

# INTELLIGENTE SPEICHER, ALGORITHMISIERTE GEISTESWISSENSCHAFT. "DIGITAL HUMANITIES" UND IHRE CHANCEN FÜR LITERATUR- UND MEDIENWISSENSCHAFT

[Vortrag zum Workshop *Lektüre im Zeitalter digitaler Medien*, Humboldt Universität zu Berlin, 29. / 30. September 2017]

Für eine intelligente Adressierung strukturierter "big data" (*alias* Literatur)  
Kritik am (anglophonen) Begriff der "Digital Humanities"  
Computerpoesie (wieder) lesbar machen  
Kybernetische Lektüre im Zeitalter digitaler Medien  
Maschinelle Philologie  
Texte als Statistik lesen  
Kittler nicht historisieren, aber indizieren  
P. S.: "Digital Humanities" - am Ende ein Nebeneffekt der NSA?

## **Für eine intelligente Adressierung strukturierter "big data" (*alias* Literatur)**

Nachdem die Geburt der deutschsprachigen Medienwissenschaft mit der programmatischen, aber nicht realisierten "Austreibung des Geistes aus der Geisteswissenschaft" verbunden gewesen war, wird dies erst mit der effektiven Programmierung von Hochleistungsrechnern tatsächlich vollzogen. Der Diskurs der DH ist ein blasser, geradezu untoter Wiedergänger der heroischen Epoche kybernetischen Denkens. Nicht eine andere Epistemologie, sondern das schiere Mooresche Gesetz trennt DH von diesem Vorlauf. Heute vermögen Prozessorkerne jene Datenmengen zu verarbeiten, deren Algorithmen in den Diagrammen der kybernetischen Informationsästhetik längst definiert waren - für die Kalkulierbarkeit von Literaturen zumal. Ebenso wußte die Kybernetik um die Grenzen einer rein explorativen Durchmusterung von Symbolketten: "Jedes rein *stochastische* Vorgehen ist jedoch für Maschinen ebenso sinnlos wie für das menschliche Verhalten."<sup>1</sup> Wenn weiterhin von Literatur und nicht schlicht von Markovketten die Rede sein soll, läuft dies einerseits auf eine veritable Medienphilologie, andererseits die Notwendigkeit von "Humanities of the Digital" hinaus (Jan Claas van Treeck), ausgehend von der Frage: "What did the Digital to the Humanities?"<sup>2</sup> Ferner resultiert daraus ein vom binären Computer nahegelegter, erweiterter Literaturbegriff. Im Sinne des kleinsten denkbaren Alphabets sind für die Algorithmen der "visual analytics" respektive "sonic analytics" Töne und Bilder ebenso alphanumerische *litterae*. Die Herausforderung von Textwissenschaft in Zeiten von DH wird darin liegen, den Literaturbegriff wieder zu begrenzen, statt der im Namen von "big data" verheißenen Entgrenzung.

## **Kritik am (anglophonen) Begriff der "Digital Humanities"**

---

1 Rul Gunzenhäuser, Zur Synthese von Texten mit Hilfe programmgesteuerter Zifferrechenanlagen [\*1963], reprinted in: Kaleidoskopien Bd. 5 (2004), 170-183 (173)

2 So kommuniziert von Sumdandro Chattapadhyay, Treffen 3. Juli 2014, am Standort der Medienwissenschaft der Humboldt-Universität zu Berlin

Das Lehnwort DIGITAL HUMANITIES klingt zunächst befremdlich an einer nach den sprachgewaltigen Gebrüdern Humboldt benannten Universität. An dergleichen *alma mater* hat G. W. F. Hegel als Philosoph ausdrücklich zur Arbeit an der "Strenge des Begriffs" aufgerufen. "Digital Humanities" mögen besser in "Algorithmisierte Geisteswissenschaft" übersetzt und überführt werden - keine Rückkehr in Diltheys akademische Galaxie, sondern gekoppelt mit techniknaher Medienwissenschaft versteht sie sich als operative Epistemologie des Digitalen.

