

„Kulturtechnik Digitalität“. Über den sich auflösenden Zusammenhang von Buch und Bibliothek und die Arbeit von Bibliotheken unter den Bedingungen digitaler Vernetzung

„Digitality as Cultural technology“. About the dissolving connection between book and library and the work of libraries under the conditions of digital networking

Sybille Krämer

Zusammenfassung

Ausgangspunkt ist die Diagnose, dass die „Lesbarkeit der Welt“ (Blumberg) sich transformiert in die „Maschinenlesbarkeit des Datenuniversums“. Der enge Zusammenhang von Buch und Bibliothek löst sich auf. Was bedeutet dies für die Arbeit der Bibliotheken? Die Bibliotheken werden Teil eines epistemologischen wie kulturtechnischen Umbruches; sie sind nicht einfach Dienstleister, sondern werden zu Akteuren, welche die Bedingungen der Möglichkeit einer Mündigkeit und Souveränität im Umgang mit Daten zu schaffen haben. Was das bedeutet, zeigt sich an den Aufgaben im Rahmen der Medienbereitstellung und -verwaltung sowie des Forschungsdatenmanagements – zwei Herausforderungen der Bibliotheksarbeit.

Schlagwörter: *Lesbarkeit; Digitalität; Maschinenlesbarkeit;*
Kulturtechnik der Verflachung; Blackboxing

Abstract

The starting point is the diagnosis that the “readability of the world” (Blumenberg) is transformed into the “machine readability of the data universe”. The close connection between book and library is dissolved. Libraries are becoming part of an epistemological and cultural-technical change; they are not simply service providers, but become actors who have to create the conditions for the possibility of maturity and sovereignty in dealing with data. What this means can be seen in the tasks involved in “media provision and administration” as well as “research data management” – two challenges of library work.

Keywords: *legibility; digitality; machine readability; cultural technique of flattening; black boxing*

„A NEW MEDIUM REQUIRES A NEW LITERACY.“
(Douglas Rushkoff, 2010)

Die folgenden Überlegungen gehen zurück auf meinen Keynote-Vortrag zur Eröffnung des Österreichischen Bibliothekartages 2019 in Graz am 19. September 2019. Dieser Essay wahrt die meinen mündlichen Vortrag leitende Thesenform.

1 Die Verbindung von Literalität und Buch löst sich auf im Übergang von der skripturalen zur digitalen Literalität.

Jahrhundertelang war die Institution der Bibliothek orientiert am Schriftstück, am Buch als einem Medium, das die Gestalt eines materialen, abgrenzbaren Objektes hat, welches gesammelt, aufgestellt und ausgehändigt werden kann. Es lohnt daran zu erinnern, dass *bibliothéke* (βιβλιοθήκη) griech. „Buch-Behälter“ bedeutet, Bibliotheken somit als räumliches Gefäß für Büchersammlungen thematisiert werden; Johann Amos Comenius hat 1658 dann dieses Wort als Terminus für „Bücherei“ erstmals verwendet.¹ Indem Bibliotheken in Buchform verkörperte Texte sammeln, ordnen, klassifizieren, katalogisieren und somit zugänglich machen, sind sie

¹ Friedrich Kluge, Alfred Götz: Etymologisches Wörterbuch der deutschen Sprache. Berlin: de Gruyter 1953, S. 109.

zutiefst verwoben mit den Kulturtechniken der Literalität. „Literalität“ wird hier verstanden als die – mehr oder weniger verbreitete – Befähigung zu schreiben und zu lesen, welche die Epoche der handschriftlichen Manuskriptkultur, aber auch das gedruckte Buch der Gutenberg-Ära umfasst. Keine Frage also, dass das Medium der Schrift und das *Schriftstück* als sinnlich-anschauliches, handgreifliches Ding mit dem Bibliotheks-wesen zutiefst verbunden ist. Doch in der Digitalisierung löst sich die enge Liaison von Literalität, Buch und Bibliothek auf. Die traditionelle Fokus-sierung auf Schriftwerke, welche auch das Raumkonzept von Bibliotheken nachhaltig bestimmt hat, erodiert. Der physische Raum, verstanden als Sammelstelle geordneter Bücher, macht dem virtuellen Raum vernetzter Datenkorpora Platz. Dass dies allerdings nicht bedeutet, dass Bibliotheken als reale Räume obsolet werden, unterstreicht Ben Kaden.² Worauf es uns hier ankommt ist, dass die Digitalisierung eine Metamorphose der Litera-lität erfordert und evoziert. Es geht um die Transformation von einer „skripturalen“ zu einer „digitalen Literalität“. Bibliotheken stehen in die-sem Wandel an vorderer Front; sie sind Vorreiter und Wegbereiter einer „Digitalität als Kulturtechnik“.

- 2 Die Digitalisierung operationalisiert Zeichenwelten zu Daten-sammlungen, welche Maschinen „lesen“ und – sofern erwünscht – automatisch verarbeiten können. Wir werden zu Zeugen eines Überganges von der „Lesbarkeit der Welt“ (Blumenberg) zur „Maschinenlesbarkeit vernetzter Daten“.

Das Buch ist immer schon mehr als eine Chiffre für Texte eingespannt zwischen Buchdeckeln; es kondensierte zur Metapher für unser Welt- und Selbstverhältnis. Ob für den Kirchenvater Augustinus Gott neben der Bibel auch das *Buch der Natur* verfasst hat,³ ob für Nikolaus von Kues „die Dinge“ selbst zu „Bücher(n) der Sinne“ werden,⁴ ob für Galilei das Buch

2 Aus: Notizen zur Bibliothekswissenschaft. Teil 1 und 2. In: LIBREAS. Library Ideas 18 (2011), S. 6.

3 Herbert Maria Nobis: Buch der Natur. In: Historisches Wörterbuch der Philoso- phie. Band 1. Darmstadt: Wiss. Buchgesellschaft 1971, S. 957–960, hier S. 958.

