

## Digital Humanities auf dem Weg zu einer Wissenschaftsmethodik: Transparenz und Fehlerkultur\*

Charlotte Schubert

**Abstract:** The methods of the Digital Humanities have been subject to massive criticism for quite some time. There are two main accusations that are made again and again: On the one hand, the Digital Humanities do not lead to new results, but present the familiar in a different guise. On the other hand, the methods of the Digital Humanities even lead to false results. Furthermore, it is concluded that the reproducibility/replication and thus the scientific soundness of the results is questionable. This article deals with this issue of (assumed or actual) incorrectness by analyzing this accusation and presents a proposal for a critical approach to errors that can secure the Digital Humanities their place in scientific methodology.

### I. Keine neuen Ergebnisse, aber dafür Fehlerhaftigkeit?<sup>1</sup>

An sich sollte die Definition der Digitalen Geisteswissenschaften, wie sie etwa Sybille Krämer ausgehend von dem Begriff der ‚Verdatung‘ jüngst zusammengefaßt hat, für sich sprechen können:<sup>2</sup> Krämer hat den Neuigkeitswert von Erkenntnissen in den Digital Humanities daran gemessen, daß Erkenntnisse und Einsichten zu gewinnen sind, die mit nichtdigitalen Methoden entweder ganz schwierig oder überhaupt nicht zu erreichen sind.<sup>3</sup> Der Vorwurf gegenüber den Digital Humanities ist aber nun, sie führten nicht zu neuen Ergebnissen, sondern würden lediglich Bekanntes in anderem Gewand präsentieren.<sup>4</sup> Dieser Vorwurf läßt sich jedoch leicht entkräften. Denn ob ein Ergebnis neu oder altbekannt ist, ist unter Fachkollegen recht schnell zu klären. Die Neuigkeitsanforderung, die die Kritiker an die Digital Humanities richten, hat sich längst als erfüllbar erwiesen. In diversen Projekten und Publikationen sind dazu

---

\* Das hier behandelte Thema geht auf einen Vortrag zurück, den ich am 21.11.2019 im Rahmen des Symposiums „Wozu Digitale Geisteswissenschaften. Innovationen, Revisionen, Binnenkonflikte“ (DFG-geförderte Symposienreihe „Digitalität in den Geisteswissenschaften“) gehalten habe. Das Symposium fand an der Leuphana Universität Lüneburg unter der Leitung von Sybille Krämer, Claus Pias und Martin Huber statt. An dieser Stelle möchte ich Sybille Krämer für ihre vielen Anregungen und weiterführenden Hinweise zu dem Thema sehr herzlich danken.

1 Thomas (2016), 525: „Yet, paradoxically, the 20-year surge in the digital humanities – from 1993 to 2013 – has produced relatively little interpretive or argumentative scholarship. In this first phase of the digital humanities, scholars produced innovative and sophisticated hybrid works of scholarship, blending archives, tools, commentaries, data collections, and visualizations. For the most part in the disciplines, however, few of these works have been reviewed or critiqued. Because the disciplines expect interpretation, argument, and criticism, it could be argued that digital humanists have not produced enough digital interpretive scholarship, and what we have produced has not been absorbed into the scholarly disciplines.“ Vgl. ebf. van Zundert (2016), 331–347.

2 Krämer (2018), 5–11, online: <https://doi.org/10.11588/dco.2017.0.48490> (17.7.2021).

3 Krämer (2018), 6.

4 Da (2019a) in <https://www.chronicle.com/article/The-Digital-Humanities-Debaacle/245986> (17.7.2021). Vgl. Da (2019b), 601–639 und mit einer Übersicht zu den Reaktionen auf den Beitrag von Nan Z. Da vgl. Schubert (2019a) <https://doi.org/10.11588/dco.2019.2.72004> (17.7.2021).

Ergebnisse vorgelegt worden, die von der Autorschaftszuweisung<sup>5</sup> über Text- und Diskursanalysen<sup>6</sup> bis hin zu Simulationsexperimenten reichen.<sup>7</sup>

Die Neuigkeitsanforderung hat jedoch noch eine zweite Seite. Ein in den Naturwissenschaften gängiges methodisches Vorgehen, die Bestätigung eines Ergebnisses mit einer zweiten, methodisch anderen, von der ersten unabhängigen Methode, ist in den Geisteswissenschaften bisher nicht üblich. Die Bestätigung eines bereits bekannten Ergebnisses, das mit den Methoden der klassischen, historisch-kritischen Analyse erzielt wurde, mit einer anderen Methode, nämlich aus dem Digital-Humanities-Bereich, die Erkenntnisse auf einem unabhängig von der klassisch-hermeneutischen Methode operierenden Weg bestätigen kann, bietet den Geisteswissenschaften eine völlig neue Perspektive. Sie kann den Geisteswissenschaften einen Weg eröffnen, über das Belegen zu einer dem Beweisen im naturwissenschaftlichen Sinn gleichartigen und intersubjektiv stärkeren Bestätigungspraxis zu kommen. Ein solches Vorgehen – das Bestätigen eines bekannten Ergebnisses durch Methoden der Digital Humanities –, führt nicht zu neuen Erkenntnissen. Es handelt sich jedoch um einen Proof of Concept für die Digital Humanities und deren Beitrag zu den Geisteswissenschaften. Die Bedeutung dieses Vorgehens spielt derzeit eine größere Rolle in den Digital Humanities als in den mit klassischen Methoden vorgehenden geisteswissenschaftlichen Fächern. In den momentanen fachlichen Diskussionen über Sinn und Nutzen der Digital Humanities auf der Seite der Geisteswissenschaften wird die Möglichkeit, die Bestätigungsfunktion des Proof of Concept durch die Digital Humanities zu nutzen, noch nicht oft genutzt. Um diesem Verfahren den Stellenwert zuzuweisen, der ihm einen methodischen Rang als eigene Methode verleiht, sind einige anspruchsvolle Voraussetzungen zu erfüllen:

- Ein Proof of Concept, der durch die Entwicklung und/oder Anwendung der Methoden der Digital Humanities erzielt worden ist, muß sich auf fachlich relevante Ergebnisse oder Diskussionen beziehen.
- Ein Proof of Concept muß auf einem methodisch reflektierten Konzept beruhen, das transparent und nachvollziehbar ist.
- Ein Proof of Concept muß skalierbar sein, so daß die spezifischen, quantitativen Ausweitungsmöglichkeiten, die die Transformation in die ‚Verdatung‘ mit sich bringt, nutzbar gemacht werden.<sup>8</sup>

Unter diesen Voraussetzungen kann der Proof of Concept für beide Seiten ein Gewinn werden: Für die Geisteswissenschaften als eine neue und erweiterte Bestätigungspraxis und für die Digital Humanities als Anerkennung ihres methodischen Beitrags in den Geisteswissenschaften. Dies ist keineswegs Zukunftsmusik, sondern bereits in der wissenschaftlichen Praxis erprobt und mit durch die Fachcommunity akzeptierten Beiträgen zu belegen. Insbesondere die klassischen Altertumswissenschaften, die auf 2500 Jahre Erfahrung in Wandel und Möglichkeiten des kritischen Umgangs mit Texten zurückblicken können, sind prädestiniert, um hier methodisch beispielgebend zu wirken. *Pars pro toto* sei dazu auf die von Werner Riess und seinen Mitarbeitern publizierten Beiträge zu den Machbarkeitsstudien für ERIS verwiesen.<sup>9</sup>

---

5 Der spektakulärste Fall einer erfolgreichen Autorschaftszuweisung ist die Identifizierung des Autors Robert Galbraith als J.K. Rowling durch Juola (2015), 101–113.

6 Z.B. Schubert / Weiß (2015), 447–471 und Schubert (2018), 79–92.

7 Z.B. Warnking (2015) und ders. (2016), 45–90; Schäfer (2019), 22–33, online: <https://doi.org/10.11588/dco.2019.1.60564> (17.7.2021).

