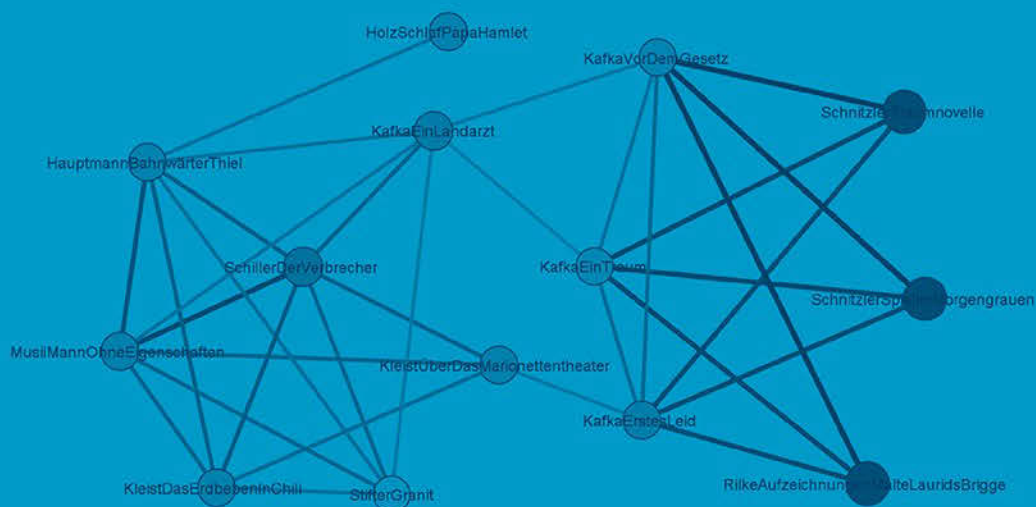


Alexa Lucke, Hermann Johannes (Hg.)

LITERATUR- WISSENSCHAFT UND INFORMATIK

Transdisziplinäre Perspektiven,
digitale Methoden und
selbstlernende Algorithmen



[transcript] <digital humanities>

Aus:

Alexa Lucke, Hermann Johannes (Hg.)

Literaturwissenschaft und Informatik

Transdisziplinäre Perspektiven, digitale Methoden und selbstlernende Algorithmen

April 2024, 256 S., kart., 8 SW-Abb., 26 Farbbabb.

47,00 € (DE), 978-3-8376-7003-5

E-Book: kostenlos erhältlich als Open-Access-Publikation

PDF: ISBN 978-3-8394-7003-9

Die transdisziplinäre Zusammenarbeit stellt die Fächer Literaturwissenschaft und Informatik vor die Herausforderung, ihre Theoriebildungen und Modelle neu zu perspektivieren. Die Beiträger*innen dokumentieren exemplarisch die spannungsreiche Verschränkung von digitalen und klassischen Methoden, fachlichen Standards sowie unterschiedlichen Forschungsgegenständen der beiden höchst differenten Wissenschaftskulturen. Aus der Sicht beider Disziplinen eruieren sie die epistemologischen Mehrwerte digitaler Modellierung oder evaluieren Tools für die Analyse literarischer Texte – und stoßen so durch die wechselseitigen methodologischen Perspektiven eine Revision facheigener Fragestellungen an.

Alexa Lucke (Dr. phil.) hat im Fach Neuere deutsche Literaturwissenschaft an der Universität Siegen promoviert. Als wissenschaftliche Mitarbeiterin ist sie im Verbund-Projekt »LiGeDi – Literaturgeschichten erarbeiten. Gemeinsam im Digitalen« der Universitäten Paderborn, Bielefeld und Wuppertal beschäftigt.

Hermann Johannes (Prof. Dr.-Ing.) hat an der Fachhochschule Südwestfalen im Bereich Wirtschaftsinformatik gelehrt. Seine Forschungsschwerpunkte sind Business Intelligence, Data Mining und Bestandsoptimierung.

Weitere Informationen und Bestellung unter:
www.transcript-verlag.de/978-3-8376-7003-5

Inhalt

Methodologische Potenziale und Herausforderungen einer transdisziplinären Zusammenarbeit – Eine Einleitung

Alexa Lucke (Literaturwissenschaft) 7

Topic Modeling am Beispiel von Franz Kafka

Eine Arbeitsskizze

Laura Kraft (Literaturwissenschaft) 35

Über den Nutzen des *Natural Language Processing* in der computationalen Analyse literarischer Texte des ›Naturalismus‹ und der ›literarischen Moderne‹

Henning Femmer (Informatik) und Alexa Lucke (Literaturwissenschaft) 47

Computergenerierter Zufall als kreatives Moment in Malerei und Literatur

Potentiale und Grenzen von Machine-Learning-Modellen am Beispiel von GPT

Andreas de Vries (Informatik) 93

Die Kunst im Zeitalter der künstlichen Intelligenz

Walter Benjamins Kunsttheorie im 21. Jahrhundert

Denis Stevanovic (Literaturwissenschaft) 123

Algorithmizität, Autorschaft und Paratext in generativer Kunst und Literatur

Alexa Lucke (Literaturwissenschaft) 139

Verwenden Menschen und Automaten die gleiche Sprache?

Literarische Fiktionen künstlichen Menschseins

Judith Schönhoff (Literaturwissenschaft/Komparatistik) 167

Softwaretools für die Literaturtextanalyse – Ein Überblick

Hermann Johannes (Informatik) 191

Gebiete und Publikationsformen in der Informatik

Jennifer Krieger (Informatik) 247

Adressen 253

Methodologische Potenziale und Herausforderungen einer transdisziplinären Zusammenarbeit – Eine Einleitung

Alexa Lucke (Literaturwissenschaft)

Im Zuge einer viel diskutierten digitalen Transformation der (Geistes-)Wissenschaften stellt sich unweigerlich die Frage nach den Potenzialen und Herausforderungen einer inter- und transdisziplinären Zusammenarbeit zwischen Literaturwissenschaft und Informatik und nach der Vereinbarkeit ihrer theoretischen Ansätze, Methoden und Modelle.

Die Digitalisierung und Algorithmisierung von Forschungssettings betreffen inzwischen sämtliche geisteswissenschaftliche Disziplinen, in denen sie äußerst ähnlich kontroverse Debatten über theorie- vs. datenbasierte Forschungsmethodiken ausgelöst haben. Im Kontext der *Digital Humanities* (DH) wird bereits vermehrt von ›Lab‹-Kulturen gesprochen und einem *laboratory turn*¹. Traditionelle Methoden werden in den geisteswissenschaftlichen Fächern dabei nicht nur um digitale ergänzt und erweitert, auch ein wissenschaftstheoretischer Diskurs richtet sein Augenmerk zunehmend auf die neuen epistemologischen Implikationen ihrer unterschiedlichen Kombinationen.² Der Einsatz von digitalen bzw. computationellen Methoden in den Geisteswissenschaften wird in den DH von einem selbst erklärten Anspruch an eine neue Theoriebildung digitaler Forschung begleitet, der insbesondere philosophisch und wissenschaftstheoretisch begründet wird. Der Anspruch an die eigene Theoriebildung zeigte sich beispielsweise in der Debatte über die Brauchbarkeit des Begriffs *Algorithmizität*.³ Dieser Begriff steht nach

- 1 Vgl. Pawlicka-Deger, U.: The Laboratory Turn: Exploring Discourses, Landscapes, and Models of Humanities Labs. In: *Digital Humanities Quarterly* 14, No. 3 (2020). <http://www.digitalhumanities.org/dhq/vol/14/3/000466/000466.html> [30.10.2023].
- 2 Vgl. zu erkenntnistheoretischen Überlegungen in den DH exemplarisch T. Gengnagel (2022): Vom Topos des Scheiterns als konstituierender Kraft. Ein Essay über Erkenntnisprozesse in den Digital Humanities. In: *Fabrikation von Erkenntnis – Experimente in den Digital Humanities*. Hg. von Manuel Burghardt, Lisa Dieckmann, Timo Steyer, Peer Trilcke, Niels Walkowski, Joëlle Weis, Ulrike Wuttke. Wolfenbüttel 2021–2022. (=Zeitschrift für digitale Geisteswissenschaften/Sonderbände, 5) text/html Format. DOI: 10.17175/sbo05_011 [30.10.2023]
- 3 Vgl. Stalder, F. (2016): Kultur der Digitalität, Berlin und Gramelsberger, G. (2022): Schlaglichter. Algorithmizität. In: *Feministische Studien* 40 (2), S. 227–228 (<https://doi.org/10.1515/fs-2022-0033>)

Felix Stalder für die Vorordnung von kulturellen (Denk-)Prozessen durch Algorithmen und maschinelle Prozesse.⁴

Die Veränderungen im Bereich von neuen digitalen Forschungspraktiken können angesichts einer verkürzten Zeitachse in Bezug auf die neueren Entwicklungen kaum adäquat aufgearbeitet werden. Die Geschichtswissenschaft hat beispielsweise andere fachliche Anforderungen an digitale Programme als die Komparatistik; die Computerlinguistik, die sich bereits seit Jahrzehnten mit der Anwendung von computationellen Methoden beschäftigt, operiert mit anderen Ansätzen und Methodiken als die Literaturwissenschaft. In Bezug auf die einzelnen traditionellen geisteswissenschaftlichen Fächer ist eine emergente Bildung von Paralleldiskursen zu konstatieren, wie auch die neben der traditionellen Literaturwissenschaft sich entwickelnde und »konkurrierende« digitale Literaturwissenschaft: die vor allem unter der Bezeichnung *computational literary studies* bekannt ist.

Aus dieser – äußerst knapp skizzierten – gegenwärtigen Situation in den DH resultieren Fragen nach den relationalen methodologischen Dynamiken und epistemologischen Mehrwerten einer transdisziplinären Zusammenarbeit zwischen Literaturwissenschaft und Informatik.⁵ Die Herausforderung der Kompatibilität von Forschungsmethodiken zwischen qualitativen, hermeneutischen und theoretisch-deduktiven Verfahrensweisen auf der Seite geistes- und kulturwissenschaftlicher Fachtraditionen und statistischen, quantitativen und induktiv-explorativen⁶ auf der Seite der Informatik wird im vorliegenden Band thematisiert; und es wird versucht, die dazwischenliegenden »gaps« pragmatisch zu handhaben. Mit Rabea Kleymann plädieren die hier versammelten Beiträge dafür, die methodischen und disziplinären

12.07.2023). Vgl. dazu die aktuelle Debatte über den Algorithmizitätsbegriff auf der DHd-Konferenz 2023. (<https://dhtheorien.hypotheses.org/1249>, 14.05.2023): Die Debatte antwortete auf den Vorwurf einer »Theorielosigkeit« (»end of theory«).

- 4 Vgl. Stalder: Kultur der Digitalität. In der AG *Digital Humanities Theories* des DHd-Verbands wurde in diesem Zusammenhang die Frage aufgeworfen, ob *Computational Thinking* (Begriff aus der Informatik) etwa »das Gleiche wie ein algorithmischer Erkenntnisprozess oder das Gleiche wie ein regelgeleitetes Vorgehen (wie etwas Dilthey oder Descartes es beschreiben)?« sei (<https://dhtheorien.hypotheses.org/1249>, 30.3.2023). Vgl. zum *Computational Thinking* auch Denning, P. J., Matti, T. (2019): *Computational Thinking*. Cambridge: MIT Press.
- 5 Vgl. auch Gius, E., Jacke, J. (2015): Informatik und Hermeneutik. Zum Mehrwert interdisziplinärer Textanalyse. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum/Thomas Stäcker. 2015 (=Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: 10.17175/sb001_006.
- 6 Dass zwischen explorativen und hypothesengetriebenen Verfahren ein Trennstrich eingezeichnet wird, erscheint angesichts der Fülle an existierenden Ansätzen etwas vereinfachend (vgl. die Annotationen der Verfasserin zum Beitrag *Mixed Methods* von J. Schröter (2023) auf https://via.hypotheses.org/https://zfdg.de/wp_2023_008). In induktiv-explorative Verfahrensweisen fließt »entgegen einem verbreiteten Mißverständnis – durchaus theoretisches Vorwissen ein« (U. Flick (1996): *Qualitative Forschung. Theorie, Methoden, Anwendung in Psychologie und Sozialwissenschaften*, Reinbek bei Hamburg, S. 10). »In der Praxis besteht keine scharfe Trennung zwischen datengetriebenen und hypothesengetriebenen Verfahren. Zumeist existieren bestimmte Annahmen und eine völlig freie Suche nach 'irgendwelchen' Auffälligkeiten in den Daten wird allgemein als wenig sinnvoll erachtet« (J. Kaufmann, W. Mülder (2023): *Grundkurs Wirtschaftsinformatik*, Wiesbaden, S. 393. https://doi.org/10.1007/978-3-658-37937-7_14).

