

Manuel Burghardt
Lisa Dieckmann
Timo Steyer
Peer Trilcke
Niels-Oliver Walkowski
Joëlle Weis
Ulrike Wuttke (Hg.)

Fabrikation von Erkenntnis

Experimente in den Digital Humanities



Zeitschrift für digitale Geisteswissenschaften

Beitrag aus:

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. (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5) text/html Format. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Vorwort

Autor*in:

Manuel Burghardt

Kontakt: burghardt@informatik.uni-leipzig.de

Institution: Universität Leipzig

GND: [1237665523](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0003-1354-9089](https://orcid.org/0000-0003-1354-9089)

Autor*in:

Lisa Dieckmann

Kontakt: lisa.dieckmann@uni-koeln.de

Institution: Universität zu Köln

GND: [1077268289](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0002-1708-7371](https://orcid.org/0000-0002-1708-7371)

Autor*in:

Timo Steyer

Kontakt: t.steyer@tu-braunschweig.de

Institution: Universitätsbibliothek Braunschweig

GND: [1053806175](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0003-0218-2269](https://orcid.org/0000-0003-0218-2269)

Autor*in:

Peer Trilcke

Kontakt: trilcke@uni-potsdam.de

Institution: Universität Potsdam

GND: [139145117](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0002-1421-4320](https://orcid.org/0000-0002-1421-4320)

Autor*in:

Niels-Oliver Walkowski

Kontakt: niels-oliver.walkowski@uni.lu

Institution: Universität Luxemburg

GND: [1023378671](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0003-3043-3010](https://orcid.org/0000-0003-3043-3010)

Autor*in:

Joëlle Weis

Kontakt: weis@hab.de

Institution: Herzog August Bibliothek

GND: [1233399721](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0002-0080-4362](https://orcid.org/0000-0002-0080-4362)

Autor*in:

Ulrike Wuttke

Kontakt: wuttke@fhpotdam.de

Institution: Fachhochschule Potsdam

GND: [1107808405](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0002-8217-4025](https://orcid.org/0000-0002-8217-4025)

DOI des Artikels:

[10.17175/sb005_001](https://doi.org/10.17175/sb005_001)

Nachweis im OPAC der Herzog August Bibliothek:

[1764795245](#)

Erstveröffentlichung:

01.09.2021

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen: Medienrechte liegen bei den Autor*innen

Letzte Überprüfung aller Verweise: 11.08.2021

GND-Verschlagwortung:

[Digital Humanities](#) | [Experiment](#) |

Zitierweise:

Manuel Burghardt, Lisa Dieckmann, Timo Steyer, Peer Trilcke, Niels-Oliver Walkowski, Joëlle Weis, Ulrike Wuttke: Vorwort. 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. (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5) text/html Format. DOI: [10.17175/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_001](https://doi.org/10.17175/sb005_001).

Manuel Burghardt, Lisa Dieckmann, Timo Steyer, Peer
Trilcke, Niels-Oliver Walkowski, Joëlle Weis, Ulrike Wuttke

Fabrikation von Erkenntnis: Experimente in den Digital Humanities

Vorwort

Die Labor-Metapher ist in den Digital Humanities (DH) weit verbreitet¹ und wird aktuell verstärkt im Kontext der wissenschaftssoziologischen laboratory studies in der Tradition von Karin Knorr Cetina² auf vielfältige Weise diskutiert.³ Häufig betonen DH-Labore dabei den Aspekt des wissenschaftlichen Austauschs und der Kollaboration (im Sinne eines collaborative),⁴ der in den interdisziplinären Konstellationen der DH geradezu genre-prägend ist. Darüber hinaus ist das Labor traditionell eng verknüpft mit der Funktion als Experimentierstätte.

In DH-Labs werden dementsprechend geistes- und kulturwissenschaftliche Daten in Experimentalsettings aufbereitet, transformiert, kodiert und mithilfe von speziellen Instrumenten und Verfahren⁵ – also digitalen Tools und Methoden – analysiert. Die regelmäßige Reflexion des wechselseitigen Einflusses von digitalen Methoden und bestehenden Theorien sowie auch Epistemen eröffnet im DH-Lab zusätzlich die Möglichkeit erkenntnistheoretischer Versuche. Dabei fungieren DH-Labs auch als dynamische Werkstätten des Wissens. Grundlegende Praktiken der Geisteswissenschaften entfalten im Ideen- und Schreiblabor ihre kreative Kraft: Die Dokumentation wie die Kommunikation von Forschungsprozessen und -ergebnissen können dank neuer Publikationsformate an sich schon zum Experiment und zur interdisziplinären Ideenschmiede werden.

Unter dem Titel *Fabrikation von Erkenntnis: Experimente in den Digital Humanities* greifen wir in diesem Sonderband die eingangs genannten Aspekte auf und präsentieren Beiträge, die jeweils auf unterschiedliche Weise das experimentelle Potenzial der DH ergründen. Der Band nimmt mit dem Thema »Experimente« auch das Rahmenthema der vDHd2021 auf, das Alternativformat für die verschobene Jahrestagung des DHd-Verbands. Zwar ist der Band losgelöst von den Beiträgen der vDHd2021, dennoch greift er den Anspruch der DHd-Jahrestagungen auf, zur Sichtbarkeit aktueller DH-Aktivitäten im deutschsprachigen Raum beizutragen, wobei er selbst einen experimentellen Ansatz verfolgt. Der Band erscheint als Open-Access-Publikation und im Rahmen eines Co-Publishing Modells bei **Melusina Press** (Universität Luxemburg) und bei der **Zeitschrift für digitalen Geisteswissenschaften (ZfdG)** (Forschungsverbund Marbach Weimar Wolfenbüttel). Zudem setzt sich der Gesamtband aus drei, teils selbst noch experimentellen Beitragsformaten zusammen:

¹ Earhart 2015; Lane 2016.

² Knorr-Cetina 1991.

³ Vgl. Pawlicka-Deger 2020.

⁴ Vgl. Siemens / Siemens 2012.

⁵ Vgl. Hannaway 1986, S. 585.

- Fachartikel zu experimentellen Methoden und Verfahren der DH
- Daten-Experimente zur Publikation von Datensätzen
- Code-Experimente zur Publikation von ausführbaren Notebooks

Während die Fachartikel in beiden Publikationsorganen erscheinen, werden die Daten-Experimente bei der ZfdG und die Code-Experimente bei Melusina Press veröffentlicht. Dieses Modell des Co-Publishings ermöglicht überhaupt erst die experimentellen Formate und das Begutachtungsverfahren (s. u.), die diesem Band zugrunde liegen. Durch gemeinsame Publikationsrichtlinien, durch die eindeutige Referenzierung des Gesamtbandes mittels einer DOI:⁶ und durch den Nachweis aller Beiträge auf beiden Publikationsplattformen ist trotz des Co-Publishing-Konzepts stets die Struktur des Gesamtbandes erkenn- und auffindbar.

Der experimentelle Charakter des Bandes findet zudem im Begutachtungsverfahren seinen Ausdruck. Neben dem bekannten geschlossenen Peer Review-Verfahren für die Beiträge aus den Bereichen Daten- und Code-Experimente sowie einzelne Fachartikel, steht den Autor*innen der Fachartikel darüber hinaus mit einem Open-Peer-Review-Verfahren (OPR) eine derzeit im Open-Science-Kontext viel diskutierte Alternative zur Verfügung. Während im erstgenannten Verfahren die Begutachtung der Publikation vorgelagert ist (pre publication), findet im OPR-Verfahren die Begutachtung des Beitrags öffentlich anhand einer Preprint-Version statt (post publication). Die Community und zwei vom Herausgeber*innen-Team eingeladene Expert*innen können dabei wortgenaue Hinweise und Kommentare zur Preprint-Version abgeben. Auch werden die Expert*innen eine zusammenfassende Einschätzung des von ihnen zu begutachtenden Beitrages verfassen. Die Autor*innen können jederzeit auf die Kommentare und Monita eingehen, der Prozess wird vom Herausgeber*innen-Team moderiert. Ziel des Verfahrens ist eine konstruktive Kommunikation zwischen Autor*innen sowie Gutachtenden unter Einbeziehung der Community. Nach Abschluss der vierwöchigen Begutachtungsphase können die Autor*innen bei einem positiven Ergebnis eine überarbeitete Version ihres Beitrages verfassen, die dann Eingang in den Band findet. Die begutachtete Version bleibt unabhängig von der Annahme oder Ablehnung als Preprint inklusive der Gutachten öffentlich zugänglich. Falls sowohl Autor*innen wie auch die Expert*innen einverstanden sind, werden auch die Namen der Gutachtenden veröffentlicht.

Die Herausgeber*innen sehen in dieser Ausdifferenzierung von Begutachtungsverfahren großes Potential für einen fairen, konstruktiven und auf die individuellen Bedürfnisse der jeweiligen Autor*innen zugeschnittenen Review-Prozess.⁷

Den Autor*innen, den Gutachtenden sowie der DH-Community danken wir an dieser Stelle für die Offenheit für das neue, für viele noch unbekanntes Review-Verfahren. Den Leser*innen wünschen wir erkenntnisreiche Lektüren und viel Vergnügen beim Entdecken aktueller Experimentierstätten der Digital Humanities.

⁶DOI: [10.26298/vra6-mn32](https://doi.org/10.26298/vra6-mn32)

⁷Nach der Publikation aller Beiträge wird an dieser Stelle noch auf die Erfahrungen genauer eingegangen werden, die im Rahmen des Begutachtungsprozesses gesammelt worden sind.

Bibliographische Angaben

Amy E. Earhart: The digital humanities as a laboratory. In: *Between Humanities and the Digital*. Hg. von Patrik Svensson / David Theo Goldberg. Cambridge u. a. 2015, S. 391–400. [[Nachweis im GBV](#)]

Owen Hannaway: Laboratory Design and the Aim of Science: Andreas Libavius versus Tycho Brahe. In: *Isis – Journal of the History of Science Society* 77 (1986), H. 4, S. 585–610. [[online](#)] [[Nachweis im GBV](#)]

Karin Knorr-Cetina: *Die Fabrikation von Erkenntnis: Zur Anthropologie der Naturwissenschaft*. Revidierte und erweiterte Fassung, 1. Auflage. Frankfurt/Main 1991. (= Suhrkamp Taschenbuch Wissenschaft, 959) [[Nachweis im GBV](#)]

Richard J. Lane: *The Big Humanities: Digital Humanities/Digital Laboratories*. London u. a. 2016. [[Nachweis im GBV](#)]

Urszula Pawlicka-Deger: The Laboratory Turn: Exploring Discourses, Landscapes, and Models of Humanities Labs. In: *Digital Humanities Quarterly* 14 (2020), H. 3. [[online](#)]

Lynne Siemens / Raymond Siemens: Notes from the Collaboratory: An Informal Study of an Academic Lab in Transition. In: *Digital Humanities 2012: conference abstracts* (Digital Humanities Conference, Hamburg, 16.07.–22.07.2012) Hamburg 2012. [[online](#)] [[Nachweis im GBV](#)]

Zeitschrift für digitale Geisteswissenschaften

Beitrag aus:

Zeitschrift für digitale Geisteswissenschaften

Titel:

Qualitative Sprünge in der Qualitätssicherung? Potenziale digitaler Open-Peer-Review-Formate

Autor*in:

Yuliya Fadeeva

Kontakt:

yuliya.fadeeva@uni-due.de

Institution:

Universitätsbibliothek Duisburg-Essen

GND:

1047025450

ORCID:

0000-0002-3119-7892

DOI des Artikels:

[10.17175/sb005_002_v2](https://doi.org/10.17175/sb005_002_v2)

Nachweis im OPAC der Herzog August Bibliothek:

[1839571004](#)

Erstveröffentlichung:

01.09.2021

Version 2.0:

21.03.2023

Lizenz:

Sofern nicht anders angegeben



Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise:

24.08.2021

Format:

PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:

[Digital Humanities](#) | [Peer Review](#) | [Open Access](#) | [Qualitätssicherung](#) | [Wissenschaftskommunikation](#) |

Empfohlene Zitierweise:

Yuliya Fadeeva: Qualitative Sprünge in der Qualitätssicherung? Potenziale digitaler Open-Peer-Review-Formate. In: Zeitschrift für digitale Geisteswissenschaften (2023). HTML / XML / PDF. DOI: [10.17175/sb005_002_v2](https://doi.org/10.17175/sb005_002_v2).

Yuliya Fadeeva

Qualitative Sprünge in der Qualitätssicherung? Potenziale digitaler Open-Peer-Review-Formate

Abstracts

Open Access stößt innerhalb der Geisteswissenschaften nach wie vor auf gemischte Reaktionen. Insbesondere wird Open-Access-Publikationen oft eine mangelnde Qualitätssicherung vorgeworfen. Doch das Gegenteil ist der Fall, jedenfalls dann, wenn der Öffnungsgedanke konsequent auf Prozesse und Verfahren der Qualitätssicherung angewendet wird. Durch die Öffnung und Digitalisierung des Peer-Review-Prozesses werden signifikante Verbesserungen in der Qualität geisteswissenschaftlicher Arbeiten möglich. Gleichzeitig wird deutlich, wie sich der Begriff der wissenschaftlichen Arbeit durch die Ablösung von tradierten Publikationspraktiken verändert.

Open Access is still met with mixed reactions and hesitation from scholars in the humanities. In particular, a lack of quality assurance is often assumed. But quite the opposite is true, at least when the idea of openness is thoroughly applied to quality assurance processes and procedures. By opening up and digitising the peer review process, significant improvements in the quality of works in the humanities become possible. It also becomes clear how the concept of academic work is changing through the movements inside traditional publication practices.

1. Einleitung

Dieser Text¹ befasst sich konkret mit der wissenschaftlichen Qualitätssicherung geisteswissenschaftlicher Arbeiten, verfolgt aber zudem mehrere Ziele: Auf einer *übergeordneten* Ebene zielt er zum einen auf eine Verknüpfung neuer mit bekannten Themen und zum anderen auf eine Verbindung theoretischer Fragen mit praktischen Aufgaben. Zu den neuen Themen gehören die Reichweite und Auswirkung der Digitalisierung auf wissenschaftliche Praktiken, aber auch die Öffnung dieser Praktiken und dadurch stimulierte innerdisziplinäre Reflexionsprozesse. Besonders relevant sind in diesem Zusammenhang Themen wie die wissenschaftliche Arbeit bzw. Publikation oder das Wissenschaftsethos. Die praktischen Aufgaben lassen sich unterteilen in Modelle für Open Peer Review und Erfahrungen mit ihrer Umsetzung auf der einen Seite und spezifische Aufgaben der Open-Access-Publikationspraxis in den Geisteswissenschaften auf der anderen Seite. Open Access bezeichnet den kosten- und schrankenlosen Zugang zu und Nutzung wissenschaftlicher Erkenntnisse, vor allem wissenschaftlicher Literatur, aber auch anderer Forschungsergebnisse wie Forschungsdaten, Code, Korpora etc. Ist eine wissenschaftliche Publikation Open Access, dann ist sie »digital, online, free of charge, and free of most copyright and licensing restrictions.«² Es gibt jedoch mehrere Verwendungsweisen des Ausdrucks, darunter fallen auch teilweise konträre Positionen, die sich hauptsächlich hinsichtlich der Achsen kostenfrei (für Leser*innen / Autor*innen) sowie frei für die Verwendung (offen lizenziert) unterscheiden - diese Unterscheidung bezeichnet Suber als »gratis« und »libre«.³ Wichtig für Open Access im politischen Sinne ist die Kombination beider Achsen, also die kostenlose Zugänglichkeit *und* die freie Verwendungs-erlaubnis.

Theoretische Fragen betreffen den Begriff einer wissenschaftlichen Arbeit und seine Verbindung zum Begriff der Qualität sowie die wissenschaftssoziologische Reflexion der gerade stattfindenden Veränderungen der Wissenschaftskommunikation durch die Open-Access-Bewegung. Wir – Akteur*innen innerhalb der Digital Humanities – nehmen im aktuellen Publikations- und Wissenschaftsgeschehen eine doppelte Rolle ein, zum einen aus der Perspektive der wissenschaftlichen Analyse der Umbrüche sowie der möglichen weiteren Folgen für die Zukunft der Wissenschaft. Zum anderen sind wir selbst als Teilnehmende in vielen überlappenden Rollen in dieses Geschehen involviert. Wir gestalten es mit und können uns nur zum Teil davon distanzieren, um die Perspektive des »desinteressierten Beobachters«⁴ im Sinne Alfred Schütz' einzunehmen.

Auf *konkreter* Ebene geht es um die Diskussion des Open Peer Reviews, also der unterschiedlichen Weisen, die Praxis der Begutachtung zu öffnen, in Bezug auf die Qualitätssicherung wissenschaftlicher Arbeiten. Ich argumentiere dafür, dass durch Open Peer Review eine Qualitätssteigerung möglich ist. Diese Steigerung zeigt sich aktuell in der Erprobung unterschiedlicher digitaler Methoden. Die Verknüpfung dieser Themenkomplexe –

¹ Ich danke Dorothee Graf, Lisa Klaffki, Stefan Reining und Timo Steyer für hilfreiche Kommentare zu unterschiedlichen Vorversionen dieses Textes.

² Suber 2019.

³ Vgl. Suber 2019.

⁴ Schütz 1971, Teil I, Kap. IV, S. 41–43.

1. aktueller Aufgaben der Open-Access-Publikationspraxis in den Geisteswissenschaften,
2. begrifflicher und wissenschaftssoziologischer Themen, die durch den Transformationsprozess hin zu Open Access ans Licht treten und
3. der Methoden der Qualitätssicherung

– wird im inter- und transdisziplinären ›Labor‹ der Digital Humanities auf *genuin neuartige, qualitative Weise* gebündelt und adressiert. In diesem Kontext verstehe ich Digital Humanities in einem umfassenden, pluralistischen Sinne, wie ihn Fitzpatrick beschreibt: »The particular contribution of the digital humanities [...] lies in its exploration of the difference that the digital can make to the kinds of work that we do as well as to the ways that we communicate with one another.«⁵

Was heißt in diesem Zusammenhang ›genuin neuartige, qualitative Weise‹? Schließlich ist der gesamte disziplinäre Bereich der Digital Humanities vergleichsweise ›neu‹. Doch es geht nicht um die Feststellung, dass es sich um ein junges Phänomen handelt oder darum, dass die Digital Humanities wesentlich durch den Bezug auf ein neuartiges und sich rasant entwickelndes Kulturgut – das Digitale – gekennzeichnet werden sollen. Hier ist das Zusammenspiel folgender Eigenschaften gemeint:

- Digitale Emergenz: In den genannten Bereichen (1)–(3) ist eine digitale Komponente insofern konstitutiv, als dass die Fragestellung vor der Entwicklung digitaler Technologien nicht möglich war. Erst durch diese Technologien werden Open Access und bestimmte Weisen der Qualitätssicherungsverfahren möglich.
- Qualitative Sprünge: Digitale Verfahren führen zu qualitativen Verbesserungen der wissenschaftlichen Arbeit, d. h. es geht nicht um rein quantitative Steigerung von Effizienz, die Größen der Datensamples oder die Algorithmisierung von Verfahren.⁶
- Wissenschaftssoziologische und -theoretische Implikationen: Digitale Potenziale können unter anderem auch tiefgreifende Veränderungen der Wissenschaft als solcher bewirken, indem sie die Art und Weise, wie wissenschaftliche Praxen bisher funktioniert haben und dadurch begriffliche und normative Grenzen bestimmten, visualisieren und zum Untersuchungsgegenstand machen. Knorr-Cetinas Arbeiten,⁷ das *strong programme* von David Bloor und Barry Barnes⁸ oder die *Bath School* mit Harry Collins⁹ kehren das Narrativ eines naiven wissenschaftlichen Realismus um. Die soziologische Betrachtung der Wissenschaft fragt, inwiefern die Praxen wissenschaftlicher Forschung maßgeblich für Theorien, deren Terminologien und methodischen Forderungen sind. In anderen Worten: Inwiefern ist die konkrete wissenschaftliche Praxis nicht die – bestenfalls fehlerlose – Ausführung und Manifestation einer eindeutigen Axiomatik, strengen Methodik und Terminologie, sondern eine durch und durch von kontingenten, sozial-gesellschaftlichen, ökonomischen und kreativen Faktoren geprägte Manufaktur (im besten Sinne des Wortes)?

In gleicher Weise werfen digitales Publizieren und insbesondere der bisher nur formal,¹⁰ aber nicht inhaltlich definierte Begriff des Open Access¹¹ wichtige Fragen auf. Diese teilen sich in zwei Bereiche, mit Grundsatzfragen auf der einen Seite:

- Was ist eine (digitale) Publikation?
- Wie hängt der Begriff einer wissenschaftlichen Arbeit (als Produkt) mit dem Qualitätsbegriff zusammen?
- Welche Verbindung besteht zu den einzelnen Methoden der Qualitätssicherung?

Auf der anderen Seite stehen Fragen, die den zukunftsgerichteten, transformativen Charakter der digitalen Veränderungen betreffen:

- Wie lässt sich Qualität außerhalb der bisherigen, an Verlage und ihre jeweilige Reputation gebundenen, Publikationspraxis bestimmen?
- Was ändert Open Access an bisherigen Verständnissen und welche positiven wie negativen Folgen sind naheliegend oder bereits zu beobachten?

Auf die konkrete Thematik der Qualität (geistes-)wissenschaftlicher Texte richtet sich der Fokus des Textes. Speziell für Geisteswissenschaftler*innen ist die Qualitätssicherung eine Frage, die in Bezug auf eine Open-Access-Publikation eigener Arbeiten noch verunsichert. Hartnäckig hält sich das Vorurteil minderwertiger Qualität bzw. unzureichender Verfahren der

⁵ Fitzpatrick 2011. Zudem schließe ich – *for the sake of the argument* – unter die Rede von digitalen Verfahren und Technologien auch die ubiquitäre Verfügbarkeit des Internets auf institutioneller und individueller Basis ein.

⁶ Ein Beispiel für einen solchen Sprung findet sich in der gemeinsamen Entwicklung, Kommentierung, und im Austausch von Arbeiten auf digitalen Plattformen. Zur Zeit stechen im geisteswissenschaftlichen Bereich leider die kommerziellen Angebote von [academia.edu](https://www.academia.edu) und [researchgate.com](https://www.researchgate.com) hervor, vgl. ResearchGate 2022 und Academia 2022.

⁷ Knorr-Cetina 1981; Knorr-Cetina 1999.

⁸ Bloor 1976; Barnes / Bloor 1982.

⁹ Collins 1985.

¹⁰ Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities der Max-Planck-Gesellschaft 2003.

¹¹ Graf / Fadeeva 2020.

Qualitätssicherung.¹² Dieses Vorurteil entspringt zum einen der unglücklichen Assoziation von Open Access mit unseriösen und unklaren Praktiken – die Bandbreite reicht von online verfügbarer *grauer Literatur*¹³ diverser Art bis zum *predatory publishing*. Unter predatory publishing wird eine wissenschaftlich unredliche und oft betrügerische Form der Publikationspraxis verstanden. Dabei wird Wissenschaftler*innen die Publikation ihrer Arbeiten gegen die Zahlung einer (oft sehr hohen) Gebühr angeboten. Dabei findet keine oder sehr dürftige Qualitätsprüfung und -sicherung der Arbeiten statt.¹⁴

Zum anderen speist sich das Vorurteil aus dem (Misstrauen erweckenden) Neuheitscharakter der zahlreichen Open-Access-Publikationsformen und -modelle. Da Reputation im geisteswissenschaftlichen Bereich häufig an die Tradition eines Verlags bzw. einer Zeitschrift gebunden wird, haben es neue Modelle und Marken grundsätzlich schwer, sich zu etablieren.¹⁵ Der erste Aspekt des Vorurteils trägt so den Charakter einer negativen Konnotation, der zweite den einer konservativen Form der disziplinären Vertrauensbildung. Beide Aspekte betreffen weniger die tatsächlichen Eigenschaften der Publikationsform Open Access, bleiben aber durch die Komplexität und Unübersichtlichkeit der Thematik weiter im Raum stehen. Argumentative wie aufklärende Arbeit ist nötig, um diese Startschwierigkeiten für Open Access in den Geisteswissenschaften zu überwinden.

Am besten kann diese Argumentation durch möglichst transparente und detaillierte Explikation und natürlich stete Verbesserung der Qualitätssicherungsverfahren erfolgen. Die Open-Access-Publikationslandschaft ist komplex und divers, so dass weitere Eingrenzungen notwendig sind. Bereits eine oberflächliche Beschäftigung mit der Qualitätssicherung in Open Access zeigt zwei Dinge: Erstens, die Methoden der Qualitätssicherung, insbesondere der Peer Review, sind bei Open Access nicht grundsätzlich anders als im Closed Access.¹⁶ Zweitens, es gibt gute Gründe, Peer Review zu öffnen, um damit eine signifikante Qualitätsverbesserung zu befördern. Auf diese Punkte werde ich nun ausführlicher eingehen.

2. (Open) Peer Review

2.1 Begriffe, Methode und Öffnungsmöglichkeiten

Peer Review, die Begutachtung durch themenkundige Fachkolleg*innen, gilt als die wichtigste Methode der Qualitätssicherung, ist aber auch eine kontrovers diskutierte Methode¹⁷ mit einer langen Geschichte.¹⁸ Zunächst sollen die gebräuchlichen Ausdrücke differenziert werden: Im Deutschen gibt es die *Rezension*, eine (zumeist in einer Fachzeitschrift) von einer fachkundigen Person publizierte kritische Diskussion einer (in der Regel kürzlich erschienenen) längeren Arbeit: einer Monographie, eines Sammelbandes, Handbuchs etc. Die Rezension fasst die wichtigsten Thesen bzw. Beiträge der Arbeit zusammen und gibt eine umfassende Beurteilung ihrer argumentativen Stringenz, Überzeugungskraft, Relevanz, Aktualität etc. Durch die begründete Einordnung der Arbeit als Beitrag zum Diskurs empfiehlt eine Rezension sie als lesenswert oder rät davon ab, sie weiter zur Kenntnis zu nehmen. Das *Gutachten* legt weitere Kriterien an, um Arbeiten (aber auch Personen, Projekte, Anträge, Bewerbungen etc.) hinsichtlich eines externen Ziels (z. B. Publikation, Förderentscheidung, Stellenvergabe) zu bewerten. Das Gutachten beurteilt die Eignung der Arbeit für das Erreichen dieses externen Ziels. Der englische *Review* ist weniger eng bestimmt hinsichtlich seiner Funktion und des Stils. Er schließt die Rezension und das Gutachten ein sowie weitere, mehr oder weniger strukturierte Formen der begründeten Stellungnahme, z. B. als Sammlung einzelner Kommentare ohne allgemeine Bewertung des Gesamttexts. Ich verwende »Review«, wenn ich diese nicht eindeutig bestimmte Form meine, hinter der sowohl strukturierte Gutachten (*review report*) als auch einzelne Kommentare zu Texten im Sinne der *open participation* (auf die ich weiter unten eingehe) stehen können. Der passendere deutsche Ausdruck wäre vielleicht die (*wissenschaftliche*) *Stellungnahme*.

Die sorgfältige und begründete kritische Stellungnahme zu einer Arbeit innerhalb des eigenen Spezialgebiets ist eine wissenschaftliche Kernkompetenz. Dazu gehört auch die wissenschaftsethische Forderung,¹⁹ für die Objektivität und Belastbarkeit der Stellungnahme mit der eigenen wissenschaftlichen Reputation zu bürgen. Dahinter steht der Anspruch,

¹² Fitzpatrick / Santo 2012; Kleineberg / Kaden 2017.

¹³ Als graue Literatur gelten solche Texte, die nicht im traditionellen Sinne veröffentlicht wurden, also weder selbständig als Buch noch unselbständig als Beitrag in einem Sammelwerk im weitesten Sinne oder einem Periodikum erschienen sind. Berichte, Vorversionen später verändert veröffentlichter Arbeiten, Forschungsdaten, häufig auch Korrespondenz oder solche Texte, die aus unterschiedlichen Gründen nicht veröffentlicht wurden, werden als graue Literatur bezeichnet. Mit der Digitalisierung des wissenschaftlichen Publizierens wird diese Bezeichnung zunehmend unbefriedigend. Die Veröffentlichung *durch einen Verlag bzw. eine Zeitschrift*, das zwischen einer Publikation und grauer Literatur unterscheidet, wird zunehmend durch die digitale Möglichkeit der Veröffentlichung ergänzt oder ersetzt.

¹⁴ Siehe auch Science Media Center (Hg.) 2018 und Helmholtz 2022 (Hg.).

¹⁵ Auf diesen Punkt komme ich später zurück.

¹⁶ Z. B. Fadeeva 2021.

¹⁷ Batagelj et al. 2017; Ross-Hellauer 2017; Tennant / Ross-Hellauer 2020; Hirschauer 2004.

¹⁸ Biagioli 2002; Hirschi 2018.

¹⁹ Z. B. Merton 1973.

den wissenschaftlichen Anforderungen nach bestem Wissen und Gewissen²⁰ entsprochen zu haben. Das impliziert auch die Verpflichtung, eigene Urteile und Thesen bei Bedarf zu verteidigen und sie maximal transparent, nachvollziehbar und objektiv zu formulieren. Es kann gute Gründe geben, diese Verpflichtung im Einzelfall auszusetzen und Stellungnahmen zu anonymisieren oder geheim zu halten, z. B. um repressives oder diskriminierendes Verhalten innerhalb asymmetrischer Machtstrukturen zu verhindern. Aber das sind begründungspflichtige Ausnahmen zu einer wissenschaftlichen Grundhaltung der Transparenz und Nachvollziehbarkeit.

Peer Review bezeichnet als Terminus die Praxis der Begutachtung durch Fachkolleg*innen.²¹ Diese Stellungnahme unterstützt Institutionen bzw. Gremien, die über die Publikation, aber auch die Bewilligung von Drittmitteln oder auch Beförderungen und Stellenvergaben entscheiden, bei ihrer Auswahl. Üblicherweise findet der Review-Prozess geschlossen statt, ohne dass Namen von Personen, die entstandenen Reviews, die Erstversionen der im Review-Verfahren veränderten Arbeiten oder weitere interne Kommunikations- oder Entscheidungsprozesse publiziert werden.²² Am häufigsten sind Anonymisierungen der Reviewenden gegenüber den Autor*innen (*single blind*) oder die beidseitige Anonymisierung zwischen diesen Parteien (*double blind*), wobei in der Regel Herausgeber*innen die Zuweisung zwischen der Arbeit und qualifizierten Reviewenden übernehmen. Erfolgt auch diese Zuteilung anonymisiert, handelt es sich um ein Review-Verfahren in *triple blind*.

Werden einzelne Schritte dieses Prozesses bekannt gegeben, in Form einer Rückmeldung an Autor*innen oder durch Veröffentlichung, wird von *Open Peer Review* gesprochen, häufig jedoch ohne weitere Differenzierung. Ross-Hellauer²³ untersucht multidisziplinär verschiedene Definitionen des Open Peer Reviews in der Literatur. Seine sehr aufschlussreiche Arbeit enthält leider nur 13 (10,7 %) Quellen aus den Geistes- und Sozialwissenschaften und ist damit im Vergleich zur Betrachtung der STEM-Disziplinen (51,6 %) und des interdisziplinären Bereichs (37,7 %) weniger aussagekräftig.²⁴ Er findet unter 122 Definitionen sieben Merkmale, die als charakteristisch für den Aspekt der Öffnung gelten und jeweils Lösungen für Probleme der geschlossenen Review-Praxis bieten. Die sieben Merkmale sind, in absteigender Gewichtung:²⁵

- *open identities* – Beteiligte sind einander namentlich bekannt
- *open reports* – Reviews werden zusammen mit der Arbeit veröffentlicht
- *open participation* – die Fachcommunity kann sich am Review beteiligen
- *open pre-review manuscript* – die Erstversion einer Arbeit wird vor einem Review verfügbar gemacht, z. B. über einen Preprint Server
- *open interaction* – Beteiligte, sowohl Autor*in und Reviewende als auch Reviewende untereinander, können in Austausch treten
- *open final version commenting* – die publizierte finale Version kann kommentiert werden²⁶
- *open platforms* – Review-Services unabhängiger Anbieter

Ross-Hellauer schlussfolgert, dass Open Peer Review ein Sammelbegriff (*umbrella term*) ist. Das heißt, dass es keine eindeutige Kombination aus Merkmalen gibt, die für Open Peer Review notwendig ist; einige Merkmale scheinen allein hinreichend zu sein. Verschiedene Kombinationen aus den erkannten sieben Merkmalen sind möglich, wobei Ross-Hellauer aus der untersuchten Literatur 22 relevante Eigenschaftskombinationen (*unique combinations*) destilliert.²⁷ Besonders häufig (über 95 % der Definitionen in allen untersuchten Disziplinen) werden laut Ross-Hellauer *open identities* und *open reports* – einzeln oder in Kombination – als Kennzeichen von Open Peer Review verwendet.²⁸ *Open participation*, das interdisziplinär zweithäufigste Merkmal, »should perhaps be considered a core trait«²⁹ der Geistes- und Sozialwissenschaften. Ich werde mich im weiteren Verlauf auf die von Ross-Hellauer präzisierten Merkmale beziehen und von Open Peer Review im Sinne der allgemeinen Nutzung offener Formate sprechen.

2.2 Probleme im Peer Review und offene Lösungsformate

Zu zentralen Problemen aktueller Review-Praxis zählen insbesondere in den Geisteswissenschaften gerade solche Faktoren, die sich sowohl schwer eliminieren lassen, weil sie als funktionaler Teil des Review-Prozesses selbst gelten, als auch – paradoxerweise – zu seinen Stärken gezählt werden. Dazu gehören die Geschlossenheit der internen Abläufe, vor allem der Gutachten,

²⁰ Z. B. im Sinne der DFG-Leitlinien zur Sicherung guter wissenschaftlicher Praxis.

²¹ Als Peer gilt jemand mit in etwa vergleichbarer Kompetenz und Erfahrung auf dem fraglichen Gebiet.

²² In manchen Fällen werden die Reviews an Autor*innen als Hinweise zur Überarbeitung weitergegeben, Ablehnungen können durchaus kommentarlos erfolgen, was für Autor*innen frustrierend sein kann.

²³ Ross-Hellauer 2017.

²⁴ Ross-Hellauer 2017, Abb. 3. Der überwiegende Teil der begutachteten Arbeiten besteht aus Zeitschriftenartikeln: Ross-Hellauer 2017, Abb. 4.

²⁵ Ross-Hellauer 2017, S. 8–11.

²⁶ Diese Form hat Ähnlichkeit zur formalen Rezension, ohne auf Langschriften festgelegt zu werden.

²⁷ Ross-Hellauer 2017, Abb. 6

²⁸ Ross-Hellauer 2017, S. 15.

²⁹ Ross-Hellauer 2017, S. 15.

Entscheidungsprozesse und die (asymmetrische)³⁰ Anonymisierung. Dadurch sollen möglichst objektive und unabhängige Urteile erreicht werden, was jedoch aus verschiedenen Gründen gegenteilige Effekte hat, z. B. hinsichtlich impliziter wie expliziter Vorurteile, subjektiver Beurteilungen oder intransparenter, langwieriger Verfahren.³¹

Ross-Hellauer³² bündelt die Literatur zu Problemen der geschlossenen Review-Praxis in sechs Kategorien. Die ersten drei Kategorien betreffen problematisches Verhalten der Reviewenden: Unzuverlässigkeit und Inkonsistenz, mangelnde Verantwortlichkeit und Manipulationsgefahr im Prozess, unterschiedliche Formen von Voreingenommenheit (*social / publication bias*). Die restlichen Kategorien beziehen sich eher auf Probleme im Prozess: unnötig lange Wartezeiten und ›Schleifen‹ im Einreichungszyklus, weil Texte wiederholt gleiche Vorgänge bei verschiedenen Publikationsorganen durchlaufen; Fehlen entsprechender Anreize im akademischen Belohnungssystem, wodurch Wissenschaftler*innen die aufwändige und zeitintensive Review-Arbeit selten priorisieren können; Verschwendung interner Wissenschaftskommunikation, die vor allem für Nachwuchswissenschaftler*innen sehr aufschlussreich und wertvoll sein kann. Diesen Problemen stellt Ross-Hellauer die jeweils relevanten Arten der offenen Verfahren als Lösungsmöglichkeiten gegenüber,³³ vor allem *open identities*, *open report* und *open participation*. Seine Untersuchung zeigt eindeutig die Vorteile der offenen Review-Verfahren.

Die Wirksamkeit der Öffnung einzelner Verfahren misst Ross-Hellauer empirisch, u. a. anhand der Aufdeckungsquote von Fehlern durch Reviews und Angaben zur Review-Bereitschaft unter offenen Bedingungen. Für die Geisteswissenschaften sind jedoch andere Kriterien - argumentative Stringenz, Originalität der Gedankenentwicklung, explanative Stärke oder gründliche Auseinandersetzung mit Einwänden - relevanter, die zum einen nicht in der gleichen Weise empirisch mess- und nachweisbar sind wie im STEM-Bereich³⁴. Zum anderen geht es um normative Überlegungen, also wie eine Praxis sein bzw. geändert werden *sollte*. Es bedarf einer spezifisch geisteswissenschaftlichen Differenzierung der Funktionen der Review-Praxis bzw. der Anforderungen an sie. Eine wichtige Funktion ist das sogenannte *Gatekeeping* mit dem Anspruch, durch Selektion nur die besten Arbeiten zur Publikation zuzulassen. Genauso wichtig ist auch, dass Peer Review der tatsächlichen Verbesserung der Arbeiten dient, indem es konkretes Feedback zur Überzeugungskraft und Konsistenz der Argumentation gibt, alternative Blickwinkel bietet, Literatur empfiehlt oder auf explikationsbedürftige Passagen hinweist.

Gerade in dieser zweiten Funktion zeigt sich das Potenzial digitaler Technologien, um die bisher ›analoge‹ und geschlossene Praxis des Peer Reviews als kollaborative, offene Methode in qualitativer Hinsicht zu erweitern. Das geschieht zum einen durch die Entwicklung und Bereitstellung von Review-Plattformen bzw. weiterer Open Source Software³⁵, die eine breite Partizipation am Review-Prozess erlauben. Dadurch wird sowohl die Publikation von Reviews gleichzeitig mit dem Text selbst möglich als auch der Post-Publication-Review, der keine Gatekeeping-Funktion mehr hat. Zum anderen würde eine Veränderung innerhalb der anonymisierten, versteckten und nicht kreditierten Review-Arbeit dazu führen, sie genauso zum Teil der wissenschaftlichen Debatte zu machen wie den begutachteten Text selbst. So können Wissenschaftler*innen mehr Ressourcen in die weitere Verbesserung von Texten (der Reviews und der Arbeiten selbst) investieren. Wie kann das aussehen?

3. Qualität im Labor

3.1 Digitales Labor und die Methode der Öffnung

Digitale Peer-Review-Portale funktionieren in mehreren Hinsichten als ein Labor für die qualitative Verbesserung der bisherigen Praxis und ihrer Ergebnisse. Mir geht es weniger darum, für eine konkrete Plattform oder ein einziges Modell des Open Peer Reviews zu argumentieren. Ich schließe mich dem Urteil Ross-Hellauers³⁶ und dem von Fitzpatrick / Santo³⁷ an, dass es kein einzelnes Passpartout-Modell gibt bzw. geben kann, das allen Formaten und Ansprüchen an die Praxis optimal gerecht wird.³⁸ Für die Verbesserung der Ergebnisse des Review-Verfahrens, d. h. der begutachteten geisteswissenschaftlichen Arbeiten, gehe ich auf

³⁰ Obwohl die Formate *double* und *triple blind* gerade dieser Asymmetrie vorbeugen sollen, wird oft darauf hingewiesen, dass eine Anonymisierung der Autor*innen nicht immer möglich ist. Durch inhaltliche Ausrichtung bzw. die Zugehörigkeit zu einer Schule, durch persönliche Merkmale wie Schreibstil und Zitationsverhalten oder auch bei geringer Anzahl aktiver Wissenschaftler*innen in einem Spezialgebiet ist die Identifikation zumindest bestimmter Personen für *Peer Reviewer* häufig nicht schwierig.

³¹ Fitzpatrick 2010; Hirschauer 2015, 2004; Peters / Ceci 1982.

³² Ross-Hellauer 2017, S. 4f.

³³ Ross-Hellauer 2017, S. 13f.

³⁴ STEM steht für die englische Bezeichnung der Disziplinen sciences, technology, engineering und mathematics.

³⁵ Z. B. Hypothesis 2022

³⁶ Ross-Hellauer 2017.

³⁷ Fitzpatrick / Santo 2012, S. 3.

³⁸ Gleiches gilt für andere Bereiche des Open-Access-Publikationszyklus – Geschäfts- und Finanzierungsmodelle (Falkenstein-Feldhoff / Graf 2020; Speicher et al. 2018), Lizenzierungen (Müller 2017), Publikationsarten (Adema 2019), technische Umsetzung (AG Digitales Publizieren 2021, Kap. 2.2-3.1; Bülte 2020), Aufgaben- und Rollenverteilung (Graf / Fadeeva 2020) etc.

das von Ross-Hellauer als Kernmerkmal geistes- und sozialwissenschaftlicher Definitionen bezeichnete Merkmal *open participation* ein. Es impliziert die weiteren Merkmale *open pre-review manuscript* und *open report*,³⁹ scheint aber mit der Anonymisierung der Reviewenden kompatibel. Inwiefern kann bei *open participation* von Laborbedingungen die Rede sein, insbesondere in Anbetracht der nicht-empirischen Natur meiner Argumentation?

Eine Lesart von Labor besteht im Experimentieren mit der Vielfalt digital-technischer Optionen, z. B. mit dem Working Paper der AG Digitales Publizieren,⁴⁰ das über Hypothes.is kommentiert werden konnte, dem offenen Post-Publication-Review-Format für den Band *Fabrikation von Erkenntnis*,⁴¹ zu dem auch dieser Beitrag gehört. Zum Verständnis als Labor gehören auch die emergenten Erkenntnisse, die durch den Einsatz und die Kombination computergestützter Technologien entstehen (Versionierung, Hypertext, Durchsuchbarkeit, Transparenz, simultaner Zugang / Bearbeitung / Interaktion einer sehr großen Community, Überprüfbarkeit und nicht zuletzt freier Zugang zu Quellen).

Eine weitere Lesart hängt mit der Tradition wissenschaftlichen Arbeitens in den Geisteswissenschaften zusammen, zu der wesentlich die Entwicklung und Veränderung von Ideen gehört. Im STEM-Bereich gilt häufig, Ergebnisse zuerst zu erzielen und durch die Publikation zu sichern, d. h. an den eigenen Namen und die eigene Institution zu binden; eine Entdeckung bringt der ersten Forschungsgruppe das meiste Renommee und weitere Verwertungsmöglichkeiten wie Patente und Förderungen. Diese Sorge in Bezug auf die Schnelligkeit der Veröffentlichung gibt es im geisteswissenschaftlichen Publizieren nicht, auch weil die Konstellation viel seltener ist, dass mehrere Wissenschaftler*innen in Bezug auf die Forschung am selben Gegenstand (um erwartbar gleiche Ergebnisse) konkurrieren.

Zeit spielt in einer anderen Hinsicht eine wichtige Rolle für die Qualitätssteigerung eines geisteswissenschaftlichen Textes. Diese Hinsicht kann metaphorisch als ›Reifung‹ bezeichnet werden. Geisteswissenschaftliche Texte haben den Anspruch, relevante Beiträge zu komplexen, synchronen und diachronen Diskursen zu sein, bei denen es um innovative und kreative Blickwinkel auf Probleme geht, um die Entwicklung neuer Argumente, konsistenter Narrative und Interpretationen.⁴² Diese Ideen, Argumente und Interpretationen werden von verschiedenen Seiten betrachtet, durchdacht, mit anderen diskutiert, gegen mögliche Einwände abgewogen. Sie ›gären‹ eine ganze Weile im diskursiven Raum, werden verschriftlicht, überarbeitet, besprochen, revidiert, geschärft etc.⁴³ Dieser Reifungsprozess erfolgt selten hermetisch abgeschlossen ›in‹ einem Einzelindividuum⁴⁴ – weder ist das ein adäquates Verständnis des Bewusstseins noch der Ideengeschichte. Der Austausch mit anderen – im Forschungsbereich meist Kolleg*innen –, die Präsentation als Vortrag in Kolloquien, Lesekreisen oder Seminaren, die Rückmeldung zu Entwürfen, die möglichst vielen hilfsbereiten Peers gegeben werden, prägen die Entwicklungsstufen eines Textes.⁴⁵ Im bisherigen Publikationsprozess kommt ein Text erst zu einem viel späteren Zeitpunkt zur Veröffentlichung, in einer häufig gegenüber dem Erstentwurf stark veränderten Form.⁴⁶

Open Peer Review stellt die technische und methodische Fortsetzung und Erweiterung dieser Praxen dar, leistet aber auch durch die Nutzung der Interaktionsmöglichkeiten einen Beitrag zur Erweiterung des Autorschaftskonzepts⁴⁷ und der kollaborativen Wissensproduktion. Die Involvierung einer breiten, hinsichtlich ihrer Zusammensetzung offenen Community hat noch einen weiteren Vorteil. Dieser Vorteil besteht in der Mischung heterogener Perspektiven innerhalb der Teilnehmenden am Review-Prozess. Das können durchaus auch Perspektiven sein, die außerhalb des eigenen Denkkollektivs⁴⁸ liegen bzw. jenseits der bereits bekannten Alternativen. Das heißt natürlich nicht, dass jeder Kommentar anregend und sinnvoll sein muss. Aber gerade solche Fragen, Kritiken und Kommentare, die aus vorher nicht erwogenen oder präsenten Perspektiven kommen, können – ganz im Sinne der Serendipität – hilfreich dabei sein, neue Verbindungen zu sehen, in andere Richtungen zu denken oder zumindest eine klarere Darstellung der eigenen Position zu formulieren.⁴⁹

³⁹ Denkbar wäre auch, dass Kommentare und komplette Reviews nicht publiziert, sondern nur ausgewählten Personen verfügbar gemacht werden.

⁴⁰ AG Digitales Publizieren 2021.

⁴¹ Burghardt et al. 2021

⁴² Vgl. Fitzpatrick 2012, S. 7 für die Analyse der Bewertungspraktiken der Disziplinen.

⁴³ Vgl. Beißwengers Darstellung der Geisteswissenschaft(en) als Diskurs über Texte, in Fadeeva 2020, Kap. 2.1.1.

⁴⁴ Selbst wenn ein solch kreatives Genie als Idealbild der Aufklärung seine Gedanken mit niemandem teilt und in Isolation verschriftlicht, sind sie Beiträge zu bestehenden Diskursen, sie schließen an existierende Werke an, stoßen sich von diesen ab, erweitern, korrigieren oder kritisieren sie.

⁴⁵ »Open review is not a radically new scholarly practice driven by technological innovation; rather, it has developed out of a long history of humanities-based scholarly endeavors, taking forms including the presentation of conference papers, the formation of working groups, and other forms of process-oriented scholar-to-scholar communication.« Vgl. Fitzpatrick / Santo 2012, S. 3.

⁴⁶ Danksagungen für Diskussionen und thematische Unterstützung sind ein fester Bestandteil von Langschriften. Bemerkenswert ist, wie viele debattenprägende Artikel in Fachzeitschriften, etwa in der analytischen Philosophie, Danksagungen an eine lange Liste von renommierten Kolleg*innen, Studierenden, ›anonymen Referees‹ oder Teilnehmenden an Diskussionsrunden enthalten.

⁴⁷ Ernst 2015, Ernst 2018; AG Digitales Publizieren 2021, Kap. 2.1.

⁴⁸ Fleck 1980.

⁴⁹ *Open participation* steht der Forderung Feyerabends nach methodologischer Diversität und Experimentierfreude nahe, vgl. Feyerabend 1993.

Digitale Plattformen bieten die Möglichkeit, asynchron umfangreichen, detaillierten und nachverfolgbaren Austausch mehrerer Personen zu punktgenau identifizierbaren Textpassagen zu initiieren; zwischen beinahe beliebig vielen Interessierten. Zudem bietet die digitale Aufarbeitung verschiedener Ressourcen an einem Ort (Einbindung von Quellen, multimediale Formate, Nebeneinander von Text und Kommentar, schnelle Orientierung durch Durchsuchbarkeit etc.) einen wichtigen pragmatischen Überblick, der mit gedruckten Exemplaren in dieser Weise nicht möglich ist.⁵⁰

Open participation auf digitalen Plattformen erweitert die lange Tradition gemeinsamer, kritisch-kontroverser Textdiskussion um die Ablösung von der Kopräsenz. Die involvierten Personen müssen nicht mehr zur gleichen Zeit am gleichen Ort sein, es gibt keine durch das Format der Veranstaltung (von Lesekreis bis Vorlesung im Stil von Michael Sandell) bestimmte Mindest- oder Maximalzahl. Damit werden limitierende Faktoren wie Termin- und Ortsbindung, die persönliche kognitive und motivationale Tagesform sowie der einmalige Ereignischarakter der Diskussion erheblich gemildert. Interessierte können die Dauer des Reviewprozesses auf der entsprechenden Plattform für die Entwicklung ihrer eigenen Kommentare und Kritikpunkte nutzen.⁵¹ Für Autor*innen bedeutet das eine effektivere Zeitznutzung, wenn ihr Text gleichzeitig von allen Interessierten gelesen und kommentiert⁵² werden kann. Die kommerzielle Plattform academia.edu⁵³ ist ein Beispiel dafür, aktuelle Arbeiten zugleich von einer großen Menge an Kommentator*innen dialogisch mit dem*der Autor*in und anderen Interessierten zu diskutieren.

3.2 Spekulative Anmerkungen

Die Öffnung einzelner Bereiche der Wissenschaftskommunikation, insbesondere Open Access, steht im Kontext komplex ineinandergreifender Faktoren, u. a. globaler ökonomischer und wissenschaftspolitischer Entwicklungen, multinational agierender Großverlage und der fortschreitenden Digitalisierung. Mittel- und langfristige Folgen der Interaktions- und Rückkopplungsfunktionen des »Web 2.0«⁵⁴ und der diversen Technologien zur Bearbeitung geisteswissenschaftlicher Themen, wie Kirschenbaum⁵⁵ die Digital Humanities beschreibt, sind noch nicht absehbar. Zu künftigen Entwicklungen im Publikationsprozess möchte ich abschließend einige spekulative Bemerkungen machen. Der Review-Prozess könnte durch die Öffnung für breite Partizipation seine Funktion verändern und weniger auf Gatekeeping konzentriert sein, sondern auf die gemeinsame Verbesserung und Entwicklung von Arbeiten. Diese Transformation betrifft auch einen wissenschaftlichen Kernbegriff, die wissenschaftliche Arbeit (als Produkt), der bisher an die Publikation in einem bestimmten Format und damit verbundene implizite wie explizite Voraussetzungen gebunden war.

Ein Beispiel für eine solche Voraussetzung ist die Materialität des Drucks.⁵⁶ Mit dem Wegfall der Notwendigkeit, eine Arbeit zu einem bestimmten Zeitpunkt für »fertig« zu erklären, um sie als gedruckte Publikation verbreiten zu können, werden dauerhaft bearbeitbare Formate möglich, wie sie schon in Form von Wikis, Blogs und Living-Books bzw. Reviews existieren. Sie bilden nicht nur »unfertige«, aktualisierbare, dynamische Wissensquellen, die durch Versionierung so etwas wie die Dokumentation der Ideengeschichte darstellen. Sie explizieren zudem interne Funktionsweisen der Wissenschaft und zeigen, durch welche Faktoren – Argumente, Kommunikationsprozesse, soziale und wissenschaftspolitische Einflüsse, wissenschaftliche Moden und Paradigmen – sich Diskussionen entwickeln. Diese Zusammenhänge sind nicht nur durch die Offenbarung einer irreduziblen Kontingenz im Wissenschaftsgeschehen spannend, einem zentralen Thema der Wissenschaftssoziologie.

Die zahlreichen digitalen Möglichkeiten – insbesondere Versionierung, Langzeitverfügbarkeit, Auffindbarkeit – machen es möglich, auch jene Prozesse systematisch zu beobachten, in deren Entstehen und Wirken *wir uns selbst befinden*. Wissenschaftshistorische und -soziologische Arbeiten sind aus einer gewissen zeitlichen Distanz zum Untersuchungsgegenstand entstanden;⁵⁷ ihr Gegenstand wurde oft erst retrospektiv als solcher konstituiert. Wir haben diesen Abstand des Blicks »nach hinten« nicht, wir befinden uns mitten im Geschehen und beeinflussen es in verschiedene – absichtlich-erwünschte oder unbewusst-unerwünschte – Richtungen. Wir haben jeweils eigene Interessen in der wissenschaftspolitisch geförderten Transformation hin zu Open Science, auf die es auch kritische Perspektiven gibt.⁵⁸ Die Möglichkeit, den gesamten Prozess nachverfolgbar und für die spätere

⁵⁰ Was natürlich nicht heißt, dass Print dadurch ersetzt oder obsolet wird.

⁵¹ Als Beispiele für Open-Peer-Review-Ansätze nennt das Working Paper der AG Digitales Publizieren (AG Digitales Publizieren 2021, Kap. 4.2.5) die *Zeitschrift für digitale Geisteswissenschaften (ZfdG)*, *Language Science Press*, *Kunstgeschichte*, *Informationspraxis* und *ScienceOpen*.

⁵² Ein großer Vorteil digitaler Plattformen wie Hypothes.is und Annotationstools wie textlabor (Beißwenger 2020) besteht darin, Kommentare und Markierungen vom ursprünglichen Manuskript separierbar zu halten. Reviewende können in ihrem jeweiligen digitalen Profil das Manuskript auf eigene idiosynkratische Weisen bearbeiten. Diese Bearbeitungen sind aber nicht nur als Bearbeitungen einzelnen Nutzer*innen zugeordnet und dadurch selbst adressierbar, sondern auch ausblendbar. So bleibt das Original jederzeit erkennbar für eine frische Betrachtung.

⁵³ Vgl. die Kritik von Fitzpatrick (2015) und Hall (2015).

⁵⁴ Münker 2009.

⁵⁵ Kirschenbaum 2010.

⁵⁶ Vgl. Sahle 2010.

⁵⁷ Shapin 1975; Bloor 1976; Latour / Woolgar 1986; Latour 1988; Collins / Pinch 1993.

⁵⁸ Z. B. Knöchelmann 2019.

Untersuchung und Analyse nachhaltig verfügbar zu halten, bietet eine wertvolle Ergänzung zur Position der unmittelbar am Geschehen und seinen Kommunikationsprozessen Beteiligten. Digitale Technologien machen damit die Dokumentation der kommunikativen und extrakommunikativen Perspektiven⁵⁹ und ihrer Überlappungen möglich.

Neben der Metaperspektive sind noch zwei Dinge interessant. Digitale Technologien versetzen uns in die einzigartige Situation, die Daten über alle Schritte und Regungen dieser Transformation in bisher unerreichter Genauigkeit erheben und bei Bedarf direkt verfügbar machen zu können. Das befördert die wissenschaftliche Kollaboration für eine umfassende, gründliche Analyse eher als die Konkurrenz um schnelle Ergebnisse.

Der zweite Punkt betrifft den disziplinären Gegenstand. Science and Technology Studies, Laboratory Studies, die klassischen wissenschaftshistorischen Untersuchungen stehen schon seit geraumer Zeit im wissenschaftssoziologischen Fokus. Die Digital Humanities machen diese Art akribischer Untersuchung für die Geisteswissenschaften möglich. Die Pointe liegt also weniger darin, für die Existenz unterschiedlichster soziologischer Faktoren in der Wissenschaft zu argumentieren, sondern im neuartigen, digital erweiterten Einblick in den geisteswissenschaftlichen Arbeitsprozess. Die bereits heute technisch möglichen Formen von Open Peer Review lassen diesen Einblick und die methodologischen, kommunikations- und wissenschaftstheoretischen Implikationen mit Spannung erwarten.

4. Fazit

Für Autor*innen bedeutet die Öffnung des Review-Prozesses, insbesondere *open participation* und *open reports*, eine sprunghafte Vergrößerung der *erreichbaren*, relevanten Zielgruppe für die Diskussion ihrer Texte. Weil dieser kommunikative Austausch für die Entfaltung geisteswissenschaftlicher Themen besonders wichtig und fruchtbar ist, steigt durch die Öffnungsprozesse auch das Potenzial für die qualitative Verbesserung der Arbeiten. Reviewende können sich miteinander austauschen, was im geschlossenen Verfahren unüblich bzw. nur in Einzelfällen möglich war, und ihre Zeit und Aufmerksamkeit mit entsprechender beruflicher Anerkennung dafür einsetzen. Die aufwendige und zeitintensive Review-Arbeit kann dadurch gewinnen, dass sie transparenter, schneller und vielstimmig durchgeführt werden kann. An dieser Stelle zeigt sich eine weitere signifikante Eigenschaft des digitalen Publizierens, auf die Fitzpatrick⁶⁰ hinweist: Die Ressourcenknappheit besteht beim digitalen Publizieren nicht mehr im begrenzten Platz zum Drucken eines Textes. Die relevanten Ressourcen sind Zeit und Aufmerksamkeit. Es sind die Ressourcen der Reviewenden und der Lesenden, die durch ihre positive wie negative Leseentscheidung die weitere Wissenszirkulation (oder ihr Ausbleiben) steuern. Es sind aber auch die Ressourcen der Personen im Publikationsprozess, die häufig unbezahlt⁶¹ herausgeberische Aufgaben, Redaktion, Layout, Satz und den umfassenden digitalen Bereich übernehmen.

In Bezug auf die Bereitschaft von Wissenschaftler*innen, sich im Open Peer Review zu engagieren, ist die Motivation ein entscheidender Faktor. Um diese zu steigern, ist die Anerkennung der Review-Arbeit als karriererelevante Leistung zentral. In diesem Lichte sollte die (erwartungsgemäß) gesunkene Review-Bereitschaft bei *open identities*, die Ross-Hellauer konstatiert, gewertet werden: Nur diesen Faktor abzufragen, während die anderen Bereiche des Prozesses geschlossen bleiben, ist nicht zielführend. *Ceteris paribus* ist die Aufhebung der Anonymisierung klarerweise nicht attraktiv, da Review-Arbeit ihre Nachteile behalten würde, zuzüglich der Befürchtungen negativer Reaktionen seitens der (einflussreichen) Autor*innen. Diese Befürchtung ist durchaus begründet und nicht von der Hand zu weisen. Insbesondere im Hinblick auf den Schutz von Nachwuchswissenschaftler*innen ist dies ein Problem, dem schwierig vorzubeugen ist. Zum einen begünstigen die Förder- und Aufstiegsstrukturen innerhalb der Wissenschaft asymmetrische Abhängigkeitsverhältnisse, die vor allem für den wissenschaftlichen Nachwuchs nachteilig sein können. Entsprechende Handlungen können sehr subtil sein und sind innerhalb der Netzwerkstrukturen schwierig nachzuhalten. Zum anderen ist die Vergrämung über eine Kritik der eigenen Arbeit eine naheliegende Reaktion, die mehr oder weniger feindselig, offen oder bewusst ausfallen kann, trotz aller Professionalität.

Der Umgang mit *open identities* ist aber auch eine Frage der Praxis einzelner Disziplinen. So können Befürchtungen über negative Folgen der namentlichen Bekanntgabe von Gutachter*innen, die Ross-Hellauer für die Fächer in seiner Untersuchung nennt, in anderen Disziplinen weitaus weniger relevant sein.

Für diese Problematik gibt es keine einfache Lösung. Neben strukturellen Änderungen und Antidiskriminierungsmaßnahmen ist es sinnvoll, sowohl auf größere Transparenz zu setzen, um den Spielraum für direkte Vergeltungshandlungen zu verkleinern, als auch die Modularität der Merkmale im Blick zu behalten. So könnten die relevanten Merkmale für die diskursive Verbesserung der Arbeit – *open report* und *open participation* – auf Wunsch auch ohne *open identities* praktiziert werden.

⁵⁹ Loenhoff / Schmitz 2012.

⁶⁰ Fitzpatrick 2011.

⁶¹ Deville et al. 2019.

Eine motivierende Änderung für die Review-Arbeit mit *open identities* entsteht als Gegengewicht erst durch weitere Öffnungen. Am wichtigsten ist dabei *open report*, weil Reviews als eigenständige Arbeiten (ähnlich der Rezension) publiziert und so im eigenen Profil aufgenommen werden können.

Zum Schluss komme ich auf das Vorurteil gegenüber der Qualitätssicherung im geisteswissenschaftlichen Open-Access-Publizieren zurück. Was können offene Formate einem auf *Alter* und *tradierten Verlagsnamen* basierendem Reputationssystem, zumindest in Bezug auf die klassischen ›analogen‹ Geisteswissenschaften, entgegenstellen? Häufig handelt es sich bei solchen Formaten um wissenschaftlich organisiertes Publizieren, um Universitätsverlage oder Initiativen, denen sowohl Alter als auch Bekanntheit fehlen. Haben die großen Verlagshäuser, die das Feld bereits dominieren, hier einen Vorteil? Sie könnten frühzeitig in Open Access investieren und bieten nun Open-Access-Versionen ihrer Publikationen gegen (teilweise sehr hohe) Gebühren an, die von Autor*innen bzw. ihren Institutionen getragen werden müssen. Tatsächlich haben neue, offene Publikationsformate zunächst einen gewissen Nachteil gegenüber renommierten Verlagen. Diesem Nachteil stehen auf positiver Seite die Diversität der Publikationslandschaft und die vielen Entwicklungen hin zur »bibliodiversity« im Publizieren gegenüber, die Pierre Mounier⁶² auf den Open-Access-Tagen 2020 forderte. Die Steigerung der Transparenz im gesamten Prozess, die durch die Öffnung verschiedener Verfahren im Review-Prozess möglich wird, bildet eine nachvollziehbare und überzeugende Alternative für die Reputationsbildung.⁶³ Qualität wird durch strenge Anforderungen und transparente Arbeitsschritte gesichert, nicht durch Namen oder Alter von Verlagen und Zeitschriften. Speziell *open participation* und *open report* exemplifizieren zentrale geisteswissenschaftliche Ideale der Qualität.

Folgt daraus, dass offene Verfahren in jedem Fall und in Bezug auf jedes Merkmal geschlossenen Verfahren vorzuziehen sind? Nein. Insbesondere für Reviewende kann die Option sinnvoll sein, im Einzelfall anonym zu bleiben. Die sieben von Ross-Hellauer bestimmten Merkmale unterscheiden sich in der Gewichtung und innerdisziplinärer Relevanz. Das spezifisch geisteswissenschaftliche Qualitätssteigerungspotenzial betrifft ebenfalls nicht alle Merkmale. Es gibt einen anderen Grund, Begutachtungspraktiken als grundsätzlich offen anzulegen, um von dieser Ausgangssituation begründete Ausnahmen zu machen, also das Verhältnis zwischen offenen und geschlossenen Verfahren umzudrehen. Das ist der Bezug von Open Peer Review zu Open Science als einem praktisch umzusetzenden Ideal. Die konsequente Öffnung der Wissenschaft – ihrer Literatur, Methoden, Daten, Ergebnisse, Prozesse und Software – beinhaltet als Konsequenz die Öffnung von Verfahren der diskursiven, kollaborativen und kontingenten Wissensentstehung und Weiterentwicklung im digitalen Publizieren.

⁶² Mounier 2020; Mounier 2018.

⁶³ Als positives Beispiel dient das Verfahren der [Open Book Publishers](#).