4 Nikolaus von Kues: Predigten von 1430–1441. In: Schriften des Nikolaus von Kues. Hg. von Ernst Hoffmann. Bd. 2. Heidelberg: F.H. Kerle Verlag 1952, S. 454, zit. nach Nobis: Buch der Natur (Anm. 3), S. 959.

der Natur in mathematischer, insbesondere geometrischer Zeichensprache geschrieben ist: Die Natur wird sinnbildlich zur Chiffrenschrift, die zu rezipieren uns aufgegeben ist. Im 20. Jahrhundert schließlich werden auch menschliche Handlungen und soziale Ereignisse, mithin die Kultur selbst als textförmig betrachtet; die Existenz unterschiedlicher Lesarten von Texten kann damit zur Metapher werden für ein „nichtessentialistisches Kulturverständnis“⁵. Hans Blumenberg fasst mit seiner *Lesbarkeit der Welt*⁶ Dimensionen dieser Entwicklung in der Fülle ihrer Verästelungen kongenial zusammen und skizziert eine Ideengeschichte, in welcher Text und Buch zur Grundverfassung dessen avancieren, was überhaupt als erkennbar und verstehbar projiziert wird. Mit seiner Leitidee einer Lesbarkeit der Welt zeigt Blumenberg sich als „Weltbibliothekar“⁷.

Die Bedingungen der Lesbarkeit ändern sich gegenwärtig. Die Begriffe „Schriftstück – Buch – Dokument – Daten – vernetzte Datensammlungen“ benennen eine diagnostizierbare Tendenz;⁸ ihr Telos besteht darin, dass Schrift, Bücher und Dokumente von Menschen nicht nur gemacht, sondern auch rezipiert und verarbeitet werden, während „Daten“ und „vernetzte Datenkorpora“ eine Zeichenebene markieren, auf der es die Maschinen sind, welche Zeichen generieren, lesen und verarbeiten (können). Mit der Digitalisierung erlangen Teile der kulturellen Memoria, also das, was als Text, Bild, Film oder Musik etc. zum Bestand des symbolischen Universums gehört, eine Verfassung, die von Menschen kaum mehr, von Maschinen dagegen umso präziser entzifferbar ist.⁹ Schon der genetische Code, die Buchstabenschrift des Lebens, ist nichts mehr, was Men-

5 Doris Bachmann-Medick: Einleitung. In: Kultur als Text. Hg. von Doris Bachmann-Medick. Frankfurt a. M.: Fischer 1996, S. 7–66, hier S. 13.

6 Hans Blumenberg: Die Lesbarkeit der Welt. Frankfurt a. M.: Suhrkamp 1986.

7 Ausdruck von Caroline Neubaur: Das Leben der Bilder und Zeichen. In: DIE ZEIT 41 (1982) <https://www.zeit.de/1982/41/das-leben-der-bilder-und-zeichen> (abgerufen am 09.09.2019).

8 Diese Tendenz ist instruktiv beschrieben in Sarah Dudek: Die Zukunft der Buchstaben in der alphanumerischen Gesellschaft. Text und Dokument unter digitalen Bedingungen. In: BFB 36 (Juli 2012), S. 189–199.

9 Hinzu kommt der Umstand, dass – insbesondere in Smartphone-Kulturen – Weltereignisse vielfach synchron aufgezeichnet werden, die Welt also beständig in Datenstrukturen überführt bzw. „datifiziert“ wird.

schenaugen intuitiv wahrnehmen, doch Maschinen bestens analysieren können.

Allerdings ist diese maschinelle Operationalisierbarkeit der Zeichen nicht der Einbruch eines radikal Neuen, sondern in den Formalisierungen im alphanumerischen Zeichenraum als Keimform längst schon angelegt. Bereits die Verschriftung des Rechnens kraft der Einführung des – von indischen Mathematikern entwickelten – dezimalen Positionssystems in Europa durch den islamischen Gelehrten Al Chwarizmi,¹⁰ bereitete den „symbolischen Maschinen“, also dem Computer in uns,¹¹ den Boden. Al Chwarizmis Eigename – latinisiert als „Algorithmus“ – wird zum Terminus für eine Regel des Zahlenrechnens, die – das erkennt der Philosoph Gottfried Wilhelm Leibniz scharfsichtig – aufzufassen ist als eine Regel rein schematischer, schriftlicher Zeichenmanipulation. Mit der Kulturtechnik des schriftlichen Rechnens ist eine Textform entstanden, die nicht nur eine wohldefinierte Domäne von Referenzgegenständen – etwa Zahlen – *darstellt*, sondern die auch Problemlösungen, diese Domäne betreffend, *herstellt* und nicht zuletzt konturiert, was als Referenzgegenstand überhaupt zählt und gilt. So bekommt die Null, die Quantifizierung von Leere, also die Ablösung des Konzeptes „Zahl“ vom Zählbaren, ein Heimatrecht im Zahlenuniversum. Und genau dies meint „Operationalisierung der Zeichen“: Ein Zeichensystem ist Medium und Werkzeug in einem: Was das Medium darstellt, wird von ihm als Werkzeug zugleich auch hergestellt. So werden Zeichen nicht nur Repräsentationen dessen, was sie bezeichnen, sondern werden zum Ausgangspunkt für *Aktionen* mit dem Bezeichneten. Der Link auf den Bildflächen vernetzter Interfaces ist ein dafür charakteristisches Gegenwartsphänomen: Eine Schrift, die – einmal aktiviert – sich selbst bewegt und etwas bewirkt. Der Link allerdings ist immer noch eine nutzerzugewandte Verbindung aus menschenlesbarer Schrift und maschinenlesbarem Code, bleibt ein menschenrezipierbarer Wink für eine Aktivität. Doch die Transformation von Texten in von Maschinen operierbaren Zeichenstrukturen geht über die Erfindung einer

10 Kurt Vogel: Mohammed ibn Musa Alchwarizmi's Algorismus. Das früheste Lehrbuch zum Rechnen mit indischen Ziffern. Aalen: Otto Zeller 1963.