8 Riess (2019), 4–27 und online: <https://doi.org/10.11588/dco.2019.2.72018> (17.7.2021) ist ein überzeugendes Beispiel für ein Proof of Concept. Vgl. auch Riess 2020 (wie in Anm. 9).

9 Riess (2020), 445–473; Diemke (2020), 57–74 und online: <https://journals.ub.uni-heidelberg.de/index.php/dco/article/view/77663/71565> (17.7.2021).

Der zweite Vorwurf, die Methoden der Digital Humanities führten zu falschen bzw. nicht validen Ergebnissen, ist der schwerwiegendere von den beiden Vorwürfen, mit denen sich die Digital Humanities auseinandersetzen müssen. Dieser Vorwurf ist nicht ganz neu, aber er ist in den letzten Jahren mit neuer und erheblicher Schärfe in der Debatte erhoben worden. Die neueren Beiträge, vor allem der Artikel von Nan Z. Da vom 27.3.2019 in „The Chronicle of Higher Education“ unter der Überschrift „The Digital Humanities Debacle. Computational methods repeatedly come up short“ ist, wie das Echo sehr deutlich zeigt, als ein Frontalangriff auf die Digital Humanities aufgefaßt worden.<sup>10</sup> Der Vorwurf lautet, eine spezifische Fehlerhaftigkeit der Digital Humanities sei deren Charakteristikum.<sup>11</sup> Aber nicht nur die Qualität der Daten, die mittlerweile als Problem- und Arbeitsfeld ausreichend thematisiert worden ist, sondern vor allem die Arbeitsweisen, mit der die zugrunde gelegten Daten erschlossen werden, lassen sehr schnell erkennen, wo das Problem liegt. Dies soll hier anhand zweier Praktiken – der Verwendung von Stoppwortlisten und von Metadaten – *pars pro toto* erläutert werden.

Sowohl die Entwicklung als auch die kompetente Anwendung einer algorithmen-gestützten Auswertung eines oder mehrerer Texte setzt eine Reflexion auf den Zusammenhang zwischen Wortlaut und Bedeutung sprachlicher Äußerungen voraus. Diese, für Textwissenschaftler selbstverständliche, Voraussetzung wird allerdings in vielen Bereichen der mit Texten arbeitenden Digital Humanities mißachtet.

### Die Verwendung von Stoppwörtern

Als Stoppwörter werden Wörter bezeichnet, die bei einer Volltextindexierung nicht beachtet werden, da sie sehr häufig auftreten und ihnen für gewöhnlich keine Relevanz für die Erfassung des Dokumenteninhalts zugebilligt wird. Die Stoppwörter befinden sich in der Regel auf einer festen oder berechneten Liste, werden bei der Textverarbeitung aus dem Text entfernt und für die Textanalyse nicht indexiert. Allen Stoppwörtern ist gemeinsam, daß sie vor allem grammatikalische/syntaktische Funktionen übernehmen und daher – so die gängige Meinung – keine Rückschlüsse auf den Inhalt des Dokumentes zulassen. Eine weitere Gemeinsamkeit ist ihre Häufigkeit: Sie treten in jedem Dokument sehr oft auf und kommen in sehr vielen Dokumenten vor, wodurch sie, wenn man sie einbezöge, bei der Erschließung der Dokumente einen hohen Aufwand verursachen würden.

---

10 Da (2019a): <https://www.chronicle.com/article/The-Digital-Humanities-Debacle/245986> (17.7.2021). Vgl. Da. (2019b), 601–639 und Kirby (2019): <https://doi.org/10.29173/iq926> (17.7.2021) sowie Rizvi (2018), 401–418 und online: <http://doi.org/10.1093/llc/fqy038> (17.7.2021). Dazu Schubert (2019a), 1–3 und online: <https://doi.org/10.11588/dco.2019.2.72004> (17.7.2021).

11 Da (2019a) in <https://www.chronicle.com/article/The-Digital-Humanities-Debacle/245986> (17.7.2021): „Not only has this branch of the digital humanities generated bad literary criticism, but it tends to lack quantitative rigor. Its findings are either banal or, if interesting, not statistically robust. The problem appears to be structural. In order to produce nuanced and sophisticated literary criticism, CLS [sc. Computational Literary Studies, C.S.] must interpret statistical analysis against its true purpose; conversely, to stay true to the capacities of quantitative analysis, practitioners of CLS must treat literary data in vastly reductive ways, ignoring everything we know about interpretation, culture, and history. Literary objects are too few, and too complex, to respond interestingly to computational interpretation — not mathematically complex, but complex with respect to meaning, which is in turn activated by the quality of thought, experience, and writing that attends it.“ Eine Zusammenstellung verschiedener Antworten auf Nan Z. Da findet sich hier: [https://demo.hedgedoc.org/s/rJ\\_YoK\\_cH](https://demo.hedgedoc.org/s/rJ_YoK_cH) (17.7.2021).

über	können	wie	von	der
nach	sollen	falls	auf	ihre
alle	könnten	in	nur	ihnen
auch	brauchen	darein	oder	darauf
an	dürfen	ist	irgendeiner	da
und	darf	es	unser	diese
noch	jeder	seinem	aus	sie

**Abb. 1: Beispiel einer Stoppwortliste (aus: <https://wiki.infowiss.net/Stopwort>, abgerufen am 17.7.2021).**

Das Beispiel zeigt einen Ausschnitt aus einer Stoppwortliste für das Deutsche mit einem Teil der üblicherweise als Stoppwörter betrachteten Wörter. Aber nun sind gerade Wort wie „sein“, „sollen“, „dürfen“ bspw. im philosophischen Bereich sinntragende Wörter. So setzt sich etwa Platon in dem Dialog *Sophistes* mit der Frage auseinander, ob man das Nicht-Seiende – da es bekanntlich nicht ist – sagen kann.<sup>12</sup> Folgende Formulierung ist eine zentrale Passage in diesem Dialog:

τὸ μὴ ὄν εἶναι (Plat. *Sophistes* 237 a3f.)  
 „daß das Nicht-Seiende sei“

Diese Formulierung besteht nun aus lauter Stoppwörtern (Artikel, Negationspartikel, Kopula) und ist daher mit den üblichen Verfahren, die Stoppwortlisten verwenden, gar nicht auffindbar! Das Problem ist hier die Annahme, daß die Stoppworte wie die Artikel und Partikel, auch εἶναι in der geläufigen Verwendung als Kopula, eine rein grammatische Funktion besitzen und nichts zum Sinn beitragen.

Das Problem ist lösbar: Wenn man Platons Werk kennt, dann ist auch die Bedeutung der Kopula εἶναι geläufig. Daher muß – und zwar von Beginn einer Konzeptionierung, Modellierung oder Analyse an – darauf geachtet werden, daß man für entsprechende Textgattungen eben nicht mit automatisch erzeugten Stoppwortlisten arbeitet, sondern sie entweder fach- und sachgerecht erarbeitet oder vollständig beiseite läßt. Die in der Verwendung von Stoppwortlisten implizierten Fehler wären sehr leicht durch an- und abschaltbare bzw. auch anpaßbare Stoppwortlisten zu beseitigen, eine Praxis, die jedoch selten bis gar nicht eingesetzt wird.<sup>13</sup>

## Die Verwendung von Metadaten

Neben der Verwendung von Stoppwörtern ist eine weitere, nicht nur allgemein verbreitete, sondern mittlerweile institutionalisierte Praxis, der Gebrauch und die Auswertung von Metadaten.<sup>14</sup> Im Bereich der bibliographischen Metadaten und insbesondere im Data Profiling in der Informatik werden Metadaten als Klassifikationen eher unhinterfragt verwendet. Insbesondere das Data Profiling ist ein Prozeß, der

12 Das nachfolgende Beispiel ist entnommen aus dem Beitrag von Rautenberg (2019), 111–123, online: <https://doi.org/10.11588/propylaeum.451> (17.7.2021).

13 Vgl. dazu die Webportale mit an- und abschaltbaren Stoppwortlisten zu Digital Plato <https://digital-plato.org/> und [www.eaqua.net](http://www.eaqua.net) (17.7.2021). In der Paraphrasensuche von Digital Plato ist die Stoppwortliste dynamisch anpaßbar.