Lücken zwischen den differenten Methodiken und Zugängen nicht etwa zugunsten einer hermeneutischen Ganzheit zu überbrücken (*bridging the gap*) oder zu schließen (*closing the gap*), sondern vielmehr die disziplinären Differenzen und die zugrundeliegende ›andere‹ Perspektive anzuerkennen.⁷ Denn es bedarf der Reflexion und Kenntnis von (historischen) philosophischen Grundlagen von disziplinären Methoden, Theorien und Wissenskonzepten, die sich beispielsweise in unterschiedlichen semantischen Verwendungen gleicher Begriffe⁸ wie ›Methode‹ oder ›Interpretation‹ ausdrücken können, um eine multiple Perspektivität von Untersuchungsgegenständen zu erreichen. In der Informatik wird mit »Methode« eine bestimmte Vorgehensweise (Verfahren, Herangehensweise) zur Lösung einer Aufgaben- oder Problemstellung, insbesondere in den Bereichen Softwareentwicklung, Datenmodellierung und -management, bezeichnet⁹, in der Literaturwissenschaft ein der Multimodalität von symbolisch-kulturellen Gegenständen angemessener (theoretischer) Zugang.¹⁰

Es geht also um die Erschließung der jeweiligen Theorietraditionen, Wissensbestände und -archive der beiden akademischen Fächer. Insofern es sich aus unserer Sicht im Grunde um ein transdisziplinäres Kontinuum von durch historische Prozesse von westlich kulturellen Differentiationen häufig komplementäre und relationale Epistemologien handelt¹¹, die die Disziplinen Literaturwissenschaft und Informatik bereits miteinander verschränken, liegt eine transdisziplinäre Zusammenarbeit nahe – besonders für die Wiedereinholung und Integration ausgeblendeter (oft philosophisch basierter) theoretischer Perspektiven innerhalb der eigenen Wissenschaft(sgeschichte).

Reflexionen der Begriffe ›Methode‹ und ›Verfahren‹ gehören zum Theoriebildungsprozess und Selbstverständnis der Digital Humanities, in denen durch den Einzug quantitativer Methoden (Franco Moretti) eine Vielzahl methodologischer Debatten ausgelöst wurde.

»Seit dem frühneuzeitlichen Rationalismus ist ›Methode‹ ein epistemologischer Grundbegriff«¹² heißt es beispielsweise im Working Paper 2 des Glossars »Begriffe der Digital Humanities« in der ZfdG (2023). Jan Horstmann u.a. bezeichnen darin den Begriff als »Einbettung eines oder mehrerer ›kriteriengeleiteter‹ und ›regelbasierter‹

7 Vgl. Kleymann, R. (2022): Datendiffraktion: Von Mixed zu Entangled Methods in den Digital Humanities. In: Fabrikation von Erkenntnis – Experimente in den Digital Humanities. Hg. von M. Burghardt, L. Dieckmann, T. Steyer, P. Trilcke, N. Walkowski, J. Weis, U. Wuttke. Wolfenbüttel 2021 – 2022. (=Zeitschrift für digitale Geisteswissenschaften/Sonderbände, 5) text/html Format. DOI: 10.17175/sb005_008.

8 Vgl. dazu auch Lucke, A. (2023): Wissenschaft und Kunst. Die Begriffe ›Vereinigung‹, ›Wechselwirkung‹ und ›Bestimmbarkeit‹ in Fichtes Wissenschaftslehre und Schillers Ästhetischen Briefen. In: Jahrbuch der deutschen Schiller-Gesellschaft 66 (2022), Göttingen, S. 101–128.

9 Vgl. dazu auch das »objektorientierte Paradigma« bei Fischer, P./Hofer, P. (2008): Lexikon der Informatik, 14. Aufl., Berlin/Heidelberg, S. 522f. (unter dem Lemma ›Methode‹).

10 Vgl. Annotationen der Verf. zu Schröter: Mixed Methods (Anm. 6).

11 Vgl. Lucke, A. (2021): Schillers *Ästhetische Briefe* als Literatur. Der Vollzug von literarischen Praktiken in der ›ästhetischen Kunst‹, Bielefeld.

12 Horstmann, J., Lück, C., Normann, I. (2023): Methode. In: AG Digital Humanities Theorie des Verbandes Digital Humanities im deutschsprachigen Raum e. V. (Hg.): Begriffe der Digital Humanities. Ein diskursives Glossar (=Zeitschrift für digitale Geisteswissenschaften/Working Papers, 2). Wolfenbüttel. 25.05.2023. HTML/XML/PDF. DOI: 10.17175/wp_2023_007

Verfahren in den Horizont eines Erkenntnisziels«. ¹³ Digitale Methoden und Verfahren wären demnach prinzipiell miteinander kombinierbar, ¹⁴ zum Beispiel mit klassischen literaturwissenschaftlichen Verfahren der Paratextanalyse oder der Narratologie. ¹⁵

Dass gerade Begriffe wie »Methode« und »Verfahren«, auch in ihrer Kombination und Verschränkung, im Kontext von digitaler Literatur ¹⁶ und Literaturwissenschaft ¹⁷ wieder neu skizziert werden, liegt ebenso an einer veränderten Perspektive auf Literatur als einem nicht mehr nur text-, sondern inzwischen auch datenförmigen Gegenstand, außerdem an dessen zunehmender Theoretisierung und Verwissenschaftlichung. ¹⁸ Traditionelle literaturwissenschaftliche Kategorien wie »Autorschaft«, »Werk«, »Rezipient«, »Medium« und »Literatur« müssen angesichts der Anwendung von neuesten KI-Technologien in den Bereichen Literatur und Literaturwissenschaft neu konzipiert, operationalisiert und analysiert werden, z. B. heuristisch als Parameter eines interdependenten Feldes digitaler Literatur ¹⁹, deren semantische Verschiebungen im Feld selbst theoretisch-reflexiv und auch performativ verhandelt werden. ²⁰ Zum Begriff der digitalen Literatur ist eine Vielzahl an Definitionen im Umlauf. ²¹

Die meisten der in der digitalen Literaturwissenschaft verwendeten digitalen Methoden wie das *Distant Reading* ²², NLP- oder KI-basierte Verfahren (z. B. *Clustering* oder *Fine-Tuning* von LLMs), die Stilometrie mit ihren diversen Distanzmaßen ²³, das Topic Modeling oder Visualisierung und Netzwerkanalyse basieren nach Jannidis et al. (2017) überwiegend auf statistischen Verfahren ²⁴ (ausgenommen: Annotationen ²⁵). Eine Ana-

13 Vgl. ebd. Als Beispiel für eine digitale Methode mit verschiedenen Verfahren (wie Distanzmaßen) wird dabei die Stilometrie angeführt.

14 Vgl. Horstmann u. a.: Methode.

15 Vgl. auch Gius/Jacke: Informatik und Hermeneutik.

16 Vgl. Bajohr, H., Gilbert, A. (2021): Platzhalter der Zukunft: Digitale Literatur II (2001→2021). In: Dies. (Hg.): Digitale Literatur II. In: TEXT+KRITIK. Zeitschrift für Literatur, Sonderband, S. 7–21.

17 Vgl. zur Differenz von digitaler Literatur und Literaturwissenschaft auch umfassend den Band von Jannidis, F. (Hg.) (2017): Digitale Literaturwissenschaft, DFG-Symposium 2017, Würzburg und darin besonders Winko, S.: Einführung. In: Jannidis: Digitale Literaturwissenschaft, S. 19–25.

18 Vgl. z. B. Bajohrs »experimentelle Literatur« (<https://www.creative-city-berlin.de/de/ccb-magazin/2023/9/6/hannes-bajohr-ki-literatur/>, 08.10.2023).

19 Vgl. Nantke, J. (2017): Tausend Tode, tausend Autoren, tausend Texte? Zur Textualität digitaler Literatur. In: Jannidis: Digitale Literaturwissenschaft, S. 173–199, S. 174.

20 Vgl. den Beitrag von Lucke im vorliegenden Band.

21 Vgl. dazu grundlegend den Band Bajohr/Gilbert: Digitale Literatur II.

22 Moretti, F. (2013): *Distant Reading*. London sowie Moretti, F. (2000): *Conjectures on World Literature*. In: *New Left Review* 1 (<https://newleftreview.org/issues/i11/articles/franco-moretti-conjectures-on-world-literature>, 08.10.2023).

23 Vgl. Büttner, A., Dimpel, F., Evert, S., Jannidis, F., Pielström, S., Proisl, T., Reger, I. (2017): »Delta« in der stilometrischen Autorschaftsattribuion. In: Zeitschrift für digitale Geisteswissenschaften. text/html Format. DOI: 10.17175/2017_006.

24 Vgl. Kleymann: Datendiffraktion. Vgl. Jannidis, F., Kohle, H., Rehbein M. (2017): *Digital Humanities*. Eine Einführung. Stuttgart.

25 Es sei denn, man bemisst diese anhand von IAA-Maßen (Inter-Annotator-Agreement-Maße) als Grad der Übereinstimmung zwischen Annotationen bzw. Annotator:innen in Bezug auf die Erstellung von so genannten »Goldstandards« (Bartsch, S., Gius, E., Müller, M., Rapp, A., Weitlin, T. (2023): Sinn und Segment. Wie die digitale Analysepraxis unsere Begriffe schärft. In: Zeitschrift für digitale Geisteswissenschaften 8 (2023). 01.06.2023. HTML/XML/PDF. DOI: 10.17175/2023_003). Der

lyse mit digitalen Methoden umfasst in der Regel aber weit mehr als nur das einfache Auszählen von Worthäufigkeiten.

Dieser Band widmet sich unter anderem den methodologischen Herausforderungen digitaler Literaturwissenschaft und dem Ansatz einer transdisziplinären Forschung – wie der (gemeinsamen) Operationalisierung von (digitalen und analogen) Forschungsfragen oder der (wechselseitigen) Interpretation und Validierung von Forschungsergebnissen. Es soll besonders um folgende Fragen gehen: Inwieweit kann die digitale Literaturwissenschaft von der Informatik als Wissenschaft mit ihren Denkmustern²⁶ profitieren oder inwieweit fungiert die Informatik hier als reine Hilfswissenschaft? Welche Rolle spielt dabei die (digitale) Hermeneutik und welche hermeneutischen Verfahren lassen sich in den transdisziplinären Forschungsprozess integrieren? Welches transformative Potential kommt dabei (selbstlernenden) Algorithmen und Programmen zu und welche neuen (methodologischen, hermeneutischen und semantischen) Möglichkeiten und Grenzen bieten diese im Kontext von digitaler Literatur und Kunst sowie digitaler Literaturwissenschaft? Inwieweit modellieren wissenschaftliche Instrumente, technologische Prozesse und Tools (neben Theorien und Paradigmen) die Daten, Gegenstände und Forschungsergebnisse und wie weit komme ich mit bestimmten Methoden, Verfahren und Tools in Bezug auf meine Fragestellung? Und was kann die Informatik von der Literaturwissenschaft lernen? Entlang dieser und anderer Fragen werden aus den verschiedenen Blickwinkeln der hier vertretenen Studien literaturwissenschaftliche, informatische und transdisziplinäre Frage- und Problemstellungen sowie die Anwendung von digitalen Methoden und Tools für die Analyse, Interpretation und Produktion datenförmiger literarischer und künstlerischer Gegenstände in exemplarischen Fallanalysen erprobt und evaluiert.

Was sind die DH und die digitale Literaturwissenschaft?

Die Genese der Digital Humanities vereint verschiedene Forschungsfelder (die Geisteswissenschaften und die Informatik), wodurch typischerweise Dynamiken von unterschiedlichen, auch gegenläufigen, Forschungsparadigmen und überlagernden, widerstreitenden oder teilweise einander ausschließenden Forschungspositionen entstehen können. Seit einigen Jahren hat sich ein (Meta-)Diskurs im Kontext der DH entwickelt, der sich insbesondere mit der Übersetzung von geistes- und kulturwissen-

Zusammenhang zwischen dem Grad der Übereinstimmung und dem Grad der intersubjektiven Gültigkeit scheint dennoch als fraglich (vgl. ebd.). Vgl. zum Goldstandard auch Gius/Jacke: Informatik und Hermeneutik. Der Goldstandard (also die Übereinstimmung von Annotationen mehrerer Personen) ist häufig die »Grundlage für Training und Evaluation von NLP-Systemen« (vgl. ebd.).

26 Vgl. Nerbonne, J. (2015): Die Informatik als Geisteswissenschaft. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von C. Baum und T. Stäcker (=Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). text/html Format. DOI: 10.17175/sb001_003: »Die Informatik beschäftigt sich mit digitaler Informationsverarbeitung, sowohl in einzelnen Prozessen (Algorithmen oder Rechnern) als auch in Organisationen. Digitale Informationsverarbeitung aber ist ein Erzeugnis des kollektiven menschlichen Geistes« (ebd.).

schaftlichen Theorien²⁷ und Fragestellungen in digitale Forschungssettings beschäftigt. Als klassische Texte und Grundlagen für den theoretischen Diskurs in den DH gelten unter anderen Franco Morettis Reflexion über den Einzug von quantitativen Methoden in die Literaturwissenschaft und seine Problematisierung des Begriffs der Weltliteratur,²⁸ Evelyn Gius' und Janina Jackes Erweiterung des hermeneutischen Zirkels,²⁹ die Verlagerung einer epistemologischen zu einer methodologischen Problematik von Julian Schröter et al.³⁰, Gabriele Gramelsbergers philosophische Grundlagen des Digitalen³¹ und die Definition des Begriffs der Algorithmizität von Felix Stalder.³² Dabei sind Grenzen und Methoden des neu etablierten Feldes DH noch keinesfalls abgesteckt.