Die Verschanzung hinter dem Anglizismus "DH" indiziert eine gewisse Befremdung, als Indiz einer kritischen Distanzierung gegenüber eigenen Traditionen der Hermeneutik an der deutschen Universität. Die Geister treten zumindest in den Plural, und auch außerhalb des wissensarchitektonischen Gefüges der Universität: "Ubiquitous scholarship is marked by [...] the pen and the smartphone, the millennia-long histories of humankind and the real-time feeds of the now"<sup>3</sup>, heißt es in einem Sammelband zu Digital Humanities, und: "The singularity of the "I-subject" has been transformed into the collaborative authorship of a 'we-subject."<sup>4</sup>

Der Wiedereintritt von *humanities* nach der Austreibung des Geistes aus den Geisteswissenschaften, ist, in geradezu Hegelscher Dialektik, die technik- und mathematiknahe Medienwissenschaft als "Aufhebung" der Dichotomie von Geist und Technologie, parallel zum Wiederaufleben der kybernetischen Hypothese einer Aufhebung der Mensch-Maschine-Differenz in und als Signalverarbeitung. Medientheorie versucht sich an einer Geisteswissenschaft des Digitalen, bzw. "Humanities of the Digital" (frei nach Jan Claas van Treeck). Für die Kernoperation der neuen (kartesischen) Methode als Denk-Weg, nämlich den Algorithmus als schrittweise, analytische Problemlösung, heißt dies: Vor aller Implementierung in techno-logische Systeme, d. h. bereits *vor* allem Quellcode im informatischen und Quellkodierung im nachrichtentechnischen Sinne) bedarf des der Klärung, welches Problem durch Berechnung überhaupt lösbar ist (Turing's "Halteproblem" und die scheinbare Tautologie der *computable numbers*). Kriterien dafür sind Endlichkeit und Code-Effizienz (Rechenzeitdauer *versus* Datenmenge und Speicherkapazität) - klassische Kompetenzfelder wissenschaftlicher Vernunft. Der Programmablaufplan (*flow chart*) ist eine Form des *diagrammatic reasoning* i. S. von Charles S. Peirce, und jedem aktiven Programmierer ist die Notwendigkeit der Kommentierung des Quellcode neben den eigentlichen Mnemonics vertraut - ein Wiedereintritt von Hermeneutik und der damit verbundenen kritischen Computerphilologie.

## **Computerpoesie (wieder) lesbar machen**

*Digitalia* haben die Eigenschaft, der sinnlichen Wahrnehmung nicht unmittelbar zugänglich zu sein; "sie müssen durch Software interpretiert und lesbar gemacht werden"<sup>5</sup>. Eine zentrale Aufgabe von Wissensdesign ist also nicht

---

3 Anne Burdick / Johanna Drucker / Peter Lunenfeld / Todd Presner / Jeffrey Schnapp, *Digital\_Humanities*, Cambridge, MA (MIT Press) 2012, 60

4 *Ibid.*, 84

5 Eric W. Steinhauer, „Wissen ohne Zukunft?“ Referat auf der Urheberrechtstagung am 9. November 2011 in Göttingen

schlicht die graphische Gestaltung großer Datenmengen unter Einsatz des Computers, sondern die Sichtbarmachung dieses Unsichtbaren in der Datenprozessierung selbst - den Prozeß eher als schlicht das finale Interface-Phänomen.

[Streng mit dem promovierten Literaturwissenschaftler McLuhan gelesen (*Understanding Media*, 1964) heißt kritische Medienanalyse, gegenüber der Oberflächen-Figur (dem sogenannten Inhalt) auf den technologischen Grund, die eigentliche Medienbotschaft abzuheben; gegenüber der literarischen Oberfläche lautet ihr "sonischer" Grund in Zeiten elektronischer Rundfunkmedien *acoustic space*. Ein Beitrag zum Begriff techniknaher Medienphilologie sucht damit zugleich dem Un-wort des "Medialen" zu widerstehen und den Textbegriff seinerseits konkret medientechnisch um die Analyse von Hardware-Konfigurationen, Schaltplänen und Quellcodes zu erweitern. Hinzu kommt ihr nachrichtentechnischer Zugriff, wie sie Claude Shannon mit seinem Text über "Entropy in English" meisterhaft vorgegeben hat.<sup>6]</sup>

"Die Texte wandern von Papyrus, Pergament oder Papier ins virtuelle Interface ab - wobei die zugrunde liegenden numerischen Daten auch für die geübtesten Leser\*innen unlesbar sind"<sup>7</sup> - die von Frider Nake so genannte "Unterfläche". Auch Filmphilologie behauptet beharrlich, was zählt, ist die Fortdauer von Film als Format, gleich in welchem technischen Dispositiv verkörpert. "Alle Kunst scheint aber an ihr Ende zu kommen, wenn auch der Text selbst verschwindet. Dieser ist im Computerzeitalter nur noch ein Oberflächeneffekt, der auf den Bildschirmen stationärer oder mobiler Endgeräte visualisiert werden kann (aber nicht muss). Die realen Datenströme zirkulieren derweil als Spannungsimpulse in und zwischen Computern."