14 Das nachfolgende Beispiel ist eine gekürzte und aktualisierte Version von Schubert (2019b), 4–21 und online: <https://doi.org/10.11588/dco.2019.1.59356> (17.7.2021).

Metadaten aus Datenbanken analysiert, um deren Metadaten gegebenenfalls zu korrigieren, der jedoch automatisch durchgeführt wird.

Grundsätzlich böte das Arbeiten mit Textdaten die Chance, verschiedene Datentypen wie die Daten historischer Ereignisse, Werkzuordnungen, geographische Informationen u.v.a.m. auf wissenschaftlicher Basis zu klassifizieren. Die Komplexität von Textdaten ist jedoch oft mit Unsicherheiten wie etwa unklaren Lebensdaten und -orten oder umstrittenen Werkidentifikationen verbunden. Daraus ergibt sich, daß die einzelnen Datentypen (chronologische Einordnung, geographische Verteilungen, Werkzuordnungen, lexikographische Anordnung) wiederum eine Vielzahl von Einzelaspekten umfassen können, d.h. daß diese Unsicherheiten bei der Aufbereitung von Textdaten einbezogen werden müßten. Dies wäre durchaus möglich, allerdings läßt die heute übliche Praxis der Metadatenauszeichnung nicht erkennen, daß die unterschiedlichen wissenschaftssystematischen und wissenschaftshistorisch-methodischen Positionen in der Praxis berücksichtigt werden.

Die diversen Interoperabilitätsbemühungen im Hinblick auf die unterschiedlichen Metadatenstandards und -ontologien verweisen an sich schon auf das Dilemma:<sup>15</sup> Die Gemeinsame Normdatei (GND) der Deutschen Nationalbibliothek, die von allen deutschen Bibliotheken und insbesondere zur Vernetzung von deren Informationsressourcen verwendet wird, führt daher für die Deutsche Nationalbibliothek, alle deutschsprachigen Bibliotheksverbände mit den angeschlossenen Bibliotheken, die Zeitschriftendatenbank (ZDB) und zahlreiche weitere Einrichtungen die Metadaten für Personen, Körperschaften, Konferenzen, Geographika, Sachschlagwörter und Werktitel, die vor allem zur Katalogisierung von Literatur in Bibliotheken dienen.<sup>16</sup>

Allerdings tritt hier eine Kontingenz zutage, die von denen der lexikographischen Ordnungsverfahren her bekannt sein müßte: Die Voraussetzungen von Etikettierungen oder Labels oder Indexeinträgen sind immer begriffsbezogene Datenkategorien, die aus fachspezifischen Kontexten stammen, die durch automatisierte Extraktion nicht zuverlässig erfaßt werden können. Gleichwohl wird der Anspruch erhoben, daß bibliographische Metadaten von den fachspezifischen Voraussetzungen und deren Kontingenzen unabhängig seien, da man bibliographisch-administrative Daten einerseits und inhaltsbeschreibende bzw. fachliche Daten andererseits unterscheiden könnte. Erstere gäben Informationen zur Verwaltung der Daten aus den Publikationen selbst, letztere hingegen beschrieben einzelne Aspekte oder Datensätze genauer, und inhaltsbeschreibende Metadaten seien grundsätzlich – disziplinspezifisch – unterschiedlich aufgebaut.

Welche Fallen jedoch trotz dieser an sich nicht falschen Unterscheidung der Metadaten in dem Verfahren lauern, ist bereits früher in dieser Zeitschrift ausführlich dargestellt worden.<sup>17</sup> Hier soll dies daher nur kurz rekapituliert und das damalige Ergebnis mit einer weiteren Methode der digitalen Textanalyse untermauert werden.

Von dem Neuplatoniker Iamblich von Chalkis (Mitte des 3. Jh. n. Chr. – ca. 320/325 n. Chr.), ist ein zehnbändiges Werk über die Lehre des Pythagoras erhalten. Dessen zweiter Band ist ein sog. *Protreptikos*, eine Mahnschrift in Form des Aufrufes, Philosophie zu betreiben. Diese Gattung war in der Antike sehr beliebt und stand zur Zeit des Iamblich bereits in einer langen Tradition berühmter Vorgänger. Insofern überrascht es nicht, in Iamblichs *Protreptikos* zahlreiche Zitate aus Werken anderer

15 Die Angaben zu den verwendeten Metadatenformaten und den internationalen Interoperabilitätsbemühungen der Deutschen Nationalbibliothek: [https://www.dnb.de/DE/Professionell/Standardisierung/Standards/standards\\_node.html#%5BAnkerMARC21%5D](https://www.dnb.de/DE/Professionell/Standardisierung/Standards/standards_node.html#%5BAnkerMARC21%5D) (17.7.2021).

16 [https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd\\_node.html](https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd_node.html) (17.7.2021).

17 Die Ergebnisse zu Iamblichs *Protreptikos* sind in Schubert (2017), 17–48 publiziert, online: <https://books.ub.uni-heidelberg.de/propylaeum/catalog/book/257> (17.7.2021).

Autoren zu finden. Aus diesem Befund ist die These entwickelt worden, daß sich in dem iamblichischen *Protreptikos* auch Teile des verlorenen, aber in der Antike außerordentlich berühmten aristotelischen *Protreptikos* finden ließen. Daher haben verschiedene Altphilologen den Versuch unternommen, aus dem iamblichischen *Protreptikos* den aristotelischen *Protreptikos* zu rekonstruieren und verschiedenste Editionen dieses – eigentlich verlorenen, also nichtexistenten – aristotelischen *Protreptikos* publiziert. Diese Editionen sind größtenteils aus Textpassagen des iamblichischen *Protreptikos* zusammengesetzt und werden in den Bibliothekskatalogen unter dem Namen des Autors Aristoteles rubriziert. So finden sie sich sowohl in den digitalen Textdatenbanken (z.B. in der altgriechischen Textdatenbank TLG<sup>18</sup>) wie auch in den Metadaten-Einträgen der Bibliotheksdatenbanken. Der entsprechende Eintrag der GND führt diesen „Protreptikos des Aristoteles“ folgendermaßen auf:

GND	
<b>Link zu diesem Datensatz</b>	<a href="http://d-nb.info/gnd/4305003-7">http://d-nb.info/gnd/4305003-7</a>
<b>Verfasser/Urheber</b>	Aristoteles
<b>Titel des Werkes</b>	Protrepticus
<b>Andere Titel</b>	Der Protreptikos (ÖB-Alternative) Protreptikos
<b>Quelle</b>	Thes. ling. Graec.
<b>Erläuterungen</b>	Definition: Fragmentarisch erhalten
<b>Zeit</b>	400 - 301 v. Chr. (UDK-Zeitcode v03)
<b>Land</b>	Griechenland (Altertum) (XS)
<b>Sprache(n)</b>	Griechisch (grc)
<b>Systematik</b>	4.7p Personen zu Philosophie
<b>Typ</b>	Werk (wit)
<b>Thema in</b>	<p><a href="#">3 Publikationen</a></p> <ol style="list-style-type: none"> <li>1. <i>Der Protreptikos des Aristoteles</i> Aristoteles. - Frankfurt, M. : Klostermann, 2014, 3., unveränd. Aufl.</li> <li>2. <i>Protreptikos</i> Aristoteles. - Darmstadt : Wiss. Buchges., [Abt. Verl.], 2005</li> <li>3. ...</li> </ol>

Abb. 2: <http://d-nb.info/gnd/4305003-7> (19.7.2021).