Bibliographische Angaben

- Academia. Hg. von Academia.edu. In: academia.edu. San Francisco, CA 2022. [\[online\]](#)
- Janneke Adema: Executive summary: Towards a Roadmap for Open Access Monographs. A Knowledge Exchange Report. 2019. In: zenodo.org. Version 1.0 vom 31.05.2019. DOI: [10.5281/ZENODO.2645038](#)
- Digitales Publizieren in den Geisteswissenschaften: Begriffe, Standards, Empfehlungen. Hg. von der AG Digitales Publizieren. Wolfenbüttel 2021. Version 1.0 vom 18.03.2021. (= Zeitschrift für digitale Geisteswissenschaften / Working Papers, 1) DOI: [10.17175/wp_2021_001](#)
- Barry Barnes / David Bloor: Relativism, Rationalism and the Sociology of Knowledge. In: Rationality and Relativism. Hg. von Martin Hollis / Stephen Lukes. Oxford 1982, S. 21–47. [\[Nachweis im GVK\]](#)
- Vladimir Batagelj / Anuška Ferligoj / Flaminio Squazzoni: The emergence of a field: a network analysis of research on peer review. In: Scientometrics 113 (2017), H. 1, S. 503–532. DOI: [10.1007/s11192-017-2522-8](#) [\[Nachweis im GVK\]](#)
- Michael Beißwenger: Innovative Möglichkeiten der Arbeit mit digitalen (Open-Access-)Publikationen in der Lehre: Kooperative Texterschließung mit dem textlabor. In: Bücher im Open Access. Ein Zukunftsmodell für die Geistes- und Sozialwissenschaften? Hg. von Dorothee Graf / Yuliya Fadeeva / Katrin Falkenstein-Feldhoff. Opladen u. a. 2020, S. 140–151. PDF. [\[online\]](#) [\[Nachweis im GVK\]](#)
- Mario Biagioli: From Book Censorship to Academic Peer Review. In: Emergences: Journal for the Study of Media & Composite Cultures 12 (2002), H. 1, S. 11–45. DOI: [10.1080/1045722022000003435](#) [\[Nachweis im GVK\]](#)
- David Bloor: Knowledge and Social Imagery. London u. a. 1976. [\[Nachweis im GVK\]](#)
- Tobias Bülte: Welche (Um-)Wege nehmen Metadaten für Open-Access-Bücher vom Verlag zum Discovery System der Bibliothek? In: Bücher im Open Access. Ein Zukunftsmodell für die Geistes- und Sozialwissenschaften? Hg. von Dorothee Graf / Yuliya Fadeeva / Katrin Falkenstein-Feldhoff. Opladen u. a. 2020, S. 43–75. PDF. [\[online\]](#) [\[Nachweis im GVK\]](#)
- 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. (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5) text/html Format. DOI: [10.26298/vra6-mn32](#)
- Kathryn B. H. Clancy / Robin G. Nelson / Julianne N. Rutherford / Katie Hinde: Survey of Academic Field Experiences (SAFE): Trainees Report Harassment and Assault. In: PLoS ONE 9 (2014), H. 7, Artikel e102172 vom 16.07.2014. DOI: [10.1371/journal.pone.0102172](#)
- Harry M. Collins: Changing Order. London u. a. 1985. [\[Nachweis im GVK\]](#)
- Harry M. Collins / Trevor Pinch: The Golem: What Everyone Should Know About Science. Cambridge, NY 1993. [\[Nachweis im GVK\]](#)
- Joe Deville / Jeroen Sondervan / Graham Stone / Sofie Wennström: Rebels with a Cause? Supporting Library and Academic-led Open Access Publishing. In: LIBER Quarterly 29 (2019), H. 1, S. 1. Artikel vom 26.09.2019. DOI: [10.18352/lq.10277](#)
- Leitlinien zur Sicherung guter wissenschaftlicher Praxis, Kodex. Hg. von Deutsche Forschungsgemeinschaft (DFG). Bonn 2019. PDF. [\[online\]](#). [\[Nachweis im GVK\]](#)
- Thomas Ernst: Vom Urheber zur Crowd, vom Werk zur Version, vom Schutz zur Öffnung? Kollaboratives Schreiben und Bewerten in den Digital Humanities. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1) DOI: [10.17175/sb001_021](#)
- Thomas Ernst: Literatur als soziales Medium in der digitalisierten Gesellschaft. Duisburg / Essen 2018, im Erscheinen.
- Yuliya Fadeeva: Open Access und Qualitätssicherung. In: budrich.de. Hg. von Barbara Budrich GmbH. Blogbeitrag vom 30.03.2021. [\[online\]](#)
- Yuliya Fadeeva: Open-Access-Monografien: »Alles Open – Chance oder Risiko?« Dokumentation der Vorträge und der Podiumsdiskussion des Projektworkshops. In: Bücher im Open Access. Ein Zukunftsmodell für die Geistes- und Sozialwissenschaften? Hg. von Dorothee Graf / Yuliya Fadeeva / Katrin Falkenstein-Feldhoff. Opladen u. a. 2020, S. 152–200. PDF. [\[online\]](#) [\[Nachweis im GVK\]](#)
- Katrin Falkenstein-Feldhoff / Dorothee Graf: Explorative Studie der Verkaufs- und Nutzungszahlen. In: Bücher im Open Access. Ein Zukunftsmodell für die Geistes- und Sozialwissenschaften? Hg. von Dorothee Graf / Yuliya Fadeeva / Katrin Falkenstein-Feldhoff. Opladen u. a. 2020, S. 89–110. PDF. [\[online\]](#) [\[Nachweis im GVK\]](#)
- Paul Feyerabend: Against Method. 3. Auflage. London u. a. 1993. [\[Nachweis im GVK\]](#)
- Kathleen Fitzpatrick: Peer-to-peer Review and the Future of Scholarly Authority. In: Social Epistemology 24 (2010), H. 3, S. 161–179. DOI: [10.1080/02691728.2010.498929](#) [\[Nachweis im GVK\]](#)
- Kathleen Fitzpatrick: The Humanities, Done Digitally. In: Debates in the Humanities. Hg. von Matthew K. Gold. 2. Auflage. Minneapolis, MN 2011. [\[Nachweis im GVK\]](#)
- Kathleen Fitzpatrick: Giving It Away: Sharing and the Future of Scholarly Communication. In: Journal of Scholarly Publishing 43 (2012), H. 4, S. 347–362. DOI: [10.3138/jsp.43.4.347](#) [\[Nachweis im GVK\]](#)
- Kathleen Fitzpatrick: Academia, not Edu. In: kfitz.info. Hg. von Kathleen Fitzpatrick. Blogbeitrag vom 26.10.2015. [\[online\]](#)
- Kathleen Fitzpatrick / Avi Santo: Open Review, A Study of Contexts and Practices. Hg. von The Andrew W. Mellon Foundation. In: mellon.org. New York, NY 2012. PDF. [\[online\]](#)
- Ludwig Fleck: Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv. Frankfurt / Main 1980. [\[Nachweis im GVK\]](#)
- Dorothee Graf / Yuliya Fadeeva: Einleitung und abschließende Evaluation des Projekts: Was bleibt nach OGeSoMo (zu tun)? In: Bücher im Open Access. Ein Zukunftsmodell für die Geistes- und Sozialwissenschaften? Hg. von Dorothee Graf / Yuliya Fadeeva / Katrin Falkenstein-Feldhoff. Opladen u. a. 2020, S. 14–42. PDF. [\[online\]](#) [\[Nachweis im GVK\]](#)
- Gary Hall: Does Academia.edu Mean Open Access Is Becoming Irrelevant? In: garyhall.info. Hg. von Gary Hall. Blogbeitrag vom 18.10.2015. [\[online\]](#)
- Predatory Journals and Predatory Publishing. In: Helmholtz Open Science. Hg. von Helmholtz-Gemeinschaft. In: Helmholtz Open Science. Potsdam 2022. [\[online\]](#)
- Stefan Hirschauer: How Editors Decide. Oral Communication in Journal Peer Review. In: Human Studies 38 (2015), H. 1, S. 37–55. DOI: [10.1007/s10746-014-9329-x](#) [\[Nachweis im GVK\]](#)
- Stefan Hirschauer: Peer Review Verfahren auf dem Prüfstand: Zum Soziologiedefizit der Wissenschaftsevaluation. In: Zeitschrift für Soziologie 33 (2004), H. 1, S. 62–83. DOI: [10.1515/zfsoz-2004-0104](#) [\[Nachweis im GVK\]](#)
- Caspar Hirschi: Wie die Peer Review die Wissenschaft diszipliniert. In: Merkur. Artikel vom 27.08.2018. PDF. [\[online\]](#)
- Hypothes.is. Hg. von Hypothesis. 2022. [\[online\]](#)
- Matthew Kirschenbaum: What Is Digital Humanities and What's It Doing in English Departments? In: ADE Bulletin 150 (2010), S. 55–61. DOI: [10.1632/ade.150.55](#) [\[Nachweis im GVK\]](#)
- Michael Kleineberg / Ben Kaden: Open Humanities? ExpertInnenmeinungen über Open Access in den Geisteswissenschaften. In: LIBREAS 32 (2017), S. 1–31. [\[online\]](#)
- Marcel Knöchelmann: Open Science in the Humanities, or: Open Humanities? In: Publications 7 (2019), H. 4, S. 65. DOI: [10.3390/publications7040065](#) [\[Nachweis im GVK\]](#)

- Karin Knorr-Cetina: *Epistemic cultures: how the sciences make knowledge*. Cambridge, MA 1999. [[Nachweis im GVK](#)]
- Karin Knorr-Cetina: *The Manufacture of Knowledge*. Oxford 1981. [[Nachweis im GVK](#)]
- Bruno Latour: *The pasteurization of France*. Cambridge, MA 1988. [[Nachweis im GVK](#)]
- Bruno Latour / Steve Woolgar: *Laboratory life: the construction of scientific facts*. Princeton, NJ 1986. [[Nachweis im GVK](#)]
- Jens Loenhoff / Walter H. Schmitz: Kommunikative und extrakommunikative Betrachtungsweisen: Folgen für Theoriebildung und empirische Forschung in der Kommunikationswissenschaft. In: *Methoden der Geisteswissenschaften: eine Selbstverständigung*. Hg. von Dirk Hartmann / Amir Mohseni / Erhard Reckwitz / Tim Rojek / Ulrich Steckmann. Weilerwist 2012, S. 35–59. [[Nachweis im GVK](#)]
- Robert K. Merton: *The Normative Structure of Science*. In: *The Sociology of Science. Theoretical and Empirical Investigations*. Chicago, IL u. a. 1973, S. 267–278. [[Nachweis im GVK](#)]
- Pierre Mounier: Supporting bibliodiversity in open science: a European challenge. In: av.tib.eu/media/49139. Filmsegment. 2020. DOI: [10.5446/4913910.5446/49139](https://doi.org/10.5446/4913910.5446/49139)
- Pierre Mounier: Towards universal open access? Why we need bibliodiversity rather than a ›silver bullet‹. In: *SciELO in Perspective*. Blogbeitrag vom 2018. [[online](#)]
- Uwe Müller: Standards und Best Practices im Kontext von Open Access. In: *Praxishandbuch Open Access*. Hg. von Konstanze Söllner / Bernhard Mittermaier. Berlin u. a. 2017, S. 53–61. DOI: [10.1515/9783110494068-007](https://doi.org/10.1515/9783110494068-007) [[Nachweis im GVK](#)]
- Stefan Münker: *Emergenz digitaler Öffentlichkeiten: die sozialen Medien im Web 2.0*. Frankfurt / Main 2009. (= Edition Unsel; 26) [[Nachweis im GVK](#)]
- Douglas P. Peters / Stephen J. Ceci: Peer-review practices of psychological journals: The fate of published articles, submitted again. In: *Behavioral and Brain Sciences* 5 (1982), H. 2, S. 187–195. DOI: [10.1017/S0140525X00011183](https://doi.org/10.1017/S0140525X00011183) [[Nachweis im GVK](#)]
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Hg. von Max-Planck-Gesellschaft. Berlin 2003. [[online](#)]
- ResearchGate. Hg. von ResearchGate GmbH. In: researchgate.net. Berlin 2008–2022. [[online](#)]
- Tony Ross-Hellauer: What is open peer review? A systematic review. In: *F1000Research* 6 (2017), S. 588. DOI: [10.12688/f1000research.11369.1](https://doi.org/10.12688/f1000research.11369.1)
- Patrick Sahle: Zwischen Mediengründenheit und Transmedialisierung. In: *Editio. Internationales Jahrbuch für Editionswissenschaft* 24 (2010), S. 23–36. [[Nachweis im GVK](#)]
- Alfred Schütz: *Gesammelte Aufsätze*. 3 Bände. Den Haag 1971–1972. Bd. 1 (1971): *Das Problem der sozialen Wirklichkeit*. [[Nachweis im GVK](#)]
- Pseudo Journale – worum es sich handelt und wie die Wissenschaft gegensteuert. Hg. von Science Media Center Germany. 19.07.2018. PDF [[online](#)]
- Steven Shapin: Phrenological knowledge and the social structure of early nineteenth-century Edinburgh. In: *Annals of Science* 32 (1975), H. 3, S. 219–243. DOI: [10.1080/00033797500200261](https://doi.org/10.1080/00033797500200261) [[Nachweis im GVK](#)]
- Lara Speicher / Lorenzo Armando / Margo Bargheer / Martin Paul Eve / Sven Fund Delfim Leão / Max Mosterd / Frances Pinter / Irakleitos Souyioultzoglou: *Operas Open Access Business Models White Paper*. Hg. von OPERAS. Version 1.0 vom 30.07.2018. In: zenodo.org. DOI: [10.5281/ZENODO.1323707](https://doi.org/10.5281/ZENODO.1323707)
- Peter Suber: *Open Access Overview*. In: *Knowledge Unbound*. Hg. von MIT Press. 14.03.2019. [[online](#)]
- Jonathan P. Tennant / Tony Ross-Hellauer: The limitations to our understanding of peer review. In: *Research Integrity and Peer Review* 5 (2020), H. 1, S. 1–14. DOI: [10.1186/s41073-020-00092-1](https://doi.org/10.1186/s41073-020-00092-1)

Zeitschrift für digitale Geisteswissenschaften

Beitrag aus:
Zeitschrift für digitale Geisteswissenschaften

Titel:
Bye, Bye, Bias! Digital-Humanities-Projekte ethisch überprüfen und gestalten mit Value Sensitive Design

Autor*in:
Katharina Leyrer

Kontakt:
katharina.leyrer@fau.de

Institution:
Friedrich-Alexander-Universität Erlangen-Nürnberg

GND:
1237790735

ORCID:
0000-0002-3051-0621

DOI des Artikels:
[10.17175/sb005_003_v2](https://doi.org/10.17175/sb005_003_v2)

Nachweis im OPAC der Herzog August Bibliothek:
1839573139

Erstveröffentlichung:
08.09.2021

Version 2.0:
21.03.2023

Lizenz:
Sofern nicht anders angegeben 

Medienlizenzen:
Medienrechte liegen bei den Autor*innen

Letzte Überprüfung aller Verweise:
07.09.2021

Format:
PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:
[Informationsethik](#) | [Digital Humanities](#) | [Mensch-Maschine-Kommunikation](#) | [Design](#) | [Bias](#) |

Empfohlene Zitierweise:
Katharina Leyrer: Bye, Bye, Bias! Digital-Humanities-Projekte ethisch überprüfen und gestalten mit Value Sensitive Design. In: Zeitschrift für digitale Geisteswissenschaften NaN (). HTML / XML / PDF. DOI: [10.17175/sb005_003_v2](https://doi.org/10.17175/sb005_003_v2).

Katharina Leyrer

Bye, Bye, Bias! Digital-Humanities-Projekte ethisch überprüfen und gestalten mit Value Sensitive Design

Abstracts

Technologien sind nicht neutral, sondern fördern bestimmte Werte und behindern andere. Das gilt auch für Technologien, die in den Digital Humanities entwickelt und genutzt werden. Daher bedürfen diese Technologien einer ethischen Untersuchung und Einordnung. Der Ansatz des Value Sensitive Design ermöglicht einerseits, systematisch zu analysieren, welche Interessen und Werte in eine bestimmte Technologie eingeschrieben werden und ist zugleich ein Framework, um Technologien wertensensitiv zu entwickeln. Der Beitrag zeigt auf, wie Value Sensitive Design als theoretischer und methodischer Ansatz in den unterschiedlichen Spielarten der Digital Humanities genutzt werden kann.

Technological systems are not neutral but promote certain values and hinder others. That is also true for technologies that are developed and used in the field of Digital Humanities. Therefore, these technologies need to be reviewed from an ethical point of view. Value Sensitive Design is a framework to systematically analyze whose interests and values are implemented in a technological system. At the same time, Value Sensitive Design facilitates accounting for human values when designing technologies. This contribution points out how Value Sensitive Design can be used as a theoretical and methodological approach in different areas of Digital Humanities activities.

1. Einleitung

Eine Gesichtserkennungssoftware, die in polizeilichen Ermittlungen in den USA zum Einsatz kommt, erzeugt überdurchschnittlich oft falsche Matches, wenn sie Bilder Schwarzer Menschen abgleicht – weil die Software mit Gesichtern *weißer* Menschen trainiert wurde. Die Technologie ruft damit rassistische Diskriminierung hervor.¹ Ein selbstlernendes System, das der Konzern *Amazon* zur Vorauswahl von Bewerber*innen einsetzte, diskriminierte auf Basis binärer Geschlechtszuschreibungen: Die Software bewertete Bewerbungen von Frauen systematisch schlechter als die von Männern.² Stereotype auf Basis von religiöser Zugehörigkeit reproduzierte das Sprachmodell *GPT-3*, zu dessen Nutzer*innen *Microsoft* gehört: Das Sprachmodell assoziierte Menschen muslimischen Glaubens mit Gewalt.³ Diese Beispiele zeigen: Technologien können bestimmte Akteur*innen systematisch benachteiligen. Solche Technologien beschreiben Friedman und Nissenbaum als »biased«, da sie »*systematically and unfairly discriminate* against certain individuals or groups of individuals in favor of others«. ⁴ Das Design von Technologien bedarf daher einer ethischen Untersuchung.⁵

Warum ist das für die Digital Humanities von Bedeutung? Jedes Digital-Humanities-Projekt hat in irgendeiner Weise mit Technologien zu tun, z. B. wenn Akteur*innen digitale Datensätze erstellen, kuratieren und nutzen, Tools und Analyseverfahren entwickeln und anwenden, Online-Communities und computervermittelte Interaktionen geisteswissenschaftlich untersuchen⁶ oder digitale Publikationen und E-Learning-Angebote realisieren.⁷ Diese Technologien können – wie alle anderen auch – Bias aufweisen. Aktuell finden ethische Aspekte bei der Konzeption und Realisierung von Digital-Humanities-Projekten, -Technologien und -Infrastrukturen jedoch wenig Beachtung.⁸ Für die gute wissenschaftliche Praxis, wie sie beispielsweise die DFG in ihren Leitlinien festschreibt, ist eine ethische Überprüfung von Forschungsvorhaben zugleich unabdingbar.⁹

¹ Vgl. Gierlinger 2020; Kretschmar 2020.

² Vgl. Holland 2018.

³ Vgl. Holland 2021.

⁴ Friedman / Nissenbaum 1996, S. 332; Hervorhebungen im Original.

⁵ Die Erkenntnis, dass Technologien nicht neutral, sondern von ihren gesellschaftlichen Rahmenbedingungen geprägt sind, ist dabei nicht neu. Bereits in den 1990er-Jahren thematisierten verschiedene techniksoziologische Ansätze, dass soziale, politische, technische und ökonomische Faktoren die Gestaltung von Technologien maßgeblich beeinflussen, u. a. Social Construction of Technology (SCOT) und Social Shaping of Technology (SST; vgl. Häußling 2019, S. 198–206).

⁶ Vgl. Roth 2019.

⁷ Vgl. Burghardt 2020.

⁸ Vgl. Rojas Castro 2020; Noble 2019. Warum spielen ethische Aspekte in DH-Projekten bislang kaum eine Rolle? Dazu gibt es nach Kenntnis der Autorin keine umfassende Untersuchung. Daher kann hier nur ein Erklärungsansatz angeboten werden, der sich auf einzelne ethische Aspekte in den DH bezieht: Noble schreibt in Bezug auf die Reproduktion kolonialer Machtverhältnisse in den Digital Humanities, dass sich die DH als »neutral« und damit »noncolonial« verstehen – dies aber nicht sind (Noble 2019, S. 28f.). Ein Grund könnte also sein, dass die ethische Bedeutung und Auswirkung von DH-Projekten den Forschenden nicht oder kaum bewusst sind. Zudem beschreibt Noble eine Art Teufelskreis: Weil sich die DH nicht eindeutig gegen rassistische und sexistische Gewalt positionieren, verlassen Frauen und Schwarze Menschen das Forschungsfeld (Noble 2019, S. 29f.). Dies könnte wiederum zur Folge haben, dass ethische Aspekte im Kontext von Rassismus und Sexismus noch weniger Beachtung finden.

⁹ Vgl. Deutsche Forschungsgemeinschaft (Hg.) 2019, Leitlinie 10.

Dieser Beitrag geht daher der Frage nach, wie Technologien, die in den Digital Humanities entwickelt und genutzt werden, ethisch überprüft und wertebasiert gestaltet werden können. Ziel des Beitrages ist es, zu zeigen, wie Digital Humanities-Projekte auf Basis des Value-Sensitive-Design-Ansatzes aus einer ethischen Perspektive reflektiert und so gestaltet werden können, dass sie die Werte aller beteiligten Akteur*innen berücksichtigen.

Zunächst wird beschrieben, welche verschiedenen Formen von Bias Technologien aufweisen können. Zudem wird skizziert, welche Technologien in den Digital Humanities eine Rolle spielen. Da ethische Reflektionen in den Digital Humanities an die Informations- und Technikethik anknüpfen können, werden daraufhin diese beiden Bereichsethiken kurz umrissen. Anschließend wird der Value-Sensitive-Design-Ansatz vorgestellt, mit dem Technologien wertebasiert analysiert und entwickelt werden können. Obwohl der Ansatz in Feldern wie der Informationstechnologie, dem Bauingenieurwesen und der Gesundheitstechnologie bereits etabliert ist, hat die Anwendung von Value Sensitive Design in den Digital Humanities experimentellen Charakter.¹⁰ Daher wird in Kapitel 5 anhand konkreter Beispiele vorgestellt, wie Value Sensitive Design in den verschiedenen Teilbereichen der Digital Humanities angewendet werden kann.

2. Wie kommt der Bias in die Technologie?

Nach Friedman und Nissenbaum kann Bias in technischen Systemen auf drei unterschiedliche Weisen entstehen.¹¹ Sie unterscheiden *Pre-Existing Bias*, *Technical Bias* und *Emergent Bias*. Mit Pre-Existing Bias sind Verzerrungen gemeint, die bereits vor der Entwicklung der Technologie existieren: Die Designer*innen oder Auftraggeber*innen schreiben – bewusst oder unbewusst – bereits existierende persönliche oder gesellschaftliche Vorurteile in das System ein. Ein Beispiel: Die automatische Zugangskontrolle zu den Umkleieräumen eines Fitnessstudios in London ordnete Personen mit einem Dokortitel automatisch als männlich ein und verwehrte ihnen den Zugang zu den Damenumkleiden – weil die Designer*innen des Systems davon ausgegangen waren, dass nur Männer einen Dokortitel haben können.¹² Auch die oben bereits erwähnte Software, die *Amazon* zur Vorauswahl von Bewerber*innen einsetzte, hat einen Pre-Existing Bias: Da das Unternehmen bislang vor allem Männer eingestellt hatte, ging das selbstlernende System davon aus, dass Frauen weniger geeignet für eine Einstellung bei *Amazon* seien.¹³

Technical Bias entsteht hingegen bei der »resolution of issues in the technical design«¹⁴, beispielsweise durch Limitationen der Hard- und Software oder wenn Algorithmen bestimmte Akteur*innen bevorzugen und andere benachteiligen. So führt die beschränkte Größe von Bildschirmen beispielsweise dazu, dass die Ergebnisliste von Suchmaschinen in verschiedene »Seiten« aufgeteilt wird, die nur nacheinander auf dem entsprechenden Bildschirm angezeigt werden können. Suchergebnisse, die auf der zweiten Seite angezeigt werden, werden jedoch weniger oft angeklickt. Bevorzugt der Algorithmus der Suchmaschine zusätzlich bestimmte Websites und zeigt diese ganz oben in der Ergebnisliste an, werden diese also deutlich öfter angeklickt – es liegt eine unfaire, systematische Diskriminierung vor. Technical Bias kann aber auch entstehen, wenn »Human Constructs«¹⁵ wie Urteile, Diskurse oder Institutionen formalisiert und quantifiziert werden. Dies passiert beispielsweise in Software, die voraussagt, wie hoch die Wahrscheinlichkeit ist, dass ein*e Straftäter*in wiederholt delinquent wird.¹⁶

Emergent Bias entsteht hingegen erst im Anwendungskontext der Technologie, also wenn Personen die Technologie nutzen. Gründe für Emergent Bias können sein, dass »new knowledge in society«¹⁷ nicht in die Technologie aufgenommen wird oder die Nutzer*innen der Technologie andere Fähigkeiten und Werte haben als die Nutzer*innengruppe, für welche die Technologie ursprünglich designt wurde. Wenn beispielsweise ein Geldautomat, der v. a. schriftbasiert mit Nutzer*innen kommuniziert (z. B. »Geben Sie unten links Ihre Karte ein«) in einer Gegend aufgestellt wird, in der viele Analphabet*innen wohnen, entsteht ein Emergent Bias: Die tatsächlichen Nutzer*innen unterscheiden sich in ihren Fähigkeiten von den angenommenen Nutzer*innen der Technologie. Vor allem Benutzeroberflächen weisen oft Emergent Bias auf, da sie mit Blick auf die Fähigkeiten und Gewohnheiten bestimmter Zielgruppen gestaltet werden. Diese können sich jedoch im Lauf der Zeit verändern.¹⁸

¹⁰ Für eine ausführliche Zusammenstellung bisheriger Anwendungsbeispiele vgl. Friedman / Hendry 2019 und Winkler / Spiekermann 2018.

¹¹ Vgl. Friedmann / Nissenbaum 1996.

¹² Vgl. Wachter-Boettcher 2017, S. 6.

¹³ Vgl. Holland 2018.

¹⁴ Friedman / Nissenbaum 1996, S. 334.

¹⁵ Friedman / Nissenbaum 1996, S. 334.

¹⁶ Vgl. Wachter-Boettcher 2018, S. 119–129.

¹⁷ Friedman / Nissenbaum 1996, S. 335.

¹⁸ Vgl. Friedman / Nissenbaum 1996, S. 335.

3. Technologien in den Digital Humanities

Technologien können also auf unterschiedliche Weise Bias hervorbringen. Was bedeutet das für die Digital Humanities?

Nach Friedman und Hendry umfassen *Technologische Systeme* sowohl *Werkzeuge* und *Technologien*, als auch *Infrastrukturen* und *Policies*. Unter *Werkzeugen* verstehen sie dabei »physical artifacts that augment human activity«¹⁹ wie z. B. eine Steinaxt oder ein Messer. Wenn darüber hinaus wissenschaftliche Erkenntnisse angewandt werden, um praktische Probleme zu lösen (beispielsweise bei Traktoren und Computern), sprechen Friedman und Hendry von Technologien – wobei die Übergänge zwischen Werkzeugen und Technologien fließend sind. Infrastrukturen beschreiben schließlich alle Einrichtungen und Strukturen, die nötig sind, um ein Vorhaben oder gesellschaftliche Aktivitäten zu realisieren, beispielsweise Straßen und Energieversorgung. Werkzeuge, Technologien und Infrastrukturen stehen dabei in wechselseitiger Abhängigkeit und ergeben in ihrem Zusammenspiel sogenannte Technologische Systeme. Im Folgenden wird der Begriff Technologien als Kurzform für Technologische Systeme genutzt und beschreibt dabei Werkzeuge, Technologien und Infrastrukturen genauso wie deren Zusammenspiel. Schließlich fassen Friedman und Hendry auch Policies, also Regelwerke oder Strategien, unter den Technologie-Begriff, da Policies als »some combination of tool and infrastructure«²⁰ verstanden werden und menschliches Handeln ebenfalls prägen.²¹

Welche Technologien werden in den DH also genutzt und angewendet? Eine systematische Orientierung bieten dabei die Spielarten der DH, die Roth und Burghardt unterscheiden: *Digitized Humanities*, *Numerical Humanities*, *Humanities of the Digital* und *Public Humanities*.²²

Mit dem Begriff Digitized Humanities beschreibt Roth »the creation, curation, and use of digitized data sets«²³ in den Geistes- und Sozialwissenschaften. Akteur*innen in den Digitized Humanities »digitize, store, process, gather, connect, manage, make available, mine and visualize«²⁴ Texte, Bilder und Multimedia-Artefakte. Dafür entwickeln und nutzen sie Software. In dieser Spielart der DH sind also beispielsweise Datenmanagement- und Datenspeichersysteme, aber auch Visualisierungs-, Analyse-, Scraping- und Optical Character Recognition (OCR)-Tools sowie Benutzeroberflächen von Bedeutung. Darüber hinaus spielen digitale Bibliotheken und Archive als Infrastrukturen eine Rolle, genauso wie Policies für die Erhebung, Kuratierung, Archivierung und Verfügbarmachung von Datensätzen.

In der Kategorie der Numerical Humanities geht es hingegen darum, soziale Prozesse zu formalisieren, indem mathematisch abstrahierte Modelle erstellt werden. Ziel ist dabei, »to capture the possibly general mechanisms at the root of the observed data«²⁵. Prozesse softwaregestützt zu realisieren, die andernfalls langwierig oder schwierig umsetzbar sind – wie beispielsweise bei OCR oder statistischen Analysen – versteht Roth dabei nicht als Numerical Humanities; vielmehr zeichnet sich diese Spielart dadurch aus, dass »numerical models of human or social behavior per se«²⁶ entwickelt werden. Hier kommen vor allem selbstlernende Systeme zum Einsatz, beispielsweise Mustererkennungssoftware, die nach Gesetzmäßigkeiten in bestimmten Datensätzen sucht.

Schließlich beschreibt Roth die Humanities of the Digital als weitere Spielart der DH und meint damit die geisteswissenschaftliche Forschung zu computervermittelter Kommunikation, Mensch-Maschine-Interaktion und Online-Gemeinschaften. Online-Communities werden als sozio-technische Systeme verstanden, die sich durch eigene Regeln und Settings auszeichnen und die es zu untersuchen gilt. So gehören neben Studien zur Rolle von digitaler Kommunikation im Alltag auch Untersuchungen zu Wikipedia, Blogs, Sozialen-Netzwerk-Seiten und Gaming-Plattformen zu den Humanities of the Digital.²⁷ In dieser Spielart geht es – anders als in den Digitized Humanities und den Numerical Humanities – weniger darum, Technologien anzuwenden und zu entwickeln; technologische Artefakte und Infrastrukturen sind vielmehr der Untersuchungsgegenstand.

¹⁹ Friedman / Hendry 2019, S. 19.

²⁰ Friedman / Hendry 2019, S. 20.

²¹ Vgl. Friedman / Hendry 2019, S. 19f.

²² Vgl. Roth 2019; Burghardt 2020. Die Spielarten von Roth und Burghardt werden in diesem Beitrag als Grundlage gewählt, da sie u. a. auf Basis einer empirischen Untersuchung von Zeitschriftenartikeln und Konferenzbeiträgen einen aktuelleren Überblick über die verschiedenen Teilbereiche der DH-Forschung geben als Standardwerke wie Jannidis et al. (Hg.) 2017 oder Schreibman et al. (Hg.) 2015.

²³ Roth 2019, S. 616.

²⁴ Roth 2019, S. 616.

²⁵ Roth 2019, S. 618.

²⁶ Roth 2019, S. 618.

²⁷ Vgl. Roth 2019, S. 623–625.

Den drei von Roth beschriebenen Teilbereichen der DH fügt Burghardt noch einen vierten hinzu, die Public Humanities. Diese umfassen die Themenfelder des digitalen Publizierens, der Wissenschaftskommunikation und des E-Learning.²⁸ Hier spielen Infrastrukturen wie Zeitschriften, Portale, Verlage und deren Policies eine Rolle, aber auch Soft- und Hardware, die für die Rezeption digitaler Publikationen und E-Learning-Ressourcen benötigt werden, z. B. E-Book-Formate oder User Interfaces.

4. Ethik in den DH: Informations- und Technikethik

Bisher wurde also deutlich, dass in Digital-Humanities-Projekten verschiedenste Technologien eine Rolle spielen und dass jede Technologie verschiedene Formen von Bias enthalten kann. Damit kann auch jede Technologie, die in DH-Projekten entwickelt oder angewendet wird, *biased* sein und bedarf einer ethischen Überprüfung. Wie aber können wir den ethischen Herausforderungen begegnen, die sich aus den verschiedenen Technologien in den Digital Humanities ergeben?²⁹

Allgemein versucht Ethik die Frage zu beantworten, »wie menschliches Handeln möglich wird, das als gut, richtig und wünschenswert empfunden werden kann«³⁰. Dabei gibt es Handlungsbereiche, die spezifische ethische Probleme hervorbringen, welche nur in ihrem jeweiligen konkreten Kontext gelöst werden können. Um diese Probleme zu adressieren, haben sich verschiedene Bereichsethiken entwickelt, beispielsweise die Informations-, die Technik-, die Medien- und die Computerethik. Diese Bereichsethiken sind nicht scharf voneinander abgrenzbar und haben viele Berührungspunkte.³¹

Um ethische Probleme in den Digital Humanities zu diskutieren, kann vor allem an Ansätze der Informations- und Technikethik angeknüpft werden. Die Informationsethik befasst sich mit ethischen Fragestellungen, die im Zusammenhang mit Informations- und Kommunikationstechnologien³² und »aus den Anwendungen des Internets, des Informationsmanagements und der Computernutzung insgesamt hervorgehen«³³ – und damit also mit den Technologien, die in den DH genutzt und entwickelt werden (vgl. Kapitel 3). Im Fokus steht dabei, wie sich Informations- und Kommunikationstechnologien auf Umwelt und Gesellschaft auswirken.³⁴ Als zentrale Grundwerte der Informationsethik identifiziert Rösch Informationsfreiheit, Informationsgerechtigkeit, Privatheit, Geistiges Eigentum, Informationsqualität und Informationsökologie.³⁵

Darüber hinaus sind in den Digital Humanities auch ethische Aspekte von Bedeutung, die im Zusammenhang mit der sozialen Praxis, der Genese und der Einbettung von Technologien entstehen. Diese sind Gegenstand der Technikethik: Deren Aufgabe ist es, ethische Aspekte, die durch technische Innovationen und den »wissenschaftlich-technischen Fortschritt[...]«³⁶ entstehen, zu reflektieren und Entscheidungen über Technik auf Basis ethischer Argumente zu ermöglichen. Technikethik beschäftigt sich auch mit der »sozialen Praxis der Technik und ihrer Einbettung, der Praxis der Genese technischer Produkte und Technologien«³⁷. Anders als bei der Informationsethik gehen die zentralen Werte der Technikethik dabei deutlich über den Bezugsrahmen der Technik selbst hinaus und umfassen neben Effektivität, Effizienz und Sicherheit auch Gesundheit, menschliches Wohlbefinden, (ökologische) Nachhaltigkeit, Gerechtigkeit, Demokratie, Inklusivität³⁸ und Geschlechtergerechtigkeit.³⁹

In der Informations-, der Technik- und der Computerethik haben sich verschiedene Methoden und Frameworks entwickelt, um die ethische Gestaltung von Projekten und Technologien zu unterstützen, beispielsweise *Embedded Ethics* oder *Partizipatives Design*.⁴⁰ Kern des Embedded-Ethics-Ansatzes ist es, eine oder mehrere Ethiker*innen an allen Entscheidungen im gesamten Verlauf eines Projektes zu beteiligen.⁴¹ Das Konzept des Partizipativen Designs legt hingegen den Fokus darauf, potentielle Nutzer*innen in den Design-Prozess einzubinden.⁴²

²⁸ Vgl. Burghardt 2020.

²⁹ An dieser Stelle herzlichen Dank an Frauke Schade, die als eine der Reviewerinnen dieses Textes ein Kapitel zu Informationsethik angeregt hat.

³⁰ Rösch 2021, S. 3.

³¹ Vgl. Rösch 2021, S. 33.

³² Vgl. Burgess et al. 2019, S. 3.

³³ Heesen 2016, S. 2.

³⁴ Vgl. Rösch 2021, S. 58.

³⁵ Vgl. Rösch 2021, S. 55.

³⁶ Grunwald / Hillerbrand 2021, S. 5.

³⁷ Grunwald / Hillerbrand 2021, S. 5.

³⁸ Vgl. van de Poel 2021, S. 129.

³⁹ Vgl. Ernst 2021, S. 110.

⁴⁰ Eine Übersicht über Ansätze, die Werte beim Design von Technologien einbeziehen, findet sich z. B. bei Simon 2016 und bei Friedman / Hendry 2019, S.15f.

⁴¹ Vgl. MacLennan et al. 2020.

⁴² Vgl. Simonsen / Robertson (Hg.) 2013.

Einen Schwerpunkt auf Informations- und Kommunikationstechnologien legt der Ansatz des Value Sensitive Design,⁴³ der von Batya Friedman bereits in den 1990er-Jahren entwickelt wurde⁴⁴ und sowohl in der Technik-, als auch in der Informationsethik eine bedeutende Rolle spielt.⁴⁵ Friedman et al. kritisierten, dass die Forschung zur ethischen Gestaltung von Technologien bislang lediglich einzelne Werte wie Privatsphäre, Autonomie oder informierte Einwilligung in den Blick genommen hatte.⁴⁶ Daher schlugen sie mit Value Sensitive Design ein »overarching theoretical and methodological framework«⁴⁷ vor, mit dem einerseits systematisch analysiert werden kann, welche Werte in eine bestimmte Technologie eingeschrieben sind. Zugleich ermöglicht das Framework, Technologien wertesensitiv zu entwickeln, indem die Werte aller betroffenen Akteur*innen im gesamten Design-Prozess reflektiert und berücksichtigt werden.⁴⁸

Mit Value Sensitive Design können Forscher*innen Technologien und ihre Anwendung also nicht nur ethisch reflektieren, sondern auch im Sinne normativer Ethik »verändernd, regulierend«⁴⁹ auf die Entwicklung von Technologien und deren Rahmenbedingungen einwirken. Seit seiner Entstehung vor über 30 Jahren wurde der Value-Sensitive-Design-Ansatz stetig weiterentwickelt und in unterschiedlichen Bereichen angewendet, u. a. in der Informationstechnologie, der Informatik, der Philosophie, der Gesundheitstechnologie und der Pädagogik.⁵⁰

5. Value Sensitive Design: Grundlagen und methodisches Vorgehen

Um Technologien wertebasiert zu entwickeln sieht Value Sensitive Design eine integrative und iterative Methodologie mit drei Untersuchungsabschnitten vor: den *konzeptionellen Überlegungen*, den *empirischen Untersuchungen* und den *technischen Investigationen*.

5.1 Drei Untersuchungsabschnitte von Value Sensitive Design

In den konzeptionellen Überlegungen wird herausgearbeitet, um welche zentralen Themen, Technologien und Werte es geht (siehe auch Beispiele in Kapitel 5.2). Im Mittelpunkt stehen dabei die Fragen:

- Welche Technologie wird analysiert oder gestaltet?
- Welche Stakeholder sind direkt oder indirekt von der jeweiligen Technologie betroffen?
- Welche Werte und Wertkonflikte spielen eine Rolle?

Zentral ist im Value Sensitive Design dabei der *Stakeholder*-Begriff: Dieser bezieht – anders als der Begriff *User* – auch diejenigen Akteur*innen mit ein, die zwar nicht direkt mit der Technologie interagieren, auf die sich die Technologie aber dennoch auswirkt: die sogenannten *indirekten Stakeholder*. Akteur*innen, die direkt mit der jeweiligen Technologie interagieren, werden dagegen als *direkte Stakeholder* verstanden. Stakeholder können dabei sowohl Individuen, Gruppen, Organisationen und Gesellschaften als auch zukünftige Generationen sein. In den konzeptionellen Überlegungen wird also analysiert, welche Stakeholder direkt oder indirekt von der jeweiligen Technologie betroffen sind und auf welche Weise.⁵¹

Auch die Werte und Wertkonflikte, die im Kontext der Technologie eine Rolle spielen, werden im Rahmen der konzeptionellen Untersuchungen expliziert und konkret beschrieben. Ein Wert ist dabei, »what a person or group of people consider important in life«⁵². Da der Value-Sensitive-Design-Ansatz den Werte-Begriff damit sehr weit fasst, sieht er vor, auf Basis von Literatur und bisheriger Forschung zu spezifizieren, was genau unter einem bestimmten Wert im jeweiligen Kontext verstanden wird.⁵³ Als

⁴³ Vgl. Grunwald / Hillerbrand 2021, S. 469.

⁴⁴ Vgl. Friedman 1996.

⁴⁵ So wird der Ansatz sowohl in Handbüchern für Technikethik (Vgl. z. B. Grunwald / Hillerbrand 2021), als auch in Handbüchern für Informationsethik vorgestellt (Vgl. z. B. Heesen 2016).

⁴⁶ Vgl. Friedman et al. 2008, S. 69f.

⁴⁷ Friedman et al. 2008, S. 70.

⁴⁸ Vgl. Friedman et al. 2008, S. 69.

⁴⁹ Rösch 2021, S. 14.

⁵⁰ Für eine ausführliche Zusammenstellung bisheriger Anwendungsbeispiele vgl. Friedman / Hendry 2019 und Winkler / Spiekermann 2018.

⁵¹ Vgl. Friedman / Hendry 2019, S. 35–44.

⁵² Friedman et al. 2008, S. 70.

⁵³ Vgl. Friedman et al. 2008, S. 72.

Ausgangspunkt bieten Friedman und Hendry hierfür eine Liste mit 13 Werten, die beim Design technischer Systeme oft eine Rolle spielen und von ethischer Bedeutung sind. Die Autor*innen weisen aber ausdrücklich darauf hin, dass diese Liste nicht vollständig ist.⁵⁴ Die Liste umfasst:

- »Human welfare«
- »Ownership and property«
- »Privacy«
- »Freedom from bias«
- »Universal Usability «
- »Trust«
- »Autonomy«
- »Informed consent«
- »Accountability«
- »Courtesy«
- »Identity«
- »Calmness«
- »Environmental sustainability«⁵⁵

Human welfare meint beispielsweise körperliches und psychisches Wohlbefinden. Der Wert *Courtesy* bezieht sich darauf, Akteur*innen mit Höflichkeit und Rücksichtnahme zu begegnen. *Identity* beschreiben Friedman und Hendry als »people's understanding of who they are over time, embracing both continuity and discontinuity«⁵⁶. Leider kann an dieser Stelle keine ausführliche Diskussion der Werte erfolgen, die Friedman und Hendry aufzählen. Eine Kurzdefinition und ausführliche Literaturhinweise finden sich aber zu jedem der Werte bei Friedman et al.⁵⁷

Wie oben erwähnt ist bei Value Sensitive Design jedoch entscheidend, dass auf Basis von Literatur herausgearbeitet wird, was unter einem bestimmten Wert im Kontext einer konkreten Technologie verstanden wird.⁵⁸ Eine umfangliche Konzeptualisierung der Werte, die in den Digital Humanities eine Rolle spielen, kann also nicht unabhängig vom konkreten Kontext der jeweiligen Technologie geleistet werden. In Kapitel 5 wird gezeigt, dass in den Digital Humanities über die hier genannten Werte hinaus beispielsweise auch Geschlechtergerechtigkeit, Barrierefreiheit, globale Gerechtigkeit im Sinne von Teilhabe und Freiheit von Diskriminierung⁵⁹, aber auch die in der Informationsethik zentralen Werte Informationsfreiheit und -gerechtigkeit eine Rolle spielen.

Im zweiten Untersuchungsabschnitt, den empirischen Untersuchungen, wird mittels qualitativer und quantitativer Methoden analysiert, wie die betroffenen Akteur*innen die jeweiligen Werte im Kontext der Technologie bewerten und welche Prioritäten sie bei Wertkonflikten setzen (siehe auch Beispiele in Kapitel 5.2). Dieser Teilbereich des Value Sensitive Design sieht darüber hinaus vor, zu einem späteren Zeitpunkt im Design-Prozess empirisch zu prüfen, inwiefern die Technologie die entsprechenden Werte in der Anwendung tatsächlich fördert.

Die technischen Investigationen haben ebenfalls zwei Aufgaben: Erstens, zu analysieren, wie sich bestimmte bereits existierende Charakteristika einer Technologie auf verschiedene Werte auswirken. Zweitens wird in den technischen Investigationen überlegt, wie die Technologie so gestaltet werden kann, dass sie die Werte unterstützt, die in den konzeptionellen und empirischen Untersuchungen herausgearbeitet wurden (siehe auch Beispiele in Kapitel 5.2).⁶⁰ Friedman und Hendry betonen dabei, dass Technologien eng mit den sozialen Strukturen verknüpft sind, in die sie eingebettet sind. Um bestimmte Werte zu fördern, muss daher nicht in jedem Fall das technische Design angepasst werden. Eine Lösung kann auch darin bestehen, Policies oder soziale Strukturen zu verändern.⁶¹

Die drei Untersuchungsabschnitte sind nicht voneinander unabhängig, sondern beeinflussen sich wechselseitig. Die Reihenfolge der Untersuchungsabschnitte ist daher nicht vorab festgelegt; vielmehr ist eine iterative Wiederholung der Investigationen vorgesehen (vgl. Abbildung 1).⁶²

⁵⁴ Vgl. Friedman / Hendry 2019, S. 22-29.

⁵⁵ Alle Friedman / Hendry 2019, S. 28.

⁵⁶ Friedman / Hendry 2019, S. 28.

⁵⁷ Vgl. Friedman et al. 2008, S. 90f.

⁵⁸ Vgl. Friedman et al. 2008, S. 72.

⁵⁹ Vgl. Vereinte Nationen Generalversammlung 2014, S. 8.

⁶⁰ Vgl. Friedman et al. 2008, S. 73.

⁶¹ Friedman / Hendry 2019, S. 20.

⁶² Winkler / Spiekermann 2018, S. 3.

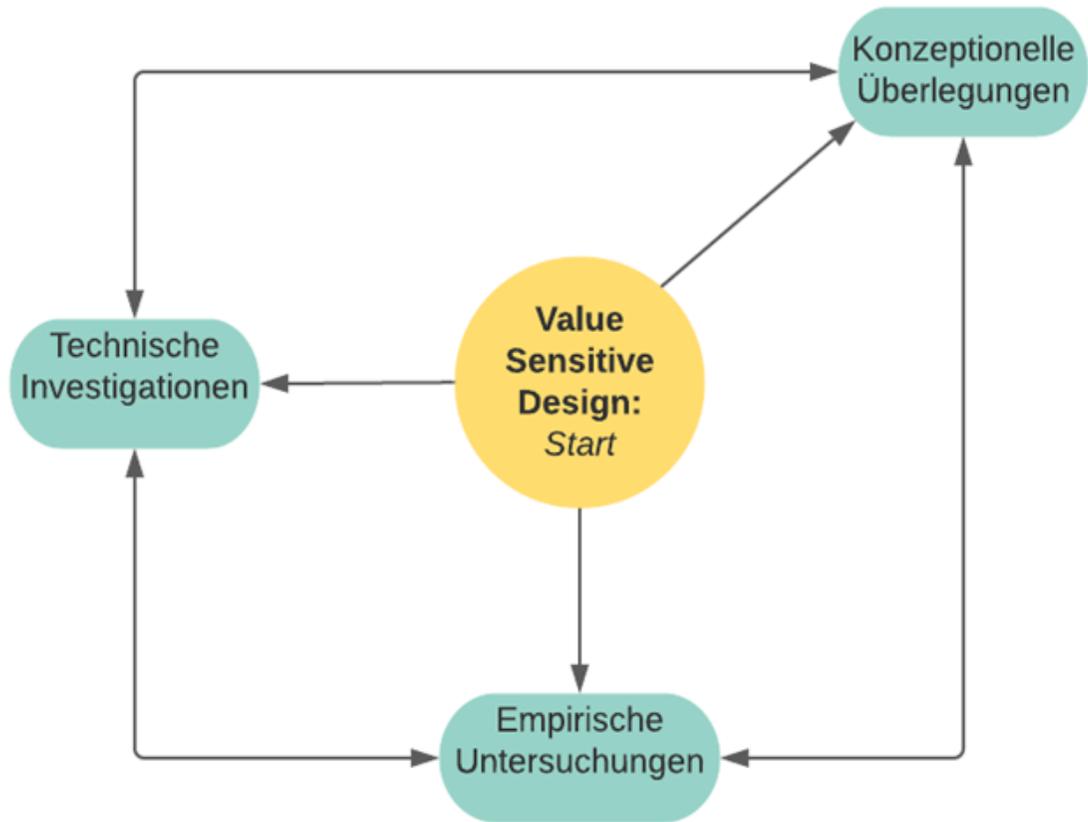


Abb. 1: Die drei Untersuchungsschritte von Value-Sensitive-Design-Analysen. [Leyrer 2021]

5.2 Anwendung von Value Sensitive Design: Zwei Beispiele

So weit, so theoretisch: Wie aber sieht es mit Value-Sensitive-Design-Analyse in der Praxis aus? Im Folgenden werden zwei Projekte vorgestellt, die eine Technologie auf Basis von Value Sensitive Design entwickelt haben.

5.2.1 ›Room with a view‹: Virtuelle Fenster in Büroräumen

In dem Projekt ›Room with a view‹ ging es darum, in fensterlosen Büroräumen Bildschirme aufzustellen, auf denen eine Live-Übertragung des Platzes vor dem Bürogebäude zu sehen war. Friedman et al. begleiteten das Projekt mit Value Sensitive Design.⁶³

In den konzeptionellen Überlegungen stellten Friedman et al. die These auf, dass der Blick auf ein ›virtuelles Fenster‹ – also die Live-Übertragung einer Straßenszene auf Plasma-Bildschirme – die Gesundheit, das Wohlbefinden und die Kreativität der Beschäftigten steigert. Basis dafür waren verschiedene psychologische Studien, die zeigen, dass der Blick auf Natur-Szenen positive physiologische und psychologische Auswirkungen hat und beispielsweise Stress und Krankheiten reduziert.⁶⁴ Direkte Stakeholder der Technologie sind also die Beschäftigten in den fensterlosen Büroräumen, deren Werte Gesundheit und Kreativität gefördert werden sollen. Die ›virtuellen‹ Fenster haben aber auch Auswirkungen auf Personen, die sich auf dem Platz vor dem Bürogebäude aufhalten: Sie werden gefilmt und auf den Bildschirmen in den Büros abgebildet. Sie sind damit indirekte Stakeholder der Technologie. Die Live-Übertragung des Platzes auf Bildschirme in den Bürogebäuden könnte damit die Gefährdung der Privatsphäre der indirekten Stakeholder mit sich bringen.⁶⁵

⁶³ Friedman et al. 2008, S. 75–80.

⁶⁴ Friedman et al. beziehen sich hier auf Ulrich 1984 und Ulrich 1993 sowie auf Kahn 1999.

⁶⁵ Vgl. Friedman et al. 2008, S. 75–80.

Ausgehend von diesen konzeptionellen Überlegungen führten Friedman et al. verschiedene empirische Untersuchungen durch, um zu prüfen, inwiefern die Technologie der ›virtuellen Fenster‹ die unterschiedlichen Werte der direkten und indirekten Stakeholder tatsächlich fördert oder behindert. So untersuchten sie zunächst in einem experimentellen Setting, welche Effekte ein Bildschirm, auf dem eine Straßenszene vor dem Bürogebäude live übertragen wird, auf die Büroangestellten hat. Hierzu erhoben die Forscher*innen verschiedene Parameter wie die Blickrichtung, den Puls und verschiedene Leistungsdaten der Büroangestellten und führten anschließend qualitative Interviews. Dabei kamen sie zu dem Ergebnis, dass die Live-Übertragung auf dem Bildschirm eine ähnlich erholsame Wirkung hat wie ein echtes Fenster. Einen größeren physiologischen Erholungseffekt wiesen echte Fenster im Vergleich zu ›virtuellen Fenstern‹ nur dann auf, wenn die Teilnehmenden länger als 30 Sekunden auf das jeweilige ›Fenster‹ blickten.⁶⁶

Zusätzlich untersuchten Friedman et al. die Interessen der indirekten Stakeholder – also der Passant*innen auf dem Platz vor dem Bürogebäude –, indem sie diese zu ihrer Einstellung zu Privatsphäre im öffentlichen Raum allgemein und zur Live-Übertragung ihres Bildes auf die Büro-Bildschirme befragten. Dabei wurde deutlich, dass die Passant*innen die Werte Privatsphäre, informierte Einwilligung und Sicherheit sehr hoch bewerteten. Mit dieser Vielfalt empirischer Methoden untersuchten die Forscher*innen also, welche Werte für die verschiedenen betroffenen Akteur*innen im Kontext der ›Virtuelles Fenster‹-Technologie wichtig sind. Dabei wurde deutlich, dass ein Wertkonflikt vorliegt: Während die Technologie die Werte der direkten Stakeholder steigert – also Gesundheit, Wohlbefinden und Kreativität –, gefährdet sie die Werte der indirekten Stakeholder – Privatsphäre, Sicherheit und informierte Einwilligung.⁶⁷

Wie kann mit einem solchen Wertekonflikt umgegangen werden? In den technischen Investigationen empfehlen die Forscher*innen, Gebäude grundsätzlich mit Fenstern zu bauen, da der Blick auf eine Natur-Szene die Werte der Büroangestellten fördert und durch ›virtuelle Fenster‹ nur teilweise ersetzt werden kann.⁶⁸ Um im Falle von Gebäuden, die nun mal bereits ohne Fenster gebaut wurden, dennoch einen positiven Effekt für Büroangestellte durch ›virtuelle Fenster‹ zu erzielen, ohne die Werte von Passant*innen auf dem Platz vor dem Bürogebäude als indirekte Stakeholder zu gefährden, könnte darüber hinaus eine Natur-Szene auf den Bildschirm übertragen werden, in der keine Menschen zu sehen sind.



Abb. 2: Beispiel für eine Natur-Szene, auf der keine Menschen abgebildet werden. Eignet sich auch für ein kurzes Experiment: Entspannt Sie der Anblick dieser Natur-Szene? [Free Fotos 2016]

⁶⁶ Vgl. Friedman et al. 2008, S. 75–80.

⁶⁷ Vgl. Friedman et al. 2008, S. 75–80.

⁶⁸ Vgl. Friedman et al. 2008, S. 75–80.

Das Beispiel macht deutlich, wie die drei Untersuchungsabschnitte des Value-Sensitive-Design-Ansatzes angewendet werden können und, dass das Einbeziehen der indirekten Stakeholder und deren Interessen bedeutsame Implikationen für das Design einer Technologie haben kann. Um noch stärker zu verdeutlichen, wie auf Basis von Value Sensitive Design konkrete Design-Anforderungen für die Gestaltung einer Technologie formuliert werden können, wird im Folgenden ein zweites Beispiel skizziert.

5.2.2 Die Groupware CodeCOOP für den Wissensaustausch in Unternehmen

Die Entwicklung und Umsetzung des Groupware-Systems *CodeCOOP* begleiteten Miller et al. aus einer Value-Sensitive-Design-Perspektive. Das Groupware-System wurde mit dem Ziel entwickelt, den Austausch von Wissen unter den Mitarbeiter*innen des Software-Unternehmens *LEO-R* zu fördern. Mitarbeiter*innen sollten Informationen und Code einfach und effizient über die Groupware miteinander teilen können. Zu den Features des Groupware-Systems gehören u. a. Diskussionsforen, eine Volltext-Suche und ein »code repository«⁶⁹. Um zu gewährleisten, dass Mitarbeiter*innen das System nach seiner Einführung tatsächlich nutzen, sollten Wertkonflikte dabei von Anfang an erkannt und gelöst werden.⁷⁰

In einem ersten Schritt untersuchten Miller et al. im Rahmen der konzeptionellen Investigationen, welche Stakeholder von dem Groupware-System betroffen sind. Direkte Stakeholder sind die Mitarbeiter*innen des Software-Unternehmens *LEO-R*, die direkt mit dem System interagieren.⁷¹ Als indirekte Stakeholder identifizierten Miller et al. Manager*innen und Führungskräfte des Unternehmens. Anschließend untersuchten die Autor*innen der Studie, welche potentiellen Vorteile und Gefahren das Groupware-System für jede Stakeholder-Gruppe mit sich bringt und welche Werte dies betrifft. Auf Basis von Forschungsliteratur zu vergleichbaren Groupware-Systemen und in Zusammenarbeit mit einer Expert*innen-Gruppe aus Unternehmens-Mitarbeiter*innen stellten Miller et al. fest, dass vor allem Reputation, Zusammenarbeit, Privatsphäre, Anonymität, Vertrauen, und »awareness«⁷² im Kontext des Systems eine Rolle spielen.⁷³

Im zweiten Abschnitt, den empirischen Investigationen, führten Miller et al. eine anonyme Online-Befragung mit den Mitarbeiter*innen des Software-Unternehmens durch. Dabei wurden die Mitarbeiter*innen gefragt, wie sie die Vor- und Nachteile der Groupware und die damit verbundenen Werte einschätzen und gewichten. Darüber hinaus planten Miller et al., Manager*innen und Führungskräfte des Software-Unternehmens zu ihren Interessen und Werten zu befragen; diese sagten ihre Teilnahme an Interviews jedoch aus Zeitgründen ab.⁷⁴ Die Autor*innen der Studie bemängeln daher das unzureichende Einbeziehen der indirekten Stakeholder in ihrer Value-Sensitive-Design-Analyse.⁷⁵

In den technischen Investigationen, dem dritten Abschnitt der Value-Sensitive-Design-Analyse, übersetzten Miller et al. die Werte und Interessen der Stakeholder in Design-Anforderungen für die Groupware. Dazu nutzten sie die Methode der *Value Dams and Flows*. Die Methode sieht vor, technische Funktionen und Policies zu vermeiden, wenn sie von einem Teil der Stakeholder als sehr problematisch bewertet werden. Solche Funktionen und Policies werden als *Value Dams* bezeichnet. Value Dams werden bereits dann ausgesprochen, wenn ein geringer Anteil der Stakeholder eine bestimmte Funktion oder Policy stark ablehnt. Als *Value Flows* werden hingegen technische Funktionen und Policies bezeichnet, die ein großer Teil der Stakeholder als wünschenswert beurteilt. Sie werden in die Gestaltung des Systems aufgenommen. Zwischen den Value Dams und Flows können Konflikte entstehen, die systematisch adressiert und gelöst werden müssen.⁷⁶

Für das Groupware-System *CodeCOOP* legten Miller et al. fest, dass Features, die mehr als die Hälfte der Umfrage-Teilnehmer*innen als deutlichen Vorteil bewerteten, als Value Flows eingeordnet werden. So war mehr als 50% der System-Nutzer*innen wichtig, zu sehen, mit welcher Häufigkeit eigene Inhalte von anderen Mitarbeiter*innen genutzt werden. Ähnlich wichtig war den Nutzer*innen, dass angezeigt wird, wie andere Nutzer*innen einzelne Posts bewerten bzw. ranken. Diese Features definierten Miller et al. daher als Value Flows, die im Design der Groupware umgesetzt werden sollen, um die »awareness«⁷⁷ der System-Nutzer*innen zu unterstützen.⁷⁸

⁶⁹ Miller et al. 2007, S. 282.

⁷⁰ Vgl. Miller et al. 2007, S. 281f.

⁷¹ Vgl. Miller et al. 2007, S. 283f.

⁷² Miller et al. 2007, S. 284.

⁷³ Vgl. Miller et al. 2007, S. 284.

⁷⁴ Vgl. Miller et al. 2007, S. 284.

⁷⁵ Vgl. Miller et al. 2007, S. 290.

⁷⁶ Vgl. Miller et al. 2007, S. 284.

⁷⁷ Miller et al. 2007, S. 284.

⁷⁸ Vgl. Miller et al. 2007, S. 284f.

Zugleich sahen mehr als 11% der Teilnehmer*innen der Online-Befragung einen deutlichen Eingriff in ihre Privatsphäre, wenn geloggt wird, welche Fragen sie in den Diskussionforen posten und welche Suchanfragen sie stellen. Um die Privatsphäre der System-Nutzer*innen zu schützen, legten Miller et al. das Loggen von Fragen und Suchanfragen als Value Dams fest. Damit entsteht ein Konflikt zwischen der Förderung der »awareness« der Mitarbeiter*innen und dem Schutz ihrer Privatsphäre. Miller et al. formulierten daher die Designanforderung, dass zwar angezeigt wird, wie oft einzelne Inhalte genutzt und wie sie gerant wurden; dass dabei aber nicht sichtbar ist, wer den Inhalt gesucht, genutzt oder gerant hat.⁷⁹

Schließlich formulierten Miller et al. auch für die Unternehmens-Policies in Bezug auf das Groupware-System Value Dams und Flows, denn »groupware systems do not exist independent of complex social systems«⁸⁰. So gaben über 70% der Teilnehmer*innen der Online-Befragung an, dass sie es als deutlichen Vorteil ansehen, wenn das Beantworten von Fragen und Beitragen von Code ihre Reputation im Unternehmen stärkt. Zugleich nahmen 13% der Befragten als Gefahr war, dass sich ihre Reputation verschlechtern könnte, wenn sie inkorrekte Antworten oder fehlerhaften Code in der Groupware posten. Diesen Wertkonflikt lösten Miller et al., indem sie empfahlen, dass der Beitrag zum Groupware-System ausschließlich dann in die jährliche Evaluation der Mitarbeiter*innen einfließt, wenn die Bewertung positiv ausfällt.⁸¹

Das hier vorgestellte Groupware-System wurde zwar im Kontext eines Unternehmens entwickelt; zugleich können Informationssysteme, die den Austausch von Wissen, die Dokumentation von Code und die Zusammenarbeit verschiedener Akteur*innen fördern, auch im Kontext der Digital Humanities eine zentrale Rolle spielen.

6. Value Sensitive Design in Digital-Humanities-Projekten

In den Digital Humanities ist der Ansatz aber bisher nach Kenntnis der Autorin nicht rezipiert worden. Im Folgenden soll daher vorgestellt werden, wie Value Sensitive Design in unterschiedlichen DH-Projekten angewendet werden kann. Für jede Spielart der DH nach Roth und Burghardt werden dazu Anknüpfungspunkte für Value Sensitive Design aufgezeigt.

6.1 Pre-Existing Bias in den Digitized Humanities abmindern

Wie oben bereits ausgeführt, fasst Roth unter Digitized Humanities die Aktivitäten, die digitale Datensätze in den Geistes- und Sozialwissenschaften erstellen, kuratieren und nutzen, indem Texte, Bilder und Multimedia-Artefakte digitalisiert, verarbeitet, verknüpft, verfügbar gemacht und visualisiert werden, beispielsweise in digitalen Archiven.⁸² Diese Aktivitäten haben direkte und indirekte Auswirkungen auf verschiedene Interessensgruppen und deren Werte: So zeigen ethische und wissenschaftstheoretische Untersuchungen, »dass bereits im Prozess der Datenformatierung als auch der weiteren Verarbeitung von Daten und der aus ihnen gezogenen Schlussfolgerungen mannigfaltige Wertentscheidungen getroffen werden«.⁸³

Direkte Stakeholder von Technologien, die in den Digitized Humanities erstellt und genutzt werden, sind einerseits diejenigen, die an DH-Projekten und -Infrastrukturen mitwirken und diese gestalten, beispielsweise DH-Forscher*innen, Bibliotheksmitarbeiter*innen und IT-Dienstleister*innen. Aber auch die Nutzer*innen von DH-Infrastrukturen sind direkte Stakeholder der Technologien in den Digitized Humanities, indem sie beispielsweise aus wissenschaftlichem oder privaten Interesse die Inhalte nutzen, die digitale Archive und Bibliotheken bereitstellen. Als indirekte Stakeholder der Technologien in den Digitized Humanities können darüber hinaus all jene verstanden werden, die in den Texten, Bildern und Multimedia-Artefakten der DH-Projekte und -Infrastrukturen abgebildet werden – genauso wie jene Gruppen, die in diesen Projekten und Infrastrukturen eben nicht repräsentiert werden. Im Folgenden wird gezeigt, dass die Werte, die in den Digitized Humanities eine Rolle spielen, unter anderem Geschlechtergerechtigkeit, Barrierefreiheit, globale Gerechtigkeit im Sinne von Teilhabe und Freiheit von Diskriminierung⁸⁴ umfassen.

Der Value-Sensitive-Design-Ansatz kann in den Digitized Humanities also genutzt werden, um bei der Zusammenstellung digitaler Datensätze und beim Aufbau digitaler Archive und Infrastrukturen die Interessen aller betroffenen Stakeholder zu beachten.⁸⁵ Dadurch kann verhindert werden, dass Pre-Existing Bias reproduziert wird – oder sogar dazu beigetragen werden, diesen

⁷⁹ Vgl. Miller et al. 2007, S. 285.

⁸⁰ Miller et al. 2007, S. 286.

⁸¹ Vgl. Miller et al. 2007, S. 285f.

⁸² Vgl. Roth 2019, S. 616.

⁸³ Simon 2016, S. 363.

⁸⁴ Vgl. Vereinte Nationen Generalversammlung 2014, S. 8.

⁸⁵ Zu den direkten Stakeholder gehören dabei diejenigen, die digitale Archive und Infrastrukturen erstellen und nutzen. Indirekte Stakeholder sind hingegen diejenigen, die in den digitalen Archive und Infrastrukturen abgebildet oder repräsentiert werden (wollen).

abzumindern. Zur Verdeutlichung folgen einige Beispiele, wie Value Sensitive Design in Digitized Humanities-Projekten dazu beitragen kann, Diskriminierung auf Basis von kolonialen Machtstrukturen, binären Geschlechtszuschreibungen und körperlichen Fähigkeiten zu lindern.

6.1.1 Gender Bias ausgleichen

Digitale Datensätze, die in den DH erstellt, kuratiert und genutzt werden, entstehen meist auf Basis bereits vorhandener Sammlungen, beispielsweise von Bibliotheken und Archiven. Zentral ist daher für die ethische Gestaltung von Projekten in den Digitized Humanities, zu Beginn zu analysieren, inwiefern die Datenbasis des jeweiligen Projekts einen Bias aufweist.⁸⁶ So weisen beispielsweise die Bestände von Archiven meist einen Gender Bias auf. So konstatiert Buettner für die *Monacensia im Hildebrandhaus*, dem Literaturarchiv der Stadtbibliothek München:

»Mit einem Blick auf die Archivbestände der Monacensia stellen wir fest, dass sich das literarische Gedächtnis bislang einseitig aus Erinnerungen von Männern speist. Das künstlerische Werk von Frauen ist in der Vergangenheit landauf landab als weniger relevant eingeschätzt worden. [...] Das literarische Gedächtnis der Stadt München reproduziert so – genau wie viele Museumssammlungen, Theaterspielpläne, Verlags-, Radio- und Ausstellungsprogramm [sic!] – eine unvollständige Wirklichkeit.«⁸⁷

Bei der Erstellung digitaler Datensätze und Archive gilt es also, diesen Pre-Existing Gender Bias so gut wie möglich auszugleichen. Auf Basis von Value Sensitive Design können bei der Erstellung digitaler Datensätze auf Basis von Archiven wie der Monacensia in einem ersten Schritt – den konzeptionellen Überlegungen – die betroffenen Stakeholder und Stakeholder-Gruppen identifiziert werden. Neben den Mitarbeiter*innen des Archivs bzw. des DH-Projekts können auch die Nutzer*innen der Monacensia als direkte Stakeholder begriffen werden. Eine bedeutende Gruppe indirekter Stakeholder sind hingegen Frauen im Allgemeinen und Akteurinnen des literarischen Lebens in München im Besonderen: Sie sind in den Beständen der Monacensia unterrepräsentiert, sodass ihre Werte Teilhabe und Repräsentation gefährdet sind.

In einem zweiten Schritt, den empirischen Investigationen, müsste untersucht werden, welche Werte den verschiedenen Stakeholder-Gruppen in Bezug auf das digitale Archiv tatsächlich wichtig sind. Dazu könnten beispielsweise Akteurinnen des literarischen Lebens in München wie Autorinnen und Literaturwissenschaftlerinnen im Rahmen von qualitativen Interviews befragt werden. Auch die Werte der direkten Stakeholder-Gruppen, also der Mitarbeiter*innen des Archivs bzw. des DH-Projekts und der potentiellen Nutzer*innen des digitalen Archivs müssten im Rahmen empirischer Untersuchungen erhoben werden, sodass potentiell auftretende Wertkonflikte identifiziert werden können.

Wie können die Werte Teilhabe und Repräsentation von Frauen gefördert werden, wenn auf Basis von Archivbeständen mit Gender Bias ein digitaler Datensatz oder ein digitales Archiv erstellt wird? Im Rahmen des dritten Analyseschritts, den technischen Investigationen, werden Strategien entwickelt, um Daten von und über Frauen in ausreichender Quantität und Qualität in den Datensatz bzw. das digitale Archiv aufzunehmen – oder, sollte dies nicht möglich sein, Frauen trotz der strukturellen Unterrepräsentation im Datensatz sichtbar zu machen. Die Monacensia im Hildebrandhaus rief beispielsweise das mehrjährige partizipative Kulturerbeprojekt *#femaleheritage* ins Leben, »um die literarischen Texte und Dokumente von Frauen ins literarische Gedächtnis ein- bzw. zurückzuschreiben«.⁸⁸ Das Projekt zeigt in Blogartikeln und Sonderausstellungen Archivlücken und Forschungsdesiderate zu Frauen in München auf, erstellte ein Online-Dossier mit Dokumentarfilmen und Texten zu jüdischen Schriftstellerinnen und sammelte in einer Blogparade über 200 Beiträge in Podcasts, Blogs und der Wikipedia, in denen Frauen im Kontext von Literatur, Kultur, Politik und Gesellschaft gewürdigt werden.⁸⁹

⁸⁶ Herzlichen Dank an Swantje Dogunke, eine der Reviewerinnen dieses Aufsatzes, für diesen wichtigen Hinweis.

⁸⁷ Buettner 2020, Hervorhebung im Original.

⁸⁸ Buettner 2020.

⁸⁹ Buettner 2021.

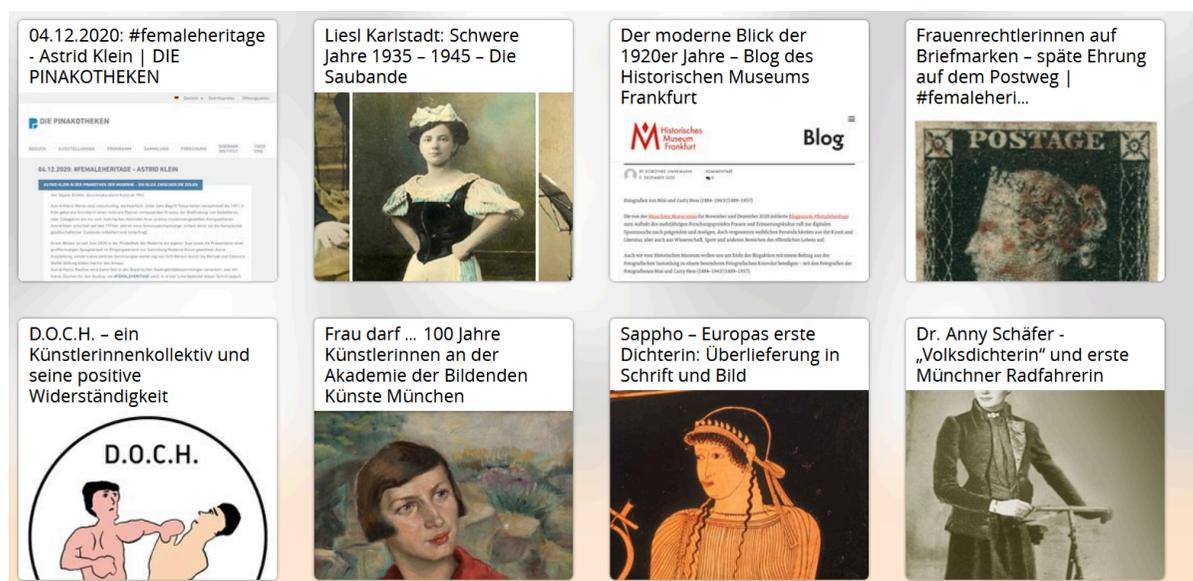


Abb. 3: Ausschnitt aus einer Zusammenstellung der knapp 200 Blogartikel, die im Kontext der #femaleheritage-Blogparade entstanden sind. [Praske 2021]

Das Beispiel zeigt, dass Value Sensitive Design dazu beitragen kann, Pre-Existing Bias in Archiv- und Datenbeständen zu ermitteln und Strategien zu entwickeln, um diesen Bias in digitalen Datensätzen und Archiven abzumindern.

6.1.2 Digitale Archive postkolonial gestalten

Darüber hinaus können digitale Archive koloniale Machtstrukturen reproduzieren. Das macht Risam am Beispiel der *Networked Infrastructure for Nineteenth-Century Electronic Scholarship* (NINES) deutlich: Das Projekt dient der Vernetzung von DH-Projekten zum 19. Jahrhundert in den USA und Großbritannien, indem es über 900.000⁹⁰ Projekte, Publikationen und digitale Sammlungen zusammenstellt, zentral durchsuchbar macht und Qualitätssicherung durch Peer Review anbietet. NINES ist damit eine bedeutende Infrastruktur in den Digital Humanities. Die Autor*innen, die in NINES zu finden sind, sind allerdings eine »narrowly prescribed group of primarily white writers«⁹¹. Texte Schwarzer und indigener Autor*innen sind hingegen stark unterrepräsentiert. Zudem sind englischsprachige Texte, die im 19. Jahrhundert in den von Großbritannien kolonialisierten Gebieten entstanden sind, nicht zu finden. Thematisiert wird dieser Mangel auf der Projektseite nicht.⁹² NINES schreibt damit »colonial dynamics of knowledge production«⁹³ fort, durch die »colonial violence«⁹⁴ in der Digital-Humanities-Forschung reproduziert wird.

Risam fordert daher, digitale Archive postkolonial zu gestalten und koloniale Formen der Wissensproduktion zu überwinden. Dafür kann Value Sensitive Design als Ausgangspunkt dienen: So sieht der Ansatz vor, bei der Identifikation der Stakeholder – also derjenigen, die direkt oder indirekt von einer Technologie betroffen sind – besonders Individuen und Gruppen zu berücksichtigen, die von Diskriminierung betroffen sind. Personen, die von rassistischer und (neo-)kolonialer Diskriminierung betroffen sind, können somit als Stakeholder von DH-Projekten wie NINES identifiziert werden: Als Gruppe, die in den Inhalten der Infrastruktur NINES unterrepräsentiert ist, sind sie indirekte Stakeholder. Als Mitarbeiter*innen oder Nutzer*innen von DH-Projekten wie NINES können sie zugleich aber auch direkte Stakeholder sein. Mit Value Sensitive Design kann im Rahmen der konzeptionellen und empirischen Untersuchungen zunächst geklärt werden, welche verschiedenen Stakeholder-Gruppen konkret von der DH-Infrastruktur betroffen sind, welche Werte die unterschiedlichen Stakeholdergruppen als wichtig erachten und welche Wertkonflikte dabei entstehen. Auf Basis dieser Analysen können die Werte von Stakeholder-Gruppen, die von rassistischer und (neo-)kolonialer Diskriminierung betroffen sind, aktiv in das Design von digitalen Archiven und Infrastrukturen einbezogen werden.

