbeitet werden zu können. Die Vektorisierung und Faktorisierung, die Bestimmung von Variablen und der Ausschluss von konfundierenden Variablen, die Verlässlichkeit von Metadaten und deren Standardisierung sagen immer eines: dass die vertrauten historisch-hermeneutischen Methoden um solche der exakten Wissenschaften zu erweitern oder gar in diese exakten Begriffe und Konzepte zu überführen sind. Computer verstehen nichts und können auch nicht lesen oder beobachten. Ihre Leistungen hängen an der Genauigkeit und methodische Strenge, die Menschen durch das experimentelle Design in die Untersuchung einbringen. Computer sind nur so exakt wie die Wissenschaftler es sind, die die Forschung betreiben. Und dafür sind nun eher formale und eher quantitative Denkweisen notwendig, die in den Geisteswissenschaften nicht oder nur etwa in Fächern wie etwa der Linguistik eingeübt werden. Die Lernkurve ist entsprechend steil, wenn jemand Texte mit der Statistiksoftware R untersuchen möchte. Ohne Modellierung von Fragestellungen ist hier nichts zu erreichen und ich übertreibe nicht, wenn ich die Modellierung von Fragestellungen das Schibboleth der Geisteswissenschaft des 21. Jahrhunderts nenne. An ihm hängt wesentlich, ob die digitale Transformation der Geisteswissenschaften gelingt.

Eng damit zusammen hängt dann der für die Infrastruktur so wesentliche Punkt der Stellen. Wollen die Geisteswissenschaften mehr sein als nur ausstehende Beobachter, die warnend den Finger erheben angesichts der fortschreitenden Digitalisierung, wird es darauf ankommen, ob die Fächer Stellen zur computergestützten Modellierung ihrer Arbeit einrichten oder nicht. Denn nicht Datenbanktechniken oder Datenstandards, Rechnerkapazitäten oder Statistik sind so schwierig in den Fächern zu verankern, sondern die Umschichtung von Stellen und die Gewöhnung daran, dass nicht der Einzelwissenschaftler, sondern Teams die Geisteswissenschaften des 21. Jahrhunderts ausmachen. Man braucht eine Machine Learnerin und vielleicht noch einen Datenbank-Spezialisten, schreibt ein Paper zu Robert Musil mit vier anderen Kollegen, die nicht alle Germanistik als Fachhintergrund haben, sondern etwa auch Informatik, man braucht eine andere Infrastruktur der Stellen ähnlich der in den Naturwissenschaften. Damit werden auch Geisteswissenschaften teurer. Politikwissenschaften suchen *Twitter*-Ana-

lysten und die Archäologen brauchen für ihre Forschung Radarspezialisten. Der Erfolg so unterschiedlicher Projekte wie die Forschung zur sefardischen Polemikliteratur in der Frühen Neuzeit oder die Forschung zur Verbreitung der St. Gallener Hymnik im Europa des Mittelalters hängt daran, ob man genügend Fachleute für solche Vorhaben findet, die eben nicht mehr nur aus dem jeweils eigenen Fach wie Iberoromanistik oder Musikwissenschaft kommen. Dafür braucht es Stellen, genauer eine andere Stellenstruktur, die Technikerstellen nicht nur als Hilfswissenschaftler behandelt. Sie sind der Schlüssel für die Geisteswissenschaften des 21. Jahrhunderts. Die Modellierung von Forschungsfragen braucht andere Forschungsteams als es noch etwa Grimms Wörterbuch gebraucht hat. Die Geisteswissenschaften des 21. Jahrhunderts werden vielfach nicht mehr nur Geisteswissenschaften bleiben können, wenn sie den Anforderungen computergestützter Modellierung entsprechen wollen und das hängt letztlich an den Stellenprofilen. Entweder setzen die philosophischen Fakultäten auf die Einzelwissenschaftler oder auf Teams. Und das bedeutet, sie müssen die Fächer anders als bisher vernetzen, um arbeitsfähige Gruppen aufzubauen.