"Unkörperlich" sind digitale Informationen nie: Sie bedürfen immer der zumindest zeitweiligen Implementierung in realer (Elektro-)Physik, um wirksam (und maschinen- sowie menschenlesbar) zu werden. An die Stelle des prinzipiell dauernden Drucktexts tritt der regenerative Text *auf Zeit*, das fortwährend Ephemäre, "the enduring ephemeral" (Wendy Chun). Der Vergleich des mathematischen Algorithmus mit der Kulturtechnik Kochrezept ist problematisch sowohl hinsichtlich der thermodynamischen Entropie und der elektronischen Materialität seiner Verkörperung (eben jenseits des performativen Menschen), welche die Rechenvorschrift erst in die Welt, d. h. in die Zeit bringt. Um *executable* zu sein, bedarf es für Software einer real operativen Maschine, denn ihre logischen Grundstrukturen lauten Sequenz, Entscheidung, und Schleife.

Aleida Assmann zufolge ist die Wissensgeschichte professionellen Lesens modellhaft eine Abfolge kultursymbolischer Deutungsrahmen, "die den Vorgang des Lesens steuern, indem sie definieren, welche Sinnbildungsverfahren jeweils zulässig sind"<sup>8</sup>; heilige Text wie die Bibel waren nur dann wahrhaft verständlich, wenn der Leerraum zwischen seinen Zeilen mit anderen Quellen oder Stimmen

---

6 xxx

7 Exposé (Stephan Brändle / Maximilian Pöttsch) zur Beitrag zum Workshop *Lektüre im Zeitalter digitaler Medien*, Humboldt-Universität zu Berlin, 29. / 30. September 2017

aufgefüllt wurde. An dieser Stelle kürzt Medienarchäologie drastisch ab, um sich nicht in kulturgeschichtlichen und -wissenschaftlichen Vorgeschichten und Diskursen zu verlieren, und schaut auf die herrschende Textsorte der Gegenwart: den Quellcode von Software, die nicht mehr von Priestern und Gläubigen, sondern von der CPU in Mikroprozessoren gelesen wird, um folglich für Menschengenau überhaupt erst lesbar zu sein. Dieser primäre nicht-menschliche "Interpretant" (i. S. von Peirce) ist algorithmischer Natur.<sup>9</sup>

Dem entspricht exemplarisch die sogenannte "ergodische" Literatur, die nicht fest geschrieben steht (außer im Quellcode), sondern im menschlichen Gelesenwerden sich ständig neu konfiguriert und entfaltet, etwa John Cayleys Hypercard-basierter Poesiegenerator *The Speaking Clock*.<sup>10</sup>

[Im Sommer 2016 wurde dieses Programm durch Johannes Maibaum auf einem antiken PowerPC-Macintosh im Signallabor des Instituts für Musikwissenschaft und Medienwissenschaft der Humboldt-Universität zu Berlin wieder in Funktion gesetzt und in Form eines Screencasts im Streaming Media-Kanal YouTube dokumentiert.<sup>11</sup>]

## **Kybernetische Lektüre im Zeitalter digitaler Medien**

Das Verhältnis von Digital Humanities und Literatur im engeren Sinne ist nur auf den ersten Blick ein privilegiertes. Daß sich "unter dem zunehmenden Einfluss digitaler Medien" der Umgang mit Texten "nachhaltig" verändert<sup>12</sup> ist eine schwache literaturwissenschaftliche These. Demgegenüber emergiert als Funktion nachrichtentechnischen Wissens eine genuin autonome, neuartige Form des Begriffs von "Literatur", seit Zeiten der kybernetischen Informationsästhetik. So wurde im Hörspiel Max Benses, *Der Dialog der Terry Jo*, aus stochastisch organisierten Buchstabenmengen tatsächlich eine dramatische Form, mithin Literatur. Dies aber erfordert keine primär literatur-, sondern zunächst einmal radikal textwissenschaftliche Analyse, geradezu Medienphilologie (für die Datenstrings) und Medienarchäologie (für die sie prozessierenden Maschinen).

---

8 Exposé, unter Bezug auf: Aleida Assmann: *Im Dickicht der Zeichen*. Berlin 2015

9 Frieder Nake, *Das algorithmische Zeichen und die Maschine*, in: Hansjürgen Paul / Erich Latniak (Hg.), *Perspektiven der Gestaltung von Arbeit und Technik*. Festschrift für Peter Brödner, München / Mering (Rainer Hampp) 2004, 203-223

10 Siehe Espen Aarseth: "Aporia and Epiphany in 'Doom' and 'The Speaking Clock.'" In: *Cyberspace Textuality. Computer Technology and Literary Theory*, hg. v. Mary-Laure Ryan, Bloomington 1999, 31-41

11 <https://youtu.be/ZHyni32CyQA> (Version mit deutschen Untertiteln); <https://youtu.be/GY7NMslGm94> (Version mit englischen Untertiteln); die Funktionsweise des Programms wird erläutert und ein Beispiellauf des Programms nachvollziehbar