⁹⁰ Vgl. NINES (Hg.) 2021.

⁹¹ Risam 2019, S. 51.

⁹² Vgl. NINES (Hg.) 2021.

⁹³ Risam 2019, S. 51.

⁹⁴ Risam 2019, S. 51.

6.1.3 Barrierefreiheit fördern

Schließlich kann auf Basis des Value-Sensitive-Design-Ansatzes auch sichergestellt werden, dass digitale Archive und Infrastrukturen möglichst barrierefrei sind. Ein bedeutender Aspekt ist dabei die Gestaltung der Benutzeroberfläche: Diese sollte nach den Prinzipien der *Web Accessibility* so gestaltet sein, dass auch Personen mit visuellen, auditiven und physischen Einschränkungen das Online-Angebot nutzen können (ausführlicher hierzu Kapitel 6.4). Aber auch finanzielle Aspekte spielen für die barrierearme Nutzung von und die Beteiligung an digitalen Archiven und Infrastrukturen eine Rolle: So macht Rojas Castro darauf aufmerksam, dass Institutionen mit geringen finanziellen Ressourcen keine eigenen Digitalisate zu digitalen Infrastrukturen beitragen können, wenn sie ihre Dokumente dafür zwingend mit einem kostenpflichtigen Persistenten Identifikator wie *DOI* oder *Handle* versehen müssen. Dies führt wiederum zu weiteren Verzerrungen: Wenn beispielsweise ressourcenarme Institutionen im globalen Süden eigene Digitalisate in zentralen DH-Infrastrukturen aus finanziellen Gründen nicht beitragen können, entsteht ein regionaler Bias.⁹⁵

Auf Basis von Value Sensitive Design können Personen mit verschiedenen physischen Einschränkungen genauso wie Institutionen mit geringen finanziellen Ressourcen als Stakeholder von digitalen Archiven und Infrastrukturen in den Fokus genommen werden. Im Rahmen empirischer Untersuchungen kann erhoben werden, welche Interessen und Werte diese Stakeholder-Gruppen haben – beispielsweise Barrierefreiheit, Freiheit von Bias, Teilhabe und (Informations-)Gerechtigkeit – und an welchen Stellen diese möglicherweise mit den Interessen weiterer Stakeholder-Gruppen in Konflikt geraten. Darauf aufbauend können im Rahmen der technischen Investigationen Strategien entwickelt werden, um die Benutzeroberflächen und Policies der DH-Projekte so zu gestalten, dass sie auch die Interessen von Personen mit verschiedenen physischen Einschränkungen genauso wie von Institutionen mit geringen finanziellen Ressourcen berücksichtigen.

6.2 Design-Normen für KI-Anwendungen in den Numerical Humanities

Anders als die Digitized Humanities legt die Spielart der Numerical Humanities den Schwerpunkt darauf, formale Modelle zu erstellen und menschliches Verhalten mathematisch zu abstrahieren. Hier kommen vor allem KI-Anwendungen wie Machine Learning Software zum Einsatz. Machine Learning Software zeichnet sich dadurch aus, dass sie sich autonom und adaptiv weiterentwickelt, indem sie mit ihrem Umfeld interagiert. Machine Learning Systeme können daher aus verschiedenen Gründen Bias enthalten: Wenn die Trainingsdaten biased sind, wird dieser Bias unweigerlich in die Machine Learning Software übertragen. Beispiel hierfür ist die eingangs erwähnte Gesichtserkennungssoftware, die bei Schwarzen Personen eine deutlich höhere Rate an falsch positiven Treffern aufweist als bei *weißen* Personen – weil die Trainingsdaten vorrangig Bilder *weißer* Personen enthielten. Mithilfe des Value-Sensitive-Design-Ansatzes kann also sichergestellt werden, dass der Datensatz, auf dessen Basis eine Machine Learning Software lernt, keinen Bias aufweist (vgl. auch Kapitel 6.1).

Machine Learning Systeme können aber auch bei unverzerrten Trainingsdaten einen Algorithmus hervorbringen, der »might conceivably be described as following a rule that is somehow biased«. ⁹⁶ Diese Form von Bias ist »emergent und opaque«, ⁹⁷ er ist also nicht vorhersehbar und für Menschen schwer zu erkennen. Eine ethische Überprüfung von Machine Learning Technologien bringt daher besondere Herausforderungen mit sich. Umbrello und van de Poel schlagen vor, das Value-Sensitive-Design-Framework für KI-Anwendungen weiterzuentwickeln und um ein »set of AI-specific design principles« ⁹⁸ zu ergänzen. Drei Aspekte sind dabei zentral:

1) Als Design-Vorgaben, die für alle KI-Anwendungen gelten, schlagen Umbrello und van de Poel die *AI for Social Good-Prinzipien* (AI4SG) vor. Diese spezifizieren die Prinzipien der *EU High Level Expert Group on AI* und übertragen sie in sieben konkrete Anforderungen für das Design von KI-Anwendungen:

- Die Anwendung wird kontinuierlich und zunächst in kleineren Szenarien empirisch getestet;
- die Designer*innen sichern die Anwendung gegen die Manipulation der Datenbasis und die übermäßige Abhängigkeit von einzelnen Variablen ab;
- die Nutzer*innen-Autonomie wird gestärkt, indem die User*innen verschiedenen Auswahlmöglichkeiten bekommen;
- die Anwendung ist transparent und wird angepasst an die jeweilige Nutzer*innen-Gruppe erklärt;
- die Privatsphäre von »data subjects« ⁹⁹ wird geschützt;

⁹⁵ Rojas Castro 2020.

⁹⁶ Umbrello / van de Poel 2021, S. 4.

⁹⁷ Umbrello / van de Poel 2021, S. 4.

⁹⁸ Umbrello / van de Poel 2021, S. 2.

⁹⁹ Umbrello / van de Poel 2021, S. 5.

- die Datenbasis, mit der die Anwendung lernt, muss frei von Bias sein, v. a. bezüglich Kategorien wie Race, Gender und Altersklassen und
- die Zuschreibung von Bedeutung darf nicht willkürlich erfolgen und nicht zu eng gefasst werden.¹⁰⁰

2) Die AI4SG-Prinzipien zielen jedoch vorrangig darauf ab, KI-Anwendungen so zu gestalten, dass sie keinen Schaden hervorrufen. Der Value-Sensitive-Design-Ansatz sieht darüber hinaus aber auch vor, gesellschaftlich erwünschte Werte und Ziele aktiv zu fördern sowie die Werte aller beteiligten Stakeholder zu berücksichtigen. Umbrello und van de Poel schlagen daher vor, zwischen »values promoted by design and values respected by design«¹⁰¹ zu unterscheiden. Welche Werte gesellschaftlich erwünscht sind und von KI-Technologien gefördert werden sollen, kann beispielsweise anhand der *Sustainable Development Goals* (SDG) der UN abgeleitet werden, die von den Vereinten Nationen erarbeitet und ihren 193 Mitgliedsstaaten verabschiedet wurden.¹⁰² Zu den Zielen, die in den SDG festgehalten sind, gehören u. a. Geschlechtergerechtigkeit, die Bekämpfung von Armut und der Zugang zu Bildung.¹⁰³

3) Schließlich schlagen die Autoren vor, dass die Value-Sensitive-Design-Analyse den gesamten Lebenszyklus einer KI-Anwendung begleitet. Sollten im Lauf der Anwendungszeit einer Technologie Problematiken und Wertverletzungen auftauchen, kann sie so entsprechend angepasst werden. Die Erweiterung des Value-Sensitive-Design-Frameworks sieht daher vor, über den gesamten Lebenszyklus hinweg Prototypen der KI-Anwendung zu erstellen und diese im jeweiligen Anwendungsfeld zu testen. Dabei werden nicht nur die technischen Funktionen geprüft, sondern auch die Auswirkungen der KI-Anwendung auf Werte und Gesellschaft. Diese Tests sollten mit der Anwendung in kleinem Rahmen beginnen und dann sukzessive den Bezugsrahmen ausweiten.¹⁰⁴

Umbrello und van de Poel illustrieren den Ablauf einer Value-Sensitive-Design-Analyse von KI-Anwendungen am Beispiel der Corona-Datenspende-App, die im April 2020 vom Robert-Koch-Institut (RKI) vorgestellt wurde. Über die App können Bürger*innen mit Fitnessarmbändern und Smartwatches dem RKI ihre Gesundheitsdaten zur Verfügung stellen, um die Ausbreitung des COVID-19-Virus nachvollziehbar zu machen. In drei Schritten analysieren Umbrello und van de Poel die Werte, die bei der Entwicklung und dem Einsatz der Corona Datenspende-App eine Rolle spielen:

- Als gesellschaftlich erwünschten Wert, den die Anwendung fördert (»Values promoted by design«¹⁰⁵) identifizieren sie öffentliche Gesundheit. Dieser Wert lässt sich dem dritten SDG, Gesundheit und Wohlbefinden, zuordnen.
- Darüber hinaus analysieren die Autoren, welche »values respected by design«, also Design-Vorgaben, die für alle KI-Anwendungen gelten, im Kontext der Corona-Datenspende-App besondere Bedeutung einnehmen. So spielt zum Beispiel die Verhinderung von Schaden (»Nonmaleficence«¹⁰⁶) eine hervorgehobene Rolle, vor allem in Bezug auf Datensicherheit und Privatsphäre. Aber auch Erklärbarkeit und Autonomie sind Werte, die beim Design der App beachtet werden müssen, d. h. es muss eine »balance between human decision-making power and its abdication to AI systems«¹⁰⁷ gewährleistet werden, die Funktionsweise der App und des KI-Systems müssen verständlich sein und mindestens eine Person muss für die Funktionsweise verantwortlich zeichnen.
- Schließlich analysieren Umbrello und van de Poel, welche Interessen und Werte die Stakeholder der Corona-Datenspende-App haben. Dabei nehmen sie die Nutzer*innen der App als direkte Stakeholder in den Fokus. So ist beispielsweise Freiwilligkeit entscheidend, also dass das Nutzen der App nicht verpflichtend ist. Damit einher geht der Wert Fairness, der gewährleistet, dass niemand stigmatisiert oder diskriminiert wird, der*die die App nicht nutzt. Zudem könnte das Tracing der Nutzer*innen über die App eine Gefahr für die Versammlungsfreiheit darstellen, was die Autonomie der Nutzer*innen und die »Nonmaleficence« der KI-Anwendung gefährdet. Auch der Wert Gesundheit ist aus Nutzer*innen-Perspektive von besonderer Bedeutung: Wenn die App zu einem falschen Sicherheitsgefühl führt, könnten Nutzer*innen Gesundheitsrisiken eingehen.¹⁰⁸

Anschließend zeigen Umbrello und van de Poel beispielhaft, wie die verschiedenen identifizierten Werte in Design-Anforderungen übersetzt werden können. Um den Wert Verhinderung von Schaden (»Nonmaleficence«¹⁰⁹) zu unterstützen, müssen bei der Gestaltung der App vor allem der Schutz der Privatsphäre und informierte Einwilligung beachtet werden. Daraus ergibt sich erstens die Anforderung, dass die Nutzungsbedingungen der App verständlich formuliert werden müssen, und zweitens, dass die erhobenen Daten pseudonymisiert, lokal gespeichert und nach einem bestimmten Zeitraum wieder gelöscht werden müssen. Um den Wert der Erklärbarkeit zu unterstützen, muss für Nutzer*innen zudem transparent gemacht werden, welche Daten und

¹⁰⁰ Vgl. Umbrello / van de Poel 2021, S. 4–6.

¹⁰¹ Umbrello / van de Poel 2021, S. 6.

¹⁰² Vgl. Beisheim 2015.

¹⁰³ Vgl. Umbrello / van de Poel 2021, S. 6f.

¹⁰⁴ Vgl. Umbrello / van de Poel 2021, S. 6.

¹⁰⁵ Umbrello / van de Poel 2021, S. 9.

¹⁰⁶ Umbrello / van de Poel 2021, S. 10.

¹⁰⁷ Umbrello / van de Poel 2021, S. 9.

¹⁰⁸ Vgl. Umbrello / van de Poel 2021, S. 8-10.

¹⁰⁹ Umbrello / van de Poel 2021, S. 10.

Datensets gesammelt und erstellt werden, wie diese verwendet, gespeichert und gelöscht werden.¹¹⁰ Abschließend empfehlen die Autoren die Erstellung eines Prototyps der Corona-Datenspende-App. Dieser soll zunächst mit einer limitierten Anzahl an Nutzer*innen getestet werden. Dabei sollen nicht nur die technischen Funktionen, sondern auch mögliche Auswirkungen auf das Verhalten der Nutzer*innen, die Gesellschaft und die Werte der betroffenen Stakeholder überprüft und das Design der App gegebenenfalls nachgebessert werden.¹¹¹

Das erweiterte Framework von Value Sensitive Design für KI-Anwendungen kann auch in den Numerical Humanities genutzt werden, um selbstlernende Systeme wertebasiert zu gestalten oder bereits existierende KI-Anwendungen ethisch zu überprüfen. Zentral sind dabei neben den Werten der betroffenen Stakeholder auch die »values promoted by design and values respected by design«¹¹², also einerseits die in den *Sustainable Development Goals* formulierten Werte und Ziele und andererseits die Werte, die den *AI for Social Good-Prinzipien* (AI4SG) zugrunde liegen, also Freiheit von Bias, Autonomie, Privatsphäre und Transparenz.

6.3 Akteur*innen und ihre Interessen in den Humanities of the Digital analysieren

Als dritte Spielart der Digital Humanities beschreibt Roth die Humanities of the Digital: Die geisteswissenschaftliche Forschung zu »computermediated interactions and societies«,¹¹³ also zu soziotechnischen Systemen wie Online-Communities, Blogs, Sozialen-Netzwerk-Seiten und Gaming-Plattformen. Hier geht es also weniger darum, Technologien zu entwickeln und anzuwenden, sondern diese als Untersuchungsgegenstände in den Blick zu nehmen. Geschlechtergerechtigkeit sowie globale Gerechtigkeit im Sinne von Teilhabe und Freiheit von Diskriminierung seien hier nur beispielhaft als Werte genannt, die dabei fokussiert werden können.

Value Sensitive Design kann dabei als theoretischer Ausgangspunkt dienen, um Online-Communities, Soziale-Netzwerk-Seiten und Gaming-Plattformen aus einer ethischen Perspektive zu untersuchen und sich unterschiedlichen Fragestellungen zu nähern: Welche direkten und indirekten Stakeholder hat eine Soziale-Netzwerk-Seite wie *Instagram*, eine Online-Community wie *Wikipedia* oder eine Gaming-Plattform wie *Steam*? Welche Werte und wessen Interessen werden von den Nutzungsbedingungen, Policies und dem technischen Design gefördert bzw. behindert? Wie könnte die jeweilige Plattform oder Community gestaltet sein, um die Werte und Interessen aller Stakeholder zu beachten?

Dass Online-Communities, Soziale-Netzwerk-Seiten und Gaming-Plattformen die Interessen und Werte verschiedener Stakeholder in unterschiedlichem Maß fördern oder behindern, wird anhand von Beispielen deutlich. So weist die Community-betriebene Online-Datenbank *Wikidata* einen Gender Bias auf: Von den 8,2 Millionen Personen, die Wikidata verzeichnet, sind nur 23 Prozent Frauen.¹¹⁴ Die mit Wikidata verknüpfte Enzyklopädie *Wikipedia* enthält zudem einen »Eurocentric bias«,¹¹⁵ da historische Artikel in mehreren Sprachversionen der *Wikipedia* vor allem die Geschichte europäischer Länder thematisieren.¹¹⁶ Die Nutzungsbedingungen der Sozialen-Netzwerk-Seite *Facebook* erlauben Bilder nackter Oberkörper, die als männlich gelesen werden, verbieten aber Bilder nackter Oberkörper, die als weiblich gelesen werden. West beschreibt dies als »gendered policies on images of female nudity«¹¹⁷. Eine umfassende ethische Analyse der jeweiligen Communities und Sozialen-Netzwerk-Seiten, in der alle direkten und indirekten Stakeholder und ihre Werte berücksichtigt werden, steht jedoch noch aus.

Wie könnte die Analyse einer Online-Community wie *Wikipedia* anhand von Value Sensitive Design beispielsweise aussehen? Im Rahmen der konzeptionellen Untersuchungen werden zunächst der Kontext der Online-Community beschrieben und die betroffenen Stakeholder ausgemacht. Die Online-Enzyklopädie *Wikipedia* zeichnet sich dadurch aus, dass sie alle Inhalte kostenfrei zur Verfügung stellt und es allen Nutzer*innen ermöglicht, die Inhalte des Lexikons zu bearbeiten.¹¹⁸ Direkte Stakeholder sind dabei diejenigen, die selbst redaktionell an der *Wikipedia* mitarbeiten, genauso wie alle, die Inhalte der Online-Enzyklopädie rezipieren. Auch die *Wikimedia Foundation*, die als Non-Profit-Organisation die *Wikipedia* trägt und finanziert,¹¹⁹ kann als direkter Stakeholder verstanden werden. Indirekte Stakeholder sind aber auch Individuen und Gruppen, die in den Artikeln und Inhalten der *Wikipedia* abgebildet werden, genau wie jene, die nicht repräsentiert werden. Welche Werte diese unterschiedlichen Stakeholder-Gruppen priorisieren und wo potentiell Wertkonflikte auftreten, muss im Rahmen von empirischen Investigationen untersucht werden, beispielsweise in Online-Befragungen oder qualitativen Interviews. Im dritten

¹¹⁰ Vgl. Umbrello / van de Poel 2021, S. 10-11.

¹¹¹ Vgl. Umbrello / van de Poel 2021, S. 12.

¹¹² Vgl. Umbrello / van de Poel 2021, S. 6.

¹¹³ Roth 2019, S. 623.

¹¹⁴ Vgl. Denezh (Hg.) 2020.

¹¹⁵ Samoilenko 2017.

¹¹⁶ Samoilenko 2017.

¹¹⁷ West 2017, S. 31.

¹¹⁸ Vgl. Greb 2021.

¹¹⁹ Vgl. Liao 2015, S. 38.

Schritt der Analyse, den technischen Investigationen, wird analysiert, welche technischen Features und organisatorischen Strukturen der Online-Community Wikipedia die Interessen und Werte der verschiedenen Stakeholder-Gruppen fördern oder behindern. Eine Value-Sensitive-Design-Analyse ermöglicht damit ein Verständnis von Problematiken wie dem Eurocentric oder Gender Bias in der Online-Community Wikipedia, das neben verschiedenen Stakeholder-Interessen auch technische und organisatorische Aspekte einbezieht.

6.4 Mehr Barrierefreiheit in den Public Humanities schaffen

Den drei von Roth beschriebenen Teilbereichen der Digital Humanities fügt Burghardt eine vierte Spielart hinzu: Die *Public Humanities*, also die Themenfelder des digitalen Publizierens, der Wissenschaftskommunikation und des E-Learnings. Auch hier gilt es, die Interessen unterschiedlicher Stakeholder zu berücksichtigen, um Bias und Diskriminierung zu vermeiden. Auf Basis von Value Sensitive Design kann beispielsweise sichergestellt werden, dass digitale Publikationen und E-Learning-Angebote möglichst barrierefrei gestaltet werden. Barrierefreiheit bezieht sich dabei auf Barrieren physischer Art, aber auch finanzieller und sprachlicher Art. Damit spielen in den Public Humanities neben Barrierefreiheit auch die in der Informationsethik zentralen Werte Informationsfreiheit und Informationsgerechtigkeit eine Rolle.¹²⁰

Genau wie in den Digitized Humanities sollten Benutzeroberflächen und Formate von digitalen Publikationen und E-Learning-Angeboten nach den Prinzipien der Web Accessibility gestaltet sein, sodass sie von Personen mit auditiven, visuellen, körperlichen, kognitiven und sprachlichen Einschränkungen genutzt werden können. Orientierung bieten dafür die Standards und Empfehlungen der *W3C Web Accessibility Initiative*,¹²¹ die beispielsweise vorsehen, Bilder mit einem Alternativtext zu versehen, Audio-Inhalte und Videos mit Transkriptionen bzw. Untertiteln auszustatten und die Navigation auf Websites per Tastatur zu ermöglichen. Stakeholder, deren Interessen damit verfolgt werden, sind dabei neben Menschen mit Behinderung auch weitere Gruppen: Personen, deren Fähigkeiten sich altersbedingt verändern; Nutzer*innen, die über eine langsame Internetverbindung verfügen; und Personen, die sich in limitierenden Situationen befinden, beispielsweise in einer Umgebung, in der sie Audio-Inhalte nicht anhören können oder hellem Sonnenlicht ausgesetzt sind. Die W3C Web Accessibility Initiative empfiehlt, Aspekte der Barrierefreiheit bereits zu Beginn eines Projektes zu bedenken, um zu vermeiden, Designs und Inhalte noch einmal neu gestalten zu müssen.¹²²

Neben Empfehlungen für die Gestaltung und Entwicklung von Websites und Online-Publikationen weist die W3C Web Accessibility Initiative auch darauf hin, dass Texte klar, prägnant und in verständlicher Sprache formuliert werden sollen.¹²³ Auch Sprache kann also eine Barriere sein, die es Stakeholdern schwierig bis unmöglich macht, eine Ressource zu nutzen. Dabei spielen nicht nur die Verständlichkeit von Texten, sondern auch Fremdsprachenkenntnisse eine Rolle. Ein Beispiel: Die Website *The Programming Historian* publiziert Tutorials, »that help humanists learn a wide range of digital tools, techniques, and workflows«¹²⁴. Die englischsprachige Version der Website ist im *Directory of Open Access Journals* (DOAJ) verzeichnet. Das DOAJ lehnte es zunächst jedoch ab, auch die spanisch- und französischsprachigen Versionen der Tutorials in ihr Verzeichnis aufzunehmen, da diese als nicht »original enough«¹²⁵ angesehen wurden. Für Personen, die über geringe Englischkenntnisse verfügen und daher gezielt nach französisch- oder spanischsprachigen Ressourcen suchen, waren die entsprechenden Tutorials von The Programming Historian über das DOAJ damit nicht auffindbar. Seit April 2021 verzeichnet das DOAJ aber alle drei Sprachversionen der Website und trägt damit zur barrierefreien Auffindbarkeit der Tutorials von The Programming Historian für Personen ohne oder mit geringen Englischkenntnissen bei.¹²⁶

Schließlich können auch finanzielle Aspekte ausschlaggebend dafür sein, dass digitale Publikationen oder E-Learning-Ressourcen von bestimmten Individuen oder Gruppen nicht genutzt werden können. Der Zugang zu wissenschaftlichen Zeitschriften, Monographien und Sammelbänden ist sowohl für Einzelpersonen als auch für Institutionen in der Regel mit Kosten verbunden – mit Ausnahme von Publikationen, die *Open Access* zur Verfügung gestellt werden. Aber auch hier können finanzielle Barrieren entstehen, und zwar für die Autor*innen, die eigene Beiträge veröffentlichen wollen: Sogenannte *Gold-Open-Access-Zeitschriften*, die ihre Beiträge direkt bei der Veröffentlichung für Leser*innen kostenfrei zur Verfügung stellen, erheben teilweise *Article Processing Charges* (APCs) – also Gebühren, die Autor*innen für die Veröffentlichung ihres Beitrages entrichten müssen.¹²⁷ Für

¹²⁰ Vgl. Rösch 2021, S. 74, 87.

¹²¹ Vgl. W3C Web Accessibility Initiative, Henry 2021.

¹²² Vgl. W3C Web Accessibility Initiative, Henry 2019.

¹²³ Vgl. W3C Web Accessibility Initiative, White 2020.

¹²⁴ The Programming Historian (Hg.) 2008-2021.

¹²⁵ Rojas Castro 2020.

¹²⁶ Vgl. Directory of Open Access Journals (Hg.) 2021.

¹²⁷ Vgl. Frick / Kaier 2020.

Wissenschaftler*innen, die an weniger ressourcenstarken Institutionen affiliert sind, hat dies zur Folge, dass sie ihre Beiträge nur in einem Teil der entsprechenden Journals publizieren können, die in ihrem jeweiligen Fachbereich eine Rolle spielen – nämlich denjenigen ohne APCs. Damit sind ihre Möglichkeiten, am wissenschaftlichen Diskurs teilzunehmen, eingeschränkt.

Wie also können digitale Publikationen auf Basis von Value Sensitive Design wertebasiert gestaltet werden? Zunächst wird im Rahmen der konzeptionellen Überlegungen untersucht, welche Stakeholdergruppen im Kontext digitaler Publikationen eine Rolle spielen. Zu den direkten Stakeholdern gehören neben den Autor*innen und Leser*innen wissenschaftlicher Texte auch die Organisationen, die digitale Publikationen bereitstellen, und ihre Mitarbeiter*innen. Diese Stakeholder-Gruppen müssen in weitere Untergruppen ausdifferenziert werden, deren Interessen bei der Gestaltung von digitalen Publikationen und ihren Policies besonderer Berücksichtigung bedürfen. Zu Stakeholdergruppen mit spezifischen Interessen gehören beispielsweise Wissenschaftler*innen, die an ressourcenarmen Institutionen affiliert sind, aber auch Personen mit geringen Englischkenntnissen und Nutzer*innen, die über eine langsame Internetverbindung verfügen sowie Menschen mit auditiven, visuellen oder körperlichen Einschränkungen.

In einem zweiten Schritt, den empirischen Untersuchungen, muss erhoben werden, welche Interessen und Werte die verschiedenen Stakeholder-Gruppen als wichtig bewerten, beispielsweise im Rahmen von Interviews oder Online-Befragungen. Vor allem in Bezug auf Open Access ist dabei wahrscheinlich, dass Werte wie Informationsfreiheit und Informationsgerechtigkeit, die Autor*innen und Leser*innen wichtig sind, in Konflikt mit den wirtschaftlichen Interessen der Organisationen stehen, die digitale Publikationen bereitstellen, vor allem wenn diese in privatwirtschaftlichen Strukturen organisiert sind.¹²⁸ Dieser Interessenskonflikt kann zugunsten der Autor*innen, Leser*innen und Wissenschaftler*innen, die an ressourcenarmen Institutionen affiliert sind, gelöst werden, indem Fachcommunities eigene Publikationsorgane anbieten, die nicht privatwirtschaftlich, sondern community-basiert organisiert sind. Ein Beispiel ist die *Zeitschrift für digitale Geisteswissenschaften (ZfdG)*, die vom *Forschungsverbund Marbach Weimar Wolfenbüttel* und dem *DHd-Verband (Digital Humanities im deutschsprachigen Raum)* herausgegeben wird.¹²⁹ Alle Artikel der Zeitschrift erscheinen Open Access, ohne dass Autor*innen für die Veröffentlichung ihrer Texte Gebühren entrichten müssen.¹³⁰

Um darüber hinaus auch die Interessen von Stakeholder-Gruppen wie Personen mit geringen Englischkenntnissen, Nutzer*innen, die über eine langsame Internetverbindung verfügen und Menschen mit auditiven, visuellen oder körperlichen Einschränkungen in das Design von digitalen Publikationen einzubeziehen, sind weitere empirische und technische Investigationen nötig, die hier nicht ausgeführt werden können. Deutlich wurde aber, dass der Value-Sensitive-Design-Ansatz auch in den Public Humanities dazu beitragen kann die Policies und das technische Design von digitalen Publikationen und E-Learning-Ressourcen so zu gestalten, dass Barrieren finanzieller, sprachlicher und physischer Art abgebaut werden, indem die Interessen und Werte unterschiedlicher Stakeholder berücksichtigt werden.

6.5 Menschenrechte und Umweltschutz in allen DH-Spielarten fördern

In den bisherigen Abschnitten wurde gezeigt, wie Value Sensitive Design in den vier von Roth und Burghardt benannten Spielarten der DH angewendet werden kann, um Bias und Diskriminierung zu vermeiden. Abschließend soll hier auf eine Problematik eingegangen werden, die alle DH-Projekte unabhängig von ihrer Spielart aufwerfen. Noble ruft Akteur*innen in den DH dazu auf, sich die »materiality of the digital humanities and its impact in the world«¹³¹ stärker bewusst zu machen. So führt der Abbau von Mineralien, die für die Herstellung digitaler Technologien benötigt werden, in Regionen des globalen Südens zu Konflikten und Ausbeutung, während Nutzer*innen im globalen Norden die Profiteur*innen dieser Technologien sind. Zudem entstehen sowohl bei der Produktion, als auch bei der meist unsachgemäßen Entsorgung von Hardware Umweltschäden. Noble fordert daher: Die DH-Community sollte nicht nur reflektieren, wie sie zu einer gerechteren Verteilung von IT-Infrastrukturen weltweit beitragen kann, sondern auch »design against – or outright resist – the exploitive forms of labor and hazardous environmental practices in which information and communication technologies are implicated«¹³². Auch das Manifest *Digital Humanities and the Climate Crisis*, das im Sommer 2021 von einem internationalen Zusammenschluss DH-Forschender veröffentlicht wurde, ruft Akteur*innen in den DH dazu auf, die ökologischen und sozialen Auswirkungen ihres Handelns zu prüfen.¹³³

¹²⁸ Vgl. Rösch 2021, S. 147.

¹²⁹ Vgl. ZfdG 2021a.

¹³⁰ Vgl. ZfdG 2021b.

¹³¹ Noble 2019, S. 32.

¹³² Noble 2019, S. 30.

¹³³ Baillot 2021.

Auch hier kann Value Sensitive Design als Ausgangspunkt dienen, um Werte wie Umweltschutz und ökologische Nachhaltigkeit, globale Gerechtigkeit und Gesundheit beim Design von Projekten und Technologien in den DH mitzudenken. So können die Personen, die am Herstellungsprozess von Hardware und IT-Technologien beteiligt sind, genauso als indirekte Stakeholder eines jeden DH-Projektes begriffen werden wie diejenigen, die von den Umweltschäden durch unsachgemäße Hardware-Entsorgung betroffen sind. Wie können die Werte und Interessen dieser Stakeholder also im Design berücksichtigt werden? Dies gilt es, mit empirischen Untersuchungen und technischen Investigationen zu erforschen. Hier zwei Ideen: DH-Projekte könnten sich grundsätzlich zum Ziel setzen, Software und Infrastrukturen so zu entwickeln, dass sie möglichst wenig (Hardware-)Ressourcen in Anspruch nehmen, wie es zum Beispiel die 2014 gegründete Arbeitsgruppe *Minimal Computing* der *Global Outlook DH* propagiert.¹³⁴ Bei der Beschaffung von Hardware sollte zudem darauf geachtet werden, dass diese sozial- und umweltverträglich hergestellt wurde. Orientierung bieten dabei Label wie *TCO Certified* für soziale Kriterien oder der *Blaue Engel*, das *EU-Ecolabel* und das *Nordic Ecolabel* für ökologische Aspekte.¹³⁵ Wie öffentliche Einrichtungen soziale und ökologische Kriterien in Vergabeverfahren verankern können, verdeutlichen u. a. die *Fachkonferenzen für sozial verantwortliche IT-Beschaffung* und die Nichtregierungsorganisation *Electronics Watch* auf ihren Websites.¹³⁶ Darüber hinaus könnten Menschenrechts- und Umweltschutz-Aspekte nicht nur bei der Beschaffung von Hardware, sondern auch bei der Organisation von Konferenzen und Dienstreisen im Rahmen von DH-Projekten handlungsleitend sein, beispielsweise indem bei Tagungen fair gehandelte und ökologisch erzeugte Verpflegung angeboten und auf Flugreisen so weit wie möglich verzichtet wird.

7. Zusammenfassung, Limitationen und Ausblick

Dieser Beitrag zeigt auf, dass Technologien nicht neutral sind, sondern immer bestimmte Werte fördern oder behindern. Zugleich werden in jedem Digital-Humanities-Projekt Technologien angewendet oder entwickelt. Um zu verhindern, dass diese Technologien Bias reproduzieren oder hervorbringen, schlägt dieser Beitrag die Anwendung von Value Sensitive Design vor: Dieses theoretische und methodische Framework bietet eine dreiteilige Vorgehensweise, um die Werte aller betroffenen Stakeholder einer Technologie im gesamten Design-Prozess zu berücksichtigen. Anhand der vier von Roth und Burghardt geprägten Spielarten von DH wurden verschiedene Anknüpfungspunkte und Anwendungsbeispiele aufgezeigt, wie Value Sensitive Design in DH-Projekten gewinnbringend eingesetzt werden kann.

Dabei unterliegen sowohl dieser Beitrag, als auch der Value-Sensitive-Design-Ansatz verschiedenen Limitationen. So können Technologien, Datensätze und Forschungsprojekte auch Verzerrungen aufweisen, die keine systematische und unfaire Diskriminierung hervorrufen. Solche Verzerrungen werden zwar nicht als Bias im Sinne von Nissenbaum und Friedman verstanden,¹³⁷ sind aber trotzdem problematisch, da sie zu verzerrten Forschungsergebnissen führen. Um diese Art von Bias zu verhindern, ist Value Sensitive Design weniger geeignet; mögliche Strategien zur Vermeidung solcher Verzerrungen sind vielmehr Methodenvielfalt, Triangulation und unterschiedliche experimentelle Settings.

Darüber hinaus ließen sich sicherlich noch weitere Aspekte und Beispiele finden, wie Value Sensitive Design in DH-Projekten genutzt werden kann. Gleichmaßen gibt es wie in Kapitel 4 beschrieben neben Value Sensitive Design weitere Methoden und Frameworks, die die ethische Gestaltung von Projekten und Technologien unterstützen, wie *Embedded Ethics* oder *Partizipatives Design*. In beiden Ansätzen hängt die ethische Bewertung und Gestaltung aber stärker von wenigen, ausgewählten Akteur*innen ab: im Falle von *Embedded Ethics* von den Werten und Perspektiven der Ethiker*innen im jeweiligen Projekt; bei *Partizipativem Design* von den Werten und Perspektiven ausgewählter Nutzer*innen. Aus meiner Sicht hat Value Sensitive Design demgegenüber den entscheidenden Vorteil, die Perspektiven, Werte und Ziele verschiedener direkt und indirekt beteiligter Akteur*innen einzubeziehen und diese systematisch zu analysieren. Dieser Beitrag erhebt aber keinerlei Anspruch auf Vollständigkeit möglicher ethischer Ansätze, sondern zeigt Ideen und Anregungen auf, wie DH-Projekte in der Praxis ethisch überprüft und wertebasiert gestaltet werden können. Eine Diskussion und Ergänzung der hier vorgestellten Aspekte sind daher wünschenswert.

Um ein DH-Projekt auf Basis von Value Sensitive Design zu gestalten, sind je nach Projekt und Kontext zusätzliche zeitliche, finanzielle und personelle Ressourcen notwendig. Diese müssen bei der Planung von Projekten und der Beantragung von Mitteln eingeplant werden. Darüber hinaus muss entweder innerhalb des Projektes oder an zentraler Stelle Expertise zur theoretischen und praktischen Umsetzung von Value Sensitive Design vorhanden sein. Drei Ideen, um dies für jedes DH-Projekt zu gewährleisten: In den Curricula von DH-Studiengängen könnten Technik- und Informationsethik sowie konkrete methodische Ansätze wie Value Sensitive Design verpflichtend verankert werden, sodass möglichst viele zukünftige Akteur*innen in den DH die

¹³⁴ Vgl. *Minimal Computing* (Hg.).

¹³⁵ Vgl. Rat für Nachhaltige Entwicklung (Hg.) 2019.

¹³⁶ Vgl. *Electronics Watch* (Hg.) [2020]; *Konferenzen für sozial verantwortliche IT-Beschaffung* (Hg.) 2020.

¹³⁷ Vgl. Friedman / Nissenbaum 1996, S. 332.

Grundlagen wertebasierten Projekt- und Technologie-Designs kennen. Auch die Idee der *Data Stewards*, also von Expert*innen, die Forschende beim Management von Forschungsdaten unterstützen,¹³⁸ kann als Anregung dienen: So sind parallel dazu *Value Stewards* oder *Ethic Stewards* denkbar, die DH-Forschende bei der ethischen Überprüfung und wertebasierten Gestaltung ihrer Projekte und Technologien beraten. Ähnlich wie in den Bereichen Forschungsdatenmanagement oder Open Access könnten auch Forschungsförderinstitutionen eine entscheidende Rolle einnehmen: In ihren Förderrichtlinien könnten sie von Forscher*innen fordern, eine Strategie zu entwickeln, um ihr Projekt ethisch zu reflektieren und zu gestalten – und diese Strategie bereits in ihrem Förderantrag zu skizzieren.

Darüber hinaus ist es auch mithilfe von Value Sensitive Design nicht möglich, Bias vollständig auszuschließen. Wie genau eine Value-Sensitive-Design-Analyse durchgeführt wird, welche Stakeholder einbezogen, welche empirischen Erhebungen durchgeführt und wie darauf aufbauend das technische Design und Policies konkret umgesetzt werden, unterscheidet sich mit großer Wahrscheinlichkeit von Forscher*in zu Forscher*in: Eigene Werte, Erfahrungen und Interessen fließen immer bewusst oder unbewusst in die Anwendung des Value-Sensitive-Design-Frameworks ein. Dies ist nicht nur unvermeidlich, sondern auch in Ordnung: Friedman und Hendry resümieren, »we need not require perfection, but commitment to practice – and through practice, progress«. ¹³⁹ Value Sensitive Design garantiert also keine »ethisch perfekten« Technologien – aber das Framework bietet Unterstützung dabei, Bias so gut wie möglich zu vermeiden und die Auswirkungen von Technologien auf die Gesellschaft zu bedenken, offenzulegen und zu diskutieren. Zudem ermöglicht Value Sensitive Design, Technologien nicht nur nach klassischen Kriterien wie Effektivität oder Robustheit zu bewerten – sondern auch nach den Werten, die sie fördern. Das ist ein guter Start, um Technologien und Projekte so zu gestalten, dass sie nicht nur die Digital Humanities, sondern auch die Welt ein wenig besser machen.

¹³⁸ Vgl. Fdm:nrw (Hg.) 2020.

¹³⁹ Friedman / Hendry 2019, S. 180.

Bibliografische Angaben

- Anne Baillot: »Digital Humanities und die Klimakrise« – ein Manifest. In: Digital Intellectuals. Hg. von Anne Baillot. Blogbeitrag vom 24.08.2021. [\[online\]](#).
- Marianne Beisheim: Die »Agenda 2030 für nachhaltige Entwicklung«. In: Zeitschrift für die Vereinten Nationen und ihre Sonderorganisationen 63 (2015), H. 6, S. 255–260. [\[online\]](#) [\[Nachweis im GVK\]](#)
- Anke Buettner: Gedächtnis mit Schlagseite – die Monacensia und #femaleheritage. In: Münchner Stadtbibliothek. Blog. Blogbeitrag vom 11.11.2020. [\[online\]](#)
- Anke Buettner: Kulturerbe-Projekt #femaleheritage: Vom Spuren legen und Fahrten lesen – Ausblick & Nachlese der Blogparade. In: Münchner Stadtbibliothek. Blog. Blogbeitrag vom 08.03.2021. [\[online\]](#)
- John T. F. Burgess / Emily Knox / Robert Hauptman: Foundations of information ethics. Chicago 2019. [\[Nachweis im GVK\]](#)
- Manuel Burghard: Theorie und Digital Humanities – Eine Bestandsaufnahme. In: Digital Humanities Theorie. Hg. von AG Digital Humanities Theorie. Blogbeitrag vom 15.05.2020. [\[online\]](#)
- Gender Gap in Wikidata. Hg. von Denelezh. In: denelezh.wmcloud.org/gender-gap. Analyse-Tool abgerufen am 21.09.2020. [\[online\]](#)
- Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Kodex. Hg. von Deutsche Forschungsgemeinschaft. Bonn 2019. In: zenodo.org/record/3923602. Version 1.0 vom 15.09.2019. DOI: [10.5281/zenodo.3923602](#)
- Directory of Open Access Journals. Hg. von DOAJ. [\[online\]](#)
- Werkzeugsammlung Öffentliche Beschaffung. Hg. von Electronics Watch. In: electronicswatch.org. Amsterdam. [\[online\]](#)
- Waltraud Ernst: Feministische Technikphilosophie. In: Handbuch Technikethik. Hg. von Armin Grunwald und Rafaela Hillebrand. Stuttgart 2021, S. 114–118. [\[Nachweis im GVK\]](#)
- Dokumentation vom 04.09.2020. Hg. von Fachkonferenzen für sozial verantwortliche IT-Beschaffung. In: faire-beschaffung.de. Wustermark 2020. [\[online\]](#)
- Claudia Frick / Christian Kaier: Publikationskosten für Zeitschriftenartikel abseits von Open-Access-Publikationsfonds – Lost in Transformation? In: o-bib. Das offene Bibliotheksjournal 7 (2020), H. 2, S. 1–15. Artikel vom 05.05.2020. DOI: [10.5282/O-BIB/5586](#)
- Batya Friedman: Value-sensitive design. In: Interactions 3 (1996), H. 6, S. 16–23. [\[Nachweis im GVK\]](#)
- Batya Friedman / David Hendry: Value Sensitive Design. Shaping technology with moral imagination. Cambridge MA u. a. 2019. [\[Nachweis im GVK\]](#)
- Batya Friedman / Peter Kahn / Alan Borning: Value Sensitive Design and Information Systems. In: The handbook of information and computer ethics. Hg. von Kenneth Einar Himma / Herman T. Tavani. Hoboken, NJ 2008, S. 69–101. [\[Nachweis im GVK\]](#)
- Batya Friedman / Helen Nissenbaum: Bias in computer systems. In: ACM Transactions on Information Systems 14 (1996), H. 3, S. 330–347. PDF. [\[online\]](#) [\[Nachweis im GVK\]](#)
- Marisa Gierlinger: Warum Tech-Konzerne der Gesichtserkennung abschwören. In: Süddeutsche Zeitung. Artikel vom 12.06.2020. [\[online\]](#)
- Minimal Computing: About. Hg. von GO::DH Minimal Computing Working Group. In: go-dh.github.io. [\[online\]](#)
- Verena Greb: 20 Jahre Wikipedia. Acht Dinge, die es zu wissen lohnt. In: Deutsche Welle vom 15.01.2021. [\[online\]](#)
- Handbuch Technikethik. Hg. v. Armin Grunwald und Rafaela Hillerbrand. 2. Auflage. Stuttgart 2021.
- Roger Häußling: Techniksoziologie. Eine Einführung. 2., überarbeitete und aktualisierte Auflage. Opladen u. a. 2019. [\[Nachweis im GVK\]](#)
- Shawn Lawton Henry: Introduction to Web Accessibility. Hg. von W3C Web Accessibility Initiative. In: w3.org/WAI/standards-guidelines vom 11.07.2019. Website. [\[online\]](#)
- Shawn Lawton Henry: W3C Accessibility Standards Overview. Hg. von W3C Web Accessibility Initiative. In: w3.org/WAI/standards-guidelines vom 30.04.2021. [\[online\]](#)
- Handbuch Medien- und Informationsethik. Hg. v. Jessica Heesen, Stuttgart 2016. [\[Nachweis im GVK\]](#)
- Martin Holland: Amazon: KI zur Bewerbungsprüfung benachteiligte Frauen. In: Heise Online. Artikel vom 11.10.2018. [\[online\]](#)
- Martin Holland: Analyse: KI-Sprachmodell GPT-3 hegt tief verankerte Vorurteile gegen Muslime. In: Heise Online vom 25.01.2021. [\[online\]](#)
- Digital Humanities: Eine Einführung. Hg. von Fotis Jannidis / Hubertus Kohle / Malte Rehbein. Stuttgart 2017 [\[Nachweis im GVK\]](#)
- Peter H. Kahn: The human relationship with nature. Development and culture. Cambridge, MA. 1999. [\[Nachweis im GVK\]](#)
- Daniél Kretschmar: Rassismus in Algorithmen. Gesichtserkennung für die Tonne. In: taz.de. Artikel vom 01.07.2020. [\[online\]](#)
- Data Stewardship. Hg. von Landesinitiative für Forschungsdatenmanagement. In: fdm.nrw. Duisburg-Essen 2020. [\[online\]](#)
- Han-Teng Liao: Harnessing the power of collaborative filtering. In: China Perspectives (2015), H. 4, S. 35–49. [\[Nachweis im GVK\]](#)
- Stuart McLennan / Amelia Fiske / Leo Anthony Celi / Ruth Müller / Jan Harder / Konstantin Ritt / Sami Haddadin / Alena Buyx: An embedded ethics approach for AI development. In: Nature Machine Intelligence 2 (2020), H. 9, S. 488–490. [\[Nachweis im GVK\]](#)
- Jessica K. Miller, Batya Friedman, Gavin Jancke: Value tensions in design. The Value Sensitive Design, development and appropriation of a corporation's groupware system, in: Proceedings of the 2007 International ACM Conference on Supporting Group Work, S. 281–290. DOI: [10.1145/1316624.1316668](#)
- What is NINES? Hg. von NINES Nineteenth-century Scholarship Online. In: nines.org/about. [\[online\]](#)
- Safiya Umoja Noble: Toward a Critical Black Digital Humanities. In: Debates in the Digital Humanities. Hg. von Matthew K. Gold, Lauren F. Klein. Minneapolis, MN 2019, S. 27–35. [\[Nachweis im GVK\]](#)
- Ibo van de Poel: Werthaltigkeit der Technik. In: Handbuch Technikethik. Hg. von Armin Grunwald und Rafaela Hillebrand. Stuttgart 2021, S. 132–136. [\[Nachweis im GVK\]](#)
- Tanja Praske: Blogparade »Frauen und Erinnerungskultur #femaleheritage« der Monacensia (11.11.–09.12) mit Nachklapp. In: pearltrees.com. Inhalt vom 16.02.2021. [\[online\]](#)
- The Programming Historian. Hg. von The Programming Historian. 2008–2021. [\[online\]](#)
- Wie die Beschaffung fairer IT-Hardware gelingt. Hg. von Rat für Nachhaltige Entwicklung. In: nachhaltigkeitsrat.de/aktuelles. Artikel vom 25.06.2019. [\[online\]](#)
- Roopika Risam: New digital worlds. Postcolonial digital humanities in theory, praxis, and pedagogy. Evanston, IL 2019. [\[Nachweis im GVK\]](#)
- Hermann Rösch: Informationsethik und Bibliotheksethik. Grundlagen und Praxis. Berlin, Boston 2021. [\[Nachweis im GVK\]](#)
- Antonio Rojas Castro: FAIR enough? Building DH resources in an unequal world. (Digital Humanities Kolloquium, BBAW, 07.08.2020). In: vimeo.com/445147368. Video. [\[online\]](#)
- Camille Roth: Digital, digitized, and numerical humanities. In: Digital Scholarship in the Humanities 34 (2019), H. 3, S. 616–632. Artikel vom 05.11.2018. DOI: [10.1093/llc/fqy057](#) [\[Nachweis im GVK\]](#)

- Anna Samoilenko: Multilingual historical narratives on Wikipedia. 2017. In: GESIS Data Archive. Datensatz. Version 1. DOI: [10.7802/1411](https://doi.org/10.7802/1411)
- Judith Simon: Values in Design. In: Handbuch Medien- und Informationsethik. Hg. von Jessica Heesen. Stuttgart 2016, S. 357–364. [[Nachweis im GVK](#)]
- Routledge international handbook of participatory design. Hg. von Jesper Simonsen / Toni Robertson. New York, NY u. a. 2013. [[Nachweis im GVK](#)]
- Roger S. Ulrich: Biophilia, biophobia, and natural landscapes. In: The biophilia hypothesis. Hg. von Stephen R. Kellert. Washington, DC u. a. 1993, S. 73–137. [[Nachweis im GVK](#)]
- Roger S. Ulrich: View through a window may influence recovery from surgery. In: Science 224 (1984), H. 4647, S. 420–421. [[Nachweis im GVK](#)]
- Steven Umbrello / Ibo van de Poel: Mapping Value Sensitive Design onto AI for social good principles. In: AI and Ethics (2021). Artikel vom 01.02.2021. DOI: [10.1007/s43681-021-00038-3](https://doi.org/10.1007/s43681-021-00038-3) [[Nachweis im GVK](#)]
- A new companion to digital humanities. Hg. von Susan Schreibman / Ray Siemens / John Unsworth. Chichester u. a. 2016. [[Nachweis im GVK](#)]
- Sara Wachter-Boettcher: Technically wrong. Sexist apps, biased algorithms, and other threats of toxic tech. New York, NY u. a. 2017. [[Nachweis im GVK](#)]
- Vereinte Nationen Generalversammlung: Aktivitätenprogramm für die Durchführung der Internationalen Dekade der Menschen afrikanischer Abstammung. Resolution der Generalversammlung, verabschiedet am 18. November 2014.
- Sarah Myers West: Raging Against the Machine: Network Gatekeeping and Collective Action on Social Media Platforms. In: Mobile Media and Communication 5 (2017), H. 3, S. 28–36. DOI: [10.17645/mac.v5i3.989](https://doi.org/10.17645/mac.v5i3.989) [[Nachweis im GVK](#)]
- Kevin White / Shadi Abou-Zahra / Shawn Lawton Henry: Writing for Web Accessibility – Tips for Getting Started. Hg. von W3C Web Accessibility Initiative. In: w3.org/WAI/tips/writing/ vom 01.12.2020. [[online](#)]
- Till Winkler / Sarah Spiekermann: Twenty years of Value Sensitive Design: a review of methodological practices in VSD projects. In: Ethics and Information Technology 23 (2021), S. 17–21. Artikel vom 21.08.2018. DOI: [10.1007/s10676-018-9476-2](https://doi.org/10.1007/s10676-018-9476-2) [[Nachweis im GVK](#)]
- ZfdG (2021a): I wie Impressum. In: Zeitschrift für digitale Geisteswissenschaften. 2021. [[online](#)]
- ZfdG (2021b): P wie Publizieren. In: Zeitschrift für digitale Geisteswissenschaften. 2021. [[online](#)]

Abbildungsverzeichnis

Abb. 1: Die drei Untersuchungsschritte von Value-Sensitive-Design-Analysen. [Leyrer 2021].

Abb. 2: Beispiel für eine Natur-Szene, auf der keine Menschen abgebildet werden. Eignet sich auch für ein kurzes Experiment: Entspannt Sie der Anblick dieser Natur-Szene? [Wasserfall, Tal, Berge, Landschaft, Wald, Bachlauf. Hg. von Free Fotos. In: pixabay.com. Bild vom 08.01.2016]. [online]

Abb. 3: Ausschnitt aus einer Zusammenstellung der knapp 200 Blogartikel, die im Kontext der #femaleheritage-Blogparade entstanden sind. [Praske 2021]

Beitrag aus:

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. (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5) text/html Format. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Bomber's Baedeker – vom Text zum Bild zur Datenquelle

Autor*in:

Felix Bach

Kontakt: fbach9310@gmail.com

Institution: Leibniz-Institut für Europäische Geschichte (IEG) | Hochschule Darmstadt
GND: [124168099X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64863-p0011-9) ORCID: [0000-0001-9517-270X](https://orcid.org/0000-0001-9517-270X)

Autor*in:

Stefan Schmunk

Kontakt: stefan.schmunk@h-da.de

Institution: Hochschule Darmstadt
GND: [1028340028](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64863-p0011-9) ORCID: [0000-0001-9706-9757](https://orcid.org/0000-0001-9706-9757)

Autor*in:

Cristian Secco

Kontakt: stcrsecc@stud.h-da.de

Institution: Leibniz-Institut für Europäische Geschichte (IEG) | Hochschule Darmstadt
GND: [1241301123](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64863-p0011-9) ORCID: [0000-0002-5023-015X](https://orcid.org/0000-0002-5023-015X)

Autor*in:

Thorsten Wübbena

Kontakt: wuebbena@ieg-mainz.de

Institution: Leibniz-Institut für Europäische Geschichte (IEG)
GND: [123312396](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64863-p0011-9) ORCID: [0000-0001-8172-6097](https://orcid.org/0000-0001-8172-6097)

DOI des Artikels:

[10.17175/sb005_004](https://doi.org/10.17175/sb005_004)

Nachweis im OPAC der Herzog August Bibliothek:

[1770855890](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64863-p0011-9)

Erstveröffentlichung:

22.09.2021

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autor*innen

Letzte Überprüfung aller Verweise: 16.09.2021

GND-Verschlagwortung:

[Bomber's Baedeker](#) | [Datentransformation <OCR>](#) [<XML>](#) [<Python>](#) | [Digitalisierung](#) | [Informatik](#) | [Optische Zeichenerkennung](#) |

Zitierweise:

Felix Bach, Stefan Schmunk, Cristian Secco, Thorsten Wübbena: Bomber's Baedeker – vom Text zum Bild zur Datenquelle. 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. (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5) text/html Format. DOI: [10.17175/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_004](https://doi.org/10.17175/sb005_004).

Felix Bach, Stefan Schmunk, Cristian Secco, Thorsten Wübbena

Bomber's Baedeker – vom Text zum Bild zur Datenquelle

Abstracts

Das zweibändige Druckwerk *The Bomber's Baedeker. A Guide to the Economic Importance of German Towns and Cities* wurde während des Zweiten Weltkrieges vom britischen Foreign Office und dem Ministry of Economic Warfare erstellt. Darin sind Städte des Deutschen Reichs mit mehr als tausend Einwohner*innen und Informationen zu deren kriegswichtigen Infrastrukturen, Industrie- und Produktionsanlagen aufgeführt. Weltweit existieren nur noch vier nachgewiesene Exemplare und keines davon war bislang für die wissenschaftliche Nutzung digital zugänglich. Der *Bomber's Baedeker* wurde 2019 in der Bibliothek des Leibniz-Instituts für Europäische Geschichte (IEG) (wieder-)entdeckt, in Kooperation mit der Universitätsbibliothek Mainz digitalisiert und im Rahmen einer institutionsübergreifenden Kooperation vom Bereich Digitale historische Forschung | DH Lab des IEG und von der Hochschule Darmstadt, u. a. in Lehrveranstaltungen mit Studierenden, erschlossen und aufbereitet, so dass der *Bomber's Baedeker* nun als offene, maschinenlesbare Datenquelle unter Einhaltung der FAIR-Prinzipien genutzt, analysiert und weiterbearbeitet werden kann.

The two-volume printed work *The Bomber's Baedeker. A Guide to the Economic Importance of German Towns and Cities* was produced by the British Foreign Office and the Ministry of Economic Warfare during the Second World War. It lists towns and cities of the German Reich with more than a thousand inhabitants and information on their war-related infrastructure, industrial and production facilities. Only four verified copies still exist worldwide and none of them has been digitally accessible for scholarly use until now. In 2019, *The Bomber's Baedeker* was (re-)discovered in the library of the Leibniz Institute of European History (IEG), digitised in cooperation with the University Library of Mainz and made accessible and processed in a cross-institutional cooperation between the Digital Historical Research Unit | DH Lab of the IEG and the Darmstadt University of Applied Sciences, including in courses with students, so that *The Bomber's Baedeker* can now be used, analysed and further processed as an open, machine-readable data source in compliance with the FAIR principles.

1. Einleitung

»Handbuch für den Feuersturm« war der Titel der Frankfurt Allgemeinen Zeitung im Juni 2019¹ zum zweibändigen Werk *The Bomber's Baedeker. A Guide to the Economic Importance of German Towns and Cities*.² Anlass für die Berichterstattung war die vollständige Digitalisierung der zweiten Auflage der seltenen britischen Quelle aus dem Jahr 1944 (809 Seiten, Quartformat, 6 ganzseitige Karten). Diese Ausgabe ist weltweit nur noch in vier Exemplaren nachweisbar, so dass es durchaus als Glückstreffer bezeichnet werden kann, dass dieses Exemplar in der Forschungsbibliothek des Leibniz-Instituts für Europäische Geschichte (IEG) entdeckt und in einem gemeinsamen Projekt für eine digitale Nutzung vorbereitet wurde. Der Titel *Bomber's Baedeker* verwundert zugleich ein wenig, denn inhaltlich werden keine touristischen und sehenswerten Reiseziele identifiziert und beschrieben,

¹Vgl. Burger 2019.

²*The Bomber's Baedeker* 1944.

wie sie die namensgebenden roten Reiseführer Karl Baedekers seit dem 19. Jahrhundert³ liefern, sondern es werden ausschließlich militärische, infrastrukturelle und ökonomisch-kriegswichtige Ziele der Royal Air Force für den Bombenkrieg gegen das Deutsche Reich im Zweiten Weltkrieg identifiziert. Das zweibändige Druckwerk enthält Informationen zu Orten mit mehr als tausend Einwohner*innen, sofern dort aus Sicht des britischen Foreign Office und des Ministry of Economic Warfare kriegswichtige Industrie- und Produktionsanlagen vorhanden waren. Die Daten zur geografischen Lage, Einwohnerzahl, Entfernung zu London sowie umfassende Beschreibungen von mehr als 500 Städten in Deutschland wurden im Bestreben zusammengestellt, eine möglichst effektive und effiziente Auswahl von potenziellen Zielen zu ermöglichen.⁴

Für das Lesen der Quelle ist die digitale Bereitstellung als gemeinfreies Digitalisat hinreichend, für eine Auswertung der enthaltenen Daten kann dies aus Sicht der Digitalen Geisteswissenschaften aber nur als erster Schritt betrachtet werden. Nach der Transformation zu einer digitalen Bilddatei ist die Weiterverarbeitung zu maschinenlesbaren Daten für zahlreiche Methoden der Digital Humanities ein wichtiger Schritt in der Aufbereitung. Im Fall des Bomber's Baedeker wurde dies durch eine Verbesserung der Zeichenerkennung und Überführung der Inhalte in ein standardisiertes Datenformat durchgeführt, so dass dadurch digitale Analysen ermöglicht werden. Seit 2019 findet daher im Rahmen einiger kleinerer Projekte eine intensive wissenschaftliche Auseinandersetzung und zugleich eine datenspezifische Aufbereitung statt, so dass die Qualität der maschinenlesbaren Daten stetig verbessert wurde.

2. Hintergrund

Woher genau die Bezeichnung Bomber's Baedeker stammt und wie deren etymologische Entwicklung zu deuten ist, lässt sich nicht mit hundertprozentiger Sicherheit belegen. Vermutlich ist diese Namensgebung eine Reaktion auf die im Frühjahr 1942 durch die Deutsche Luftwaffe ausgeführten Angriffe auf Exeter, Bath, Norwich, York und Canterbury.⁵ Als Begründung der Auswahl der Ziele wurden von deutscher Seite in einer Pressekonferenz des Auswärtigen Amtes im April 1942 die Auszeichnung dieser Städte im Baedeker-Reiseführer für England angegeben.⁶ Verständlicherweise sorgte dies in der britischen Bevölkerung für einen ungeheuren öffentlichen Furor, da durch diese Aussage deutlich wurde, dass eben nicht ausschließlich militärische Ziele bzw. Industriestandorte, sondern vielmehr bewusst historisch bedeutsame Städte von deutscher Seite als Ziele ausgewählt wurden.⁷

³Zu einer medienkulturhistorischen Analyse von Reiseführern vgl. Müller 2012.

⁴Vgl. Hohn 1994, S. 213–230.

⁵Vgl. Rothnie 1991, S. 142f.

⁶Mit den Worten »Now the Luftwaffe will go for every building which is marked with three stars in Baedeker« wurde Legationsrat Gustav Braun von Stumm in der britischen Presse zitiert. Vgl. Knuth 2006, S. 165.

⁷Eine virtuelle Ausstellung des Imperial War Museums zu den »Baedeker Raids« findet sich bei Google Arts and Culture.

Diese Vorgänge und insbesondere die enorme öffentliche Empörung dürften dazu geführt haben, dass die Mitarbeiter*innen des Foreign Office und des Ministry of Economic Warfare, welche die Analyse der militärisch und wirtschaftlich bedeutsamen Ziele im Deutschen Reich bereits 1942 – eben zum Zeitpunkt der sogenannten ›Baedeker Raids‹ – durchführten, ihrerseits wiederum den Namen Bomber's Baedeker für die eigenen Aufstellungen wählten. Die erste Auflage des Bomber's Baedeker erschien dann auch ein Jahr später und deckt sich mit den alliierten Absprachen zu einem gemeinsamen ›Bomber Command‹ zwischen Großbritannien und den USA auf der Casablanca-Konferenz im Januar 1943, auf der u. a. festgelegt wurde: Vordringliches Ziel des Bomber Command ist die fortschreitende Zerstörung des deutschen militärischen, industriellen und wirtschaftlichen Systems, um die Moral des deutschen Volkes bis zu einem Grad zu untergraben, wo seine Fähigkeit zum bewaffneten Widerstand entscheidend geschwächt ist.⁸

Zeitgleich wurde im britischen Unterhaus das Dehousing Paper verabschiedet, in dem als strategisches Ziel der britischen Bomberverbände die gezielte Zerstörung von Wohngebieten vorgesehen wurde – einer Doktrin mit Namen ›Moral Bombing‹, an der bis zum Ende des Zweiten Weltkrieges festgehalten wurde und die bereits im Frühjahr 1942 zum 1.000-Bomber-Angriff gegen Köln führte.⁹ Es ist allerdings festzuhalten, dass die Auswahl der tatsächlichen Ziele dem Bomber Command und dem Air Ministry oblag und deshalb nicht eindeutig nachvollziehbar ist, welche Rolle der Bomber's Baedeker insbesondere bei der Wahl der Angriffsziele tatsächlich besaß – vor allem, weil in diesem ausschließlich infrastrukturelle und wirtschaftliche Ziele angegeben waren. Die Datengrundlage für die Erhebung durch das Foreign Office und das Ministry of Economic Warfare bildeten – neben Informationen aus der Feindaufklärung, Adressbüchern, Luftaufnahmen, Berichten von Emigrant*innen etc. – vor allem die Unterlagen der britischen Rückversicherer. Bei diesen waren seit Mitte der 1930er-Jahre die Brandversicherungen der deutschen Versicherungsunternehmen abgesichert. Da es sich um eine Pflichtversicherung für alle Gebäude im Deutschen Reich handelte, stellte dies eine vollständige Datenbasis dar, um eine umfassende Quartiersanalyse aller deutschen Städte zu erstellen. Darüber hinaus konnten hierüber die Standorte aller Firmen identifiziert werden und über deren Namen und Eigentümer*innen war zugleich zu erfahren, was dort höchstwahrscheinlich produziert wurde. Auch konnte diesen Unterlagen Informationen über die Bausubstanz der Gebäude entnommen werden. Auf dieser Basis war es möglich, eine höchst detaillierte Topographie jeder einzelnen deutschen Stadt zu erstellen und zugleich die Unterschiede in der quartierbezogenen Bausubstanz zu erfassen.¹⁰ De facto war genau dies der Schlüssel für die alliierten Luftangriffe gegen Deutschland, bei denen gezielt Städte bzw. Stadtteile mit leicht entzündbaren und brennbaren Baustoffen angegriffen wurden.¹¹ Folgt man dieser Argumentation, so wird daraus deutlich, dass Bomber's Baedeker eher für strategische Planungen genutzt wurde und weniger eine taktische Bedeutung besaß.

⁸TNA London, AIR 41/5, Directive 21 January 1943, International Law of the Air, 1939–1945, Confidential supplement to Air Power and War Rights, 1946, zitiert nach: Böhm 2015, S. 147.

⁹Vgl. Longmatte 1983.

¹⁰Vgl. Hohn 1994, S. 213–230.

¹¹Vgl. zur Geschichte des Bombenkrieges gegen deutsche Städte: Overy 2014; Müller 2004; Boog et al. (Hg.) 2001.

Während in der ersten Auflage des Bomber's Baedeker von 1943 nur 392 Städte mit einer Größe über 15.000 Einwohner berücksichtigt wurden, beinhaltet die zweite Auflage von 1944 insgesamt 518 Städte und umfasst auch Kleinstädte ab einer Größe von 1.000 Einwohnern. Dieser Umstand ist u. a. darauf zurückzuführen, dass ab Mitte 1943 eine Verlagerung von kriegswichtigen Produktionsstätten aus den Städten erfolgte.¹² Für die Datenerhebung und #aufbereitung im Rahmen des zugrunde liegenden Projektes wurde die zweite Auflage von 1944 verwendet, die neben der Nennung der jeweiligen Stadt mit entsprechenden Breiten- und Längenangaben und der Flugdistanz (in Meilen) zu London einleitend auch eine kurze Beschreibung der Stadt mit geographischen Markern (in der Nähe liegende Flüsse, Berge, Seen, Wälder etc.) sowie eine Kategorisierung der Ziele gibt.¹³ Folgende Kategorisierung wird im Bomber's Baedeker für alle Städte angewandt:¹⁴

- Transportwesen
- Infrastruktur (Wasser, Elektrizität etc.)
- Festbrennstoffe (Bergbau, Brennstofflager etc.)
- Flüssigbrennstoffe (Raffinerien, Brennstofflager etc.)
- Eisen- und Stahlindustrie
- Sonstige metallverarbeitende Betriebe
- Flugzeuge und Motoren
- Werften
- Sonstige Industrie- und Rüstungsbetriebe
- Chemie- und Munitionsbetriebe
- Textil-, Seide-, Zellstoff- und Papierbetriebe
- Gummi- und Reifenhersteller
- Lederindustrie
- Nahrungsmittelindustrie

Die aus Sicht der beiden Ministerien kriegswichtige Bedeutung dieser Kategorien wurde zudem auf einer Skala von 1 bis 3 bewertet, wobei anzumerken ist, dass nicht alle Kategorien bei allen Städten zu finden sind.

3. Datenerhebung – vom Bild zum XML

Die hier am Bomber's Baedeker durchgeführte Transformation des Inhalts eines gedruckten Buchs in maschinenlesbare Daten – auf Grundlage einer zuvor erstellten digitalen 1:1-Abbildung (Repräsentant) – ist ein wichtiger Baustein im Prozess der Datenaufbereitung und zumeist die Voraussetzung für die Anwendung zahlreicher Methoden und Verfahren der Digital Humanities.

¹²Vgl. Schmunk 2005, S. 59, 65f.

¹³Der Unterschied zwischen der ersten und zweiten Auflage besteht vor allem darin, dass in der zweiten Auflage weitere Städte aufgenommen und somit der Datenbestand vergrößert wurde.

¹⁴Vgl. The Bomber's Baedeker 2019 (1944), [Preface](#).

Wie oben beschrieben, besitzt die Bibliothek des IEG mit dem zweibändigen Werk eines der wenigen noch verfügbaren Exemplare. Mit dieser Situation geht auch eine entsprechende Verantwortung einher, zum einen in konservatorischer Hinsicht und zum anderen im Hinblick auf die Verfügbarmachung des Inhalts für die Forschung. Beiden Aspekten kann mithilfe der Digitalisierung begegnet werden. Deshalb wurde im Rahmen eines gemeinsamen Vorhabens zwischen IEG und der Universitätsbibliothek der Johannes Gutenberg-Universität Mainz im dortigen Servicezentrum Digitalisierung und Fotodokumentation der Bomber's Baedeker im Jahr 2019 digitalisiert. Die angefertigten Digitalisate (in den Formaten ›.jpg‹ und ›.pdf‹) stehen seitdem in Gutenberg Capture, dem Online-Portal der Universitätsbibliothek zur digitalen Erschließung und Bereitstellung von Quellenmaterial für die Wissenschaft, zur Verfügung.¹⁵

Im Zuge der Digitalisierung wurde durch die Universitätsbibliothek auch eine erste Erschließung des Textes mittels Optical Character Recognition (OCR) durchgeführt.

An dieser Stelle setzte die Machbarkeitsstudie an, die im Wintersemester 2019 / 20 an der Hochschule Darmstadt durchgeführt wurde, und in deren Rahmen das Konzept für einen automatisierten Prozess der XML-isierung und einer OCR-Verbesserung des Bomber's Baedeker entstand. Dieses Konzept war ein erster, methodologisch wichtiger Baustein, der dann im Rahmen einer Zusammenarbeit zwischen der Hochschule Darmstadt, der Bibliothek des IEG und dem Bereich Digitale historische Forschung | DH Lab (ebenfalls IEG) umgesetzt wurde. Die ersten experimentellen Ansätze der Transformation der Quelle zu Daten wurden hier realisiert und auf Grundlage dieser datafication – also der Umwandlung von Informationen in maschinenlesbare, quantifizierbare Daten zum Zweck der Aggregation und Analyse – und der entsprechenden Verfügbarmachung ist nun eine weitergehende Bearbeitung und Analyse dieser Daten möglich.

3.1. OCR

Die Inhalte des Bomber's Baedeker werden bereits im Druck stark strukturiert dargestellt und sind daher grundsätzlich sehr gut geeignet, um die darin enthaltenen Informationen in eine maschinenlesbare, objektorientierte und strukturierte Form zu bringen. Eine nähere Betrachtung ergibt, dass jeweils zwei Hauptabschnitte pro Stadt aufgeführt werden, die sich in folgende Muster aufgliedern:

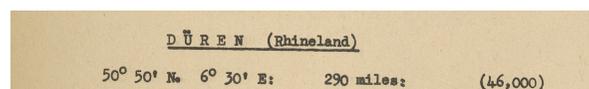


Abb. 1: ›Kopfteil‹ (Düren), Ausschnitt [The Bomber's Baedeker. Guide to the economic importance of German towns and cities, London 1944, S. 176. [Public Domain Mark 1.0]

¹⁵ The Bomber's Baedeker 2019 (1944).

I. Kopfteil

- Name der Stadt (Großbuchstaben, Sperrsatz), dahinter in Klammern: die Verwaltungseinheit und gegebenenfalls die Region.
- In der nächsten Zeile: Geokoordinaten der Stadt im Format ›00° 00' N. 00° 00' E:‹. Hier sind Variationen zu beobachten. So tauchen auch einstellige Angaben auf und es existieren nicht immer Nachkommastellen in den Geokoordinaten, also z. B. ›00° N. 00° E:‹.
- Im Anschluss an die Koordinaten folgt die Entfernung zu London in Meilen, im Format ›000 miles:‹.
- Die letzte Information in der zweiten Zeile des Kopfteils gibt die Einwohnerzahl in Klammern wieder.