11 Sybille Krämer: Symbolische Maschinen. Darmstadt: Wiss. Buchgesellschaft 1988, S. 2–4.

sich selbst bewegenden, auto-operativen Schrift¹² weit hinaus; sie weist in Richtung eines Paradigmenwechsels vom „Lesen von Texten“ zum „Rechnen mit Daten“. Allerdings darf dies nicht als eine Ablösung und Ersetzung missverstanden werden, so wenig wie die Mündlichkeit durch die Erfindung und Verbreitung der Schrift oder durch das gedruckte Buch ersetzt wurde. Ein neues Medium löst die alten Medien nicht einfach ab, sondern konfiguriert deren kulturellen Rollen neu.

3 Die maschinelle Operationalisierbarkeit der Zeichen zehrt von der „Kulturtechnik der Verflachung“, von einer artifiziellen Flächigkeit, die kein kognitives Hemmnis, vielmehr eine Produktivkraft ist für mannigfaltige Formen der Kreativität in Wissenschaft, Kunst und Technik.

Wir sind vertraut mit einer Rhetorik, die den Tiefgang im Denken nobilitiert und – was an der Oberfläche bleibt – diskreditiert. Doch die Assoziation des menschlichen Geistes mit dem, was in die Tiefe geht und somit tiefgründig ist, steht in einem bemerkenswerten Spannungsverhältnis zu der schlichten Beobachtung, dass Kulturen kaum denkbar sind, in denen der Gebrauch markierter, ornamentaler, bebildeter und schließlich beschrifteter Oberflächen nicht augenfällig wird. Von der Höhlenmalerei über die Hauttätowierung bis eben zu den Fernsehbildschirmen, Computerscreens und elektronischen Interfaces zieht sich das Band „artifizieller Flächigkeit“¹³. Zwar gibt es empirisch keine Flächen; doch Oberflächen von immer noch voluminösen Körpern werden durch Praktiken der Gravur, Bebilderung, Beschriftung so behandelt, als ob sie *keine* Tiefe hätten. Alles, was in deren Rezeption zählt, ist von oben sichtbar. Für moderne Civilisationen ist Anschauung und Gebrauch von Oberflächenfiguren unverzichtbar. Wissenschaften, Künste, Architektur und Technik verdanken ihre Komplexität, Realisierbarkeit und Tradierbarkeit den vielfältigen Möglichkeiten, anhand von Texten, Bildern, Karten, Katalogen,

12 Gernot Grube: Autooperative Schrift – und eine Kritik der Hypertexttheorie. In: Schrift. Kulturtechnik zwischen Auge, Hand und Maschine. Hg. von Gernot Grube, Werner Kogge und Sybille Krämer. München: Fin 2005, S. 81–114.

13 Sybille Krämer: Figuration, Anschauung, Erkenntnis. Umrisse einer Diagrammatologie. Frankfurt a. M.: Suhrkamp 2016, S. 11–13.

Entwurfszeichnungen, Diagrammen, Graphen, etwas Unsichtbares oder ein noch nicht Existierendes auf der Fläche nicht nur wahrnehmbar, sondern auch manipulierbar und explorierbar zu machen. Im Wechselspiel von Auge, Hand und Hirn wird das Papier zum Entwicklungslabor von Gedanken, zum Spielfeld künstlerischer Komposition oder zum Experimentierraum für Ingenieursleistungen. In der ubiquitären Smartphone-Nutzung erlebt der Einsatz kleiner mobiler, zumeist rechteckiger Flächen seinen Kulminationspunkt.

Schon die mit der artifiziellen Flächigkeit verbundene medientechnische Amputation des „Dahinter“ und „Darunter“, also jener Domäne, die – gemessen an unseren Körperachsen vorne/hinten, rechts/links, oben/unten – eine unüberschaubare Region abgibt, signalisiert Überblick und Kontrolle und inauguriert damit ein Transparenzversprechen: Kulturelles Wissen wird übertragbar, Gedankenfolgen werden nachvollziehbar und kritisierbar, mathematische Regeln werden durch die Formeln symbolischer Algebra lehr- und lernbar, Schlagwortregister machen Wörter in Büchern auffindbar, technische Konstruktionsprinzipien und betriebswirtschaftliche Organisationsprinzipien werden einsehbar und verstehbar. Um nur ein für dieses Transparenzversprechen plastisches Beispiel zu geben: Indem von der artifiziellen Flächigkeit zehrende Signatursysteme und Zettelkataloge Büchern Adressen verleihen, bleiben Bibliotheken keine schwer bewältigbaren Bücher-Labyrinthe, sondern werden zu Anlaufstellen für Dokumentensuche.

- 4 Die Ambivalenz des Digitalen zeigt sich in der Janusköpfigkeit und Zwiespältigkeit, welche das Phänomen artifizieller Flächigkeit unter den Bedingungen elektronischer Vernetzung annimmt im Widerstreit zwischen Transparenz und Zugänglichkeit von Wissen einerseits und Intransparenz, „Blackboxing“ und Datenmissbräuchen andererseits.