12 Exposé zum interdisziplinären Symposium *Lektüre im Zeitalter digitaler Medien* an der Humboldt-Universität zu Berlin, 29. / 30. September 2017 (Initiative und Durchführung: Stephan Brändle & Maximilian Pöttsch)

Was sich derzeit im Namen der sogenannten Digital Humanities vollzieht, ist in vielen Hinsichten eine ebenso gleichursprüngliche wie evolutionär verästelte Wiedereinkehr der klassischen kybernetischen Methoden, Hypothesen und Werkzeuge. Die Informationsästhetik der deutschsprachigen Nachkriegskybernetik ist ganz und gar noch nicht historisiert (im Sinne Lacans, der damit die Insistenz unmemorisierbarer Traumata meint); sie ist vielmehr im Begriff der Digital Humanities "aufgehoben". Was seinerzeit technisch kaum realisierbar war, weil es an den Prozessorkapazitäten röhrenbasierter Großcomputer scheiterte, ist nun mit High Performance Computing machbar: die algorithmische Durchmusterung von *big data*.

Die kybernetische Ästhetik hat solch distantes Lesen nicht *avant la lettre* von Digital Humanities, sondern *avec les lettres* längst praktiziert: mit der rechenleistungsbedingten epochalen Einschränkung auf die *medium size data* zumeist eines oder weniger Werke. In der Tat, jeder Text ist selbst schon buchstäblich eine "große Daten"menge, wenn nicht als Satzfolge, sondern buchstäblich entziffert, als statistische Menge gelesen im Sinne Markovs und Shannons, mit Übergangswahrscheinlichkeiten in Einzel- und Zweiergruppen, dann in Trigrammen usf.. Die Urszene dieser Form von zugleich "close" und "distant reading" war Markovs Analyse von Puschkins Poem *Eugen Onegin*, mithin also ausgerechnet Literatur, für Literaturwissenschaft eine nicht weniger frohe Botschaft als die Ursprung des Vokalalphabets aus dem Geist der homerischen Gesänge.<sup>13</sup>

Computergestützte Textanalyseverfahren zeitigen bislang eher triviale Resultate; der Datenberg kreißt und gebiert eine Wissensmaus. Entscheidend ist, überhaupt erst auf die Fragen kommen, die algorithmische *text analytics* entbirgt. Tatsächlich eröffnen binär kodierte Informationen "ganz neue Möglichkeiten der Verarbeitung, auch literarischer Werke" (Exposé), etwa die Entropie eines Textes zu messen, wie Manovich und Wannhoff es anhand von (zunächst allerdings noch singulären) abendfüllenden Spielfilmen, nach Einzelkadern sortiert, vollzogen haben.<sup>14</sup> Die eigentliche Stunde von und für eine Medienwissenschaft, die den Computer als das modellbildende Medium zum Ausgang ihrer Forschung und Lehre nimmt, schlägt erst, wenn diese Form von *cultural analytics* (Manovich) auf die Datenmenge von einhundert Jahren Film selbst exponentiell erweitert wird.<sup>15</sup>

[DH unterliegen in solchen Operationen ihrerseits nach wie vor dem Primat des Visus in der alphabetisierten Kultur (McLuhan); im nicht-numerischen Computing aber gibt es keinen zwingenden Grund mehr, visuelle Diagramme zu privilegieren. Stattdessen erschließen sich durch die Sonifikation hochdimensionaler "big data", also durch die zeitliche Abfolge, dem Ohr als Erkenntnisorgan Muster gerade dort, wo sie im zweidimensionalen graphischen *plotting* noch zufällig erscheinen, etwa in Form von Autokorrelationen. David Worrall hat den höchst zeitkritischen Hochfrequenzhandel an der elektronischen Börse (wie es im Frequenzbegriff schon nahegelegt ist)

---

13 Dazu W. E. / Friedrich Kittler (Hg.), Die Geburt des Vokalalphabets aus dem Geist der Poesie, München (Fink) 2007

14 Siehe Matthias Wannhoff, xxx, *online* [www.medientheorien.huberlin.de/hausarbeiten](http://www.medientheorien.huberlin.de/hausarbeiten) xxx

15 Eine These von Philipp Sander, Berlin (September 2017)

sonifiziert wie Benoit Mandelbrot einst Fraktale; in beiden Verfahren zeitigt das Rechenmedium einerseits optische, andererseits klangliche Figuren.<sup>16</sup> Tatsächlich verbirgt sich bereits hinter der Errechnung von Mandelbrot- und Julia-Mengen die musikalische "Algorhythmik" (Miyazaki) des Mikroprozessors.]