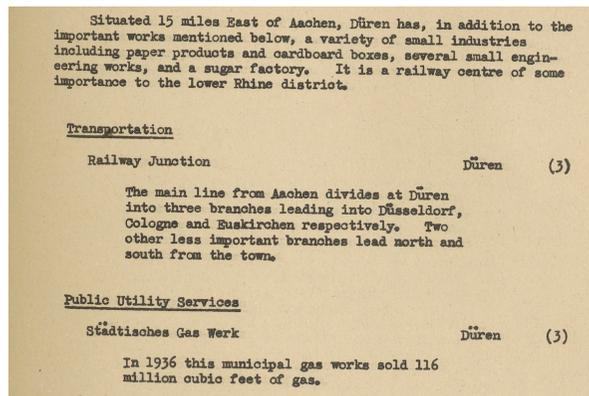


Abb. 2: ›Beschreibungsteil‹ (Düren), Ausschnitt [The Bomber's Baedeker. Guide to the economic importance of German towns and cities, London 1944, S. 176. [Public Domain Mark 1.0]

II. Beschreibungsteil

- Eine knappe Beschreibung der wichtigsten Charakteristika der jeweiligen Stadt (z. B. grobe geographische Verortung, Verkehrsinfrastruktur, industrielle Bedeutung).
- Strukturierung nach Kategorien (z. B. ›Transportation‹, ›Liquid Fuels and Substitutes‹ usw.)
- Es folgt eine detaillierte Auflistung der wichtigsten Ziele innerhalb der Kategorie, pro potenzielles Ziel unterteilt in drei Punkte:
 -
 - Name und Beschreibung des Ziels.
 - Standort.
 - Priorität (in absteigender Reihenfolge ›1‹ bis ›3‹ sowie ›-‹ für nicht erwähnenswerte bzw. nicht einschätzbare Ziele).

Ein Blick in eines der bei Gutenberg Capture zur Verfügung stehenden Digitalisate im PDF-Format zeigt, dass hier bereits ein OCR-Verfahren eingesetzt wurde. Jede Seite verfügt über einen Text-Layer, der es ermöglicht, im Dokument zu suchen. Die Ergebnisse des verwendeten OCR-Verfahrens sind aber insgesamt nur für eine einfache Suche verwendbar, wenig verlässlich und letztlich nicht für eine direkte Umwandlung im Sinne des Vorhabens

geeignet. Nach entsprechender Analyse¹⁶ kann von einer OCR-Genauigkeit von ungefähr 80 % ausgegangen werden, wobei die für die effiziente Umsetzung des Projektes benötigte Erkennungsqualität aber mindestens bei mehr als 95 % liegen sollte.¹⁷ Die vorliegenden OCR-Daten waren daher nicht für eine Prozessierung nutzbar und die OCR musste erneut durchgeführt werden. Angesichts dieses neuen Zwischenschritts ergab sich auch die Situation einer erneuten, genauen Betrachtung des digitalen Ausgangsmaterials. Zwei Probleme zeigten sich direkt, zum einen der Vergilbungsgrad der Seiten und die daraus resultierende zu geringe Helligkeit des Digitalisats, so dass relevante Details nicht von der OCR erkannt werden konnten (false negative) beziehungsweise irrelevante Artefakte auf dem Papier erkannt wurden (false positive). Zum anderen zeigten sich einige Schwächen im Kontrast des Schriftbilds, was angesichts des Alters der Schreibmaschinenseiten des physischen Originals und dessen Nutzungsszenarien nicht weiter verwundert.¹⁸

Nach der Entscheidung für eine erneute OCR-Behandlung fiel die Wahl zunächst auf die freie Software Tesseract.¹⁹ Allerdings erwies sich diese im operativen Einsatz zum Zeitpunkt Ende 2019 als nicht sehr geeignet bei der Behandlung größerer Textmengen, da eine Aufteilung des Konvoluts notwendig geworden wäre. Im Falle des Bomber's Baedeker, mit mehr als 800 Seiten, wäre hier ein erheblicher Zeitaufwand entstanden.

Als Alternative kam stattdessen das Werkzeug FineReader der Firma ABBYY zum Zuge,²⁰ mit dem sich zum Zeitpunkt der OCR-Verbesserung unkompliziert auch PDF-Dateien größeren Umfangs verarbeiten ließen. Darüber hinaus bietet diese Software die Option, mehrere Sprachen gleichzeitig verarbeiten zu lassen, was in der vorliegenden Situation von großem Nutzen war, da der Löwenanteil des Textes zwar auf Englisch²¹ verfasst ist, aber aufgrund der behandelten Gegenstände natürlich auch zahlreiche deutsche Begriffe, wie z. B. Städtenamen etc., zu finden sind.

Die vollumfängliche Verarbeitung dauerte drei Stunden und das Ergebnis wurde in einer TXT-Datei gespeichert. Durch dieses Vorgehen konnte die Genauigkeit – ohne Veränderung oder Bearbeitung des Digitalisats – gesteigert werden. Letztlich war aber auch die damit erzielte Gesamtgenauigkeit nicht ausreichend, so dass eine Bildbearbeitung der digitalen Vorlage notwendig wurde, um die anvisierten Ziele in eine realistische Nähe rücken zu lassen. Um eine optimale Erkennung der gedruckten Zeichen zu erreichen, war es nötig, den Kontrast zu erhöhen. Hierzu wurde die in MacOS integrierte Applikation Fotos genutzt, mit der dann auf sehr einfache Art Bildkorrekturen erstellt, kopiert und in der Gesamtheit auf alle Seiten angewendet wurden. Nach dieser Optimierung konnten die Ergebnisse des OCR-Verfahrens

¹⁶Mit einem Zufallsgenerator wurden 500 beliebige Zeichen des Gesamttexts ausgewählt und diese auf Richtigkeit überprüft. Die Auswahl des Validierungsverfahren, dem sog. Bernoulli-Experiment, basiert auf einer Empfehlung der [DFG-Praxisregeln Digitalisierung](#), DFG (Hg.) 2016.

¹⁷Dies ist u. a. darauf zurückzuführen, dass der Text selbst mit einer Schreibmaschine erstellt und dann hektographiert wurde, so dass einzelne Stellen schlecht lesbar und für die OCR nicht bzw. kaum interpretierbar waren.

¹⁸Es lassen sich auch Phänomene wie schwächer werdende Farbbänder beobachten, die z. B. auch innerhalb einer Seite für starke Kontrastschwankungen sorgen.

¹⁹[Tesseract-OCR](#).

²⁰[ABBYY FineReader](#).

²¹Zum Zeitpunkt der Durchführung dieses Arbeitsschritts im Projekt verfügte Tesseract nicht über die entsprechenden Funktionalitäten.

erneut und sehr deutlich gesteigert werden, so dass nun eine Genauigkeit von ca. 95 % vorlag.²² Da auf diesem Weg keine weiteren großen Verbesserungen der OCR-Qualität zu erwarten waren, wurde fortan zur weiteren Steigerung der Datenqualität auf ein Python-Skript gesetzt, so dass zum Zeitpunkt der Veröffentlichung des Data-Papers im September 2021 eine Zeichengenauigkeit von mehr als 99 % erreicht werden konnte.

3.2. Python-Skript

Die Struktur des für das weitere Vorgehen geschriebenen Python-Skripts kann in drei Schritte unterteilt werden. Schritt Eins stellt das Preprocessing dar. Hier werden wiederholt auftretende Fehler aus dem OCR-Verfahren verbessert, die in der nachfolgenden Ausführung weiterer Schritte im Skript zu Fehlern oder fehlerhafter Erfassung der Objekte führen könnten. Zusätzlich werden strukturelle Abweichungen an die im Bomber's Baedeker vorwiegend genutzten Normen angepasst und es werden Funktionen ausgeführt, die folgende Muster in den oben beschriebenen Kopf- und Beschreibungsteilen erkennen: Städtenamen, Land / Region, Koordinaten, Entfernung zu London, Bevölkerungszahl, Informationstext und strategische Ziele der beschriebenen Stadt. Darüber hinaus wird auf Seitenebene die referenzierende URL in Gutenberg Capture erzeugt.

Die oben genannten Funktionen liefern eine Liste von erkannten Inhalten und hier setzt der zweite Schritt an. Die Anzahl der Inhalte ist aufgrund der vorgegebenen Struktur immer dieselbe und jeder Listeneintrag korrespondiert mit dem entsprechenden Eintrag in der Vorlage. Der vierte Eintrag in der generierten Liste der Bevölkerungszahlen zum Beispiel ist aus dem vierten Eintrag in der Liste der erkannten Städte entstanden. Für die weitere Verarbeitung werden diese Listen nun in ein sogenanntes Dictionary umgewandelt und hier wird jedem Wert ein Schlüssel zugeordnet, wodurch die Daten nun so gespeichert werden können, dass klar ist, welche Informationen sie enthalten. Auf diese Weise haben wir ein Dictionary mit Listen erstellt, die schon alle benötigten Informationen enthalten.

Im dritten und letzten Schritt wird dieses Dictionary dann in eine XML-Datei eingefügt.²³ Damit hier nicht in eine komplett leere Datei geschrieben wird, wurde die Grundstruktur vorher schon erstellt. Beim Prozess des Exports werden alle Inhalte aus dem Data-Dictionary eingelesen, sowie die Seitenzahlen aus dem Originaldokument. Nun kann eine XML-Baumstruktur erzeugt werden, an die zuerst die Städtenamen angefügt werden. Anschließend wird jeder Stadt-Eintrag mit den dazugehörigen Informationen gefüllt. Als letzter Schritt wird die XML-Datei exportiert und die entstandenen HTML-Entitäten werden aufgelöst.²⁴

²²Beim Präprozessieren sollten zukünftig alternative Verfahren zur Kontrastkorrektur durchgeführt werden – Abby Finereader unterstützt dies ebenfalls.

²³Ein mit XML ausgezeichnete Text bietet sich hier als Zielformat an, da entsprechend strukturierte Dateien für weitere Möglichkeiten der Nachnutzung offenbleiben und jedes spezifische Format an dieser Stelle ohne konkrete Aufgabenstellung produziert worden wäre.

²⁴Das hier genutzte Skript ist samt Dokumentation im GitHub-Repositorium des Leibniz-Instituts für Europäische Geschichte unter folgender URL aufzurufen: https://github.com/ieg-dhr/bombers_baedeker/.

3.3. Bereitstellung

Neben dem Python-Skript für die OCR-Optimierung finden sich in dem Projekt-Repository auch die aus seiner Anwendung entstandenen Daten: für jeden Band des Bomber's Baedeker liegt eine XML-Datei vor, in welcher der entsprechend formal strukturierte und angereicherte Text enthalten ist.²⁵ Zukünftige Versionen mit verbesserter Datenqualität werden dort ebenfalls publiziert.

Bei der Bereitstellung des Datensatzes und des Skripts wurden die FAIR Data Principles umfänglich berücksichtigt.²⁶ Für die Erfüllung dieser Prinzipien sorgen die Zugriffsmöglichkeiten per GitHub-Repository und die Veröffentlichung der erzeugten Daten in Zenodo, womit ein persistenter Identifikator (DOI) einhergeht.²⁷ Für die rechtliche Sicherheit bei der Wiederverwendung von Daten und Skript sorgen die gewählten Lizenzen (Creative Commons Lizenz, CC BY-SA 4.0²⁸ für die Daten sowie die GNU General Public License²⁹ für das Python-Skript).

4. Forschungs- und Nachnutzungspotenzial

Wie oben bereits geschildert, handelte es sich bei dieser Bereitstellung der nachnutzbaren Daten des Bomber's Baedeker um ein Vorhaben, welches mit einem überschaubaren Ressourceneinsatz durchgeführt wurde. Der Fokus lag dabei primär auf der persistenten Verfügbarmachung des Datenbestands für die scientific community. Eine basale, aber dennoch nicht zu vernachlässigende Verbesserung, die sich für die Forscher*innen direkt aus dem digitalen Angebot ergibt, spiegelt sich im vereinfachten Umgang mit dem Text wider. In den bereitgestellten Daten lassen sich Such- und Analyseszenarien durchspielen, deren Umsetzung ausschließlich auf dieser digitalen Grundlage möglich sind.

Die bislang durchgeführten Arbeiten haben die Grundlagen geschaffen, um in einem nächsten Schritt ein digitales Editionsprojekt durchzuführen. Allein im Bereich der Auszeichnung und Anreicherung mit Normdaten – z. B. bei den Städte- oder Firmennamen – besteht großes Potenzial für weitere Analysen. Neben der weiteren Aufbereitung der Daten sehen wir ein breites Feld an möglichen Forschungsfragen, die an den Text bzw. die Daten gestellt werden können. So wäre ein Abgleich der im Bomber's Baedeker vorgeschlagenen Ziele im Deutschen Reich mit den tatsächlich bombardierten Städten und Einrichtungen eine interessante Forschungsfrage, um einerseits Rückschlüsse auf die praktische Anwendung im Hinblick auf die Zielvorgaben der Royal Airforce zu geben. Andererseits können die detaillierten Standortinformationen der Firmen aus wirtschaftshistorischer Perspektive erstmals einen annähernd vollständigen Überblick über die Branchenverteilung in den 1930er- und 1940er-

²⁵ Das Verzeichnis ›`bomber_output_ext`‹ enthält das XML-Ergebnis des datengenerierenden Skripts.

²⁶ FAIR Data Principles, GOFair (Hg.) 2016–2021.

²⁷ Bach et al. 2021.

²⁸ CC BY-SA 4.0, Creative Commons – Namensnennung – Weitergabe unter gleichen Bedingungen – 4.0 International, Creative Commons (Hg.) 2021.

²⁹ GNU General Public License, Free Software Foundation, Inc. (Hg.) 2007.

Jahren geben. In beiden Fällen wäre noch zu prüfen, wie exakt die vermerkten Geokoordinaten angegeben sind und ob sich gegebenenfalls ein geographisches Muster aus den ermittelten Abweichungen herauslesen lässt.

Unsere Arbeiten ermöglichen nun einen datengetriebenen Analyseansatz. Während mit einer traditionellen Methodik oftmals Karten ausschließlich rein visuell analysiert werden und zugleich deren Richtigkeit und Entstehungskontexte nur partiell hinterfragt werden können, z. B. ob die Ersteller*innen – wie in Abb. 3 dargestellt – eindeutige und gültige Geoangaben verwendet haben, besitzen datengetriebene Ansätze der Visual Analytics weitaus größere Möglichkeiten und können aufzeigen, auf welcher Datenbasis die entsprechenden Angaben generiert wurden.

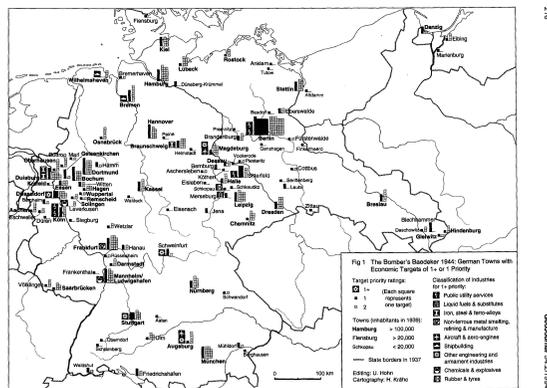


Abb. 3: Klassische Visualisierung der Ziele nach höchster Priorität aus dem Bomber's Baedeker. [Uta Hohn: The Bomber's Baedeker – Target Book for Strategic Bombing in the Economic warfare against German Towns 1943–1954. In: Hohn 1994, S. 216.]

Dies ist allerdings erst dann möglich, wenn die entsprechenden Angaben maschinenlesbar vorliegen und so aufbereitet werden, dass sie visuell interpretiert werden können. Die Darstellung in Abb. 4 wurde – basierend auf den im Bomber's Baedeker gemachten Angaben der Geokoordinaten – mittels Tableau³⁰ und dem DARIAH-DE Geo-Browser³¹ angefertigt. Hier ist erkennbar, dass die Ersteller*innen des Bomber's Baedeker eine klassische (Raster-)Papierkarte als Vorlage benutzt haben müssen, mit deren Hilfe die entsprechenden Geokoordinaten in die Textfassung übertragen wurden. Auffällig ist, dass die validen Geokoordinaten aufgrund des damaligen technischen Standes der Flugzeug-Leitsysteme keine Bedeutung besaßen, da die Flugzeuge mittels Radio-Leitstrahlen navigierten. Die Geokoordinaten hatten dementsprechend nur einen untergeordneten Informationsgehalt. Auch können auf diese Weise fehlerhafte Angaben im Bomber's Baedeker schneller identifiziert werden – wie beispielsweise, dass Mannheim versehentlich mit den Geokoordinaten von Kiew versehen wurde. Alleine durch dieses Beispiel wird deutlich, dass die Angabe der Geokoordinaten nur eine zusätzliche faktisch rein kartographische Information darstellt.

³⁰ Tableau, Tableau Software (Hg.) 2003–2021.

³¹ DARIAH-DE Geo-Browser, DARIAH-DE (Hg.) 2021.

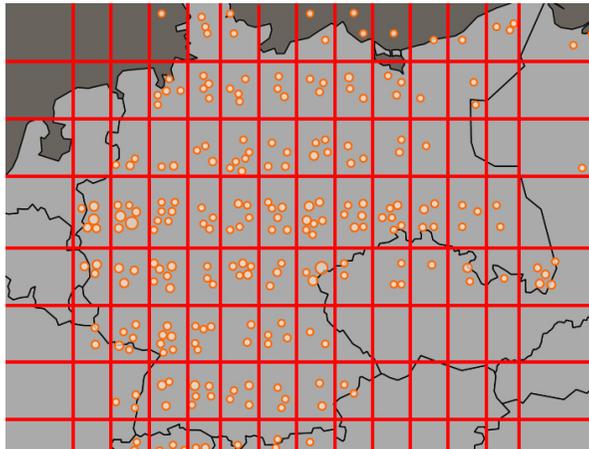


Abb. 4: Visualisierung der im Bomber's Baedeker benannten Zielstädte basierend auf den angegebenen zeitgenössischen Geokoordinaten mittels DARIAH-DE Geo-Browser und Tableau. [Bach / Secco 2021]

Wichtiger war zum Zeitpunkt der Erstellung in den 1940er-Jahren also vielmehr die Entfernung zu London und über welche Planquadrate die Flugzeuge fliegen mussten, um ihre Ziele zu erreichen. Wetterbedingungen, die Abwehrmaßnahmen der Deutschen Luftwaffe, aber auch welche Ziele in der Nähe lagen, die alternativ angefliegen werden konnten, wenn das eigentliche Ziel nicht erreicht werden konnte, spielten eine größere Rolle, als die richtige Angabe der Geokoordinaten.

Allein durch dieses Beispiel wird deutlich, dass die basale Erschließung dieses Datensatzes in XML und die dadurch hergestellte Maschinenlesbarkeit ungeheures Forschungspotenzial verspricht. Momentan arbeiten wir daran, dieses Beispiel auf die Ebene der einzelnen Städte herunterzubrechen, um zu analysieren und einzuschätzen, wie gut und valide die Alliierten tatsächlich über die Standorte von Infrastrukturen und Industrieanlagen informiert waren. Der Abgleich mit zeitgenössischen Stadtplänen und Adressbüchern ermöglicht diesbezüglich die Durchführung valider Datenanalysen.

5. Ausblick

Wie nicht zuletzt an diesem Beispiel der Verfügbarmachung des Bomber's Baedeker nachzuvollziehen ist, stellen wissenschaftliche Bibliotheken bereits eine große Anzahl von Digitalisaten bereit. In einigen Fällen wurden bereits OCR-Verfahren eingesetzt, um den Nutzer*innen weitergehende Möglichkeiten zu bieten. Leider liegen die angebotenen OCR-Daten nicht immer in einer Qualität vor, die eine digitale Bearbeitung ohne weiteren Aufwand für die Vorverarbeitung erlaubt. Diese Situation belastet die ohnehin knappen Ressourcen in Forschungsprojekten.

Es wäre wünschenswert, wenn die hier festgehaltene Herausforderung der Erstellung von maschinenlesbaren Datensets zukünftig von Bibliotheken, Forschungsdateninfrastrukturinitiativen und Wissenschaftler*innen gemeinsam adressiert

werden könnte. Auf diese Weise kann eine höhere, leichter verarbeitbare Datenqualität in den Angeboten der Bibliotheken zur Unterstützung der digital forschenden Wissenschaftler*innen generiert werden.

Nachdem die Bereitstellung des Volltextes für die Nutzer*innen mehr oder weniger fest im Portfolio der Bibliotheken verankert ist, wäre der nächste Schritt, das Bewusstsein für die Notwendigkeit einer hohen Datenqualität (über die Metadaten hinaus) zu schärfen. Trotz vieler guter Beispiele für Ansätze und Umsetzungen sind hier weitere Aktivitäten notwendig, um das Mindset und die aktuelle Praxis zu verändern. Die geisteswissenschaftlichen Initiativen in der Nationalen Forschungsdateninfrastruktur (NFDI) und seine relevanten Teilnehmenden wären die idealen Multiplikator*innen, um die Standards der Datenbereitstellung für digitale Texte voranzutreiben.

Bibliographische Angaben

Felix Bach / Cristian Secco / Stefan Schmunk / Thorsten Wübbena: The Bomber's Baedeker. A Guide to the Economic Importance of German Towns and Cities. In: zenodo.org. Data set vom 26.07.2021. DOI: [10.5281/zenodo.5138504](https://doi.org/10.5281/zenodo.5138504)

Das Deutsche Reich in der Defensive – Strategischer Luftkrieg in Europa, Krieg im Westen und in Ostasien 1943 bis 1944/45. Hg. von Horst Boog / Gerhard Krebs / Detlef Vogel. Stuttgart u. a. 2001. (= Das Deutsche Reich und der Zweite Weltkrieg, 7) [[Nachweis im GBV](#)]

Martin Böhm: Die Royal Air Force und der Luftkrieg 1922–1945. Personelle, kognitive und konzeptionelle Kontinuitäten und Entwicklungen. Paderborn 2015. [[Nachweis im GBV](#)]

Rainer Burger: Handbuch für den Feuersturm. In: Frankfurter Allgemeine Zeitung. Artikel vom 26.06.2019. [[online](#)]

Creative Commons-Lizenzen. Hg. von Creative Commons. Mountain View, CA 2021. [[online](#)]

DARIAH-DE Geo-Browser. Hg. von DARIAH-DE. Version 3.6.7. Göttingen 2021. [[online](#)]

DFG-Praxisregeln ›Digitalisierung‹. Hg. von Deutsche Forschungsgemeinschaft. Bonn 2016. PDF. [[online](#)]

FAIR Data Principles. Hg. von GOFair. In: go-fair.org. Leiden u. a. 2016–2021. [[online](#)]

GNU General Public License. Hg. von Free Software Foundation, Inc. Version 3 vom 29.06.2007. [[online](#)]

Uta Hohn: The Bomber's Baedeker-target book for strategic bombing in the Economic Warfare against German Towns 1943–45. In: GeoJournal 34 (1994), H. 2, S. 213–230. [[Nachweis im GBV](#)]

Rebecca Knuth: Burning Books and Leveling Librarians. Extremist Violence and Cultural Destruction. Westport, CT 2006. [[Nachweis im GBV](#)]

Norman Longmatte: The Bombers: The RAF offensive against Germany 1939–1945. London u. a. 1983. [[Nachweis im GBV](#)]

Rolf-Dieter Müller: Der Bombenkrieg 1939–1945. Berlin 2004. [[Nachweis im GBV](#)]

Susanne Müller: Die Welt des Baedeker. Eine Medienkulturgeschichte des Reiseführers 1830–1945. Frankfurt/Main u. a. 2012. [[Nachweis im GBV](#)]

Richard Overy: Der Bombenkrieg: Europa 1939–1945. Berlin 2014. [[Nachweis im GBV](#)]

Niall Rothnie: The Baedeker Blitz. Hitler's Attack on Britain's Historic Cities. Shepperton 1992. [[Nachweis im GBV](#)]

Stefan Schmunk: Entweder KZ oder ordentliche Deutsche. Die Luftwaffe und der Arbeitseinsatz 1942–1944. Darmstadt 2005. In: researchgate.net. DOI: [10.13140/rg.2.2.20030.08003](https://doi.org/10.13140/rg.2.2.20030.08003)

Tableau. Hg. von Tableau Software. In: tableau.com. Seattle, WA 2003–2021. [[online](#)]

The Bomber's Baedeker. Guide to the economic importance of German towns and cities. (Foreign Office & Ministry of Economic Warfare). 2 Bände. London 1944. In: Gutenberg Capture. Hg. von Universitätsbibliothek Mainz. Online-Ausgabe. Mainz 2019. URN: [urn:nbn:de:hebis:77-vcol-20056](https://nbn-resolving.org/urn:nbn:de:hebis:77-vcol-20056)

The Bomber's Baedeker Guide to the economic importance of German towns and cities. London 1944. Softwareskripte und Dokumentation. Hg. von IEG Mainz. In: github.com. 2021. [[online](#)]

TNA London, AIR 41 / 5, Directive 21 January 1943, International Law of the Air, 1939–1945, Confidential supplement to Air Power and War Rights, 1946, zitiert nach: Böhm, Martin: Die Royal Air Force und der Luftkrieg 1922–1945. Paderborn 2015, S. 147.

Baedeker Raids. The story of the historic towns and cities in Britain targeted by the German Air Force in Spring 1942. Hg. von Imperial War Museums. London 05.07.2019–05.01.2020. In: Google Arts and Culture. Virtuelle Ausstellung. 2021. [[online](#)]

Abbildungsverzeichnis

Abb. 1: ›Kopfteil‹ (Düren), Ausschnitt [The Bomber's Baedeker. Guide to the economic importance of German towns and cities, London 1944, S. 176. Public Domain Mark 1.0; [online](#)].

Abb. 2: ›Beschreibungsteil‹ (Düren), Ausschnitt [The Bomber's Baedeker. Guide to the economic importance of German towns and cities, London 1944, S. 176. Public Domain Mark 1.0; [online](#)].

Abb. 3: Klassische Visualisierung der Ziele nach höchster Priorität aus dem Bomber's Baedeker. [Uta Hohn: The Bomber's Baedeker – Target Book for Strategic Bombing in the Economic warfare against German Towns 1943–1954. In: Hohn 1994, S. 216.]

Abb. 4: Visualisierung der im Bomber's Baedeker benannten Zielstädte basierend auf den angegebenen zeitgenössischen Geokoordinaten mittels DARIAH-DE Geo-Browser und Tableau. [Bach / Secco 2021]

Beitrag aus:

Fabrikation von Erkenntnis – Experimente in den Digital Humanities (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5)

Titel:

Best Practices für die Gestaltung virtueller Museumsräume

Autor*in:

Vera Piontkowitz

Kontakt: vera.piontkowitz@uni-leipzig.de

Institution: Universität Leipzig

GND: [1237790166](#) ORCID: [0000-0003-3605-3609](#)

Autor*in:

Manuel Burghardt

Kontakt: burghardt@informatik.uni-leipzig.de

Institution: Universität Leipzig

GND: [1237665523](#) ORCID: [0000-0003-1354-9089](#)

DOI des Artikels:

[10.17175/sb005_005_v2](https://doi.org/10.17175/sb005_005_v2)

Nachweis im OPAC der Herzog August Bibliothek:

[185820383X](#)

Erstveröffentlichung:

19.11.2021

Version 2.0:

01.09.2023

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise:

10.02.2023

Format:

PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:

[Heuristik](#) | [Informationswissenschaft](#) | [Usability](#) | [Virtual environment](#) | [Virtuelles Museum](#) |

Empfohlene Zitierweise:

Vera Piontkowitz / Manuel Burghardt: Best Practices für die Gestaltung virtueller Museumsräume. In: Manuel Burghardt / Lisa Dieckmann / Timo Steyer / Peer Trilcke / Niels Walkowski / Joëlle Weis / Ulrike Wuttke (Hg.): Fabrikation von Erkenntnis – Experimente in den Digital Humanities (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5). Wolfenbüttel 2021–2022. 19.11.2021. Version 2.0 vom 20.08.2023. HTML / XML / PDF.

DOI: [10.17175/sb005_005_v2](https://doi.org/10.17175/sb005_005_v2)

Änderungen in Version 2.0 (01.09.2023):

Es wurden folgende Änderungen vorgenommen: Inhaltliche und formale Anpassungen, wie sie von den Gutachten angemerkt worden sind. Aktualisierung und Ergänzung der bibliografischen Angaben.

Vera Piontkowitz, Manuel Burghardt

Best Practices für die Gestaltung virtueller Museumsräume

Abstracts

Die Covid-19-Pandemie belegt eindrucksvoll die Relevanz von virtuellen Museumsräumen, welche den Zugang zu Kunst- und Kulturartefakten auch in Zeiten physischer Einschränkungen ermöglichen. Neben der unstrittigen Relevanz solcher virtueller Angebote steht als weiteres wichtiges Kriterium deren Akzeptanz, die im Wesentlichen von Kriterien wie Usability und User Experience abhängt. Bislang gibt es für die Umsetzung von benutzerfreundlichen Virtual Environments (VE) nur generische Design-Guidelines, wie etwa die *Heuristiken-Sammlung* von Sutcliffe und Gault 2004. In diesem Beitrag untersuchen wir im Rahmen einer heuristischen Evaluation die Anwendbarkeit dieser allgemeingültigen VE-Heuristiken auf virtuelle Museumsräume. Dazu wählen wir sechs exemplarische virtuelle Ausstellungen aus und evaluieren sie bzgl. der bestehenden Heuristiken. Im Ergebnis entstehen so einerseits Best Practices zur konkreten Umsetzung virtueller Museumsräume. Andererseits diskutieren wir anhand konkreter Beispiele die Grenzen der bestehenden Heuristiken. In der Folge benennen wir zwei Heuristiken, die für den Anwendungsbereich virtueller Museen nicht geeignet scheinen. Weiterhin ergänzen wir zwei Heuristiken, welche aus den Anwendungsbeispielen abgeleitet werden konnten und argumentieren für einen weiteren Ausbau spezifischer Heuristiken und Best Practices für virtuelle Museumsräume, um so systematisch die Wissensvermittlung im Sinne des *Public-History*-Ansatzes zu verbessern.

The Covid-19 pandemic demonstrates the relevance of virtual museums, which provide access to art and cultural heritage even in times when museums are closed. Besides their undisputable relevance, another important factor is the acceptance of virtual museums, which primarily depends on usability and user experience. So far, there are only generic guidelines for the design of user-friendly virtual environments, such as the twelve heuristics presented by Sutcliffe and Gault 2004. This paper presents an evaluation of those VE heuristics and their applicability to virtual museums. To do so, we select six exemplary virtual exhibitions and evaluate them with respect to the existing heuristics. Based on the results, we present best practices for the design of virtual museums. Furthermore, we discuss limitations of the existing heuristics. Two existing heuristics do not seem to be suitable for virtual museums. In addition, we complement two heuristics which could be derived from the examples and argue for a further development of domain-specific heuristics and best practices for virtual museums in order to systematically improve knowledge transfer in the sense of the *public history* approach.

1. Zur Relevanz virtueller Museen und Ausstellungen

Die Fotografie hat die Malerei nicht ersetzt, Instagram hat die Fotografie nicht platt gemacht und virtuelle Museumsrundgänge werden nicht die Museen killen.¹

Mit diesen Worten verfehlet die Kuratorin und Kolumnistin Anika Meier die Rolle virtueller Museumsrundgänge und verteidigt sie vor all jenen Kritiker*innen, die behaupten, eine virtuelle Museumserfahrung könne niemals den realen Besuch ersetzen. Diese Sichtweise sei laut Meier ein großes Missverständnis: Bei virtuellen Rundgängen gehe es nicht darum, Museen zu ersetzen, sondern darum, einen neuen Zugang zu Kunst und Kulturgütern zu schaffen.

Im deutschsprachigen Raum wird seit den 1990er Jahren über eine einheitliche Definition des *virtuellen Museums* diskutiert.² Während der Begriff oft für begleitende oder ergänzende Angebote physischer Museen verwendet wird, definiert das *Virtual Multimodal Museum* (ViMM) virtuelle Museum in Abgrenzung zum physischen Museum folgendermaßen: »A virtual museum is a digital entity that draws on the characteristics of a museum, in order to complement, enhance, or augment the museum through personalization, interactivity, user experience and richness of content.«³ Bei dieser Definition ist besonders hervorzuheben, dass das virtuelle Museum unabhängig vom physischen Museum existieren und damit als ein eigenständiges (und nicht nur als unterstützendes oder ergänzendes) Angebot fungieren kann. Angelehnt an diese Definition werden im Rahmen der vorliegenden Studie unter virtuellen Museen oder Museumsrundgängen sämtliche Angebote online zugänglicher Ausstellungen von Kunst und Kulturgütern verstanden, die als eigenständiges Angebot fungieren, wobei die Präsentation der Exponate und Informationen vielfältig ist und von einfachen Fotografien bis hin zu videospieldartigen 360-Grad-Erfahrungen reicht. Bei virtuellen Museen liegt der Fokus in Abgrenzung zu beispielsweise Online-Sammlungen in der Vermittlung, weniger in der Erforschung oder Bewahrung

¹ Meier 2020.

² Vgl. Biedermann 2021; Niewerth 2020.

³ Polycarpou 2018.

von Exponaten.⁴ Weiterhin ist der Begriff des virtuellen Museums nicht auf die Institution ›Museum‹ beschränkt – auch andere Gedächtnisinstitutionen wie Bibliotheken, Archive und Galerien (sogenannte *GLAM-Institutionen*) bieten derartige Ausstellungen an.

Bereits lange bevor die Covid-19-Pandemie viele Museen und Kulturinstitutionen dazu zwang, ihre Tore für Besucher*innen zu schließen,⁵ wurde das Internet als Ausstellungsplattform von Museen genutzt. Die Pandemie zeigt jedoch stärker und dringlicher denn je den Bedarf, aber auch die Potenziale virtueller Rundgänge auf. Neben einer hohen Zugänglichkeit durch die Verfügbarmachung von Exponaten im Internet sowie die Überwindung geografischer und zeitlicher Barrieren bieten virtuelle Museen die Möglichkeit, die Museumserfahrung interaktiv zu gestalten und Nutzer*innen beispielsweise durch digitales Storytelling zu aktivieren. Virtuelle Museen haben keine Einschränkungen, was die Größe der Ausstellungsräume betrifft und ermöglichen Kurator*innen und Museumsschaffenden, Räumlichkeiten nach den unterschiedlichen Bedürfnissen der Nutzer*innen zu gestalten und neuartige Präsentationsformen von Exponaten zu testen und umzusetzen. Weiterhin lassen sich fremde Exponate als Digitalisate in eine virtuelle Ausstellung einbinden. Nicht zuletzt können historische Stätten und zerstörte Orte, die nicht besucht werden können, für Nutzer*innen rekonstruiert und zugänglich gemacht werden.⁶ Beispielhaft für die Relevanz virtueller Museen steht hier das Virtual Museum of Iraq: Dieses virtuelle Museum entstand aus der Not heraus, als während des Irakkrieges im Jahr 2003 das Gebäude des irakischen Nationalmuseums erheblichen Schaden nahm und unzählige Artefakte gestohlen wurden. Im Rahmen des Projektes des Consiglio Nazionale delle Ricerche,⁷ des italienischen Außenministeriums und des irakischen Nationalmuseums wurden einige Exponate der Sammlung digitalisiert und in einen virtuellen Raum eingebunden, in dem Nutzer*innen 3D-Modelle der Artefakte, ergänzt durch Audio- und Videoinhalte, betrachten konnten. Das Projekt trug maßgeblich dazu bei, das historisch-archäologische Erbe des Iraks für die ganze Welt zugänglich zu machen, als das Museum selbst für Besucher*innen schließen musste. Das Virtual Museum of Iraq wurde so zu einem wichtigen Element der Sicherung des kulturellen Erbes des Landes.⁸

Während die Relevanz virtueller Museumsräume spätestens seit Covid-19 außer Frage steht,⁹ so ist aus Perspektive der *public history* für den Aspekt der Wissensvermittlung ein weiterer zentraler Faktor die *Akzeptanz*.¹⁰ Die Akzeptanz virtueller Ausstellungen hängt dabei wesentlich von deren Gestaltung unter Berücksichtigung des Kriteriums der *Usability* ab. Immer dann, wenn wir das Gefühl haben, dass ein Produkt oder eine Software einfach zu benutzen ist und vielleicht sogar Spaß macht, ist dies auf die inhärente Usability zurückzuführen. Gute Usability ist typischerweise das Ergebnis vieler unterschiedlicher Faktoren und deswegen nicht immer klar benennbar.¹¹ Deshalb schlägt etwa Nielsen vor, Usability nicht als abstraktes, eindimensionales Konzept zu behandeln, sondern in mehrere konkrete Komponenten wie Erlernbarkeit, Effizienz, Einprägsamkeit, Fehlerquote und Zufriedenheit zu zerlegen.¹² Auf Basis dieser einzeln benennbaren Komponenten wurden im Bereich der *human computer interaction* (HCI) in der Vergangenheit unterschiedliche Guidelines, Heuristiken und Best Practices abgeleitet, um die Erstellung von Benutzeroberflächen mit einer möglichst hohen *Usability* und *User Experience* zu unterstützen. Dabei gibt es einerseits generische Guidelines, wie beispielsweise die Usability-Heuristiken von Nielsen¹³ oder die *acht goldenen Regeln* des Interface-Designs von Shneiderman et al.¹⁴ Andererseits gibt es eine ganze Reihe von domänen-spezifischen, angepassten Varianten von Design-Guidelines und Heuristiken, etwa für die Bereiche *mobile design* oder *e-commerce*.¹⁵

Auch für die Domäne virtueller Ausstellungen finden sich Best Practices im Bereich der Benutzeroberflächen: Beispielsweise untersucht Johnson verschiedene Präsentationsformen von 3D-Objekten im Internet und leitet aus der Evaluation von fünf Projekten Best Practices ab, die Museen und andere Gedächtnisinstitutionen dabei unterstützen sollen, 3D-Objekte online frei verfügbar zu machen.¹⁶ Vergleichbare Empfehlungen für die Umsetzung von Desktop-VR-Anwendungen im musealen Bereich sucht man bislang jedoch vergebens. Dabei liegt es nahe, generische Heuristiken aus dem übergeordneten Feld der VR-User-Interfaces¹⁷ für den Bereich virtueller Museumsräume zu adaptieren.

⁴ Vgl. Deutscher Museumsbund e. V. 2017.

⁵ Vgl. Späth 2020.

⁶ Vgl. Lugin et al. 2018, S. 1; Styliani et al. 2009, S. 524.

⁷ Zu Deutsch: Nationaler Forschungsrat.

⁸ Vgl. Parrinello / Dell'Amico 2019, S. 1033–1034.

⁹ Tatsächlich wird die wichtige Rolle virtueller Museumsräume und deren schrittweise Integration in das Standardangebot von Museen und verwandten Kultureinrichtungen bereits wesentlich früher beschrieben (vgl. Catton / Smith 2021; Schweibenz 2019) und durch die Covid-19-Pandemie weiter zementiert.

¹⁰ Vgl. Kelley 1978.

¹¹ Vgl. Barnum 2011, S. 1.

¹² Vgl. Nielsen 1993, S. 26.

¹³ Vgl. Nielsen 1994.

¹⁴ Vgl. Shneiderman et al. 2018.

¹⁵ Vgl. Joyce / Lilley 2014.

¹⁶ Vgl. Johnson 2016.

¹⁷ Beispielsweise Sutcliffe / Gault 2004.

In unserem Beitrag verwenden wir zwölf Heuristiken von Sutcliffe und Gault,¹⁸ um bestehende virtuelle Ausstellungen zu evaluieren und Best Practices für deren Umsetzung abzuleiten. Das Vorgehen orientiert sich dabei am Prinzip einer heuristischen Evaluation,¹⁹ d. h. es werden sechs exemplarische virtuelle Ausstellungen systematisch exploriert und gegen die genannten Heuristiken evaluiert. Im Ergebnis entstehen so Best Practices für den Anwendungsbereich virtueller Ausstellungs- und Museumsräume, welche künftig für die Evaluation bestehender Angebote und vor allem auch für die Umsetzung neuer Ausstellungen eingesetzt werden können.

2. Methodisches Vorgehen

In diesem Abschnitt wird das methodische Vorgehen bei der heuristischen Evaluation dargestellt. Dabei werden zum einen die zugrundeliegenden Heuristiken und zum anderen die evaluierten Projekte näher erläutert. Zudem sollen vorab einige grundlegende Begrifflichkeiten im Kontext des (3D-)User-Interface-Designs kurz erörtert werden.

2.1 Grundbegriffe des User-Interface-Designs

Während ein *User Interface* (UI) das Medium beschreibt, mithilfe dessen die Kommunikation zwischen Nutzer*in und Computer stattfindet, werden Interaktionen in einem 3D-User-Interface in einer virtuellen 3D-Umgebung bewältigt. Eine 3D-Welt, welche aus einer First-Person-Perspektive erkundet wird, bezeichnet man als *Virtual Environment* (VE). Ein VE, welches nicht mithilfe eigens für den Zweck entwickelter Ein- und Ausgabegeräte (beispielsweise Controller und VR-Headset) betreten wird und stattdessen den Computerbildschirm als Fenster in die virtuelle Realität nutzt, wird als eine *Desktop-VR-Anwendung* bezeichnet.²⁰ Solche Desktop-VR-Anwendungen, die von Museen, aber auch anderen Institutionen frei im Internet zugänglich gemacht werden, sind Gegenstand der nachfolgenden Evaluationsstudie.

2.2 Ausgangslage: Heuristiken nach Sutcliffe und Gault

Die Ermittlung von Best Practices im Anwendungsgebiet virtueller Museen stützt sich zunächst auf allgemeine Gestaltungsgrundsätze für (3D-)User-Interfaces. Solche *Heuristiken*, also Gestaltungsgrundsätze bzw. Usability-Prinzipien für Benutzeroberflächen, dienen deren Evaluation.²¹ Basierend auf der Methode der *Usability Inspection*,²² bei der User Interfaces von Expert*innen getestet werden, ist die *heuristische Evaluierung* eine Methode, bei der überprüft wird, ob eine Benutzeroberfläche mit allgemeinen Heuristiken der entsprechenden Domäne übereinstimmt und welche Probleme sich bei der Benutzung eines Interfaces ergeben könnten.²³ Solche Heuristiken wurden mit dem Hinweis, diese für explizite Anwendungsgebiete anzupassen und zu ergänzen, von Nielsen und Molich und in überarbeiteter Form von Nielsen veröffentlicht.²⁴ Diese Heuristiken beziehen sich ausschließlich auf 2D-User-Interfaces und sind deshalb nicht oder nur teilweise für die Nutzung in anderen Anwendungsszenarien geeignet.

Die Heuristiken dienen trotz ihres weit zurückliegenden Publikationsdatums in vielen Anwendungsgebieten nach wie vor als Grundlage für die Erstellung von neuen, domänenspezifischen Heuristiken, welche die Eigenheiten und spezifischen Probleme des jeweiligen Anwendungsfelds berücksichtigen, beispielsweise für die sprachliche Interaktion von Mensch und Maschine oder für Augmented-Reality-Anwendungen für das Smartphone.²⁵ Für die Gestaltung und Evaluation von VEs präsentieren Sutcliffe und Gault, angelehnt an Nielsen, zwölf Heuristiken (vgl. Tabelle 1).

Wengleich die Heuristiken nach Sutcliffe und Gault für die Evaluation von VEs bereits vor über 15 Jahren veröffentlicht wurden, erweisen sie sich nach wie vor als aktuell: Beispielsweise dienen sie der Evaluation verschiedener italienischer virtueller Museumsrundgänge²⁶ oder der Entwicklung einer VR-Anwendung zur Behandlung von Nyktophobie²⁷.

¹⁸ Vgl. Sutcliffe / Gault 2004.

¹⁹ Vgl. Nielsen 1992.

²⁰ Vgl. LaViola et al. 2017, S. 6–8.

²¹ Vgl. Nielsen / Molich 1990.

²² Vgl. Nielsen / Mack (Hg.) 1994.

²³ Vgl. Nielsen 1994, S. 25–26.

²⁴ Vgl. Nielsen / Molich 1990, S. 339; Nielsen 1994, S. 30.

²⁵ Vgl. Almeida Pacheco et al. 2019; Murad et al. 2019.

²⁶ Vgl. Kabassi et al. 2019.

²⁷ Vgl. Paulus et al. 2019.

Nr.	Name	Beschreibung
1	<i>Natural engagement</i>	Interaktionen sollen sich den Erwartungen der Nutzer*innen an die Interaktion in der realen Welt so weit wie möglich annähern. Die Interpretation dieser Heuristik hängt von der Anforderung der Natürlichkeit der Anwendung sowie dem Gefühl der Präsenz und des Interesses der Nutzer*innen ab. ²⁸
2	<i>Compatibility with the user's task and domain</i>	Das VE und das Verhalten von Objekten sollen so genau wie möglich den Erwartungen der Nutzer*innen an Objekte der realen Welt, ihrem Verhalten und den Möglichkeiten für die Durchführung von Aufgaben entsprechen.
3	<i>Natural expression of action</i>	Die Darstellung des Selbst im VE soll es den Nutzer*innen ermöglichen, auf natürliche Weise zu handeln und zu erforschen und normale körperliche Bewegungen nicht einschränken. Diese Gestaltungsqualität kann durch die verfügbaren Geräte eingeschränkt sein. Wenn das haptische Feedback fehlt, leidet unweigerlich der natürliche Ausdruck.
4	<i>Close coordination of action and representation</i>	Die Darstellungen des im VE manifestierten Selbst sollen den Handlungen der Nutzer*innen getreu sein. Die Reaktionszeit zwischen einer Bewegung der Nutzer*innen und der Aktualisierung der Anzeige des VE soll weniger als 200 ms betragen, um Probleme mit <i>Cybersickness</i> ²⁹ zu vermeiden.
5	<i>Realistic feedback</i>	Die Auswirkungen der Aktionen der Nutzer*innen auf Objekte der virtuellen Welt sollen sofort sichtbar sein und den physikalischen Gesetzen sowie den die Wahrnehmung betreffenden Erwartungen der Nutzer*innen entsprechen.
6	<i>Faithful viewpoints</i>	Die visuelle Darstellung der virtuellen Welt soll der normalen Wahrnehmung der Nutzer*innen entsprechen, und der Blickpunktwechsel durch Kopfbewegungen soll ohne Verzögerung wiedergegeben werden.
7	<i>Navigation and orientation support</i>	Für Nutzer*innen soll es immer möglich sein, ihre Position im VE auszumachen und zu bekannten, voreingestellten Positionen zurückzukehren. Unnatürliche Aktionen, wie beispielsweise das Durchfliegen eines Raumes, können hilfreich sein, müssen aber in einem Kompromiss mit der Natürlichkeit beurteilt werden (Heuristik 1 und Heuristik 2).
8	<i>Clear entry and exit points</i>	Die Mittel zum Betreten und Verlassen einer virtuellen Welt sollen klar kommuniziert werden.
9	<i>Consistent departures</i>	Wenn Designkompromisse eingegangen werden, sollen sie konsistent und klar gekennzeichnet sein.

Tab. 1 : Heuristiken nach Sutcliffe und Gault. [Sutcliffe / Gault 2004]

²⁸ Die Beschreibungen der Heuristiken wurden von den Verfasser*innen mithilfe der Übersetzungssoftware **DeepL** aus dem Englischen übersetzt.²⁹ Cybersickness oder die VR-Krankheit meint ein Unwohlsein, das beim Nutzen virtueller Umgebungen auftreten kann. Sie tritt auf, wenn eine Aktion der Nutzer*innen nicht mit dem (visuellen) Feedback der Anwendung übereinstimmt (vgl. LaViola et al. 2017, S. 425).

10	<i>Support for learning</i>	(Inter-)Aktive Objekte sollen markiert werden und sich gegebenenfalls selbst erklären, um das Erlernen von VEs zu fördern.
11	<i>Clear turn-taking</i>	Wo Systeminitiativen eingesetzt werden, sollen diese klar signalisiert, weiterhin sollten Konventionen für das turn-taking festgelegt werden.
12	<i>Sense of presence</i>	Die Präsenz im virtuellen Raum sollte sich für Nutzer*innen so natürlich und real wie möglich anfühlen.

Tab. 1 : Heuristiken nach Sutcliffe und Gault. [Sutcliffe / Gault 2004]

Die grundlegende Eignung der in Tabelle 1 vorgestellten Heuristiken wurde anhand zweier Fallstudien, einerseits mit einem Experten für Mensch-Computer-Interaktion, andererseits mit einer Gruppe Bachelorstudierender, evaluiert.³⁰ Wenngleich die Einschätzung der Begutachtenden grundlegend positiv war, so ergaben sich in der Gesamtschau einige weitere Erkenntnisse für die Optimierung der Heuristiken. Vor der konkreten Anwendung der Heuristiken sollten diese, angepasst an das jeweilige VE, gefiltert werden. Die Autoren schlagen dafür folgende Einteilung vor:³¹

- Heuristiken 1, 2, 3, 4, 5, 6, 7 und Heuristik 12 sind elementare Heuristiken, welche in jedem Fall für die Evaluation von VEs hinzugezogen werden können.
- Heuristiken 8 und 9 hingegen eignen sich nur dann, wenn es sich beim VE um eine Desktop-VR-Anwendung handelt.
- Die Einbindung von Heuristik 10 hängt von der Komplexität und Größe des VE ab.
- Heuristik 11 eignet sich nur für kollaborative VEs, also solche VEs, in denen sich mehrere Nutzer*innen gemeinsam bewegen.

Die Evaluation zu den Heuristiken nach Sutcliffe und Gault zeigt, dass die Sammlung zwar viele wichtige Aspekte abbildet, allerdings nicht gleichermaßen für die ganze Bandbreite von VE-Szenarien eingesetzt werden kann. Im Rahmen dieses Beitrags dienen die Heuristiken deshalb zunächst als Ausgangspunkt, um bestehende virtuelle Ausstellungen systematisch zu evaluieren und auf dieser Basis Best Practices für deren benutzerfreundliche Gestaltung abzuleiten.

2.3 Heuristische Evaluation

Für die Erarbeitung von allgemeingültigen Best Practices zur Entwicklung und Gestaltung virtueller Museumsräume werden ausgewählte VEs aus dieser Domäne genauer betrachtet und Beispiele zur konkreten, positiven Umsetzung der Heuristiken³² präsentiert. Der Erkenntnisgewinn stützt sich dabei auf ein deduktives Vorgehen, angelehnt an die heuristische Evaluation nach Nielsen.³³ Die Heuristiken werden dabei als Schablone auf die Anwendungen gelegt, woraufhin untersucht wird, ob und wie diese realisiert wurden. Mithilfe textueller Beschreibungen werden die Möglichkeiten zur Umsetzung der Heuristiken diskutiert und anschließend für jede Heuristik zusammengefasst. Unterstützend wird für die Beurteilung der Ansätze der Anwendungen einschlägige Literatur zu 3D-UIs hinzugezogen. Weiterhin wird die Relevanz jeder Heuristik für die gegebene Domäne diskutiert.

2.4 Projektauswahl

Bei der Erarbeitung von Best Practices wurden insgesamt sechs VEs aus dem GLAM-Bereich ausgewählt. Die Projekte wurden weiterhin so ausgewählt, dass möglichst vielfältige, unterschiedliche Ansätze zur Umsetzung virtueller Museumsräume abgedeckt werden. Konkret wurde bei der Wahl der Projekte auf Diversität im Hinblick auf die nachfolgenden Selektionskriterien geachtet:

- *Gesamtansatz*: eher klassisch vs. eher experimentell
- *Grad der Immersion*: eher immersiv vs. weniger immersiv
- *Bewegung im Raum*: eher frei vs. eher eingeschränkt
- *Interaktionsmöglichkeiten*: eher viel vs. eher wenig

³⁰ Vgl. Sutcliffe / Gault 2004, S. 837–845.

³¹ Vgl. Sutcliffe / Gault 2004, S. 846.

³² In der nachfolgenden Bearbeitung werden unter dem Begriff Heuristiken, wenn nicht anders gekennzeichnet, die Heuristiken nach Sutcliffe und Gault verstanden.

³³ Vgl. Nielsen 1994.

Alle ausgewählten Projekte sind Desktop-VR-Anwendungen, die frei im Internet zur Verfügung stehen. Um die Projekte nutzen zu können, brauchen die Nutzer*innen also lediglich einen Computer mit Bildschirm, Tastatur und Maus und einem Audioausgang. Zum Teil unterstützen die Anwendungen zwar die Nutzung von VR-Headsets, dies wurde in der Analyse aber nicht weiter berücksichtigt, da die Nutzung der Anwendungen eine hohe Zugänglichkeit gewährleisten soll. Nachfolgend werden die Projekte kurz vorgestellt. Die Beschreibungen basieren dabei, wenn nicht anders gekennzeichnet, auf der eigenen Betrachtung und Nutzung der Anwendungen. Tabelle 2 zeigt eine Übersicht über die ausgewählten Anwendungen sowie eine Aufschlüsselung der Selektionskriterien.

			Selektionskriterien			
Name	Kurzname ³⁴	Veröffentlichungsjahr	Gesamtansatz	Grad der Immersion	Bewegung im Raum	Interaktionsmöglichkeiten
WDR Zeitkapsel ³⁵	Zeitkapsel	2018	Eher experimentell	Eher immersiv	Eher eingeschränkt	Eher viel
Google Arts and Culture ³⁶	Google	2011	Eher klassisch	Wenig immersiv	Eher frei	Eher wenig
Nach Berlin ³⁷	Berlin	2020	Eher experimentell	Wenig immersiv	Eher eingeschränkt	Eher wenig
Walter's Cube ³⁸	Walter	2016	Eher klassisch	Eher immersiv	Eher frei	Eher wenig
Digitale Kunsthalle ³⁹	Kunsthalle	2019	Eher klassisch	Eher immersiv	Eher frei	Eher viel
Museum Schnütgen ⁴⁰	Schnütgen	2020	Eher klassisch	Wenig immersiv	Eher frei	Eher wenig

Tab. 2 : Übersicht über die zu untersuchenden Projekte. [Piontkowitz / Burghardt 2021]

1. Projekt: *WDR-Zeitkapsel*

Die WDR-Zeitkapsel ist ein Projekt, welches den Nutzer*innen eine virtuelle Zeitreise ermöglichen soll. Es besteht die Möglichkeit, ein Wohnzimmer im Jahr 1968 zu besuchen oder an Bord eines Flugzeuges im Jahr 1960 zu reisen. Die virtuellen Orte bieten den Nutzer*innen eine 360°-Perspektive mit vielen Interaktionsmöglichkeiten. Bei bestimmten Interaktionen werden innerhalb der Umgebung kurze Videosequenzen abgespielt, die in die 360°-Erfahrung eingebunden sind. Der WDR bietet darüber hinaus Unterrichtsmaterialien für Lehrer*innen, um die Zeitkapsel in den Schulunterricht einzubinden.⁴¹

2. Projekt: *Google Arts and Culture*

In diesem Projekt von Google werden virtuelle Rundgänge für eine Vielzahl von Museen und historischen Stätten angeboten. Einerseits haben Nutzer*innen die Möglichkeit, Museen und historische Stätten in der Funktionsweise von Google Street View zu besuchen, andererseits gibt es ein großes Angebot hochauflösender Fotografien von Exponaten, die in eigens für dieses Medium konzipierten Ausstellungen eingebunden sind.

3. Projekt: *Nach Berlin*

Die virtuelle Ausstellung *Nach Berlin* ist ein Projekt zum 75. Jahrestag des Kriegsendes. Die Ausstellung bietet 360°-Ansichten an unterschiedlichen Orten, darunter der Reichstag, das Brandenburger Tor, das ehemalige KZ Sachsenhausen sowie der Alexanderplatz. Bei der 360°-Erfahrung werden Bilder der Orte, wie sie heute aussehen, vermischt mit historischen Bildern vom Kriegsende. Audiosequenzen, Bilder und kurze Texte versorgen die Nutzer*innen mit Informationen zu unterschiedlichen Aspekten des Krieges und des Kriegsendes.

4. Projekt: *Walter's Cube*

Walter's Cube bietet Galerien und Museen die Möglichkeit, Ausstellungen als 3D-Räume zu erfassen und auf verschiedenen Plattformen zu veröffentlichen. Nutzer*innen können sich frei in den virtuellen Räumen von Walter's Cube bewegen und unterschiedliche Kunstwerke wie Malereien oder 3D-modellierte Skulpturen betrachten. Das Ziel der Organisation ist es, zeitliche und geographische Barrieren abzubauen und Ausstellungen frei zugänglich zu machen.

5. Projekt: *Digitale Kunsthalle*

³⁴ Um die einzelnen Projekte in der nachfolgenden Ableitung von Best Practices zu referenzieren, werden die in Tabelle 2 eingeführten Kurznamen verwendet.

³⁵ *WDR Zeitkapsel*, WDR (Hg.) 2018.

³⁶ *Google Arts & Culture*, Google (Hg.) 2021.

³⁷ *Nach Berlin. 75 Jahre Kriegsende*, Kulturprojekte Berlin (Hg.) 2020.

³⁸ *Online Viewing Room*, Walters Cube (Hg.) 2020.

³⁹ *Digitale Kunsthalle*, ZDF (Hg.) 2016.

⁴⁰ *Rundgang, Museum Schnütgen*, Stadt Köln (Hg.) 2020

⁴¹ Vgl. WDR 2018.

Die *digitale Kunsthalle* ist ein Projekt des ZDF, das wechselnde Ausstellungen in einem virtuellen Museum zeigt. Gezeigt werden neben Bildern und Malereien auch 3D-modellierte Objekte, welche durch die freie Bewegung im Raum und das Heranzoomen mit dem Mausrad genauer betrachtet werden können. Eine Einbindung von Informationstexten und Videos vermittelt den Nutzer*innen weitere Hintergründe zur Ausstellung und ihren Exponaten.

6. Projekt: *Museum Schnütgen*

Die Dauerausstellung des Kunstmuseums Museum Schnütgen ist als virtueller 360°-Rundgang online zugänglich. Nutzer*innen können sich an bestimmten Punkten in der Ausstellung in einer 360°-Perspektive umsehen und sich mit einem der Audioguides (für Kinder oder für Erwachsene) über die Ausstellung und die Exponate informieren. Zusätzlich stehen für eine Vielzahl der Exponate Fotografien in guter Qualität zur Verfügung, die in die virtuelle Ausstellung eingebunden sind.

3. Best Practices für die Gestaltung virtueller Museumsräume

In diesem Abschnitt stellen wir die Ergebnisse der heuristischen Evaluation von den genannten virtuellen Museumsräumen vor und leiten jeweils Best Practices für die konkrete Umsetzung einzelner Heuristiken ab. Dabei diskutieren wir die Anwendbarkeit und mögliche Interpretationsspielräume der bestehenden VE-Heuristiken für das Szenario musealer Räume. In Ergänzung zu den nachfolgenden Heuristiken sollten virtuelle Ausstellungen grundsätzlich einen möglichst barrierefreien Zugang haben. Gemeint ist damit beispielsweise die Verfügbarmachung von Untertiteln für auditive Elemente und Vorlesefunktionen für textuelle Informationen, aber auch die Möglichkeit, Texte in leichter Sprache anzeigen zu lassen.

3.1 Natürlichkeit der Anwendung (Natural engagement)

Bei der Untersuchung der ausgewählten Projekte auf ihre Natürlichkeit und Nähe zur realen Welt stellt sich die Frage, welcher Anspruch an Natürlichkeit bei der Gestaltung eines virtuellen Museums gestellt werden sollte. Hier bietet sich ein Vergleich zu einer anderen Domäne an: Ist das VE beispielsweise eine Trainingssoftware für Chirurgie, so sollten sich die Aktionen und Interaktionen von Nutzer*innen klar an der realen Welt orientieren und diese nachahmen. Virtuelle Museen dagegen haben niedrigere Anforderungen an die Natürlichkeit, da hier nicht der Anspruch besteht, eine möglichst reale Museumserfahrung zu imitieren, sondern vielmehr die Vorteile und Stärken des interaktiven Mediums genutzt werden sollen. Dieser Kompromiss wird von Sutcliffe und Gault auch betont,⁴² indem sie verdeutlichen, dass die Interpretation dieser Heuristik von den Anforderungen an die Natürlichkeit der Anwendung abhängig gemacht werden soll. In der zu betrachtenden Domäne sollte nur dann eine Nähe zur realen Welt aufgebaut werden, wenn dies auch direkte Vorteile für die Nutzung der Anwendung mit sich bringt. Es lassen sich bestimmte Funktionen festmachen, die einerseits das Maß der Natürlichkeit der Anwendung erhöhen und andererseits auch zur intuitiven Nutzung der Anwendung beitragen.

Wie in der *Kunsthalle* umgesetzt, ermöglicht die freie Bewegung im Raum den Nutzer*innen, sich ohne Einschränkungen umzusehen, verschiedene Positionen und Blickwinkel einzunehmen und so beispielsweise Skulpturen im Museum ganzheitlich betrachten zu können. Die Nutzer*innen können sich einerseits mit den Pfeiltasten fortbewegen, andererseits können sie mit der Maus einen beliebigen Punkt im Raum anklicken, um dorthin zu gelangen. Hält der*die Nutzende eine beliebige Maustaste gedrückt und bewegt die Maus, kann die Sicht in alle Richtungen verändert werden. Diese handgesteuerte Lenkung hat einerseits den Nachteil, dass die Nutzer*innen zwei Richtungen, also Blick- und Bewegungsrichtung, kontrollieren müssen, was die Lenkung komplizierter macht. Andererseits birgt diese Art der Lenkung den Vorteil, dass Blick- und Bewegungsrichtung nicht die gleiche sein müssen und sich durch die Kontrolle zweier Richtungen schneller ein räumliches Verständnis bei den Nutzer*innen bildet.⁴³

Zusammenfassung: Die Natürlichkeit der Anwendung ist keine hohe Priorität bei virtuellen Museen und sollte lediglich dann umgesetzt werden, wenn dies die Usability erhöht oder Möglichkeiten für das Vermittlungskonzept bietet. Durch die freie Bewegung im Raum mit einer handgesteuerten Lenkung können Nutzer*innen Kunstwerke und Exponate aus unterschiedlichen Blickwinkeln betrachten und das VE besser erkunden.

⁴² Vgl. Sutcliffe / Gault 2004, S. 833.

⁴³ Vgl. LaViola et al. 2017, S. 340–341.

3.2 Konformität mit der Anwendungsdomäne und den Nutzererwartungen (Compatibility with the user's task and domain)

Bei Betrachtung dieser Heuristik sollte vorerst definiert werden, welche Aufgaben Nutzer*innen potenziell in einem virtuellen Museum durchführen bzw. welche Ziele sie erreichen möchten, und wie sie diese Aufgaben in einem realen Museum durchführen würden. Deggim et al. nennen verschiedene Anwendungen für die Nutzung virtueller Museen:⁴⁴ den Zugang zu Informationen und Zusammenhängen, die Präsentation von Inhalten, wie sie gegebenenfalls mit klassischen Ausstellungsmethoden nicht möglich ist, die Entwicklung eines Interesses am Themengebiet sowie die Möglichkeit zeit- und ortsunabhängiger Museumsbesuche. Letzteres ist bei allen ausgewählten Projekten gegeben, da diese online zur Verfügung stehen. In besonderem Maße wird dies jedoch im *Walter-Projekt* umgesetzt. Zum einen handelt es sich hier um 3D-modellierte Ausstellungen, die nachträglich verändert werden können und deshalb nicht in einer Zeit »stehenbleiben«. Weiterhin macht sich die Firma Walter's Cube die Orts- und Zeitunabhängigkeit explizit zur Aufgabe: »Our mission is to eliminate geographic and temporal barriers between exhibitions and the audiences.«⁴⁵

Exponate betrachten zu können und sich über diese und darüberhinausgehende Zusammenhänge zu informieren, ist eine elementare Aufgabe, die in virtuellen Museen unbedingt durchführbar sein sollte. *Schnütgen* ermöglicht den Nutzer*innen die Bewältigung dieser Aufgaben, wobei eine starke Orientierung an realen Museumserfahrungen auffällt. Die Möglichkeit zur Informationsakquise wird hier durch unterschiedliche Medien gewährleistet: Nutzer*innen können sich mithilfe eines Audioguides durch die Ausstellung führen lassen. Ein zusätzlicher Audioguide für Kinder vermittelt die Inhalte in leicht verständlicher, kindgerechter Weise. Die Stationen für die Audioguides werden durch runde Buttons mit entsprechenden Symbolen gekennzeichnet. Wird der Button angeklickt, öffnet sich oben rechts im Bildschirm ein kleines Fenster, das den Titel der Tonspur, die Länge der Audiosequenz, die Länge der bisher angehörten Audiosequenz sowie einen Play- und Pause-Button anzeigt. Darüber hinaus sind ausgewählte Exponate mit Buttons versehen, die kennzeichnen, dass zu diesem Objekt eine hochauflösende Fotografie sowie ein Informationstext oder eine Beschreibung des Exponats zur Verfügung steht. Detailansichten, die mit Informationstexten versehen sind, finden sich auch im Projekt *Kunsthalle*. Hier klicken Nutzer*innen die Exponate an, wobei sich ein neues Fenster öffnet, welches entweder mit Fotografien der Exponate aus unterschiedlichen Blickwinkeln und passenden Informationstexten gefüllt ist oder die Nutzer*innen mit zusätzlichen Medien wie eigens für die Ausstellungen gedrehten Videos versorgt (vgl. Abbildung 1).

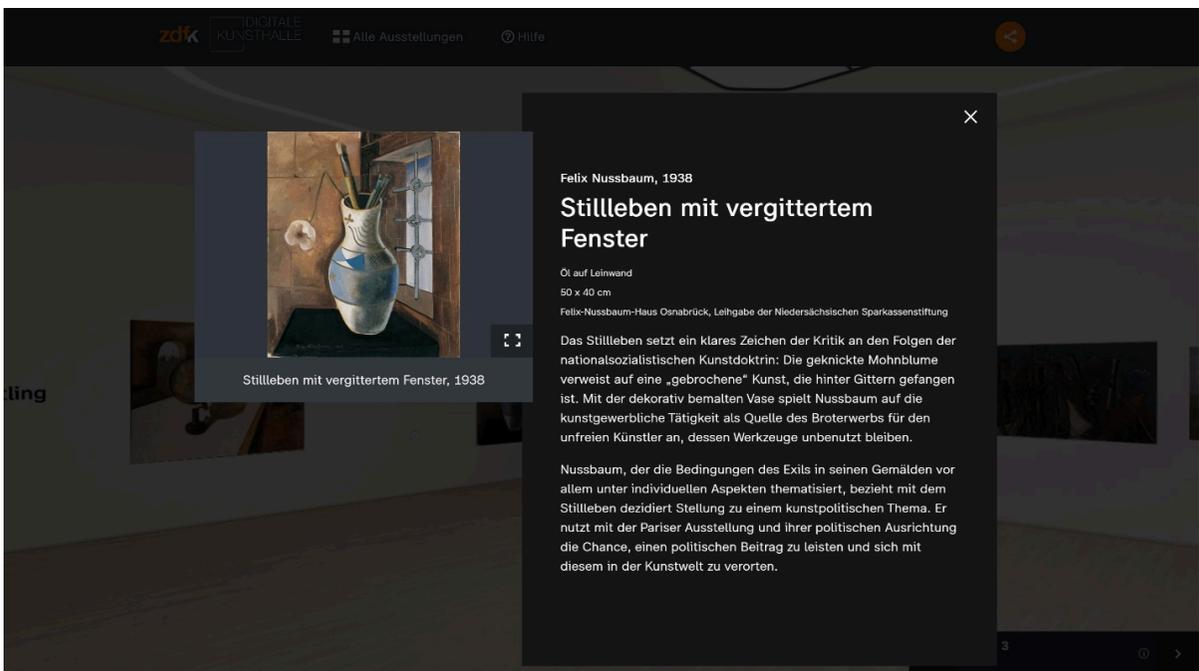


Abb. 1: Detailansicht eines Kunstwerks mit zugehörigem Informationstext im Projekt *Kunsthalle*. In: *Digitale Kunsthalle*. [ZDF (Hg.) 2016]

⁴⁴ Vgl. Deggim et al. 2017, S. 454–455.

⁴⁵ Walter's Cube (Hg.) 2020.

Weiterhin umfasst diese Heuristik auch das Verhalten von virtuellen Objekten, welches den Erwartungen der Nutzer*innen weitestgehend entsprechen sollte. Beispielhaft lässt sich dies im Projekt *Zeitkapsel* festmachen. Im Szenario *1968 – Schöner Wohnen* haben Nutzer*innen etwa die Möglichkeit mit Elektrogeräten in authentischer Weise zu interagieren. Beim Fernseher können Nutzer*innen ein Programm auswählen und sich dieses anschauen, das Radio bietet die gleiche Möglichkeit. Mit einem Klick auf den Plattenspieler können Nutzer*innen zwischen zwei Platten wählen und diese dann abspielen. In der digitalen Kunsthalle zeigt sich beispielhaft, wie multimediale Inhalte für die Nutzer*innen intuitiv eingebunden werden können: In der Ausstellung *Geraubte Kunst – Jüdische Sammlungen im Nationalsozialismus* befindet sich im zweiten Raum ein 3D-modellierter Schreibtisch, auf dem unterschiedliche Objekte präsentiert sind, unter anderem ein Laptop. Der Bildschirm des Laptops zeigt ein Standbild eines Interviews, welches mit einem Klick auf das Gerät abgespielt wird.