Die früher bereits publizierte Visualisierung auf der Basis der Metadaten und der N-Gramm-Analyse hat gezeigt,<sup>19</sup> wie die Metadaten als Kernstück der bibliographischen Klassifizierung zu falschen Ergebnissen führen müssen, da aufgrund der bibliographischen Angaben natürlich die Aristoteles-Zuweisung hervorsteht.<sup>20</sup> Ohne Zugrundelegung der vorgegebenen Autorzuweisung für den iamblichischen *Protreptikos* bzw. ohne Einbeziehung der rekonstruierten Editionen ergibt sich jedoch, daß der Text des Iamblich keinerlei textuelle Verbindung zu dem Werk des Aristoteles hat, jedoch sehr maßgebliche zu

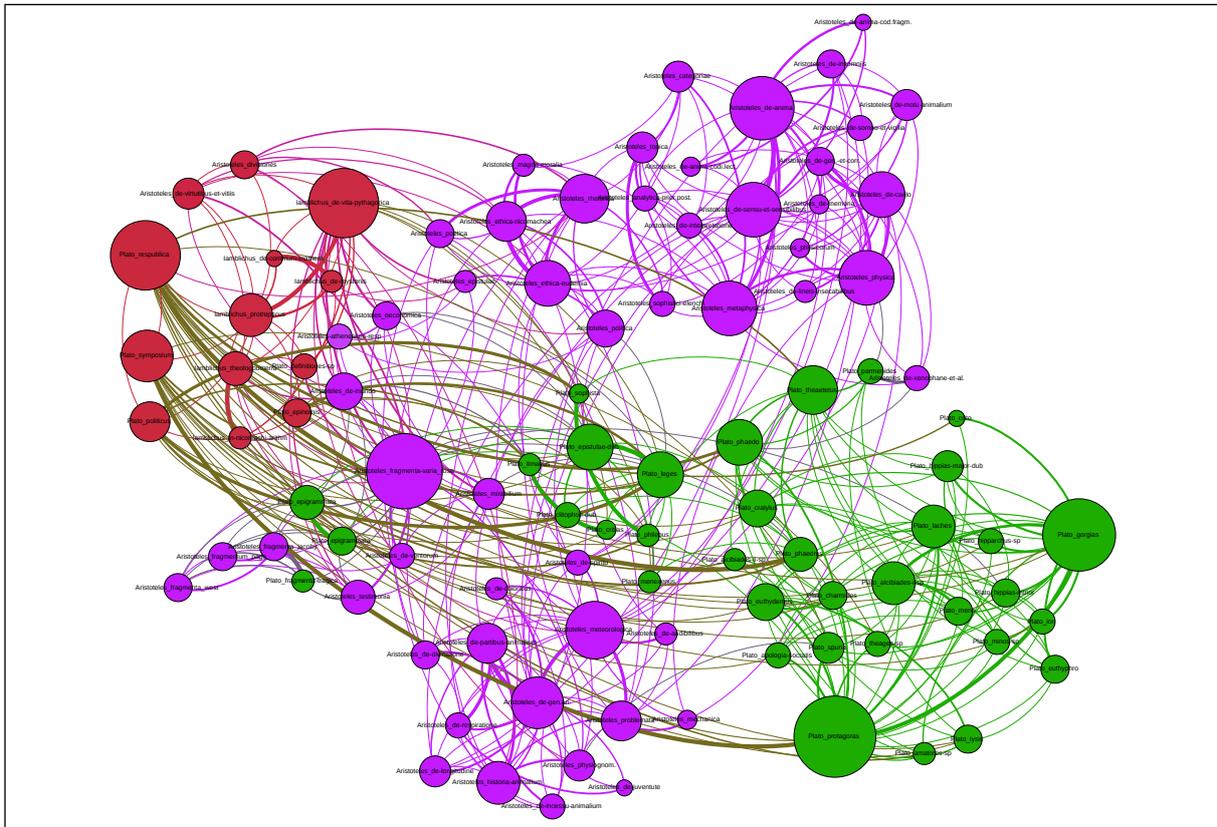
18 <http://stephanus.tlg.uci.edu> (17.7.2021).

19 Vgl. Schubert (2019b), 4–21 und online: <https://doi.org/10.11588/dco.2019.1.59356> (17.7.2021).

Zugrunde liegt die Extraktion der Parallelstellensuche aus eAQUA (N-Gramm-Analyse auf der Basis von fünf exakt gleichen Wort-N-Grammen), deren Metadaten mit Hilfe des Netzwerkvisualisierungsprogramms Gephi (<https://gephi.org/> [17.7.2021]) visualisiert wurden. Die verwendeten Metadaten sind Autor- und Werkname – zwei Metadatenkategorien, die nicht nur das grundlegende Gerüst aller Metadatenmodelle sind, die sich auf Textdaten beziehen, sondern heute praktisch immer unhinterfragt eingesetzt werden.

20 Dazu Abb. 4 in Schubert (2019b), 4–21 und online: <https://doi.org/10.11588/dco.2019.1.59356> (17.7.2021).

demjenigen Platons.<sup>21</sup> Auch eine Überprüfung des früheren Ergebnisses mit einer anderen Methode, der Buchstaben-Trigramm-Analyse aus der Stilometrie, bestätigt dieses Ergebnis.<sup>22</sup>



**Abb. 3: Visualisierung der Buchstaben-Trigramm-Analyse aus StyloAH in Gephi; rot markiert ist das Netzwerk aus nächsten Nachbarn des Knotens iamblich's *Protreptikos* mit den entsprechenden Verbindungen zwischen den Knoten.**

Die Visualisierung der mit StyloAH durchgeführten Untersuchung eines Textkorpus mit Gephi,<sup>23</sup> das die Werke des Iamblich, Platons und Aristoteles (ohne die moderne Rekonstruktionsedition des *Protreptikos*) in einem Textkorpus zusammenstellt, um Verbindungen zu erkennen, zeigt die Nähe des iamblichischen *Protreptikos* zu Platon, insbesondere zu der platonischen *Politeia*, während für Aristoteles lediglich eine Verbindung zwischen dem iamblichischen *Protreptikos* und der heute als unecht klassifizierten aristotelischen Schrift *De virtutibus et vitiis* sowie der pseudo-aristotelischen Sammlung der *Divisiones* angezeigt wird.<sup>24</sup>

Dieses Ergebnis verweist auf zwei Aspekte: Zum einen wird das frühere Ergebnis der kombinierten Metadaten-/Wort-N-Grammanalyse durch eine buchstabenbasierte N-Grammanalyse erneut bestätigt und erhöht somit die Plausibilität mit Bezug auf die oben genannte Forderung nach dem Einsatz von zwei (oder sogar mehr) unabhängig voneinander operierenden Methoden. Zum anderen kann die Un-

21 Dazu Abb. 6a in Schubert (2019b), 4–21 und online: <https://doi.org/10.11588/dco.2019.1.59356> (17.7.2021) und dies. (2017), 17–48, online: <https://books.ub.uni-heidelberg.de/propylaeum/catalog/book/257?lang=en> (17.7.2021).

22 Diese Methode ist ausführlich und mit Literaturhinweisen beschrieben in: Schubert (2020), 305–327.

23 Die Buchstaben N-Gramme (Buchstaben N=3) wurden mit StyloAH (M. Eders Stylo mit der Programmerweiterung durch H. Kahl: v.0.7.4.5: <https://github.com/ecompsHONgit/stylo> [17.7.2021]) durchgeführt: Die Visualisierung erfolgte mit Gephi 0.9.1 (<https://gephi.org/>) im Layout OpenOrd. Zu OpenOrd: Martin et al. (2011), online: <https://github.com/gephi/gephi/wiki/OpenOrd> (17.7.2021).

