

DIE DIGITALE NACHHALTIGKEIT DER GEISTESWISSENSCHAFTEN

Nachdem die Digitalen Geisteswissenschaften in der globalen Forschungslandschaft als Ansatz kenntlich sind und an vielen Universitäten gelehrt werden, entscheidet die Entwicklung in den kommenden zehn Jahren über die Nachhaltigkeit des gesamten Prozesses. Etabliert haben sich bisher zwei unterschiedliche Arten der universitären Institutionalisierung: Auf der einen Seite stehen die Digital Humanities, die sich als eine »Transformationswissenschaft« mit dem Ziel beschreiben, Know How für die »geisteswissenschaftlichen Fachdisziplinen« zu entwickeln.¹ Auf der anderen Seite haben sich innerhalb der vorhandenen Fächer digitale Forschungsprojekte und Lehrprogramme durchgesetzt. Beide Richtungen stehen im Wettbewerb miteinander und müssen sich zugleich eng austauschen, weil sie die gleichen Schwierigkeiten zu bewältigen haben. Dabei lassen sich drei besondere Herausforderungen für Nachhaltigkeit unterscheiden: Zugänglichkeit, Kompetenz und Ausbildung.

Im wissenschaftlichen Alltag tritt, wie in der generellen gesellschaftlichen Alltäglichkeit, das Zugangsproblem am offensichtlichsten zutage. Die Umwelt digitaler geisteswissenschaftlicher Forschung besteht aus der überwältigenden Mehrheit von Kolleginnen und Kollegen, die selbstverständlich ihre digitalen Endgeräte, das Internet und soziale Netzwerke nutzen und trotzdem methodisch bei dem bleiben, was sie in vordigitaler Zeit gelernt haben und was auch ohne digitale Medien funktionieren würde. Entgegen einer verbreiteten Annahme sind ihnen gegenüber die Studierendengenerationen der *digital natives* keineswegs im Vorteil, wenn es um den Zugang zu den Methoden geht, mit denen die Digitalen Geisteswissenschaften arbeiten. Digitale Methoden im engeren Sinn beruhen auf dem Dreischritt aus Formalisierung, Quantifizierung und Modellierung der Forschungsgegenstände, also von Texten, Bildern, Objekten oder sozialen Interaktionen. Wer auf diese Art mit den Gegenständen der Geisteswissenschaften verfährt, braucht nach wie vor auch die historisch-hermeneutischen Fähigkeiten, über die wir klassischerweise verfügen, und zwar nicht erst, wenn es um die Interpretation von Ergebnissen geht, sondern schon

1 Fotis Jannidis, Hubertus Kohle und Malte Rehbein, Digital Humanities. Eine Einführung. Stuttgart 2017.

dort, wo beispielsweise Texte für Korpora ausgewählt und nach nachvollziehbaren Standards vereinheitlicht werden müssen. Notwendig sind aber auch bestimmte Fähigkeiten mathematisch-naturwissenschaftlicher Prägung, etwa aus den Bereichen der Statistik, der Logik empirischen Denkens und des Programmierens. Jeden Monat gehen neue digitale Analysetools an den Start, die nicht schwieriger zu bedienen sind als ein Textverarbeitungsprogramm und deren Benutzerfreundlichkeit den Zugang sehr leicht macht. Über den dadurch ersparten Aufwand kann man sich jedoch leicht täuschen. Die Visualisierungen, für deren rasche Ausgabe viele Tools ausgelegt sind, bleiben immer voraussetzungsreich und erklärungsbedürftig. Und zu erklären sind sie nur dann, wenn genau angegeben werden kann, wie sie zustandekommen. Andernfalls werden Scheinevidenzen produziert. Wer den unverzichtbaren Blick unter die Benutzeroberfläche riskiert und nachvollzieht, auf Grundlage welcher abstrakten Repräsentation der Gegenstände was gezählt, in Relation gesetzt und berechnet worden ist, für den gibt es im Grunde keine einfachen Tools mehr.

In der nächsten Zeit wird es darauf ankommen, dieses für einen Innovationsbereich ganz normale Zugangsdilemma nicht so aufzulösen, dass sich das damit verbundene Kompetenzproblem verschärft. Noch immer dominieren gerade in der Verbundforschung Tandemprojekte aus Informatik und Geisteswissenschaft, die als Brückentechnologie notwendig waren, um die Digital Humanities aufzuleisen. Inzwischen aber zeichnet sich ab, dass diese Struktur eine Arbeitsweise fördert, in der die Geisteswissenschaften *content* bereitstellen, während die Informatik die Methodenanwendung bestimmt. Mancher guten Einzeluntersuchung zum Trotz können die so erreichten Ergebnisse nicht nachhaltig sein, weil dazu transparente Operationalisierungen gehören, die sich von Beginn an aus gemeinsamen Fragen ableiten. Diese Fragen zu formulieren ist eine Aufgabe für Geisteswissenschaften, die sich den Problemen, die digitale Methoden in ihren Bereichen aufwerfen, stellen wollen.