4. Infrastrukturen und Umwelten der Geisteswissenschaften

Veränderungen der Fächer durch die Digitalisierung werden nicht durch die Digitalisierung der Gegenstandsfelder noch durch die computationelle Modellierung und Formalisierung von Methoden jeweils alleine nahegelegt. Das gilt, drittens, auch für die Veränderungen der Akteure und Umwelten der Geisteswissenschaften. Es ist schon deutlich geworden, dass eine, wenn nicht die wesentliche Veränderung der digitalen Geisteswissenschaften darin liegt, dass sie eher von Forschergruppen umgesetzt werden. Das bedeutet nicht, dass es nicht auf Einzelarbeit ankäme und das einsame Nachdenken bedeutungslos geworden wäre. Auch im [CERN](#) kommt es darauf immer noch an, so auch in der Teamarbeit in den Geisteswissenschaften des 21. Jahrhunderts. Aber Korpusaufbau und Datenbanktechnik, Textmining und Archivkenntnisse, Machine Learning und Edition meinen verschiedene Kompetenzen und sind selten in einer Person vereint. Daher ist die Arbeit in Gruppen eine der potentiellen Veränderungen, die nicht wenig Irritation in den gerne am Genie-Paradigma angelehnten Geisteswissenschaften auslöst.

Mit der Veränderung der Akteure meine ich aber nicht nur die Betonung von Forschergruppen, sondern auch die andere Vernetzung, die erst neue Forschung ermöglicht. Für die Geisteswissenschaften sind besonders Bibliotheken nicht nur ein Ort, wo man Bücher findet und sie in Ruhe liest, sie sind ein Forschungspartner und das meint mehr als nur die Infrastruktur von Digitalisierungszentren, die zumeist an Bibliotheken angesiedelt sind. Es meint auch die gemeinsame Entwicklung von Datenstandards und Austauschformaten bzw. -technologien wie [IIIF](#), es meint das gemeinsame Schreiben von Forschungsanträgen, weil nur zusammen mit Bibliotheken ein Projekt vom Papyrus in einer Sammlung wie in der UB Basel bis zur Machine Learning-gestützten Untersuchung der Zusammenhänge dieser Dokumente gelingen kann.

Eine weitere Veränderung gibt es auch auf der Seite der Publikation der Ergebnisse. Ergebnisse sind nicht mehr nur der Aufsatz oder das Buch, sondern auch der Code oder die Datenbank. Publikationen auf [Github](#), Preprints auf [ArXiv.org](#), Datenbanken im Netz, sind Teil der Publikation, wenn Replikation und Reliabilität ernstgenommen werden sollen. Open Science ist daher die Kurzformel für einen Umbau, der vor allem in den etablierten Geisteswissenschaften auf

Ablehnung stößt. Für die nicht-etablierten Geisteswissenschaften sind hingegen solche Publikationsformen essentiell, weil die Leistungen auch in zitierfähigen Editionen, Datenbanken, rdf-Ontologien, Trainingssets oder Scripts bestehen können. Mehr noch entsteht ein Mehrwert daraus, wenn Daten nachgenutzt werden können, rekombiniert und mit anderen Befunden verglichen werden können. Für unser Projekt zur Geschichte des Tagebuchschreibens in Zeiten des Kriegs bringt ein einzelnes Tagebuch wenig; viel dagegen, wenn etwa Anne Franks Tagebuch mit anderen Tagebüchern verglichen werden kann, und damit erst Eigentümlichkeit des Schreibens Jugendlicher unter solchen Belastungen wie im Versteck in der Prinsengracht gesehen und Ähnlichkeiten bestimmt werden können. Die korpusbasierte Arbeit entsteht eben erst, wenn wir bereit sind die Daten zu teilen, ob das Editionen oder einfach nur Textfiles sind, und dann über Portale mit der Community nicht nur für uns behalten. In Fächern wie der Archäologie ist der Mehrwert längst evident, wenn Sammlungen wie die des *British Museum* ihre Bestände über das Netz teilen. Eine solche Sammlung wie die des *British Museums* über das Internet zu teilen ist Infrastruktur in dem elementaren Sinn, um den es mir hier geht.