24 Zur Unechtheit der Schrift *De virtutibus et vitiis*: Flashar (2004), 207 und 274f. Zu den bei Diogenes Laertius III 80ff. überlieferten *Divisiones*: Flashar (2004), 96f.

tersuchung als Beispiel für einen Weg der Fehleranalyse im Hinblick auf die Entwicklung einer Fehlerkritik dienen, um den Umgang mit Metadaten zu systematisieren. Denn Metadaten sind heute nicht nur strukturgebend, sondern sie sind auch zu strukturvorgebenden Daten geworden. Sie prägen die Informationen und erzeugen so selbst eine Kontextualisierung, wie eben die Aufnahme einer modernen Rekonstruktionsedition in ihrer Weiterwirkung in modernen Bibliotheksdatenbanken zeigt. Demgegenüber steht die Forderung, daß sich auch die Praxis der Metadaten den Bedingungen zu fügen hat, die für jede wissenschaftlich-kritische Praxis gelten: Jedes Metadatenmodell muß theoretisch begründet und kritisch systematisiert werden. Auf einer solchen Grundlage genügt nicht nur eine Visualisierung von Textanalysen dem wissenschaftlichen Anspruch, sondern ist überhaupt erst eine wissenschaftliche Arbeit mit Metadaten möglich. Eine Analyse wie die hier kurz referierte, muß die Grundlage für eine wissenschaftlich seriöse Implementierung von Metadaten sein und diese wiederum ist die Aufgabe einer kritischen Wissenschaftspraxis, die jedoch heutzutage in den bibliographischen Metadaten – insbesondere, wenn sie automatisch erzeugt werden – nicht vorliegt und so ihrerseits zu Fehlern bei der digitalen Auswertung dieser Daten beiträgt.

## II. Auf dem Weg zu einer Praxis konstruktiver Fehlerkultur

Seit einigen Jahren ist zu beobachten – empirisch belegt in einer zeitlichen Korrelation zum Fortschreiten der Digitalisierung – daß die Anzahl zurückgezogener Aufsätze und Publikationen ansteigt: In der Mehrzahl geschehen diese Zurücknahmen auf Grund experimenteller und unbeabsichtigter Fehler, es handelt sich also bei weitem nicht nur um Plagiate. Die Biochemie und Molekularbiologie, Zellbiologie und Onkologie führen das Feld der zurückgenommenen Artikel an. Dazu passt die Beobachtung, daß Journale mit sehr hohem Impact ebenfalls in den vordersten Reihen dieser Rücknahmen zu finden sind.<sup>25</sup> Die Datenbank Retraction Watch gibt einen sehr zeitnahen Überblick der zurückgezogenen Artikel, die nach Disziplinen aufgeschlüsselt sind.<sup>26</sup> Für den Bereich Molecular Biology verzeichnet die Datenbank mit Datum v. 17.7.2021 insg. 127 zurückgezogene Artikel:

Retraction or Other Notices Title/Subject(s)/Journal --- Publisher/Affiliation/Retraction Watch Post URL(s) <i>127 Items(s) Found</i>	Reason(s)	Author(s)	Original Paper Date/PubMedID/DOI	Retraction or Other Notices Date/PubMedID/DOI	Article Type(s) Nature of Notice	Countries/ Paywalled? Notes
The Effects of Vitamin D Supplementation on Signaling Pathway of Inflammation and Oxidative Stress in Diabetic Hemodialysis: A Randomized, Double-Blind, Placebo-Controlled Trial (BLS) Biology - Molecular, (HSC) Medicine - Dentistry; (HSC) Medicine - Pharmacology; (HSC) Nutrition; <i>Frontiers in Pharmacology --- Frontiers</i>	-Concerns/Issues About Data -Investigation by Company/Institution	Hamed Haddad Kashani Elabe Seyed Hosseini Hossein Nikzad Alireza Soleimani Maryam Soleimani mohammad Reza Tamadon Fariba Kenechlohi Zatollah Asami	02/02/2018 29456507 10.3389/fphar.2018.00050	10/11/2020 33041833 10.3389/fphar.2020.602201	Clinical Study Expression of concern	Iran No
Anatomical Sciences Research Center, Kashan University of Medical Sciences, Kashan, Iran Department of Internal Medicine, Kashan University of Medical Sciences, Kashan, Iran Research Center for Biochemistry and Nutrition in Metabolic Diseases, Kashan University of Medical Sciences, Kashan, Iran Department of Internal Medicine, Semnan University of Medical Sciences, Semnan, Iran Department of Urology, School of Medicine, Alborz University of Medical Sciences, Karaj, Iran <a href="http://retractionwatch.com/2020/11/10/journals-flag-concerns-in-three-dozen-papers-portion-researchers/">http://retractionwatch.com/2020/11/10/journals-flag-concerns-in-three-dozen-papers-portion-researchers/</a>						
Partitioning of adipose lipid metabolism by altered expression and function of PPAR isoforms after bariatric surgery (BLS) Biochemistry; (BLS) Biology - Molecular; (HSC) Medicine - Gastroenterology; (HSC) Medicine - Surgery; <i>International Journal of Obesity --- Springer - Nature Publishing Group</i>	-Error in Image	Cyrus Jahansouz Hongliang Xu Ann V. Hertzel Scott Kiziv Kaylee A. Steen Rosio Faneza Federico J Serrot Nicholas Kvalheim Girija Lathra Kristin Ewing Daniel B. Leslie Sayed Keramuddin David A. Bernlohr	08/14/2017 28894292 10.1038/sj.ijo.2017.197	08/06/2019 31388095 10.1038/s41366-019-0425-x	Clinical Study Correction	United States No
Department of Surgery, University of Minnesota, Minneapolis, MN, USA Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN, USA Park Nicollet Bariatric Surgery Center, St. Louis Park, MN, USA CentraCare Bariatric Center, St. Cloud, MN, USA						
The Role of the Carbohydrate Recognition Domain of Placental Protein 13 (PP13) in Pregnancy Evaluated with Combining PP13 and the PP13 Variant (BLS) Biochemistry; (BLS) Biology - Cellular; (BLS) Biology - Molecular; (HSC) Medicine - Obstetrics/Gynecology; <i>PLoS One --- PLOS</i>	-Concerns/Issues About Data -Concerns/Issues About Image -Manipulation of Images	Marei Sammar Shahar Nisimblatt Ron Gonon Berthold Huppertz Sveinbjorn Gizarason George Choi Hamutal Meiri	07/31/2014 25079598 10.1371/journal.pone.0102832	07/25/2019 31344142 10.1371/journal.pone.0220517	Clinical Study Retraction	Austria Israel United States No
Prof. Ephraim Katzir Department of Biotechnology Engineering, ORT Braude College, Karmiel, Israel Department of Obstetrics and Gynecology, Bnai Zion Medical Center, and Technion-Israel Institute of Technology, Haifa, Israel Institute of Cell Biology, Histology and Embryology/Biobank Graz, Medical University of Graz, Graz, Austria Department of Obstetrics, Gynecology and Reproductive Sciences, University of Vermont College of Medicine, Burlington, Vermont, United States of America Telemarpe Ltd, Tel Aviv, and Hy-Laboratories, Rehovot, Israel						
Diagnostic and prognostic value of circulating miR-18a in the plasma of patients with gastric cancer (BLS) Biology - Cancer; (BLS) Biology - Molecular; (HSC) Medicine - Cardiovascular; (HSC) Medicine - Gastroenterology; (HSC) Medicine - Oncology; (HSC) Medicine - Pathology; (HSC) Medicine - Urology/Nephrology; <i>Tumor Biology (Tumour Biology) - Official Journal of the International Society of Oncology and Biomarkers (ISOBM) --- Springer</i>	-Fake Peer Review -Investigation by Journal/Publisher	Zhong-Xue Su Jian Zhao Zhong-Hou Rong Ya-Guang Wu Wen-Mao Gong Cheng-Kun Qin	11/22/2014 25416437 10.1007/s13277-014-2516-6	04/20/2017 31344142 10.1007/s13277-017-5487-6	Clinical Study Retraction	China No
Department of Hepatobiliary Surgery, Shandong Provincial Hospital Affiliated to Shandong University, Jinan, Shandong Department of Hematology, Shandong Provincial Hospital Affiliated to Shandong University, Jinan, Shandong <a href="http://retractionwatch.com/2017/04/20/new-record-major-publisher-retracting-100-journals-cancer-journal-fake-peer-review/">http://retractionwatch.com/2017/04/20/new-record-major-publisher-retracting-100-journals-cancer-journal-fake-peer-review/</a>						

Abb. 4: <http://retractiondatabase.org> mit den Suchparametern Biology – Molecular/Clinical Study (17.7.2021).