Gegenüber der Idee von separaten Digital Humanities kann ein in den bestehenden Fächern angesiedeltes digitales Programm von Vorteil sein, wenn es darum geht, Analysen so auf vorhandene Forschungsprobleme auszurichten, dass der Nutzen der erzielten Ergebnisse sich unmittelbar kommuniziert. Dabei geraten Wissenschaftlerinnen und Wissenschaftler unvermeidlich in Situationen, die ihre Kompetenzen überfordern. Für den Umgang damit empfiehlt sich als Faustregel, theoriebasierten Untersuchungen gegenüber einem explorativen Vorgehen den Vorzug zu geben. Selbstverständlich ist Exploration immer wieder notwendig, etwa wenn neue Tools getestet werden oder innerhalb einer bekannten Versuchsanordnung der Einfluss bestimmter Einstellungen auf das Ergebnis überprüft wird. Weil wir aber wissen, dass nicht erst der Einsatz von

Analysesoftware, sondern schon die Art und Weise, wie die Gegenstände für die Analyse präpariert werden, mitbestimmt, was am Ende herauskommt, sollten wir in jedem Untersuchungsschritt begründen können, warum gerade so und nicht anders verfahren wird. Das fällt viel leichter, wenn wir ausgehend von der Forschungsfrage eine dafür angemessene Operationalisierung wählen und versuchen, den Einsatz digitaler Methoden neben der zwangsläufigen Rücksicht auf Verfügbarkeit und individuelle Kompetenz aus unserer Frage abzuleiten. Explorative Analysen laufen immer Gefahr, solange zu explorieren, bis das Ergebnis gefällt. Theoriebasierte Studien verfügen über einen Reflexionsanker, der es im Erfolgsfall erlaubt, einem Ergebnis auch tatsächlich Erklärungslasten anzutragen, und uns andernfalls in die Lage versetzt, aus dem Scheitern fachlich zu lernen.

Zu den Zumutungen der Digitalen Geisteswissenschaften gehört ohne Zweifel, dass sie sich nicht wie jeder andere *turn* mitvollziehen lassen. Man kann sich durch den akademischen Buchmarkt informieren lassen, aber um digitale Methoden richtig einsetzen zu können, reicht es nicht, seinen Moretti gelesen zu haben. Wie unterschiedlich die Wissenschaftskulturen bereits geworden sind, zeigt sich in den Publikationsstrategien. Während in den Digital Humanities die Konferenzen und Zeitschriften den Takt angeben, ist die Mehrheit der Geisteswissenschaften Buchkultur geblieben. Solche habituell einschneidenden Entwicklungen begünstigen einseitige Wahrnehmungen, die Tradition und Digitalisierung zum kulturkonservativen Gegensatzpaar aufbauen und damit einer mit dem *cultural turn* für überwunden gehaltenen Denkweise den Weg ebnen, die es immer schwieriger macht, voneinander zu lernen. Inzwischen befinden wir uns in einem Stadium der institutionellen Ausdifferenzierung, wo es nicht mehr selbstverständlich ist, dass man wissenschaftlich miteinander im Gespräch bleibt.

Unter solchen Umständen ist es Sache der Forschungspraxis, unnötige Dichotomien aufzulösen. Beim Einsatz digitaler Methoden lassen sie sich durch ein Denken in *trade-offs* am besten überwinden, zum Beispiel in der literaturwissenschaftlichen Korpusanalyse. Korpusanalysen lassen sich theoretisch als »strukturalistische Tätigkeit«² im Sinne von Roland Barthes beschreiben: Sie zerlegen die zum Korpus gehörigen Texte zunächst in einzelne Einheiten, um daraus mit dem Ziel struktureller Erkenntnis etwas Neues zusammenzusetzen. Barthes nennt das »Simulacrum«,³ wir sprechen heute von Modellbildung. In einem Zwischenschritt zählen wir die diskreten Einheiten, um aus diesen Daten unser Modell zu errechnen. Barthes Begriff wirkt heute aufgesetzt und benennt doch etwas Wesentliches. Am Anfang jeder quantitativen Korpusanalyse steht die Ent-

2 Roland Barthes, »Die strukturalistische Tätigkeit«, in: *Texte zur Literaturtheorie der Gegenwart*, 215-223. Stuttgart 2003, S. 215-223.

3 Ebd. S. 217 und S. 221.

scheidung für eine bestimmte abstrakte Repräsentation von Text. Gerechnet wird nicht mit den Korpus-texten selbst, sondern mit einem numerischen Stellvertreter, der sie vergleichbar macht. Diese ursprüngliche Abstraktion wird häufig übersehen.⁴ Unabhängig davon, ob ich Worthäufigkeiten (Distanzmaß), die Kontinuität des Wortgebrauchs (Kontrastmaß) oder die Wahrscheinlichkeit semantischer Muster (Topic-Gewicht) zugrunde lege, in jedem Fall werden die Texte über eine festgelegte Anzahl entsprechender Werte repräsentiert und verglichen. Der große Vorteil dieser Art des Vergleichens liegt auf der Hand: Ich kann praktisch unbegrenzt viele Texte nebeneinanderhalten und habe sofort ein Ergebnis. Je nach Fragestellung können die Werke aus einer Epoche oder Gattung oder über Epochen- und Gattungsgrenzen hinweg sein. Die Unterschiede weiblichen und männlichen Schreibens lassen sich ebenso ermitteln wie jene zwischen der literarischen und nicht-literarischen Welt einer Zeit. Die einzige, in der Praxis allerdings eklatante Einschränkung für Untersuchungen ergibt sich daraus, ob vergleichbare, d. h. nach den gleichen Standards als digitale Volltexte aufbereitete Korpora zur Verfügung stehen oder mit vertretbarem Aufwand erschlossen werden können.