Schließlich ändern sich auch Umwelten. Es ist keine Seltenheit mehr, dass Mathematiker sich mit geisteswissenschaftlichen Fragestellungen befassen, das haben schon Thomas de Morgan oder Corwin Mendenhall getan, als sie sich mit der Frage der Urheberschaft der Werke Shakespeares oder der Autorschaft der Briefe des Paulus befassten.¹⁰ Das war im 19. Jahrhundert und vielleicht vielfach nur Liebhaberei. Heute nutzen Mathematiker statistische Modelle zur Überprüfung der Pollyanna-Hypothese oder zur Untersuchung der Plotstruktur von Romanen. Ganz andere Ansätze wie die DNA-Analyse greifen in die Bereiche der Geschichtswissenschaft oder Linguistik über. Die Fakultätsgrenzen des 20. Jahrhunderts lösen sich gleich an mehreren Stellen auf. Taxonomische Modelle wie Splitstrees werden für die Analyse von Märchenmotiven genutzt, eine Fragestellung in den Geisteswissenschaften, die schon mehr als hundert Jahre alt ist und von der Finnischen Märchenschule im *Aarne-Thompson-Uther-Index* installiert wurde.¹¹ Jetzt sind solche taxonomischen Modelle der Biologie auf einmal mitten in den Geisteswissenschaften zu finden. Noch findet eine solche Forschung am Rande der Mehrheits-Geisteswissenschaften statt, etwa in der *Max-Planck-Gesellschaft*.

Die Liste der Beispiele lässt sich leicht verlängern. Deutlich wird dabei, dass die Geisteswissenschaften des 21. Jahrhundert mindestens nicht notwendig die des 20. Jahrhunderts sein müssen. Im Gegenteil sind andere Arbeitsformen längst erprobt, andere Formen der Zusammenarbeit vielfach umgesetzt, andere Publikationswege mehr als nur ein Versprechen und andere Grenzziehungen im Streit der Fakultäten möglich.

5. Geisteswissenschaften für das 21. Jahrhundert

Keine der Veränderungen alleine drängt eine Transformation der Geisteswissenschaften auf und legt eine Neujustierung der bisherigen Infrastrukturen dafür nahe. Veränderungen in den Wissenschaften sind eben von vielen Faktoren abhängig, vor allem in welchem Verhältnis Aufwand und Ertrag für einzelne Wissenschaften zueinander stehen. Zeit in das Erlernen statisti-

¹⁰ vgl. Gryzbek, Peter (Hg.): Contributions to the Science of Text and Language. Word Length Studies and Related Issues. Dordrecht: Springer 2006.

¹¹ vgl. Tehrani, Jamshid: The Phylogeny of Little Red Riding Hood. In: PLoSONE 8(11) (2013), e78871.

scher Methoden mit R zu investieren, wird nur der aufbringen, der sich einen im Fach angesehenen Gewinn davon versprechen kann. Auf [ArXiv.org](https://arxiv.org) wird nur die publizieren, die weiß, dass ihre Arbeit dort auch von Fachkollegen gelesen wird. Warum die Zeit mit Studenten und Kollegen verbringen, wenn ich am besten in Ruhe mein Buch schreiben kann, gerade wenn der Horizont der Fachwahrnehmung eher im Feuilleton denn im Fach zu finden ist. Den *Leibniz-Preis* vergibt die *DFG* in den Geisteswissenschaften selten an risikobehaftete Forschung. Es hat schlicht viele Vorteile, der etablierten Wissenschaftskonvention zu folgen. Gegebenenfalls sind die Sanktionen scharf, die Fächer gegenüber Minderheiten erlassen können. Es gibt also nur wenige Gründe, dass sich etwas ändern wird.

Aber vielleicht sind unter den Geisteswissenschaftlern und -wissenschaftlerinnen auch genügend, die allen Widerständen zum Trotz gerade deshalb Wissenschaften als Beruf gewählt haben, weil sie auch die Umwege oder wenig begangenen Wege einschlagen wollen und dabei das Risiko eingehen, dass dies auch Holzwege sein können. Wer die breiten Straßen verlässt, macht Entdeckungen und die brauchen dann andere Infrastrukturen. Wenn also Gegenstand des Fachs nicht nur der Kanon ist, sondern eher die Population, wenn die Methoden auch quantitative und experimentelle Ansätze umfassen, wenn andere Formen des Arbeitens, vor allem solche des Zusammenarbeitens und das auch über Fakultätsgrenzen hinweg und mit anderen Akteuren wie den Bibliotheken die eigene Neugierde anleiten, weil man voneinander lernen kann für die jeweils eigene Forschung, dann ändern sich auch Infrastrukturen.