Jedwede Mensch-Maschinen-Interaktion, jeglicher Medieneinsatz ist von Zwiespältigkeiten, von den stets gegebenen Optionen eines Fortschreitens wie des Rückschreitens, der Entstehung von Problemlösungen wie neuer Problemlagen gezeichnet. Es kommt auf die Einbettung in historisch-soziale Praktiken an, die bestimmen und verstehbar machen, was Technik- und Mediennutzung konkret erbringen oder wodurch sie bedroht werden.

Gleichwohl gibt es – der Annahme einer schlichten Zweck-Mittel-Relation von Instrumenten und Verfahren zum Trotze – eine Strukturlogik und Eigensinnigkeit von Werkzeugen und Medien, deren Materialität und Machart immer auch vorgeben, wozu, auf welche Weise und wie sie zu gebrauchen sind. Vor diesem Horizont kann gefragt werden: Was geschieht, wenn die beschriftete Papierfläche sich zum elektronischen Interface fortbildet? Vor dem Screen schreiben und lesen Nutzer*innen in vertraut tastendrückender oder bildschirmwischender Manier; doch hinter dem Screen breitet sich rhizomartig ein Universum miteinander interagierender Protokolle, Programme und Maschinen aus, welches von denen, die vor dem Bildschirm arbeiten und kommunizieren, nicht mehr überschaubar, geschweige denn kontrollierbar und beherrschbar ist. *Die medientechnisch einst eliminierte Tiefenstruktur kehrt zurück.* Die operative Bildlichkeit des Interface schlägt um in die apparative Unbildlichkeit von „Subfaces“ – ein Begriff, den Frieder Nake vor Jahren einführte.¹⁴ Die Digitalität zeigt ihren Januskopf.

5 Die Maschinenprozessierbarkeit von Datenkorpora setzt voraus, implizite – also latente bleibende – Dimensionen von Text- und Bildoberflächen als Markierung und Textur explizit und manifest zu machen.

Menschliche Zeichenrezeption wird gesteuert von kollektivem Hintergrundwissen. Auf den beschrifteten und bebilderten Oberflächen werden Unterschiede wahrnehmbar, die semantisch bedeutsam sind und die Interpretation steuern. Einen Text lesend wissen wir, dass Überschriften und Fließtext zu unterscheiden sind, dass bestimmte Ausdrücke als Eigennamen gelten, dass die schriftbildlichen Markierungen von Groß- und Kleinschreibung grammatisch-semantische Vorgaben verkörpern. Alles das geschieht beim Lesen unmerklich und intuitiv. Doch die Verwandlung des Textdokumentes in ein maschinenprozessierbares Datenkonvolut, sofern diese mehr anzielt, als nur eine Buchstabenfolge zu codieren, setzt voraus, solches Hintergrundwissen (aber auch Kontextinforma-

¹⁴ Frieder Nake: Surface, Interface, Subface. Three cases of interaction and one concept. In: Paradoxes of Interactivity. Hg. von Uwe Seifert, Jin Hyun Kim und Anthony Moore. Bielefeld: transcript 2008, S. 92–109.

tionen) mithilfe von Auszeichnungssprachen in einer Oberflächenkonfiguration darzustellen, für die Maschine – metaphorisch ausgedrückt – „transparent“ zu machen. Implizites ist in Explizites, latent Bleibendes in Manifestes zu überführen. Übrigens bildet die Explikation von Implizitem einen Kern aller Formalisierung: Wenn das Kommutativgesetz der Addition symbolisch als „ $a+b=b+a$ “ angeschrieben wird, tritt mit solcher Formel ein *Muster* hervor, welches buchstäblich zeigt, dass eine Vielzahl unterschiedlicher Gleichungen Exemplare dieses allgemeingültigen Musters sind und ergo mit dieser Transformationsregel bearbeitet werden können. Auf das Phänomen des Musters kommt es hier an; denn was Maschinen überaus gut können, ist Muster zu dechiffrieren. Computer sind *forensische Maschinen*, denen es gelingt, etwas für Menschenaugen Verborgenes – unter gewissen Bedingungen – sichtbar und explizit zu machen. Sie fungieren als Mikroskop, wie als Teleskop symbolischer Strukturen. Computer können sowohl im Kleinen, am individuellen Objekt Spuren dechiffrieren, die nicht der Region menschlicher Wahrnehmbarkeit angehören; und sie können auch im Großen in unübersehbar umfangreichen Datenmengen für Menschen nicht erfassbare Muster zutage fördern. Konkret bedeutet das „Explizitmachen“ eines Impliziten: Textdokumente werden dadurch maschinenlesbar, dass sie mithilfe standardisierter Auszeichnungssprachen zu codieren sind. Keine Frage, dass die Arbeit an der Standardisierung solcher Auszeichnungssysteme (SGML; TEI)¹⁵ von größter Relevanz ist, wenn Datensammlungen mit neuen Betriebssystemen und Anwendersoftware nicht untergehen sollen. Und die Frage, welche Merkmale eines menschenrezipierbaren Textes überhaupt ausgezeichnet, also codiert werden sollen, ist keineswegs trivial: Daher ist menschliche Entscheidungskompetenz und Urteilskraft auf nahezu jeder Ebene der Verwandlung von menschenrezipierbaren Zeichen in maschinenlesbare Daten gefordert. So hat sich etwa das Editionsprojekt der Schriften und Notate Ludwig Wittgensteins im norwegischen Bergen gegen TEI/SGML als Auszeichnungssys-