Medienarchäologie betreibt eine *Verunsinnlichung* von Literaturanalyse durch radikale Mathematisierung, um nicht in die hermeneutische Falle zu tappen, daß erst die menschliche Wahrnehmung und Kognition als Nachrichtensenke zählt, sondern - vorgeschaltet - bereits der technische Empfänger der kanalkodierten Signale.

Eine durch und durch algorithmisierte Literaturwissenschaft, die sich des Computers nicht schlicht als ergänzendes Werkzeug bedient, bricht mit traditionellen Theorien philologischer Lektüren. Doch werden diese damit nicht "obsolet" (Exposé). Epistemologische DISKONTINUIERUNG (ein wissensstrategisches Plädoyer Foucaults in Anschluß an Bachelard) besteht buchstäblich zum Großteil aus KONTINUITÄT.<sup>17</sup> Dies gelingt kaum aus dem Innern der literaturwissenschaftlich orientierten Philologien, vielmehr allein in Distanz zum Fach, bspw. als Medienwissenschaft.

[Auf Foucault als Textkorpus selbst angewendet: Peer Trilcke, Frank Fischer: Fernlesen mit Foucault? Praxis des *distant reading* und zur Operationalisierung von Foucaults Diskursanalyse. *foucaultblog* 2016; <http://www.fsw.uzh.ch/foucaultblog/featured/141/fernlesen-mit-foucault>]

## Maschinelle Philologie

Das *re-entry* geistvoller Fragestellungen in die algorithmische Analyse, also geistvolles Vorgehen, aber ist - seit Turings und Lacans Definition von *mind* - längst kein Privileg des Menschen mehr. Die "Lektüre" kodierter Symbolketten ist nicht mehr exklusiver Vollzug des menschlichen Sinnesapparats; neben den körpergebundenen Kulturtechniken des Lesens und Schreibens gelingt vielmehr dem Lese-/Schreibkopf der *turingmaschine* in Verbindung von Anweisungstabellen das, wonach eine Geisteswissenschaft jenseits der Hermeneutik verzweifelt strebte: die (zumindest zeitweilige Suspendierung von vorschneller kultureller Sinnggebung, also der *Kalkül* "anstelle eines phänomenologischen oder hermeneutischen Verstehens"<sup>18</sup>. Die Turingmaschine basiert ausdrücklich auf Lese- und Schreiboperationen diskreter Symbole (*scanning*). Das Lesen alphabetischer Codes war (anders als Signalverarbeitung in Lebewesen) die längste Zeit ein Proprium menschlicher Kultur. Ein nicht-menschliches *close reading*, eine andere Philologie erlaubt erstmals der medienaktive, archäologische Blick des Computers selbst. Der *bit*-genaue, also ebenso "textkritische" Algorithmus operiert im Dienst all jener Datenmengen,

---

16 Dazu Florian Dambois, Angeschlagene Moderne, in: Andi Schoon / Axel Volmar (Hg.), Das geschulte Ohr, Bielefeld (transcript) 2012, 165-170 (152)

17 Ein Argument von Walter Seitter, Wien

18 Helmar Frank, Kybernetische Pädagogik, in: Simon Moser (Hg.), Information und Kommunikation [1967], 111-120; Wiederabdruck in: Kybernetische Pädagogik Bd. 1 (1973), 477-486 (113/479)

welche das binäre Sampling von Text-, Ton- und Bildwelten und seine digitalen Speicher zur Analyse bereitstellen.

"The algorithm, a set of step-by-step instructions, is the heart of software programs, but these instructions have to be translated into a binary language that the computing hardware understands."<sup>19</sup> Dieser "Verstehens"begriff aber ist längst am Modell der Turing Maschine orientiert. Er wird damit nicht post-, sondern vielmehr infrahumanistisch in dem Sinne, daß die Turingmaschine dem Denken den Spiegel vorhält: Rechnet der Mensch (im Kopf oder ausgestattet mit kariertem Papier, einem Bleistift und einem Radiergummi), ist er selbst im Maschinenzustand.

"Jedes rein *stochastische* Vorgehen ist jedoch für Maschinen ebenso sinnlos wie für das menschliche Verhalten."<sup>20</sup> Lektüre als Dekodierung liest noch nicht Literatur, sondern ist buchstäblich(e) Auf-Lese, Sampling, ein Versammeln von Symbolketten mit Übergangswahrscheinlichkeiten.