Zusammenfassung: Es zeigt sich, dass Detailansichten von Exponaten, versehen mit hochauflösenden Fotografien der Objekte und Informationsmedien unterschiedlicher Art, den Nutzer*innen deren Betrachtung erleichtert. Weiterhin bietet es sich im virtuellen Raum an, Gebrauchsobjekte entsprechend ihrer Verwendung in der realen Welt mit Interaktionen oder multimedialen Angeboten zu versehen.

3.3 Natürliche Interaktion (Natural expression of action)

Im Rahmen dieser Studie wurden ausschließlich Projekte mit einer hohen Zugänglichkeit ausgewählt, was u. a. bedeutet, dass für deren Nutzung keine speziellen VR-Ein- und Ausgabegeräte benötigt werden. Entsprechend erfahren die Nutzer*innen der Anwendungen auch kein haptisches Feedback und können die Anwendung nicht mit natürlichen Körperbewegungen wie Gehen oder Greifen bedienen, weshalb die Relevanz dieser Heuristik in den Hintergrund rückt. Sutcliffe und Gault nehmen die Einschränkungen durch die verfügbaren Ein- und Ausgabegeräte zur Kenntnis, beschreiben aber auch, wie durch eben diese Einschränkungen der natürliche Ausdruck leidet.⁴⁶ Diese Problematik scheint in der Museumsdomäne aber wenig relevant, da hier, wie oben bereits erläutert, nicht der Anspruch besteht, eine reale Museumserfahrung nachzuahmen.

Die natürliche Darstellung des Selbst kann jedoch auch ohne haptisches Feedback in Desktop-VR-Anwendungen angedeutet werden. LaViola et al. schlagen etwa vor, bei einer Einschränkung der zur Verfügung stehenden Geräte das haptische Feedback durch auditives und visuelles Feedback zu substituieren.⁴⁷ Beispielsweise wird den Nutzer*innen von *Walter's Cube* mit Gehgeräuschen ihre Bewegung verdeutlicht, was den natürlichen Ausdruck im VE fundiert. Auch im Projekt *Zeitkapsel* wird die Selbstwahrnehmung im VE durch auditives Feedback unterstützt. Obgleich die Nutzer*innen dieser Anwendung sich lediglich in einer 360°-Ansicht im Raum umsehen können, verändert sich beispielsweise die Lautstärke des Fernsehers im Szenario *1968 – Schöner Wohnen*, wenn sich die Nutzer*innen von diesem wegdrehen. Auditives Feedback unterstützt jedoch nicht nur die Selbstwahrnehmung, sondern kann auch förderlich für die Orientierung im VE sein. Dieser Aspekt wird in der Betrachtung von Heuristik 7: *Navigation and orientation support* näher beleuchtet. Auch die freie Bewegung im Raum, wie sie in der Kunsthalle umgesetzt ist und unter Heuristik 1: *Natural engagement* bereits erörtert wurde, trägt zum natürlichen Ausdruck und einer besseren Selbstwahrnehmung im VE bei.

Zusammenfassung: Insgesamt lässt sich festhalten, dass der natürliche Ausdruck eine*r Nutzer*in zwar durch die Ein- und Ausgabegeräte limitiert ist, sich haptisches Feedback aber durch auditives Feedback substituieren lässt und auch die freie Bewegung im Raum zu einer verbesserten Selbstwahrnehmung und Natürlichkeit beiträgt.

3.4 Schnelle Reaktionszeiten (Close coordination of action and representation)

Eine schnelle Reaktionszeit zwischen Aktionen der Nutzer*innen und der Anzeige dieser Aktionen soll vermeiden, dass Nutzer*innen unter Cybersickness leiden. Studien zeigen, dass Desktop-VR-Anwendungen jedoch deutlich seltener Cybersickness hervorrufen als VR-Headsets.⁴⁸ Dennoch verliert die Heuristik dadurch nicht an Relevanz, da eine schnelle Reaktionszeit insgesamt zur Usability der Anwendung beiträgt.⁴⁹ Reaktionszeiten hängen von unterschiedlichen Faktoren, wie beispielsweise der Serverleistung, ab und werden weniger vom Design eines VE, als vielmehr von dessen Implementierung beeinflusst.⁵⁰ Dennoch ist festzuhalten, dass mit neu aufkommenden Formaten, welche Nutzer*innen die Erstellung virtueller 3D-Räume und

⁴⁶ Vgl. Sutcliffe / Gault 2004, S. 833.

⁴⁷ Vgl. LaViola et al. 2017, S. 428.

⁴⁸ Vgl. Sharples et al. 2008; Srivastava et al. 2019.

⁴⁹ Vgl. LaViola et al. 2017, S. 111.

⁵⁰ Vgl. Yin et al. 2018.

die kollaborative Nutzung jener ermöglichen, die Performance auch von der Anzahl der eingebundenen Medien und der sich zeitgleich im virtuellen Raum befindenden Nutzer*innen abhängt.⁵¹ Da es sich bei den betrachteten virtuellen Museen nicht um solche Räume handelt, soll diese Heuristik für die weitere Bearbeitung nicht berücksichtigt werden.

Zusammenfassung: Diese Heuristik ist für Desktop-VR-Anwendungen, wie sie in dieser Studie untersucht werden, nicht relevant.

3.5 Realistisches Feedback (Realistic feedback)

Objekte im VE sollten in ihrem Verhalten physikalischen Gesetzen folgen und darauf bezogen den Erwartungen der Nutzer*innen gerecht werden. Der Fokus liegt hier auf Realismus, und weniger auf den Erwartungen der Nutzer*innen an die Nutzung der Objekte zur Durchführung von Interaktionen, wie es bei Heuristik 2: *Compatibility with the user's task and domain* der Fall ist. Die Anforderungen an realistisches Verhalten von Objekten sind in der vorliegenden Domäne eher nebensächlich. Beispielsweise lassen sich in keinem der ausgewählten Projekte Beispiele für Objekte finden, die ergriffen und anschließend fallen gelassen werden können.

Bestimmte Faktoren werden dennoch den Anforderungen an eine realistische Darstellung gerecht. Beispielsweise spiegelt das auditive Feedback in der *Zeitkapsel*-Anwendung, das unter Heuristik 3: *Natural expression of action* bereits beschrieben wurde, realistisches Verhalten von Objekten bzw. die Interaktion mit diesen wider. Weiterhin fällt bei der *Kunsthalle* auf, dass hier Schatteneffekte genutzt wurden, die die realistische Darstellung des Raumes und der Objekte unterstützen. Unterschiedliche Lichtquellen lassen die 3D-modellierten Objekte im virtuellen Museum unterschiedliche Schatten werfen. Die Lichtquellen selbst spiegeln sich teilweise im Boden. Diese Tiefenhinweise bieten den Nutzer*innen eine verbesserte Wahrnehmung des dreidimensionalen Raumes insgesamt und ergänzen die Vorstellung der Positionen der Objekte im Raum.⁵² In der gleichen Anwendung wird der Realismus dadurch verstärkt, dass Gegenstände, Wände und Möbel innerhalb der 3D-modellierten Museumsräume Hindernisse darstellen. Das bedeutet, dass Nutzer*innen nicht durch diese Objekte hindurchlaufen können. So wird die Nähe zur realen Welt gewährleistet. Nutzer*innen werden Raum für Raum durch die Ausstellung geleitet und können sich besser orientieren, da ihr Orientierungssinn nicht durch physisch unmögliche Handlungen wie das Durchfliegen von Wänden gestört wird. Darüber hinaus besteht die Möglichkeit, die Reihenfolge der Räume und Exponate so zu gestalten, dass diese zum Verständnis des Themenkomplexes beiträgt, ohne dass Nutzer*innen diese Reihenfolge missachten können.

Zusammenfassung: Objekte, die Interaktion durch Bewegung zulassen, sollten grundlegenden physikalischen Gesetzen folgen. Realismus kann in Desktop-VR-Anwendungen durch auditives Feedback sowie Schatteneffekte verbessert werden, wodurch auch die dreidimensionale Wahrnehmung der Nutzer*innen unterstützt wird. Wird die freie Bewegung dadurch limitiert, dass Objekte und Wände Hindernisse darstellen, so verbessert dies die Orientierung in der Anwendung: Indem die Wahlmöglichkeiten der Nutzer*innen eingeschränkt sind und sie weniger Entscheidungen treffen müssen, wird der Umgang mit dem System vereinfacht.⁵³

3.6 Zuverlässige Blickpunkte (Faithful viewpoints)

Diese Heuristik behandelt Perspektivwechsel innerhalb des VE. Diese sollen ohne Verzögerung entsprechend den Kopfbewegungen der Nutzer*innen wiedergegeben werden. Hier wird deutlich, dass diese Heuristik nur bedingt auf Desktop-VR-Anwendungen angewendet werden kann, da kein VR-Headset zur Verfügung steht, um den Blickpunkt mit Kopfbewegungen zu wechseln. Jedoch lassen sich auch in Desktop-VR-Anwendungen die Blickpunkte mit Maus und Tastatur als Eingabegeräte verändern.

Projekt *Google* ermöglicht es den Nutzer*innen, sich in alle Richtungen ohne Einschränkungen umzusehen. Dazu halten die Nutzer*innen die linke Maustaste gedrückt und ändern den Blickpunkt, indem sie das Bild mit der Maus umherschoben, wobei der Bildschirmrand eine Limitation darstellt: Ist dieser mit dem Cursor erreicht, muss der Cursor erst wieder in eine neue Position gebracht werden, damit der Blickwinkel weiter verändert werden kann. Durch die Betätigung des Mousrades oder das Zusammen- bzw. Auseinanderziehen der Finger auf dem Touchpad kann zusätzlich heran- und weggezoomt werden. Die Möglichkeit zur Änderung des Blickpunktes bietet sich vor allem in Anwendungen an, in denen historische Orte präsentiert werden. Ein Beispiel aus der *Google*-Anwendung ist ein virtueller Rundgang durch Schloss Versailles, bei dem es viele Kunstwerke zu entdecken gilt, die sich nicht auf Kopfhöhe der Nutzer*innen befinden, sondern beispielsweise die Decke zieren. Durch die

⁵¹ Vgl. Eriksson 2021.

⁵² Vgl. Preim / Dachsel 2015, S. 261–263.

⁵³ Vgl. LaViola et al. 2017, S. 431.

Zoom-Funktion lassen sich in den Gebäuden und Kunstwerken Details ausmachen, die sonst nicht zu erkennen wären. Auch die Kunsthalle bietet Nutzer*innen die Möglichkeit, sich frei im Raum umzuschauen. Die Steuerung unterscheidet sich minimal von Google: Die Veränderung des Blickwinkels wird durch Gedrückthalten einer beliebigen Maustaste gesteuert. Dabei wendet sich der Blick in die Richtung, in die der Cursor bewegt wird. Bei Google ist es umgekehrt. Hier bewegt sich der Blickwinkel immer in die entgegengesetzte Richtung des Cursors. Es ist unklar, welche der beiden Varianten der Blickpunktänderung hier stärker konventionalisiert ist, da beide Interaktionsvarianten jeweils auch in zahlreichen anderen Anwendungen vorzufinden sind.

Zusammenfassung: Grundsätzlich sollten bei der Gestaltung von virtuellen Museumsangeboten freie Wechsel des Blickwinkels ermöglicht werden. Das ist insbesondere dann wichtig, wenn die Ausstellung von Exponaten nicht auf eine durchschnittliche Kopfhöhe limitiert ist, sondern überall im Raum zu finden sind. Durch das Heranzoomen mit Mausrad oder Touchpad können zusätzlich Details exploriert werden.

3.7 Unterstützung bei Navigation und Orientierung (Navigation and orientation support)

Die Orientierung im VE sollte insofern durch eine intuitive Navigation unterstützt werden, dass Nutzer*innen ihre Position stets ausmachen und gegebenenfalls zu bekannten Positionen steuern können. Dabei sollten sich unnatürliche Aktionen wie das Durchfliegen von Räumen mit dem Anspruch des VE nach Natürlichkeit die Waage halten. In der digitalen Kunsthalle wird etwa die Orientierung in der Ausstellung durch eine Anzeige am rechten unteren Bildschirmrand unterstützt, welche die Nutzer*innen darüber informiert, in welchem Raum sie sich befinden. Die Räume sind mit einer Nummer und einem Titel versehen. Nutzer*innen werden in einen anderen Raum teleportiert, wenn sie einen der Pfeile links und rechts der Anzeige anklicken. So können sie, ohne selbst in die entsprechenden Räume zu navigieren, einfach in andere Räume innerhalb der Ausstellung springen. Eine ähnliche Umsetzung der Unterstützung wird den Nutzer*innen beim Projekt *Nach Berlin* geboten. Hier sind am unteren Bildschirmrand nebeneinander alle Orte, an die sich die Nutzer*innen begeben können, mit Namen und einem Bild gelistet. Damit deutlich ist, an welchem Ort sich die Nutzer*innen befinden, sind die inaktiven Orte grau unterlegt. So wird der aktive Ort hervorgehoben. Die gelisteten Orte fungieren einerseits als Information über den aktuellen Aufenthaltsort der Nutzer*innen im VE, andererseits kann nur durch Anklicken der Orte die Position gewechselt werden.

Beim Museum *Schnütgen* wurde bei der Navigationshilfe auf eine klassische Karte gesetzt (vgl. *Abbildung 2*). Durch das Klicken auf ein entsprechendes Symbol am rechten unteren Bildschirmrand öffnet sich eine Karte, welche die Ausstellungsräume zeigt. Zusätzlich werden auf der Karte alle Punkte angezeigt, die »angeflogen« werden können. Auch der aktuelle Standort der Nutzer*innen wird auf der Karte rot markiert und mit dem Winkel der Blickrichtung dargestellt. Standortänderungen können durch das Anklicken der verfügbaren Punkte erfolgen. Diese Hilfestellung bietet sich hier besonders an, da die Bewegung in der Anwendung auf das Anklicken einzelner Punkte, von denen aus sich die Nutzer*innen umsehen können, beschränkt ist. Auch auditive Stimuli können die Nutzer*innen bei der Orientierung im VE unterstützen, indem Geräusche und ihre Lautstärke einen Ort oder die Distanz zu einem Ort signalisieren.⁵⁴ Stehen den Nutzer*innen Kopfhörer zur Verfügung, profitieren sie in der *WDR-Zeitkapsel* von Raumklang. So können sie ausmachen, aus welcher Richtung Geräusche kommen, beispielsweise Gespräche von Personen oder der Klang eines Fernsehers oder Plattenspielers.

Zusammenfassung: Nutzer*innen sollten bei der Navigation und Orientierung im VE unterstützt werden. Je nach Anwendung bieten dafür unterschiedliche Möglichkeiten an, wie z. B. die Einbindung einer Karte oder die Auflistung der Räume der Ausstellung. Durch auditives Feedback kann die Orientierung zusätzlich verbessert werden, indem Orte oder Distanzen zu diesen so signalisiert werden.

⁵⁴ Vgl. LaViola et al. 2017, S. 363.

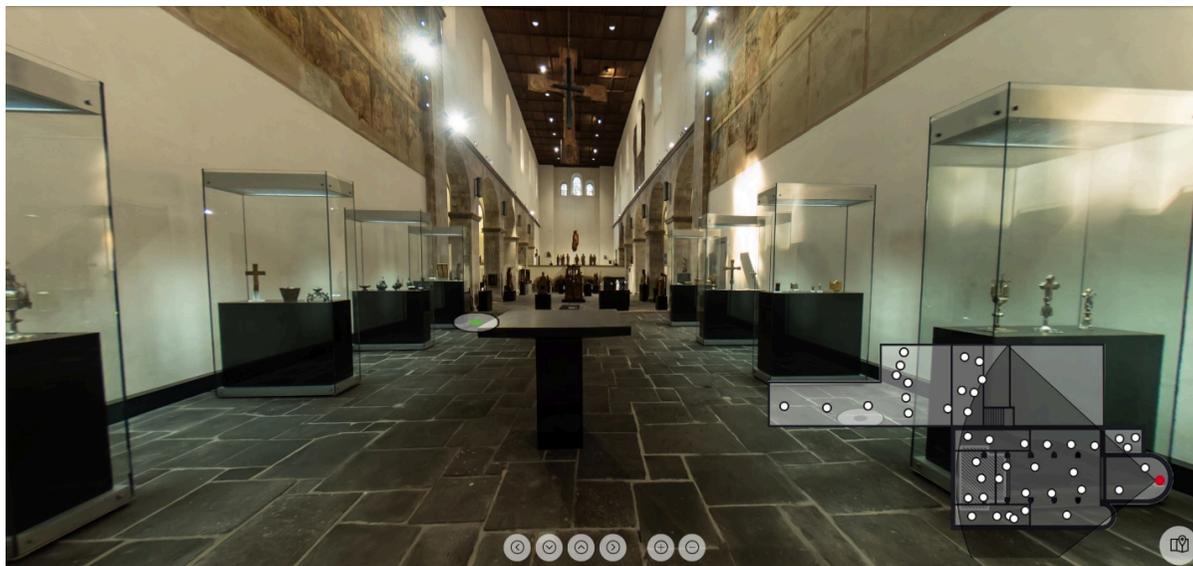


Abb. 2: Navigationshilfe im Museum Schnütgen. In: Museum Schnütgen. 360° Rundgang. [Stadt Köln (Hg.) 2020]

3.8 Eindeutige Ein- und Ausstiegspunkte (Clear entry and exit points)

Die Ein- und Ausstiegspunkte der VEs sollen eindeutig sein, sodass Nutzer*innen wissen, wie sie ein VE betreten und auch wieder verlassen können. Die Relevanz dieser Heuristik für Desktop-VR-Anwendungen wurde von den Autoren betont.⁵⁵ Die Einbindung eindeutiger Ein- und Ausstiegspunkte wurde in den Projekten unterschiedlich umgesetzt.

In der *Zeitkapsel* ist der Einstieg für Nutzer*innen in das VE in ein Storytelling eingebaut. Um in die Anwendung selbst zu gelangen, betätigen die Nutzer*innen einen ›Starten‹-Button. Daraufhin kann zwischen einer normalen oder einer hohen Auflösung gewählt werden. Nun befinden sich die Nutzer*innen bereits im VE, müssen sich aber noch entscheiden, in welche virtuelle Welt sie reisen möchten (*1960 – Willkommen an Bord* *1968 – Schöner Wohnen*). Dazu werden ihnen von einem ›Dealer‹ unterschiedliche Zeitreise-Kapseln angeboten, von denen sie eine einnehmen sollen. So ist der Einstieg in das VE interaktiv gestaltet und in das Storytelling der Anwendung eingebunden. Die virtuellen Welten können gewechselt werden, indem über das Menü, das am rechten oberen Bildschirmrand aufgerufen werden kann, eine andere Zeitkapsel geschluckt bzw. die Episode gewechselt wird. Das Aufrufen des Menüs ist jederzeit möglich, wodurch Nutzer*innen die Möglichkeit haben, das anfängliche Storytelling zu überspringen und sofort in eine andere virtuelle Welt einzutauchen.

Weniger interaktiv, dafür aber leichter verständlich, wird der Ein- und Ausstieg in und aus dem VE in der *Kunsthalle* gestaltet. Auf der Startseite der Digitalen Kunsthalle sind alle aktuell zugänglichen Ausstellungen gelistet. Mit einem Klick auf den Button ›Ausstellung besuchen‹ gelangen die Nutzer*innen, nachdem die virtuelle Ausstellung geladen wurde, in das gewünschte VE. Die Möglichkeiten zum Verlassen einer Ausstellung bzw. zum Wechseln zu einer anderen sind ebenfalls klar gekennzeichnet. Einerseits befindet sich im ersten und letzten Raum jeder Ausstellung an der 3D-modellierten Wand ein ›Alle Ausstellungen‹-Button, welcher die Nutzer*innen zurück auf die Startseite führt. Andererseits ist dieser Button auch über das Menü am oberen Bildschirmrand zugänglich.

Im Sinne des *Natural engagement* beginnt die virtuelle Museumstour von Schnütgen im Foyer des Museums. Mit dem Anklicken des sich am Boden des Museums befindlichen Buttons ›Zum 360°-Rundgang‹, der sich an der Stelle befindet, wo in der realen Welt auch der Eingang zu der Ausstellung wäre, gelangen die Nutzer*innen in das virtuelle Museum.

Zusammenfassung: Ein- und Ausstiegspunkte sollten klar gekennzeichnet sein und idealerweise an mehreren Stellen zur Verfügung stehen. Beispielsweise können im VE selbst Buttons zum Verlassen oder Ändern der virtuellen Welt eingebunden werden, die gleichzeitig auch über das Menü zugänglich sind. Abhängig von der Art und Zielgruppe der Anwendung bietet sich eine Einbindung der Ein- und Ausstiegspunkte ins Storytelling an.

⁵⁵ Vgl. Sutcliffe / Gault 2004, S. 845.

3.9 Konsistenz bei Designkompromissen (Consistent departures)

Immer dann, wenn im virtuellen Raum etwas anders umgesetzt wird als in der realen Welt, sollte dies eindeutig gekennzeichnet und vor allem konsistent umgesetzt werden. Typische Beispiele sind hier etwa die Substitution von Modalitäten, beispielsweise das Ersetzen von realem ›Tasten / Fühlen‹ durch ›Hören / Sehen‹ im virtuellen Raum, oder die Verwendung von effektiven Navigationsmechanismen, etwa das Fliegen durch den virtuellen Raum oder die Nutzung von Teleportation. In den evaluierten Projekten finden sich vor allem Beispiele aus dem letztgenannten Bereich der Navigation, allerdings nicht im Sinne erweiterter Navigationsmöglichkeiten, sondern eher im Sinne von Einschränkungen der Bewegung im virtuellen Raum.

Bei Google Arts and Culture etwa ist eine freie Bewegung im Raum nur bedingt möglich. Die Nutzer*innen können mit der Maus anzeigen, in welche Richtung sie sich bewegen möchten. Der Cursor stellt in der Anwendung einen Pfeil dar, der in die Richtung zeigt, in die sich die Nutzer*innen beim Mausklick bewegen werden. Bei der Bewegung zu einem ausgewählten Punkt zeigt eine kurzzeitige Unschärfe des Bildes den Nutzer*innen an, dass diese zu einem Punkt navigieren. So wird deutlich, dass es sich bei der Fortbewegung nicht um ein flüssiges Laufen im VE, sondern um ein schrittweises Weiterkommen zu einem ausgewählten Punkt handelt. Im Sinne dieser Heuristik wird die Anzeige dieser eingeschränkten Navigation allerdings in der gesamten Anwendung konsistent umgesetzt. Beim Museum *Schnütgen* ist die Fortbewegung vergleichbar eingeschränkt, da sich die Nutzer*innen hier nur zu bestimmten, eindeutig gekennzeichneten Punkten bewegen können. Hierbei handelt es sich um grüne Punkte, die mit einem weißen Kreis umrandet sind. Befindet sich der Cursor auf einem der Punkte, so wird den Nutzer*innen in einer Vorschau gezeigt, welchen Blickpunkt sie einnehmen, wenn sie den Punkt anklicken. Wurde ein Punkt bereits angeklickt, verändert sich die Farbe von grün zu grau. So wird den Nutzer*innen signalisiert, dass sie sich an diesem Punkt bereits umgesehen haben. Auch in diesem Beispiel wird die farbkodierte Information konsistent innerhalb der gesamten Anwendung umgesetzt.

Zusammenfassung: Besonders in Anwendungen, in denen die Bewegung im virtuellen Raum nicht so frei wie in der realen Welt gegeben ist, ist es wichtig, den Nutzer*innen eindeutig und konsistent zu signalisieren, an welche Punkte sie sich bewegen können. Dies kann etwa durch eine Darstellung des Cursors als Pfeil und durch die Unschärfe des Bildes, während die Nutzer*innen an einen bestimmten Punkt ›reisen‹, gewährleistet werden.

3.10 Unterstützung beim Lernen (Support for learning)

Diese Heuristik fordert, dass die Nutzer*innen beim Erlernen der Funktionen des VE durch entsprechende Erklärungen und Hilfestellungen unterstützt werden. In der *Zeitkapsel*-Anwendung wird dies umgesetzt, indem die Interaktionsmöglichkeiten innerhalb des VE unterschiedlich gekennzeichnet werden. Unterschieden wird dabei zwischen erklärenden Audiosequenzen, die in das Storytelling der Anwendung eingebunden sind, dem Heranzoomen an Objekte wie dem Fernseher, Bücherregal oder Radio und anderen Interaktionsmöglichkeiten wie beispielsweise der Entstehung eines Gesprächs zwischen zwei Personen oder dem Klingeln des Telefons. Für alle Elemente gibt es Symbole, welche die Möglichkeiten der Interaktion andeuten und gleichzeitig als Buttons fungieren. Die Buttons verändern beim Näherkommen des Cursors ihre Größe und werden so deutlicher erkennbar. Entsprechend werden die Nutzer*innen subtil auf die Interaktionsmöglichkeiten hingewiesen. Zusätzlich ist über das Menü, das über einen Button in der rechten oberen Ecke geöffnet wird, eine Hilfeseite zugänglich, welche die Funktionen des VE kurz erörtert.

Beim Starten der *Berlin*-Anwendung öffnet sich ein kurzes Tutorial, das die Hauptfunktionen der Anwendung erklärt. Erst, wenn die Nutzer*innen auf ›Weiter‹ klicken, gelangen sie in die eigentliche Anwendung (vgl. *Abbildung 3*). So werden die Nutzer*innen bereits vor der Nutzung mit der Anwendung vertraut gemacht. Auch in diesem Beispiel signalisieren unauffällige, aber pulsierende Punkte, dass weitere Informationen oder Interaktionsmöglichkeiten zur Verfügung stehen. Ein kurzes Tutorial zu den Bedienelementen der Anwendung beim Start bietet auch *Walter's Cube*. Dieses Tutorial kann mit einem Klick auf ein Fragezeichen-Symbol am rechten unteren Bildschirmrand stets wieder aufgerufen werden.

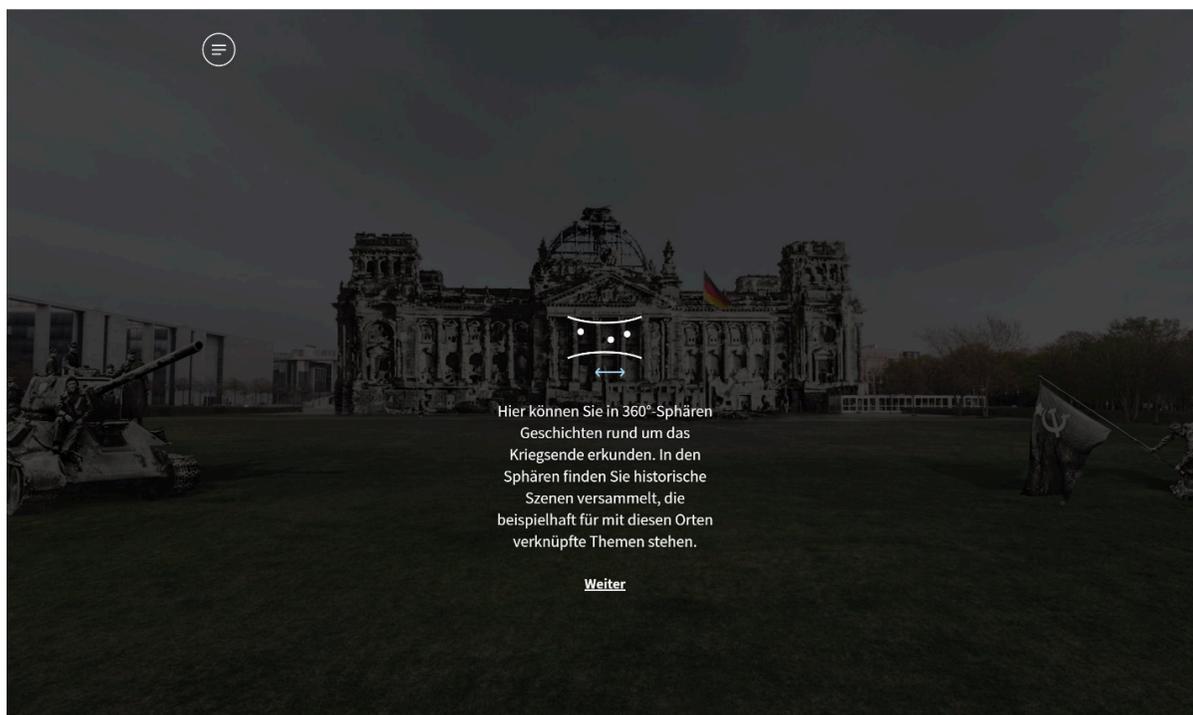


Abb. 3: Einstiegstutorial des Projekts Nach Berlin. In: Nach Berlin. 75 Jahre Kriegsende. [Kulturprojekte Berlin (Hg.) 2020]

Zusammenfassung: Eine Hilfeseite, die den Nutzer*innen die Bedienung und Hauptfunktionen der Anwendung verständlich und kurz erläutert, ist obligatorisch für VEs der Domäne Museum, da so auch die Zugänglichkeit erleichtert wird. Diese Hilfestellung sollte schnell zu finden sein, beispielsweise über das Menü oder über einen Button, der direkt zur Hilfeseite führt. Mithilfe unterschiedlich gestalteter Buttons kann signalisiert werden, dass unterschiedliche Interaktionsmöglichkeiten zur Verfügung stehen. Durch ein unauffälliges Design stören sie die Nutzer*innen nicht in der Exploration des VE, machen aber durch ihre Dynamik, beispielsweise der Veränderung ihrer Größe, auf sich aufmerksam. Weiterhin wird den Nutzer*innen auf diese Weise verdeutlicht, dass nur bestimmte Objekte die Möglichkeit einer Interaktion bieten.

3.11 Eindeutiges turn-taking (Clear turn-taking)

Laut Sutcliffe und Gault ist diese Heuristik nur für kollaborative VEs relevant,⁵⁶ also VEs, in denen sich mehrere Nutzer*innen gleichzeitig befinden und miteinander interagieren. Da eine solche Mehrbenutzer*innen-Funktion in keinem der ausgewählten Projekte implementiert ist, wird Heuristik 11 aus der vorliegenden Evaluation ausgeschlossen.

Zusammenfassung: Diese Heuristik betrifft ausschließlich kollaborative VEs und ist damit für diese Studie nicht relevant.

3.12 Gefühl von Präsenz (Sense of presence)

Diese Heuristik fordert, dass die Wahrnehmung des Aufenthalts im VE und das damit verbundene Gefühl von Präsenz so natürlich und nah wie möglich an der realen Welt sein sollte. Ob oder inwieweit dieser Anspruch auf die Domäne zutrifft, wurde bereits unter *Abschnitt 3.1 Natural engagement* diskutiert. Jedoch lässt sich im Projekt *Zeitkapsel* ein Element ausmachen, welches die Wahrnehmung der Natürlichkeit verstärkt: die Interaktion mit Personen. In die genannte Anwendung ist ein interaktives Storytelling eingebunden, wobei reale Personen mit den Nutzer*innen interagieren und auf deren Aktionen reagieren. Bereits beim Einstieg in das VE interagieren die Nutzer*innen mit einem ›Dealer‹, der Zeitkapseln verkauft, die virtuelle Zeitreisen ermöglichen sollen (vgl. *Abschnitt 3.8 Clear entry and exit points*). Der ›Dealer‹ bewegt sich zum Standpunkt der Nutzer*innen und spricht diese direkt an. Im Storytelling wird suggeriert, dass die Nutzer*innen innerhalb der Szenarien nicht von den sich darin befindenden Personen gesehen werden können. Trotzdem schaut beispielsweise beim Einstieg in das Szenario *1968 – Schöner Wohnen* das Ehepaar, das sich in ihrem Wohnzimmer im Jahr 1968 befindet, in Richtung der

⁵⁶ Vgl. Sutcliffe / Gault 2004, S. 846.

Nutzer*innen und scheint deren Präsenz zu bemerken. Spielerisch werden die Nutzer*innen dazu aufgefordert, Zeitfehler in den Szenarien zu entdecken. Zum Beispiel steht in eben diesem Szenario das Buch *Fifty Shades of Grey* aus dem Jahr 2011 im Bücherregal des Ehepaares. Entdecken die Nutzer*innen diesen Zeitfehler, erscheint der ›Dealer‹ als Hologramm, entschuldigt sich bei den Nutzer*innen für den Fehler und tauscht das Buch durch ein anderes aus.

Zusammenfassung: Vor allem interaktive Elemente steigern das Gefühl der Immersion bei den Nutzer*innen. Sie dienen weiterhin der Unterhaltung und sollen die Nutzer*innen aktivieren und dazu motivieren, das VE weiter zu explorieren.

4. Diskussion

Bei der Beschäftigung mit den Heuristiken stellte sich heraus, dass die Interpretation und damit zusammenhängend die Abgrenzung einiger Heuristiken Probleme für die Bearbeitung darstellen. Davon betroffen sind vor allem jene Heuristiken, die sich weitestgehend auf die Natürlichkeit der Anwendung und die Nähe zur realen Welt beziehen (Heuristiken 1–6 und 12). Bei diesen Heuristiken sind die Unterschiede graduell und beziehen sich auf das, was Gegenstand der Natürlichkeit ist, also beispielsweise das Verhalten von Objekten, die Wahrnehmung des Selbst oder die Durchführung von Aufgaben. Diese Schwierigkeit wurde von den Autoren in der Evaluation ihrer Heuristiken bereits erkannt und thematisiert: Hier gaben viele der Begutachter*innen an, Probleme bei der Interpretation der Heuristiken zu haben.⁵⁷ Die Problematik konnte jedoch als Interpretationsspielraum verstanden und genutzt werden, um die Heuristiken der Domäne entsprechend auszulegen. Dennoch stellt sich die Frage, welcher Anspruch an Natürlichkeit und Nähe zur realen Welt an Desktop-VR-Anwendungen im musealen Bereich gestellt werden sollte. Anstatt eine reale Museumserfahrung zu imitieren, können die Potenziale dieser Präsentationsform genutzt werden, um Exponate anders auszustellen, zu vermitteln und die Nutzer*innen über Themenkomplexe zu informieren. Dabei werden an vielen Stellen Kompromisse mit der Natürlichkeit der Anwendung eingegangen.

Nicht alle Heuristiken haben sich als geeignet für die Ausarbeitung von Best Practices erwiesen. Heuristik 4 (*Close coordination of action and representation*) wurde in der Bearbeitung nicht berücksichtigt, da sie weniger mit der Gestaltung als der Implementierung einer Desktop-VR-Anwendung in Zusammenhang steht. Auch Heuristik 11 (*Clear turn-taking*) wurde für die Generierung von Best Practices als nicht relevant eingestuft, da sie sich lediglich für kollaborative VEs eignet. Im Gegensatz dazu erwiesen sich die anderen Heuristiken als äußerst produktiv und gut anwendbar. Aus ihnen konnten Best Practices für die Gestaltung von virtuellen Museumsräumen abgeleitet werden, die nachfolgend nochmals zusammengefasst werden.

4.1 Zusammenfassung von Best Practices auf Basis der Heuristiken

Nachfolgend findet sich eine Zusammenfassung der wichtigsten Erkenntnisse aus der Evaluation ausgewählter VEs in Form von verallgemeinerbaren Best Practices:

- **Freie Bewegung unterstützen:** Durch eine freie Bewegung im Raum und handgesteuerte Navigation können Nutzer*innen das VE erforschen und sich Exponate aus unterschiedlichen Blickwinkeln anschauen, wobei die Kollision mit Objekten oder Wänden vermieden werden sollte. Kommen in VEs unrealistische Navigationselemente zum Einsatz – sei es durch Einschränkungen oder erweiterte Möglichkeiten – muss dies eindeutig gekennzeichnet und vor allem konsistent umgesetzt werden. Außerdem sollen den Nutzer*innen zum Betreten oder Verlassen des VE eindeutige Ein- und Ausstiegspunkte an mehreren Stellen zur Verfügung stehen.
- **Navigationshilfen anbieten:** Bei der Orientierung sollten grundsätzlich Navigationshilfen wie Karten die Nutzer*innen unterstützen. Auch auditives Feedback kann hier eine Hilfestellung sein, da Geräusche einen bestimmten Ort (beispielsweise den Standort eines Radios) bzw. die Distanz zu diesem signalisieren können.
- **Interaktion unterstützen und kenntlich machen:** Detailansichten mit hochauflösenden Fotografien der Exponate sowie die Möglichkeit des Heranzoomens an Objekte erleichtern deren genaue Betrachtung. Stehen den Nutzer*innen Interaktionsmöglichkeiten mit Objekten oder virtuellen Personen zur Verfügung, sollte dies stets deutlich gemacht werden. Dies kann beispielsweise durch Buttons angezeigt werden, wobei unterschiedliche Buttons auf verschiedene Arten der Interaktion hinweisen können. Eine Hilfeseite oder ein Tutorial sollten die wichtigsten Funktionen und Bedienelemente in verständlicher und leicht zugänglicher Form präsentieren.

⁵⁷ Vgl. Sutcliffe / Gault 2004, S. 846.

Zusammenfassend lässt sich sagen, dass aus vielen der generischen Heuristiken für VEs auch Best Practices auf den Museumsbereich übertragen werden konnten. Gleichzeitig fiel bei der heuristischen Evaluation der sechs ausgewählten Projekte auf, dass die bestehenden Heuristiken nicht alle Aspekte, die für den musealen Bereich relevant sind, vollständig abdecken. Nachfolgend ergänzen wir deshalb zwei weitere Heuristiken, die induktiv aus der Evaluation der Projektbeispiele abgeleitet wurden. Diese ergänzenden Heuristiken decken im Wesentlichen die Aspekte *Multimedialität* und *Interaktivität* ab und sollen nachfolgend näher erläutert werden.

4.2 Zusätzliche Heuristik: Multimediale Vermittlungselemente

Im Rahmen der Betrachtung von Heuristik 2 (*Compatibility with the user's task and domain*) wurden einige multimediale Elemente innerhalb der Projekte bereits beschrieben. Allerdings beschränkten sich die Beispiele, die als Best Practices angeführt wurden, auf multimediale Bestandteile der Projekte, welche Elemente einer realen Ausstellung imitieren. In den Projekten finden sich jedoch beispielsweise Audiosequenzen oder Videos, die in anderer Weise, wie es im realen Museum so nicht möglich wäre, präsentiert werden. Deggim et al. sehen die »große Stärke [von virtuellen Museen] in der ergänzenden Präsentation [...], die mit klassischen Ausstellungsmethoden nicht möglich [ist]«. ⁵⁸ Ansätze dazu finden sich auch in den betrachteten Projekten. In der WDR-Zeitkapsel haben die Nutzer*innen im Szenario *1968 – Schöner Wohnen* die Möglichkeit, ein Bücherregal im Wohnzimmer, in dem sie sich befinden, genauer zu untersuchen. Fahren sie mit der Maus über das Regal, werden ausgewählte Bücher ein Stück weit aus dem Regal hinausgezogen und ein kurzer Auszug aus dem Buch vorgelesen. Bei den Büchern handelt es sich um Werke, die um die entsprechende Zeit veröffentlicht wurden, beispielsweise das Buch *Deutschstunde* von Siegfried Lenz aus dem Jahr 1968 (vgl. Abbildung 4).



Abb. 4: Beispiel für Multimedialität in der WDR-Zeitkapsel. In: Zeitkapsel. [WDR (Hg.) 2018]

In der Kunsthalle in der Ausstellung *Felix Nussbaum – Leben und Werke* entpuppt sich ein Bild, das sich im ersten Raum an der Wand befindet, als kurzer Dokumentarfilm. Mit einem Klick auf das Bild wird das Video automatisch an der virtuellen Ausstellungswand abgespielt. Zwar steht dies im Kontrast zur oben genannten Heuristik, da laut dieser das Verhalten von virtuellen Objekten so nah wie möglich an Objekte der realen Welt angelehnt sein soll, jedoch werden so die Vorteile von virtuellen Museen gegenüber realen Ausstellungen ausgenutzt. Im gleichen Projekt in der Ausstellung *BEETHOVEN – Welt.Bürger.Musik* können Nutzer*innen an Hörstationen auditive Eindrücke von Beethovens Kompositionen erhalten. Dies unterscheidet sich auf den ersten Blick nicht von Hörstationen im realen Museum. Im VE können die Nutzer*innen die Musik aber weiter hören, während sie sich in der kompletten Ausstellung umsehen, die Hörerfahrung ist also nicht an einen bestimmten Punkt in der Ausstellung gekoppelt bzw. darauf beschränkt. Durch ein Fenster, welches sich mit dem Anklicken einer Hörstation öffnet, können die Nutzer*innen die Musik jederzeit pausieren, vor- oder zurückspulen und beenden.

⁵⁸ Deggim et al. 2017, S. 455.

Die Beispiele veranschaulichen unterschiedliche multimediale Präsentationsformen, die die Potenziale des Online-Mediums nutzen und sich so von Ausstellungsformen realer Museen absetzen. Daraus ergibt sich folgende Beschreibung für die neu entwickelte Heuristik *Einbindung multimedialer Vermittlungselemente* :

- Im Virtual Environment sollen multimediale Elemente so eingebunden werden, dass sie helfen, den Nutzer*innen Inhalte zu vermitteln. Dabei kann die Nähe des Verhaltens von Objekten zur Realität dann umgangen werden, wenn das Vermittlungsangebot und der Zugang zu diesem profitieren.

Diese Heuristik ist nicht auf die genannten Beispiele beschränkt; vielmehr sind die Gestalter*innen virtueller Museumsräume und vergleichbarer VEs dazu aufgerufen, die Potenziale des Mediums auszuschöpfen und neue Formen der Präsentation in Desktop-VR-Anwendungen zu entwickeln.

4.3 Zusätzliche Heuristik: Interaktive Wissensvermittlung

Interaktive Medienstationen sind längst Teil moderner Ausstellungskonzeption. Durch Spiele oder digitales Storytelling unterstützen sie bei der Vermittlung von Ausstellungsinhalten und aktivieren die Besucher*innen. Durch ihre Attraktivität müssen Besucher*innen allerdings häufig warten, bis sie eine Medienstation nutzen können.⁵⁹ Mit diesem Problem sind Nutzer*innen virtueller Museen nicht konfrontiert. Auch erfordert die Einbindung interaktiver Medien keine Station, stattdessen können sie direkt mit Exponaten gekoppelt und beispielsweise beim Klick auf ein Exponat geöffnet werden.

Die Ausstellung *Berechenbar – Unberechenbar* der *Kunsthalle*-Anwendung zeigt eine künstlerische und wissenschaftliche Auseinandersetzung mit Berechenbarkeit und Computern. Bei einem der Kunstwerke handelt es sich um den ›CellularAutomataExplorer‹ – eine Simulation, die nach einfachen Regeln die Populationsentwicklung von Lebewesen nachahmt. Nutzer*innen haben hier die Möglichkeit, in dem Raster des Automaten festzulegen, welche Zellen lebendig sind, woraufhin der Automat generationsweise nach den festgelegten Regeln ermittelt, welche Zellen überleben, neu geboren werden oder sterben. Die Nutzer*innen verfolgen den Algorithmus dann sowohl visuell als auch auditiv. Ein weiteres Kunstwerk, *Hypercam*, setzt sich künstlerisch mit der Praxis der Videokonferenzen auseinander und will auf die Künstlichkeit von derartigen digitalen Treffen aufmerksam machen. Mithilfe von *Hypercam* können Nutzer*innen ihre Videobilder verzerren und mit Schiebereglern in verschiedenen Dimensionen manipulieren. So werden sie interaktiv direkt in das Kunstwerk eingebunden und können es selbst explorieren.

Bei der Entwicklung eines virtuellen Museums für ein historisches Gebäude legten Deggim et al. den Fokus auf eine interaktive Visualisierung der Baugeschichte des Alt-Segeberger Bürgerhauses.⁶⁰ Diese Visualisierung ist für die Nutzer*innen des virtuellen Museums als ein Modell im 3D-modellierten Modell des Bürgerhauses zugänglich. Für die Betrachtung stehen sieben unterschiedliche Bauphasen zur Verfügung, wobei die Nutzer*innen im Menü auswählen, welche Bauphase sie betrachten möchten. Um die Veränderungen zwischen Bauphasen besser nachzuvollziehen, wird der Übergang von einer Bauphase in eine andere durch eine Animation dargestellt. Zusätzlich können das Dach des Modells entfernt und die einzelnen Räume angeklickt werden, um mehr über die Nutzung der Räume zu erfahren. Die Autoren betonen, dass die Einbindung des Bürgerhaus-Modells in den virtuellen Rundgang ein Beispiel für ein Vermittlungsmedium ist, das in dieser Form nur in einem virtuellen Museum eingebunden werden kann und so die Vorteile des Mediums illustriert.

Auch die Interaktion mit Personen, wie sie bereits unter der Heuristik 12 *Sense of presence* dargestellt wurde, ist für diese Heuristik relevant. Ein wie im Projekt *Kunsthalle* umgesetztes digitales Storytelling kann unterhaltende und den Lernprozess fördernde Elemente kombinieren und greift so das Konzept des »Edutainment« auf.⁶¹ Das Konzept bezeichnet die Verschmelzung von lehrenden und unterhaltenden Aspekten und ist besonders im musealen Kontext relevant, da ein Museum einerseits eine Lernumgebung ist, die andererseits (normalerweise) freiwillig besucht wird und entsprechend auch unterhaltend sein soll. Eine Anwendung dieses Konzeptes, beispielsweise in Form von digitalem Storytelling oder der Einbindung von (Lern-)Spielen und Simulatoren, kann helfen, dass Nutzer*innen aktiviert werden, dass sie sich besser an Inhalte erinnern und sich diese stärker einprägen und dass sie überhaupt ein ausgeprägteres Interesse für Exponate oder Themenkomplexe aufbauen.

⁵⁹ Vgl. Landesstelle für die nichtstaatlichen Museen in Bayern et al. (Hg.) 2019, S. 99.

⁶⁰ Vgl. Deggim et al. 2017, S. 459. Bei dem Projekt handelt es sich um eine nicht frei zugängliche Desktop-VR-Anwendung, weshalb bewusst eine Entscheidung gegen die Aufnahme in die Projektauswahl getroffen wurde. Aufgrund der gelungenen Einbindung eines interaktiven 3D-Modells wird es an der Stelle dennoch als Beispiel für Interaktivität in virtuellen Museen aufgeführt.

⁶¹ Vgl. Ioannidis et al. 2013, S. 421-422.

Virtuelle Museen haben das Potenzial, solche interaktiven Medien direkt an Exponate und Kunstwerke zu koppeln oder Nutzer*innen sogar in diese einzubinden. Die Beschreibung für die Heuristik *Interaktive Wissensvermittlung* lautet zusammengefasst wie folgt:

- Die Einbindung interaktiver Wissensvermittlung in das Virtual Environment hilft, um Nutzer*innen zu aktivieren und bei der Erfassung von Inhalten zu unterstützen.

5. Fazit

Die vorliegende Studie untersucht die Anwendbarkeit von bestehenden Heuristiken zur Gestaltung von VEs auf die Domäne virtueller Museumsräume. Dabei wurden die Heuristiken im Rahmen einer heuristischen Evaluation auf sechs beispielhafte Projekte angewandt. So sollten einerseits die Eignung der Heuristiken für den Bereich virtueller Museen überprüft und andererseits – die Heuristiken als Schablone nutzend – konkrete Umsetzungen innerhalb der Projektbeispiele als Best Practices extrahiert werden. Von den ursprünglichen zwölf Heuristiken von Sutcliffe und Gault sind letztlich nur zehn für den Bereich virtueller Museumsräume geeignet. Zusätzlich wurden zwei weitere, domänen-spezifische Heuristiken aus den Projektbeispielen abgeleitet. Im Ergebnis entstand so eine Sammlung von Heuristiken, die es künftigen Gestalter*innen im GLAM-Bereich ermöglichen sollen, virtuelle Museumsanwendungen mit einem hohen Maß an Usability und einer guten User Experience zu konzipieren. Die aus den Projektbeispielen abgeleiteten Best Practices geben dabei weitere Hinweise zur konkreten Umsetzung der teilweise abstrakten Heuristiken.

Wenngleich wir davon überzeugt sind, dass die vorliegende Diskussion und Erweiterung von Heuristiken sowie die abgeleiteten Best Practices hilfreich für das Design künftiger Museumsanwendungen sein werden, so gibt es gleichzeitig einige Limitationen dieser Studie, die nachfolgend diskutiert werden. Eine wesentliche Einschränkung stellt die Auswahl von insgesamt sechs Projekten, stellvertretend für den gesamten Bereich musealer VR-Anwendungen, dar. Zwar wurden die Projekte nach unterschiedlichen Diversifikationskriterien so ausgewählt, dass eine möglichst repräsentative Auswahl entstanden ist, ein Anspruch auf vollständige Abbildung aller Gestaltungsphänomene in diesem Bereich kann auf dieser Basis allerdings nicht erhoben werden. Eine Auswahl nach beispielsweise inhaltlichen Kategorien oder den Zielgruppen der Anwendungen könnte die Vergleichbarkeit der Projekte bei zukünftigen Untersuchungen erhöhen. Da vier der sechs Projekte aus dem deutschsprachigen Raum kommen, und mit *Walter's Cube* und *Google Arts and Culture* nur zwei internationale Projekte untersucht wurden, ist der Aspekt interkultureller Designimplikationen in dieser Studie nicht weiter berücksichtigt.⁶² Eine weitere Einschränkung der Studie liegt zudem sicherlich in der Fokussierung auf frei zugängliche Desktop-VR-Anwendungen, die – trotz zunehmender Verbreitung von VR-Technologie im Consumer-Bereich – in den nächsten Jahren trotz alledem weiterhin der Standard-Anwendungsfall bleiben dürften.

Die Relevanz von barrierefreien Zugängen virtueller Ausstellungen wurde im Rahmen dieser Studie zwar erwähnt, in der heuristischen Evaluation fand dieser Aspekt aber wenig Beachtung. Eine Evaluation und Überarbeitung der Heuristiken bezüglich des Einbezugs von Barrierefreiheit und eine sich daran anschließende heuristische Evaluation der Projekte nach den neu gewonnenen Erkenntnissen sind unabdingbar, um die Liste an Best Practices zu vervollständigen.

Insgesamt wirft die Beschäftigung mit virtuellen Museumsräumen auch die Frage auf, welche alternativen Präsentationsformen von Kunst- und Kulturobjekten das Medium Internet bietet. Viele Gedächtnisinstitutionen in Deutschland und auf der ganzen Welt befinden sich längst im Prozess der Digitalisierung ihrer Sammlungen und Exponate und vereinigen sich in Projekten wie der *Deutschen Digitalen Bibliothek* und *Europeana*,⁶³ um ihre Digitalisate online frei zur Verfügung zu stellen. Neben der Möglichkeit, die Digitalisate und Informationen für Forschungsvorhaben zu nutzen, können so auch institutionsübergreifende Online-Ausstellungen kuratiert werden, die Sammlungen und Objekte unterschiedlichster Museen und anderer Gedächtnisinstitutionen aus ganz Europa zeigen. Die in diesem Beitrag diskutierten Heuristiken und die daraus abgeleiteten Best Practices für die Gestaltung virtueller Ausstellungen wurden eingangs in Hinblick auf ein möglichst hohes Maß an Akzeptanz bei den Besucher*innen motiviert. Jenseits bloßer Akzeptanz im Sinne hoher Besuchszahlen haben Design-Guidelines für virtuelle Museumsräume darüber hinaus die wichtige Funktion der Unterstützung der Wissensvermittlung für breite Teile der nicht-akademischen Welt, wie sie etwa im Bereich der *Public History* gefordert wird.⁶⁴ Sieht man sich die Ursprünge der Public History-Bewegung näher an, wird ein weiterer Aspekt deutlich, der das große Potenzial benutzerfreundlicher, virtueller Ausstellungen verdeutlicht. So beschreibt Conard in ihrer Herleitung der intellektuellen Wurzeln der Public History etwa Benjamin Shambaugh

⁶² Vgl. Heimgärtner 2019.

⁶³ Bei der Deutschen Digitalen Bibliothek und Europeana handelt es sich um virtuelle Metabibliotheken, deren Ziel ist, das kulturelle und wissenschaftliche Erbe Deutschlands bzw. Europas in Form von Bild-, Text-, Audio- und Videodateien frei im Internet verfügbar zu machen.

⁶⁴ Vgl. Dean / Wojdon 2017.

Ansatz der *Applied History*, also eine konkrete Anwendung von Geschichte.⁶⁵ Dabei ging es Shambaugh vordergründig um die praktische Anwendung von Geschichtswissen, um aktuelle politische, soziale und wirtschaftliche Probleme zu lösen. Virtuelle Museumsräume bieten hier offenkundig vielfältige Möglichkeiten, um Kultur und Geschichte in simulierten Umgebungen direkt erfahr- und anwendbar zu machen und liefern damit einen wichtigen Grundstein für ein praktisches Geschichtsverständnis, dass sich hoffentlich auch auf aktuelle politische und soziokulturelle Herausforderungen anwenden lässt.

⁶⁵ Vgl. Conard 2001.

Bibliographische Angaben

- Beatriz de Almeida Pacheco / Marcelo Guimarães / Ana Grasielle Correa / Valeria Farinazzo Martins: Usability Evaluation of Learning Objects with Augmented Reality for Smartphones. A Reinterpretation of Nielsen Heuristics. In: Human-Computer Interaction. Hg. von Vanessa Agredo-Delgado / Pablo H. Ruiz. (HCI-Collab: 4, Popayán, 23.-27.04.2018) Cham 2019, S. 214–228. (= Communications in Computer and Information Science, 847) [\[Nachweis im GVK\]](#)
- Carol M. Barnum: Usability testing essentials: Ready, set-- test. Burlington, MA 2011. [\[Nachweis im GVK\]](#)
- Bernadette Biedermann: Virtual museums as an extended museum experience: Challenges and impacts for museology, digital humanities, museums and visitors - in times of (Coronavirus) crisis. In: Digital Humanities Quarterly 15 (2021), H. 3. [\[online\]](#)
- Freyja Catton / Laura Smith: Museums Without Walls. In: Pathfinder: A Canadian Journal for Information Science Students and Early Career Professionals 2 (2021), H. 2, S. 72–85. DOI: [10.29173/pathfinder31](#)
- Rebecca Conard: Benjamin Shambaugh and the intellectual foundations of public history. Iowa City, IA 2001. [\[Nachweis im GVK\]](#)
- David Dean / Joanna Wojdon: Public History and History Didactics – A Conversation. In: Public History Weekly 5 (2017), 9. Artikel vom 09.03.2017. [\[online\]](#)
- Simon Deggim / Felix Tschirschwitz / Thomas P. Kersten: Entwicklung eines virtuellen Museums für ein historisches Gebäude am Beispiel des Alt-Segeberger Bürgerhauses. (DGPF). In: Kulturelles Erbe erfassen und bewahren - von der Dokumentation zum virtuellen Rundgang. Hg. von Thomas P. Kersten. (Wissenschaftlich-Technische Jahrestagung der DGPF: 37, Würzburg, 08.–10.03.2017). Würzburg 2017, S. 454–463. (= Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation, 26) DOI: [10.24407/GBV:1015708250](#)
- Tommy Eriksson: Failure and Success in Using Mozilla Hubs for Online Teaching in a Movie Production Course. In: 2021 7th International Conference of the Immersive Learning Research Network (ILRN: 7, online, 17.05.–10.06.2021). Eureka, CA 2021. DOI: [10.23919/iLRN52045.2021.9459321](#)
- Das erweiterte Museum. Medien, Technologien und Internet. Hg. von Landesstelle für die nichtstaatlichen Museen in Bayern / Regina Franken-Wendelstorf / Sybille Greisinger / Christian Gries / Astrid Pellengahr. Berlin u. a. 2019. (= MuseumsBausteine, 19). DOI: [10.1515/9783422981010](#) [\[Nachweis im GVK\]](#)
- Google Arts and Culture. Startseite. Hg. von Google Inc. 2021. [\[online\]](#)
- Rüdiger Heimgärtner: Intercultural User Interface Design. Cham 2019. [\[Nachweis im GVK\]](#)
- Yannis Ioannidis / Katerina El Raheb / Eleni Toli / Akrivi Katifori / Maria Boile / Margaretha Mazura: One object many stories. Introducing ICT in museums and collections through digital storytelling. In: 2013 Digital Heritage International Congress. Hg. von Alonzo C. Addison. (DigitalHeritage, Marseille, 28.10.–01.11.2013) Piscataway, NJ 2013, S. 421–424. [\[Nachweis im GVK\]](#)
- Tanya M. Johnson: Let's Get Virtual. An Examination of Best Practices to Provide Public Access to Digital Versions of Three-Dimensional Objects. In: Information Technology and Libraries 35 (2016), No. 2, S. 39–55. DOI: [10.6017/ital.v35i2.9343](#) [\[Nachweis im GVK\]](#)
- Ger Joyce / Mariana Lilley: Towards the Development of Usability Heuristics for Native Smartphone Mobile Applications. In: Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Hg. von Aaron Marcus. (DUXU: 3, Heraklion, 22.–27.06.2014) Cham 2014. DOI: [10.1007/978-3-319-07668-3_45](#) [\[Nachweis im GVK\]](#)
- Katerina Kabassi / Alessia Amelio / Vasileios Komianos / Konstantinos Oikonomou: Evaluating Museum Virtual Tours. The Case Study of Italy. In: Information 10 (2019), No. 11, S. 351. DOI: [10.3390/info10110351](#)
- Robert Kelley: Public History: Its Origins, Nature, and Prospects. In: The Public Historian 1 (1978), H. 1, S. 16–28. Artikel vom 01.10.1978. [\[Nachweis im GVK\]](#)
- Nach Berlin. 75 Jahre Kriegsende. Hg. von Kulturprojekte Berlin. Berlin 2020. [\[online\]](#)
- Joseph J. LaViola Jr. / Ernst Kruijff / Ryan P. MacMahan / Doug A. Bowman / Ivan Poupyrev: 3D user interfaces. Theory and practice. 2. Auflage. Boston 2017. [\[Nachweis im GVK\]](#)
- Jean-Luc Lugrin / Florian Kern / Ruben Schmidt / Constantin Kleinbeck / Daniel Roth / Christian Daxer / Tobias Feigl / Christopher Mutschler / Marc Erich Latoschik: A Location-Based VR Museum. In: 10th International Conference on Virtual Worlds and Games for Serious Applications. Hg. von IEEE. (VS-Games: 10, Würzburg, 05.–07.09.2018) Piscataway, NJ 2018. [\[Nachweis im GVK\]](#)
- Anika Meier: Digitale Ausstellungen. Virtuelle Museumsräume killen nicht die Museen. In: Monopol (2020). Kolumne vom 05.05.2020. [\[online\]](#)
- Christine Murad / Cosmin Munteanu / Benjamin R. Cowan / Leigh Clark: Revolution or Evolution? Speech Interaction and HCI Design Guidelines. In: IEEE Pervasive Computing 18 (2019), H. 2, S. 33–45. [\[Nachweis im GVK\]](#)
- Museumsaufgaben. In: museumsbund.de. Hg. von Deutscher Museumsbund e.V. März 2017. [\[online\]](#)
- Jakob Nielsen: Heuristic Evaluation. In: Usability inspection methods. Hg. von Jakob Nielsen / Robert L. Mack. New York, NY 1994, S. 25–62. [\[Nachweis im GVK\]](#)
- Usability inspection methods. Hg. von Jakob Nielsen / Robert L. Mack. New York 1994. [\[Nachweis im GVK\]](#)
- Jakob Nielsen: Usability engineering. Boston, MA 1993. [\[Nachweis im GVK\]](#)
- Jakob Nielsen: Finding usability problems through heuristic evaluation. In: Striking a balance. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Hg. von Penny Bauersfeld / John Bennett / Gene Lynch. (CHI'92, Monterey, CA, 03.–07.05.1992) New York, NY 1992, S. 373–380. [\[Nachweis im GVK\]](#)
- Jakob Nielsen / Rolf Molich: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Hg. von John Whiteside / Jane Carrasco Chew. (CHI'90, Seattle, WA, 01.–05.04.1990). Reading, MA 1990, S. 249–256. [\[Nachweis im GVK\]](#)
- Dennis Niewerth: Virtuelle Museen. In: Handbuch Virtualität. Hg. von Dawid Kasprovicz / Stefan Rieger. Wiesbaden 2020, S. 521–532. [\[online\]](#) [\[Nachweis im GVK\]](#)
- Sandro Parrinello / Anna Dell'Amico: Experience of Documentation for the Accessibility of Widespread Cultural Heritage. In: Heritage 2 (2019), H. 1, S. 1032–1044. DOI: [10.3390/heritage2010067](#)
- Erick Paulus / Yusuf Firdaus / Miria Suryani / Ino Suryana: Development and Evaluation on Night Forest Virtual Reality as Innovative Nyctophobia Treatment. In: Journal of Physics: Conference Series 1235 (2019). DOI: [10.1088/1742-6596/1235/1/012003](#)
- Christiana Polycarpou: The ViMM Definition of a Virtual Museum. In: ViMM. Blogbeitrag vom 10.01.2018. [\[online\]](#)
- Bernhard Preim / Raimund Dachselt: Interaktive Systeme. 2 Bände. 2. Auflage. Berlin 2010. Bd. 2 (2015): User Interface Engineering, 3D-Interaktion, Natural User Interfaces. [\[Nachweis im GVK\]](#)
- Werner Schweibenz: The virtual museum: An overview of its origins, concepts, and terminology. In: The Museum Review 4 (2019), H. 1. [\[online\]](#)
- Sarah Sharples / Sue Cobb / Amanda Moody / John R. Wilson: Virtual reality induced symptoms and effects (VRISE). Comparison of head mounted display (HMD), desktop and projection display systems. In: Displays 29 (2008), H. 2, S. 58–69. DOI: [10.1016/j.displa.2007.09.005](#) [\[Nachweis im GVK\]](#)
- Ben Shneiderman / Catherine Plaisant / Maxine Cohen / Steven Jacobs / Niklas Elmqvist: Designing the User Interface. Strategies for Effective Human-Computer Interaction. 6. Auflage. Boston u. a. 2018. [\[Nachweis im GVK\]](#)
- Sebastian Späth: Museen in Deutschland. Hinter verschlossenen Türen. In: Die Zeit. Artikel vom 02.04.2020. [\[online\]](#)

Priyanka Srivastava / Anurag Rimzhim / Palash Vijay / Shruti Singh / Sushil Chandra: Desktop VR Is Better Than Non-ambulatory HMD VR for Spatial Learning. In: *Frontiers in Robotics and AI* 6 (2019). DOI: [10.3389/frobt.2019.00050](https://doi.org/10.3389/frobt.2019.00050)

Museum Schnütgen. 360° Rundgang. Hg. von Stadt Köln. Köln 2020. [[online](#)]

Sylaiou Styliani / Fotis Liarokapis / Kostas Kotsakis / Patias Petros: Virtual museums, a survey and some issues for consideration. In: *Journal of Cultural Heritage* 10 (2009), H. 4, S. 520–528. [[Nachweis im GVK](#)]

Alistair Sutcliffe / Brian Gault: Heuristic evaluation of virtual reality applications. In: *Interacting with Computers* 16 (2004), H. 4, S. 831–849. [[Nachweis im GVK](#)]

Walter's Cube: About Us. Hg. von Walter's Cube. In: walterscube.com. New York, NY 2020. [[online](#)]

Jugendliche reisen virtuell in die Vergangenheit. Unterrichtsvorschläge. Hg. von WDR. In: WDR. Schule digital. Unterrichtsvorschläge. Köln 2018. [[online](#)]

Changqing Yin / Zhaohui Chen / Yonghao Hu / Kexin Yu: Fine-grained Transmission Optimization of Large-scale Web VR Scenes. In: 2018 IEEE International Conference on Progress in Informatics and Computing. Hg. von Yinglin Wang / Yaoru Sun / Xing Wu. (PIC: 6, Suzhou, 14.–16.12.2018) Piscataway, NJ 2018, S. 209–214. [[Nachweis im GVK](#)]

Digitale Kunsthalle. Willkommen in der Digitalen Kunsthalle. Hg. von ZDF. Mainz 2021. [[online](#)]

Abbildungs- und Tabellenverzeichnis

Tab. 1: Heuristiken nach Sutcliffe und Gault. [Sutcliffe / Gault 2004]

Tab. 2: Übersicht über die zu untersuchenden Projekte. [Piontkowitz / Burghardt 2021]

Abb. 1: Detailansicht eines Kunstwerks mit zugehörigem Informationstext im Projekt Kunsthalle. In: Digitale Kunsthalle. [ZDF (Hg.) 2016]

Abb. 2: Navigationshilfe im Museum Schnütgen. In: Museum Schnütgen. 360° Rundgang. [Stadt Köln (Hg.) 2020]

Abb. 3: Einstiegstutorial des Projekts Nach Berlin. In: Nach Berlin. 75 Jahre Kriegsende. [Kulturprojekte Berlin (Hg.) 2020]

Abb. 4: Beispiel für Multimedialität in Projekt 1. In: Zeitkapsel. [WDR (Hg.) 2018]

Beitrag aus:

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. (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5) text/html Format. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Publishing an OCR ground truth data set for reuse in an unclear copyright setting. Two case studies with legal and technical solutions to enable a collective OCR ground truth data set effort

Autor*in:

David Lassner

Kontakt: lassner@tu-berlin.de

Institution: Technische Universität Berlin, Machine Learning Group | The Berlin Institute for the Foundations of Learning and Data (BIFOLD)

GND: [1246941414](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0001-9013-0834](https://orcid.org/0000-0001-9013-0834)

Autor*in:

Julius Coburger

Kontakt: julius.coburger@gmx.de

Institution: Technische Universität Berlin, Machine Learning Group

GND: [124694197X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0003-4502-7955](https://orcid.org/0000-0003-4502-7955)

Autor*in:

Clemens Neudecker

Kontakt: clemens.neudecker@sbb.spk-berlin.de

Institution: Staatsbibliothek zu Berlin – Preußischer Kulturbesitz

GND: [1246943069](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0001-5293-8322](https://orcid.org/0000-0001-5293-8322)

Autor*in:

Anne Baillot

Kontakt: anne.baillot@univ-lemans.fr

Institution: Le Mans Université | École normale supérieure de Lyon, Interactions, Corpus, Apprentissages, Représentations - ICAR

GND: [1065904681](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0002-4593-059X](https://orcid.org/0000-0002-4593-059X)

DOI des Artikels:

[10.17175/sb005_006](https://doi.org/10.17175/sb005_006)

Nachweis im OPAC der Herzog August Bibliothek:

[1780168195](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9)

Erstveröffentlichung:

10.12.2021

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise: 02.12.2021

GND-Verschlagwortung:

Zitierweise:

David Lassner, Julius Coburger, Clemens Neudecker, Anne Baillot: Publishing an OCR ground truth data set for reuse in an unclear copyright setting. Two case studies with legal and technical solutions to enable a collective OCR ground truth data set effort. 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. (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5) text/html Format. DOI: [10.17175/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_006](https://doi.org/10.17175/sb005_006).

David Lassner, Julius Coburger, Clemens Neudecker, Anne Baillot

Publishing an OCR ground truth data set for reuse in an unclear copyright setting. Two case studies with legal and technical solutions to enable a collective OCR ground truth data set effort

Abstracts

In dieser Arbeit stellen wir einen OCR-Trainingsdatensatz für historische Drucke vor und zeigen, wie sich im Vergleich zu unspezifischen Modellen die Erkennungsgenauigkeit verbessert, wenn sie mithilfe dieser Daten weitertrainiert werden. Wir erörtern die Nutzbarkeit dieses Datensatzes anhand von zwei Experimenten, die die rechtliche Grundlage zur Veröffentlichung digitalisierter Bilddateien am Beispiel von deutschen und englischen Büchern des 19. Jahrhunderts betrachten. Wir präsentieren ein Framework, mit dem OCR-Trainingsdatensätze veröffentlicht werden können, auch wenn die Bilddateien nicht zur Wiederveröffentlichung freigegeben sind.

We present an OCR ground truth data set for historical prints and show improvement of recognition results over baselines with training on this data. We reflect on reusability of the ground truth data set based on two experiments that look into the legal basis for reuse of digitized document images in the case of 19th century English and German books. We propose a framework for publishing ground truth data even when digitized document images cannot be easily redistributed.

1. Introduction

Digital access to Cultural Heritage is a key challenge for today's society. It has been improved by Optical Character Recognition (OCR), which is the task by which a computer program extracts text from a digital image in order to draw the text from that image and present it in a machine-readable form. For historical prints, off-the-shelf OCR solutions often result in inaccurate readings. Another impediment to accessing digitized cultural heritage data consists in the fact that cultural heritage institutions provide online access to massive amounts of digitized images of historical prints that have not been (or have been poorly) OCRed. Solutions to improve this situation would benefit a wide range of actors, be they scholars or a general audience. Many actors would indeed profit greatly from methods conceived to extract high quality machine-readable text from images.

The results of an OCR method can be improved significantly by using a pre-trained model and fine-tuning it on only a few samples that display similar characteristics.¹ To that end, there has been a growing effort from the Digital Humanities community to create and publish data sets for specific historical periods, languages and typefaces aiming at enabling scholars to fine-

¹ See Liebl / Burghardt 2020; Reul et al. 2017; Springmann et al. 2018.

tune OCR models for their collection of historical documents.² In Germany, the DFG-funded OCR-D initiative brings together major research libraries with the goal to create an open source framework for the OCR of historical printed documents, including specifications and guidelines for OCR ground truths.³

In order to improve OCR results, images and the corresponding transcriptions are collected in such a way that each pair (image and text) only represents one line of text from the original page. This is called a ground truth data set and is precisely what we will focus on in the following.