25 Dollfuß (2015): <https://www.egms.de/static/de/journals/mbi/2015-15/mbi000336.shtml#ref2> (17.7.2021).

26 <http://retractiondatabase.org> (17.7.2021).

Es geht im Kern darum, daß Fehler gemacht werden, und vor allem, daß Ergebnisse nicht reproduzierbar sind.<sup>27</sup> Daher stellt sich die Frage, wie man mit diesem Befund umgehen soll. Für die Digital Humanities hat Nan Z. Da in ihrem Beitrag gerade diese Fehlerhaftigkeit und mangelnde Reproduzierbarkeit kritisiert. Sie geht jedoch noch weiter: Ihrer Ansicht nach werden Fehlklassifizierungen sogar zum Objekt des Interesses gemacht, Ungenauigkeiten und Sonderfälle würden theoretisiert, und alles dies werde Grundlage für Forschungsförderung und Publikationen.<sup>28</sup>

Eine sehr nachdrückliche Gegenposition von Seiten der Digital Humanities ist dazu von Christof Schöch formuliert worden,<sup>29</sup> der den Anspruch an die Replizierbarkeit folgendermaßen ansetzt: „The typology describes the relationship between an earlier study and its replication in terms of four key variables: the research question, the method of analysis (including the implementation of that method) and the dataset used.“<sup>30</sup> Etwas anders setzt Fotis Jannidis an,<sup>31</sup> der, ausgehend davon, daß die Literary Studies schon immer eine empirisch-quantitative Seite gehabt hätten, in diesem Feld lediglich einen inhärenten Wechsel zweier Methoden, der hermeneutischen und der quantitativen, konstatiert, die sich im Grunde genommen auch überlappen würden.<sup>32</sup> Diese hier sehr kurz und auch nur schematisch skizzierten Überlegungen zeigen, daß das Problem in den Digital Humanities erkannt und auch angegangen wird, wenngleich die Diskussion doch noch sehr dem Abwehrmodus verhaftet zu sein scheint.

Ganz anders ist die Position der einflußreichsten Forschungsförderorganisation in Deutschland im Hinblick auf die Replizierbarkeitsproblematik.<sup>33</sup> In ihrer diesbezüglichen Stellungnahme aus dem Jahr 2017 schließt die Deutsche Forschungsgemeinschaft (DFG) ausdrücklich alle Wissenschaftsbereiche ein und hält fest: „Replikation ist ein sehr wichtiges Verfahren zur Prüfung experimentalwissenschaftlich und quantitativ begründeter empirischer Wissensansprüche in der Medizin, in den Natur-, Lebens-, Ingenieur- sowie den Sozial- und Verhaltenswissenschaften und auch den Geisteswissenschaften.“ Interessant ist, daß die DFG hier zwar die Geisteswissenschaften einbezieht, jedoch mit der einschränkenden Begrenzung auf die quantitativ begründeten, empirischen Wissenschaften. Das 2018 veröffentlichte

27 Eine einleuchtende Begriffsbestimmung zu Replikation und Reproduktion findet sich in DFG (2018), 12: ([https://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/2018/180507\\_stellungnahme\\_replizierbarkeit\\_sgkf.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2018/180507_stellungnahme_replizierbarkeit_sgkf.pdf), abgerufen 17.7.2021):

„Die Begriffsbildung der Wörter Replikation, Replizierbarkeit bzw. der ebenfalls oft verwendeten Begriffe Reproduktion, Reproduzierbarkeit sowie die Abgrenzungen der Bedeutung zwischen diesen Ausdrücken ist noch nicht abgeschlossen. Die Begriffe werden auch im angelsächsischen Sprachraum nicht einheitlich verwendet.

In diesem Papier werden nur die Begriffe Replikation und Replizierbarkeit verwendet, mit folgender Bedeutung:

**Replikation** ist die Wiederholung einer Untersuchung/eines Experiments/einer Studie, die den Anspruch auf Wiederholbarkeit erhebt, meint aber auch allgemein die ergebnisoffene Möglichkeit, etwas zu wiederholen bzw. ein Experiment noch einmal durchführen zu können.

**Replizierbarkeit** meint entsprechend die Möglichkeit bzw. Fähigkeit, Ergebnisse innerhalb des Fehlerrahmens wiederholend zu bestätigen.“

28 Da (2019a): „CLS [sc. Computational Literary Studies, C.S.] routinely relies on these concepts to provide plausible explanations or theoretical motivations for results that are nothing more than a description of the data. In their project on *The Sorrows of Young Werther*, for example, Andrew Piper and Mark Algee-Hewitt compared a standard visualization of the repetition of 91 words in Goethe’s oeuvre with theoretical paradigms as different as those of Gilles Deleuze, Alain Badiou, Bruno Latour, and Michel Foucault.“

29 Schöch et al. (2020), online: <https://hcommons.org/deposits/item/hc:30439/> (17.7.2021).

30 Schöch, Contribution 1 in Schöch et al. (2020) (wie Anm. 29), S. 2 der PDF-Version.

31 Jannidis, Contribution 3 in Schöch et al. (2020) (wie Anm. 29), S. 6 der PDF-Version.

32 Da die klassische Philologie sich bereits seit dem 19. Jahrhundert solcher empirisch-quantitativen Analysen bedient hat und auch die Alte Geschichte, bspw. im Bereich der Epigraphik und Numismatik, mit ähnlichen Ansätzen arbeitet, ist Jannidis in dem Punkt natürlich zustimmen, daß es solche empirischen Studien auch in geisteswissenschaftlichen Fächern schon lange gibt.

33 DFG (2017) zur Replizierbarkeit von Forschungsergebnissen, online: [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/2017/170425\\_stellungnahme\\_replizierbarkeit\\_forschungsergebnisse\\_de.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2017/170425_stellungnahme_replizierbarkeit_forschungsergebnisse_de.pdf) (17.7.2021).

DFG-Papier „Replizierbarkeit von Ergebnissen in der Medizin und Biomedizin. Stellungnahme der Arbeitsgruppe ‚Qualität in der Klinischen Forschung‘ der DFG-Senatskommission für Grundsatzfragen in der Klinischen Forschung“ (2018) bezieht ausdrücklich nur die Medizin und Biomedizin ein.<sup>34</sup>

Aufschlußreich ist es, dazu einige Zitate aus der Veröffentlichung der DFG zur Frage der Replizierbarkeit von Forschungsergebnissen anzusehen.<sup>35</sup> Die drei Kernaussagen der Stellungnahme aus 2017 sind:

- „Replizierbarkeit ist kein generelles Kriterium wissenschaftlicher Erkenntnis.“
- „Nicht-Replizierbarkeit ist kein genereller Falsifikationsbeweis.“
- „Die Feststellung der Replizierbarkeit oder Nicht-Replizierbarkeit eines wissenschaftlichen Ergebnisses ist ihrerseits ein wissenschaftliches Ergebnis.“

Die Frage der Replizierbarkeit ist, vor dem Hintergrund der massiven Kritik an den Digital Humanities, von grundsätzlicher Bedeutung für ihre Position im Verhältnis zu den klassischen Geisteswissenschaften. Es ist nicht zu bestreiten, daß ein hermeneutisches Vorgehen sich nicht replizieren läßt in dem Sinn, daß unter gleichen Bedingungen (wie etwa im Labor, bei Experimenten, bei empirischen und/oder quantitativen Studien) ein Ergebnis sozusagen wiederholbar ist. Vielmehr müssen das Ergebnis und Verfahren des hermeneutischen Vorgehens in ihrer Sinnhaftigkeit überzeugen und erhalten dadurch den gleichen wissenschaftlichen Stellenwert wie die Wiederholbarkeit durch Replikation.