- Seit der Epoche von Lochkarten und -streifen enthalten kodierte Speichermedien "die Informationen in einer für Datenverarbeitungsanlagen lesbaren Form und sind in bestimmten Grenzen beliebig oft verwendbar."<sup>21</sup> Damit löst sich die maschinelle Operation a) durch interne Autonomie und b) durch für Menschen nicht mehr leistbare Operationen in Zeit und Genauigkeit und Gedächtnis von körpergebundenen Kulturtechniken (insofern sie, mit Marcel Mauss definiert, in *Körpertechniken* gründen). Diese Differenz konkretisiert sich mit Schreibmaschinen- und Druckschrift, die buchstäblich "zwischen" Hand und Schrift treten (Heidegger, "Parmenides"). Ein Argument in Semen Karsakovs 1832er Entwurf einer komputativen *Ideenmaschine* (Velminski / Ernst 20xx) lautet: Sobald Sprache in alphabetische Zeichen entäußert wird, ist sie als symbolische Ordnung mechanisierbar.

Der Begriff von "Digital Humanities" ist nicht allein der matte Abglanz, sondern auch das Vergessen ihres Vorweglaufs unter dem schönen Namen der "Informationsästhetik". Maschinelle Philologie heißt Nutzbarmachung der elektronischen Datenverarbeitung in der textwissenschaftlichen Quellenkunde; so definierte es Gundlach 1968, also auf dem Höhepunkt des kybernetischen Denkens <232>. Dies meint nicht schlicht eine Maschinisierung von Lektüre im traditionellen Sinn, sondern ebenso seine Mathematisierung (die eigentliche Bedeutung radikaler Medienarchäologie); mit Gödel und Turing wird der Algorithmus zum Inbegriff der "maschinalen" Operation selbst.

- Das qualitative *close reading* weicht nicht schlicht einem quantitativen *distant reading* von "big data"; deren Wert entfaltet sich erst in Kopplung mit einer neuen Qualität, der algorithmisierten Geisteswissenschaft.<sup>22</sup> Gerade *computational analytics* will nicht länger das Kind (den Geist) mit dem Bade

---

19 Burdick et al. 2012: 53

20 Rul Gunzenhäuser, Zur Synthese von Texten mit Hilfe programmgesteuerter Zifferrechenanlagen [1963], in: Kaleidoskopien Bd. 5 (2004), 170-183 (173)

21 Siegfried Kühlewind / Klaus Schwedler, Datenträger. Fachkunde für Datenverarbeiter, 6., bearb. Aufl. Berlin (Verlag Die Wirtschaft) 1969, 7

22 Siehe Stephen Ramsay, Reading Machines. Toward an Algorithmic Criticism, Urbana 2011

(den *streaming data*) ausschütten. Literaturwissenschaft unter hochtechnischen Bedingungen erfordert daher nicht nur neue, sondern durchaus vertraute Fertigkeiten im Umgang mit Software und anderen analytischen Werkzeugen, nämlich Quellenkritik im Doppelsinne der Informatik (Quellcodes) und ihrer materialen Verkörperung als technische Medien. Gerade der "kalte Blick" der Medienarchäologie bahnt den Weg für ein *re-entry* geisteswissenschaftlicher Kompetenz: die epistemologische Frage nach Wesen (objektorient) und den Vollzugsweisen (prozeßorientiert) des Hochtechnischen.

## Texte als Statistik lesen

Geisteswissenschaften werden als *digital humanities* zu *science*. Damit gemeint ist die Verwendung von Meßwerten als Spannungswerten (elektronisch) respektive Zahlen ("digital"), um Hypothesen zu überprüfen. Die in Anschlag gebrachten Algorithmen selbst legen dabei neue Hypothesen erst nahe. Mit Zahlen lassen sich Muster und Bezüge ausdrücken, für die den klassischen Geisteswissenschaften bislang die Worte respektive Formeln und *termini technici* fehlten. Harold Bloom sagte 1973 voraus, sogenannte Intertextualitätsforschung werde einmal vom Computer abgelöst werden. "Ein traditioneller analoger Gelehrter" argumentiert aus Sicht der Naturwissenschaften "anekdotisch". Diagramme, Statistiken, Visualisierung - Reduktion auf Quantifizierung. Der Versuch einer - analog zur Biologie - evolutionären Theorie der Kultur wird sich "nicht in Worten, sondern in Gleichungen ausdrücken". Im Unterschied zur Natur aber erschaffen Menschen Dinge mit "Bedeutung", die der geisteswissenschaftlichen Erklärung bedarf - "ein Bereich des Verstehens, den Algorithmen und Zahlen nicht erfassen". Dem gegenüber stehen nun *deep learning*-Algorithmen, die durch die prozessorteknische Ermöglichungsbedingung massiver iterativer Rechenleistungen etwa stilometrische "Bedeutungen aus Daten" abzuleiten vermögen. Durch solch operative Hermeneutik lassen sich "neue Bedeutung von algorithmischen Mikroskopen entdecken <...>, die noch gar nicht gebaut sind."<sup>23</sup>