15 Die *Standard Generalized Markup Language* (SGML) ist ein von Betriebssystem und Programmen unabhängiger Standard zur Textcodierung, während die *Text Encoding Initiative* (TEI) auf eine Initiative speziell von Philologen zurückgeht und es Textwissenschaftlern ermöglicht, sich von den vorgegebenen Formaten der Textspeicherung (z. B. MS-Word, HTML) freizumachen.

teme entschieden.¹⁶ Doch das sind nur Beispiele für den Sachverhalt, dass gerade die Operationalisierung zur Maschinenlesbarkeit nur um den Preis vielfältiger Anstrengungen menschlicher Subjekte – aber auch kultureller Institutionen wie der Fachinformationsdienste und der Bibliotheken – zu erringen ist. Wenn die Übertragung in Maschinencode nach diesen Maßgaben erfolgt und zum Grundprinzip der Webrepräsentationen wird, kann sich das „web of documents“ in ein „web of linked data“ (World Wide Web Consortium¹⁷) verwandeln. Ein Zustand wäre dann erreicht, bei dem (im Prinzip) alle Worte in Texten miteinander vernetzt sind und also extrahiert und zitiert, indexikalisiert und analysiert, annotiert und neu angeordnet werden und dabei – jedenfalls in der *Vision* von Kevin Kelly – tiefer mit dem Gesamt unserer Kultur in Verbindung treten könnten, als es jemals zuvor möglich war.¹⁸ Prosaischer und auf das Buch als vernetztes Textdokument bezogen: Suchtechniken kompensieren im „web of linked data“ die „schwindende Autorität“ des Buches, indem sie aus einzelnen Büchern das Eingangsportal in eine *Universalbibliothek* machen.¹⁹ Allerdings sind solche umfangreichen Mengen von Datenkorpora ihrerseits nur noch mit algorithmischen Suchverfahren zugänglich zu machen: „What do you do with a million books“, fragt 2006 Gregory Crane.²⁰ Aufschlussreich ist bei dieser Datifizierung des kulturellen Gedächtnisses in Form der Linked Open Data – wie Sarah Dudek fest-

16 Vgl. die Selbstbeschreibung der Bergen Electronic Edition von Wittgensteins Nachlass: http://wab.uib.no/wab_BEE.page (abgerufen am 11.12.2019).

17 Das World Wide Web Consortium (W3C) ist ein internationales Standardisierungsgremium, das 1994 gegründet wurde.

18 „Turning inked letters into electronic dots that can be read on a screen is simply the first essential step in creating this new library. The real magic will come in the second act, as each word in each book is cross-linked, clustered, cited, extracted, indexed, analyzed, annotated, remixed, reassembled and woven deeper into the culture than ever before. In the new world of books, every bit informs another; every page reads all the other pages.“ Kevin Kelly: „Scan This Book!“. In: The New York Times, 14.05.2006, <http://www.nytimes.com/2006/05/14/magazine/14publishing.html> (abgerufen am 11.12.2019), zit. n. Dudek, Die Zukunft der Buchstaben (Anm. 8), S. 198.

19 David Shields: Reality Hunger: Ein Manifest. München: Beck 2011.

20 Gregory Crane: What Do You Do with a Million Books? In: [D-Lib Magazine](https://www.researchgate.net/publication/28106851) 12 (March 2006), 3, online verfügbar unter: <https://www.researchgate.net/publication/28106851> (abgerufen am 11.12.2019).

stellt –, dass sich ein Brennpunkt bibliothekarischer Arbeit, der in der Titelaufnahme besteht, die auf ein einzelnes Dokument in Form eines singulären Exemplars verweist, nun auf die Beschreibung des semantischen *Inhalts* eines Textes übergeht, soweit dieser durch Daten beschreibbar ist.²¹ Verschiebt sich der Fokus bibliothekarischer Arbeit von der Materialität des Buches als Ding zur Virtualität der Datenstruktur semantischer – also von der Buchform abgelöster – Einheiten? Doch bei all dem ist nicht zu vergessen: Auch wenn „Big Data“ in den Textwissenschaften erstmals die Möglichkeit eröffnet, gerade nichtkanonische Texte in umfänglicher Weise mit einzubeziehen und zu analysieren, so erzeugt die Verdatung von Texten zugleich bemerkenswerte Ausschlüsse: Was nicht in ein Digitalisat verwandelt wurde, existiert dann nicht (mehr).

6 Das für die Digitalisierung von Forschungsmethoden erforderliche Explizitmachen der zumeist unthematisiert bleibenden Dimensionen geisteswissenschaftlicher Arbeit kann eine kritische Selbstreflexion und -korrektur der geisteswissenschaftlichen Verabsolutierung von Interpretation und Hermeneutik stimulieren.

Mit der Digitalisierung der Bibliotheksbestände seit den 1990er-Jahren haben die Bibliotheken überhaupt erst die Voraussetzungen für die Entstehung der Digital Humanities geschaffen. Das Verhältnis zwischen hermeneutisch orientierten Geisteswissenschaften und Digital Humanities ist allerdings angespannt, denn nicht wenige traditionelle Geisteswissenschaftler sehen in der umfassenden, gerade auch finanziellen Förderung der digitalen Neuankömmlinge eine stromlinienförmige Szientifizierung der interpretierenden Geisteswissenschaften. Alle Geisteswissenschaftler sind heute von der *Digitalität als Kulturtechnik* geprägt: Sie nutzen E-Mails, digitale Ressourcen, kooperieren über Files etc. Von diesen Dimensionen der Bildschirmarbeit als Professionalisierungsgrundlage geisteswissenschaftlicher Tätigkeit sind die „Digitalen Geisteswissenschaften“ bzw. „Digital Humanities“ zu unterscheiden. Denn diese sind die Sammelnamen für Sparten geistes- und kulturwissenschaftlicher Arbeit, deren Forschungsgegenstände nicht einfach nur digitalisiert sind – sei es als Retro-Digitalisierung oder als digital-born-materials –, sondern deren For-