Besides the fact that creating transcriptions of images manually is tedious work, another major issue arises from this type of collective effort in that the institutions that produce the scan often claim some form of copyright to it. For example, on the first page of any of their PDFs, Google Books »[...] request[s] that you use these files for personal, non-commercial purposes«⁴. As a consequence, a scholar aiming to create an OCR ground truth data set would not know with certainty whether the rights to redistribute the textline images derived from the PDF can be considered as granted.

In this paper, we present an OCR ground truth data set with an unclear copyright setting for the image data. We discuss the legal background, show the relevance of the data set and provide in-depth analysis of its constitution and reuse by investigating two different approaches to overcome the copyright issues.

In order to address these issues, we compare in the following two ways to publish the OCR ground truth data set with image data.

- As Google Books works with cultural heritage institutions (CHIs) to digitize books, we asked permission from the CHIs to redistribute the image data.
- We published a data set formula, which consists of the transcriptions, links to the image sources, and a description on how to build the data set. For this process, we provide a fast, highly automated framework that enables others to reproduce the data set.

2. Legal background and its interpretation at CHIs

Clarifying the copyright situation for the scans of a book collection requires to take into account, for each book, the cultural heritage institution owning the book (usually a library), and, in the case of private-public partnerships, also the scanning institution (e. g. Google Books) involved in its digitization. For Google Books, there exist different contracts between CHIs and

² See Padilla et al. 2019. For manuscripts, just recently the Transcriptiones platform launched, see [transcriptiones](#), ETH-Library 2020. For French texts from the 18th to the 21st century there exists HTR-United, see [htr-United](#), Chagué / Clérice 2021. The slightly different approach of just publishing fine-tuned models for different settings is proposed by Transkribus, see [Transkribus](#), READ-COOP 2021, or Kraken 2021 [ocr_models](#), OCR/HTR model repository 2021.

³ See Engl 2020.

⁴ Google Inc. 2021, cited after Ruiz 2011.

Google, and not all of them are open to public inspection. However, based on comparing the ones that are available, we assume that other contracts are to some extent similar (see List of Contracts). The contracts contain information on the ›Library Digital Copy‹ for which non-profit uses are defined under Section 4.8 (cf. British Library Google Contract), which states that a

»Library may provide all or any portion of the Library Digital Copy, that is [...] a Digital Copy of a Public Domain work to (a) academic institutions or research libraries, or (b) when requested by Library and agreed upon in writing by Google, other not-for-profit or government entities that are not providing search or hosting services substantially similar to those provided by Google.«⁵

When trying to unpack this legal information against the use case presented here, multiple questions arise. What are the legal possibilities for individual scholars regarding the use of the Library Digital Copy of a Public Domain work? How can there be limitations in the use of a Public Domain work? Is the use case of OCR model training substantially similar to any search or hosting services provided by Google? Would and can libraries act as brokers in negotiating written agreements about not-for-profit use with Google?

In the continuation of Section 4.8, additional details are specified with regard to data redistribution by ›Additional institutions‹ where

»[a written agreement with Google] will prohibit such Additional institution from redistributing [...] portions of the Library Digital Copy to other entities (beyond providing or making content available to scholars and other users for educational or research purposes.«⁶

This brings up further questions but also opens the perspective a bit, since there appear to be exceptions for »scholars and other users for educational or research purposes«⁷, which is a precise fit of the use case we present here. Now what does this mean in practice? Digital Humanities scholars are not necessarily legal experts, so how do libraries that have entered public-private-partnerships with Google for digitization of Public Domain works implement these constraints? Schöch et al. discuss a wide range of use cases in the area of text and data mining with copyright protected digitized documents, but they do not cover the creation and distribution of ground truth.⁸ In other scenarios that involve copyrighted texts published in derived formats, one question typically preventing redistribution is whether it is possible to re-create the (copyright-protected) work from the derived parts. In the case of textline ground truth, it is however likely that this would constitute a violation of such a principle. In this unclear setting, scholars are in need of support and guidance by CHIs.

⁵ British Library Google Books Agreement in Ruiz 2011.

⁶ British Library Google Books Agreement in Ruiz 2011.

⁷ British Library Google Books Agreement in Ruiz 2011.

⁸ See Schöch et al. 2020.

Institution	Total # books	Total # pages	Response time (# working days)	Allowed to publish as part of the paper	Allowed to license	Alternative source	Responsible	Citation needed
Bayerische Staatsbibl.	4	12	3	yes	yes	yes	yes	yes
Biblioteca Statale Isontina Gorizia	1	3	-	-	-	-	-	-
Bodleian Library	11	20	2	yes, alternative	already CC-BY-NC	yes	yes	yes
British Library	1	35	4	no	no	no	yes	-
Harvard University, Harvard College Library	1	3	0	yes	yes	yes	no	yes
New York Public Library	5	29	3	-	-	no	no	no
Austrian National Library	2	6	10	yes, alternative	no	yes	yes	yes
Robarts - University of Toronto	2	3	-	-	-	-	-	-
University of Illinois Urbana-Champaign	6	4	0	yes	yes	no	yes	yes
University of Wisconsin - Madison	8	24	2	yes	yes	no	no	no

Tab. 1: Responses of library institutions to our request to grant permission to publish excerpts of the scans for which they were contractors of the digitization. Most institutions responded within a few working days and except for the fact that most acknowledged the public domain of the items, the responses were very diverse. Many answered that they are either not responsible or only responsible for their Library Copy of the PDF. [Lassner et al. 2021]

We have asked ten CHIs for permission to publish image data that was digitized based on their collection in order to publish them as part of an OCR ground truth data set under a CC-BY license. As shown in Table 1, the institutions gave a wide variety of responses. Many institutions acknowledged that the requested books are in the public domain because they were published before the year 1880. However, there is no general consensus on whether the CHIs are actually responsible for granting these rights, especially if one wants to use the copy from the Google Books or Internet Archive servers. Some institutions stated that they are only responsible for their Library Copy of the scan and granted permission to publish only from that source. Only two institutions, the Bayerische Staatsbibliothek and University of Illinois Urbana-Champaign stated that they are responsible and that we are allowed to also use the material that can be found on the Google Books or Internet Archive servers.

This case study underlines the lack of a clear and simple framework of reference that would be recognized and applied, and would reflect on good practices in the relationships between CHIs and digital scholarship. The lack of such a framework is addressed among others by the DARIAH initiative of the Heritage Data Reuse Charter⁹ that was launched in 2017. Another approach towards such a framework is that of the ›digital data librarian‹.¹⁰

3. Description of the data set

In the data set that we want to publish in the context of our OCR ground truth, we do not own the copyright for the image data.¹¹ We therefore distinguish between the data set formula and the built data set. We publish the data set formula which contains the transcriptions, the links to the images and a recipe on how to build the data set.

The data set formula and source code are published on Github¹² and the version 1.1 we are referring to in this paper is mirrored on the open access repository Zenodo.¹³ The data set is published under a CC-BY 4.0 license and the source code is published under an Apache license.

3.1 Origin

The built data set contains images from editions of books by Walter Scott and William Shakespeare in the original English and in translations into German that were published around 1830.

⁹ See Baillot et al. 2016. For additional information on the DARIAH Heritage Data Reuse Charter, see [data-reuse](#), DARIAH 2021.

¹⁰ See Eclevia et al. 2019.

¹¹ The current version of the data set can be found at [ocr-data/data](#), OCR-Data 2021.

¹² See [ocr-data](#), OCR-Data 2021.

¹³ See Lassner et al. 2021.

The data set was created as part of a research project that investigates how to implement stylometric methods that are commonly used to analyze the style of authors with the goal of analyzing that of translators. The data set was organized in such a way that other variables like authors of the documents or publication date can be ruled out as a confounder of the translator style.

We found that 1830 Germany was especially suitable for the research setting we had in mind. Due to an increased readership in Germany around 1830, there was a growing demand in books. Translating foreign publications into German turned out to be particularly profitable because, at that time, there was no copyright regulation that would apply equally across German-speaking states. There was no general legal constraint to regulate payments to the original authors of books or as to who was allowed to publish a German translation of a book. Therefore, publishers were competing in translating most recent foreign works into German, which resulted in multiple German translations by different translators of the same book at the same time. To be the first one to publish a translation into German, publishers resorted to what was later called translation factories, optimized for translation speed.¹⁴ The translators working in such ›translation factories‹ were not specialized in the translation of one specific author. It is in fact not rare to find books from different authors translated by the same translator.

3.2 Method

We identified three translators who all translated books from both Shakespeare and Scott, sometimes even the same books. We also identified the English editions that were most likely to have been used by the translators. This enabled us to set up a book-level parallel English-German corpus allowing us to, again, rule out the confounding author signal.

As the constructed data set is only available in the form of PDFs from Google Books and the Internet Archive or the respective partner institutions, OCR was a necessary step for applying stylometric tools on the text corpus. To assess the quality of off-the-shelf OCR methods and to improve the OCR quality, for each book, a random set of pages was chosen for manual transcription.

3.2.1 Preparation

Following the OCR-D initiative's specifications and best practices,¹⁵ for each book, we created a METS¹⁶ file that contains the link to the source PDF as well as the chosen pages. The following example presents an excerpt from one of the METS files:

¹⁴ See Bachleitner 1989.

¹⁵ See [ocr-d spec](#), OCR-D 2021.

¹⁶ See [METS](#), The Library of Congress 2021.

```
...
< mets:fileGrp USE="IMG">
  < mets:file ID="pdf_2jMFAAAMAA3_28" MIMETYPE="application/pdf">
    < mets:Flocat LOCTYPE="URL" xlink:href="http://books.google.com/books?id=2jMFAAAMAA3#page=28"/>
  </mets:file>
  < mets:file ID="pdf_2jMFAAAMAA3_103" MIMETYPE="application/pdf">
    < mets:Flocat LOCTYPE="URL" xlink:href="http://books.google.com/books?id=2jMFAAAMAA3#page=103"/>
  </mets:file>
</mets:fileGrp>
...
```

Fig. 1: Excerpt of a METS file as used in our data set. For each book, we created one METS file. The link to the resource contains the identifier and the page number. [Lassner et al. 2021]

The PDFs have been downloaded from the URLs in this METS file, and the page images have been extracted from the PDF, deskewed and saved as PNG files.¹⁷

3.2.2 Transcription

For transcription, the standard layout analyzer of Kraken 2.0.8 (depending on the layout either with black or white column separators) has been used and the transcription was pre-filled with either the German Fraktur or the English off-the-shelf model and post-corrected manually. To ensure consistency, some characters were normalized: for example, we encountered multiple hyphenation characters such as - and # which were both transcribed by -.

3.2.3 Size

In total, the data set contains 5,354 lines with 224,745 characters. It consists of German and English books from 1815 to 1852. A detailed description of the characteristics of the data set is shown in Table 2.

3.3 Reproducibility and Accessibility

The data set formula has been published as a collection of PAGE files and METS files.¹⁸ The PAGE files contain the transcriptions on line-level and the METS files serve as the container linking metadata, PDF sources and the transcriptions. There exists one METS file per item (corresponding to a Google Books or Internet Archive id) and one PAGE file per PDF page. The following excerpt of an example PAGE file shows how to encode one line of text:

```
...
< TextLine id="textline_2">
  <Coords points="457,124 457,1712 534,1712 534,124"/>
  <TextEquiv>
    <Unicode>wenn von ctarker Faut ein Stoß über das Schlü</Unicode>
  </TextEquiv>
</TextLine>
...
```

Fig. 2: Excerpt from the PAGE file showing the bounding box of the line on the page image and the corresponding text string. [Lassner et al. 2021]

¹⁷ The process is implemented in the pdfs.py submodule [pdfs.py:23](#) and it uses the command line tools `imagemagick` and `pdftimages`, see [OCR-Data 2021](#).

¹⁸ See [Pletschacher / Antonacopoulos 2010](#).

The `<TextLine>` `<TextLine>` contains the absolute pixel coordinates where the text is located on the preprocessed PNG image and the `<TextEquiv>` `<TextEquiv>` holds the transcription of the line.

As shown above, the METS files contain links to the PDFs. Additionally, the METS files contain links to the PAGE files as shown in the following excerpt.

```
<mets:fileGrp USE="GT">
  <mets:file ID="gt_2jMfAAAAAAAJ_28" MIMETYPE="text/xml">
    <mets:Flocat LOCTYPE="URL" xlink:href="data/xml_output/2jMfAAAAAAAJ_28.page"/>
  </mets:file>
  <mets:file ID="gt_2jMfAAAAAAAJ_3103" MIMETYPE="text/xml">
    <mets:Flocat LOCTYPE="URL" xlink:href="data/xml_output/2jMfAAAAAAAJ_3103.page"/>
  </mets:file>
  <mets:file ID="gt_2jMfAAAAAAAJ_3132" MIMETYPE="text/xml">
    <mets:Flocat LOCTYPE="URL" xlink:href="data/xml_output/2jMfAAAAAAAJ_3132.page"/>
  </mets:file>
  ...
</mets:fileGrp>
```

Fig. 3: Excerpt from the METS file as used in our data set. For each book, we created one METS file. This part of the METS file contains the references to the PAGE files. [Lassner et al. 2021]

As one can see, there are links from one METS file, namely the one encoding works by Walter Scott's, Volume 2, published by the Schumann brothers in 1831 in Zwickau, identified by the Google Books id `2jMfAAAAAAAJ_2jMfAAAAAAAJ`, to multiple pages (and PAGE files).

Finally, the METS file contains the relationship between the URLs and the PAGE files in the `<mets:structMap>` `<mets:structMap>` section of the file:

```
<mets:structMap>
  <mets:dIv ID="map_001">
    <mets:fptr FILEID="gt_2jMfAAAAAAAJ_28"/>
    <mets:fptr FILEID="pdf_2jMfAAAAAAAJ_28"/>
  </mets:dIv>
  ...
</mets:structMap>
```

Fig. 4: Excerpt from the METS file as used in our data set. For each book, we created one METS file. Together with the links to the image resources shown in Figure 1, and the links to the PAGE files, the METS file holds the connection between the text lines and the page images. [Lassner et al. 2021]

In order to reuse the data set, a scholar may then obtain the original image resources from the respective institutions as PDFs, based on the links we provide in the METS files. Then, the pair data set can be created by running the `>make pair_output<` command in the `>pipelines/<` directory. For each title, it extracts the PNG images from the PDF, preprocesses them, extracts, crops and saves the line images along respective files containing the text of the line.

Although the image data needs to be downloaded manually, the data set can still be compiled within minutes.

4. Framework for creating, publishing and reusing OCR ground truth data

We have published the framework we developed for the second case study, which enables scholars to create and share their own ground truth data set formulas when they are in the same situation of not owning the copyright for the images they use. This framework offers both directions of functionality:

- Creating an XML ground truth data set from transcriptions to share it with the public (data set formula) and
- Compiling an XML ground truth data set into standard OCR ground truth data pairs to train an OCR model (built data set).¹⁹

As already described in the Sections 3.2 and 3.3 there are multiple steps involved in the creation, publication and reuse of the OCR data set. In this Section, we would like to show that our work is not only relevant for scholars who want to reuse our data set but also for scholars who would like to publish a novel OCR ground truth data set in a similar copyright setting.

4.1 Creation and Publication

1. Corpus construction: selection of the relevant books and pages
2. Creation of the METS files²⁰
3. Transcription of the pages
4. Creation of the PAGE files²¹
5. Publication of the METS and the PAGE files

4.2 Reuse

1. Download of the METS and PAGE files
2. Download of the PDFs as found in the METS files
3. Creation of the pair data set²²
4. Training of the OCR models²³

In the Section 3.3, the steps listed in Reuse have been described. The download of the transcriptions and the PDFs has to be done manually but for the creation of the pair data set and the training of the models, automation is provided with our framework. We would like to also automatize the download of the PDFs; this, however, remains complicated to implement.

¹⁹ The documentation how to create a new or reproduce an existing data set can be found at [README.md](#), OCR-Data 2021.

²⁰ See [mets_page_template.xml](#), OCR-Data 2021.

²¹ See [create_xml_files.py](#), OCR-Data 2021.

²² See [extract_pair_dataset.py](#), OCR-Data 2021.

²³ See [train_ocr_model.py](#), OCR-Data 2021.

The first reason for this is a technical one: soon after starting the download, captchas appear (as early as by the 3rd image), which hinders the automatization. Another reason is the Google Books regulation itself. Page one of any Google Books PDF states explicitly:

»Keine automatisierten Abfragen. Senden Sie keine automatisierten Abfragen irgendwelcher Art an das Google-System. Wenn Sie Recherchen über maschinelle Übersetzung, optische Zeichenerkennung oder andere Bereiche durchführen, in denen der Zugang zu Text in großen Mengen nützlich ist, wenden Sie sich bitte an uns. Wir fördern die Nutzung des öffentlich zugänglichen Materials für diese Zwecke und können Ihnen unter Umständen helfen.«²⁴

Finding a way to automatize download could hence not be realized in the context of this project and will have to be addressed in future work.²⁵

Additionally, we provide useful templates and automation for the creation of a novel OCR ground truth data set. As already described, we used the Kraken transcription interface to create the transcription. In Kraken, the final version of the transcription is stored in HTML files. We provide a script to convert the HTML transcriptions into PAGE files in order to facilitate interoperability with other OCR ground truth data sets.

Finally, the pair data set can be created from the PAGE transcriptions and the images of the PDFs and the OCR model can be trained.

5. Relevance of the data set

In order to evaluate the impact that the data set has on the accuracy of OCR models, we trained and tested model performance in three different settings. In the first setting, we fine-tuned an individual model for each book in our corpus using a training and an evaluation set of that book and tested the performance of the model on a held-out test set from the same book. In Table 2, we show how this data set has dramatically improved the OCR accuracy on similar documents compared to off-the-shelf OCR solutions. Especially in cases where the off-the-shelf model (baseline) shows a weak performance, the performance gained by fine-tuning is large.

In the second and third setting, we split the data set into two groups: English Antiqua, German Fraktur. There was also one German Antiqua book that we did not put into any of the two groups. For the second setting, we split all data within a group randomly into train set, evaluation set and test set and trained and tested an individual model for each group. In Table 3, the test performance of this setting is shown. For both groups, the fine-tuning improves

²⁴When downloading any book PDF from Google Books one page is prepended to the document. On this page, the cited usage statement is presented. As an example, please consider [Walter Scott's Werke](#), see Google Inc. 2006.

²⁵ Our progress on this topic will be documented in issue 2 of our [github repository](#), see OCR-Data 2021.

the character accuracy by a large margin over the baseline accuracy. This experiment shows that overall, the fine-tuning within a group improves the performance of that group and that patterns are learned across individual books.

Google Books or Internet Archive identifier	baseline model	Train # lines	Test # lines	Train # chars	Test # chars	baseline character accuracy	fine-tuned character accuracy	δ
rDUJAAAAQA	_best	82	11	3520	493	99.8	100.0	0.2
chroniclesof	_best2sc20		3	836	97	100.0	100.0	0.0
anneofgeier	_best3sc20		3	805	138	100.0	100.0	0.0
_QgOAAAAQA	_best	60	8	2659	359	95.54	100.0	4.46
chroniclesof	_best3sc40		5	1766	185	99.46	99.46	0.0
zviTtwEAC	faktur_1_66st		9	3396	519	98.27	99.23	0.96
quentindun	enabde2sc30oft		5	1748	241	99.17	99.17	0.0
3pVMAAAA	faktur_1_92st		12	4830	598	96.49	99.16	2.67
2jMfAAAA	faktur_1_65st		20	7386	939	93.5	98.94	5.44
t88yAQAA	faktur_1_84st		11	3345	436	94.5	98.85	4.35
HCRMAAA	faktur_1_125st		16	5100	579	92.23	98.79	6.56
zDTMtgeA	faktur_1_76st		10	4277	560	93.93	98.75	4.82
DNUwAQAA	faktur_1_76st		10	4147	517	94.58	98.45	3.87
H9UwAQAA	faktur_1_76st		10	4017	533	97.19	98.31	1.12
AdiKyqdlp	faktur_1_77st		10	2827	405	92.84	98.27	5.43
J4knAAAA	_best	20	3	851	104	97.12	98.08	0.96
aNQwAQAA	faktur_1_52st		7	2752	309	95.79	98.06	2.27
XtEyAQAA	faktur_1_86st		11	3489	383	94.52	97.91	3.39
D5pMAAAA	faktur_1_88st		12	4557	546	93.22	97.8	4.58
8AQoAAAA	faktur_1_74st		9	3130	434	94.93	97.7	2.77
Fy4JAAAA	_best	20	3	743	125	96.0	97.6	1.6
anneofgeier	_best2sc42		6	1747	204	98.04	97.55	-0.49
u4cnAAAA	faktur_1_76st		10	3936	553	91.5	97.11	5.61

Tab. 2: Performance comparison of baseline model and fine-tuned model for each document in our corpus. For almost all documents there is a large improvement over the baseline even with a very limited number of fine-tuning samples. The sum of lines and characters depicted in the table do not add up to the numbers reported in the text because during training we used an additional split of the data as an evaluation set that had the same size as the test set respectively. [Lassner et al. 2021]

1VUJAAAA	QA_best	85	11	3899	455	94.73	96.7	1.97
quentindun	en_best	20	3	708	86	95.35	95.35	0.0
4zQfAAAA	fraktur_1_best	59	20	6817	932	87.98	94.74	6.76
7JVMAAAA	fraktur_1_best	12	12	4604	616	65.91	94.32	28.41
YAZXAAAA	fraktur_1_best	512	219	66253	8327	80.17	93.61	13.44
8dAyAQAA	fraktur_1_best	12	12	3448	380	87.11	93.42	6.31
PzMJAAAA	QA_best	61	8	2294	234	90.17	92.74	2.57
wggOAAAA	QA_best	19	3	716	94	91.49	92.55	1.06
WjMfAAAA	fraktur_1_best	33	23	7363	814	71.62	91.52	19.9
MzQJAAAA	QA_best	36	5	1265	201	88.56	90.55	1.99
fAoOAAAA	QA_best	40	6	1675	121	86.78	87.6	0.82
kggOAAAA	QA_best	40	6	1572	243	82.72	82.72	0.0
oNEyAQAA	fraktur_1_best	73	10	2874	386	68.39	79.02	10.63
htQwAQAA	fraktur_1_best	78	10	3990	464	69.18	78.02	8.84

Tab. 2: Performance comparison of baseline model and fine-tuned model for each document in our corpus. For almost all documents there is a large improvement over the baseline even with a very limited number of fine-tuning samples. The sum of lines and characters depicted in the table do not add up to the numbers reported in the text because during training we used an additional split of the data as an evaluation set that had the same size as the test set respectively. [Lassner et al. 2021]

Document Group	baseline model	Train # lines	Test # lines	Train # chars	Test # chars	baseline character accuracy	fine-tuned character accuracy	δ
English Antiqua	en_best	650	82	26793	3406	94.19	96.21	2.02
German Fraktur	fraktur_1_best	649	432	145928	17577	85.89	95.99	10.1

Tab. 3: Performance comparison of baseline model and fine-tuned model trained on a random splits of samples within the same group. [Lassner et al. 2021]

Left-out identifier	baseline model	Train # lines	Test # lines	Train # chars	Test # chars	baseline character accuracy	fine-tuned character accuracy	δ
---------------------	----------------	---------------	--------------	---------------	--------------	-----------------------------	-------------------------------	----------

Tab. 4: Model performance evaluated with a leave-one-out strategy. Within each group (German Fraktur and English Antiqua), an individual model is trained on all samples except from the left-out identifier on which the model is tested afterwards. The performance of the fine-tuned model is improved in each case, often by a large margin. [Lassner et al. 2021]

chroniclesofbabel3sc686		50	28134	2182	99.22	99.59	0.37
H9UwAQAAFraktur_1_6594		96	159088	5130	96.74	99.57	2.83
aNQwAQAAFraktur_1_6822		65	161053	3397	97.0	99.53	2.53
chroniclesofbabel2sc709		25	29226	1017	99.02	99.51	0.49
zDTMtgeACFraktur_1_6594		96	159131	5430	95.05	99.43	4.38
anneofgeiersties3sc708		26	29144	1062	98.68	99.34	0.66
t88yAQAAFraktur_1_6586		105	160286	4181	91.13	99.28	8.15
anneofgeiersties2sc684		53	28053	2181	98.3	99.27	0.97
DNUwAQAAFraktur_1_6594		96	159113	5228	95.26	99.01	3.75
D5pMAAAFraktur_1_6580		111	159386	5660	93.69	99.01	5.32
3pVMMAAAFraktur_1_6577		115	158561	6036	94.68	98.99	4.31
zviTtwEACFraktur_1_6806		83	159741	4384	95.76	98.97	3.21
8AQoAAAAFraktur_1_6800		89	160966	3926	94.7	98.9	4.2
1VUJAAAAQAAbest	635	107	25735	4839	96.88	98.8	1.92
AdiKyqdlp4Fraktur_1_6593		97	160065	3736	92.34	98.47	6.13
rDUJAAAAQAAbest	639	103	26265	4419	97.85	98.42	0.57
quentindurwab92sc687ft		49	28274	2223	97.35	98.34	0.99
HCRMMAAAFraktur_1_6539		157	158250	6378	91.28	98.28	7.0
J4knAAAAQAAbest	708	26	29219	1089	97.15	98.07	0.92
2jMfAAAAFraktur_1_6573		197	155342	9181	92.43	98.04	5.61
XtEyAQAAFraktur_1_6583		108	160349	4322	87.69	97.59	9.9
quentindurwab91sc708ft		26	29284	940	96.38	97.13	0.75
wggOAAAAQAAbest	710	24	29362	869	92.52	96.89	4.37
_QgOAAAAQAAbest	664	75	27117	3320	94.43	96.66	2.23
fAoOAAAAQAAbest	685	51	28128	2007	94.72	96.61	1.89
4zQfAAAAFraktur_1_6570		199	156399	8681	88.68	96.37	7.69
PzMJAAAAQAAbest	662	77	27724	2817	90.7	95.49	4.79
u4cnAAAAFraktur_1_6595		95	159827	4889	91.31	95.21	3.9
7JVMMAAAFraktur_1_6580		112	159080	5816	71.35	94.62	23.27
8dAyAQAAFraktur_1_6580		111	159841	4271	84.45	94.24	9.79

Tab. 4: Model performance evaluated with a leave-one-out strategy. Within each group (German Fraktur and English Antiqua), an individual model is trained on all samples except from the left-out identifier on which the model is tested afterwards. The performance of the fine-tuned model is improved in each case, often by a large margin. [Lassner et al. 2021]

htQwAQAA	Fraktur_1_b592	98	158623	4996	88.42	94.14	5.72	
YAZXAAAA	Fraktur_1_b509	2190	89328	82910	80.68	92.92	12.24	
MzQJAAAA	Antiqua_best	691	45	28714	1622	84.9	89.52	4.62
kggOAAAA	Antiqua_best	685	51	28216	1983	85.64	87.56	1.92
Fy4JAAAA	Antiqua_best	709	25	29424	943	78.9	85.15	6.25
oNEyAQAA	Fraktur_1_b598	92	160955	3589	66.31	84.79	18.48	

Tab. 4: Model performance evaluated with a leave-one-out strategy. Within each group (German Fraktur and English Antiqua), an individual model is trained on all samples except from the left-out identifier on which the model is tested afterwards. The performance of the fine-tuned model is improved in each case, often by a large margin. [Lassner et al. 2021]

In the third setting, we trained multiple models within each group, always training on all books of that group except one and using only the data of the left-out book for testing. In all settings, we also report the performance of the off-the-shelf OCR model on the test set for comparison.

As depicted in Table 4, the performance of fine tuning improves character accuracy each time even for the held-out book. This shows that the fine-tuned model indeed did not overfit on a specific book but captures patterns of a specific script. We should note, that in some cases of the third experiment different volumes occur as individual samples, for example, the second volume of Anne of Geierstein by Scott was not held-out when tested for the third volume of Anne of Geierstein. Scripts in different volumes are often more similar than scripts of the same font type which might improve the outcome of this experiments in some cases.

For all three experiments, the Kraken OCR engine with a German Fraktur model and an English model was used as baselines. They were provided by the maintainers of Kraken.²⁶

In the context of the research project for which this data set was created, the performance gain is especially relevant as research shows that a certain level of OCR quality is needed in order to be able to obtain meaningful results on downstream tasks. For example, Hamdi et al. show the importance of OCR quality on the performance of Named Entity Recognition as a downstream task.²⁷ With additional cross training of sub-corpora we are confident that we will be able to push the character accuracy beyond 95% on all test sets that will enable us to perform translatorship attribution analysis.

More generally, the results show that in a variety of settings, additional ground truth data will improve the OCR results. This advocates strongly for the publication of a greater range of, and especially more diverse, sets of open and reusable ground truth data for historical prints.

²⁶ See Kiessling 2019. For baselines and fine-tuning version 3.0.4 of the Kraken engine was used that can be found at [kraken release 3.0.4](#), Kiessling 2021.

²⁷ See Hamdi et al. 2020.

The data set we thus created and published is open and reproducible following the described framework. It can serve as a template for other OCR ground truth data set projects. It is therefore not only relevant because it shows why the community should create additional data sets: it also shows how to create the data sets and invites to new publications bound to bring Digital Humanities research a step forward.

The data pairs are compatible with other OCR ground truth data sets such as e. g. OCR-D²⁸ or GT4HistOCR²⁹. Using the established PAGE-XML standard enables interoperability and reusability of the transcriptions. Using open licenses for the source code and the data, and publishing releases at an institutional open data repository ensures representativeness and durability.

6. Conclusion

The work we realized in order to constitute the data set we need for our stylometric research provided not only a ground truth data set, but also a systematic approach to the legal issues we encountered in the extraction of information from the scanned books we rely on as a primary source. While we have been successful at automating many work steps, improvements could still be envisioned.

In future work, we would like to enrich the links to the original resource with additional links to mirrors of the resources in order to increase the persistence of the image sources, whenever available also adding OCLC IDs as universal identifiers.³⁰ We would also like to look into ways to automate the download of the PDFs from Google Books, the Internet Archive or CHIs. Also, we would like to extend the framework we proposed here. It could serve for hybrid data sets with parts where the copyright for the image data is unclear (then published as data set formula), and others with approved image redistribution (which could then be published as a built data set). It could be used for example for the datasets from Bayerische Staatsbibliothek and University of Illinois Urbana-Champaign.

Finally, we would like to encourage scholars to publish their OCR ground truth data set in a similarly open and interoperable manner, thus making it possible to ultimately increase accessibility to archives and libraries for everyone.

Acknowledgements

This work has been supported by the German Federal Ministry for Education and Research as BIFOLD.

²⁸ See Baierer et al. 2019.

²⁹ See Springmann et al. 2018.

³⁰ OCLC is a registry of IDs referencing items in libraries, see worldcat.org, OCLC 2021.

List of contracts

The contracts between

- a number of US-based libraries and Google is available [here](#),
- the British Library and Google is available [here](#),
- the National Library of the Netherlands and Google is available [here](#),
- the University of Michigan and Google is available [here](#),
- the University of Texas at Austin and Google is available [here](#),
- the University of Virginia and Google is available [here](#),
- Scanning Solutions (for the Bibliotheque Municipale de Lyon) and Google is available [here](#),
- University of California and Google is available [here](#).

Bibliographic references

- Norbert Bachleitner: »Übersetzungsfabriken«: das deutsche Übersetzungswesen in der ersten Hälfte des 19. Jahrhunderts. In: Internationales Archiv für Sozialgeschichte der deutschen Literatur 14 (1989), i. 1, pp. 1–50. [[Nachweis im GBV](#)]
- Anne Baillot / Mike Mertens / Laurent Romary: Data fluidity in DARIAH – pushing the agenda forward. In: Bibliothek Forschung und Praxis 39 (2016), i. 3, pp. 350–357. DOI: [10.1515/bfp-2016-0039](https://doi.org/10.1515/bfp-2016-0039) [[Nachweis im GBV](#)]
- Konstantin Baierer / Matthias Boenig / Clemens Neudecker: Labelling OCR Ground Truth for Usage in Repositories. In: Proceedings of the International Conference on Digital Access to Textual Cultural Heritage (DATeCH2019: 3, Brussels, 08.–10.05.2019) New York, NY 2019, pp. 3–8. [[Nachweis im GBV](#)]
- HTR-United. In: GitHub.io. By Alix Chagué / Thibault Clérice. 2021. [[online](#)]
- Marian Ramos Eclevia / John Christopher La Torre Fredeluces / Carlos Jr Lagrosas Eclevia / Roselle Saguibo Maestro: What Makes a Data Librarian? An Analysis of Job Descriptions and Specifications for Data Librarian. In: Qualitative and Quantitative Methods in Libraries 8 (2019), n. 3, pp. 273–290. [[online](#)]
- Elisabeth Engl: Volltexte für die Frühe Neuzeit. Der Beitrag des OCR-D-Projekts zur Volltexterkennung frühneuzeitlicher Drucke. In: Zeitschrift für Historische Forschung 2 (2020), n. 47, pp. 223–250. [[Nachweis im GBV](#)]
- Transcriptiones. A platform for hosting, accessing and sharing transcripts of non-digitised historical manuscripts. Ed. by ETH-Library. Zürich 2020. [[online](#)]
- Ahmed Hamdi / Axel Jean-Caurant / Nicolas Sidère / Mickaël Coustaty: Assessing and Minimizing the Impact of OCR Quality on Named Entity Recognition. In: Digital libraries for open knowledge. International Conference on Theory and Practice of Digital Libraries. (TPDL: 24, Lyon, 25.–27.08.2020) Cham 2020, pp. 87–101. [[Nachweis im GBV](#)]
- The Heritage Data Reuse Charter. In: DARIAH.eu. 2021. [[online](#)]
- Informationen und Richtlinien. Ed. by Google Inc. In: Google Books. Walter Scott: Großvater's Erzählungen aus der Geschichte von Frankreich. Ed. by Georg Nicolaus Bärmann. Neue Folge. Zweiter Theil. Zwickau 1831. Digitalisiert am 15.11.2006. PDF. [[online](#)]
- Benjamin Kiessling: Kraken – an Universal Text Recognizer for the Humanities. In: Digital Humanities 2019 Conference papers. (DH2019, Utrecht, 08.–12.07.2019) Utrecht 2019. [[online](#)]
- Kraken 3.0.4. In: GitHub.io. Ed. by Benjamin Kiessling. 2021. [[online](#)]
- David Lassner / Julius Coburger / Clemens Neudecker / Anne Baillot: Data set of the paper »Publishing an OCR ground truth data set for reuse in an unclear copyright setting«. In: zenodo.org. 2021. Version 1.1 from 07.05.2021. DOI: [10.5281/zenodo.4742068](https://doi.org/10.5281/zenodo.4742068)
- METS. Metadata Encoding & Transmission Standard. Home. Ed. by The Library of Congress. Washington D.C. 04.10.2021. [[online](#)]
- Bernhard Liebl / Manuel Burghardt: From Historical Newspapers to Machine-Readable Data: The Origami OCR Pipeline. In: Proceedings of the Workshop on Computational Humanities Research. Ed. by Folgert Karsdorp / Barbara McGillivray / Adina Nerghe / Melvin Wevers. (CHR2020, Amsterdam, 18.–20.11.2020), Aachen 2020, pp. 351–373. (= CEUR Workshop Proceedings, 2723) URN: [urn:nbn:de:0074-2723-3](https://nbn-resolving.org/urn:nbn:de:0074-2723-3)
- OCR-Data. In: GitHub.io. 2021. [[online](#)]
- OCR-D. Specifications. In: OCR-D.de. Wolfenbüttel 2021. [[online](#)]
- OCR/HTR model repository. In: Zenodo.org. 2021. [[online](#)]
- WorldCat. Ed. by OCLC. Dublin 2021. [[online](#)]
- Thomas Padilla / Laurie Allen / Hannah Frost / Sarah Potvin / Elizabeth Russey Roke / Stewart Varner: Final Report – Always Already Computational: Collections as Data. In: zenodo.org. Version 1 from 22.05.2019. DOI: [10.5281/zenodo.3152935](https://doi.org/10.5281/zenodo.3152935)
- Stefan Pletschacher / Apostolos Antonacopoulos: The PAGE (Page Analysis and Ground-Truth Elements) Format Framework. In: Proceedings of the 20th International Conference on Pattern Recognition. Ed. by IEEE. (ICPR: 20, Istanbul, 23.–26.08.2010) Piscataway, NJ 2010, vol. 1, pp. 257–260. [[Nachweis im GBV](#)]
- Public AI models in Transkribus. Ed. by READ-COOP. Innsbruck 2021. [[online](#)]
- Christian Reul / Christoph Wick / Uwe Springmann / Frank Puppe: Transfer Learning for OCRopus Model Training on Early Printed Books. In: Zeitschrift für Bibliothekskultur 5 (2017), i. 1, pp. 32–45. In: zenodo.org. Version 1 from 22.12.2017. DOI: [10.5281/zenodo.4705364](https://doi.org/10.5281/zenodo.4705364)
- Javier Ruiz: Access to the Agreement between Google Books and the British Library. In: Open Rights Group. Ed. by The Society of Authors. Blogpost from 24.08.2011. [[online](#)]

Christof Schöch / Frédéric Döhl / Achim Rettinger / Evelyn Gius / Peer Trilcke / Peter Leinen / Fotis Jannidis / Maria Hinzmann / Jörg Röpke: Abgeleitete Textformate: Text und Data Mining mit urheberrechtlich geschützten Textbeständen. In: Zeitschrift für digitale Geisteswissenschaften 5 (2020). DOI: [10.17175/2020_006](https://doi.org/10.17175/2020_006)

Uwe Springmann / Christian Reul / Stefanie Dipper / Johannes Baiter: Ground Truth for training OCR engines on historical documents in German Fraktur and Early Modern Latin. In: The Journal for Language Technology and Computational Linguistics 33 (2018), i. 1, pp. 97–114. PDF. [\[online\]](#)

List of Figures with Captions

Tab. 1: Responses of library institutions to our request to grant permission to publish excerpts of the scans for which they were contractors of the digitization. Most institutions responded within a few working days and except for the fact that most acknowledged the public domain of the items, the responses were very diverse. Many answered that they are either not responsible or only responsible for their Library Copy of the PDF. [Lassner et al.2021]

Abb. 1: Excerpt of a METS file as used in our data set. For each book, we created one METS file. The link to the resource contains the identifier and the page number. [Lassner et al. 2021]

Abb. 2: Excerpt from the PAGE file showing the bounding box of the line on the page image and the corresponding text string. [Lassner et al. 2021]

Abb. 3: Excerpt from the METS file as used in our data set. For each book, we created one METS file. This part of the METS file contains the references to the PAGE files. [Lassner et al. 2021]

Abb. 4: Excerpt from the METS file as used in our data set. For each book, we created one METS file. Together with the links to the image resources shown in Figure 1, and the links to the PAGE files, the METS file holds the connection between the text lines and the page images. [Lassner et al. 2021]

Tab. 2: Performance comparison of baseline model and fine-tuned model for each document in our corpus. For almost all documents there is a large improvement over the baseline even with a very limited number of fine-tuning samples. The sum of lines and characters depicted in the table do not add up to the numbers reported in the text because during training we used an additional split of the data as an evaluation set that had the same size as the test set respectively. [Lassner et al. 2021]

Tab. 3: Performance comparison of baseline model and fine-tuned model trained on a random splits of samples within the same group. [Lassner et al. 2021]

Tab. 4: Model performance evaluated with a leave-one-out strategy. Within each group (German Fraktur and English Antiqua), an individual model is trained on all samples except from the left-out identifier on which the model is tested afterwards. The performance of the fine-tuned model is improved in each case, often by a large margin. [Lassner et al. 2021]

Beitrag aus:

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. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Visuelle modellbasierte Experimente in den (digitalen) Geisteswissenschaften

Autor*in:

Zoe Schubert

Kontakt: zoe.schubert@uni-koeln.de

Institution: Universität zu Köln, Institut für Digital Humanities | Stiftung Preußischer Kulturbesitz
GND: [124777337X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0001-9043-3632](https://orcid.org/0000-0001-9043-3632)

Autor*in:

Jan Wieners

Kontakt: jan.wieners@uni-koeln.de

Institution: Universität zu Köln, Institut für Digital Humanities
GND: [1067253424](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0002-7176-1048](https://orcid.org/0000-0002-7176-1048)

Autor*in:

Øyvind Eide

Kontakt: oeide@uni-koeln.de

Institution: Universität zu Köln, Institut für Digital Humanities und Center for Data and Simulation Science
GND: [1084655063](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0002-7766-6287](https://orcid.org/0000-0002-7766-6287)

DOI des Artikels:

[10.17175/sb005_007](https://doi.org/10.17175/sb005_007)

Nachweis im OPAC der Herzog August Bibliothek:

[1782376437](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9)

Erstveröffentlichung:

17.01.2022

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise: 20.12.2021

GND-Verschlagwortung:

[Datenmodell](#) | [Experiment](#) | [Medienwissenschaft](#) | [Transformation](#) | [Visualisierung](#) |

Zitierweise:

Zoe Schubert, Jan Wieners, Øyvind Eide: Visuelle modellbasierte Experimente in den (digitalen) Geisteswissenschaften. 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/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_007](https://doi.org/10.17175/sb005_007).

Zoe Schubert, Jan Wieners, Øyvind Eide

Visuelle modellbasierte Experimente in den (digitalen) Geisteswissenschaften

Abstracts

Dieser Beitrag stellt diverse visuelle Experimente vor, bei denen insbesondere mediale Transformationsprozesse und sowohl Vorgang als auch Ergebnis von Modellierungsprozessen im Fokus stehen. Bei den ausgewählten Beispielen entstehen durch den Einsatz und die Entwicklung von Softwaresystemen unterschiedliche visuelle Artefakte. Zentral ist dabei, dass die Experimente nur auf Grundlage von Modellen und in Verbindung mit Modellierungsprozessen durchgeführt werden können. Dies ist substantiell für den Erkenntnisgewinn und die analytischen Betrachtungen insgesamt, wie in diesem Beitrag aufgezeigt wird. Darüber hinaus wird dargelegt, dass sich visuelle Experimente in den digitalen Geisteswissenschaften nie ohne Modelle unterschiedlichster Arten durchführen lassen.

This article presents a number of visual experiments with a special focus on processes of media transformation and modelling. In the selected examples, the development and use of computer systems are used to create different visual artefacts. These experiments can only be performed with the use of models in modelling processes. This chapter shows how important modelling and the use and development of software is for the analytical reflections and thus for the gaining of new knowledge. We also show how these points tend towards a more general pattern: visual experiments in the digital humanities cannot be performed without the use of models.

1. Kollaborative Experimente in geteilten Umgebungen – von Medienprodukten und Modellen

Die im Folgenden in den Blick genommenen Experimente werden alle in einem (virtuellen) Setting umgesetzt, welches eine gemeinsam geteilte Umgebung für die Teilnehmenden schafft: Es ist möglich, kollaborativ an demselben Experiment zu arbeiten, auch wenn dies mit unterschiedlichen Auffassungen hinsichtlich der Bedeutung des Experiments getan wird. Es geht dementsprechend nicht um reine Gedankenexperimente, wenngleich solche das Teilen der individuellen Resultate (und eben nur dieser) ermöglichen.

Jene Experimente befassen sich mit medialen Transformationen: text- und filmbasierte Karten sowie VR-Systeme, die Theateraufführungen replizieren. Studierende und Forscher*innen implementieren digitale visuelle Systeme mit dem Zweck, nicht-digitale Medienprodukte tiefgreifend zu verstehen. Darüber hinaus werden Unterschiede zwischen digitalen und prä-digitalen Medien reflektiert.

Es lässt sich ableiten, dass die hier vorgestellten Experimente in Verbindung und basierend auf existierenden, gemeinsam rezipierten Objekten und Phänomenen (Texte, Filme und Ausführungen) durchgeführt werden. Aus digital-geisteswissenschaftlicher Perspektive lassen sie sich als intersubjektiv rezipierbare Prozesse konstatieren, welche mit digitalen Objekten

durchgeführt werden, die über materielle Eigenschaften verfügen. Dieser Materialität der Objekte im Digitalen wird auf sehr unterschiedliche Art Ausdruck verliehen, sie sind allerdings in jedem der aufgezeigten Fälle als Teil von Kommunikationsprozessen identifizierbar und zu verstehen. Die Objekte lassen sich Lars Elleström zufolge als Medienprodukte benennen und beschreiben.¹ Sie repräsentieren hier für sich genommen andere geisteswissenschaftliche Objekte und Prozesse: Entsteht im Rahmen eines Experiments beispielsweise eine literarische Karte, die räumliche Verbindungen in einem literarischen Werk repräsentiert, lässt sich diese wiederum in einem weiteren Experiment zur literarischen Interpretation verwenden. Die Karte repräsentiert dann im Kontext des folgenden Experiments narrative und räumliche Verhältnisse des ihr zugrundeliegenden Texts. Ein weiteres Beispiel lässt sich aus der wissenschaftlichen Arbeit innerhalb der Archäologie ableiten. Es wird hier mit 3D-Modellen prähistorischer Objekte experimentiert, weil diese 3D-Modelle prähistorische Objekte repräsentieren. Diese beispielhaft skizzierten Repräsentationen sind Modelle. Sie zeigen auf, dass visuelle Experimente in den digitalen Geisteswissenschaften (DH) mit Modellen geisteswissenschaftlicher Objekte und Prozesse vorgenommen werden.

Im Kontext der in diesem Artikel geschilderten Experimente werden Medienprodukte in den Blick genommen. Jene Medienprodukte fußen auf den vorgenommenen unterschiedlichen Konfigurationen medialer Modalitäten. Es ließen sich demnach ähnliche Fragen zur Gestaltung der modellierten Medienprodukte (Ishiguro, Brecht, Murakami; Roman, Theatertext, Novelle) sowie den in den zugehörigen Experimenten umgesetzten Modellen stellen. Die Vorgehensweise in dem vorliegenden Beitrag fokussiert sich auf den Prozess der experimentellen Modellierung im Sinne von medialen Transformationen mit einem Fokus auf mediale Modalitäten.

Das Verhältnis zwischen Modellierung und Experimenten wird im Kontext vieler wissenschaftlicher Auseinandersetzungen diskutiert.² Modellierung lässt sich als ein Grundbegriff der digitalen Geisteswissenschaften verstehen.³ Im *praktischen Entwicklungsbereich* der DH dienen Modelle als notwendige Bestandteile der Vorbereitung und Durchführung von Softwareentwicklungsprojekten. Bei der *Datenanalyse* innerhalb der DH ist die Modellierung von Daten ein notwendiger Bestandteil zur Erstellung und Verwendung von Datensätzen. Im *theoretischen Verständnis* von DH ist Modellierung ein analytischer Kernbegriff. Basierend auf den beschriebenen Verbindungen zwischen Modellen und Experimenten, sowie der wichtigen Rolle von Modellen in den Geisteswissenschaften werden im Folgenden drei praktische Experimente genauer beschrieben und ausgewertet. Sie alle betrachten Softwareentwicklung als eine Forschungsmethode, die zu einem tieferen Verständnis von Medienprodukten beiträgt. Zentral sind dabei die Unterschiede zwischen einzelnen Medien und mediale Transformationen. Zwei Aspekte werden hierbei in den Fokus gestellt:

1. Digitale Medialität (Materialität) der Modelle und
2. Die Verbindungen zwischen Modellen und modellierten Objekten und Prozessen, was essentiell für die Möglichkeit und Relevanz der Experimente ist.

¹ Elleström 2021

² Vgl. z. B. Peschard / Fraassen (Hg.) 2018.

³ McCarty 2015; Ciula et al. 2018.

Verbindungen zwischen Medialität und Materialität wurden bereits aufschlussreich und tiefgreifend im Kontext der Semiotik analysiert.⁴ Die Medialität als Materialität ist eng mit der Frage zur Ähnlichkeit der Gestalt des Modells und des im Experiment modellierten Systems verbunden; insgesamt wird das theoretische Verständnis von Experimenten in den digitalen Geisteswissenschaften vorläufig konstatiert. Die drei folgenden Beispiele zur Verwendung von Modellen im Kontext von Experimenten werden dafür ausführlicher beschrieben.

2. Kartenbasierte Experimente in der Textanalyse

Sowohl Texte als auch Karten können zum Ausdruck konkreter räumlicher Verhältnisse und Verbindungen sowie zur allgemeinen Erfassung eines Raumes herangezogen werden. Sie unterscheiden sich dabei allerdings offenkundig in ihrer Ausdrucksform. Im Folgenden werden Experimente beschrieben, die entsprechende Unterschiede in den Mittelpunkt stellen und darauf aufbauen, was wiederum dazu dient, ein tiefgreifendes Textverständnis zu entwickeln. Es geht dabei weniger um einen ausgewogenen Vergleich zwischen Karten und Text, sondern vielmehr darum, Kartierung als Methodik und damit als Teil von Untersuchungen zur Textualität anzuwenden. Unabhängig von diesem zentralen Ziel wird durch den Modellierungsprozess selbst auch gleichzeitig ein tieferes Verständnis des Modells, hier also der Karte, sowie dessen Medialität gefördert. Besonders bedeutsam ist es, nachvollziehbar aufzuzeigen, wie das Modellieren von Karten als ein experimentelles Forschungswerkzeug dienen kann.

Für die experimentellen Untersuchungen werden in einem ersten Schritt textuelle Phänomene identifiziert, die räumliche Informationen ausdrücken. Für diese werden dann kartografische, modellhafte Entsprechungen etabliert. Durch systematische, schrittweise vorgenommene Formalisierungen können Textstellen anschließend zu Karten medial transformiert werden. Als Grundlage solcher Transformationsprozesse können sehr spezifische Aussagen im Text verwendet werden, wie in Tabelle 1 veranschaulicht.

	1	2	3	4	5
α	Aussage im Text	Räumliches Verhältnis	Mathematisch ausgedrücktes räumliches Verhältnis	Raumkoordinaten	Kartographisches Modell
β	A befindet sich nordöstlich von B	B \rightarrow Richtung Nord-Ost \rightarrow A	Richtung B \rightarrow A: 45°	A: (1140, 820) B: (520, 200)	\odot A \odot B

Tab. 1: Kritische schrittweise Formalisierung. [Eide et al. 2021 nach Eide 2015]

⁴ Ciula / Eide 2017.

Von elementarer Bedeutung hinsichtlich der Relevanz der hier beschriebenen experimentellen Untersuchungen ist es, ein Verständnis dafür zu entwickeln, dass β_5 nicht die exakt gleiche Aussage wie β_1 trifft. Vielmehr unterscheiden sie sich ebenso, wie α_1 und α_5 voneinander abweichen. Dies bedeutet, dass die durch systematische Formalisierung dokumentierten Unterschiede der medialen Produkte herangezogen werden können, um generelle mediale Differenzen zwischen verschiedenartigen Medientypen zu analysieren. Die Zwischenschritte bei der medialen Transformation dienen dazu, feingranular die Veränderungen und Unterschiede bei jeder schrittweisen Übertragung deutlicher zu machen. Diese Möglichkeit, konkrete textuelle Ausdrücke (die Quellen des Transformationsprozesses) zu verwenden, macht den experimentellen Charakter der modellbasierten Untersuchungen aus. Bei einer vigilant durchgeführten schrittweisen Formalisierung ist das Endresultat – Spalte 5 – nicht das primäre zu erreichende und einzig zu analysierende Ziel, es handelt sich dabei vielmehr um ein experimentelles, variables und nicht allein richtiges Resultat der experimentellen medialen Transformationen. Elementare Forschungsergebnisse lassen sich insbesondere in den durchgeführten Mikro-Transformationen $\beta_1 \# \beta_2$, $\beta_2 \# \beta_3$, usw. ableiten. Hier können Relikte identifiziert werden: Es sind solche einzelnen informativen Elemente eines Ausdrucks, die bei dem Übergang von einem Stadium zum nächsten transformationsbedingt verloren gehen. Elleström merkt an, dass Medien durch eine systematische Analyse der Konfiguration medialer Modalitäten besser verstanden und entsprechend genauer definiert werden können.⁵ Dieser Annahme folgend kann durch das Identifizieren, Systematisieren und Analysieren der beschriebenen Relikte ein besseres Verständnis über Medien und deren Abgrenzung und Unterscheidung voneinander entwickelt werden. Der Prozess der Modellierung ist hier auch ein Prozess der medialen Transformation⁶ und gestaltet sich transitiv: Wenn sich ein Verhältnis zwischen den Modellen *und* ein medial transformatives Verhältnis zwischen β_1 und β_2 und zwischen β_2 und β_3 bestimmen lässt, dann kann jenes Verhältnis auch zwischen β_1 und β_3 , usw. bestimmt werden. Durch diese Transitivität der einzelnen Modellierungsschritte zur medialen Transformation sind die Ergebnisse des Experiments – die final erstellten Karten – auch Modelle textlicher Phänomene.

Der beschriebene Modellierungsprozess impliziert zahlreiche und mannigfaltige Auslegungs- und Entscheidungsmomente, da Texte für eine eindeutige Transformation häufig unterspezifiziert sind. Werden jene Momente der Exegese und Interpretation explizit als Parameterwerte verstanden, so lassen sie sich ändern, ebenso wie sich der Transformationsprozess experimentell durch Variationen der Parameterwerte der Modelle ändern lässt. Es ergibt sich als Grundlage ein experimentelles Setup durch systematische Modellierung für experimentelle visuelle Untersuchungen von Textualität. Durch dieses kann außerdem ein tieferes Verständnis von konkreten Texten, wie auch von Textualität im Allgemeinen, gewonnen werden.

⁵ Elleström 2010.

⁶ In Medienwissenschaften *erforscht* man normalerweise mediale Transformationen. Bei der praxisorientierten experimentellen Modellierung in den digitalen Geisteswissenschaften und der analytischen schrittweise vorgenommenen Formalisierung, handelt es sich hingegen eher um das *Vornehmen und Durchführen* einer medialen Transformation.

Durch experimentelle Modellierung, wie sie hier beschrieben wird, können die etwaigen Grenzen der Darstellung, spezifische Eigenschaften oder typische Merkmale in der Kommunikation unterschiedlicher Medien untersucht werden. Für Text und Karte lassen sich zwei grundsätzliche Fragen formulieren:

1. Welche Elemente eines sprachlichen Ausdrucks können durch eine Karte nicht ausgedrückt werden?
2. Was muss spezifiziert sein / werden, wenn man eine Karte anfertigen möchte?

Die erste Frage adressiert sprachliche Aspekte, die nicht kartographiert werden können, wie beispielsweise besondere Formen der Ambiguität und Negation.⁷ Die andere Frage bezieht sich z. B. auf geometrische Präzision: Abstände und Richtungen sind in schriftlicher Form als Text beinahe nie so exakt beschrieben, wie es für geometrische Ausdrücke auf Karten notwendig ist.

In einem derzeit laufenden experimentellen Forschungsvorhaben geht es um Surrealismus und traumähnliche Beschreibungen in der Literatur.⁸ Eine zentrale Frage, die sich dabei stellt, ist die Frage, ob sich geometrische Darstellungen verwenden lassen, um surrealistische Literatur besser zu verstehen. Gibt es, neben Karten, weitere Möglichkeiten für die Verwendung visueller Ausdrucksformen in experimentellen Untersuchungen, um die durch Text ausgedrückten Räumlichkeit besser zu verstehen? Kazuo Ishiguro beschreibt beispielsweise in seinem Roman *The Unconsolable* viele surreale Relationen zwischen dem Protagonisten, anderen Personen und der Landschaft. Das 10. Kapitel handelt von einer Reise beginnend in einem Hotel in der Stadt bis hin zu einem Abendessen auf dem Lande. Nach vielen Vorkommnissen bei dem Abendessen findet der Protagonist heraus, dass es am Ende des Speisesaals eine Türe gibt, die zur Hotelloobby des ursprünglichen Hotels in der Stadt führt. Die Entfernungen sind also nicht konsistent – was zuvor Meilen waren, sind plötzlich lediglich Meter.⁹ Solche inkonsistenten Räume sind im Surrealismus gewöhnlich und können in der ›neuen Welle‹ von Science Fiction-Literatur der 1960er und 70er Jahre sowie in vielen anderen literarischen Formen beobachtet werden. Zweidimensionale Karten mit einer geometrisch ausgedrückten ›irgendwo und nirgendwo-Perspektive‹¹⁰ drücken alle Richtungen und Abstände präzise aus. Die Präzision von Karten ist eine medial spezifische Eigenschaft und nicht etwa eine Folge aus der Tatsache heraus, dass sie sich auf eine extern existierende Wahrheit beziehen und versuchen diese abzubilden. Karten können in diesem Sinne die geschilderten textuell ausgedrückten Umstände, die für menschliche Leser*innen zugänglich sind, nicht ausdrücken.¹¹

⁷ Für eine ausführliche Beschäftigung mit nicht-kartographierbaren Aspekten vgl. Eide 2015.

⁸ Eine Einführung dieser Forschung ist zugänglich in Eide 2016.

⁹ Diese Interpretation wird, gemeinsam mit weiteren Lesarten, in Eide 2016 erläutert.

¹⁰ Die Vogelperspektive im Kontext der Kartographie bezieht sich nicht auf einen singular aufgefundenen Blick eines Vogels, sie zeigt keine Linearperspektive. Eher repräsentiert sie den visuellen Eindruck eines fliegenden Vogels – oder, da Karten durch technische Methoden erstellt werden, Bildserien von Flugzeugen oder Satelliten, die mit Kameras ausgestattet sind.

¹¹ Eide 2016.

Die Grenzen der Ausdrucksmöglichkeiten von Karten müssen jedoch nicht allgemeine Grenzen aller visuellen Ausdrucksformen bedeuten. Durch den Einsatz unterschiedlicher Perspektiven, der Ästhetik aus Film und Computerspielen, sowie das Zurückgreifen auf Erfahrungen aus VR-Systemen lässt sich experimentell erforschen, welche erweiterten Darstellungsmöglichkeiten und Grenzen für surrealistische textuelle Beschreibungen durch mediale Transformationen erreicht werden können. Visuelle Ausdrücke des Surrealismus, wie man sie bei den paradoxen Formen von Magritte bis hin zu perspektivischen Spielen bei Escher beobachten kann, belegen, dass es mit – im technischen Sinne – statischen Bildern viele Möglichkeiten gibt, um surreale Umstände und Bewegungen darzustellen. Wie im Kontext der zuvor geschilderten analytischen schrittweisen Formalisierung bereits angemerkt: Es geht weniger darum, durch Modellierung ein System zur Visualisierung zu erstellen oder lediglich um das visuelle Artefakt, es geht vielmehr um das Aufsetzen experimenteller Kontexte, die zur Untersuchung von ästhetischen Formen in unterschiedlichen Medien und ihrer (fehlenden) Übersetzbarkeit durch Transformation herangezogen werden.

3. Experimentelle Analysen mit Theaterstücktexten durch mediale Transformation in die virtuelle Realität

Der Zugang zu virtuellen Welten in einer virtuellen Realität (VR) wird durch die ständig wachsende Begeisterung für VR-Technologien und den entsprechenden zahlreichen VR-Lösungen in den letzten Jahren immer niedrigschwelliger. So ist ein Smartphone neuerer Generation in Kombination mit einem Cardboard und beispielsweise einem Browser oder einer App ausreichend, um ein VR-Erlebnis für die Nutzer*innen zu ermöglichen. Diese Entwicklung trägt dazu bei, dass sich neben diversen kommerziellen VR-Lösungen unter anderem auch geisteswissenschaftliche VR-Entwürfe konstatieren lassen.¹² Nicht nur der erleichterte Zugang ist beachtlich – ebenso ist das Erschaffen beziehungsweise Erstellen einer virtuellen Welt mit weniger Aufwand verbunden, was auf eine Vielzahl nachnutzbarer 3D-Modelle und für diesen Zweck entwickelte Softwarelösungen zurückzuführen ist.

Die geschilderten Umstände dienen als Voraussetzung für die in diesem Abschnitt vorgestellten experimentellen medialen Transformationen: der praktischen Entwicklung von VR-Anwendungen, die Theaterstücktexte in eigens dafür erstellte audiovisuelle, virtuelle Welten übertragen. Die hier thematisierten resultierenden praktischen Umsetzungen sind in unterschiedlichen Kontexten in den letzten Jahren am Institut für Digital Humanities der Universität zu Köln entstanden. Einerseits wurde die mediale Transformation von Theaterstücken initial im Kontext praktischer Bachelor-Lehrveranstaltungen des Studiengangs Medieninformatik erprobt, andererseits wurde der Ansatz im Rahmen von Forschungsarbeiten von Wissenschaftler*innen des Instituts weiterentwickelt. Der Leitgedanke ist, dass die Transformation eines Theaterstücktextes in eine performative Aufführung zunächst intendiert ist (das Stück ist für die Bühne geschrieben), dies aber nicht zwangsläufig eine Exklusivität hinsichtlich der medialen Transformation zur Folge hat (ein Theaterstücktext, kann nur als

¹² Bekannte Anwendungsbeispiele finden sich im Kontext der Archäologie, der Rekonstruktion und Dokumentation (insbesondere bezugnehmend auf das Kulturerbe). Dabei geht es meist um die Vermittlung von Wissen oder die Visualisierung von (3D-)Datensätzen.

Vorlage für eine Bühnenaufführung genutzt werden). Es ergibt sich die These, dass in diesen Texten eine Transformativität eingeschrieben ist, welche auch eine gute Voraussetzung für eine mediale Transformation in weitere audiovisuelle Medienprodukte schaffen kann. Fernsehspiele und filmische Umsetzungen von Theaterstücken, sowie aufgezeichnete Aufführungen dienen unter anderem in der Medienkulturwissenschaft als Forschungsgegenstand für die geschilderten Überlegungen. Aus wissenschaftlich geleiteter Perspektive scheint es von hohem Mehrwert für die Untersuchungen, nicht nur das resultierende Produkt, sondern vielmehr auch den Entstehungsprozess, den Transformationsprozess und damit auch das Modellieren eines audiovisuellen Stücks in den Blick zu nehmen.

3.1 Der Zugang zu virtuellen Welten: eingesetzte Technologien

Für die Umsetzung medialer Transformationen von Theaterstücktexten in eine virtuelle Welt ist zunächst pragmatisch betrachtet die Auswahl der einzusetzenden Technologien zu klären, da diese die Darstellungsmöglichkeiten selbst definieren und damit als ein ausschlaggebender, aber auch limitierender Faktor betrachtet werden können. Für einen verhältnismäßig unkomplizierten Zugang zu VR mit einem Webbrowser ohne Plugins oder zusätzliche Software, geraten JavaScript-Frameworks in Verbindung mit HTML5, WebGL und WebXR, in den Blick. Wie eingangs erläutert kann mit einem Google Cardboard¹³ und einem Smartphone neuerer Generation außerdem auch die Anschaffung von kostspieligem Equipment vermieden werden, was sich im hier geschilderten Kontext anbietet. Zwar ist dieses Equipment als limitierender Faktor hinsichtlich eines möglichst immersiven VR-Erlebnisses für die Nutzer*innen zu betrachten, kann aber unterstützend für eine anti-illusionistische Inszenierung angesehen werden, wie sie im Experiment beschrieben wird.

So sind hauptsächlich lediglich HTML- und JavaScript-Kenntnisse die Voraussetzung um erste basale VR-Umgebung zu realisieren. Die benötigten 3D-Modelle können selbst modelliert, oder aber auch nachgenutzt werden. Um die Einstiegshürde für den Einsatz in der Lehre möglichst gering zu halten eignet sich das seit 2016 entwickelte und kostenfrei bereitgestellte WebVR-Framework A-Frame¹⁴. Es zeichnet sich durch eine sehr unkomplizierte Handhabung aus und ermöglicht einen schnellen Einstieg. Die umfangreiche 3D-Engine BabylonJS¹⁵ wird wegen weitreichender Einsatzmöglichkeiten und Funktionen in aufbauenden generischen Ansätzen im Kontext von Forschung verwendet.

¹³ Google Cardboard, vgl. Google 2021.

¹⁴ A-Frame, vgl. Mozilla 2020.

¹⁵ BabylonJS, vgl. Cathue 2013.

3.2 Experimentelle Ansätze in der Lehre und Forschung am Institut für Digital Humanities der Universität zu Köln

Im Rahmen der Vorbereitungen für unterschiedliche praktische Übungen im Bachelorstudiengang Medieninformatik, der im Verbund mit dem Fach Medienkulturwissenschaft belegt werden kann, entstand im Kontext des Themenfelds Mediatransformation die Idee zu folgendem experimentellen Ansatz: Bertolt Brechts Theaterstücktexte sollen durch noch zu erforschende Transformationsprozesse in die Virtuelle Realität übertragen werden. Beispielhaft können dafür klassische Aufführungen und filmische Umsetzungen in den Blick genommen werden, die ebenso auf Brechts Stücktexten basieren. Die Ästhetik des epischen Theaters von Brecht, in dem die Wirklichkeit nicht als Illusion abgebildet werden soll und das inszenierte Spiel stets als solches erkennbar bleibt, ergänzt sich mit den (limitierten) technischen Darstellungsmöglichkeiten.¹⁶ Die Illusion von Immersion im (virtuellen) Theater konnte somit während der Experimente eigens thematisiert werden.

Die Studierenden begannen damit, vergleichend die unterschiedlichen Darstellungs- und Ausdrucksmöglichkeiten bei Theateraufführungen und im Film, welche insbesondere in Bezug auf Brechts Werke von Joachim Lang analysiert wurden,¹⁷ sowie in VR auszuarbeiten. Dabei entstanden bereits wertvolle und innovative Ansätze zur Transformation, welche bei der Implementierung Beachtung finden sollten. Beispielhaft dafür ist die folgende Umsetzung, bei der eine Szene in einer Fabrik aufgegriffen wurde: der/die Zuschauer*in respektiv Nutzer*in sieht die virtuelle Umgebung aus der Perspektive eines Arbeiters am Fließband. Im Hintergrund sind Fabrikgeräusche durch Kopfhörer wahrnehmbar. Wird der Blick durch eine Kopfbewegung vom Fließband abgewendet, wird der Fabriklärm unmittelbar unangenehm lauter und passt sich erst wieder einer normalen Lautstärke an, wenn die für den Arbeiter vorgeschriebene Blickrichtung auf das Fließband zeigt. Dieses entwickelte Feature ist spezifisch auf die mediale Darstellung und Interaktion in der VR ausgerichtet. Es soll die als anstrengend und monoton dargestellte Arbeitsweise in der Fabrik unterstreichen und Empathie mit dem Arbeiter erzeugen. Das Beispiel zeigt, dass durch die Umsetzung in einer virtuellen Welt und den zugrundeliegenden Transformationsprozess eine tiefgreifende und erkenntnisversprechende experimentelle Auseinandersetzung mit dem Stücktext, aber auch den verwendeten Medien ermöglicht wird. Auch wenn es bisher wenig Konventionen zu VR-Umsetzungen gibt, ist genau diese Tatsache sehr förderlich für eine kreative und analytische Auseinandersetzung mit dem Gegenstand. Die Auswahl und Erzeugung der 3D-Modelle selbst war der erste wichtige Schritt zur virtuellen praktischen Umsetzung: Es müssen Modelle für das Bühnenbild aber auch für den Aufführungsraum selbst ausgewählt oder erstellt werden. Dabei verschwimmen die Grenzen zum Theater, zu Computerspielen und zum Film: Sollen ein klassischer Zuschauerraum und die Bühne selbst im virtuellen durch 3D-Modelle repräsentiert werden, auch wenn sie in ihrer eigentlichen Funktionalität nicht notwendig erscheinen oder haben sie eine zu starke symbolische Bedeutung, als dass sie im Virtuellen ausgelassen werden können?

¹⁶ Für eine ausführliche Analyse des ausgewählten Themenschwerpunkts vgl. Wieners et al. 2018.

¹⁷ Vgl. Lang 2006.

Aufbauend auf den Ergebnissen der Lehrveranstaltungen wird aktuell im Kontext einer Dissertation ein generischer Ansatz zur semi-automatisierten Transformation von Stücktexten erarbeitet. Nutzer*innen soll es durch Interaktion mit einem dafür entwickelten Editor ermöglicht werden, individuelle virtuelle Aufführungen zu erstellen und / oder als Zuschauer*innen in VR daran teilzunehmen.

4. Von der Kurzgeschichte zum Computergame – Haruki Murakamis Scheunenabbrennen als VR-Anwendung

Ein weiteres Praxisbeispiel medialer Transformation soll im Folgenden besprochen werden. Es geht dabei um die Beschäftigung mit der Umsetzung von textuell verfassten Narrationen als Computer- und Videospiele in VR. So generiert sich die praktische Beschäftigung mit Spielmechanismen in VR in der schulischen und universitären Lehre zum einen als wertvolles Moment zur spielerischen Annäherung an komplexe Lerninhalte wie Verfahren künstlicher Intelligenz,¹⁸ zum anderen als haptisches Werkzeug zur praxisfokussierten kritischen Beschäftigung mit medialer Transformationspraxis: Wie verändert sich eine Narration, wenn sie anhand eines Spieles erzählt wird? Welche Möglichkeiten und Grenzen bietet spielerisches Erzählen – und was ist das überhaupt, ein Spiel?

Wenn Jean Piaget gleich zu Beginn seiner *Theorie der geistigen Entwicklung* ausführt, um Objekte zu erkennen, müsse »das Subjekt auf sie einwirken und infolgedessen transformieren: Es muss sie von der Stelle bewegen, verbinden, in Beziehung zueinander setzen, auseinander nehmen und wieder zusammensetzen«¹⁹, so formuliert Piaget einen zentralen Mechanismus in der experimentellen spielerischen Beschäftigung mit unterschiedlichsten (Forschungs-)Objekten: Forschungsgegenstände wollen angefasst, müssen angedacht und durchdacht werden, Thesen und Antithesen müssen formuliert, Annahmen getroffen, Modelle gebildet, verändert, angepasst, mitunter gar verworfen werden – kurzum: Experimentelles Interagieren gießt das Fundament für theoretischen und praktischen Kenntniserwerb.