Diese Infrastrukturen haben die folgenden Merkmale:

- Sie ermöglichen den flexiblen, je an die Fragestellung angepassten Aufbau von Korpora.
- Sie bieten Schnittstellen, zur Weiternutzung der Korpora.
- Sie erlauben den Export für verschiedene Miningverfahren, dem Rechnen auf virtuellen Maschinen etwa.
- Sie befördern die Kollaboration etwa durch Schnittstellen für Zusammenarbeit mit Bibliotheken, über [Gitlab](https://gitlab.com) oder Veröffentlichungen über [Github](https://github.com).
- Sie erlauben persistente Verlinkungen und stabile Versionierungen von Daten.
- Sie sichern sensible Daten.
- Und sie ermöglichen es, dass noch wenig erprobte Ansätze Pilotstudien fahren können.

Aber entscheidend ist das alles nicht. Entscheidend ist erstens, ob die philosophischen Fakultäten andere, eher an Forschergruppen orientierte Stellenstrukturen etablieren oder nicht. Zweitens ist wesentlich, ob die eher formalen Modellierungen von geisteswissenschaftlichen Fragestellungen nicht zuletzt auch in der Lehre der Fächer verankert werden.

Ich weiß nicht, ob in etwa fünf Jahren Infrastrukturen für die vermutlich auch dann noch kleine Gruppe von experimentierfreudigen Geisteswissenschaftlern und -wissenschaftlerinnen hinreichend etabliert sein werden, damit nicht jeder Schritt einer Grundsatzentscheidung gleich kommt. In einzelnen Fächern ist das ja schon gelungen, in meinem Fach der Literaturwissenschaft ist das noch die Zukunft. Und die hat bekanntlich den Vor- oder Nachteil unbekannt zu sein.

Francis Bacon hatte vor mehr als 400 Jahren verschiedene Idola unterschieden, die uns Menschen daran hindern, Erkenntnisse zu gewinnen, solche der eher unbewussten Täuschungen, der Dogmen, der Gewohnheit, aber auch des Verstandes. Es sind wohl letztere, die uns davon

abhalten, in unseren Philosophischen Fakultäten überhaupt nur die Diskussion zu beginnen, ob andere Infrastrukturen für die Geisteswissenschaften des 21. Jahrhunderts möglich sind. Mein Beitrag ist ein schwaches Plädoyer dafür diese Diskussion um die Geisteswissenschaften des 21. Jahrhunderts zu beginnen, bevor das stille Sterben der Geisteswissenschaft diese Fächer aufgelöst haben wird.

Literaturverzeichnis

Bacon, Francis: The Advancement of Learning. In: Francis Bacon. The Major Works. Edited with an Introduction and Notes by Brian Vickers. Oxford: Oxford University Press 1996 [1605], S. 120-299.

Burrows, John. F.: Computation into Criticism. A Study of Jane Austen's Novels and an Experiment in Method. Oxford: Clarendon Press 1987.

Diemer, Alwin: Art. Geisteswissenschaften. In: Historisches Wörterbuch der Philosophie 3 (1974), S. 211-215.

Gaukroger, Stephen: Francis Bacon and the Transformation of Early-Modern Philosophy. Cambridge: Cambridge University Press 2010.

Gryzbek, Peter (Hg.): Contributions to the Science of Text and Language. Word Length Studies and Related Issues. Dordrecht: Springer 2006.

Herrmann, Berenike / Lauer, Gerhard: Korpusliteraturwissenschaft. In: Osnabrücker Beiträge zur Sprachtheorie 92 (2018), S. 127-156.

Holliger, David. A.: The Humanities and the Dynamics of Inclusion since World War II. Baltimore: Johns Hopkins University Press 2006.

Lauer, Gerhard (Hg.): Wilhelm von Humboldt. Schriften zur Bildung. Stuttgart: Reclam 2017.

Mueller, Martin: <https://scalablereading.northwestern.edu/> (Zugriffsdatum: 05.12.2018).

Piper, Andrew: Fictionality. In: Journal of Cultural Analytics (2016), DOI: [10.31235/osf.io/93mdj](https://doi.org/10.31235/osf.io/93mdj) bzw. <http://culturalanalytics.org/2016/12/fictionality/> (Zugriffsdatum: 05.12.2018).

Rebenich, Stefan: Theodor Mommsen und Adolf Harnack. Wissenschaft und Politik des ausgehenden 19. Jahrhunderts. Berlin: de Gruyter 2012.

Tehrani, Jamshid: The Phylogeny of Little Red Riding Hood. In: PLoS ONE 8(11) (2013), e78871. <https://doi.org/10.1371/journal.pone.0078871> (Zugriffsdatum: 05.12.2018).