One of the most exciting things about the field is that it is somewhat uncharted territory. From an outside, or even semi-outside perspective it sometimes looks as if ›digital‹ has become a prefix that enables unfundable projects to become fundable again, where the dear old humanities with their close readings, their semiotics, and idle enjoyment of the aesthetic value of theory and works of art can, under the guise of future-oriented digitality, benefit from the present discursive and monetary euphoria for specific new technologies and methods.³³

Die Unerschöpflichkeit von Definitionen der DH lässt sich auf der Internetseite *whatisdigitalhumanities.com* ablesen³⁴ und in etwa mit der Frage vergleichen: Was ist Kunst (wahlweise auch: Literatur, Religion etc.)? Die heterogenen Konstellationen antagonistischer Begriffspaare wie digital/analog, digital/präsent, digital/nicht-digital usw. haben zu einer unspezifischen Breite allein schon des Begriffs *digital* geführt.³⁵

-
- 27 Es interessiert dabei besonders die Frage, welches Verständnis von ›Theorie‹ den Forschungen in den DH zugrunde liegt (z.B. Theorie als Praxis im Sinne von *doing theory*, z.B. Reckwitz, A., Rosa, H. (2021): *Spätmoderne in der Krise. Was leistet die Gesellschaftstheorie*. Berlin, S. 25ff. und Kleyermann, R. (2023): *Theorie*. In: AG DH Theorie (Hg.): *Begriffe der Digital Humanities*. (Anm. 12).
- 28 Vgl. Moretti: *Conjectures on World Literature*. Moretti prägte auch den Begriff des *Great Unread* im Anschluss an Margaret Cohen, der die riesigen Mengen an »primary source material that is not analyzed by scholars« bezeichnet (Darren Reid (2019): *Distant Reading, ›The Great Unread‹, and the 19th-Century British Conceptualizations of the Civilizing Mission. A Case Study*. In: *Journal of Interdisciplinary History of Ideas*, Bd. 15, S. 1. <https://journals.openedition.org/jihi/435> [30.10.2023]).
- 29 Gius, E./Jacke, J. (2017): *The Hermeneutic Profit of Annotation: On Preventing and Fostering Disagreement in Literary Analysis*, In: *International Journal of Humanities and Arts Computing* 11.2, S. 233–254, S. 240.
- 30 Schröter, J., Du, K., Dudar, J., Rok, C., Schöch, C. (2021): *From Keyness to Distinctiveness – Triangulation and Evaluation in Computational Literary Studies*. In: *Journal of Literary Theory* 15 (1–2), S. 81–108.
- 31 Gramelsberger, G. (2023): *Philosophie des Digitalen zur Einführung*, Hamburg.
- 32 Stalder: *Kultur der Digitalität*.
- 33 Sonnenberg-Schrank, B. (2019): *Pathological Philology. Desire, Lack & the Digital Humanities*. In: Kelemen, P., Pethes, N. (Hg.): *Philology in the Making. Analog/Digital Cultures of Scholarly Writing and Reading*. Bielefeld, S. 57–69, S. 58.
- 34 *whatisdigitalhumanities.com* [30.10.2023].
- 35 Zu einer Definition des Begriffs und seiner historisch-semanticen und genealogischen Herleitung in der Literaturwissenschaft vgl. Jannidis u.a.: *Digital Humanities. Eine Einführung*, S. 59ff.

Einen ersten Überblick über wesentliche Bereiche der *Digital Humanities* gibt die Einleitung eines Sonderhefts des *Digital Journal of Philology*, herausgegeben von Jan Horstmann und Frank Fischer.³⁶ Dort wird die für die digitale Literaturwissenschaft maßgebliche Bezeichnung *Computational Literary Studies* (CLS) genannt, die sich gegenüber den vormals *Digital Literary Studies* – ursprünglich ein von James O'Sullivan von 2016 begründetes Journal – durchgesetzt hat³⁷, insbesondere da, wo es um computergestützte Methoden geht.

Die Polysemie der »digital humanities« wird dort auf dreifache Weise genealogisch hergeleitet – von erstens »digitized humanities« als »Aufbau, [...] Verwaltung und Verarbeitung digitalisierter Archive«, zweitens »numerical humanities« als »mathematische Abstraktionen geisteswissenschaftlicher Inhalte sowie formale Modelle« und drittens »humanities of the digital« als Untersuchung von »Online-Kommunikation und -Communitys bzw. allgemein Inhalte, die born-digital sind«. ³⁸ Bei der dritten Kategorie geht es um die Beobachtung von Digitalisierung in der Gesellschaft.

Die vorliegende Herausgeberschaft klammert im Gegensatz zu Horstmanns und Fischers Band den die Digitalisierung analoger Dokumente betreffenden ersten Bereich der digitalen Editionen aus, einem wesentlichen Zweig innerhalb der DH,³⁹ und behandelt auch nicht das inzwischen ebenso unüberschaubare Feld der Social-Media-Forschung im Anschluss an einen *digital turn*, der sich mit der dritten Kategorie von Roth deckt. In den hier versammelten Beiträgen steht die zweite Kategorie der Klassifikation Roths im Vordergrund, die die »Bedeutung für klassische literaturwissenschaftliche Fragestellungen« sowie die »Vermittlungsversuche zwischen traditionellen Ansätzen und digitalen Methoden und deren »epistemischen Ausgangslagen«⁴⁰ fokussiert.⁴¹ Denn häufig werden diese als inkompatibel wahrgenommen.⁴² Dass »computationale Verfahren häufig nicht traditionelle Fragen der Literaturwissenschaft beantworten – auch wenn dies als eigentliches Ziel gesetzt war –, sondern dass sie vielmehr etablierte Forschungsdiskurse um weitere Fragemöglichkeiten ergänzen«⁴³, scheint sich häufig noch als etwas ernüchterte Synthese in der Literaturwissenschaft abzuzeichnen.

36 Vgl. Horstmann, J./Fischer, F.: »Einleitung«. In: Dies. (Hg.): *Digital Methods in Literary Studies*. Sonderheft # 6 von *Textpraxis. Digital Journal for Philology* 1 (2022). URL: <https://www.textpraxis.net/en/jan-horstmann-frank-fischer-einleitung>, DOI: <https://doi.org/10.17879/64059433528>. S. 1.

37 Vgl. Horstmann/Fischer: Einleitung, S. 1.

38 Vgl. ebd., S. 1. Diese Unterscheidung basiert auf dem Aufsatz von Camille Roth: »Digital, Digitized, and Numerical Humanities«. In: *Digital Scholarship in the Humanities* 34.3 (2019), S. 616–632, hier S. 616. DOI: 10.1093/llc/fqy057.

39 An dieser Stelle seien nur einige Datenbanken literarischer Texte und digitalisierter Texteditionen genannt, die in verschiedenen Textformaten (TEI, XML, TXT, TCF, HTML) vorliegen: Zeno.org, Deutsches Textarchiv (DTA), TextGrid Repository, DraCor, KOLIMO, Gutenberg-Projekt, u.v.m.

40 Horstmann/Frank: Einleitung, S. 4.

41 Vgl. zur mathematischen Operationalisierung von philosophischen Fragen und einer operativen Epistemologie auch G. Gramelsberger (2020): *Operative Epistemologie. (Re-)Organisation von Anschauung und Erfahrung durch die Formkraft der Mathematik*. Hamburg: Meiner. DOI: 10.28937/978-3-7873-3900-6.

42 Vgl. Horstmann/Frank: Einleitung, S. 4.

43 Vgl. Horstmann/Frank: Einleitung, S. 4.

Dynamische Relationen zwischen Literaturwissenschaft und Informatik

Um jenseits eines inzwischen spezialisierten Expert:innen-Diskurses in den DH die Frage zu stellen, welche Rolle der Informatik in der transdisziplinären Zusammenarbeit mit der Literaturwissenschaft idealerweise zukommen sollte – als Hilfswissenschaft oder als treibender Motor für zukünftige Denkweisen in den *Humanities* – sollen im Folgenden einige dynamische Relationen zwischen Literaturwissenschaft und Informatik in Bezug auf (mögliche gemeinsame) Themen, Forschungsfelder und Modelle skizziert werden.

In der Arbeit mit selbstlernenden Algorithmen hat man es anscheinend zunehmend mit *black boxes* als Teil des Forschungssettings zu tun. Daher wird gerade die informatische Expertise, aber auch das *out-of-the-box* Denken, das in den *Humanities* und in Bezug auf die Künste eine lange Theorietradition aufweist, im Bereich von weniger kontrollierbaren maschinellen Prozessen im Zeitalter des *Computing* ›unserer‹ Realität immer relevanter. In der Literatur und anderen Künsten hat man es mit ästhetisch-kulturellen Symbolsystemen als hochkomplexen multimodalen Artefakten zu tun, deren Codierungs-,⁴⁴ Kontextualisierungs- oder Semantisierungsprozesse nur mit entsprechendem theoretischen Vorwissen sichtbar(er) werden. Auch für die Analyse und Erklärbarkeit des Outputs von vortrainierten Machine-Learning-Modellen mit integrierter Transformer-Architektur sind fachliche Expertisen von hoher Relevanz, um damit zumindest teilweise Rückschlüsse auf den Input und die maschinelle Bedeutungsverarbeitung ziehen zu können – wie jüngst in einer Studie der Informatik gezeigt werden konnte.⁴⁵

Es gilt also immer noch kulturwissenschaftliche Ansätze wie Theorien zur Sinn- und Bedeutungskonstruktion zu vermitteln, wenn ›unsere Welt‹, ›unsere Realität‹ und ›unsere Erfahrungen‹ maßgeblich von digitalen Technologien und Differentiationen, Datenförmigkeiten und -kulturen sowie Miningprozessen geprägt sind, die sich in der gegenwärtigen Wahrnehmung als höchst fluktuativ, dynamisch, relational und variabel erweisen. Die real stattfindenden Wechselwirkungen und gegenseitigen Einflüsse zwischen Technologien, Theorien und Erfahrungen sind vielfältig: Einerseits transformieren neue Technologien einschneidend unsere sozialen, kulturellen, ästhetischen, medialen und materialen Erfahrungen und Vorstellungen, die bereits als *mixed realities* (Chris Salter)⁴⁶ bezeichnet werden. Demgegenüber affizieren und befeuern auch die ästhetisch-kulturellen, literarisch und medial vorgeprägten, fiktionalen, theoretischen und historischen Kon- und Präfigurationen unseres Wissens, unserer Vorstellungen und Imaginationen zugleich das Kreieren neuer Technologien und ihre gesellschaftlichen Deutungsdimen-

44 Vgl. Koschorke, A. (2004): Codes und Narrative. Überlegungen zur Poetik der funktionalen Differenzierung. In: Walter Erhart (Hg.): Grenzen der Germanistik. Rephilologisierung oder Erweiterung? Stuttgart, Weimar: Metzler, S. 174–185.

45 Vgl. Herm, L.-V., Heinrich, K., Wanner, J., Janiesch, C. (2023): Stop Ordering Machine Learning Algorithms by their Explainability! A User-Centered Investigation of Performance and Explainability. In: International Journal of Information Management 69 (2023), 102538, <https://doi.org/10.1016/j.ijinfomgt.2022.102538>

46 Vgl. Salter, C. (2022): Sensing Machines. How Sensors Shape Our Every Day Life, Cambridge. (<https://blog.zhdk.ch/immersivearts/>, 08.10.2023)

sionen.⁴⁷ Diese (affektiven) epistemischen Dynamiken dabei nicht nur mit historischen Kontexten zu rahmen oder abzugleichen, sondern diese historisch-kulturellen Kontexte selbst wiederum auf computationeller Grundlage (unter Berücksichtigung der eigenen Forschungsperspektive) sorgfältig zu überprüfen und empirisch zu belegen, sind die neuen Herausforderungen in den digitalen Geisteswissenschaften. Ob beispielsweise die Ergebnisse des *Distant Reading*, wie das ›Lesen‹ mit quantitativ-statistischen Methoden bezeichnet wird, auch tatsächlich valide sind, muss wiederum ›händisch‹ durch traditionelles Methodenwerkzeug (z.B. *Close Reading* = Textstellenlektüre) und die Expertise der Literaturwissenschaft überprüft werden. Gerade im Bereich der Literaturgeschichte könnte dieses Vorgehen zu neuen Erkenntnissen und unter Umständen zu alternativen Literaturgeschichtsschreibungen führen.⁴⁸

In diesem hier grob abgesteckten Rahmen kommen den notwendigerweise neu (und gemeinsam) zu operationalisierenden Fragestellungen von kulturwissenschaftlich orientierter Literaturwissenschaft und Informatik sowie der Validität ihrer Ergebnisse eine hohe gesellschaftliche Bedeutung zu. Gerade in den DH, einer digitalen Literatur- und Medienwissenschaft werden daher Konzepte und Ansätze wie Mixed-Methods-, Multi-Methods- oder Entangled-Methods- sowie diffraktive Designs diskutiert, die sich entweder einer multiplen Perspektivität öffnen und/oder die Reflexion und Diffraktion⁴⁹ der eigenen partialen Perspektive und ihrer Situiertheit im Anschluss an Donna Haraway⁵⁰ und Karen Barad⁵¹ im Forschungsprozess berücksichtigen. Ins Auge gefasst wird dabei die disziplinäre Verschränkung mit der eigenen Theorie- und Methodentradition und der Art und Weise des eigenen Schließens und Rasonierens.⁵²

47 Vgl. Koch, L., Nanz, T., Pause, J. (2018): Imagined Scenarios of Disruption. A Concept. In: Koch, L., Nanz, T., Pause, J. (Hg.): *Disruption in the Arts. Textual, Visual, and Performative Strategies for Analyzing Societal Self-Descriptions*. Berlin/Boston, S. 63–81.