21 Dudek, Die Zukunft der Buchstaben (Anm. 8), S. 194.

schungsfragen nur noch mithilfe computergenerierter, komputationaler Methoden zu bearbeiten sind, unter transdisziplinärer Einbeziehung von angewandter Informatik und statistischer Mathematik. Genau genommen sind das Fragen, deren Beantwortung nicht nur den Computer als Werkzeug und Gerät, sondern eigens für das Forschungsfeld entwickelte quantifizierende Methoden erfordern. Der Methodenstreit zwischen qualitativer, interpretierender und quantifizierender, statistischer Forschung, der die Sozialwissenschaft im 20. Jahrhundert beschäftigte²², scheint nun auch die Geisteswissenschaften erreicht zu haben. Allerdings gibt es niemals – auch nicht in den Naturwissenschaften, vielleicht noch nicht einmal in der Mathematik – *rein* quantifizierende Methoden, sondern auf nahezu allen Ebenen wissenschaftlicher Arbeit, ganz besonders in der Hypothesenbildung und der Deutung der Ergebnisse, ist Interpretation im Spiel. Die Frage zu beantworten, unter welchen Bedingungen, auf welche Weise und warum überhaupt der Einsatz rechnerischer Verfahren geisteswissenschaftlich zulässig und sinnvoll ist, ist komplex und wird erst anhand guter Beispiele für anschließbare Forschungsergebnisse zu beantworten sein. Gegenwärtig scheint sich allerdings eine gewisse Ernüchterung abzuzeichnen, ob der immense Aufwand, den fruchtbare Einsätze digitaler Forschungsverfahren verlangen, in einem vertretbaren Verhältnis steht zu den damit errungenen Resultaten.

Hier ist nicht der Ort zu diesem Methodenstreit Stellung zu nehmen, sondern auf etwas zu verweisen, das in der Hitze des Gefechts gerne übersehen wird. Wenn computergenerierte Forschungsmethoden, wie die Entwicklung von Auszeichnungssprachen und von Verfahren des Text

22 Während der Streit um Quantifizierung innerhalb der Sozialwissenschaft in den 60er-Jahren des 20. Jahrhunderts im Positivismusstreit zwischen Popper und Adorno einen Höhepunkt fand, hat seit den 80er-Jahren eine Rückbesinnung auf die notwendige Interdependenz von quantifizierenden und qualitativen Aspekten eingesetzt, die schließlich 2003 zur Gründung der Sektion *Qualitative Methoden* in der Deutschen Gesellschaft für Soziologie führte. Von der Gegenwart sozialwissenschaftlicher Forschung her gesehen ist dies nicht mehr als eine historische Reminiscenz. Auch für die Geisteswissenschaften scheint diese Integration komputationaler Methoden in den Methodenkanon geboten, ohne dabei zu vergessen, dass die quantifizierbaren und zugleich fruchtbaren Fragestellungen nur partielle und oft schmale Felder im Gesamtspektrum geisteswissenschaftlicher Arbeit besetzen.

Mining²³ oder Topic Modeling²⁴ eingesetzt werden, kommen damit Tätigkeiten geisteswissenschaftlicher Arbeit in den Blick, die im Selbstverständnis traditioneller Geisteswissenschaft gerne übersehen, wenn nicht gar verdrängt werden. Wir dürfen nicht vergessen, dass noch vor aller Interpretation auf der geisteswissenschaftlichen Agenda immer schon Tätigkeiten stehen wie das Suchen, Sammeln, Ordnen, Annotieren, Vergleichen und Rekonstruieren von Dokumenten und Objekten. Die Geisteswissenschaften handeln eben nicht einfach vom „Geist“, sondern von dessen kulturellen und zeichenhaften Verkörperungen, deren raumzeitlich situiertes Auftreten – sofern das forscherisch von Belang ist – immer auch vermessen werden kann. Die Frage, wie diese geisteswissenschaftlichen Basisarbeiten, die sogenannten „scholarly primitives“, die alle Geisteswissenschaften von der Archäologie über die Kunst- und Musikwissenschaften bis zu den Philologien vereinigen, computergestützt umsetzbar sind, wurde nicht zufällig zu einer Springquelle des „Humanities Computing“.²⁵ Die Entdeckung der Materialität der Kommunikation und der physischen Gegenständlichkeit kultureller Artefakte als unabdingbares Fundament geisteswissenschaftlichen Tuns im letzten Drittel des vergangenen Jahrhunderts fällt nicht zufällig zusammen mit der „Datifizie-

-
- 23 Beim Text Mining, verwandt dem Data Mining, werden große Textkonvolute auf Muster hin durchsucht, um einander ähnliche „Datensätze“ aufzuspüren und dann nach Kriterien zu sortieren und zu klassifizieren, um auf inhaltliche Zusammenhänge und Differenzen zwischen Texten zu schließen. Diese Muster waren den Texten implizit und werden jetzt durch die algorithmische Bearbeitung explizierbar gemacht.
 - 24 Das Topic Modelling bezieht sich auf eine Gruppe quantitativer Analysen von Texten, bei denen mithilfe statistischer Analysen des Wortinventars semantische Rückschlüsse auf Themata und Topoi eines Textes getroffen werden. Es geht also um einen Zugang, der sich auf den Bedeutungsgehalt von Texten bezieht und doch – mithilfe hochkomplexer Rechenverfahren (Probabilistic Latent Semantik Analysis, Latent Dirichlet Allocation) – dazu statistische Methoden einsetzt. Sina Bock, Keli Du, Michael Huber, Stefan Pernes, Steffen Pielström: Der Einsatz quantitativer Textanalyse in den Geisteswissenschaften. Göttingen: Dariah-DE 2016 (= Dariah – DE Working Papers Nr. 18).
 - 25 John Unsworth: Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This. Conference Paper, January 2000, online verfügbar unter: <https://www.researchgate.net/publication/205836758> (abgerufen am 12.12.2019).