Im Unterschied dazu ist bei Ergebnissen, die mit den Methoden der Digital Humanities erzielt worden sind, Replikation bzw. Reproduzierbarkeit möglich, denn damit ist die Wiederholung einer Analyse mit dem gleichen Datensatz und den gleichen Methoden gemeint. Hierfür könnten die von der DFG 2018 für den Bereich der Medizin und Biomedizin formulierten Grundsätze, die als Voraussetzungen der Replizierbarkeit dienen, durchaus übertragen werden.<sup>36</sup>

- „Validität von Modellen und Standardisierung von Methoden“
- „Adäquate statistische Planung“
- „Sorgfältiges Management von Forschungsdaten und Materialien“
- „Umfassende Darstellung der Methoden und Analysen“

Insofern sollte Replizierbarkeit in den Digital Humanities als hinreichend betrachtet werden, wenn Verfahren und Ergebnis auf diesen Voraussetzungen beruhen. Wenn diese Form der Replizierbarkeit nicht gegeben und auch noch nicht als Standard etabliert ist, dürften auch weiterhin Zweifel an der Validität aufkommen, und – wie die Diskussion, die hier skizziert ist, deutlich zeigt, – die Ergebnisse als fehlerhaft klassifiziert werden. Hinzu kommt darüber hinaus die kritische Begründung der konzeptionellen Voraussetzungen, die aber von der eigentlichen Replizierbarkeit unterschieden werden sollte und den hier genannten Voraussetzungen der Replizierbarkeit vorausgehen muss.<sup>37</sup>

Damit stellt sich die grundlegende Frage, wie angesichts der laufenden Auseinandersetzungen mit Fehlerhaftigkeit in den Digital Humanities umgegangen werden soll. Wenn man bspw. in der Alten Geschichte mit dem klassischen Methodeninstrumentarium arbeitet, dann verwendet man bei der Arbeit mit Textquellen Editionen mit einem kritischen Apparat, der auf der Sichtung der Handschriften durch

34 DFG (2018): [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/2018/180507\\_stellungnahme\\_replizierbarkeit\\_sgkf.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2018/180507_stellungnahme_replizierbarkeit_sgkf.pdf) (17.7.2021).

35 DFG (2017) zur Replizierbarkeit von Forschungsergebnissen, online: [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/2017/170425\\_stellungnahme\\_replizierbarkeit\\_forschungsergebnisse\\_de.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2017/170425_stellungnahme_replizierbarkeit_forschungsergebnisse_de.pdf) (17.7.2021).

36 DFG (2018) zu Replizierbarkeit von Ergebnissen in der Medizin und Biomedizin: [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/2018/180507\\_stellungnahme\\_replizierbarkeit\\_sgkf.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2018/180507_stellungnahme_replizierbarkeit_sgkf.pdf), S. 8 (17.7.2021).

37 Anders Jannidis, Contribution 3 in Schöch et al. (2020).

die Editoren beruht. Man stellt Quellenreferenzen und verwendete Literatur zusammen und versucht dies in der Regel in vollständiger Weise, um die Argumente zu belegen und so auch zu plausibilisieren. Unterlaufen Fehler oder vielleicht sogar nur vermeintliche Fehler, so ist dies Ausgang eines wissenschaftlichen Diskurses, der im Fach geführt wird. Dabei kommt es durchaus zu Schärfen gegenüber denjenigen, denen man Fehler nachweist oder auch nachzuweisen glaubt, jedoch wird niemand das ganze Methodeninstrumentarium oder das Fach selbst infrage stellen.

In den Digital Humanities beobachten wir eine andere Art der Diskussion um Fehler, wie die zitierten Beiträge, aber auch viele andere zeigen: Hier wird dem ganzen Bereich eine grundsätzliche Schwäche und Fehlerhaftigkeit attestiert, denn Fehler, Ungenauigkeiten und Nicht-Wiederholbarkeit werden als ‚Schuld‘ eines falschen Vorgehens klassifiziert. Dies ist ganz offensichtlich der Versuch, die wissenschaftliche Entwicklung der Digital Humanities in ihrem Methodenbereich mit der Figur einer Schuldkultur zu verbinden.

Demgegenüber sollte man für die zweifelsfrei sichtbaren Probleme und Fehler eine konstruktive Fehlerkultur für die Digital Humanities entwickeln und offensiv mit den Fehlern umgehen. Daher wird hier ein methodisches Vorgehen vorgeschlagen, durch das eine konstruktive Fehlerkultur in den Digital Humanities ermöglicht und eine Brücke zu den klassisch hermeneutischen Arbeitsweisen gebaut würde.

Grundsätzlich ist zuallererst eine Taxonomie der Fehler nötig,<sup>38</sup> denn je nach Arbeitsphase oder -stand sind die Fehler andere und ist auch anders damit umzugehen. Die Fehler müssen klassifiziert werden nach Grundannahmen, Experimentphasen und Komplexitätskontext.

Darauf muß eine Fehlerkritik aufsetzen, die die Fehler entsprechend ihrer Klassifikation einordnet und in Iterationen so integriert, so daß die Veränderungen in Daten, Metadaten und Parameter immer – auch nachträglich – zugeordnet und transparent gemacht werden können.

Schließlich muß das Ergebnis selbst in jeder Hinsicht transparent gemacht werden, so daß der Weg dahin jederzeit nachverfolgt werden kann. Am Ende muß auch Ergebnisstabilität erzielt werden, so daß die Digital Humanities aus dem Projektstadium zum „Dauerbetrieb“ kommen.

Letztlich würde dies auf die Notwendigkeit der Standardisierung der Methoden führen und dies könnte mit der Einführung von SOPs (Standard Operating Procedures) in den Digital Humanities erreicht werden, die sie vergleichbar mit anderen technischen Bereichen machen. Somit müssen nicht nur die einzelnen Schritte des Preprocessing dokumentiert werden, sondern vor allem müssen die Voraussetzungen und die der Datenaufbereitung zugrundeliegende Systematik offengelegt und in ein Verhältnis zur jeweiligen, fachspezifischen Wissensordnung gesetzt werden.

Dies sind anspruchsvolle Anforderungen, die aber den Digital Humanities die Augenhöhe sowohl mit anderen technisch orientierten Fächern wie vor allem mit den klassischen Geisteswissenschaften ermöglichen werden.

---

38 Schöch in Schöch, Contribution 1 in Schöch et al. (2020), S. 2f. (der PDF-Version) schlägt eine Typologie der Replikation vor, für die er auch eine Systematik visualisiert. M.E. ist es jedoch notwendig, bevor die Digital Humanities dies angehen können, die Fehler kritisch zu analysieren und dann erst die Replikationssystematik zu entwickeln.