48 Vgl. dazu auch den Beitrag von Femmer und Lucke im vorliegenden Band. Das ist zum Beispiel der Fall, wenn Epochenklassifizierungen nicht mit historisch-empirischen Datenlagen übereinstimmen und andere Gruppenzuordnungen (u.a. durch das unüberwachte Lernen) erforderlich werden.

49 Vgl. Eickelmann, J. (2020): Digitale Medien und Methoden. Jennifer Eickelmann zu Diffraktion als Methode. In: *Open Media Studies. Blog zu Open Access und Open Science in der Medienwissenschaft*. Marburg (aktualisiert am 27.09.2023). <https://mediastudies.hypotheses.org/2230>; vgl. auch Eickelmann, J. (2017): »Hate Speech« und Verletzbarkeit im digitalen Zeitalter. Phänomene mediatisierter Missachtung aus Perspektive der Gender Media Studies, Bielefeld und Eickelmann, J., Meis, M. (2023): Diffraktive Ethnographie Sozialer Medien: Diskurs – Ästhetik – Materialität. In: S. Stollfuß et al. (Hg.): *Handbuch Digitale Medien und Methoden*. Wiesbaden, S. 1–25 (https://doi.org/10.1007/978-3-658-36629-2_14-2, 08.11.2023).

50 Haraway, D. (2007): Situiertes Wissen. Die Wissenschaftsfrage im Feminismus und das Privileg einer partialen Perspektive. In: Hark, S. (Hg.): *Dis/Kontinuität. Feministische Theorie*, Wiesbaden, S. 305–322.

51 Barad, K. (2013): Diffraktionen: Differenzen, Kontingenzen und Verschränkungen von Gewicht. In: *Geschlechter Interferenzen. Wissensformen – Subjektivierungsweisen – Materialisierungen*. Hg. von Corinna Bath/Hanna Meißner/Stephan Trinkaus/Susanne Völker. Berlin, S. 27–67 und Karen Barad, K. (2015): *Verschränkungen*. Berlin.

52 Vgl. Kleymann: Datendiffraktion.

›Fremde‹ Wissenschaftskulturen?

Die mit Blick auf Forschungsgegenstand, Theorie und Methode sehr unterschiedlichen akademischen Fachdisziplinen Literaturwissenschaft und Informatik werden in diesem Band in einen gleichermaßen herausfordernden wie auch gewinnbringenden Dialog gebracht, bei dem die zwischen beiden Disziplinen auftretenden *obstacles épistémologiques* (Gaston Bachelard) die Chance bieten, die eigenen disziplinären Wissens- und Erkenntnisgrenzen zu überschreiten. Die Überwindung dieses »Erkenntnishindernisses« erscheint dabei als ein Akt der (Selbst-)Erkenntnis für die Form des eigenen Wissens (und seiner Denkstile) sowie seiner Voraussetzungen. Diese verweisen auf gemeinsame Wurzeln des Denkens,⁵³ bevor Differenzierungsprozesse und Systemabschließungen⁵⁴ die modernen Disziplinen Natur- und Geisteswissenschaften gegen Ende des 19. Jahrhunderts »endgültig« voneinander trennten (Wilhelm Dilthey).⁵⁵

Nicht nur für die Literaturwissenschaft ist die Informatik eine ›fremde‹ Wissenschaft – beispielsweise Programmierungen als Teil von Software-Engineering, Computersprachen und Entwicklungsumgebungen, Algorithmen, Datenstrukturen, -management und -formate, Data-Mining-Verfahren, KI-Systeme, Datenbanken, computerlinguistische Sprachverarbeitung sowie nicht zuletzt statistischer und mathematischer Verfahren (z.B. Clusteranalysen, Vektorisierung). Auch die Informatik sieht sich mit den Anforderungen von äußerst heterogenen (hermeneutischen, strukturalistischen, poststrukturalistischen, diskursanalytischen, systemtheoretischen, medien- und kulturwissenschaftlichen) Zugängen und Methoden der Literaturwissenschaft⁵⁶ konfrontiert, die auf einer inhaltlichen und semantischen Ebene häufig Sinn und Bedeutung von Texten in ihren jeweiligen historischen Kontexten verhandeln.

Solche hermeneutischen, interpretativen und verstehenden Verfahren stehen häufig noch konträr, auch komplementär, zu den quantitativen Datenauswertungen, deterministischen und rechnerischen Verfahren, statistischen Häufigkeitsmessungen und Programmen der Informatik.⁵⁷

53 Vgl. zu den ersten (und unzureichenden) mathematischen Formalisierungsansätzen philosophischer Inhalte Lucke: Schillers *Ästhetische Briefe* als Literatur, S. 30. Vgl. auch Nerbonne: Die Informatik als Geisteswissenschaft.

54 Vgl. Niklas Luhmann (1998): Gesellschaftliche Struktur und semantische Tradition. In: Ders.: Gesellschaftsstruktur und Semantik. Studien zur Wissenssoziologie der modernen Gesellschaft. Bd. 1, Frankfurt a.M. S. 9–71, S. 27ff.

55 Vgl. Roland Borgards/Harald Neumeyer/Nicolas Pethes/Yvonne Wübben: Vorwort. In: Borgards u.a. (Hg.): Literatur und Wissen. Ein interdisziplinäres Handbuch. Stuttgart, S. 1.

56 Zur historischen Anbahnung der methodenpluralistischen Situation der Literaturwissenschaft, vgl. Konle, L., Jannidis, F., Martus, S. (2021): Disruptionen der Literaturwissenschaft am Beispiel der DVjs. Methodische Validierung durch Simulation und Anwendung. In: Fabrikation von Erkenntnis: Experimente in den Digital Humanities. Sonderband 5 der ZfdG. Hg. v. von M. Burghardt, L. Dieckmann, T. Steyer, P. Trilcke, N.-O. Walkowski, J. Weis, U. Wuttke, Melusina Press, S. 2. DOI: <https://doi.org/10.17175/sb005>

57 Evelyn Gius spricht auch von non-deterministischen Zugängen der literaturwissenschaftlichen, speziell hermeneutischen, Verfahren und von deterministischen (=entscheidbaren) der Informatik (insb. des Data Minings), vgl. Gius/Jacke: Informatik und Hermeneutik. Eine Ausnahme dazu bildet der KI-Ansatz.

Demgegenüber geht es besonders in literarischen Texten eher um das Einzelne und Besondere, wie das einzelne Phänomen oder eine spezielle, mal individuelle, mal kontextabhängige Bedeutung. Aber auch das Gegenteil ist der Fall: Unentscheidbarkeit, Vieldeutigkeit bis hin zur Beliebigkeit oder das Unbestimmte, das Inkommensurable (Heidegger), Offene (Umberto Eco) wie auch Widersprüchliche und Paradoxe (Friedrich Schiller, Dieter Mersch) perspektivieren oft im selben Text die Grenzen der (eigenen) Verstehbarkeit und des Nicht- bzw. Vorbegrifflichen, die in der Literaturwissenschaft qualitativ offene Zugangsweisen zu (literarischen) Texten und ihren intermedialen Bezügen zu anderen Künsten notwendig machen. Die Untersuchungsgegenstände bestimmen häufig die Untersuchungskategorien⁵⁸, denn die Vielzahl und Vielfalt der Gegenstände erstrecken sich von Autorschaft(skonstitutionen), Gattungsfragen, Themenkomplexen und Motivtraditionen über ästhetisch-poetologische Programme, Epochisierungen und kulturellen Wandel bis hin zu ›Lektüren‹, Erzähltechniken, rhetorischen und sprachlichen Strategien, künstlerischen Praktiken, medial-materialen Bedingungen u.v.m. Einen Text in seinem spezifischen Kontext (wie Werk-, Epoche-, Gattungskontext, sozio-kultureller Problematik usw.) zu befragen und zu verstehen entscheidet über die Wahl des Zugangs. Je nach Fragestellung (z.B. nach der Gattung) wählt man die passende Methode (z.B. die Erzähltextanalyse) aus. Bei der Analyse von spezifischen und seltenen Phänomenen in einem literarischen Text, sog. *sparse data* oder *rare events*⁵⁹, wäre im Kontext eines digitalen Forschungssettings beispielsweise die Überlegung ratsam, ob statt des Mittelwerts eher der Median in der statistischen Analyse verwendet werden sollte.⁶⁰

Ähnliche Modelle in beiden Disziplinen finden sich im Umgang mit Differenz-, Logik- und Formkalkülen⁶¹ oder mit Übersetzungen von einer rein hypothetischen, geistigen und gedanklichen (Vor-)Arbeit in sprachliche oder zeichenbasierte Operationen.⁶² Weitere Ähnlichkeiten finden sich in der (binären) Modellierung von Codes und Zeichen, sprachlichen und syntaktischen Strukturen, (latenten) Tiefen- und Oberflächenstrukturen bzw. Relationen von Sichtbarkeiten und Unsichtbarkeiten.

Aus Sicht der Informatik sind wegen komplexer kulturwissenschaftlicher Fragestellungen und Analysen der Literaturwissenschaft teilweise multimodale Ansätze oder Verfahren aus dem Bereich des *Machine Learnings* gefordert, die bei dieser Anwendung gleichzeitig auf die Probe gestellt werden. Denkbare zukünftige Anwendungsgebiete von LLMs in Kombination mit hermeneutischen Verfahren in der Literaturwissenschaft wären vielleicht die rhetorische Figurenanalyse, metaphorisches Schreiben⁶³,

58 Vgl. <https://fortext.net/tools/tols/catma>.

59 Vgl. dazu Gius/Jacke: Informatik und Hermeneutik.

60 Denn nach dieser Logik hätte der statistische Ausreißer (auch der Median) einen höheren heuristischen Wert als der statistische Mittelwert.

61 Dirk Bäcker (1993): *Kalkül der Form*. Frankfurt a.M. und Sybille Krämer (1991): *Berechenbare Vernunft: Kalkül und Rationalismus im 17. Jahrhundert*, Berlin, Boston: de Gruyter. DOI: 10.1515/9783110847079.

62 Vgl. Gramelsberger: Philosophie des Digitalen. Vgl. auch Nerbonne: Die Informatik als Geisteswissenschaft.

63 Vgl. dazu den SFB 1475 »Metaphern der Religion« – Religiöse Sinnbildung in sprachlichen Prozessen <https://ceres.rub.de/de/forschung/projekte/sfb1475/>, 08.10.2023). Der SFB verfügt über ein

oder kulturelle und ästhetische Differentiationsprozesse.

Der Perspektive der Informatik ist also im Kanon der geisteswissenschaftlichen Theorien ein (neuer) Platz zuzuweisen, um neue Forschungsperspektiven, digitale Technologien und Verfahren einerseits bereit zu stellen, andererseits auch angemessen kommentieren zu können, was sich den herkömmlichen Zugängen der Geisteswissenschaften entzieht.

Probleme der Messbarkeit: Operationalisierung, Formalisierung und Validierung

Die Vorteile von computer- und rechengestützten Methoden in den Geistes- und Kulturwissenschaften liegen auf der Hand: Digitale Analysen und Klassifizierungen großer Korpora sind möglich (»Literature as Big Data«⁶⁴), ebenso quantitative Erkenntnisse, die vom Menschen so nicht produziert werden können, so dass sich literaturwissenschaftliche Forschung in die Richtung einer evidenz- und empiriebasierten Forschung (weiter)entwickeln kann. Es können viele Texte, Epochen, Gattungen bzw. Genres oder Autor:innen (z.B. im Hinblick auf Ähnlichkeit oder Disruption⁶⁵) miteinander verglichen und gleichzeitig darüber hinaus mehrere Kontexte (sog. Referenzkorpora) herangezogen werden.⁶⁶ Groß angelegte maschinelle und skalierbare Vergleiche zu vielen weiteren Textkorpora (sog. Makroanalysen) stellen kein Problem mehr dar. So fortschrittlich alles auf den ersten Blick wirkt, ergeben sich daraus nicht zu unterschätzende erkenntnistheoretische Implikationen für die Literaturwissenschaft wie zum Beispiel die Problematik, dass eine datenbasierte Textanalyse literarische und literaturwissenschaftliche Daten tendenziell undifferenzierter erscheinen lässt.⁶⁷

Die Messbarkeit, definiert als »Zuordnen von Zahlen zu Objekten nach bestimmten Regeln«⁶⁸, führt weitere in den Naturwissenschaften, aber auch den empirischen Sozialwissenschaften längst bekannte Probleme mit sich, denn »die Schwierigkeiten einer solchen Begriffsbildung liegen natürlich in den »bestimmten Regeln«.⁶⁹

Bevor man versucht, eine Eigenschaft zu messen, sollte man sich zunächst überlegen, was eigentlich eine Eigenschaft ist. [...] Man kann aber festhalten, daß eine Eigenschaft von Objekten durch Beziehungen (Relationen) zwischen diesen Objekten festgelegt

methodisches Instrumentarium, das in allen Teilprojekten genutzt wird: Metaphernannotation, hermeneutische Interpretation und computergestützte Analysen.

64 Sonnenberg-Schrank: *Pathological Philology*, S. 60.

65 Vgl. Konle u.a.: *Disruptionen der Literaturwissenschaft am Beispiel der DVJs*.

66 Vgl. zum Beispiel das Projekt des Goethe-Wörterbuchs (Thesaurus-Wörterbuch) der Hamburger Arbeitsstelle der Akademie der Wissenschaften zu Göttingen zur historischen Semantik und Erschließung des Individualwortschatzes von Goethe (<https://adw-goe.de/forschung/forschungsprojekte-akademienprogramm/goethe-woerterbuch/>, 08.10.2023).