Wenn hochkulturelle Texte genuin *medienphilologisch* verhandelt werden, sind sie nicht mehr auf ihre literarische Form beschränkt, sondern algorithmisch eben auch in ihrer Statistik und Stochastik erkennbar. Ausgerechnet der Diskurstifter der "Geisteswissenschaften" Wilhelm Dilthey sah im Buchdruck ein zählendes Verfahren zur Messung wissenschaftlich-ästhetischer Strömungen angelegt: "Es bedarf nur der Ausnutzung des" - *nota bene* - "gesamten Bücherbestandes unserer Bibliotheken nach statistischen Methoden. Durch eine solche wird man einmal das ganze Causalverhältniss einer geistigen Bewegung <...> in den wesentlichen Gliedern vorstellig machen können"<sup>24</sup>

---

23 So der Evolutionsbiologie Armand Marie Leroi, Cicero zählen. Algorithmus oder Kritik?, in: Süddeutsche Zeitung, 6. März 2015

24 Wilhelm Dilthey, Archive der Litteratur in ihrer Bedeutung für das Studium der Geschichte der Philosophie, in: Archiv für Geschichte der Philosophie, II. Band 3. Heft, Berlin (Reimer) 1889, 343-367 (355)



Digital Humanities sind dort noch sehr altgeisteswissenschaft, wo sie wissenschaftlich dem Narrativ eines Ursprungsmoments erliegen. Das betrifft die schon kanonische Geburtsstunde der *computational philology* selbst. Der Jesuitenpater Roberto Busa erstellte mit Unterstützung von IBM eine Konkordanz der Schriften Thomas von Aquins<sup>25</sup> - welcher einst folgenreich als scholastischer Aristoteles-Übersetzer den Begriff *to metaxy* ins lateinische *medium* überführte.<sup>26]</sup>

Die kybernetische Informationsästhetik hat hier weitere Wege gewiesen.<sup>27</sup> "Wie die Alpen aus zunehmender Flughöhe sichtbar werden, so auch die Archivmassen"<sup>28</sup> ; dem entspricht das statistische *mapping* von Kant-Texten, sowie der panoptische Datenblick als "distant reading"<sup>29</sup> in den Digital Humanities. Immanuel Kants *Kritik der Urteilskraft* läßt sich nachrichtentechnisch lesen - nämlich mit signalverarbeitendem, textstatistischem Blick.<sup>30</sup> Dies ist wahrhaft Philologie von (Seiten der) Medien, denn zunächst ist hier der Computer der aktiven Philologe. Um eine Philologie der Medien im anderen Sinne handelt es sich, wenn der Computer in seiner Hard- und Software selbst zum Gegenstand philologischer Kritik wird.

Als Digital Humanities *avant la lettre*: nahm die Mathematisierung geisteswissenschaftlicher Disziplinen bereits 1967 ständig zu. "Vor allem die Entwicklung der Kybernetik hat diesen Prozeß enorm beschleunigt", heißt gleich eingangs in Manfred Kiemles Informationsästhetik.<sup>31</sup> Den tatsächlichen techno-logischen Maschine ist also das mathematische Verhältnis zur Welt vorgelagert, das Heidegger ("Zeit des Weltbilds") mit Descartes ansetzt. Das Wesen des Technischen ist daher zunächst nichts "technisches", sondern der mathematische *logos* der Technologie. Es war der Mathematiker Andrej Markov, der die nach ihm benannten Prozesse bzw. Ketten als Modell (und seitdem spezielle Klasse stochastischer Prozesse) für die Zufallsevolution eines gedächtnislosen Systems formulierte bzw. (als bislang latentes, "implizites" Wissen) entdeckte, wo "the likelihood of a given future state, at any given moment, depends only on its present state, and not on any past states"<sup>32</sup>.

---

25 Burdick et al. 2013: 123

26 Dazu der Beitrag von Wolfgang Hagen, in: Rösler / Münker (Hg.), Was ist ein Medium?, Frankfurt/M. (Suhrkamp) 2xxx, xxx-xxx

27 Siehe Wilhelm Fucks, Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen, Köln / Opladen (Westdeutscher Verlag) 1955

28 Walter Seitter, Vortrag „Archäologie des Wissens. Die Rhetorisierung der Bibliothek“, Fakultät Medien, Bauhaus-Universität Weimar, 2. Februar 1998

29 Franco Moretti, Distant Reading, Konstanz 2016

30 Axel Roch, Adressierung von Texten als Signale über Bilder, in: Verstärker. Von Strömungen, Spannungen und überschreibenden Bewegungen, Jg. 2, Nr. 2, Mai 1997, hg. v. Markus Krajewski u. Harun Maye; *online* [http://www.culture.hu-berlin.de/verstaerker/vs002/roch\\_adresse.html](http://www.culture.hu-berlin.de/verstaerker/vs002/roch_adresse.html)