## Literatur

- Da (2019a): Nan Z. Da, The Digital Humanities Debacle, The Chronicle of Higher Education 2019. Online: <https://www.chronicle.com/article/The-Digital-Humanities-Debacle/245986>
- Da (2019b): Nan Z. Da., The Computational Case against. Computational Literary Studies, Critical Inquiry 45/3 (Spring 2019), 601–639. Online: <https://www.journals.uchicago.edu/doi/abs/10.1086/702594?journalCode=ci>
- DFG (2017): DFG-Positionspapier zur Replizierbarkeit von Forschungsergebnissen, Stellungnahme der Deutschen Forschungsgemeinschaft, Bonn 2017. Online: [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/2017/170425\\_stellungnahme\\_replizierbarkeit\\_forschungsergebnisse\\_de.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2017/170425_stellungnahme_replizierbarkeit_forschungsergebnisse_de.pdf)
- DFG (2018): DFG Positionspapier: „Replizierbarkeit von Ergebnissen in der Medizin und Biomedizin. Stellungnahme der Arbeitsgruppe ‚Qualität in der Klinischen Forschung‘ der DFG-Senatskommission für Grundsatzfragen in der Klinischen Forschung“, Bonn 2018. Online: [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/2018/180507\\_stellungnahme\\_replizierbarkeit\\_sgkf.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2018/180507_stellungnahme_replizierbarkeit_sgkf.pdf)
- Diemke (2020): J. Diemke, Alkibiades, Pyrrhos und Alexander: Eine Untersuchung zu Emotionen und Gewalt in den Viten Plutarchs unter Verwendung digitaler Methoden, DCO 6,2 (2020), 57–74. Online: <https://journals.ub.uni-heidelberg.de/index.php/dco/article/view/77663/71565>
- Dollfuß (2014): H. Dollfuß, Analyse zurückgezogener Publikationen in der bibliografischen Datenbank Web of Science von 2004 bis 2014, MS Med Bibl Inf 2015;15(1–2):Doc09. Online: <https://www.egms.de/static/de/journals/mbi/2015-15/mbi000336.shtml#ref2>
- Flashar (2004): H. Flashar, das Werk des Aristoteles, in: H. Flashar (Hrsg.), Die Philosophie der Antike Bd. 3: Ältere Akademie. Aristoteles. Peripatos, Basel 2004.
- Juola (2015): P. Juola, The Rowling Case: A Proposed Standard Analytic Protocol for Authorship Questions, Digital Scholarship 30 (2015), Supplement 1, 101–113.
- Kirby (2019): J. S. Kirby, J. S. (2019), How NOT to create a digital media scholarship platform: the history of the Sophie 2.0 project, IASSIST Quarterly, 42,4 (2019). Online: <https://doi.org/10.29173/iq926>
- Krämer (2018): S. Krämer, Der ‚Stachel des Digitalen‘ – ein Anreiz zur Selbstreflexion in den Geisteswissenschaften? Ein philosophischer Kommentar zu den Digital Humanities in neun Thesen, DCO 4,1 (2018), 5–11. Online: <https://doi.org/10.11588/dco.2017.0.48490>
- Martin et al. (2011): W. Martin, M. Brown / R. Klavans / K. Boyack, OpenOrd: An Open-Source Toolbox for Large Graph Layout, 2011, Proceedings of SPIE – The International Society for Optical Engineering SPIE 2011. Online: <https://github.com/gephi/gephi/wiki/OpenOrd>
- Rautenberg (2019): J. Rautenberg, Negation in Platons Sophistes und die Grenzen automatisierter Paraphrasensuche, in: Schubert, Ch. et al. (Hrsgg.): Platon Digital: Tradition und Rezeption, Heidelberg 2019, 111–123 (Digital Classics Books, Band 3). Online: <https://doi.org/10.11588/propylaeum.451>

- Riess (2019): W. Riess, A Digital Analysis of Maritime Acts of Violence Committed by Alcibiades as Described by Thucydides, Xenophon, and Plutarch, DCO 5,2 (2019), 4–27. Online: <https://doi.org/10.11588/dco.2019.2.72018>
- Riess (2020): W. Riess, Prolegomena zu einer digitalen althistorischen Gewaltforschung: Gewaltmuster bei Solon, Alcibiades und Arat im Vergleich, Klio 102 (2020), 445–473.
- Rizvi (2018): P. Rizvi, The interpretation of Zeta test results, Digital Scholarship in the Humanities 34/2 (2018), 401–418. Online: <http://doi.org/10.1093/lc/fqy038>
- Schöch et al. (2020): Chr. Schöch et al., Replication and Computational Literary Studies, Panel at the Digital Humanities Conference 2020 (DH2020), Ottawa, Canada, July 20–25, 2020. Online: <https://hcommons.org/deposits/item/hc:30439/> und <https://dh2020.adho.org/>
- Schäfer (2019): Ch. Schäfer, Die Kontrolle des Meeres: Alcibiades und die Sizilische Expedition, DCO 5,1 (2019), 22–33 und online: <https://doi.org/10.11588/dco.2019.1.60564>
- Schubert / Weiß (2015): Ch. Schubert / A. Weiß, Die Hypomnemata bei Plutarch und Clemens: Ein Textmining-gestützter Vergleich der Arbeitsweise zweier ‚Sophisten‘, Hermes 143 (2015), 447–471.
- Schubert (2017): Ch. Schubert, Die Arbeitsweise Iamblichs im Protreptikos, in: S. Brandt / Ch. Schubert: Der Protreptikos des Iamblich: Rekonstruktion, Refragmentisierung und Kontextualisierung mit Textmining, Digital Classics Books, Heidelberg 2017, 17–48. Online: <https://books.ub.uni-heidelberg.de/propylaeum/catalog/book/257>
- Schubert (2018): Ch. Schubert, Eine Thukydides-Paraphrase in der Totenrede des Tiberius auf Augustus: Cassius Dios Sichtweise des augusteischen Prinzipats, Antike und Abendland 64 (2018), 79–92.
- Schubert (2019a): Ch. Schubert, Plädoyer für eine Fehlerkultur in den Digital Humanities, DCO 5,2 (2019). Online: <https://doi.org/10.11588/dco.2019.2.72004>
- Schubert (2019b): Ch. Schubert, Visualisierung von Textdaten: Die Falle der Metadaten am Beispiel von Iamblichs Protreptikos, DCO 5,1 (2019), 4–21. Online: <https://doi.org/10.11588/dco.2019.1.59356>
- Schubert (2020): Ch. Schubert, Zur Standortbestimmung des Digitalen in den Altertumswissenschaften. Textanalyse am Beispiel des Corpus Hippocraticum und des hippokratischen Eides, Gymnasium 127 (2020), 305–327.
- Thomas (2016): William G. Thomas III, The Promise of the Digital Humanities and the Contested Nature of Digital Scholarship in: Susan Schreibman, Ray Siemens, John Unsworth (Hrsgg.), A New Companion to Digital Humanities, Malden / Oxford 2016, 524–537.
- Van Zundert (2016): J. J. van Zundert, Screwmenetics and Hermeneumericals: the Computationality of Hermeneutics, in: S. Schreibman / R. Siemens / J. Unsworth (Hrsgg.), A New Companion to Digital Humanities, Malden/ Oxford 2016, 331–347. Online: <http://onlinelibrary.wiley.com/doi/10.1002/9781118680605.ch23/summary>
- Warnking (2015): P. Warnking, Der römische Seehandel in seiner Blütezeit: Rahmenbedingungen, Seewege, Wirtschaftlichkeit, Rahden / Westf. 2015.

Warnking (2016): P. Warnking, Roman Trade Routes in the Mediterranean Sea: Modelling the routes and duration of ancient travel with modern offshore regatta software, in: Ch. Schäfer (Hrsg.), Connecting the Ancient World. Mediterranean Shipping, Maritime Networks and their Impact, Rahden / Westf. 2016, 45–90.

## Internetseiten

DNB:

[https://www.dnb.de/DE/Professionell/Standardisierung/Standards/standards\\_node.html](https://www.dnb.de/DE/Professionell/Standardisierung/Standards/standards_node.html)

[https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd\\_node.html](https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd_node.html)

<http://d-nb.info/gnd/4305003-7>

Gephi 0.9.1:

<https://gephi.org/>

Infowiss:

<https://wiki.infowiss.net/Stopwort>

Retraction Database:

<http://retractiondatabase.org>

StyloAH v.0.7.4.5:

<https://github.com/ecomps-hONgit/stylo>

TLG-Datenbank:

<http://stephanus.tlg.uci.edu/>

Webportal Digital Plato:

<https://digital-plato.org/>

Webportal eAQUA:

[www.eaqua.net/](http://www.eaqua.net/)

Zusammenstellung der Antworten auf Nan Z.Da:

[https://demo.hedgedoc.org/s/rJ\\_YoK\\_cH](https://demo.hedgedoc.org/s/rJ_YoK_cH)

## Abbildungsnachweise

Abb. 1: Stopwortliste, Ausschnitt aus: <https://wiki.infowiss.net/Stopwort>

Abb. 2: Katalog der DNB, Ausschnitt aus: <http://d-nb.info/gnd/4305003-7>

Abb. 3: Ch. Schubert

Abb. 4: Retraction Database, Ausschnitt aus: <http://retractiondatabase.org>

## Autoreninformation<sup>39</sup>

**Prof. em. Dr. Charlotte Schubert**

Universität Leipzig  
Beethovenstr. 15  
04107 Leipzig  
Tel: 0341/9737071  
E-Mail: [schubert@uni-leipzig.de](mailto:schubert@uni-leipzig.de)

---

39 Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autoren. Alle Inhalte dieses Beitrages unterstehen, soweit nicht anders gekennzeichnet, der Lizenz CC BY 4.0.