67 Vgl. Horstmann/Fischer: *Einleitung*, S. 2.

68 Heinrich Wottawa: *Psychologische Methodenlehre. Eine orientierende Einführung*. Weinheim/München 1993, S. 60

69 Wottawa: *Psychologische Methodenlehre*, S. 61.

wird. So spricht man von der Eigenschaft ›rot‹, weil es Objekte gibt, die zusammen die Menge roter Objekte bilden. Dieser Menge steht eine Menge der ›nicht roten‹ Objekte gegenüber.⁷⁰

Ein Problem der Validität⁷¹ ist zum Beispiel das der adäquaten Formalisierung von kulturwissenschaftlichen Kategorien bzw. der Operationalisierung der (kultur- oder literaturwissenschaftlichen) Fragestellung in ausgewählte Indikatoren – in der Sprache der Informatik: in generierte Daten – für das, was man inhaltlich vollständig erfassen und numerisch messen möchte.⁷²

Operationalizing means building a bridge from concepts to measurement, and then to the world. In our case: from the concepts of literary theory, through some form of quantification, to literary texts.⁷³

In der Informatik muss man Daten für Programme formal aufbereiten; das ist ein notwendiger Transformationsprozess, bei dem Operationalisierungs- und Validierungsprobleme auftreten können. Diese Problematik ist in der Informatik gut bekannt und erforscht.

Für die Informatik wären daher also Analysen von spezifischen Anforderungen (*Requirements*) relevant, insbesondere auch vor dem Hintergrund methodologischer Debatten in der Informatik über die Validität von Ergebnissen. Mit dem Entwerfen von Programmen ist es hier nicht getan, es folgt eine Überprüfung, ob beispielsweise die *Findings* des Programms auch als brauchbare Ergebnisse im Sinne der Aufgaben- und Fragestellung taugen oder womöglich verzerrt, falsch (Stichwort: Bias-Effekte) oder überinterpretiert bzw. gar nicht aussagekräftig sind.⁷⁴

70 Wottawa: Psychologische Methodenlehre, S. 61.

71 »Validität bezieht sich auf das Ausmaß, in dem die gemessene Variable das ihr zugrundeliegende Konstrukt vollständig erfaßt.« (Manstead, A. und Semin, G. (1997): Methoden der Sozialpsychologie: Ideen auf dem Prüfstand. In: Stroebe, W. et al.: Sozialpsychologie, Berlin/Heidelberg, S. 100.

72 Die Gültigkeit von Theorien vor dem »Ableitung von Hypothesen« oder der »Formalisierung des gedanklichen Ansatzes« ist aber gerade nicht so einfach empirisch überprüf- und testbar, da in die Formalisierungs- und Operationalisierungsarbeit häufig zusätzliche Annahmen miteinfließen, die nicht in der Theorie enthalten sind (vgl. H. Wottawa (1993): Psychologische Methodenlehre. Eine orientierende Einführung, Weinheim/München, S. 18). Hier müsse man mit Mitteln der formalen Logik operieren und prüfen, ob die Hypothesen aus den Theorien ableitbar seien (sonst könne man nicht entscheiden, ob die »Theorie ungünstig zu beurteilen ist, oder ob die Theorie gültig und nur diese zusätzlichen Annahmen nicht erfüllt waren«, ebd. Vgl. auch die Annotationen der Verfasserin zu Schröter: Mixed Methods (Anm. 6)).

73 Moretti, F. (2013): Operationalizing: or, the function of measurement in modern literary theory. In: Pamphlets of the Stanford Literary Lab 6 (2013), S. 1–13, S. 1.

74 Vgl. den Beitrag von Femmer/Lucke im vorliegenden Band. Vgl. zu Bias-Effekten auch den Beitrag von Lucke im vorliegenden Band.

Paradigmenwechsel in der Literaturwissenschaft?

Viele sprechen schon von einem Paradigmenwechsel.⁷⁵ In jüngerer Zeit mehren sich auch insbesondere Studien in den DH-Wissenschaften, die nahezu standardisierte Forschungsdesigns und Publikationsformate verwenden, die sich offensichtlich an den Natur- und Ingenieurwissenschaften orientieren und neben der Fragestellung Methoden und Materialien sowie Aufbereitung und Auswertung von Forschungsdaten und Ergebnissen aufführen.⁷⁶ Derartige und andere den Diskurs der DH begleitende Forschungspraktiken werden häufig von Seiten der Geistes- und Literaturwissenschaften noch mit Skepsis betrachtet.⁷⁷ Beim digitalen Methodenimport scheint es um mehr zu gehen als die bloße Implementierung von digitalen Tools in den (*Digital*) *Humanities* oder in der digitalen Lehre; ihre Bedeutung jedoch gleich auf der Höhe einer Transformation⁷⁸ oder eines Paradigmenwechsels in der Literaturwissenschaft anzusetzen, ist umstritten.⁷⁹ Auch die Integration der Informatik in andere Wissenschaften wie die Biologie war mit großen Veränderungen verbunden und hat eine neue Teildisziplin hervorgebracht: die Bioinformatik. Aufhalten lässt sich der Einzug von digitalen Tools und Methoden in die Literaturwissenschaft wohl eher nicht, zu diskutieren und zu gestalten ist jedoch die sinnvolle und praktikable Integration dieser in den Forschungsprozess von literatur- und kulturwissenschaftlicher Modellierung und Theoriebildung.

Eine mögliche neue Perspektive für die Geisteswissenschaften deutet sich in der folgenden Aussage von Thomas Weitin an:

Die in den Geisteswissenschaften im Ringen um Aufmerksamkeit weitgehend habitualisierte Orientierung an Originalität lässt uns das manchmal fast vergessen. Die unvermeidliche Empirie digitaler Methoden bietet womöglich eine Gelegenheit, die Kategorien origineller Erkenntnis selbst neu zu verhandeln.⁸⁰

75 Vgl. Weitin: Digitale Literaturgeschichte, S. 2.

76 Vgl. im CfP für die Tagung der DHD 2024 (Digital Humanities im deutschsprachigen Raum): »Eine gute Einreichung folgt den Prinzipien guter wissenschaftlicher Arbeit und beschreibt in inhaltlich und formal strukturierter Weise Forschungsfrage, Material, Methode und Ergebnisse.« (<https://dih-hum.de/aktuelles/call-for-papers-dhd2024>, 13.07.2023)

77 Vgl. Krämer, S., Huber, M. (2018): Dimensionen Digitaler Geisteswissenschaften. In: Zeitschrift für digitale Geisteswissenschaften. Wolfenbüttel. text/html Format. DOI: 10.17175/sb003_013, 08.10.2023.

78 Interessanterweise bezeichnen die DH sich selbst als Transformationswissenschaft, vgl. Weitin: Digitale Literaturwissenschaft, S. 1.

79 Dabei rückt der Begriff *Paradigma* selbst in den Blick und wird ›semantisch weiter ausdifferenziert‹, vgl. Kleymann: Datendiffraktion.

80 Weitin: Digitale Literaturgeschichte, S. 4.

Die Rolle der (digitalen) Hermeneutik⁸¹

In der kontroversen Debatte um die erkenntnistheoretischen Auswirkungen beim Einsatz von digitalen Methoden stellt sich mitunter die Frage, warum ausgerechnet die Hermeneutik, die als klassische literatur- und geisteswissenschaftliche Methode der Interpretation bekannt ist, als Sparringspartnerin der DH ins Feld geführt wird.⁸² Dabei fordern die Digital Humanities nicht nur hermeneutische, sondern auch anti-hermeneutische Ansätze (wie dekonstruktive Ansätze) heraus. In der Literaturwissenschaft war die Hermeneutik u.a. aufgrund des Problems des Verstehens durch den Strukturalismus in den 1960ern und späteren Poststrukturalismus teilweise abgelöst worden.

Für die Konzentration in den DH auf die Methode der Hermeneutik gibt es einige mögliche Erklärungen:

1. Der Streit zwischen dem Strukturalismus, der für die mathematische Struktur der Sprache steht (z.B. als zählbare Seme⁸³) und der Hermeneutik als Interpretation, die sich einer mathematischen Auszählung gegenüber sperre, beschäftigte bereits Heidegger.⁸⁴ Die Interpretation gleiche einer unerschöpflichen Analyse, was laut Beda Allemann mit Heidegger ein Argument für die Interpretation sei, da das interpretatorische Verfahren auch Unausgesprochenes, Nicht-Empirisches bzw. Inkommensurables zeige. Das Werk sei immer etwas, das sich entziehe.⁸⁵ Nach Heidegger sei also das (literarische) Werk nicht (nur) empirisch zu erforschen, da es sich seiner Auffassung nach um ein Konstrukt handle, welches hinter bzw. jenseits seiner empirischen Zugänglichkeit liege. Hier befindet man sich allerdings schon im Grenzbereich zur Metaphysik.⁸⁶

2. Eine weitere Erklärung für den Fokus auf hermeneutische Verfahren lässt sich in einer Formalisierungs- und Operationalisierungsproblematik dieses Verfahrens auffinden:

Die Modellierung geisteswissenschaftlicher Artefakte geht dabei natürlich stets einher mit der Formalisierung von Forschungsfragen, die – angepasst an die Modellie-

81 Vgl. zur Rolle der Hermeneutik in der Literaturwissenschaft und im Projekt heureCLÉA auch Gius/Jacke: Informatik und Hermeneutik.

82 Jan Horstmann und Frank Fischer fragen zum Beispiel »nach der Relation von quantitativ-numerischen Analysemethoden und hermeneutischem Erkenntnisinteresse«, vgl. Horstmann/Fischer: Einleitung, S. 3.

83 Ein Sem ist ein linguistischer Begriff für die kleinste Einheit der Bedeutung eines Worts: So beinhaltet »Mutter« z.B. das Sem »weiblich«. »Königin«, »Frau«, »Löwin« und »Tochter« enthalten auch das Sem »weiblich«.

84 Vgl. die Rezeption Heideggers im Anschluss an Beda Allemann (1954): Hölderlin und Heidegger, Zürich/Freiburg.

85 Übrigens ist das ein dekonstruktiver Gedanke *par excellence*.

86 An dieser Stelle soll nicht verschwiegen werden, dass Heidegger versuchte, neben den beiden Forschungsrichtungen des positiv-empirischen Strukturalismus und einer konstruktivistischen Interpretation etwas Drittes vorzuschlagen, dass keine »Aneignung« von Texten bzw. Theorien vorsehe, sondern als eine Art »Hin-Hören« bezeichnet werden könnte.

rungskategorien – von vornherein viel spezifischer formuliert werden müssen als man dies in klassischen hermeneutischen Verstehensprozessen erwartet.⁸⁷

Als eine eher qualitative Methode des ›Verstehens‹ repräsentiert die Hermeneutik eine diametral entgegengesetzte Forschungspraxis in den Geistes- und Kulturwissenschaften gegenüber den Formalisierungsvorgaben in der Informatik. Die ›Algorithmisierbarkeit‹ von literatur- und kulturwissenschaftlichen Forschungskategorien erscheint komplex, wenn die Gegenstände in der Literatur und anderer Künste offene Lesarten und mehrdeutige oder selbstreferentielle Sinnbezüge aufweisen. Eine Konsequenz dessen ist die derzeitige Rückbesinnung auf differenzfähige philologische und ästhetische Kategorien wie ›Epoche‹, ›Autor:in‹, ›Werk‹, ›Ähnlichkeit‹ (Foucault), ›Geschlecht‹ (wie bei der Autorschaftsattribution⁸⁸) oder ›Stil‹ (wie in der Stilometrie) in den *Digital Humanities*, die sich insbesondere hinsichtlich ihrer Unterscheidbarkeit (*distinctiveness*⁸⁹) für die Formalisierung von literaturwissenschaftlichen Analysekategorien bei der computationalen Analyse literarischer Texte zu eignen scheinen.

3. Im Kontext der DH wird außerdem vermehrt auf das Operationalisierungs- und Validierungspotential von (manuellen) hermeneutischen Verfahren für quantitative Verfahren und ihre Ergebnisse verwiesen:

This paper aims to solve these methodological issues concretely for the concept of distinctiveness and thus to lay the methodological foundation permitting to operationalize quantitative procedures in order to use them not only as rough exploratory tools, but in a hermeneutically meaningful way for research in literary studies.⁹⁰

Epistemologische vs. methodologische Problematik

Neben der epistemologischen Grundsatzdebatte in den DH wird auch Kritik an der ›epistemologischen Kritik‹ geübt: Diese stammt aus der DH-Community selbst. Häufig wird dabei auf die lange Tradition des Methodenstreits in den Sozial- und Sprachwissenschaften verwiesen (›paradigm wars‹).⁹¹

From the epistemological perspective, the question is: Are quantitative and qualitative research dealing with the same world of objects at all or do both fields of research construct their own worlds, which are not interconnected, and which do not refer to the respective other world?⁹²

87 M. Burghardt (2023): Kritische Überlegungen zum Algorithmizitätsbegriff. Blogpost, <https://dththeorien.hypotheses.org/1316>.