31 Manfred Kiemle, Ästhetische Probleme der Architektur unter dem Aspekt der Informationsästhetik, Quickborn (Verlag Schnelle) 1967, 9

32 Editorische Anmerkung zu: Abraham A. Moles, Cybernetics and the Work of Art [\*1965], in: Margit Rosen (Hg.), A Little-Known Story about a Movement, a Magazine, and the Computer's Arrival in Arts. New Tendencies and Bit International, 1961-1973, Karlsruhe (ZKM) /

Diese Ergodik ist eine Herausforderung des Historismus. Vorgelagert ist diese Offenlegung des stochastischen Prinzips in einer quasi-archivischen Operation: in der statistischen Analyse von diskreten Zeichen: Markov-Ketten, ob nun alphabetisch, oder auch musikalische Notenzeichen. Eine solche Identifizierung von Übergangswahrscheinlichkeiten erlaubt im Umkehrverfahren die computerbasierte Synthese literarischer oder musikalischer Kompositionen: eine kybernetische Ästhetik, deren Feuer in den Herzen der Medienarchäologen (im Unterschied zu ihrer wissensgeschichtlichen Historisierung) noch brennt, als Differenz von systemübergreifender Operativierung *versus* schlicht anthropozentrischer Performanz.

### **Kittler nicht historisieren, aber indizieren**

Ein prominenter Exorzist des Geistes, Friedrich Kittler, schreibt ganz am Ende seines akademischen Denkwegs, in Werk *Musik & Mathematik Bd. 2: Eros*, von "Rekursionen" (*in*) der Mediengeschichte, konkret: die Wiedereinkehr der Funktionalität des altgriechischen Alphabets, das einmal als *ein* und derselbe Zeichensatz für musikalische Notation, mathematische Rechnung und eben auch zur Darstellung einer hochvokalisierten Sprache eingesetzt wurde, im Computercode.

Kittler, der Autor emergierender Medienwissenschaft, wird nun seinerseits zum Gegenstand des "Kittler Bot" im Literaturarchiv Marbach, wo Kittlers Nachlaß (Manuskripte, selbstgelötete Synthesizer, Computerprogramme) liegen. Der sogenannte "Indexer" erschließt sein digitales Erbe. Der "Indexer", welchen diesen digitalen Nachlaß suchmaschinenhaft erschließt, erlaubt die gezielte Extrahierung der Kommentare Kittlers zu dem von ihm geschriebenen Quellcode seiner operativen Texte. So interessiert sich das philologische Archiv wieder (nur) für den "literarischen" Anteil.

*Humanities* zweiter Ordnung, nämlich in bester Tradition der Kritik antiker Textüberlieferung im frühneuzeitlichen Humanismus, ist *Medienphilologie von Software*.

### **P. S.: "Digital Humanities" - am Ende ein Nebeneffekt der NSA?**

*Distant reading* heißt immer auch panskopische Überwachung als *dataveillance*. Googles proprietärer *Ngram viewer* als bevorzugtes Werkzeug von Literaturanalyse in Zeiten von DH korreliert mit der Epoche der IT-basierten positiven respektive negativen Rasterfahndung in der Epoche Herold im Bundesamt für Verfassungsschutz.<sup>33</sup>

- "Digital Humanities", verstanden als die algorithmische Durchmusterung großer Datengruppen mit Erkenntniskriterien wie der Entropie als Informationsmaß jenseits der klassischen Identifizierung von individuellen Werken, wird von der NSA als *profiling* längst praktiziert. Eskalation der empirisch-statistischen Soziologie: "[...] sammelt man Daten und durchsucht

---

Cambridge, Mass. (MIT) 2011, 217-225 (225)

33 Dazu Josef Vogl, Grinsen mit Katze, in: Claus Pias (Hg.), xxx 1998

diese mithilfe von Algorithmen nach Regelmäßigkeiten, die man dann in die Zukunft extrapoliert."<sup>34</sup> *Predictive analytics* ist eine Kernoperation mit / über *big data*. So tendieren die "digital humanities" dazu, sich unwillkürlich der Datenästhetik solcher Logistik anzupassen und deren Software-Heeresgerät für schöngeistige Themen zu mißbrauchen; dem gegenüber erhalten gewisse altmodische, langsame Praktiken und kritische Distanzen der klassischen Geisteswissenschaften (als Humanismus im akademischen Sinne) eine neue, geradezu ideologiekritische Attraktivität.

---

34 Eintrag "Predictive Analytics" in the *Glossary* of Geiselberger / Moorstedt (editing) 2013: 301