rung“, also der Virtualisierung eben dieser Materialien. Vor diesem Horizont gesehen besteht der „Stachel der Digitalisierung“²⁶ gerade darin, dass der Zwang zum digitalen Explizieren dessen, was Geisteswissenschaften zwar immer schon tun, aber bisher wenig thematisierten, zu einer Revision am geisteswissenschaftlichen Selbstbild beitragen kann. Die Mühen der Übertragung von Text, Bild und Musik in maschinenlesbare Korpora und der Entwicklung angemessener Algorithmen ihrer maschinellen Bearbeitung dokumentiert, dass die hermeneutische Verabsolutierung von Interpretation als Leitkategorie zu kurz greift, weil der Umgang mit Materialien und Objekten immer schon die Bedingung der Möglichkeit geisteswissenschaftlicher Arbeit abgab, diese präformierte und durchdrang. Die Arbeit von Bibliotheken (aber auch von Archiven und Museen) verkörpert geradezu paradigmatisch diese „materialbezogenen“ Dimensionen geisteswissenschaftlichen Tuns. Diese nimmt gegenwärtig die Form der Kuratierung virtueller, vernetzter Datensammlungen an.

- 7 Das „Forschungsdatenmanagement“ sowie die „Medienbereitstellung und -verwaltung“ bilden zwei bibliothekarische Zukunftsausgaben. Das Grundproblem beider ist die Standardisierung, die Sicherung von Langfristigkeit und Nachnutzbarkeit gerade angesichts der fehlenden Einheitlichkeit bzw. projektförmigen Zersplitterung datentechnologischer Substrukturen.

Die Rolle, welche die Formalisierung für die Konsolidierung und Dynamik neuzeitlicher Wissenschaften bedeutete, spielt gegenwärtig das Erfordernis einer Standardisierung für die digitalen Informationsinfrastrukturen (Metadaten, Repositorien). Die Verwandlung von Dokumenten in maschinenprozessierbare Datenkorpora ist an ineinander greifende Standardisierungen gebunden, eine Anforderung, die auf bemerkenswerte Weise quer liegt zu der Kurzlebigkeit von Digitaltechnologien, kommerziellen Produktzyklen und Drittmittelforschung sowie

26 Sybille Krämer: Der „Stachel des Digitalen“ – ein Anreiz zur Selbstreflexion in den Geisteswissenschaften? Ein philosophischer Kommentar zu den Digital Humanities in neun Thesen. In: Digital Classics Online, Bd. 4,1. Hg. von Charlotte Schubert, 2018. Online verfügbar unter: <https://journals.ub.uni-heidelberg.de/index.php/dco/article/view/48490/41863> (abgerufen am 08.01.2020).

in Opposition steht zur grundsätzlichen Heterogenität der Datenkollektionen.

(I) Die Archivierung und Kuratierung²⁷ von Forschungsdaten bildet eine Herausforderung. Das sind jene Daten, die während des Forschens und als dessen Resultate anfallen und als Wissensressourcen langfristig zu sichern sind, weil einerseits die Nachvollziehbarkeit von Ergebnissen und andererseits die Anschlussmöglichkeit zukünftiger Forschung datentechnisch zu gewährleisten ist. Doch in den Geisteswissenschaften gibt es kaum ein disziplinenübergreifendes Verständnis, was aufzubewahrende Forschungsdaten überhaupt sind und welche Szenarien der Nachnutzung – auch international – anzustreben sind.²⁸ Die Entfaltung einer gemeinsamen Forschungsdatenkultur wird unabdingbar.²⁹

(II) Die Rolle von Bibliotheken zur Sicherung des kulturellen Gedächtnisses im Sinne des Weltschrifterbes erweitert sich mit der Lockerung der Bindung an das Medium Buch. Das bedeutet nicht nur, dass außer Büchern nun auch andere akustische wie bildliche Sparten unserer symbolischen Praktiken zu sammeln und bereitzustellen sind, sondern das impliziert auch eine gewisse Abkehr vom Sammlungsgedanken zugunsten der Idee Fachinformationsdienste anzubieten. Dazu gehört es, Metadatenschemata zu entwickeln, die den unterschiedlichen Formaten von gesammelten Objekten, Handschriften, Audiodateien, Filmen, Videos etc. gerecht werden und so katalogisierbar bzw. operationalisierbar sind, dass der Nachweis der digitalen Ressource zugleich mit ihrer Bereitstellung verbunden werden kann. All dies macht es erforderlich, dass Bibliotheken sich nicht nur als Dienstleister in der Informationsversorgung

-
- 27 Das Konzept des „Kuratierens von Daten“ erläutert Andrew Piper so: „It values the curation of data over the amount of data.“ Andrew Piper: There will be numbers. In: Journal of Cultural Analytics 5 (2016). <https://culturalanalytics.org/article/11062>; DOI: [10.22148/16.006](https://doi.org/10.22148/16.006) (abgerufen am 08.01.2020).
- 28 Selbst digitale Publikationen schöpfen die Potenziale der Nachnutzung mit computergestützten Methoden nicht aus, bemerkt Angela Dresen in Angela Dresen: Grenzen und Möglichkeiten der digitalen Kunstgeschichte. In: [kunsttexte.de 4](https://kunsttexte.de/4) (2017), S. 5. verfügbar unter: <https://edoc.hu-berlin.de/bitstream/handle/18452/19402/Angela%20Dressen%20-%20final.pdf> (abgerufen am 08.01.2020).
- 29 Heike Neuroth et al.: Aktives Forschungsdatenmanagement – Research Data Management Organiser. In: ABI Technik 38 (2018), H. 1, S. 55–64, <http://doi.org/10.1515/abitech-2018-0008> (abgerufen am 08.01.2020).

verstehen, sondern an strategischen Vorhaben mitarbeiten, die – auch von politischer Seite – der nationalen und internationalen Vernetzung dienen.