88 Vgl. Weitin, T. (2021): *Digitale Literaturgeschichte*, Berlin.

89 Vgl. Schröter u.a.: *From Keyness to Distinctiveness*.

90 Schröter u.a.: *From Keyness to Distinctiveness*, S. 81.

91 Vgl. Kleymann: *Datendiffraktion*.

92 Schröter u.a.: *Keyness and Distinctiveness*, S. 86.

Laut Schröter et al. (2021) gibt es derartige Kritik, die die erkenntnistheoretischen Konsequenzen für die Geisteswissenschaften durch ihren Digitalisierungsprozess betrifft, von Anbeginn der DH an und verhindert letztlich praktikable Lösungen in der wesentlichen Frage der Kompatibilität von traditionellen und digitalen Methoden.⁹³

Remarkably, parts of the intense debates of the last years on the relevance of digital humanities research to proper issues of traditional humanities, in particular arguments that reject such relevance in general,^[1] rest on such epistemological reasoning.⁹⁴

Demgegenüber vertreten Schröters et al. (2021) die folgende These:

This gap, which is well known in the social sciences and especially in the field of mixed methods research surrounding discussions on integration and triangulation (Hammersley 2008; Flick 2011; Kelle 2017), can be addressed either from an epistemological or from a methodological perspective. [...]

Instead, we will address the epistemic gap from its methodological perspective. In digital humanities research, the methodological dimension of this problem arises for all concepts that have undergone a quantitative and mathematical operationalization, such as ›topic‹ in topic modeling (Blei 2012), ›style‹ in stylometry (Burrows 2002; Herrmann/Dalen-Oskam/Schöch 2015) and also the concepts of keyness and distinctiveness. The methodological issue of integrating quantitative and qualitative reasoning can also be expressed in terms of validation: As we expect quantitative procedures such as candidate measures of distinctiveness to contribute to qualitative research, the idea of qualitative validation of quantitative procedures has to be developed.⁹⁵

Vorschläge für eine praktikable transdisziplinäre Zusammenarbeit

Im Anschluss an den eingeschlagenen theoretischen Perspektivwechsel im DH-Diskurs von Schröter et al. (2021) werden hier einige praktikable Vorgehensweisen in der transdisziplinären Arbeit vorgeschlagen:

1. Eine *Differenzierung von epistemologischen Problematiken* als Auslöser der Diskussion, die den Diskurs der DH seit seiner Genese begleitet, und *methodologischen Herausforderungen*

93 Vgl. Schröter u.a.: From Keyness to Distinctiveness, S. 86.

94 Schröter u.a.: Keyness and Distinctiveness, S. 86. »Interestingly, the intensity of these debates shows that scholars of computational literary studies in large parts aim at contributing to proper questions of literary studies (most comprehensively and recently Underwood 2019). So do we, but we will not contribute to the epistemological dimension of this gap, which extends to issues that have to be addressed on the level of metaphysical reasoning.« (Ebd.)

95 Schröter u.a.: Keyness and Distinctiveness, S. 86f. »However, the mere act of designating a specific quantitative procedure as a ›measure of keyness‹ or a ›measure of distinctiveness‹ does not guarantee that this procedure will give any insight into a qualitatively ambitious understanding of distinctiveness or keyness. We call this the epistemic gap between the output of a quantitative procedure and a qualitative expectation« (S. 86).

wird als hilfreich erachtet. Methodologische Debatten zwischen quantitativen und qualitativen Verfahren in den Sozial- und Humanwissenschaften halten entsprechende Lösungsansätze bereit (vgl. 2.).

Based on a structural definition of potential candidate measures for analyzing distinctiveness in the first section, we offer a systematic description of the issue of integrating quantitative procedures into a hermeneutically meaningful understanding of distinctiveness by distinguishing its epistemological from the methodological perspective.⁹⁶

Aus dieser Verlagerung der Perspektive auf die verschiedenen Paradigmen folgen beispielsweise folgende methodologische Fragestellungen: Welche Kategorien, Ergebnisse und Unterscheidungsmerkmale aus quantitativen Prozeduren haben also welche qualitative Relevanz und Bedeutung für die jeweiligen Forschungsgegenstände in der Literaturwissenschaft, in der Informatik oder in den DH? Und welche (qualitativen) Reduktionen (z.B. von Kontext oder Transparenz) werden durch mathematische Operationen vorgenommen (z.B. durch Vektor-Normalisierung⁹⁷ oder Skalieren beim K-means-Clustering mit dem System R⁹⁸)?

2. *Mixed-Methods-Designs*⁹⁹, *Multi-Methods*-, *Entangled-Methods*¹⁰⁰ oder *diffraktive Designs*¹⁰¹ und *transdisziplinäre Ansätze*:

Um der hohen multimodalen Komplexität literatur- und textförmiger Gegenstände und

96 Schröter u.a.: Keyness and Distinctiveness, S. 82. »We will address the methodological perspective of the issue of integrating the quantitative concept of distinctiveness into qualitative research by asking: How do we make sure that a quantitative procedure $f(w \mid A, B)$, as delineated in the first section, returns distinctive features which are distinctive from the point of view of qualitative meaning?« (S. 87)

97 Büttner u.a.: »Delta« in der stilometrischen Autorschaftsattribuion.

98 Beim K-means-Clustering werden die Variablen mithilfe von zwei mathematische Operationen vorverarbeitet: Skalierung der Variablen pro Attribut auf einen Mittelwert von 0 mittels Subtraktion des Mittelwerts und auf eine Standardabweichung von 1 mittels Division durch die Standardabweichung, s. Anhang im Downloadbereich auf GitHub.

99 Vgl. dazu besonders Schröter, J. (2023): Mixed Methods. In: AG Digital Humanities Theorie des Verbandes Digital Humanities im deutschsprachigen Raum e. V. (Hg.): Begriffe der Digital Humanities. Ein diskursives Glossar (=Zeitschrift für digitale Geisteswissenschaften/Working Papers, 2). Wolfenbüttel, 25.05.2023. HTML/XML/PDF. DOI: 10.17175/wp_2023_008 Das bedeutet die kombinierte Anwendung von z.B. geisteswissenschaftlichen und hermeneutischen sowie computergestützten, digitalen, quantitativen und statistischen Methoden oder auch *Triangulation*. So praktizieren es häufig die Sozialwissenschaften: Nach einer qualitativen Vorstudie wird statistisch-quantitativ weiter geforscht und überprüft, ob die Ergebnisse repräsentativ sind. Im Forschungsprozess der DH zeigt sich häufig ein umgekehrtes Verfahren: Erst erfolgt eine explorative (Vor-)Studie. Falls es Auffälligkeiten oder Abweichungen gibt, wird in den Texten manuell nachgeschaut und es werden Erklärungen (*post-hoc-explanations*) gesucht. Große Textmengen können prinzipiell zur Annahme verführen, etwas sei objektiv.

100 Kleymann: Datendiffraktion.

101 Die Konzepte der Verschränkung und Diffraktion gehen auf Karen Barad zurück, nach dem optischen Phänomen der Diffraktion benannt (vgl. Eickelmann: Digitale Medien und Methoden).

ihrer intermedialen Übergänge zu genügen, schlagen wir vor, theoretisch-deduktive und explorativ-induktive Methodiken miteinander zu verbinden und zu verschränken, wie in Mixed-Methods-Designs, außerdem inter- und transdisziplinäre Ansätze miteinander zu kombinieren. Zu Mixed-Methods-Designs existieren ebenfalls bereits Lösungen aus anderen Disziplinen wie den Sozialwissenschaften, der Statistik, der Informatik, der Computerlinguistik, der Sprachphilosophie usw.¹⁰²

Mixed-Methods-Ansätze kommen aus der Sozialforschung und setzen eine prinzipielle *Kompatibilität* (Vereinbarkeit) von quantitativen und qualitativen Methoden voraus:

Damit wenden sich Mixed-Methods-Ansätze tendenziell gegen eine Inkompatibilitätsthese, die von einer Unvereinbarkeit qualitativer und quantitativer Methoden aufgrund inkommensurabler Paradigmen ausgeht. Im Rahmen eines Methodendualismus werden von den Vertreter*innen quantitative Methoden mit dem (Post-)Positivismus und qualitative Methoden mit dem Konstruktivismus assoziiert.^[1] Ausgehend von der *Kompatibilität* wird in der Mixed-Methods-Forschung der Paradigmenbegriff semantisch weiter ausdifferenziert. So können Paradigmen neben Weltansichten (im Sinne von Glaubenssystemen) auch epistemische oder metaphysische Grundhaltungen, geteilte Überzeugungen zu Forschungsfragen sowie Durchführungen von Studien beschreiben.¹⁰³

Entangled-Methods-Ansätze (auch diffraktive Ansätze) beschreibt Rabea Kleymann im Anschluss an Karen Barad in ihrer intraaktiven Verschränkung, und zwar nicht als methodische Interaktion wie bei Mixed-Methods-Ansätzen, sondern als Intraaktion, in der Auseinandersetzung mit den (eigenen) Paradigmen, konstitutiven Differenzen und der eigenen Kontextualität, was auch Praktiken der Datenmodellierung miteinbeziehen würde.¹⁰⁴

Der Grundgedanke in Mixed-Methods- oder Entangled-Methods-Ansätzen ist, dass Methodendichotomien (wie qualitativ/quantitativ, close/distant, induktiv/deduktiv, empirisch/hermeneutisch, theoretisch/praktisch) prinzipiell überwunden oder die ›vermeintlichen‹ Gegensätze zumindest reflektiert werden.¹⁰⁵ Als Verbindung und Ver-

102 Vgl. Weitin: Digitale Literaturgeschichte, S. 3. Eine direkte Übertragung von Debatten über methodologische Schwierigkeiten der Mixed-Methods-Forschung aus den Sozialwissenschaften in die *Digital Humanities* sollte laut Schröter allerdings eher vermieden werden, da dort andere theoretische Paradigmen und Problemlagen zugrunde liegen (vgl. Schröter: Mixed Methods).

103 Kleymann: Datendiffraktion.

104 Vgl. Kleymann: Datendiffraktion. In diesem Sinne könnten nach Kleymann Methoden als Apparate verstanden werden, da sie »agentielle Schnitte [vollzögen]«. (Ebd.)

105 Vgl. Kleymann: Datendiffraktion. Die Debatte über solche Überwindungsversuche von methodologischen Differenzen, Dichotomien und gegensätzlichen Paradigmen im Kontext von Mixed-Methods-Ansätzen erinnern auffällig an die historisch-dualistischen Setzungen von ›Denken und Sein‹ in formallogischen Ansätzen des 18. Jahrhunderts im Rahmen einer zweiwertigen Aristotelischen Logik (vgl. dazu Lucke: Schillers *Ästhetische Briefe* als Literatur, Kap. 2). Die damaligen Formalisierungsversuche von Erkenntnis, insbesondere bei der ›Vereinigung von Gegensätzen‹ (Fichte), in frühen wissenschaftstheoretischen Abhandlungen der Moderne waren aber durch die Zwewertigkeit der Ansätze bereits zum Scheitern verurteilt (vgl. G. Günther (1991): Idee und Grundriß einer nicht-Aristotelischen Logik. Die Idee und ihre philosophischen Voraussetzungen, Hamburg).

schränkung von *Close* und *Distant Reading* schlägt Thomas Weitin das *Scalable Reading*¹⁰⁶ im Anschluss an Martin Müller vor, um Texte beispielsweise auf der Basis von Vektoren der häufigsten Wörter, die die Texte numerisch repräsentieren sollen, vergleichen zu können, jedoch mit der Gefahr des ›Verlusts von Kontext‹.¹⁰⁷ ›Kontextverknappung‹ sei aber auch ein Problem der traditionellen literaturwissenschaftlichen *Close Reading* Methode, die mit kontingenten Kontextualisierungen ohne Vergleichbarkeit operiere.¹⁰⁸ Die Integration von Unterschieden »zwischen manueller Einzeltext- und maschineller Korpusanalyse«¹⁰⁹ seien nicht zu unterschätzen. Evelyn Gius' Komplexitätsmodell der »Integration als eine Komplexitätsdimension von Erkenntnissen«¹¹⁰ stuft Grade der Komplexität zwischen Einfachheit und Komplexität von unterschiedlichen Phänomenen ab, die »durch unterschiedliche Methoden adressiert werden«, welche »wiederum unterschiedliche (Teil-)Phänomene« produzieren können, die bei der Interpretation wieder integriert werden müssen.¹¹¹ Der Grad der Komplexität des Phänomens bemesse sich dabei »an den zu operationalisierenden Einheiten«.¹¹² Der Forschungsgegenstand würde auf diese Weise zwar an Facettenreichtum gewinnen, aber ggfs. um den Preis (s)einer (hermeneutischen) ›Ganzheit‹.¹¹³ Zu prüfen wäre bei Mixed Methods-Ansätzen daher, ob die Kombination von Methoden eine multiperspektivische Erfassung des gleichen Phänomens (mit womöglich unterschiedlichen Ergebnissen) bietet oder ob unterschiedliche Phänomene dabei erst durch unterschiedliche Perspektiven produziert werden (können).¹¹⁴ Eine transparente Herleitung von theoretischen Implikationen und Herkunft von Modellen, Methoden und Daten sowie ein reflektierter und (logisch) plausibilisierter Umgang mit Formalisierung und Operationalisierung (insbesondere des Theorieanteils) in geeignete Indikatoren könnte dieser Herausforderung in erheblichem Maße Rechnung tragen.¹¹⁵

Im Bereich von *transdisziplinärer Zusammenarbeit* werden Mixed-Methods-Forschungsdesigns verwendet, wobei Analysekategorien, Ergebnisse und Anforderungen wechselseitig perspektiviert werden. Hermeneutische Verfahren werden diesbezüglich mit datenbasierten Methoden und bereits bestehenden Algorithmen und technischen Lösungen der Informatik für die Textanalyse kombiniert. Es werden dabei die Heuristiken digitaler Modellierungen transdisziplinär eruiert, wie die hermeneutische Validierung von digitalen Forschungsergebnissen und die Anforderungsanalyse bzw. -spezifikation, um zum einen der Formalisierung literaturwissenschaftlicher

106 Weitin: Digitale Literaturgeschichte, S. 116. Vgl. auch Armaselu, F., Fickers, A. (Hg.) (2024): *Zoomland. Exploring Scale in Digital History and Humanities*, Bd. 7 der Reihe *Studies in Digital History and Hermeneutics*, DeGruyter Oldenbourg.