Wie alles, womit wir Erkenntnis gewinnen, hat auch diese Methode ihren Preis. Wir bezahlen die umfassende Vergleichbarkeit mit dem Verlust von Kontext in derartigen Analysen. Der ist bei den worthäufigkeitsbasierten *bag of words*-Verfahren, wo keine Zusammenhänge im Satz, im Text oder gar im geläufigen Sinn eines historischen Kontextes berücksichtigt werden, mit Händen zu greifen – vor allem im Vergleich mit dem anderen Extrem des Close Reading. Wer sich hermeneutisch in einen Text vertieft, findet viele Zusammenhänge und läuft doch Gefahr, Ergebnisse zu produzieren, die nicht gut vergleichbar sind. Vergleichbarkeit und Kontextsensitivität gegenüber scheinen wir mit einem klassischen *trade-off* konfrontiert zu sein. Anders als Dichotomien sind *trade-offs* jedoch gestaltbar, und schon ein zweiter Blick zeigt uns ein differenziertes Bild. Auch auf der Basis abstrakter numerischer Repräsentationen von Text lässt sich historischer Kontext einholen, wenn wir etwa Metadaten zu Korpus-texten nutzen, um Vergleiche mit Subsets anzustellen. Analysen mit der Variable Gender tun genau das. In der automatischen Sprachverarbeitung werden Modelle zum *word embedding* entwickelt, die die semantische und grammatische Sensitivität verbessern. Und selbst die Ergebnisse rein stochastischer Semantik wie beim Topic Modeling können mit digitalen Konkordanzanalysen in konkreten Satzzusammenhängen nachvollzogen werden. Auf der anderen Seite verfügen wir in der Literaturwissenschaft seit jeher über resiliente Methoden, die uns auch beim Close Reading davor bewahren, in assoziativen Kontexten unterzugehen. Bei jeder Auto-

4 Ulrik Brandes, Garry Robins, Ann McCraime und Stanley Wasserman, »What is network science?« *Network Science* 1 (1), 2013, S. 1-15, S. 3 f.

rin und jedem Autor kristallisieren sich bestimmte Stellen heraus, die die Forschung für relevanter hält als andere.⁵ Je nach Konjunktur können dann bislang unbeachtete Passagen in Innovationsschüben erschlossen werden. Die hermeneutische Stellenlektüre ist ebenso ein zuverlässiges Instrument der Kontextverknappung wie der Apparat von Sekundärliteratur, mit dem Wissenschaftlerinnen und Wissenschaftler in der Regel die Erschließung neuer Forschungsgebiete beginnen, die Möglichkeiten neuer Gedanken moderiert.

In Zukunft wird die nachhaltige Entwicklung der Digitalen Geisteswissenschaften entscheidend von der Ausbildung des Nachwuchses abhängen. Bislang scheinen insbesondere die digitalen Master-Programme für Studierende hochattraktiv zu sein. Auch die Bachelor-Studiengänge laufen gut, machen aber deutlich, dass der Übergang an die Universität für viele Schülerinnen und Schüler, die geisteswissenschaftlich orientiert sind, durch das doppelte Anforderungsprofil digitaler Methoden nicht einfach ist. Insgesamt sind die Studierendenzahlen wie überall in den Geistes- und Naturwissenschaften jenseits des Lehramts überschaubar. Dass über digitale Programme zusätzliche Studierende für die Geisteswissenschaften gewonnen werden, ist nicht zu erwarten. Wissenschaftspolitisch klug wäre eine zeitnahe Aufnahme digitaler Analysemethoden in die Curricula der Lehramtsausbildung. Wir finden hier eine qualifizierte Kohorte von Studierenden, die neben Deutsch auch Naturwissenschaften oder Mathematik belegen. Sie könnte zum Schlüssel werden, um zu verhindern, dass unsere Kinder künftig weiter im langen 19. Jahrhundert der zwei getrennten Wissenschaftskulturen aufwachsen. Die kritische Kompetenz, die die Geisteswissenschaften in der Auseinandersetzung mit Texten und anderen Medien schulen, muss auf den Bereich von Daten ausgeweitet werden, wenn Bildung die Mündigkeit von Bürgerinnen und Bürger zum Ziel hat. Die Schulen müssen in die Lage versetzt werden, dafür den Grundstein zu legen.

5 Die Idee verdanke ich Steffen Martus, der im Rahmen des DFG-Schwerpunktprogramms »Computational Literary Studies« die Stellen-Präferenz von Laien und Wissenschaftlerinnen und Wissenschaftlern vergleicht.