- 8 Die Bibliotheken sind Teil eines epistemologischen wie kulturtechnischen Umbruches; sie sind nicht einfach Dienstleister, sondern Akteure, welche die Bedingungen der Möglichkeit einer Mündigkeit und Souveränität im Umgang mit Daten zu schaffen haben.

Wir befinden uns in einer augenfälligen Umwandlung unserer Wissensressourcen, der uns prägenden Kulturtechniken sowie der wissenschaftlichen Methoden. Bibliotheken sollten in dem unübersichtlichen Gewirr alltäglicher datentechnologischer Detailfragen nicht aus den Augen verlieren, dass ihre Rolle in informationstechnischen Dienstleistungen nicht aufgeht. In etwas abgenutztem Jargon: Die Digitalisierung ist nicht zu vollziehen, sondern zu *gestalten*. Bibliotheken sind in der Position – in Kooperation mit anderen kulturellen und politischen Institutionen – darum zu ringen, dass wir Transparenz und Kontrolle unter den Bedingungen des Übergangs von der „Lesbarkeit der Welt“ zur „Maschinenlesbarkeit der Datenkorpora“ nicht verlieren, sondern gewinnen. Die europäische Aufklärung ist – immer noch – ein *Programm*, das noch in keiner Nation zu keiner Zeit gesellschaftliche Realität werden, geschweige denn imperiale und koloniale Praktiken hat verhindern können. Doch als Maxime einer Mündigkeit im Denken bleiben die Ideen der Aufklärungen von unübertrifftener wie uneingelöster Relevanz. Sie geben einen kritischen Maßstab ab, um deutlich zu machen, was am Bestehenden fehlt und was zu korrigieren ist. Allerdings ist es an der Zeit, das Gedankengut der Aufklärung weiterzuentwickeln, hin zu einer „digitalen Aufklärung“, die eine Mündigkeit und Souveränität im Umgang mit Daten projiziert und proklamiert. Dass dies nicht nur Gedankenspiele bleiben, sondern realiter Strukturen geschaffen werden, welche Datensouveränität ermöglichen und befördern: dafür haben auch Bibliotheken Sorge zu tragen.

9 Bruchstellen, Irritationen, Ambivalenzen.

(I) Bibliothekare und Bibliothekarinnen haben ehemals ihren Beruf zumeist aus „Liebe zum Buch“ ergriffen; jetzt sind sie konfrontiert mit einer Situation, in der das Buch als Leitmedium ihrer Arbeit erodiert und sie diese Transformation zugleich proaktiv zu gestalten haben.

(II) Digitalisierung wurzelt in der Diskretisierung und setzt eine Elementarisierung und Hierarchisierung der Text- und Bildobjekte voraus. Das geisteswissenschaftliche Verständnis von Sprache, Text und Bild ist allerdings über den Leisten solcher Zerlegung und Modularisierung in Bestandteile nicht zu schlagen. Es gibt eine grundständige Kluft zwischen der Digitalisierung von Zeichenbeständen und deren genuin holistischem, kontextabhängigem, geisteswissenschaftlichem Verständnis.

(III) Der Übergang vom Druckwerk zum maschinenprozessierbaren Datenkorpus kennt alle Arten von Zwischenstufen, bei denen im Digitalen das Medium „Druckwerk“ immer noch simuliert wird (z. B. PDF). Unerachtet der komfortablen Handhabbarkeit solcher Zwischenschritte sind – im Sinne der Vernetzung von Dateien – diese Residuen von Druckerzeugnissen im Digitalen möglichst zu überwinden: Doch das ist nicht immer „smart“, steht im Gegensatz zu gesellschaftlichen Reputationssystemen (Open Access-Probleme) und wirft eine Fülle rechtlicher Fragen auf.

(IV) Auch wenn „Big Data“ im Sinne riesiger textueller und bildlicher Datenvolumina erstmals die Möglichkeit eröffnen, nichtkanonische Werke in umfänglicher Weise in Untersuchungen miteinzubeziehen, so erzeugt die Verdatung von Text, Bild, Film, Musik zugleich bemerkenswerte Ausschlüsse: Was nicht in ein Digitalisat verwandelt wurde, fällt durch das Auffanggitter des kulturellen Gedächtnisses.

(V) Es gibt ein grundsätzliches Spannungsverhältnis zwischen der Kurzlebigkeit von allem, was im weitesten Sinne mit dem Digitalen zu tun hat, und dem Erfordernis einer Langfristigkeit in der Geltung von Standards wie der Sicherung der Überlebensfähigkeit digitaler Objekte. Generell gilt, dass die Bedingungen des Digitalen, die in beständiger Fortbewegung und Umstülpung bestehen, sich schlecht vertragen mit den Zielen von Wissenschaft, die auf Nachhaltigkeit, Überprüfbarkeit etc. zielen. „Dateninnovation“ und „wissenschaftliche Erkenntnis“ haben eine mitunter entgegengesetzte Verlaufsrichtung.