107 Vgl. Weitin: Digitale Literaturgeschichte, S. 116.

108 Vgl. Weitin: Digitale Literaturgeschichte, S. 117.

109 Weitin: Digitale Literaturgeschichte, S. 116.

110 Kleymann: Datendiffraktion. Vgl. E. Gius (2019): *Computationelle Textanalysen als fünfdimensionales Problem: Ein Modell zur Beschreibung von Komplexität*. In: *LitLab Pamphlet 8* (2019), S. 1–20.

111 Vgl. Kleymann: Datendiffraktion.

112 Vgl. Kleymann: Datendiffraktion.

113 Vgl. Kleymann: Datendiffraktion.

114 Vgl. Schröter: *Mixed Methods*.

115 Vgl. die Annotationen der Verfasserin zu Schröter: *Mixed Methods*.

Forschungskategorien und zum anderen den Anforderungen kulturwissenschaftlich orientierter literaturwissenschaftlicher Fragestellungen zu genügen.¹¹⁶

Bei der explorativen Datenanalyse (EDA) werden Daten, Texte und Korpora auf bestimmte Auffälligkeiten (oder Trends) hin explorativ untersucht. Um der Gefahr spontaner Theoriebildung und einer zu langen Phase der Exploration vorzubeugen – indem so lange exploriert wird, bis man die gewünschten Ergebnisse erhält – kann es sinnvoll sein, theoretisch-deduktive Verfahrensweisen (wie Hypothesenbildung) in die EDA mit-einzubeziehen. Visualisierungen können dazu genutzt werden, um sich einen Überblick über die Daten und eventuelle Fehler zu verschaffen, die etwas über die Daten ›als Daten‹ aussagen. Bei der EDA wird häufig Clustering (z.B. *K-Means-Clustering*) eingesetzt, aber auch Topic Modeling, um z.B. etwaige Trends bei (Gesprächs-)Themen, Phänomenen, Genderaspekten oder auch Emotionen u.v.m. in bestimmten historischen Zeitabschnitten zu analysieren.¹¹⁷

Wichtig für transdisziplinäre Forschungsszenarien scheint uns dabei ein datensensibles und transparentes Vorgehen zu sein, bei dem es darauf ankommt, die Daten und ihre Modellierung stets im Blick zu behalten – woher sie kommen, wie sie zustande kommen, – sprich: generiert werden und wofür sie stehen.

3. Digitale Hermeneutik¹¹⁸ und Anforderungsanalyse bzw. -spezifikation (*Requirements Analysis*) in transdisziplinären Ansätzen:

Es wäre eine Art heuristisches Modell digitaler Hermeneutik zu entwickeln, das sowohl für die Geisteswissenschaften als auch für die Informatik operationalisierbar ist. Dabei geht es eben nicht, wie bereits erwähnt, darum, durch konstruktive Lektüren die Lücken (›gaps‹) und weißen Flächen der (historischen) Bedeutung zu überbrücken, sondern um die Frage, welche hermeneutischen Verfahren und methodologischen Aspekte auf welche Weise konkret im digitalen oder transdisziplinären Forschungsprozess genutzt werden können.

Für die Kompatibilität von hermeneutischen und quantitativ-digitalen Verfahren oder im Bereich von experimenteller Softwareentwicklung wird eine *iterative Vorgehensweise* vorgeschlagen, die transdisziplinär anschlussfähig ist, und zwar sowohl im Sinne eines *erweiterten hermeneutischen Zirkels* in den Geisteswissenschaften¹¹⁹, als auch in die Richtung von iterativen Verfahrensweisen¹²⁰ (wie z.B. dem *iterativ erweiterten Wasser-*

116 Vgl. insbesondere im Beitrag von Femmer und Lucke in diesem Band.

117 Vgl. dazu den Beitrag von L. Kraft im vorliegenden Band.

118 Vgl. den Forschungsschwerpunkt FSP digitale_kultur auf der *digitalen Hermeneutik* an der FernUniversität Hagen (<https://www.fernuni-hagen.de/forschung/schwerpunkte/digitale-kultur>) Vgl. zur digitalen Umsetzung hermeneutischer Verfahren auch die Projekte heureCLÉA und CATMA.

119 Vgl. Gius/Jacke: *The Hermeneutic Profit of Annotation*, S. 240. In einem hermeneutischen Prozess (oder Zirkel), so versteht Gadamer Heideggers Gedanken im Text *Vom Zirkel des Verstehens*, gehe es in der zirkulären Verstehensbewegung darum, theoriegeleitete Vorurteile, hartnäckige Theorien, Denkgewohnheiten oder populäre Urteile (er nennt diese ›Volksbegriffe‹) abzutragen und zu den ›Sachen selbst‹ zu kommen, nämlich anhand der ›Ausarbeitung aus den Sachen selbst her das wissenschaftliche Thema zu sichern‹.

120 Vgl. dazu auch Femmer H., Mendez-Fernández, D., Wagner, S., & Eder, S. (2017): *Rapid quality assurance with requirements smells*. *Journal of Systems and Software*, 123, S. 190–213, S. 190.

fallmodell oder *Spiralmodell* nach Boehm 1986) der Informatik. Das bedeutet konkret: Zunächst werden (vorläufige) literaturwissenschaftliche Anforderungen an technische Lösungen skizziert (Anforderungsanalyse). Nach erfolgter (literaturwissenschaftlicher oder transdisziplinärer) Validierung von Ergebnissen und Evaluation der Anforderungen werden die Anforderungen spezifiziert.¹²¹ Ein iteratives Vorgehen wie bei der Anforderungsspezifikation kann den literaturwissenschaftlichen Analysekriterien eher gerecht werden bzw. diese gegebenenfalls noch erweitern. Aus solchen Arbeitsprozessen können sich auch veränderte literaturwissenschaftliche Fragestellungen ergeben.

Eine weitere Integration von hermeneutischen Verfahren in den transdisziplinären Forschungsprozess ist die Interpretation und *hermeneutische Validierung*¹²² von Forschungsergebnissen auf der Basis von literaturwissenschaftlichem Fach- und Kontextwissen.

Annotationen bieten als Methode beispielsweise eine direkte digitale Anwendung von hermeneutischen Verfahren, indem der Text beispielsweise mit semantischen Zusatzinformationen angereichert wird.¹²³ Die Implementierung von hermeneutischen Verfahrensweisen integriert das Tool CATMA, indem es verschiedene Annotationskategorien anbietet wie »freie Annotation nach individuell definierten Kategorien«, »Mehrfachannotation einzelner Wörter und Passagen«, »überlappende Annotation« oder »widersprüchliche Annotation«.¹²⁴

Taxonomiebasierte Textarbeit, die ›top-down‹ und theorie- wie kategoriengeleitet verfährt, ist damit ebenso möglich wie die ›bottom up‹ verfahrenende und zirkuläre hermeneutische Forschung, die erst im Zuge der Exploration konkreter Texte ihre spezifischen Beschreibungsterme und -kategorien entwirft und präzisiert. CATMA kann darum für eine große Vielfalt an Forschungsansätzen genutzt werden.¹²⁵

4. Die *Reformulierung der eigenen kultur- und literaturwissenschaftlichen Fragestellung*: Der Einsatz von digitalen Methoden kann eine Revision und Reformulierung der eigenen komplexen kultur- und literaturwissenschaftlichen Fragestellung, etwa in untergeordnete Forschungsfragen, erfordern.

Die Fragestellung sollte in geeignete unterscheidbare Indikatoren für die computationale Textanalyse operationalisiert werden. Nach ersten Ergebnissen aus einer (explo-

121 Vgl. den transdisziplinären Beitrag von Femmer und Lucke im vorliegenden Band.

122 Das Konzept der Reliabilität spielt in den *humanities* häufig eine untergeordnete Rolle, da diese nicht in erster Linie den Anspruch haben, immer gleiche Ergebnisse zu erzeugen, auch nicht zu unterschiedlichen Zeitpunkten, da Kultur im Wandel begriffen wird.

123 Vgl. J. Horstmann, M. Seltmann (2023): Annotation. In: AG Digital Humanities Theorie des Verbandes Digital Humanities im deutschsprachigen Raum e.V. (Hg.): *Begriffe der Digital Humanities*. Ein diskursives Glossar (=Zeitschrift für digitale Geisteswissenschaften /Working Papers, 2). Wolfenbüttel 2023. 25.05.2023. HTML / XML / PDF. DOI: 10.17175/wp_2023_002

124 <https://fortext.net/tools/tools/catma> »Mit der nahtlosen Integration der Funktionsmodule ^[1] unterstützt CATMA die für die hermeneutische Textarbeit erforderliche iterative Arbeitsweise, indem beispielsweise erstellte Tagsets nach einer ersten Anwendungs-, Analyse- und Auswertungsphase problemlos mehrmals modifiziert und Annotationen überarbeitet werden können« (vgl. ebd.).

125 <https://fortext.net/tools/tools/catma>.

rativen oder hypothesengetriebenen) Datenanalyse kann die Fragestellung ggfs. reformuliert bzw. Anforderungen spezifiziert werden.

5. Mögliche Arbeitsschritte für ein Mixed-Methods-Forschungsprojekt:

- Phänomen entdecken bzw. erkennen im Close Reading Verfahren, bei der EDA (Explorative Datenanalyse) oder durch eine Frage- oder Problemstellung in der aktuellen Forschung
- Operationalisierung einer literatur- oder kulturwissenschaftlichen Fragestellung, Ableitung von Hypothesen und Formalisierung des gedanklichen Ansatzes nach den Prinzipien der formalen Logik (und ggfs. Reflexion von eigenen Erwartungen und Annahmen), Anforderungsanalyse
- Theoretische Reflexion des Mixed-Methods-Ansatzes, methodische und methodologische Vorüberlegungen, Triangulation¹²⁶
- Preprocessing (Download, Installation, technische Voraussetzungen, Einarbeitung in die technische und informatische Funktionsweise, Datenaufbereitung und -modellierung)
- Korpusbildung (Testkorpus erstellen, Auswahl der Daten, Texte gemeinfrei bzw. Open Source?)
- Phänomene, Forschungsfragen auf das Testkorpus anwenden (z.B. anhand von Most Frequent Words (MFW), Keywords in Context (KWIC), Stoppwortlisten, Tokenisierung, Lemmatisierung, Part of Speech-Tagging (POS-Tagging), Distanzmaße wie in der Stilometrie, Parametereinstellungen)
- Vorläufige Ergebnisse (Vergleichsdaten in der Forschung?)
- Manuelles Review, Fachexpertise, hermeneutische Validierung, Evaluation, Anforderungsspezifikation
- Ggfs. Korpuserweiterung (iteratives Vorgehen: wieder zurück zum Arbeitsschritt ›Korpusbildung‹)
- Topic Modeling, NLP-basierte Verfahren, Annotationen, Clusterverfahren, Fine-Tuning von LLMs etc.
- Visualisierung von Ergebnissen (Auffälligkeiten, Trends, Ausreißer, Bias-Effekte, Median oder Mittelwert?)
- Hermeneutische Validierung, Fachexpertise, Kontextwissen
- Grenzen und Möglichkeiten der angewendeten (kombinierten) Methoden (ggfs. aus transdisziplinärer Perspektive)
- Anforderungsspezifikation, Weiterentwicklung von Anwendungen, Tools und Methoden
- Weiterführende Fragestellungen (ggfs. Revision oder Splitten der ursprünglichen Fragestellung)

126 Vgl. Schröter: Mixed Methods.

Zielsetzung des vorliegenden Bandes

Im vorliegenden Band wird der Versuch unternommen, anhand von exemplarischen und experimentellen Anwendungen die inter- und transdisziplinären Potentiale und Mehrwerte der in den Dialog gebrachten Wissenschaften miteinander interagieren zu lassen, ohne dem Anspruch zu genügen, die methodologischen (und epistemologischen) Probleme (vollständig) zu lösen oder zu überbrücken, aber auch ohne sie auf disziplinäre Einseitigkeiten reduzieren zu wollen. Entlang der differenten Zugänge fachspezifischer und transdisziplinärer Perspektiven werden praktikable Vorgehensweisen, digitale Methoden, Datenmodellierungen, gemeinsame Schnittmengen und unterschiedliche Fragestellungen vorgestellt.¹²⁷

Es geht in den hier versammelten Beiträgen um die Frage, wie die (vermeintliche) Disparität der Methoden beider Fachdisziplinen und die Pluralität ihrer komplementären, aber auch überlappenden Verfahrenslogiken – auf der einen Seite frage- und theoriebasiert (Literaturwissenschaft), auf der anderen Seite zahlen- und datenbasiert (Informatik) – für den erkenntnisorientierten Forschungsprozess optimal genutzt werden können.