Anhand der Lehrveranstaltung *Interaktives Erzählen in VR: Barn Burning (Haruki Murakami)*,²⁰ die im Kontext des Moduls *Visuelle Programmierung* angeboten und durchgeführt worden ist, intendieren die nachfolgenden Ausführungen, einen ergänzenden Einblick zu geben in Aspekte experimenteller universitärer Interpretations- und Visualisierungspraxis. So werden studentische Projektumsetzungen besprochen, die im Rahmen der Lehrveranstaltung entstanden und experimentell die Relation zwischen Narration und Narrationsmedium erkunden – Medien wie die textuell verfasste Kurzgeschichte, filmische Umsetzung und Narration durch Computer- und Videogames.

¹⁸ Vgl. Olari et al. 2021.

¹⁹ Fatke (Hg.) 2016, S. 43f.

²⁰ Vgl. die veranstaltungsbegleitende Website: [Media Transformation – Interaktives Erzählen in VR: "Barn Burning" \(Haruki Murakami, Lee Chang-dong\)](#), vgl. Institut für Digital Humanities, Universität zu Köln 2019.

4.1 Scheunenabbrennen

Haruki Murakamis im Jahr 1983 publizierte Kurzgeschichte Naya o Yaku (Scheunenabbrennen) spielt mit Andeutungen, semantischen Leerstellen und der Imagination ihrer Rezipienten. Auf einer Hochzeitsfeier eines Bekannten lernt der Protagonist eine junge Frau kennen, die ihn durch ihre arglose, nicht der Vernunft gehorchende Einfachheit²¹ fasziniert. Gleich nach ihrem Kennenlernen jedoch offenbart sich jene Simplizität als Fassade komplexer Metaphorik. So lässt Murakami seine Protagonistin in Scheunenabbrennen mehrmals pantomimisch den Prozess des Mandarinenschälens durchführen, der das Realitätsempfinden des Protagonisten durcheinanderwirbelt, gar transzendiert:

»Und dann ›schälte sie Mandarinen‹. ›Mandarinen schälen‹ bedeutet buchstäblich Mandarinen schälen. Links vor ihr stand eine große Glasschüssel mit einem Berg von Mandarinen und rechts eine Schüssel für die Schalen – das war die Anordnung –, in Wirklichkeit war da nichts. Sie nahm eine dieser imaginären Mandarinen in die Hand, schälte sie langsam, steckte ein Stück nach dem anderen in den Mund, nahm die Haut aus dem Mund und warf sie, wenn sie mit einer Mandarine fertig war, zusammen mit der Schale in die rechte Schüssel. Diesen Vorgang wiederholte sie unendliche Male. Wenn man es erzählt, ist es vielleicht nichts Besonderes. Aber als ich es tatsächlich zehn oder zwanzig Minuten lang direkt vor mir sah – wir standen am Tresen einer Bar und plauderten, und sie fuhr fast unbewusst beim Sprechen mit diesem ›Mandarinschälen‹ fort –, war mir, als würde mir jeglicher Realitätssinn entzogen. Ein äußerst seltsames Gefühl. [...] ›Du scheinst wirklich Talent zu haben‹, sagte ich. ›Ach was, das ist doch ganz einfach. Dazu braucht man kein Talent. Man darf nur nicht denken, dass hier Mandarinen sind, sondern man muss vergessen, dass hier keine sind. Das ist alles.«²²

»Ein äußerst seltsames Gefühl.« Auch im weiteren Verlauf der Kurzgeschichte sind Andeutungen, Metaphern und semantische Auslassungen zentral: Die Protagonistin verweist und kehrt nach wenigen Monaten mit einer Person zurück, die der Kurzgeschichte als Antagonist fungiert. Im Rahmen eines geselligen Abends berichtet jener Gegenspieler geheimnisvoll, er stecke alle zwei Monate eine Scheune in Brand und habe sich bereits seine nächste Scheune in unmittelbarer Nähe auserkoren; eine konkrete Aussage darüber, welche Scheune er abzubrennen intendiert, bleibt er jedoch schuldig:

»Ich habe das Gefühl, als gäbe es in der Welt eine Menge Scheunen, die alle darauf warten, von mir abgebrannt zu werden. Die einsame Scheune am Meer oder die Scheune mitten im Reisfeld. [...] In nur einer Viertelstunde brenne ich sie sauber ab. Es ist, als hätten sie nie existiert. Niemand trauert ihnen nach. Sie ... verschwinden einfach. In einem Nu.«²³

Fortan zieht der Protagonist joggend seine Runden und überprüft regelmäßig alle Scheunen in seiner Umgebung. Die Monate und Jahreszeiten verfliegen, die Protagonistin verschwindet, doch eine abgebrannte Scheune lässt sich nicht finden. Am Ende der Kurzgeschichte treffen

²¹ Murakami 2007, S. 143.

²² Murakami 2007, S. 144.

²³ Murakami 2007, S. 157.

sich Protagonist und Antagonist zufällig; im Gespräch berichtet der Gegenspieler, er habe die Scheune bereits wenige Tage nach dem gemeinsamen Abend zerstört: »Ich habe sie natürlich abgebrannt. Bis auf den letzten Rest ist sie verbrannt. So, wie ich es versprochen habe.«²⁴

4.2 Interaktives Erzählen in VR

Der spielerischen Beschäftigung mit der Frage, wie sich eine Narration in und mit einem anderen Medium als Text erzählen lässt, ist Murakamis Kurzgeschichte aufgrund ihrer zahlreichen Andeutungen und Auslassungen äußerst dienlich: Wie lässt sich eine Geschichte erzählen, in der die Spielerinnen und Spieler selbst agieren und mit narrativen Elementen interagieren können? Welchen Nutzen und welche Möglichkeiten bietet eine eigens implementierte VR-Umgebung? Aber auch: Welche Grenzen ziehen virtuelle Realitäten im interaktiven Geschichtenerzählen? Wenngleich einige Aspekte in der theoretischen Beschäftigung mit medialen Transformationen bereits zuvor im Kontext des brechtschen VR-Theaters erschöpfend vorgetragen und besprochen wurden, so stellen die studentischen Projekte, die im Rahmen der Lehrveranstaltung über Murakamis Scheunenabbrennen umgesetzt wurden, weitere wertvolle Fragen: Welche Möglichkeiten werden Spielerinnen und Spielern geboten, um mit (bewusst platzierten) narrativen Lücken umzugehen? Lässt die individuelle Implementation der Kurzgeschichte in VR ihre Rezipienten ob ihrer selektiven Unterspezifiziertheit im Dunkeln, oder nimmt die eigene Umsetzung interpretative Prozesse vorweg, indem beispielsweise die Protagonistin als ermordet dargestellt wird?

Das spielerische Experimentieren mit der Kurzgeschichte extrahiert unter textueller Beschreibung Verborgenes aus der Erzählung und macht sie – je nach Implementation – integrierbar und anders zugänglich als textuell-linear vermittelt. Im Rahmen der Lehrveranstaltung fertigten elf Gruppen mit rund vier Gruppenmitgliedern elf unterschiedliche Projekte mit der Entwicklungsumgebung und Game-Engine Unity an. Zwar sind die elf Projekte in ihrer visuellen Darstellung äußerst heterogen, dennoch ist beachtlich, dass nahezu alle Gruppenprojekte zum einen offenbare, zum anderen verborgene Elemente der Narration fokussieren und herausarbeiten. Elemente wie das pantomimische Mandarinenschalen und der Ort der Scheune als mysteriöser unerreichbarer Ort, der – anders als anhand des Textes – nicht imaginiert, sondern bereits vollständig ausgestaltet präsentiert wird. Trotz ihrer unterschiedlichen und vielfältigen Ansätze lassen sich in den einzelnen Projekten Umsetzungs- und Narrationsmechanismen identifizieren, die verschiedenen Projekten gleichen. So stellt sich zuerst die Frage nach dem Maß an Interaktion; umgesetzt wurden zum einen reine Walking-Simulatoren wie das unten ausschnittsweise dargestellte Museum der Narration, das seine Besucherinnen und Besucher sehr nah am Text, jedoch ohne elaborierte Interaktionsmöglichkeiten stimmig durch die Narration führt (Abbildung 1).

²⁴ Murakami 2007, S. 167.



Abb. 1: Museum der Narration. [Eide et al. 2021]

Im Museum der Narration werden einzelne Elemente der Handlung ausgewählt und kondensiert dargestellt. Interessant an diesem Beispiel ist zum einen der Fokus auf handlungserweiternde Szenen wie die oben dargestellte Flughafenszene oder den Prozess des Mandarinschälens. Zum anderen veranschaulicht sich hier die enge Bindung zwischen Text und interagirbarem Bild. Anhand des in Virtual Reality erfahrbaren Museums der Narration wird Text zwar linear dargestellt, jene Linearität wird jedoch transzendiert durch das Zugleich unterschiedlichster ausgewählter Textfragmente und exemplifizierender visueller Objekte. Zum anderen wurden VR-Anwendungen mit klassischen Spielmechanismen implementiert, die Momente der Narration durch kleine Aufgaben (Quests) erschließen oder anhand eines Inventares narrationsrelevante Objekte wie eine Mandarine eingesammelt und zu späterem Zeitpunkt eingesetzt werden können, um die Handlung fortzuführen. Andere Projekte stellen die Metapher der Scheune in den Vordergrund und arbeiten mit expliziten Darstellungen ebenjener. So finden sich Spielerinnen und Spieler in einem Projekt in einer Scheune wieder und werden mit der Aufgabe betraut, in begrenzter Zeit durch Interaktion mit Objekten, die sich in den jeweiligen Scheunen befinden, die richtige Scheune zu identifizieren. Andere Projekte platzieren ihre Spielerinnen und Spieler in einen dichten dunklen Wald und stellen die Aufgabe, die korrekte Scheune anhand visueller Brotkrumen und Rätsel zu identifizieren (Abbildung 2).



Abb. 2: Brennende Scheune. [Eide et al. 2021]

Spielerische Elemente wie Rätsel, Puzzle, kleine Aufgaben und der Einsatz von Gegenständen, die im Laufe des Spieles eingesammelt, im Inventar gelagert und an einer bestimmten Stelle im Spiel mit anderen Gegenständen kombiniert werden müssen, richten den Fokus der VR-Erzählung auf handlungstragende Momente der Kurzgeschichte – oder sind eingebettet in

freie Interpretationen der Narration, wie das folgende Projekt:²⁵ Basierend auf der Scheunen-Opfer-Metapher erfahren sich die Spielerinnen und Spieler als Mörderinnen und Mörder. Das Spiel beginnt in einem schwach ausgeleuchteten Raum, der repräsentativ für das Innere der verkörperten Figur steht. Drei auf einem Tisch platzierte Ordner gewähren Zugang zu den unterschiedlichen Levelabschnitten. Jedes Level zeigt eine Scheune oder andere Umgebungen. In den unterschiedlichen Spielabschnitten sammeln die Spielerinnen und Spieler Hinweise, um sie gedanklich miteinander zu verknüpfen. Die Hinweise sind dabei so charakteristisch gestaltet, dass sich zu jeder Scheune eine anhand der Scheune repräsentierte Person erahnen lässt. Nach jedem Level kehren die Spielerinnen und Spieler in den Ausgangsraum zurück, der sich je nach Spielfortschritt verändert und so auch die eigene mentale Veränderung im Rahmen der Suche nach einem geeigneten Opfer verdeutlichen soll. Nach Abschluss aller Level werden die Spielerinnen und Spieler anhand der Ordner mit den repräsentierten Charaktereigenschaften konfrontiert und müssen sich – basierend auf den im Rahmen des Spielverlaufes erlebten Erfahrungen – für eine der dargestellten Personen entscheiden.

5. Fazit

Mit ihrem Fokus auf kartenbasierte Experimente in der Textanalyse, brechtschem VR-Theater und der Umsetzung von Computer- und Videogames in VR gaben die vorhergehenden Ausführungen einen Überblick über visuelle modellbasierte Experimente in den digitalen Geisteswissenschaften. Abschließend sei zum einen auf die Methodik der experimentellen Beschäftigung mit Karten, Theaterstücken und textuell verfassten Narrationen eingegangen, zum anderen soll im Folgenden der hermeneutische Prozess in der Beschäftigung mit medialen Transformationen besprochen werden. Als Ausgangspunkt und Grundlage des Experimentierens wurden Werkzeuge – d. h. Software – entwickelt, um mediale Transformationsprozesse zuallererst umzusetzen, anschließend zu beobachten und zu reflektieren. So veranschaulichen die durchgeführten Experimente, welche Aspekte bei der Transformation von einem (z. B. der Kurzgeschichte) in ein anderes Medium (z. B. eine in VR umgesetzte Theateraufführung oder ein in VR umgesetztes Spiel) wesentlich sind, und gestalten sich ein als zentrales Moment in Lehre und Forschung. Charakteristisch an den vorgetragenen Experimenten ist ihr Fokus auf zwei- und dreidimensionale Bildmedien, um Textmedien zu erforschen. Wenngleich der Weg vom Buch zum Film oder dem Computerspiel mitunter intuitiv ist aufgrund zahlreicher Vorlagen und bereits existierender Umsetzungen, so sind auch andere, abstraktere Formen medialer Transformationen vorstellbar, beispielsweise unter Integration von Simulationen.²⁶ Auch könnte man die Erforschung von Bildmedien, auch Karten und VR-Systemen, durch ähnliche Experimente durchführen.

Für die hier beschriebenen Experimente wurden benötigte Werkzeuge im Kontext der Softwareentwicklung eigens gestaltet und umgesetzt. Die dahinterstehenden Prozesse sind experimentelle Bestandteile im Rahmen von Forschung und Lehre, demzufolge ist auch die Softwareentwicklung selbst ein zentraler Bestandteil des hier beschriebenen wissenschaftlichen Arbeitens. Der Prozess der Modellierung wird auf unterschiedliche Art und

²⁵ MediaTransformation, vgl. Heller 2020

²⁶ Vgl. u.a. Fishwick 2017.

Weise als essentiell relevant für die Experimente beurteilt. Zum einen nimmt das Modellieren einen wichtigen Teil innerhalb der Projektplanung in der Softwareentwicklung ein. Dazu gehört das Modellieren der Anwendungen und auch der entsprechend zugrundeliegenden Daten. Zum anderen ist die Modellierung ein wichtiger Teil des Lernprozesses, der mit Experiment und dem Experimentieren einhergeht. Die enge Verbindung zwischen der Modellierung und dem Experiment wird in unterschiedlichen Disziplinen diskutiert²⁷ und sollte auch in den digitalen Geisteswissenschaften weiterverhandelt werden. In den zugehörigen Diskussionen ist der doppelte Blick auf modellbasierte Experimente notwendig, wenn sie als mediale- und semiotische Prozesse aufgefasst werden. Dieser Blickwinkel und der entsprechende Fokus sind nicht vollständig unabhängig voneinander, sie erfordern jedoch unterschiedliche Perspektiven für die Meta-Analyse solcher Prozesse.

²⁷ Vgl. z.B. Peschard / Fraassen (Hg.) 2018; Guest / Martin 2021.

Bibliographische Angaben

David Cathue / David Rousset / Microsoft: babylonjs. 2013. [\[online\]](#)

Arianna Ciula / Øyvind Eide: Modelling in digital humanities: Signs in context. In: Digital Scholarship in the Humanities 32 (2017), Suppl. 1, S. i33–i46. DOI: [10.1093/llc/fqw045](https://doi.org/10.1093/llc/fqw045) [\[Nachweis im GBV\]](#)

Arianna Ciula / Øyvind Eide / Cristina Marras / Patrick Sahle: Modelling: Thinking in Practice. An Introduction. In: Historical Social Research / Supplement 31 (2018), S. 7–29. DOI: [10.12759/hsr.suppl.31.2018.7-29](https://doi.org/10.12759/hsr.suppl.31.2018.7-29) [\[Nachweis im GBV\]](#)

Øyvind Eide: Media Boundaries and Conceptual Modelling: Between Texts and Maps. Basingstoke/Hampshire 2015. [\[Nachweis im GBV\]](#)

Øyvind Eide: Sand in the Mapmaking Machinery: The Role of Media Differences. In: Livingmaps Review 1 (2016), H. 1. [\[online\]](#)

Lars Elleström: The Modalities of Media: A Model for Understanding Intermedial Relations. In: Media borders, multimodality and intermediality. Hg. von Lars Elleström. Basingstoke u. a. 2010, S. 11–48. [\[Nachweis im GBV\]](#)

Lars Elleström: The Modalities of Media II: An expanded Model for Understanding Intermedial Relations. In: Beyond Media Borders, Volume 1: Intermedial Relations among Multimodal Media. Hg. von Lars Elleström. Cham 2021, S. 3–91. DOI: [\[Nachweis im GBV\]](#)

Jean Piaget: Meine Theorie der geistigen Entwicklung. Hg. von Reinhard Fatke. 4. Auflage. Weinheim u. a. 2016. [\[Nachweis im GBV\]](#)

Paul A. Fishwick: Modeling as the practice of representation. In: 2017 Winter Simulation Conference (WSC, Las Vegas, NV, 03. – 06.12.2016) Piscataway, NJ 2017, S. 4276–4287. [\[Nachweis im GBV\]](#)

Google Cardboard. Virtual Reality erleben – einfach, unterhaltsam und kostengünstig. Hg. von Google. 2021 [\[online\]](#)

Olivia Guest / Andrea E. Martin: How Computational Modeling Can Force Theory Building in Psychological Science. In: Perspectives on Psychological Science 16 (2021), H. 4, S. 789–802. DOI: [10.1177/1745691620970585](https://doi.org/10.1177/1745691620970585) [\[Nachweis im GBV\]](#)

Dominik Heller: MediaTransformation. In: GitHub. 2020. [\[online\]](#)

Media Transformation – Interaktives Erzählen in VR: "Barn Burning" (Haruki Murakami, Lee Chang-dong). Hg. vom Institut für Digital Humanities, Universität zu Köln. 2019. [\[online\]](#)

Kazuo Ishiguro: The Unconsoled. London 1995. [\[Nachweis im GBV\]](#)

Tarja Knuuttila / Andrea Loettgers: Modeling/Experimentation: The Synthetic Strategy in the Study of Genetic Circuits. In: The Experimental Side of Modeling. Hg. von Isabelle F. Peschard / Bas C. van Fraassen. Minneapolis, MN u. a. 2018, S. 118–147. [\[online\]](#) [\[Nachweis im GBV\]](#)

Joachim Lang: Episches Theater als Film. Bühnenstücke Bertolt Brechts in den audiovisuellen Medien. Würzburg 2006. [\[Nachweis im GBV\]](#)

Willard McCarty: Humanities computing. Basingstoke/Hampshire u. a. 2005. [\[Nachweis im GBV\]](#)

A-Frame, A web framework for building 3D/AR/VR experiences. Hg. von Mozilla. 16.12.2020 [\[online\]](#)

Haruki Murakami: Der Elefant verschwindet. Erzählungen. Köln 2007. [\[Nachweis im GBV\]](#)

Viktoriya Olari / Kostadin Cvejovski / Øyvind Eide: Introduction to machine learning with robots and playful learning. In: Proceedings from the 35th AAAI Conference on Artificial Intelligence (AAAI: 35, virtuell, 02.–09.02.2021) Palo Alto, CA 2021.

The Experimental Side of Modeling. Hg. von Isabelle F. Peschard / Bas C. van Fraassen. Minneapolis, MN u. a. 2018, S. 118–147. [\[online\]](#) [\[Nachweis im GBV\]](#)

Fictions in science: philosophical essays on modeling and idealization. Hg. von Mauricio Suarez. New York, NY 2009. [\[Nachweis im GBV\]](#)

Jan Wieners / Zoe Schubert / Øyvind Eide: Modellieren durch mediale Transformation: Das Theater Brechts in der virtuellen Realität. In: DHd 2018. Kritik der digitalen Vernunft. Konferenzabstracts. Hg. von Georg Vogeler. (DHd: 5, Köln, 26.02.–02.03.2018). DOI: [10.5281/zenodo.4622366](https://doi.org/10.5281/zenodo.4622366) [\[Nachweis im GBV\]](#)

Frances A. Yates: The art of memory. London 1966. [\[Nachweis im GBV\]](#)

Abbildungs- und Tabellenverzeichnis

Tab. 1: Kritische schrittweise Formalisierung. [Eide et al. 2021 nach Eide 2015]

Abb. 1: Museum der Narration. [Eide et al. 2021]

Abb. 2: brennende Scheune. [Eide et al. 2021]

Beitrag aus:

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. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Datendiffraktion: Von Mixed zu Entangled Methods in den Digital Humanities

Autor*in:

Rabea Kleymann

Kontakt: kleymann@zfl-berlin.org

Institution: Leibniz-Zentrum für Literatur- und Kulturforschung, Berlin

GND: [1179559096](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0003-3856-2685](https://orcid.org/0000-0003-3856-2685)

DOI des Artikels:

[10.17175/sb005_008](https://doi.org/10.17175/sb005_008)

Nachweis im OPAC der Herzog August Bibliothek:

[1783636432](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9)

Erstveröffentlichung:

20.04.2022

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise: 17.02.2022

GND-Verschlagwortung:

[Forschungsmethode](#) | [Wissenschaftstheorie](#) | [Wissensproduktion](#) | [Datenintegration](#) |

Zitierweise:

Rabea Kleymann: Datendiffraktion: Von Mixed zu Entangled Methods 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/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_008](https://doi.org/10.17175/sb005_008).

Rabea Kleymann

Datendiffraktion: Von Mixed zu Entangled Methods in den Digital Humanities

Abstracts

Das aus der Sozialforschung übernommene Forschungsprogramm von Mixed Methods prägt die Wissenskulturen der Digital Humanities (DH). Unter Mixed-Methods-Forschung wird eine Kombination von qualitativen und quantitativen Methoden verstanden, die über die Merkmale von Kompatibilität, Integration und Inferenz bestimmt werden kann. Der Beitrag beschäftigt sich mit dem Import von Mixed Methods in die DH und fragt nach Alternativen. Dabei werden in einem deskriptiven Zugang Bedingungen einer Wissensproduktion mit Mixed Methods exponiert. In einem spekulativen Zugang schlägt der Beitrag im Rückgriff auf Karen Barads Ansatz des Agentiellen Realismus ein Entangled-Methods-Forschungsdesign vor, das über die Konzepte von Intraaktion, Datendiffraktion und Interferenzen exploriert wird.

The research program of mixed methods, adopted from social research, shapes the epistemic cultures of the Digital Humanities (DH). Mixed methods research is understood as a combination of qualitative and quantitative methods, which can be defined by the characteristics of compatibility, integration and inference. The article deals with the import of mixed methods into DH and asks for alternatives. In a descriptive approach, conditions of knowledge production with mixed methods are exposed. In a speculative approach, the paper proposes an entangled methods research design, which is explored through the concepts of intra-action, data diffraction, and interference, by drawing on Karen Barad's approach of agential realism.

1. Einleitung

Die Verfügbarkeit von größeren digitalen Datenbeständen ebenso wie die Entwicklung und Bereitstellung computergestützter Tools und Forschungsinfrastrukturen ermöglichen nicht nur einen Einsatz neuer Methoden, sondern erfordern auch eine umfassende Methodenreflexion und -kritik innerhalb der Digital Humanities (DH).¹ Die gemeinsame Arbeit an einer kritischen Methodenlehre unter einem digitalen Paradigma prägt das Selbstverständnis der DH im besonderen Maße. Ein häufig genannter Anspruch der DH ist eine Erweiterung des klassischen Methodenrepertoires um computergestützte Verfahrensweisen, die auch quantifizierende Vorgehensweisen umfassen können. Teil der DH sind daher qualitative und quantitative Methoden, die häufig über binäre Setzungen wie close und distant reading oder micro- und macroanalysis verhandelt werden.

Im Zusammenhang mit der Methodendiskussion in den DH lässt sich beobachten, dass die aus der empirischen Sozialforschung stammende Mixed-Methods-Forschung Eingang in die Wissenskulturen der DH findet. Das heißt, die Mixed-Methods-Forschung wird zum Bestandteil von denjenigen »Praktiken, Mechanismen und Prinzipien, die [...] in einem

¹ Vgl. Dobsen 2019; Schwandt (Hg.) 2021.

Wissensgebiet bestimmen, wie wir wissen, was wir wissen«.² Unter Mixed Methods wird in der Sozialforschung ein programmatisches Forschungsdesign verstanden, das eine Kombination aus qualitativen und quantitativen Methoden beschreibt.³ Der (neue) Methodendualismus der DH wird nun unter anderem ins Verhältnis zur Situation der Sozialforschung in den 1980er Jahren gesetzt. Dabei geht es insbesondere um die sogenannten Paradigmenkriege.⁴ Mit Blick auf die Unterscheidung zwischen close und distant reading schreibt Herrmann über die aktuelle Lage in den DH: »This situation is somewhat reminiscent of ›paradigm wars‹ in social science, where for years, one could either adhere to a quantitative or a qualitative mindset«.⁵ Eine ähnliche Auffassung vertritt Sá Pereira: »The specter of quantitative methods haunts work in the digital humanities. [...] Studying how the social sciences manage the tension between quantitative and qualitative, however, opens new possibilities for DH practitioners«.⁶ Auch forschungspolitisch ist eine Annäherung der DH an die Problemlagen der Sozial- und Gesellschaftswissenschaften aufgegriffen worden. So heißt es in der Ausschreibung der bereits abgeschlossenen Förderlinie ›Mixed Methods in den Geisteswissenschaften?‹ der Volkswagen Stiftung, dass »die Geisteswissenschaften [...] heute vor derselben Frage [stehen] wie die Gesellschaftswissenschaften bei der Entwicklung von Mixed Methods«.⁷ Dabei ist das Mixed-Methods-Forschungsdesign, wie ich rekurrierend auf Knorr-Cetinas Ansatz behaupte, ein Resultat der »Entscheidungsgeladenheit der Wissensfabrikation«.⁸ Das Mixed-Methods-Forschungsdesign kann, so nun der Ausgangspunkt meines Beitrags, als ein spezifischer »Fabrikationsprozess[...]« von Wissen betrachtet werden, der »eine Kette von Entscheidungen und Verhandlungen [involviert], durch die die entsprechenden Resultate zustandekommen«.⁹ Gegenstand meines Beitrags sind die Bedingungen der Möglichkeit eines solchen Imports von Mixed Methods in die Wissens- und Wissenschaftskulturen der DH sowie eine Exploration eines alternativen Fabrikationsprozesses, der sich einem Methodendualismus widersetzt.

In meinem Beitrag möchte ich zwei Thesen diskutieren: Bislang sind die Bedingungen der Möglichkeit von Mixed Methods in den DH insbesondere unter praxeologischen und pragmatischen Gesichtspunkten herausgearbeitet worden.¹⁰ Die Voraussetzungen, Zielsetzungen und Formen des Rasonierens, die mit einem Mixed-Methods-Forschungsdesign einhergehen, sind hingegen kaum selbst wissenschaftstheoretischer Gegenstand der Forschung gewesen. Ich behaupte daher erstens, dass die drei wesentlichen Merkmale von Mixed Methods als Fabrikationsprozess von Wissen *Kompatibilität*, *Integration* und *Inferenzen* sind. Daher kann ein Import des Mixed-Methods-Forschungsdesigns in die Wissenskultur der DH über diese drei Merkmale verhandelt werden. Zweitens stelle ich die These auf, dass die Denkströmung der Neuen Materialismen, insbesondere die Überlegungen der Quantenphysikerin und Wissenschaftstheoretikerin Karen Barad, Ansätze für einen spekulativen Gegenentwurf zum Mixed-Methods-Ansatz in Aussicht stellen. Barads neo-

² Knorr-Cetina 2002a, S. 11.

³ Vgl. Flick 2011, S. 76f.; Fieldling 2012, S. 124f.

⁴ Baur et al. 2017, S. 7.

⁵ Herrmann 2017.

⁶ Sá Pereira 2019.

⁷ Volkswagen Stiftung 2021, Förderangebot: Interaktion qualitativ-hermeneutischer Verfahren und Digital Humanities: ›Mixed Methods‹ in den Geisteswissenschaften?.

⁸ Knorr-Cetina 2002b, S. 25.

⁹ Knorr-Cetina 2002b, S. 26.

¹⁰ Vgl. Reiter et al. (Hg.) 2020; Herrmann 2017; Sá Pereira 2019.

materialistische Konzeption, die sie als Agentiellen Realismus bezeichnet, verbindet Einsichten der Quantenmechanik mit Elementen poststrukturalistischer Theoriebildung.¹¹ Barad stellt den Agentiellen Realismus als ein »kritisches Werkzeug zur Beschäftigung mit dem Empirischen«¹² vor. Zu Barads zentralen Begrifflichkeiten gehören *Intraaktion*, *Verschränkung* (>entanglements<) und *Diffraktion* bzw. *Interferenz*.

In Rekurs auf Barads Ansatz des Agentiellen Realismus führe ich das Konzept Entangled Methods ein. Ich verstehe darunter ein spekulatives Forschungsdesign, das aktuelle kritische Ansätze der Wissensproduktion aus den Sozialwissenschaften und den Science and Technology Studies für die DH reflektiert. Von besonderer Bedeutung ist neben dem von Uprichard & Dawney eingeführten Konzept »data diffraction«¹³ auch Baths Vorschlag einer »Interferenz als Epistem-onto-logie, um informatische Artefakte im Werden zu begreifen«.¹⁴ Das Entangled-Methods-Forschungsdesign, so argumentiere ich, dient als heuristisches Werkzeug, um latente Strukturen der Wissensproduktion von Mixed Methods sichtbar zu machen. Die leitende Fragestellung lautet: Welche neuen Perspektiven eröffnen sich, wenn wir statt von Mixed von Entangled Methods sprechen? Außerdem präsentiere ich Argumente, die darlegen, inwiefern ein spekulatives Forschungsdesign zur Erweiterung und Neuverhandlung einer Relationalität von Methoden und Daten beitragen kann. Der Beitrag kombiniert daher zwei Modi der Darstellung. Meine Argumentation umfasst einen deskriptiven sowie spekulativen Zugang.

	Mixed Methods	Entangled Methods
Voraussetzung	Kompatibilität als Interaktion	Kompatibilität als Intraaktion
Zielsetzung	Datenintegration	Datendiffraktion
Art des Rasonierens	Inferenz	Interferenz

Tab. 1: Merkmale von Mixed und Entangled-Methods-Forschungsdesigns. [Kleymann 2022]

Mit der Nebeneinanderstellung der Mixed-Methods- und der Entangled-Methods-Forschungsdesigns sollen drei Ziele erreicht werden: Erstens verspricht eine Beschreibung der drei Merkmale *Kompatibilität*, *Integration* und *Inferenz* Aufschluss über die Bedingungen der Wissensproduktion im Rahmen von Mixed Methods zu geben. Diese Merkmale scheinen mit Blick auf eine Übernahme oder Aneignung von Mixed Methods für die DH zentral, indem sie als kritische Parameter das Verhältnis von Methoden und Daten sichtbar machen. Zweitens zielt das spekulative Entangled-Methods-Forschungsdesign auf eine Erweiterung des Verständnisses von Methoden und Daten in den DH. Über die Einführung der Merkmale *Intraaktion*, *Datendiffraktion* und *Interferenz* werden komplementäre Formen für ein Nachdenken über Differenz, materielle Performativität und Relationalität möglich. Drittens exploriere ich auf methodischer Ebene, inwiefern Überlegungen aus den Neuen Materialismen für die DH

¹¹ Vgl. Hoppe / Lemke 2015, S. 262.

¹² Barad 2015, S. 181.

¹³ Uprichard / Dawney 2019, S. 26f.

¹⁴ Bath 2013, S. 97.

produktiv gemacht werden können. Die in meinem Beitrag vorgestellte Kombination aus Deskription und Spekulation stellt ferner eine Einladung dar, Darstellungsformen des Wissens in den DH auszuloten.

Ich schlage die folgende Vorgehensweise vor: Im ersten Teil (Abschnitt 2.) steht eine Beschreibung von Mixed Methods anhand der Merkmale *Kompatibilität*, *Integration* und *Inferenz* in der empirischen Sozialforschung (Abschnitt 2.1) und den DH (Abschnitt 2.2) im Fokus. Im zweiten Teil (Abschnitt 3.) entfalte ich das spekulative Gegenmodell eines Entangled-Methods-Forschungsdesigns. Dazu stelle ich den Agentiellen Realismus vor (Abschnitt 3.1). Davon ausgehend spekuliere ich über die Merkmale *Intraaktion*, *Datendiffraktion* und *Interferenz* über andere Entscheidungsgeladenheiten von Wissen für die DH (Abschnitt 3.2). Es folgt eine kurze Zusammenfassung (Abschnitt 4.).

2. Geregelte Vorgehensweisen: Zur Theorie von Methodenkombinationen

Eine Methode (griech: *mét-hodos*, »der Weg auf ein Ziel hin«¹⁵) bezeichnet zunächst »ein nach Mittel und Zweck planmäßiges [...] Verfahren, das zu technischer Fertigkeit bei der Lösung theoretischer und praktischer Aufgaben führt«.¹⁶ Methoden konstituieren nicht nur ihre Untersuchungsgegenstände, sondern legen auch eine zielgerichtete Abfolge von Arbeitsschritten fest. Klaus & Buhr definieren den Begriff *Methode* daher als »ein System von (methodischen) Regeln oder auch Prinzipien, das Klassen möglicher Operationssysteme bestimmt, die von gewissen Ausgangsbedingungen zu einem bestimmten Ziel führen«.¹⁷ Mit Blick auf die Frage nach einer spezifischen Methodologie der Geistes- und Kulturwissenschaften wird deutlich, dass eine systematische Beschreibung bislang nur in Ansätzen vorliegt.¹⁸ Methodische Fragestellungen unter einem digitalen Paradigma zu adressieren, ist ein wichtiger Teil des Selbstverständnisses in den DH.¹⁹ So spricht Scheinfeldt vom »methodological moment«²⁰ für die DH, während Burdick et al. die Entstehung der DH selbst auf die Begegnungen zwischen traditionellen Geisteswissenschaften und computergestützten Methoden zurückführen.²¹ Die Methodendebatte innerhalb der DH wird einerseits in unterschiedliche zeitliche Phasen eingeteilt, die von einer Übertragung von bestehenden geisteswissenschaftlichen Methoden, über die Entwicklung eigener methodischer Werkzeuge für digital-born-Materialien, hin zur Forderung nach kritischen Methoden für die Regeln algorithmischer Architekturen in Kultur und Gesellschaft reichen.²² Andererseits wird in den DH das Verständnis des Konzepts *Methode* selbst sowohl im Rückgriff auf Begriffe wie *Theorie*, *Tool* bzw. *digitales Werkzeug* und Infrastrukturen, als auch unter praxeologischen Aspekten neu in den Blick genommen. Insbesondere das Methodenprogramm

¹⁵ Nünning / Nünning (Hg.) 2010, S. 7.

¹⁶ Mittelstraß (Hg.) 2013, S. 379.

¹⁷ Klaus / Buhr 1976, S. 792. Zit. nach Nünning / Nünning (Hg.) 2010, S. 8.

¹⁸ Vgl. Winko 2000, S. 584; Nünning / Nünning (Hg.) 2010, S. 7f.

¹⁹ Vgl. Drucker 2021, S. xii.

²⁰ Rieder / Röhle 2017, S. 110; Scheinfeldt 2012. Zit. nach Hall 2012.

²¹ Vgl. Burdick et al. (Hg.) 2012, S. 5.

²² Vgl. Rieder / Röhle 2012, S. 67f.; Berry / Fagerjord 2017, S. 103f.; Rogers 2013, S. 19; Rossiter 2014, S. 223.

und Forschungsdesign von Mixed Methods, das in den Sozialwissenschaften als »dritte methodologische Revolution«²³ bezeichnet wird, gewinnt in diesem Zusammenhang an Bedeutung.

2.1 Unvermischte Verhältnisse: Mixed-Methods-Forschung in den Sozialwissenschaften

In den Sozialwissenschaften hat die Mixed-Methods-Forschung seit den 1980er Jahren einen etablierten Platz. Die Entstehung der Mixed-Methods-Forschung wird von den Vertreter*innen erstens mit einer steigenden Komplexität von Forschungsfragen begründet, die nach Antworten »beyond simple numbers in a quantitative sense or words in a qualitative sense«²⁴ suchen. Mixed Methods wird zweitens, wie Baur et al. betonen, als »eigene Methodologie verstanden, welche den als Forschungshindernis empfundenen Streit von qualitativem und quantitativem Paradigma (»paradigm wars«) überwinden will.«²⁵ Johnson et al. definieren wie folgt:

»Mixed methods research is the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e. g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the purposes of breadth and depth of understanding and corroboration.«²⁶

Anhand der drei Merkmale *Kompatibilität*, *Integration* und *Inferenz* werden nun Voraussetzung, Zielsetzung und Art des Rasonierens von Mixed-Methods-Ansätzen in der empirischen Sozialforschung beschrieben.

Mixed-Methods-Ansätze kennzeichnen sich erstens durch die Voraussetzung der *Kompatibilität*, das heißt der Vereinbarkeit von quantitativen und qualitativen Methoden. Damit wenden sich Mixed-Methods-Ansätze tendenziell gegen eine Inkompatibilitätsthese, die von einer Unvereinbarkeit qualitativer und quantitativer Methoden aufgrund inkommensurabler Paradigmen ausgeht.²⁷ Im Rahmen eines Methodendualismus werden von den Vertreter*innen quantitative Methoden mit dem (Post-)Positivismus und qualitative Methoden mit dem Konstruktivismus assoziiert.²⁸ 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.²⁹ Im Allgemeinen werden aus dieser binären Paradigmenstruktur die folgenden Kennzeichen abgeleitet: Quantitative Methoden werden

²³ Kuckartz 2014, S. 13.

²⁴ Creswall / Plano Clark 2011, S. 21.

²⁵ Baur et al. 2017, S. 14.

²⁶ Johnson et al. 2007, S. 123.

²⁷ Vgl. Bauer et al. 2017, S. 13; Kuckartz 2014, S. 37.

²⁸ Vgl. Lincoln / Guba 2011, S. 168; Howe 1988, S. 13.

²⁹ Vgl. Morgan 2007, S. 51.

mit einer »Verallgemeinerbarkeit der Befunde, Objektivität und Wiederholbarkeit der Datenerhebung und -analyse«³⁰ in Verbindung gebracht. Sie arbeiten mit numerischen Daten und folgen oftmals einer deduktiven Logik.³¹ Qualitative Methoden zeichnen sich hingegen durch nicht standardisierte Erhebungs- und Auswertungsprozeduren aus.³² Beispiele für qualitative Forschungsverfahren in der Sozialforschung sind unter anderem Fallstudien, Interviews, ethnographische Beschreibungen oder teilnehmende Beobachtungen.³³ Ein Ziel qualitativer Forschungsmethoden ist »eine adäquate Erfassung der Sinndeutungs- und Sinnsetzungsprozesse«.³⁴ Für die Erfassung der konstitutiven Strukturen und Prozesse der Sinnkonstruktion werden seltener standardisierte Werkzeuge eingesetzt, sodass weniger strukturierte Daten vorliegen.³⁵ Prämisse der *Kompatibilität* ist also, wie ich mit Blick auf Mixed Methods als Fabrikationsprozess hervorheben möchte, eine binäre und (teils) dichotomische Setzung der beiden Methodensets. Die *Kompatibilität* wird dann als Interaktion zwischen diesen beiden Methoden präsentiert. Legitimiert wird die *Kompatibilität* meist mit dem Hinweis auf den Pragmatismus. Genauer gesagt, die *Kompatibilität* als Interaktion divergenter Methoden ist unter einer pragmatischen Vorgehensweise möglich, da sowohl von einer konkreten Forschungsfrage ausgegangen wird als auch ein Interesse an Lösungsansätzen besteht.³⁶ Teilweise werden Mixed-Methods-Ansätze infolgedessen auch naturalisiert. Während Greene von »multiple ways of seeing and hearing and making sense of the social world«³⁷ ausgeht, sprechen Creswall & Plano Clark vom »natural outlet«³⁸ der Mixed-Methods-Forschung.

Während die *Kompatibilität* als eine Voraussetzung für Mixed Methods beschrieben werden kann, stellt zweitens die *Integration* eine Zielsetzung des Forschungsdesigns dar. Fielding erklärt: »Integration is really the heart of the whole mixed methods exercise because the purpose of mixing methods is to get information from multiple sources and so the issues in bringing together the information are crucial.«³⁹ Bei Fetters & Freshwater ist die Rede von einer »1+1 = 3 Integration challenge«.⁴⁰ Die Autor*innen fügen hinzu: »We describe the integration challenge qualitatively as the imperative to produce a whole through integration that is greater than the sum of the individual qualitative and quantitative parts.«⁴¹ An anderer Stelle wird die *Integration* definiert als »linking of qualitative and quantitative approaches and dimensions together to create a new whole or a more holistic understanding achieved by either alone.«⁴² Mit der *Integration* wird demnach eine Wiederherstellung eines Ganzen beschrieben, das aus den Teilen der jeweiligen Methoden und Forschungsdaten zusammengesetzt ist. Der Begriff der Ganzheit sowie das Gefüge Ganzes / Teil bleiben weitestgehend unterbestimmt.⁴³ Was hier also im Rahmen der »integration challenge« formuliert wird, lässt sich als eine ganzheitliche

³⁰ Kelle 2008, S. 35.

³¹ Vgl. Kuckartz 2014, S. 28.

³² Vgl. Kelle 2008, S. 35; Baur et al. 2017, S. 9.

³³ Vgl. Denzin 2012, S. 82.

³⁴ Kelle 2008, S. 35.

³⁵ Vgl. Baur et al. 2017, S. 10.

³⁶ Vgl. Denzin 2012, S. 82; Feilzer 2010, S. 7.

³⁷ Greene 2008, S. 20.

³⁸ Creswall / Plano Clark 2011, S. 1.

³⁹ Fielding 2012, S. 127.

⁴⁰ Fetters / Freshwater 2015, S. 115.

⁴¹ Fetters / Freshwater 2015, S. 115.

⁴² Fetters / Molina-Azorin 2017, S. 293.

⁴³ Kaulbach et al. 1974, S. 18f.

Setzung von *Integration* lesen. So wird die *Interaktion* von Methoden und Daten stets im Bezug auf eine kohärente Einheit begriffen. Deutlich wird dies zum Beispiel an Formulierungen wie »connecting the parts to the whole«⁴⁴ oder »piecing together qualitative and quantitative data«.⁴⁵ Exemplarisch ist hier auch die Formulierung: »The main purpose for utilizing mixed methods is for completeness [...]«.⁴⁶ Die über den integrierten Methodeneinsatz hergestellte Ganzheit adressiert so eine angenommene Komplexität des Untersuchungsgegenstandes bzw. der Untersuchungsfrage.⁴⁷ Daraus folgt erstens, dass der Untersuchungsgegenstand im Rahmen von Mixed Methods als (komplexes) Ganzes vorgestellt wird. Dieser Komplexität kann zweitens über eine methoden- und datenintegrierende Vorgehensweise begegnet werden.⁴⁸ Mit Blick auf die unterschiedlichen Forschungsdaten schreiben Fetters & Molina-Azorin: »Integration in the data collection dimension indicates the intent of one type of data collection relative to the other type of data collection during mixed methods research procedures«.⁴⁹ Ausgehend vom Integrationsverständnis scheint das Verhältnis der beiden Datensammlungen im Rahmen von Mixed Methods damit auch von einer ganzheitlichen Setzung berührt zu werden.

Als ein drittes Merkmal des Mixed-Methods-Forschungsdesigns können (Meta-)*Inferenzen* genannt werden, welche die Art des Rasonierens auf Basis der *Integration* bestimmen. *Inferenzen* beschreiben Formen des Schließens bzw. der Schlussfolgerung auf Basis von vorhandenen Wissensrepräsentationen. Dazu zählen deduktive, induktive und abduktive Schlussfolgerungen.⁵⁰ Eine Metainferenz im Rahmen von Mixed Methods meint dann »an overall conclusion, explanation, or understanding developed through an integration of the inference obtained from the qualitative and quantitative strands of a mixed methods study«.⁵¹ Dabei wird die Metainferenz mit Blick auf die Dreiteilung als Abduktion begriffen, die dem formulierten pragmatischen Anspruch des Forschungsdesigns folgt. Morgan schreibt: »The pragmatic approach is to rely on a version of abductive reasoning that moves back and forth between induction and deduction – first converting observations into theories and then assessing those theories through action«.⁵² Auf diese Weise entwickeln Metainferenzen ein ganzheitliches Verständnis vom Untersuchungsgegenstand: »They create an understanding (e. g. ›Gestalt‹ or ›whole‹) [...]«.⁵³ Die Qualität dieser Metainferenz kann über Kriterien von Plausibilität, Validität oder Transfer gemessen werden.

Zusammenfassend kann festgehalten werden, dass über die Merkmale *Kompatibilität*, *Integration* und *Inferenz* wichtige Bausteine des Mixed-Methods-Forschungsdesigns, wie es in der empirischen Sozialforschung diskutiert wird, in den Blick genommen werden

⁴⁴ Denzin 2012, S. 85.

⁴⁵ Bazeley 1999, S. 279.

⁴⁶ Tashakkori / Teddlie 2008, S. 102.

⁴⁷ Vgl. Fielding 2012, S. 128; Teddlie / Tashakkori 2011, S. 287.

⁴⁸ Vgl. Uprichard / Dawney 2019, S. 20.

⁴⁹ Fetters / Molina-Azorin 2017, S. 299.

⁵⁰ Vgl. Beierle / Kern-Isberner 2014, S. 23.

⁵¹ Tashakkori / Teddlie 2008, S. 102.

⁵² Morgan 2007, S. 71.

⁵³ Tashakkori / Teddlie 2008, S. 104.

können. Die Beschreibung von Voraussetzung, Zielsetzung und Art des Rasonierens kann als Versuch einer Rekonstruktion der gemachten Selektionen verstanden werden, die der Entscheidungsgeladenheit der Fabrikation von Wissen nachkommt.⁵⁴

2.2 Maßvolle Vermischungen: Kombinierte Methodeneinsätze in der digitalen Literaturwissenschaft

Die Rede von Mixed Methods hat nicht nur Eingang in das Vokabular von digital arbeitenden Geisteswissenschaftler*innen gefunden. Vielmehr steigt auch die Anzahl von Projekten und Untersuchungen, die sich Mixed-Methods-Forschungsdesigns verpflichten. Vor diesem Hintergrund werden auch die disziplinären Grenzen zwischen den Geistes- und Sozialwissenschaften auf den Prüfstand gestellt. So plädiert Goldstone: »Let the DH be sociological!«⁵⁵ Sá Pereira hingegen erklärt: »Again, the digital humanities can become more sociological without necessarily becoming exclusively quantitative.«⁵⁶ Liu deutet ferner an, dass die DH mit den Sozialwissenschaften ein interpretatives Paradigma teilen.⁵⁷ Auch bei Dobsen rückt die Frage »What distinguishes the interpretative procedures of the DH scholar from those of the social scientist?«⁵⁸ in den Fokus. Anhand der drei Merkmale *Kompatibilität*, *Integration* und *Inferenz* möchte ich nun kursorisch nachverfolgen, wie Mixed-Methods-Ansätze in den DH, insbesondere in den digitalen Literaturwissenschaften, verhandelt werden. An welche diskursiven Interaktionen der DH-Wissenskultur sowie methodologischen Überlegungen schließen Mixed Methods als Fabrikationsprozess in den DH an?

Eine Voraussetzung für den Mixed-Methods-Ansatz in der empirischen Sozialforschung ist die Kompatibilität, die auf einer binären und dualistischen Vorstellung der Methoden beruht. Um der Kompatibilität nachzugehen, rückt zunächst die Frage nach der Binarität bzw. Dualität von quantitativen und qualitativen Methoden in den DH in den Vordergrund. Zu »quantitative[n] Ansätze[n] im weitesten Sinn« können laut Bernhart »zählende, messende, mathematische, statistische, geometrische, empirische, computergestützte und informatische Verfahren [...]«⁵⁹ gezählt werden. Diese können mit »qualitativen und hermeneutischen Analyse- und Interpretationsverfahren«⁶⁰ verbunden werden. Im Kontext seiner Definition macht Bernhart deutlich, dass das Quantifizierende nicht ein Alleinstellungsmerkmal der DH sei. Vielmehr weise gerade die Literaturwissenschaft eine lange quantifizierende Tradition auf.⁶¹ Zudem können computergestützte Verfahren nicht mit quantifizierenden Methoden gleichgesetzt werden. Morettis *distant reading*, Jockers *macroanalysis* und Pipers *cultural analytics* werden z. T. stellvertretend für den Methodeneinsatz in den DH aufgeführt. Jannidis et al. weisen in ihrer Definition daher auf die unterschiedlichen Forschungszwecke hin:

⁵⁴ Vgl. Knorr-Cetina 2002b, S. 26.

⁵⁵ Goldstone 2014, S. 172.

⁵⁶ Sá Pereira 2019.

⁵⁷ Vgl. Liu 2013, S. 414.

⁵⁸ Dobsen 2019, S. 10.

⁵⁹ Bernhart 2018, S. 208.

⁶⁰ Bernhart 2018, S. 208.

⁶¹ Vgl. Bernhart 2018, S. 208.

»Quantitative Analysemethoden grenzen sich von qualitativen Analysemethoden ab, die Bestandteile und Eigenschaften von Forschungsgegenständen beschreiben und dabei die besondere Aufmerksamkeit auf nuancierte Differenzierungen, individualisierende Detailanalysen und herausragende [...] Einzelbeispiele legen. Quantitative Analysemethoden hingegen sind in erster Linie darauf ausgerichtet, Merkmale von Forschungsgegenständen zu identifizieren und ihre Häufigkeiten zu erheben, was möglichst klare und teils auch vereinfachende Kategorisierungen erfordert.«⁶²

Zu konkreten quantitativen Methoden zählen Jannidis et al. beispielsweise statistische Analysen, maschinelles Lernen, Netzwerkanalysen und Topic Modeling.⁶³ Gleichzeitig geht die Unterscheidung zwischen quantitativen und qualitativen Methoden mit einer Reihe von Dichotomien einher. In Herrmanns Mixed-Methods-Ansatz werden einige der zu überwindenden Dichotomien aufgeführt: »close vs. distant«, »qualitative vs. quantitative«, »explanatory vs. exploratory«, »inductive vs. deductive«, »understanding vs. explaining« und »hermeneutic vs. empirical«.⁶⁴ Krautter & Willand wiederum vertreten die Position, dass im Rahmen einer Interpretation immer schon quantitative und qualitative Vorgehensweisen und damit Daten miteinander vermischt werden.⁶⁵ Die Bedeutung und Implikationen dieser (vermeintlichen) Gegensatzpaare für die DH zu reflektieren, kann als eine wesentliche Aufgabe einer Erkenntnistheorie der DH beschrieben werden. Im Vordergrund für meine Argumentation steht aber weniger die Tragfähigkeit dieser Dichotomien für die DH. Mit Blick auf den Import von Mixed Methods als Fabrikationsprozess von Wissen sind hier strukturell die binären Setzungen hervorzuheben, die bedingen, dass auch in den DH über Kompatibilität als Interaktion divergenter Methoden gesprochen werden kann.

Ausgehend von diesen binären Setzungen werden in den digitalen Literaturwissenschaften unterschiedliche Ebenen von Kompatibilität und Inkompatibilität verhandelt.⁶⁶ Im Anschluss an Nan Z. Das Kritik und die damit einhergehende Kontroverse wurde jüngst die Kompatibilitätsannahme für die digitale Literaturwissenschaft stark diskutiert. Da spricht sich gegen die Vereinbarkeit aus, wobei sie unter quantitativen Methoden den Einsatz von statistischen Tools für die Literaturanalyse versteht: »It seems unobjectionable that quantitative methods and nonquantitative methods might work in tandem«.⁶⁷ Neben der strikten Ablehnung auf der einen Seite lässt sich in den digitalen Literaturwissenschaften auf der anderen Seite ein Prozess der Ausdifferenzierung beobachten. Schruhl, Gius und weitere DH-Vertreter*innen plädieren dafür, den Fokus zunächst auf das zu legen, was überhaupt miteinander in Interaktion treten oder inkommensurabel bleiben soll.

Ich möchte daran anknüpfend exemplarisch zwei Diskurszusammenhänge beschreiben, welche die Voraussetzung der Kompatibilität in den digitalen Literaturwissenschaften verhandeln: Zum einen wird die Kompatibilitätsannahme in den DH unter

⁶² Jannidis et al. (Hg.) 2017, S. 279.

⁶³ Vgl. Jannidis et al. (Hg.) 2017, S. 279.

⁶⁴ Herrmann 2017.

⁶⁵ Vgl. Richter 2020, S. 256; Krautter / Willand 2020, S. 95.

⁶⁶ Vgl. Da 2019a.

⁶⁷ Da 2019b.

einer praxeologischen Perspektive verhandelt. Es geht also um die Frage, wie Forschungsumgebungen und Arbeitstechniken, die mit quantitativen und qualitativen Methoden einhergehen, miteinander kombiniert werden können.⁶⁸ Mixed Methods oder Formen der Triangulation sind Teil der weiteren Diskussionen um den laboratory turn in den DH.⁶⁹ Kuhn entwirft ein Modell zur Integration von Arbeitspraktiken zwischen Computerlinguistik und hermeneutisch-arbeitenden Geistes- und Sozialwissenschaften. Bezugnehmend auf die Frage nach der Kompatibilität erklärt Kuhn, dass ein Großteil der Forschung in den Geisteswissenschaften nicht danach strebe, ein Paradigma mit einem disziplinären Konsens zu bilden.⁷⁰ Es gehe eher um die Gewichtung von unterschiedlichen Zugängen. In der Beschreibung des unter anderem von Kuhn geleiteten Projektes ›textklang: Mixed-Methods-Analyse von Lyrik in Text und Ton‹ heißt es dann:

»Das methodologische Zusammenwirken all dieser Komponenten schließt an die Entwicklung eines reflektierten Mixed-Methods-Workflows in den Digitalen Geisteswissenschaften / Digital Humanities an und stellt einen wichtigen Schritt für eine geisteswissenschaftlich informierte Betrachtung multimodaler Wechselbeziehungen dar.«⁷¹

In den DH rückt vor allem die Einheit des Analyse- und Interpretationsprozesses in den Vordergrund. Jüngst hat Drucker darauf hingewiesen, dass die Methoden der Geisteswissenschaften auf eine Auseinandersetzung mit Ambiguität, Komplexität und Widersprüchlichkeit zielen: »Humanists do not approach their research as problems to be solved, but as investigations of the cultural record.«⁷² Gleichzeitig begegnen den digital arbeitenden Geisteswissenschaftler*innen aber vordergründig Tools, die entwickelt werden, um effiziente Lösungsansätze vorzuschlagen.

Zum anderen wird Kompatibilität in den digitalen Literaturwissenschaften über Formen der Skalierung reflektiert. Das von Weitin als Mixed Methods vorgestellte scalable reading verknüpft close- und distant-reading-Ansätze:

»Scalable Reading bedeutet indes nicht (nur), dass sich close und distant reading methodisch durchdringen (diese Analogie verführt womöglich dazu, die Unterschiede zwischen Text- und Daten-Interpretation zu unterschätzen), sie steht für ein integriertes Verständnis aller Akte des Lesens und der Analyse.«⁷³

Zudem richtet sich Weitin gegen eine »allzu versöhnliche Vorstellung der Verbindung traditioneller und digitaler Methoden«. ⁷⁴ Darüber hinaus macht Weitin klar, dass die Kompatibilität konkret an den Erfolg einer Operationalisierung geknüpft ist.⁷⁵ Für ihn setze

⁶⁸ Vgl. Reiter et al. (Hg.) 2020.

⁶⁹ Vgl. Pawlicka-Deger 2020.

⁷⁰ Vgl. Kuhn 2019, S. 568.

⁷¹ Universität Stuttgart 2021: Textklang.

⁷² Drucker 2021, S. X.

⁷³ Weitin 2015b, S. 9.

⁷⁴ Weitin 2015a, S. 654.

⁷⁵ Vgl. Weitin 2015b, S. 6.

ein Mixed-Methods-Ansatz, wie das scalable reading, vor allem eine Veränderung des Denkstils innerhalb der Literaturwissenschaften voraus, denen er eine »breit habitualisierte[...] Geringschätzung für die Konsolidierung bereits bestehenden Wissens«⁷⁶ zuschreibt. Der Einsatz von Mixed Methods in den DH geht daher häufig mit einer empirischen Validierung von Forschungsergebnissen einher. Daran knüpfen sich zugleich Fragen nach der Replikation von Forschungsergebnissen in den digitalen Literaturwissenschaften an.⁷⁷ Krautter & Willand kommen bei der Ausdifferenzierung der Begriffe von Quantifizierung und Skalierung zudem auf die Vereinbarkeit zwischen nahen und distanten epistemischen Grundhaltungen zu sprechen.⁷⁸ Vor diesem Hintergrund fragen sie nach Maß und Gleichgewicht von unterschiedlichen methodischen Vorgehensweisen. Auch Kuhn spricht von einer passenden Gewichtung, die bedeutsam für die Vereinbarkeit ist. Maß, Balance und Gleichgewicht scheinen innerhalb der DH-Wissenskultur mögliche geteilte Kriterien für die Kompatibilität zu sein. Mit anderen Worten, quantitativen und qualitative Methoden sind vereinbar, wenn sie die Balance halten.

Ein zweites Merkmal von Mixed Methods ist das Ziel der Integration. Für die digitale Literaturwissenschaft hat Gius vorgeschlagen, Integration als eine Komplexitätsdimension von Erkenntnissen in computergestützten Textanalysen in den Blick zu nehmen.⁷⁹ Statt der korrelativen Begriffe von Ganzes und Teil spricht Gius von Einfachheit und Komplexität. So rückt, wie ich hervorheben möchte, eine strukturelle Beziehung in den Vordergrund, die den Untersuchungsgegenstand bereits relational fasst.⁸⁰ Dabei betont Gius nicht nur die Zusammengesetztheit von Phänomenen. Vielmehr bemisst sie den Grad der Komplexität des Phänomens an den zu operationalisierenden Einheiten. Gius geht erstens davon aus, dass unterschiedliche Phänomene durch unterschiedliche Methoden adressiert werden können. Die einzelnen Methoden konstituieren aber zweitens wiederum unterschiedliche (Teil-)Phänomene, die erst im Rahmen einer Interpretation integriert werden. Auch Piper macht deutlich, dass es kein »stable, knowable whole«⁸¹ gebe. Herrmann betont eine Aufrechterhaltung von Differenzen, die gleichzeitig überschritten werden.⁸² Die einzelnen Methoden und ihre Forschungsergebnisse können zwar unter einem pragmatischen Anspruch miteinander kombiniert werden, bilden aber nicht holistisch ein neues Ganzes.⁸³ Hervorzuheben ist an dieser Stelle ein weiterer Unterschied der diskursiven Formationen. Die Integration wird als eine Zielsetzung von Mixed Methods in den DH zwar weiterhin präsentiert, aber über die Formen der Operationalisierung verhandelt. In den digitalen Literaturwissenschaften tritt die Operationalisierung als Integration unter anderem im Bild der Brücke bzw. Überbrückung (vgl. »bridging the gap«) in Erscheinung. Bei Moretti heißt es: »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.«⁸⁴ So wird

⁷⁶ Weitin 2015a, S. 654.

⁷⁷ Vgl. Schöch et al. 2020.

⁷⁸ Vgl. Krautter / Willand 2020, S. 94.

⁷⁹ Vgl. Gius 2019, S. 10.

⁸⁰ Vgl. Koschorke 2017, S. 2; Luhmann 1976, S. 940f.

⁸¹ Piper 2018, S. 8.

⁸² Vgl. Herrmann 2017.

⁸³ Vgl. Herrmann 2017. Hier liegt auch ein wesentlicher Kritikpunkt an Horstmann / Kleymann 2019. Ich danke Berenike Herrmann für diesen Hinweis.

⁸⁴ Moretti 2013, S. 1.

eine Vorstellung der Integration skizziert, die zwei zuvor getrennte Bereiche miteinander verbindet. Diese Art einer Verbindung über eine Differenz oder Lücke hinweg wird über eine vergleichende Betrachtung der beiden Datensammlungen vollzogen. Integration geht dann mit einem Aufsuchen von Ähnlichkeiten und Differenzen einher, die über eine reflexive Betrachtungsweise gewonnen werden.

Eine Funktion der Integration im Rahmen von Mixed-Methods-Ansätzen besteht drittens darin, dass auf Basis der verknüpften Datensammlungen die Bildung von formalen (Meta-)Inferenzen möglich ist. In der digitalen Literaturwissenschaft stellt jedoch die Systematisierung von Prozessen der Schlussfolgerung ein Forschungsdesiderat dar. Jüngst hat sich Piper mit den Bedingungen der Möglichkeit von Generalisierungen beschäftigt. Er skizziert eine evidenzbasierte Form von Generalisierungen.⁸⁵ Gius hingegen geht konkret auf »Deduktion, Induktion und Abduktion als Skala für die Beschreibung des Erkenntnisbeitrags«⁸⁶ von Textanalysen und -interpretationen ein. Des Weiteren scheint die Art des Rasonierens bei Mixed-Methods-Ansätzen in den DH oftmals nur assoziativ angedeutet zu werden.⁸⁷ Ein kohärenter Zusammenhang wird eher rhetorisch erzeugt, was unter anderem zu einer problematischen Auswahl von Beispielen führen kann (vgl. »cherry picking of evidence«⁸⁸).

Über die Merkmale von Kompatibilität, Integration und Inferenzen die Diskursformation von Mixed Methods in den DH zu beschreiben, stellt eine produktive Auseinandersetzung der disziplinären Grenzen zwischen Literaturwissenschaften und Sozialwissenschaften dar. Des Weiteren können neben Gemeinsamkeiten vor allem Verschiebungen und Friktionen innerhalb der diskursiven Formation beobachtet werden, die Bedingungen von Wissensproduktionen in den DH reflektieren. Resümierend kann festgehalten werden: Im Zuge der Unterscheidung von quantitativen und qualitativen Methoden in den DH wird deutlich, dass die binäre Setzung der Methoden diskursiver Teil eines Ausdifferenzierungsprozesses ist. Mit Blick auf die zitierten Forschungsansätze fällt ferner auf, dass die korrelativen Begriffe von Ganzes / Teil in den Hintergrund treten. Stattdessen gewinnt Komplexität als Strukturbegriff an Bedeutung. So wird nicht nur eine andere Konzeption des Untersuchungsgegenstandes deutlich. Vielmehr stellt sich auch die Frage, welche Aggregationsweise der Forschungsdaten eigentlich mit dem Komplexitätsbegriff verknüpft ist. Dennoch scheinen Formen des Ganzheitsdenkens immer noch an ein hermeneutisches Ideal geknüpft zu sein.

3. Forschungsdesign im Futurum II: Begegnungen mit Neuen Materialismen

Während in einem ersten deskriptiven Zugang das Mixed-Methods-Forschungsdesign in den Sozialwissenschaften und in den DH skizziert wurde, schließt sich nun eine spekulative Exploration an. Im Wechsel des Modus von der Beschreibung hin zur Spekulation möchte ich einen alternativen Entwurf zum Mixed-Methods-Forschungsdesign für die DH vorschlagen.

⁸⁵ Vgl. Piper 2020, S. 18.

⁸⁶ Gius 2019, S. 12.

⁸⁷ Vgl. Sá Pereira 2019.

⁸⁸ Weitin 2015a, S. 655.

Einen Ausgangspunkt bildet die Frage, inwiefern das Verhältnis von Methoden und Daten auch anders erzählt oder imaginiert werden könnte. Ich behaupte, dass die Denkströmungen innerhalb der Neuen Materialismen, insbesondere der Ansatz des Agentiellen Realismus von Karen Barad, Ansätze für ein spekulatives Forschungsdesign in Aussicht stellen.⁸⁹ Meinen Entwurf, den ich als Entangled-Methods-Forschungsdesign bezeichne, verstehe ich im Sinne Gramlichs als Versuch einer »spekulativen Öffnung und Pluralisierung«⁹⁰ von Fabrikationsprozessen von Wissen in den DH. Das Forschungsdesign adressiert also ein ›Denken im Futurum II‹.⁹¹

3.1 Karen Barads Agentieller Realismus

Unter den Neuen Materialismen werden Ansätze summiert, die das dynamische Zusammenspiel von diskursiven und materiellen Bedeutungsprozessen und Konfigurationen untersuchen.⁹² Ein Fokus liegt auf Vorschlägen für eine Neuverhandlung von Materialität und Materie. Ein gemeinsamer Nenner der heterogenen Bewegung sei die Überzeugung, so Hoppe & Lemke, »dass der ›linguistic turn‹ oder primär semiotisch verfahrenende Ansätze unzureichend sind, um das komplexe und dynamische Zusammenspiel sinnhaft-symbolischer Prozesse und materieller Ordnung zu erfassen«.⁹³ Zentrale Kennzeichen der Neuen Materialismen seien, erstens eine transversale Qualität, die ermögliche jenseits von Dualismen ein (post)modernes Denken zu entwickeln. Daran knüpfte sich zweitens eine »monistische Philosophie der Differenz« an, die auf dem Modell einer »affirmativen Relationalität«⁹⁴ basiere. Drittens rücken die Neuen Materialismen performative Ontologien in den Fokus, die viertens häufig auf posthumanistischen Überlegungen beruhen.

Eine wichtige Vertreter*in der Neuen Materialismen ist Karen Barad, die in ihrem Ansatz des Agentiellen Realismus Bohrs Überlegungen zur Quantenphysik mit poststrukturalistischen und feministischen Ansätzen von Foucault, Butler und Levinas verbindet.⁹⁵ Mit dem Agentiellen Realismus bezeichnet Barad ein »epistemological-ontological-ethical framework that provides an understanding of the role of human and nonhuman, material and discursive, and natural and cultural factors in scientific and other social-material practices«.⁹⁶ Barads Anspruch an einen solchen Ansatz ist eine Verschiebung von Formen der Repräsentation hin zu einem »performativen Verständnis technisch-naturwissenschaftlicher und natürlichkultureller [sic] Praktiken, einschließlich verschiedener Arten von Praktiken der Wissensproduktion«.⁹⁷ Ausgangspunkt des Agentiellen Realismus ist Barads Auseinandersetzung mit der Arbeit von Niels Bohr. Seine Arbeit stelle einen kartesischen Glauben an die Trennung von Subjekt und Objekt, Erkennendem und Erkanntem radikal in Frage.⁹⁸ Grundlegend für die

⁸⁹ Vgl. Bode 2020.

⁹⁰ Gramlich 2020, S. 10.

⁹¹ Gramlich 2020, S. 9.

⁹² Hoppe / Lemke 2021, S. 10f.

⁹³ Hoppe / Lemke 2021, S. 10.

⁹⁴ Lemke 2020, S. 556.

⁹⁵ Vgl. Hoppe / Lemke 2021, S. 59.

⁹⁶ Barad 2007, S. 26.

⁹⁷ Barad 2013, S. 57.

⁹⁸ Vgl. Barad 2018, S. 17.

Auflösung der Trennung sei, wie Hoppe & Lemke betonen, die »quantenphysikalische Einsicht, dass Position und Impuls eines Teilchens nicht gleichzeitig bestimmbar sind.«⁹⁹ Barad kommt zum Schluss, dass die »primäre ontologische Einheit [...] nicht aus unabhängigen Gegenständen mit vorgegebenen Grenzen und Eigenschaften [besteht], sondern vielmehr aus Phänomenen«.¹⁰⁰ Ein Merkmal von Phänomenen sei folglich, dass sie die »erkenntnistheoretische Unzertrennlichkeit von Beobachter und Beobachtetem oder die Ergebnisse von Messungen«¹⁰¹ markieren. Daher seien »Phänomene ontologisch primitive Relationen [...] – Relationen ohne zuvor existierende Relata«.¹⁰² In Barads Ansatz bilden Phänomene die kleinsten ontologischen Einheiten, die aber eben nicht aus unabhängigen in sich abgeschlossenen Entitäten zusammengesetzt werden. So verbindet Barads Ansatz auf einer wissenschaftstheoretischen Ebene Realismus und Konstruktivismus miteinander. Mehr noch, Phänomene entziehen sich kausalen, regelbasierten (z. B. Wenn-Dann-Beziehung) oder zeitlichen Strukturen (z. B. Vorher-Nachher-Beziehungen). Sie können nur prozessual als fortlaufende (Ent-)Faltungen von Bedeutung und Materie verstanden werden.¹⁰³

Die von Barad vorgeschlagene relationale Ontologie kann über die drei Konzepte *Intraaktion*, *Verschränkung* und *Diffraktion* näher bestimmt werden. Über den Neologismus Intraaktion, der im Kontrast zur Interaktion steht, versucht Barad die Vorrangstellung der Relation vor dem Relatum begrifflich zu fassen. Die Intraaktion beschreibt, so Barad,

»die wechselseitige Konstitution verschränkter Wirkmächtigkeiten. Das heißt, im Gegensatz zur üblichen Interaktion, die separate Wirkmächtigkeiten als ihrer Interaktion vorgängig voraussetzt, erkennt der Begriff der Intraaktion an, dass distinkte Wirkmächtigkeiten ihrer Intraaktion nicht vorgängig sind, sondern durch ihre Intraaktion hervortreten«.¹⁰⁴

Damit einher gehe eine ontologische und epistemologische Untrennbarkeit, eben eine *Verschränkung* in Barads Worten, zwischen Subjekten und Objekten bzw. Wissen und Sein. Phänomene seien durch »die ontologische Unzertrennlichkeit/Verschränkung intraagierender Agentien«¹⁰⁵ geprägt und konstituieren sich fortwährend durch agentielle Schnitte. Barad erklärt: »Intraaktionen umfassen die weitere materielle Anordnung (d. h. die Menge materieller Praktiken), die einen agentiellen Schnitt zwischen Subjekt und Objekt vollzieht [...]«¹⁰⁶ Mit dem agentiellen Schnitt beschreibt Barad folglich die Inkraftsetzung einer Grenze, die dann einzelne Dinge als Phänomene hervorbringt. So werden Subjekte und Objekte ebenso wie Untersuchungsgegenstände in Abhängigkeit von spezifischen materiellen Anordnungen temporär hervorgebracht. Solche materiellen Anordnungen stellen für Barad

⁹⁹ Hoppe / Lemke 2021, S. 61.