Dabei spielen die Aspekte Evaluation von Methoden und Tools eine wichtige Rolle. Die Software muss den Aufgaben in der Forschung gerecht werden, benutzerfreundlich und nachhaltig sein, d.h. u.a. einen nachhaltigen technischen Support beinhalten.¹²⁸ Der Forschungsprozess wird häufig direkt oder indirekt vom Software-Tool gesteuert.¹²⁹ Fehlerhafte und nicht passende Algorithmen, komplizierte Datenimports ohne Prüfung der Datenqualität, unvollständige und problematische Ausgaben, nicht transparente zugrunde liegende rechnerische Verfahren oder nicht sorgfältig ausgearbeitete Dokumentationen können den Forschungsprozess auf äußerst ungünstige Weise beeinflussen. Eine angemessene Kritik von Forschungen muss diese Aufbereitung und Auswertung durch Forschungssoftware einbeziehen, nachvollziehen und kritisch reflektieren können.

Die Beiträge in diesem Band

In den ersten beiden Beiträgen werden digitale und computationelle Methoden und Tools in Bezug auf literaturwissenschaftliche Fragestellungen exemplarisch angewendet, reflektiert und evaluiert: Bei der Anwendung und methodischen Reflexion des Topic

127 Das Projekt ist aus dem Kontext eines interdisziplinären Netzwerks aus Literaturwissenschaftler:innen und Informatiker:innen verschiedener Hochschulen in NRW hervorgegangen. Eine Kurzbeschreibung des Projekts findet sich unter Digital Humanities Projekte an der Ruhr-Universität Bochum im Rahmen des Netzwerks DH RUB (<https://dh-netzwerk.blogs.ruhr-uni-bochum.de/projekte/>, 7.11.2023).

128 Vgl. den Beitrag von H. Johannes im vorliegenden Band.

129 Vgl. Katerbow, M., Feulner, G. (2018): Handreichung zum Umgang mit Forschungssoftware. Zenodo, 27.2.2018. <https://doi.org/10.5281/zenodo.1172970> und Schmidt, S. C., Marwick, B. (2020): Tool-Driven Revolutions in Archaeological Science. In: Journal of Computer Applications in Archaeology 3: 18–32. <https://doi.org/10.5334/jcaa.29>.

Modeling bei Kafkas Texten (Laura Kraft) oder über den heuristischen Wert von NLP-basierten Verfahren in der computationellen Literaturtextanalyse (Henning Femmer/Alexa Lucke).

Laura Kraft (Literaturwissenschaft) stellt in ihrem Werkstattbericht »Topic Modeling am Beispiel von Franz Kafka: Eine Arbeitsskizze« die Methode *Topic Modeling* vor. Am Beispiel von Franz Kafka als Zeitungsleser während des Ersten Weltkriegs erörtert sie die Chancen und Grenzen der quantitativen Analysemethoden für die Bearbeitung literaturwissenschaftlicher Fragestellungen. Topic Modeling erweist sich dabei als vielversprechendes *Distant Reading*-Verfahren. Als Methode der quantitativen Bedeutungsanalyse soll Topic Modeling latente semantische Strukturen sichtbar machen können, die eine Einzeltextanalyse bisher nicht erreichen konnte. Anhand der Fokussierung auf die Aspekte *Preprocessing*, Korpus-Kenntnis und Code-Kompetenz erläutert sie die Voraussetzungen für einen erfolgreichen Einsatz von Topic Modeling, um auch eine klassische literaturwissenschaftliche Fragestellung im Rahmen einer hermeneutischen, intertextuellen und diskursanalytischen Betrachtung angemessen bearbeiten zu können. Dabei reflektiert sie auch die Rolle der Literaturwissenschaft und ihrer Methoden im Zeitalter der Digitalisierung. Der experimentelle Beitrag zeigt dabei, warum es für die Forschung interessant sein kann, wenn man mit Hilfe eines Algorithmus bei Franz Kafka zwischen den Zeilen lesen will.

Im gemeinsamen Beitrag »Über den Nutzen des *Natural Language Processing* in der computationellen Analyse literarischer Texte des Naturalismus und der (literarischen) Moderne« von Henning Femmer (Informatik) und Alexa Lucke (Literaturwissenschaft) wird der heuristische Nutzen des *Natural Language Processing* für die computationelle Analyse von literarischen Texten transdisziplinär, aus der Sicht von Literaturwissenschaft und Informatik, eruiert und diskutiert.¹³⁰ Entlang einiger exemplarischer literaturwissenschaftlicher Frage- und Problemstellungen werden ausgewählte literarische Texte, die im traditionellen Literaturkanon entweder dem Naturalismus oder der (literarischen) Moderne zugeordnet werden, anhand von sprachlichen Vorkommnissen (*Findings*), die für verschiedene literaturwissenschaftliche Kategorien stehen können, analysiert. Dazu wird ein NLP-basiertes Tool genutzt, dessen originäre Funktion die Überprüfung der Qualität von Anforderungen an ein Softwareprodukt ist (*Anforderungsanalyse/Requirements Analysis*). Das Tool ermittelt mit Hilfe von einstellbaren Filtermechanismen sogenannte Findings im Text, wie z.B. *Imprecise Phrases*, *Negative Words* oder *Vague Pronouns* und basiert auf typischen Technologien des NLP wie *Parsing*, *Lemmatizing* und *POS-Tagging*. Mit dem Tool soll empirisch überprüft werden, ob anhand der Häufigkeiten solcher sprachlichen Vorkommnisse in einem begrenzten Korpus von naturalistischen und modernen Texten tatsächlich eindeutige Epochenzugehörigkeiten abgeleitet werden können.

Die beiden daran anschließenden Beiträge nähern sich zum einen aus informatischer Perspektive, zum anderen aus literaturwissenschaftlicher Perspektive den KI-generierten Produktionen im Bereich von Kunst und Literatur. Unter einer ähnlichen Fra-

130 Als multivariate Verfahren aus dem Bereich des NLP (Natural Language Processing) können mehrere Einflussfaktoren bei der computationellen Textanalyse gleichzeitig berücksichtigt werden, was der Komplexität von literarischen Texten entgegenkommt.

gestellung kommen beide Beiträge aber zu verschiedenen plausiblen Schlüssen unter ihrer jeweiligen fachspezifischen Perspektive.

Im ersten Beitrag »Computergenerierter Zufall als kreatives Moment in Malerei und Literatur: Potentiale und Grenzen von *Machine-Learning-Modellen* am Beispiel von GPT« untersucht Andreas de Vries (Informatik) den Einfluss computergenerierten Zufalls auf kreative Prozesse. Als paradigmatisches Beispiel dafür wird eine Werktechnik Gerhard Richters betrachtet, die darin besteht, Farbfolgen quadratischer Raster durch Zufallsgeneratoren berechnen zu lassen. Des Weiteren werden die Prinzipien der gängigen Sprachmodelle der Computerlinguistik vorgestellt, die den Zufall in Form von Wahrscheinlichkeitsverteilungen zur Generierung von Texten verwenden. In diesem Zusammenhang bietet der Beitrag außerdem einen Einblick in die informatischen Grundlagen des *Natural Language Processing* und des *Machine Learnings*. Das öffentlich verfügbare Sprachmodell ChatGPT wird anhand konkreter Dialogbeispiele untersucht und daraus die Schlussfolgerung belegt, dass sein aktuelles Design einerseits das Erkennen oder Vermeiden logischer Inkonsistenzen grundsätzlich nicht ermöglicht, die Rolle des Zufalls allerdings zur Erzeugung neuartiger Texte oder nie dagewesener Aussagen führen kann. Allerdings wäre wie bei Richters Werktechnik der eigentlich kreative Prozess die Auswahl aus zufälligen Artefakten, nicht der Zufall.

Der Frage der Kreativität von KI widmet sich auch Denis Stevanovic (Literaturwissenschaft) in seinem Artikel »Die Kunst im Zeitalter der künstlichen Intelligenz: Walter Benjamins Kunsttheorie im 21. Jahrhundert«. Modelle künstlicher Intelligenz generieren Texte, Bilder, Musik und weitere Formen, die gemeinhin als Kunst aufgefasst werden können. Walter Benjamin hat einst die Frage nach dem Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit gestellt. Aufbauend auf Benjamins Ideen untersucht diese Studie die Spezifik von KI-generierten Kunstwerken und deren Position im Bereich der künstlerischen Produktion. Indem sie die einzigartigen Eigenschaften von KI-Kreationen berücksichtigt, wie ihre algorithmischen Ursprünge, maschinelles Lernen und das Zusammenspiel zwischen menschlicher Kreativität und computergetriebenen Systemen, zielen diese Überlegungen darauf ab zu bewerten, ob KI-generierte Werke als legitime Formen künstlerischen Ausdrucks betrachtet werden können.

Auch zwei weitere Beiträge widmen sich aus der Perspektive der Literaturwissenschaft weiter der Frage nach der Bedeutung von Maschinen und maschinellen Prozessen im Kontext von Literatur und Kunst.

Der Beitrag »*Algorithmizität*, Autorschaft und Paratext in generativer Kunst und Literatur« von Alexa Lucke (Literaturwissenschaft) fragt nach der Konstitution von Autorschaft und der Bedeutung von Paratexten bei KI-generierten Kunst- und Literaturproduktionen. Der ästhetische Fokus der dargestellten digitalen künstlerischen und literarischen Experimente verlagert sich auf die zugrundeliegenden Algorithmen und maschinellen Prozesse, nicht auf den damit generierten Output. Die begleitenden ästhetisch-technischen Kommentare explizieren, implizieren und mystifizieren dabei die algorithmischen Praktiken als Ausdruck von Reorganisationen von historischem Text-, Bild- oder Meta-Datenmaterial bzw. von De- und Rekontextualisierungen von historisch geprägten Wissensdiskursen. Paratexte und die Verhandlungen von Autorschaft zwischen Mensch und Maschine werden dabei häufig für ökonomische oder soziale Platzie-

rungspraktiken in den digitalen (Teil-)Öffentlichkeiten des Kunst- und Literaturmarkts genutzt.

Judith Schönhoffs Beitrag (Literaturwissenschaft/Komparatistik) »Verwenden Menschen und Automaten die gleiche Sprache? Literarische Fiktionen künstlichen Menschseins« widmet sich unter einer komparatistischen Perspektive der Frage nach »der« *differentia* von Mensch und Maschine bzw. Mensch und Automat in der westlich-europäischen Kulturgeschichte und Literatur seit der Antike. Schon die technisch fortgeschrittenen antiken Kulturen versuchten etwas zu erschaffen, dass das menschliche oder auch tierische Leben imitiert: die Automaten. Waren die realen mechanischen Geschöpfe zu dieser Zeit noch eindeutig als nicht lebendig zu erkennen, wurde dieses Manko in der Literatur jedoch ausgeglichen und die Maschinen, als handelnde und vor allem sprechende Wesen, jenseits ihrer technischen Realisierbarkeit beschrieben. Sprache war dann auch bis zum 19. Jahrhundert das entscheidende Merkmal, das die Menschenähnlichkeit der Maschine markiert. Im vorliegenden Aufsatz soll eine Übersicht über die Entwicklungsgeschichte des Automaten-Motivs gegeben werden, wobei die Frage nach der Differenz und Unterscheidbarkeit vom Menschen im Mittelpunkt steht.

Den Schluss bilden zwei Beiträge aus informatischer Sicht: Der Beitrag von Hermann Johannes (Informatik) gibt einen einführenden Überblick über die verschiedenen Bereiche der Digital Humanities, digitale Methoden sowie deren technische Umsetzungen aus Sicht der Informatik. Dabei werden die Methoden kurz vorgestellt und einige Tools in der digitalen Literaturwissenschaft gemäß ihres jeweiligen technischen Stands, Supports und Anwendungszwecks rezensiert.¹³¹

Jennifer Krieger (Informatik) gibt nach einer einführenden Definition der Informatik als Wissenschaft und einer kurzen Vorstellung ihrer Gebiete einen Überblick über die wichtigsten Konferenz- und Journalformate in der Informatik sowie einen Einblick in den standardisierten Aufbau von wissenschaftlichen Arbeiten.¹³²

Neben den fachspezifischen und transdisziplinären Perspektiven und der spannungsreichen Verschränkung von digitalen und klassischen Methoden, sich einander annähernden Gegenständen und teilweise unterschiedlich verwendeten Begriffen treffen hier zwei höchst differente Wissenskulturen aufeinander, die das Buchprojekt in einem wechselnden Neben- und produktiven Miteinander exemplarischer Studien dokumentieren möchte.

131 Im Anschluss an Homburg, T. u.a. (2020): Diskussionsbeitrag – Handreichung zur Rezension von Forschungssoftware in den Altertumswissenschaften/Impulse, URL: https://research-squirrel-engineers.github.io/Impuls_SoftwareRezensionen_DGUF/Draft.html [Zugriff: 09.08.2023]

132 Im Anhang des Beitrags von H. Johannes befindet sich noch ein Beispiel mit Pseudocode und Programm zur Primzahlerkennung, außerdem sind im Downloadbereich des Bands (https://github.com/LuckeJohannes/Literaturwissenschaft_Informatik) weitere Analysen (mit Stilometrie, K-means Clustering), ein Glossar sowie mögliche Inhalte eines DH-Studiengangs aus Sicht eines Informatikers.