¹⁰⁰ Barad 2018, S. 19.

¹⁰¹ Barad 2018, S. 19.

¹⁰² Barad 2018, S. 19.

¹⁰³ Vgl. Hahn 2019.

¹⁰⁴ Vgl. Barad 2013, S. 65.

¹⁰⁵ Barad 2018, S. 20.

¹⁰⁶ Barad 2018, S. 20.

beispielsweise wissenschaftliche Apparate dar. Barad versteht Apparate nicht als bloße Beobachtungs- und Messgeräte, sondern als »grenzziehende Praktiken – spezifische materielle (Re-)Konfigurationen der Welt –, die sich materialisieren und Relevanz erlangen.«¹⁰⁷

Um intraaktiven Verschränkungen zu beschreiben, entwickelt Barad des Weiteren eine Methodologie, die vom optischen Phänomen der *Diffraction* ausgeht. Der aus der Physik stammende Begriff beschreibt Interferenzen, die bei der Überlagerung oder Beugung von zwei Wellen entstehen. Interferenzen können zum Beispiel beobachtet werden, wenn »gleichzeitig zwei Steine in einen See geworfen werden« und die »kreisförmigen Wellen« sich ausbreiten und überlagern.¹⁰⁸ Für Barad stellt die Diffraction anknüpfend an Haraways Überlegungen einen Gegenbegriff zur Reflexion dar. Während eine Reflexion in der geometrischen Optik das Zurückwerfen von Lichtstrahlen an einer Grenzfläche meint, beschreibt die Diffraction die Beugung von Lichtwellen. Davon ausgehend behauptet Barad, dass die Reflexion mit den Konzepten von Ähnlichkeit bzw. Identität und Repräsentation verbunden sei. Die Diffraction hingegen bringe Differenz und Performativität zum Ausdruck. Eine kritische Praxis auf Basis der Diffraction bestimmt Barad daher wie folgt: »A diffractive methodology is respectful of the entanglement of ideas and other materials in ways that reflexive methodologies are not.«¹⁰⁹ Charakteristisch für eine diffraktive Methodologie, die als ein »Durch-einander-hindurch-Denken von Einsichten«¹¹⁰ vorgestellt wird, sei erstens, »a respectful engagement that attends to detailed patterns of thinking of each; fine grained details matter«.¹¹¹ Zweitens wendet sich Barad explizit gegen wissenschaftliche Vorgehensweisen, die auf die Überbrückung von Differenzen abzielen.¹¹² Es gehe ihr nicht um »bidirektionale[] Ansätze[...], die die Ergebnisse dessen addieren, was passiert, wenn jede Theorie mal an der Reihe ist, das Gegenstück zu spielen«.¹¹³ Eher werden die aus den

»verschiedenen (inter)disziplinären Praktiken hervorgehenden Verstehensweisen miteinander in Konversation [...] gebracht. Das heißt, Aspekte einer jeden dynamisch aufeinander zu beziehen und dabei auf die iterative Produktion von Grenzen, die materiell-diskursive Beschaffenheit von Praktiken der Grenzziehung [...] zu achten«.¹¹⁴

Barads neo-materialistische Konzeption des Agentiellen Realismus kann nun wie folgt zusammengefasst werden: Ausgehend von einem relationalen Impetus schlägt Barad eine performative Ontologie wie Epistemologie (vgl. »Onto-Epistemologie«) vor, die für sie aus miteinander intraagierenden Phänomenen hervorgebracht wird. Einzelne Phänomene konstituieren sich durch agentielle Schnitte, die zum Beispiel durch wissenschaftliche Messinstrumente erzeugt werden. Zur Untersuchung dieser multiplen verschränkten

¹⁰⁷ Barad 2018, S. 21.

¹⁰⁸ Bath 2013, S. 91. Vgl. Bath et al. (Hg.) 2013, S. 7f.

¹⁰⁹ Barad 2007, S. 29.

¹¹⁰ Barad 2013, S. 60.

¹¹¹ Barad 2007, S. 92.

¹¹² Vgl. Barad 2013, S. 62.

¹¹³ Barad 2013, S. 61.

¹¹⁴ Barad 2013, S. 61.

Wirklichkeit entwickelt sie eine diffraktive Methode, die sich speziell für eine Analyse von Differenzen eignet. Das diffraktive Durch-einander-hindurch-Denken verhält sich konträr zu reflexiven Methoden, die auf Ähnlichkeit bzw. Identität und Repräsentation setzen.

3.2 Spekulative Verschränkungen: Entangled Methods in den Digital Humanities

Barads Entwurf des Agentiellen Realismus bildet nun den Ausgangspunkt, um das Verhältnis von Methoden und Forschungsdaten im Rahmen eines Entangled-Methods-Forschungsdesigns zu imaginieren. Für meinen Versuch einer spekulativen Öffnung und Pluralisierung von Mixed-Methods-Ansätzen schlage ich die folgenden Perspektivwechsel vor: Was wäre, wenn die Kompatibilität in Entangled-Methods-Forschungsdesigns zwar weiterhin vorausgesetzt wird, aber eben nicht als Interaktion sondern Intraaktion von Methoden und Daten vorgestellt wird? Was wäre, wenn die Zielsetzung eines Forschungsdesigns nicht in der Idee einer Vereinheitlichung aufgeht, sondern stattdessen in einer Datendiffraktion an der Aufrechterhaltung von Differenzen interessiert ist? Wie sähe eine Art des Rasonierens aus, welche die Regeln des Schließens an einer Freilegung von Interferenzen orientiert?

Die mit dem Mixed-Methods-Forschungsdesign einhergehende Voraussetzung der Kompatibilität wird sowohl in der empirischen Sozialforschung als auch in den DH häufig als eine Form der methodischen Interaktion präsentiert.¹¹⁵ Bei Herrmann heißt es beispielsweise: »[A] mixed-methods paradigm will allow for the interaction of different mindsets in a structured way.«¹¹⁶ Komplementär wird der Wirkungsbereich des spekulativen Forschungsdesigns nun wie folgt imaginiert: Der Entangled-Methods-Ansatz erlaubt eine Intraaktion, die nach der Konstruktion von methodischen Differenzen fragt. Auf der einen Seite würden dazu die Untersuchungsgegenstände der DH als Phänomene begriffen werden. Untersuchungsgegenstände stellen dann keine Entitäten mehr dar. Vielmehr verwiesen sie auf die materiell-diskursiven Phänomene, die durch iterative Intraaktionen überhaupt erst prozessual hervorgebracht wurden. Bode erklärt: »In digital literary studies, the boundary between text and paratext always matters, both in the sense that it is materialised in our databases, and that it profoundly shapes what we can know of literary history.«¹¹⁷ Auch Trilcke & Fischer sprechen nicht vom Drama als epistemisches Ding ihrer Netzwerkanalysen. Untersuchungsgegenstand ist vielmehr die prozessierte Form des »Zwischenformat[s]«, worunter eine Kombination aus Metadaten und Strukturdatenformaten verstanden wird. Ähnlich argumentiert auch Garvin: »In a vector-space semantic model, words and documents are mutually constituted by the linear transformation of lexical space into bibliographical space.«¹¹⁹ Als Phänomene unterliefen die Untersuchungsgegenstände binären Setzungen, wie analog und digital, sowie Skalierungseffekten. Es gäbe keine Vorher-Nachher-Beziehungen

¹¹⁵ Vgl. Creswell / Plano Clark 2011, S. 114.

¹¹⁶ Herrmann 2017.

¹¹⁷ Bode 2020.

¹¹⁸ Trilcke / Fischer 2018.

¹¹⁹ Garvin 2019, S. 13.

mehr, sondern nur eine fortwährende intraaktive Entfaltung der Phänomene. Die Relationen, welche die Phänomene hervorbrächten, könnten so potenziell überall zusammen- und auseinandergeschnitten werden.

Auf der anderen Seite würden Methoden im Rahmen von Entangled Methods selbst als Apparate beschrieben werden. Als solche »spezifische[n] materielle[n] Rekonfigurationen« stellten sie eine »dynamische Menge von erweiterbaren Praktiken« dar.¹²⁰ Sie vollzögen agentielle Schnitte, die »eindeutige Grenzen und Eigenschaften von Entitäten innerhalb von Phänomenen hervorbringen«.¹²¹ Das Verstehen von Methoden als Apparate würde eine weitere Art des Sprechens darüber ermöglichen, was Methoden sowie Tools tatsächlich täten.¹²² Für Entangled Methods folgte daraus der Anspruch, dass die Versionierung von diesen bedeutungsgebenden agentellen Schnitten in ihrer Kontextualität und Situiertheit sichtbar gemacht, dokumentiert und diskutiert würde. Darüber hinaus basierte das Entangled-Methods-Forschungsdesign im Unterschied zum Mixed-Methods-Forschungsdesign auf einer Relationalität der Phänomene, auf deren Grundlage sich Grenzziehungen und Überlagerungsprozesse erst performativ ereignen.¹²³ Entangled Methods stellten einen gemeinsamen (Nicht-)Ort der DH-Wissenskultur dar, der in Kritik an diskursiven Praktiken, Mechanismen und Narrativen einer Wissensrepräsentation Umgangsformen mit Performativität und Relationalität entfalten würde. Das heißt, die Entangled-Methods-Ansätze beförderten unter anderem Praktiken einer Datenmodellierung, die über Formen der Repräsentation hinausgingen und auf eine Erweiterung eines repräsentationalen Repertoires drängten. Wie sähe dann eine Modellierung beispielsweise von Events in CIDOC CRM aus? Neben der Dokumentation von ›what has happened‹¹²⁴ träte die Auslotung der Frage ›what could have happened‹, die Zugänge zu performativen Messungen von Daten gestaltet würde.¹²⁵

Als eine Zielsetzung von Mixed-Methods-Ansätzen wird die Integration von Forschungsdaten aus unterschiedlichen Quellen genannt. Diese Zielsetzung wird innerhalb der Sozialforschung von Uprichard & Dawney bereits als eine »orthodoxy of integration«¹²⁶ zur Disposition gestellt. Das Mixed-Methods-Forschungsdesign gehe von der Komplexität und Multidimensionalität des Untersuchungsgegenstandes aus, die über die Kombination von unterschiedlichen Methoden adressiert werde. Gleichzeitig ziele aber das Forschungsdesign darauf, ein kohärentes Bild des als komplex beschriebenen Untersuchungsgegenstandes zu generieren. Als eine mögliche Alternative schlagen sie die Datendiffraktion vor: »Diffraction provides a useful alternative to integration; whereas integration assumes that mixed data can be somehow brought together to shed light on a presupposed phenomenon, diffraction emphasizes difference and entanglements«.¹²⁷ Konträr zur ›integration challenge‹ formulieren Uprichard & Dawney:

¹²⁰ Vgl. Bode 2020; Østerlund et al. 2020, S. 6.

¹²¹ Barad 2018, S. 35.

¹²² Vgl. Fox / Alldred 2018, S. 194.

¹²³ Vgl. Eickelmann 2020.

¹²⁴ Vgl. Bekiari et al. (Hg.) 2021, S. xix.

¹²⁵ Vgl. Elswit 2021, S. 389f.

¹²⁶ Uprichard / Dawney 2019, S. 20.

¹²⁷ Uprichard / Dawney 2019, S. 26.

»[The] diffractive challenge responds to the imperative to acknowledge that [...] phenomena can only be partially empirically captured. Diffraction assumes that the whole is always part of something else and that, sometimes, research thoroughly confuses and messes up what we see as the parts and wholes of what we are studying.«¹²⁸

Das von Uprichard & Dawney entwickelte Konzept der Datendiffraktion möchte ich als eine Zielsetzung eines Entangled-Methods-Ansatzes in den DH diskutieren. Wie könnte sich die DH im Rahmen von Entangled Methods einer solchen diffraktiven Herausforderung annehmen?

Herrmann betont im Rahmen ihres Mixed-Methods-Ansatzes: »[M]oving beyond here cannot possibly mean resorting to an ›everything-goes‹ approach that defocuses the nitty-gritty details of difference [...].«¹²⁹ Genau hier in einer feinkörnigen Auflösung von nuancierten Differenzen läge nun das neue Einsatzgebiet der Datendiffraktion, die auf das Ideal der Integration mit neuen Praktiken des Verbindens jenseits der Überbrückung antwortet. Latour beispielsweise zeichnet das folgende Bild:

»[Y]ou are trying to build a bridge over a rather tumultuous river. [...] Now suppose that, instead of trying to cross this river and build this bridge, you decide instead to go with the flow – that is, to get involved in a bit of canoeing, kayaking, or rafting. Then the absence of a bridge is not such a problem [...].«¹³⁰

Datendiffraktion als Ziel von Entangled Methods könnte also mit der Aufforderung einhergehen, bildlich gesprochen: in den Fluss zu steigen, um die Flussufer aus seitlicher fließender Richtung zu betrachten. Zu den Praktiken der Operationalisierung im Rahmen von Entangled Methods zählten dann neben dem Brückenbau auch Kayaking und Rafting, die sich auf praxeologischer Ebene in konkreten Arbeitsschritten manifestierten. Darüber hinaus reformulierte die Datendiffraktion Ansprüche an die Erhebung, Sammlung, Analyse und Evaluierung von Daten. Die Datendiffraktion legitimierte und stieß ein kritisches Durcheinander-hindurch-Lesen von Daten und deren spezifischen Settings an, die in ihrer lokalen Situiertheit und Kontextualität erfasst und beschrieben würden. RDF Triples ließen sich zugleich als formale Aussagen über eine Entität sowie als multiple Verschränkung von Relationen in den Blick nehmen. Auszeichnungssprachen, Normdateien und Ontologien würden im Rahmen von Entangled Methods zu Angelegenheiten eines aktiven und kollaborativen ›Zusammen-Auseinander-Schneidens‹.¹³¹ Es gäbe unter anderem Raum für Phasen eines aktiven ›ontology hijacking‹,¹³² um die damit einhergehenden Irritationen, Widerstände und logischen Fehlschlüsse als neue Verschränkungen eines interdisziplinären Austausches produktiv zu machen (vgl. FAIR und CARE Datenprinzipien). Auf diese Weise adressierte die Datendiffraktion Daten als *capta* im Sinne Druckers.¹³³ Die Gemachtheit der Forschungsdaten würde über die Datendiffraktion in ein Interferenzmuster aufgefächert.

¹²⁸ Uprichard / Dawney 2019, S. 28.

¹²⁹ Herrmann 2017.

¹³⁰ Latour 2008, S. 13.

¹³¹ Barad 2015, S. 183.

¹³² Eide / Smith-Ore 2019, S. 188.

¹³³ Vgl. Drucker 2011.

Solche Interferenzmuster bildeten ein Forum, beispielsweise über geteilte Datenprinzipien von FAIR und CARE in den DH zu verhandeln.¹³⁴ Ein Aufgabenfeld von Entangled Methods bestünde ferner in einer Exploration von Visualisierungen und Prototypen, die die Effekte von Grenzziehungen und Überlagerungen adressieren und re-kontextualisieren.¹³⁵ Hier ginge es um Anschlüsse an eine bereits in den Science and Technology Studies praktizierte Form des diffraktiven Designs.¹³⁶

Daran schließt sich eine dritte spekulative Verschränkung an, die nun die Art des Rasonierens von Entangled-Methods-Ansätzen bestimmt. Eine Funktion der Integration im Rahmen von Mixed-Methods-Ansätzen besteht darin, dass auf Basis der verknüpften Datensammlungen die Bildung von formalen (Meta-)Inferenzen möglich ist. Wie könnten unter der Prämisse eines relationalen Verständnisses von Methoden und Daten aus unterschiedlichen Datensammlungen Ableitungen gewonnen werden? Kennzeichen der Art des Rasonierens im Rahmen von Entangled Methods wäre die erkenntnistheoretische Annahme, dass sich die Inferenzen nicht auf eine objektive Referenz mit inhärenten Eigenschaften, sondern auf materiell-diskursive Phänomene bezögen. In der feministischen Technikforschung schlägt Bath vor, den Prozess des Schließens als Form der Bildung von Interferenzen zu imaginieren. Am Beispiel von Linked Open Data macht sie deutlich, dass in unterschiedlichen Datensammlungen kodiertes Wissen aus unterschiedlichen Bereichen vorliege. Daher komme es bei semantisch verknüpften Datensätzen »zu Überlagerungen von Aussagen über die Welt, die aus höchst unterschiedlichen Bereichen stammen, z. T. auf Grundlage verschiedener Epistemologien erstellt worden sind und je spezifische Ein- und Ausschlüsse bergen«. ¹³⁷ Im Rahmen von Entangled Methods ginge es um die Einübung einer Art des Rasonierens, die über die Regelbildung auf Basis des repräsentierten Wissens hinaus versuchte, neue Relationen herzustellen zwischen dem, was in den Datensammlungen noch verschränkt und eingefaltet wäre. Im Rahmen von Entangled Methods träten Ableitungen auf Basis von repräsentiertem Wissen neben Schlussfolgerungen, die auf die Kontingenz und Kontextualität eines Referenzbereichs (z. B. einer Domain) hindeuten. Es ginge um eine Art des Rasonierens, die an den konstitutiven Ausschlüssen der Bedeutungsbildung von Datensammlungen innerhalb eines kombinierten Methodeneinsatzes interessiert wäre. Zentrale Fragen wären: Was erfährt im Rahmen eines kombinierten Methodeneinsatzes wie und warum Bedeutung? Was wird bereits mit der Methodenwahl oder Toolverwendung vorausgesetzt oder ausgeschlossen? Diese Arten des Rasonierens würden zum einen hybride Formen des wissenschaftlichen Outputs erfordern, die Rekonfigurationen von Bedeutungen im besonderen Maße Rechnung trügen und sie damit zugänglich machten für eine partizipative Zusammenarbeit zugänglich machten. Zum anderen eröffneten sie den Raum für eine zukünftige computationale Methodenkritik, welche die Entwicklung neuer Werkzeuge kritisch begleiten würde. ¹³⁸

¹³⁴ Barad 2015, S. 183.

¹³⁵ Vgl. Loukissas 2019, S. 125.

¹³⁶ Vgl. Bath 2013, S. 108f.; Soon / Cox 2020, S. 187f.

¹³⁷ Bath 2013, S. 94.

¹³⁸ Vgl. Flanders 2019.

Forschungsdesigns können mit Knorr-Cetina als Kette von gemachten Entscheidungen oder als Übersetzungen von Selektionen beschrieben werden, die wiederum selbst Selektionen erfordern. Knorr-Cetina hält weiter fest: »Selektionen können genau deshalb [innerhalb einer Wissenskultur] kritisiert werden, weil sie Selektionen sind: das heißt, gerade weil sie die Möglichkeit alternativer Selektionen einschließen.«¹³⁹ Das Entangled-Methods-Forschungsdesign ist genau das: Es ist eine Entfaltung alternativer Selektionsprozesse im Rahmen von kombinierten Methodeneinsätzen, die über das Vehikel des Agentiellen Realismus in den Blick genommen werden. So sieht das Entangled-Methods-Forschungsdesign eine Vorstellung von Untersuchungsgegenständen als materiell-diskursive Phänomene vor, die von Methoden und Tools erst performativ hervorgebracht werden. Als eine Zielsetzung des Entangled-Methods-Forschungsdesigns wird die Datendiffraktion diskutiert, die zugleich mit einer Art des materiell-diskursiven Rasonierens einhergeht.

4. Schluss

Der vorliegende Beitrag nähert sich der Frage nach der Fabrikation von Erkenntnis über das Forschungsdesign von Mixed Methods in den DH. Dazu stellt der Artikel zwei Zugänge vor: In einem ersten deskriptiven Zugang werden Prämissen, Zielsetzungen und Erkenntnisbedingungen des Mixed-Methods-Forschungsdesigns anhand dreier Merkmale diskutiert. Dazu zählen

1. eine Kompatibilität zwischen divergenten Forschungsparadigmen unter dem pragmatischen Anspruch einer Forschungsfrage zu beantworten,
2. eine sinnhafte Integration der Methoden und der daraus resultierenden Forschungsdaten,
3. Möglichkeiten, Inferenzen für den Untersuchungsgegenstand auf Basis der beiden methodischen Verfahrensweisen und ihrer Ergebnisse zu bilden.

Diese drei Merkmale bilden einen Rahmen, Mixed-Methods-Ansätze in der DH insbesondere den digitalen Literaturwissenschaften, zu vermessen. In einem zweiten spekulativen Zugang wird das Forschungsdesign Entangled Methods in Rekurs auf den neo-materialistischen Ansatz des Agentiellen Realismus von Karen Barad diskutiert. Als Merkmale dieses Forschungsdesigns werden *Intraaktion*, *Datendiffraktion* und *Interferenz* eingeführt und diskutiert. Jüngst hat Risam in ihrer Definition der DH erneut die Forderung formuliert, dass eine gegenwärtige Aufgabe darin bestehe, Formen der Differenz und Heterogenität unter einem digitalen Paradigma zu verhandeln.¹⁴⁰ Auch Flanders plädiert für eine »more diverse programming culture«.¹⁴¹ Die vorgestellte Nebeneinanderstellung der Mixed-Methods- und der Entangled-Methods-Forschungsdesigns verstehe ich als einen Beitrag, eben dieser Aufgabe nachzukommen. Im Rahmen meines Beitrags konnten nicht nur nuancierte Unterschiede zwischen den Geistes- und Sozialwissenschaften, sondern auch gegenwärtige Einsichten in Methodenkonzepte, Metaphern und Praktiken der DH, insbesondere der digitalen Literaturwissenschaften, geteilt

¹³⁹ Knorr-Cetina 2002b, S. 27.

¹⁴⁰ Vgl. Risam 2021, S. 162.

¹⁴¹ Flanders 2019.

werden. Das Entangled-Methods-Forschungsdesigns als spekulativer (Nicht-)Ort ist eine Einladung, diffraktiv die Entscheidungsgeladenheiten der Wissensproduktion in den DH zu gestalten.

Bibliographische Angaben

- Karen Barad: Agentieller Realismus. Über die Bedeutung materiell-diskursiver Praktiken. 3. Auflage. Berlin 2018. [[Nachweis im GBV](#)]
- Karen Barad: Diffractionen: Differenzen, Kontingenzen und Verschränkungen von Gewicht. In: Geschlechter Interferenzen. Wissensformen – Subjektivierungsweisen – Materialisierungen. Hg. von Corinna Bath / Hanna Meißner / Stephan Trinkaus / Susanne Völker. Berlin 2013, S. 27–67. (= Geschlechter Interferenzen, 1) [[Nachweis im GBV](#)]
- Karen Barad: Meeting the Universe Halfway. Quantum Physics and the Entanglement of Matter and Meaning. Durham 2007. [[Nachweis im GBV](#)]
- Karen Barad: Verschränkungen. Berlin 2015. (= Internationaler Merve-Diskurs, 409). [[Nachweis im GBV](#)]
- Geschlechter Interferenzen. Wissensformen – Subjektivierungsweisen – Materialisierungen. Hg. von Corinna Bath / Hanna Meißner / Stephan Trinkaus / Susanne Völker. Berlin 2013. (= Geschlechter Interferenzen, 1) [[Nachweis im GBV](#)]
- Corinna Bath: Semantic Web and Linked Open Data: Von der Analyse technischer Entwicklungen zum Diffractive Design. In: Geschlechter Interferenzen. Wissensformen – Subjektivierungsweisen – Materialisierungen. Hg. von Corinna Bath / Hanna Meißner / Stephan Trinkaus / Susanne Völker. Berlin 2013, S. 69–116. (= Geschlechter Interferenzen, 1) [[Nachweis im GBV](#)]
- Nina Baur / Udo Kelle / Udo Kuckartz: Mixed Methods – Stand der Debatte und aktuelle Problemlagen. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie 69 (2017), Sup. 2, S. 1–37. [[Nachweis im GBV](#)]
- Pat Bazeley: The bricoleur with a computer, piecing together qualitative and quantitative data. In: Qualitative Health Research 9 (1999), H. 2, S. 279–287. [[Nachweis im GBV](#)]
- Christoph Beierle / Gabriele Kern-Isberner: Methoden wissenschaftlicher Systeme. Grundlagen, Algorithmen, Anwendungen. 5., überarbeitete und erweiterte Auflage. Wiesbaden 2014. [[Nachweis im GBV](#)]
- Volume A: Definition of the CIDOC Conceptual Reference Model. Hg. von Chrissy Bekiari / George Bruseker / Martin Doerr / Christian-Emil Ore / Stephan Stead / Athanasios Velios. Version 7.1 von März 2021. [[online](#)]
- Toni Bernhart: Quantitative Literaturwissenschaft: Ein Fach mit langer Tradition? In: Quantitative Ansätze in den Literatur- und Geisteswissenschaften. Systematische und historische Perspektiven. Hg. von Andrea Albrecht / Marcus Willand / Sandra Richter / Toni Bernhart. Berlin u. a. 2018, S. 207–219. DOI: [10.1515/9783110523300-009](https://doi.org/10.1515/9783110523300-009) [[Nachweis im GBV](#)]
- David M. Berry / Anders Fagerjord: Digital Humanities. Knowledge and Critique in a Digital Age. Oxford 2017. [[Nachweis im GBV](#)]
- Katherine Bode: Data beyond representation: Thoughts on the shift from computational modelling to performative materiality. In: Program of the 2020 convention. Hg. von MLA Convention. (MLA: 135, Toronto, 07.-10.01.2020) New York, NY 2020. [[online](#)]
- Digital Humanities. Hg. von Anne Burdick / Johanna Drucker / Peter Lunenfeld / Todd Presner / Jeffrey Schnapp. Cambridge, MA 2012. [[Nachweis im GBV](#)]
- John W. Creswell / Vicki L. Plano Clark: Designing and conducting mixed methods research. 2. Auflage. Los Angeles, CA u. a. 2011. [[Nachweis im GBV](#)]
- Nan Z. Da (2019a): The Computational Case against Computational Literary Studies. In: Critical Inquiry 45 (2019), H. 3, S. 601–639. [[Nachweis im GBV](#)]
- Nan Z. Da (2019b): Computational Literary Studies: Participant Forum Responses, Day 2. In: Critical Inquiry. critinq.wordpress.com. Blogbeitrag vom 02.04.2019. [[online](#)]
- Norman Kent Denzin: Triangulation 2.0. In: Journal of mixed methods research 6 (2012), H. 2, S. 80–88. [[Nachweis im GBV](#)]
- James E. Dobson: Critical Digital Humanities. The Search for a Methodology. Urbana, IL u. a. 2019. [[Nachweis im GBV](#)]
- Johanna Drucker: Humanities Approaches to Graphical Display. In: Digital Humanities Quarterly 5 (2011), H. 1. [[online](#)]
- Johanna Drucker: The Digital Humanities Coursebook. An Introduction to Digital Methods for Research and Scholarship. London u. a. 2021. [[Nachweis im GBV](#)]
- Jennifer Eickelmann: Digitale Medien und Methoden. Jennifer Eickelmann zu Diffraction als Methode. In: Open Media Studies. Blog zu Open Access und Open Science in der Medienwissenschaft. Marburg 2020. Blogbeitrag vom 30.04.2020. Version vom 28.05.2020. [[online](#)]
- Øyvind Eide / Christian-Emil Smith-Ore: Ontologies and data modeling. In: The shape of data in the digital humanities. Modeling texts and text-based resources. Hg. von Julia Flanders / Fotis Jannidis. London u. a. 2019, S. 178–196. (= Digital research in the arts and humanities) [[Nachweis im GBV](#)]
- Kate Elswit: Performative Measure. In: Uncertain Archives. Critical Keywords for Big Data. Hg. von Nanna Bode Thylstrup / Daniela Agostinho / Annie Ring / Catherine D'Ignazio / Kristin Veel. Cambridge, MA u. a. 2021, S. 389–396. [[Nachweis im GBV](#)]

- Martina Yvonne Feilzer: Doing Mixed Methods Research Pragmatically: Implications for the Rediscovery of Pragmatism as a Research Paradigm. In: Journal of mixed methods research 4 (2010), H. 1, S. 6–16. [[Nachweis im GBV](#)]
- Michael Fetters / Dawn Freshwater: The 1 + 1 = 3 Integration Challenge. In: Journal of mixed methods research 2015 (9), H. 2, S. 115–117. [[Nachweis im GBV](#)]
- Michael Fetters / José Molina-Azorin: The Journal of Mixed Methods Research Starts a New Decade: The Mixed Methods Research Integration Trilogy and Its Dimensions. In: Journal of mixed methods research 13 (2017), H. 11, S. 291–307. [[Nachweis im GBV](#)]
- Nigel G. Fielding: Triangulation and Mixed Methods Designs. In: Journal of mixed methods research 6 (2012), H. 2, S. 124–136. [[Nachweis im GBV](#)]
- Julia Flanders: Building otherwise. In: Bodies of Information. Intersectional Feminism and Digital Humanities. Hg. von Elizabeth Losh / Jacqueline Wernimont. Minneapolis, MN u. a. 2019. [[Nachweis im GBV](#)]
- Uwe Flick: Triangulation. Eine Einführung. 3., aktualisierte Auflage. Wiesbaden 2011. [[Nachweis im GBV](#)]
- Nick J. Fox / Pam Alldred: Mixed methods, materialism and the micropolitics of the research-assembly. In: International Journal of Social Research Methodology 21 (2018), H. 2, S. 191–204. [[Nachweis im GBV](#)]
- Michael Gavin: Is there a text in my data? (Part 1): on counting words. In: Journal of Cultural Analytics 5 (2020), H. 1. DOI: [10.22148/001c.11830](https://doi.org/10.22148/001c.11830)
- Evelyn Gius: Computationale Textanalysen als fünfdimensionales Problem: Ein Modell zur Beschreibung von Komplexität. In: LitLab Pamphlet 8 (2019), S. 1–20. [[Nachweis im GBV](#)]
- Andrew Goldstone: Let DH Be Sociological! In: Book of Abstracts: Digital Humanities Conference. Hg. von Ecole Polytechnique Fédérale de Lausanne. (DH 2014, Lausanne, 08–12.07.2014) Lausanne 2014, S. 171–174. PDF. [[online](#)] [[Nachweis im GBV](#)]
- Naomie Gramlich: Feministisches Spekulieren. Einigen Pfaden folgen. In: Feministisches Spekulieren. Genealogien, Narrationen, Zeitlichkeiten. Hg. von Marie-Luise Angerer / Naomie Gramlich. Berlin 2020, S. 9–32. [[Nachweis im GBV](#)]
- Jennifer C. Greene: Is Mixed Methods Social Inquiry a Distinctive Methodology? In: Journal of mixed methods research 2 (2008), H. 1, S. 7–22. [[Nachweis im GBV](#)]
- Annemarie Hahn: Every Things Matter. In: Postdigital Landscapes. Kunst und Medienbildung in der digital vernetzten Welt. Hg. von Kristin Klein / Willy Noll. Zeitschrift Kunst Medien Bildung. Artikel vom 04.10.2019. [[online](#)]
- Gary Hall: Has Critical Theory Run Out of Time for Data-Driven Scholarship? In: Debates in the Digital Humanities. Hg. von Matthew K. Gold. Minneapolis, MN 2012. [[online](#)] [[Nachweis im GBV](#)]
- Berenike J. Herrmann: In a test bed with Kafka. Introducing a mixed-method approach to digital stylistics. In: Digital Humanities Quarterly 11 (2017), H. 4. [[online](#)]
- Katharina Hoppe / Thomas Lemke: Die Macht der Materie. Grundlagen und Grenzen des agentiellen Realismus von Karen Barad. In: Soziale Welt 66 (2015), H. 3, S. 261–279. [[Nachweis im GBV](#)]
- Katharina Hoppe / Thomas Lemke: Neue Materialismen – zur Einführung. Hamburg 2021. [[Nachweis im GBV](#)]
- Jan Horstmann / Rabea Kleymann: Alte Fragen, neue Methoden – Philologische und digitale Verfahren im Dialog. Ein Beitrag zum Forschungsdiskurs um Entsagung und Ironie bei Goethe. In: Zeitschrift für digitale Geisteswissenschaften 4 (2019). Artikel vom 12.12.2019. DOI: [10.17175/2019_007](https://doi.org/10.17175/2019_007)
- Kenneth R. Howe: Against the Quantitative-Qualitative Incompatibility Thesis. In: Educational Researcher November 17 (1988), H. 8, S. 10–16. [[Nachweis im GBV](#)]
- Digital Humanities. Eine Einführung. Hg. von Fotis Jannidis / Hubertus Kohle / Malte Rehbein. Stuttgart 2017. [[Nachweis im GBV](#)]
- Burke R. Johnson / Anthony J. Onwuegbuzie / Lisa A. Turner: Toward a Definition of Mixed Methods Research. In: Journal of mixed methods research 1 (2007), H. 2, S. 112–133. [[Nachweis im GBV](#)]
- Friedrich Kaulbach / Ludger Oeing-Hanhoff / Theo Herrmann / Heinrich Beck: Ganzes/Teil. In: Historisches Wörterbuch der Philosophie. Hg. von Joachim Ritter. 12 Bde. Basel u. a. 1974, S. 3–20. [[Nachweis im GBV](#)]
- Udo Kelle: Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte. 2. Auflage. Wiesbaden 2008. [[Nachweis im GBV](#)]
- Philologisches Wörterbuch. Hg. von Georg Klaus / Manfred Buhr. 12., neubearb. und durchges. Aufl. Bd. 2. Berlin 1976. [[Nachweis im GBV](#)]
- Karin Knorr-Cetina (2002a): Wissenskulturen. Ein Vergleich naturwissenschaftlicher Wissensformen. Frankfurt / Main 2002. [[Nachweis im GBV](#)]
- Karin Knorr-Cetina (2002b): Die Fabrikation von Erkenntnis. Zur Anthropologie der Naturwissenschaft. Revidierte und erweiterte Neuauflage, 2. Auflage. Frankfurt / Main. 2002. [[Nachweis im GBV](#)]
- Albrecht Koschorke: Einleitung. In: Komplexität und Einfachheit. DFG-Symposium 2015. Hg. von Albrecht Koschorke. (DFG-Symposium, Menaggio, 2015) Stuttgart 2017, S. 1–11. [[Nachweis im GBV](#)]

Benjamin Krautter / Marcus Willand: Close, Distant, Scalable: Skalierende Textpraktiken in der Literaturwissenschaft und den Digital Humanities. In: Ästhetik der Skalierung. Hg. von Carlos Spoerhase / Steffen Siegel / Nikolaus Wegmann. Hamburg 2020, S. 77–97. (= Zeitschrift für Ästhetik und Allgemeine Kunstwissenschaft / Sonderhefte, 18) [\[Nachweis im GBV\]](#)

Udo Kuckartz: Mixed Methods. Methodologie, Forschungsdesigns und Analyseverfahren. Wiesbaden 2014. [\[Nachweis im GBV\]](#)

Jonas Kuhn: Computational text analysis within the Humanities: How to combine working practices from the contributing fields? In: Lang Resources & Evaluation 53 (2019), S. 565–602. DOI: [10.1007/s10579-019-09459-3](https://doi.org/10.1007/s10579-019-09459-3) [\[Nachweis im GBV\]](#)

Bruno Latour: What is the Style of Matters of Concern? Two lectures in empirical philosophy. Amsterdam 2008. [\[Nachweis im GBV\]](#)

Thomas Lemke: Einführung – Neue Materialismen. In: Science and Technology Studies. Klassische Positionen und aktuelle Perspektiven. Hg. von Susanne Bauer / Torsten Heinemann / Thomas Lemke. 2. Auflage. Berlin 2020, S. 551–574. [\[Nachweis im GBV\]](#)

Yvonna Sessions Lincoln / Egon G. Guba: Paradigmatic Controversies, Contradictions, and Emerging Confluences. In: The Sage handbook of qualitative research. Hg. von Norman Kent Denzin / Yvonna Sessions Lincoln et al. 2. Auflage. Thousand Oaks u. a. 2000, S. 163–188. [\[Nachweis im GBV\]](#)

Alan Liu: The Meaning of the Digital Humanities. In: PMLA 128 (2013), H. 2, S. 409–423. [\[Nachweis im GBV\]](#)

Yanni Alexander Loukissas: All data are local. Thinking critically in a data-driven society. Cambridge, MA u. a. 2019. [\[Nachweis im GBV\]](#)

Niklas Luhmann: Komplexität. In: Historisches Wörterbuch der Philosophie. Hg. von Joachim Ritter. 12 Bde. Basel u. a. 1976, Bd. 4: I-K, S. 939–942. [\[Nachweis im GBV\]](#)

Enzyklopädie Philosophie und Wissenschaftstheorie. Band 5: Log-N. Hg. von Jürgen Mittelstraß. 2., neubearbeitete und wesentlich ergänzte Auflage. Stuttgart 2013. [\[Nachweis im GBV\]](#)

Franco Moretti: Operationalizing: or, the function of measurement in modern literary theory. In: Pamphlets of the Stanford Literary Lab 6 (2013), S. 1–13. PDF. [\[online\]](#)[\[Nachweis im GBV\]](#)

David L. Morgan: Paradigms Lost and Pragmatism Regained. In: Journal of mixed methods research 1 (2007), H. 1, S. 48–76. [\[Nachweis im GBV\]](#)

Methoden der literatur- und kulturwissenschaftlichen Textanalyse. Ansätze – Grundlagen – Modellanalysen. Hg. von Vera Nünning / Ansgar Nünning. Stuttgart u. a. 2010. [\[Nachweis im GBV\]](#)

Carsten Østerlund / Kevin Crowston / Corey Jackson: Building an Apparatus: Refractive, Reflective, and Diffractive Readings of Trace Data. In: Journal of the Association for Information Systems 21 (2020), H. 1. [\[Nachweis im GBV\]](#)

Urszula Pawlicka-Deger: The Laboratory Turn: Exploring Discourses, Landscapes, and Models of Humanities Labs. In: Digital Humanities Quarterly 14 (2020), H. 3. [\[online\]](#)

Andrew Piper: Enumerations. Data and Literary Study. Chicago, IL u. a. 2018. [\[Nachweis im GBV\]](#)

Andrew Piper: Can We Be Wrong? The Problem of Textual Evidence in a Time of Data. Cambridge u. a. 2020. [\[Nachweis im GBV\]](#)

Reflektierte algorithmische Textanalyse. Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt. Hg. von Nils Reiter / Axel Pichler / Jonas Kuhn. Berlin u. a. 2020. [\[Nachweis im GBV\]](#)

Sandra Richter: Reading with the Workflow. In: Reflektierte algorithmische Textanalyse. Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt. Hg. von Nils Reiter / Axel Pichler / Jonas Kuhn. Berlin u. a. 2020, S. 143–168. [\[Nachweis im GBV\]](#)

Bernhard Rieder / Theo Röhle: Digital Methods: Five Challenges. In: Understanding Digital Humanities. Hg. von David M. Berry. Houndmills, NY 2012, S. 67–84. [\[Nachweis im GBV\]](#)

Bernhard Rieder / Theo Röhle: Digital Methods. From Challenges to Bildung. In: The Datafied Society. Studying Culture through Data. Hg. von Karin van Es / Mirko Tobias Schäfer. Amsterdam 2017, S. 109–125. [\[Nachweis im GBV\]](#)

Roopika Risam: Digital Humanities. In: Uncertain Archives. Critical Keywords for Big data. Hg. von Nanna Bonde Thylstrup / Daniela Agostinho / Annie Ring / Catherine D'Ignazio / Kristin Veel. Cambridge, MA u. a. 2021. [\[Nachweis im GBV\]](#)

Richard Rogers: Digital methods. Cambridge, MA u. a. 2013 [\[Nachweis im GBV\]](#)

Ned Rossiter: Materialities of Software. In: Advancing Digital Humanities. Research, Methods, Theories. Hg. von Paul Longley Arthur / Katherine Bode. Basingstoke 2014, S. 221–240. [\[Nachweis im GBV\]](#)

Moacir P. de Sá Pereira: Mixed Methodological Digital Humanities. In: Debates in the Digital Humanities. Hg. von Matthew K. Gold / Lauren Klein. Minneapolis, MN 2019. [\[online\]](#) [\[Nachweis im GBV\]](#)

Tom Scheinfeldt: Why Digital Humanities Is »Nice«. In: Debates in the Digital Humanities. Hg. von Matthew K. Gold. Minneapolis, MN 2012. [\[online\]](#) [\[Nachweis im GBV\]](#)

Christof Schöch / Karina van Dalen-Oskam / Maria Antoniak / Fotis Jannidis / David Mimno: Replication and Computational Literary Studies. In: The Digital Humanities Conference 2020 (DH2020, Ottawa, 20.–25.07.2020) Ottawa 2020. DOI: [10.5281/zenodo.3893428](https://doi.org/10.5281/zenodo.3893428)

Digital Methods in the Humanities. Challenges, Ideas, Perspectives. Hg. von Silke Schwanndt. Bielefeld 2021. [[Nachweis im GBV](#)]

Winnie Soon / Geoff Cox: Aesthetic Programming. A handbook of Software Studies. Open Humanities Press 2020. PDF. [[online](#)]

Abbas Tashakkori / Charles Teddlie: Quality of Inferences in Mixed Methods Research. Calling for an Integrative Framework. In: Advances in Mixed Methods Research: Theories and Applications. Hg. von Manfred Max Bergman. Los Angeles, CA 2008, S. 101–119. [[Nachweis im GBV](#)]

Charles Teddlie / Abbas Tashakkori: Mixed Methods Research. Contemporary Issues in an Emerging Field. In: The SAGE Handbook of Qualitative Research. Hg. von Norman Kent Denzin / Yvonna Sessions Lincoln. 4. Auflage. Los Angeles, CA u. a. 2011, S. 285–301. [[Nachweis im GBV](#)]

Textklang. Hg. von der Universität Stuttgart. Stuttgart 2021. [[online](#)]

Peer Trilcke / Frank Fischer: Literaturwissenschaft als Hackathon. Zur Praxeologie der Digital Literary Studies und ihren epistemischen Dingen. In: Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden. Hg. von Martin Huber und Sybille Krämer. Wolfenbüttel 2018 (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 3). DOI: [10.17175/sb003_003](https://doi.org/10.17175/sb003_003)

Emma Uprichard / Leila Dawney: Data Diffraction: Challenging Data Integration in Mixed Methods Research. In: Journal of mixed methods research 13 (2019), H. 1, S. 19–32. [[Nachweis im GBV](#)]

Interaktion qualitativ-hermeneutischer Verfahren und Digital Humanities: »Mixed Methods« in den Geisteswissenschaften? Hg. von Volkswagen Stiftung. Hannover 2021. [[online](#)]

Thomas Weitin (2015a): Digitale Literaturwissenschaft. In: Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte 89 (2015), H. 4, S. 651–656. [[Nachweis im GBV](#)]

Thomas Weitin (2015b): Thinking slowly. Literatur lesen unter dem Eindruck von Big Data. (= Konstanz LitLingLab Pamphlete, 1) URN: [urn:nbn:de:bsz:352-0-285900](https://nbn-resolving.org/urn:nbn:de:bsz:352-0-285900) [[Nachweis im GBV](#)]

Simone Winko: Methode. In: Reallexikon der deutschen Literaturwissenschaft. Neubearbeitung des Reallexikons der deutschen Literaturgeschichte. Hg. von Harald Fricker. 3., neubearbeitete Auflage. 3 Bde. Berlin 1997–2003. Bd. 2 (2000): H–O, S. 581–585. [[Nachweis im GBV](#)]

Weiterführende Literatur

Constanze Baum / Thomas Stäcker: Methoden – Theorien – Projekte. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. Wolfenbüttel 2015. (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1) DOI: [10.17175/sb001_023](https://doi.org/10.17175/sb001_023) [[Nachweis im GBV](#)]

Johanna Drucker / Bethany Nowviskie: Speculative Computing. Aesthetic Provocations in Humanities Computing. In: A companion to digital humanities. Hg. von John Unsworth / Raymond George Siemens / Susan Schreibman. Malden, MA 2004, S. 431–447. (= Blackwell companions to literature and culture, 26) [[Nachweis im GBV](#)]

Uncertain Archives. Critical Keywords for Big Data. Hg. von Nanna Bonde Thylstrup / Daniela Agostinho / Annie Ring / Catherine D'Ignazio / Kristin Veel. Cambridge, MA 2021. [[Nachweis im GBV](#)]

Abbildungs- und Tabellenverzeichnis

Tab. 1: Merkmale von Mixed und Entangled-Methods-Forschungsdesigns. [Kleymann 2022]

Beitrag aus:

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. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Explorative spatial analysis of the function of landscape in video games

Autor*in:

Dominik Kremer

Kontakt: dominik.kremer@fau.de

Institution: Friedrich-Alexander-Universität Erlangen-Nürnberg, Institut für Geographie
GND: [1059545926](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0051-9) ORCID: [0000-0003-1244-7363](https://orcid.org/0000-0003-1244-7363)

Autor*in:

Daniel-Harald Sonnenwald

Kontakt: daniel-harald.sonnenwald@stud.uni-bamberg.de

Institution: Otto-Friedrich-Universität Bamberg
GND: [1237791790](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0051-9) ORCID: [0000-0003-4971-9131](https://orcid.org/0000-0003-4971-9131)

Autor*in:

Blake Byron Walker

Kontakt: blake.walker@fau.de

Institution: Friedrich-Alexander-Universität Erlangen-Nürnberg, Institut für Geographie
GND: [1203153791](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0051-9) ORCID: [0000-0002-1983-3147](https://orcid.org/0000-0002-1983-3147)

DOI des Artikels:

[10.17175/sb005_009](https://doi.org/10.17175/sb005_009)

Nachweis im OPAC der Herzog August Bibliothek:

[1764798465](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0051-9)

Erstveröffentlichung:

20.05.2022

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise: 30.03.2022

GND-Verschlagwortung:

[Videospiele](#) | [Anthropogeographie](#) | [Feldforschung](#) | [Landschaft](#) | [Teilnehmende Beobachtung](#) |

Zitierweise:

Dominik Kremer, Daniel-Harald Sonnenwald, Blake Byron Walker: Explorative spatial analysis of the function of landscape in video games. 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/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_009](https://doi.org/10.17175/sb005_009).

Dominik Kremer, Daniel-Harald Sonnenwald, Blake Byron Walker
Explorative spatial analysis of the function of landscape in video games

Abstracts

Visuelles Landschaftserleben im Sinne der neueren Kulturgeographie bedeutet, dass soziokultureller Kontext und Umwelt erst durch individuelle Bedeutungszuweisungen zueinander in Beziehung gesetzt werden. Landschaft wird also nicht als vorgegeben betrachtet, sondern erst im Moment des Betrachtens diskursiv konstruiert. Über die individuellen Konstruktionsweisen von digitalen Landschaften ist in diesem Zusammenhang noch wenig bekannt. Sind Landschaften in Videospielen nur eine detailreiche Kulisse? Locken am Horizont Affordancen? Oder entwickeln Spieler*innen ihre jeweils eigenen Konstruktionsweisen? Unser methodischer Ansatz zur Beantwortung dieser Fragen ist doppelt experimentell. Zunächst erweitern wir die Methode des textbasierten close reading in seiner Variante des close playing auf digitale walk-alongs im Sinne der geographischen Feldforschung. In einem zweiten Schritt nutzen wir das etablierte Inventar der image schemata für die Annotation von aus Screencasts gewonnenen Keyframes. Hauptbefund ist, dass es eine gewisse Spielfertigkeit erfordert, um eine Landschaft überhaupt individuell anders inszenieren zu können als von der Spielmechanik gefordert. Zweitens kann je nach Spielstil Landschaft als eine Aura wirksam werden, die nicht nur das Spielgeschehen als entfernte Kulisse rahmt, sondern auch die Ziele der Spieler*innen in allen weiteren Interaktionen mit dem Spiel beeinflusst.

Experiencing visual landscape in the sense of newer cultural geography means relating socio-cultural context and environment by individual sense-making. Landscape is therefore not considered pre-given, it is discursively constructed in the very moment of watching. In the context of digital landscapes, little is known of the individual sense-making. Do landscapes only set the scene in video games? Are affordances set on the horizon? Or do the players create their own ways of enacting? Our methodological approach to answer these questions is doubly experimental. First, we extend the method of text-based close reading in its variant of close playing to digital walk-alongs in the sense of geographical field research. In a second step we use the established inventory of image schemata for the annotation of key frames extracted from observed screencasts. As main findings, we found that it requires a certain skill level to enact landscape any different than simply compliant to the game mechanics. Secondly, we found that depending on the style of playing, landscape can be enacted as an aura that not simply frames the game as distant scenery, but also influences the mindset for all further interactions with the game.

1. Introduction

Recently, the long-established study of video games¹ made initial steps towards embracing contemporary conceptualisations and framing of space and place². This study leverages methodological and theoretical advances in cultural geographies and the digital humanities to advocate for close playing³ through a lens of the geographic concept of landscapes,⁴ through which we identify and interrogate the functions and assemblages of spacing⁵ in different types of video games.

¹ Cf. Aarseth et al. 2003.

² Cf. Günzel 2008; Aarseth / Günzel 2019.

³ Cf. Inderst 2020.

2. Related work

2.1 Constituents of gameplay

Our analysis operationalises Aarseth's⁶ core ontology of video games:

- Player: the impersonation of the core game agent by a controlling human being
- Object: a single game element, players can interact with
- Agent: any other being a player can interact with
- Setting: an assemblage of visualized game elements, players cannot interact with without movement
- Event: game situation that occurs regularly or randomly

When conceptualised and presented as embodying *Homo ludens*,⁷ the subject in the player-game relation pursues no higher purpose than generating and extracting joy through software-based freedom of choice in an ad hoc setting. In contrast to the fluidity characterising spaces of play, the act of gaming in the digital space is defined and constrained by a set of well-defined game mechanics, parameters, and affordances. Extrapolating from Esposito, one reduces the video game to a digital apparatus designed and intended for ›ludens‹, in which settings, environments, and narratives collectively constitute a playable space for the resituated subject-avatar.⁸ A resituating of the player is functionalised by, and contingent to, forms and degrees of immersiveness and interactivity; Esposito's definition centres this relation on the interactivity of a video game as mediated through audio and visual stimuli and essentialises the digital as medial. The technical and denarrativized design of video game functionality and elements of play, and consequent research within this sphere, indirectly underscore interactivity and immersion as key elements of the player experience.⁹ Converse to the technical are the strands within game research, which leverage approaches from the social sciences to explore the societal positionality and psychosocial functionality of games, generally through the use of empirical methods such as surveys and controlled laboratory experiments.¹⁰ The resulting body of knowledge demonstrates the linkages between the player experience, game mechanics, and aesthetics, and highlights the centrality of these elements in contemporary game design.

Humanities-informed approaches situate games as constructing and construing meanings that reflect and / or undermine broader sociocultural and media discourses. Empirically, interpretative methods are most commonly used to examine and deconstruct sense-making within video games; close reading, text analysis, and audience theory are borrowed from other media disciplines such as television and film studies. Social science and humanities approaches

⁴ Cf. Rose 1996.

⁵ Cf. Schatzki 2002.

⁶ Cf. Aarseth 2012.

⁷ Cf. Huizinga et al. 1987.

⁸ Cf. Esposito 2005.

⁹ Cf. Williams / Smith 2007.

¹⁰ Cf. Lankoski / Björk 2015.

may therefore overlap, for example in the case of ethnographic or folkloristic studies, where fieldwork may include a form of observation or pseudo-ethnography of games in an attempt to understand their social and cultural meanings.¹¹

We argue that these approaches construct the video game space as a distinctly digital-geographical space, and as such, prior research examining the geographical characteristics of games themselves remain highly relevant.¹² In consideration of these differentiated approaches, we position a spatial situated game mechanics as the primary game function linking the player to his / her own, partially pre-conscious experiences of play.

2.2 Phenomenological framing

Drawing upon methodologies commonly used for real-world environments we argue that the act of experiencing digital landscapes is phenomenologically similar.¹³ Of course, bodily exhaustion caused by movement, weather conditions, or air quality cannot be experienced the same way in digital worlds.¹⁴ As the experience of landscape is bound to a predominantly visual experience¹⁵ due to the avatar being constricted to spaces accessible by locomotion, these embodied qualities do not differ significantly. Our phenomenology of experiencing digital landscapes draws on literature from the social sciences and humanities concerning

- the sense-making of place and
- the stream of consciousness whilst moving and interacting;
- thereafter imbuing our framework with conceptualisations of landscape.

Phenomenologically, sense-making of places and integration by movement are closely linked. Cresswell introduces movement as spatiotemporally enacted practices (moving your hand, dancing, travelling to a different town) by locomotion that affords the integration of places with environments.¹⁶ Mobility is socially produced movement, such that it demonstrates contingency of meaning;¹⁷ constructed meanings of movement constitute the backdrop for geographical imagination and therefore place mobility as a key means of resituating the subject in the digital landscape. Within this flow, places can attract attention and provide fields of action as long as they are stably embedded in their spatial context.¹⁸ Agnew distinguishes between this quasi-stable position (locale) and the individual meaning (sense of place).¹⁹ This meaning is closely bound to human activities enacting it,²⁰ and it is specifically those activities that allows both for observation and social communication. As places are always places of action, they

¹¹ Cf. Egenfeldt-Nielsen et al. 2015; Lankoski / Björk 2015.

¹² Cf. Aarseth et al. 2003; Ash / Gallacher 2011.

¹³ All constituents may be readily complemented easily by drawing upon environmental psychology, for example.

¹⁴ Cf. Cresswell 2006.

¹⁵ Cf. Rose 1996.

¹⁶ Cf. Cresswell 2006; Montello 1993.

¹⁷ Cf. Cresswell 2006.

¹⁸ Cf. Tuan 1977.

¹⁹ Cf. Agnew 1987.

²⁰ Cf. Relph 1976.

can facilitate some actions and impede others;²¹ they offer a field of affordances to which actions can attach.²² We therefore highlight Löw's emphasis on the highly emotional suggestive power of places.²³ If such fields of affordance are visibly used in the same way in the long run, they are affirmed and stabilised by individual routines, and they become social places. Embodiment plays a crucial role at this juncture, as physical settings as well as discourses can always constrain specific actions. Thus, individual judgements regarding which actions are possible at a certain place are not only bound to embodied preferences, but are also extensively socioculturally mediated.²⁴ The self-experiences itself always in a tension between a set of social roles / constraints and a free individual. Essentially, it is the social more than a geometric dimension of reality that evokes the feeling of proximity or distance.²⁵

The implementation of this phenomenological framework in the study of video game spaces as digital landscapes requires that we assess how streams of consciousness assemble and integrate fields of possible actions in a digital living environment. Implemented in a hermeneutic-phenomenological sense following Heidegger,²⁶ an object in the game space is real to its subject if it provides action disposition.²⁷ When considered in conjunction with the emotional aspect described above, a video game can express a high degree of reality independent of its technical presentation or environmental realism.

The agent experiences attention à la vie,²⁸ characterised by a temporary singularity in their experience of a hermeneutic-phenomenological reality. All bodily and haptic manipulations (re-)produce meaning on a site of action, thereby further constituting place. The basic assumption of movement is that reality can be re-established by (loco-)motion and that everything left behind stays unchanged for further revisit. On the move, reality becomes a rolling panorama gradually changing (stream of consciousness) while floating through a world. Actions as transformative manipulations of place are all soaked with memory of earlier experiences and manipulations and thereby modify memory for the next revisit. Crouch references explicitly the concepts of game and place when he describes acting on place as ritualized practice on pre-given codes, habitually repeated.²⁹ Those codes, usually constantly reconfigured, broken, adjusted or negotiated in real world environments, act as preconfigured, purposely placed affordances in gameplay.³⁰

In this setting, the geographic concept of landscape is right at the edge between visual discovery of fields of affordances and a rolling panorama on the move with a certain atmospheric aspect attached to it. Visual experiencing of landscape in the sense of

²¹ Cf. Schatzki 2002.

²² Cf. Gibson 1982.

²³ Cf. Löw 2017.

²⁴ Cf. Löw 2017.

²⁵ Cf. Schütz / Luckmann 2003.

²⁶ Cf. Crowell 2013.

²⁷ Cf. Schütz / Luckmann 2003.

²⁸ Cf. Schütz / Luckmann 2003, p. 6–7; Bergson 1968.

²⁹ Cf. Crouch 2003.

³⁰ Cf. visual clues in chapter 2.3.

contemporary cultural geography thus means the active creation of a relationship between social conditions and environment. Landscape is therefore not pre-given, rather, it is actively constructed in the very moment of watching or discovering while moving.³¹

2.3 Space and place in video games

In recent years, research on video games has been established as a distinguished and unique artwork in media sciences.³² An explicit focus on spatial practices in video games is set by Espen Aarseth and Stephan Günzel who coined the term ludotopia.³³ As shown above, it is the three aspects of (individual) phenomenology, the (objective) epistemology and the cultural significance of space and place that can be questioned.³⁴

Pablo Abend et al. assert that research on playful participatory practices constitutes an important part of digital media culture and art.³⁵ In that context, they argue that playing video games is not simply a use of media, but must be thought of as ongoing (re-)production while playing. Depending on the specific game, certain influences of the player on the game world can be stated according to his / her intentionality.³⁶ In addition, Günzel highlights that games do not simply enact spatial concepts, but are spatial concepts on their own in the sense of enactivism, if they produce bodily experiences (e. g. Wii, but also VR games).³⁷

Domsch stresses that although all video games are preset, rule-bound environments, the player's decisions are relevant.³⁸ Depending on the individual way of enacting, players inscribe a unique narrative into a game setting while moving and interacting in space. In opposition to sequential narratives like videos, every decision taken and performance done contributes to a gradual development of a narrative story. As a discovery journey into fictional otherness, a spatial narrative functions as evocative space meant to resemble carefully constructed existing conceptualisations of space.³⁹ In many cases, it is the explicit ambiguity of game elements that helps to evocate individual narrations. Careful constructions are often offered as visual clues. A motivation to (inter-)act is even stronger, if the player does not notice as such. Visual clues can mark both a possibility for interaction or a directional suggestion.⁴⁰

³¹ Cf. Rose 1996; Rose 2016.

³² Cf. Feige 2015.

³³ Cf. Aarseth / Günzel 2019.

³⁴ Cf. Günzel 2019.

³⁵ Cf. Abend et al. 2020.

³⁶ Cf. Domsch 2019; Günzel 2019.

³⁷ Cf. Günzel 2019.

³⁸ Cf. Domsch 2019.

³⁹ Cf. Domsch 2019.

⁴⁰ Cf. Domsch 2019.

A special form of meta-narrative is offered by the game *Gone Home* which allows it to reflect on cultural artifacts of the 1990s.⁴¹ In this game, players even have the opportunity to reconstruct prominent discourses of this period in their own way. Further, Unterhuber highlights that this process is immensely fostered and only possible by offering it as a game.⁴²

Regarding landscape, games produce their own perspectives. Landscapes can, by design, simply create a mood or atmosphere for a narrative setting,⁴³ but also open a field of open exploration as laid out above. In difference to real world environments, many video games offer shortcuts to prevent real time enactment of navigational tasks.⁴⁴

2.4 Digital walk-alongs

Recurring on the terms close and distant reading,⁴⁵ the basic process of qualitative critical analysis of texts can be easily transferred to a number of analyses, including gameplay. In that sense, for example Rudolf Inderst used the term close playing.⁴⁶ Without naming it, Joshua Tanenbaum and Jim Bizzocchi propose a similar approach in their paper on *Oblivion*.⁴⁷ In detail, they propose two categories of analysis: adaptivity & believability. Whereas adaptivity describes the capability of a game to adapt to the user's needs, believability reflects the capability of agents in the game to act human-like towards the user. Methodically, they provide an elaborated coding system by which they recorded and classified all observable interactions encountered in gameplay.

In our study, we extend close playing to digital walk-alongs in the tradition of geographic field research. Proposed by Monica Degen and Gillian Rose, walk-alongs add to close reading by distinguishing between a person enacting and a person observing.⁴⁸ In their original work, they encourage residents of various cities during an accompanied city walk, to talk as freely as possible about their feelings and evaluations on site. The background is to obtain data as close as possible to the bodily experience on site. Counterintuitively, many subjects on site report their memories of this place in detail, without even taking a look, thus revealing the effect size of individual narratives in sense-making.

Whereas in close playing players and observers are the same person and validity is only obtained by intercoder reliability, digital walk-alongs separate both roles from each other, mitigating priming biases and context switches. In addition, players do not have to be aware of the aspects observed at all, strengthening results in an experiment-like situation.

⁴¹ Cf. Unterhuber 2015.

⁴² Cf. Unterhuber 2015.

⁴³ Cf. Domsch 2019.

⁴⁴ This is implemented, for example, through portals, cf. Günzel 2019.

⁴⁵ Cf. Moretti 2013.

⁴⁶ Cf. Inderst 2020.

⁴⁷ Cf. Tanenbaum / Bizzocchi 2009.

⁴⁸ Cf. Degen / Rose 2012.



Fig. 1: Key question analysing visual material. [Rose 2016, p. 25]

The soundest approach for visual analysis in geography relates to the work of Gillian Rose.⁴⁹ Her code matrix intersects three types of modalities in the production process of visual artifacts with four sites of interpretation.

Modalities are:

- technological: by which means was the visual artifact produced?
- compositional: how was the visual artifact arranged
- social: what is the anticipated social context of the visual artifact?

Sites of interpretation are:

- site of production: how was the visual artifact influenced surroundings?
- site of image itself: how was content of the visual artifact selected?
- site of circulation: how was the visual artifact made accessible?
- site of audience: what are the anticipated surroundings, in which a visual artifact is consumed?

For each cell of the matrix, Rose proposes a number of questions that lead to an interpretation of different aspects of analysis covering content analysis, semiology and discourse analysis.⁵⁰

CONTAINER	BALANCE	COMPULSION
BLOCKAGE	COUNTERFORCE	RESTRAINT REMOVAL
ENABLEMENT	ATTRACTION	MASS-COUNT
PATH	LINK	CENTER-PERIPHERY
CYCLE	NEAR-FAR	SCALE
PART-WHOLE	MERGING	SPLITTING
FULL-EMPTY	MATCHING	SUPERIMPOSITION
ITERATION	CONTACT	PROCESS
SURFACE	OBJECT	COLLECTION

Fig. 2: Selective list of image schemata. [Johnson 1987, p. 126]

⁴⁹ Cf. Rose 2016, p. 25; Figure 1.

⁵⁰ Cf. Rose 2016.

A more formal and thus more standardized analysis of spatial relatable visual content can be done recurring on Johnson's concept of image schemata.⁵¹ Partially relying on Lynch, image schemata can be understood as a general purpose ontology identifying structural patterns of common-sense geographic knowledge both in language and views.⁵² In computer science, image schemata have been used to formalize wayfinding tasks, e. g. at airports.⁵³ The cooccurrence of the schemata PATH, LINK and SURFACE, for instance, was used to model a movement towards a target.

In our explorative study, we use a reduced tagset of 7 schemata useful for the exploration of games. We identified:

- CONTAINER: physically limited areas that provide a homogenous field of affordance.
- ENABLEMENT: an affordance to interact or react
- BLOCKAGE: a game element constraining players' actions. A special case in gameplay is, if a visualized object does not offer an anticipated opportunity for interaction or a map boundary is reached.
- PATH: directed locomotion towards a target
- ATTRACTION: A visual quality that provides a motivation for (inter-)action, i. e. a visual clue
- COUNTERFORCE: An object, agent or event that prevents a player from acting as desired
- LINK: A perceived junction between different game elements

3. Study setup

With respect to the background knowledge laid out above, we designed a study setup suitable to test our main hypotheses:

- Skill and preferred type of gameplay affect the way players spatially enact a video game.
- Enactivism: Not all observed types of spatial enactment can be explained only by interaction with game induced affordances.

3.1 Sample generation

As the preferred type of spatial gameplay cannot be easily predicted and are part of the research question in this paper, sampling was done by skill level. Skill level was explicitly assessed by possible participants and was verified during the sessions. We looked for suitable candidates on the following skill levels:

- *experienced gamers* have played and finished a large variety of video games and play regularly.

⁵¹ Cf. Johnson 1987; Figure 2 for an overview.

⁵² Cf. Lynch 1960.

⁵³ Cf. Raubal et al. 1997.

- *casual gamers* are curious to play video games in general, but only play on special occasions or do not continue playing after a first glance.
- *no-gamers* are aware of video games as means of playing, but refuse to do so for several reasons.

We used a gatekeeper approach on the authors' social network to cast possible candidates. As we intended to conduct an explorative and qualitative study, we limited the number of participants to simply covering the skill levels mentioned above with at least one example. Perhaps due to the popularity of streaming events in the community, we found much more experienced gamers than needed and were able to conduct two spare game sessions for that group. For casual and no-gamers we were able to conduct exactly one game session.

Table 1 provides an overview of the participants with respect to self-reported player type, gender, and age. As we focus on the phase of early spatial accommodation in the game, the minutes walked-along are the first minutes the specific player spent on playing the example game. As our screencast approach did not work out properly on all hardware, the number of minutes captured deviates from the minutes-along. Minutes analysed cover phases of intense interaction with landscape cut from the minutes captured.

player type and ID	gender, age	technical limitations	minutes walked-along	minutes captured	minutes analysed
b: casual gamer	m, 22	game crashed after 16 minutes	45	16	10
i: no-gamer	f, 25		38	38	18
k: experienced gamer	m, 22	video corrupted	45	0	0
n: experienced gamer	m, 23		42	42	10
u: experienced gamer	m, 24	insufficient frame rate	45	45	0

Tab. 1: Participants and amount of analysed screencast. [Kremer et al. 2022]

Our aim was to have at least one recording for each player type. Due to the aforementioned technical problems, we had to conduct three playing sessions with experienced gamers to fine tune our setup and to eventually obtain usable material.

3.2 Technical preparations

All digital walk-alongs were conducted as online sessions. This allowed us to react dynamically to the participants' schedules in times of the COVID-19 pandemic. We checked in advance that existing hardware met the requirements of the game plus sufficient resources to run a video recording in the background. To guarantee an equivalent setup, we first asked the

participants to join a **discord** server instance established for this purpose that was used for the screencast of this experiment and installed and preconfigured **OBS Studio** for redundant on-site recording in higher quality. Video connection was deactivated to save bandwidth during the screencast and resources for parallel recording. We guided the participants through the installation process of our example game. Depending on network connectivity and load the technical setup took in between 30 and 45 minutes.

3.3 Lab setup for experiment

After the first start of the game, we guided the participants through the configuration process of the game to make sure the same preconditions applied. All participants were asked to choose the survival mode of the game which focuses on surviving in a snowy landscape and doesn't offer a rich narrative. All participants were told to choose ›Pleasant Valley‹ as the starting region. As the starting location is generated by a random seed in the game, all participants started in different locations, but with comparable visual access to landscape and infrastructure. Figure 3 shows a fan-made map of the game region. The easiest level of difficulty, ›Pilgrim‹, was advised. Only the player can choose to pick another gender than his own. In the end of the configuration process, the game's main mission was shown to the user.⁵⁴ All close playing sessions were held in German, the mother tongue of all participants.



Fig. 3: Fan made map of ›Pleasant Valley‹. [XHead / stmSantana 2019]

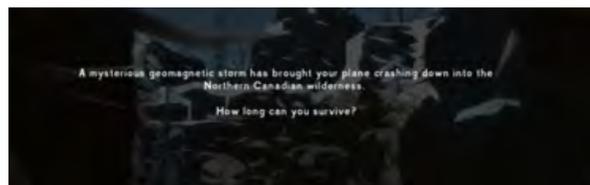


Fig. 4: Main quest of The Long Dark. [Kremer et al. 2022]

⁵⁴ Cf. Figure 4.

After the game launched, we gave the participants a short introduction into the game controls and answered their questions. After that we gave them the main instruction: »Now play the game the way you explore new games. Speak to it!«⁵⁵ No further information on the target of the study was given. While playing, we only intervened to keep the participants speaking and to ask for their mindset on interesting situations. All players reacted instantly to the loud background music especially with respect to the communication via Discord and reduced the master volume of the game. If we observed problems with game control within the running gameplay, we offered assistance only concerning the underlying interaction metaphor of the game and not for the toolbox provided. After around 45 minutes, we concluded the sessions and did a short 5-10 minutes guided interview on the experiences during the game. The main questions were:

- »What did you experience?«
- »How do you evaluate the game mechanics?«
- »Would you keep on playing, if it was your game?«

The first question was meant to reveal insights into active mindsets that might have been not discovered by observation yet. The second question gave us the opportunity to check to what degree the participants were consumed with complying to the game mechanics and to what degree there was time for random exploration or perception. The last question helped us to understand how motivated the participants were during gameplay. The findings largely verified what we had observed so far.

3.4 Example Game

As an example, we use the video game *The Long Dark*, a game with simple game mechanics that is easily accessible even to inexperienced participants. As a cultural artifact in the sense of critical Canadiana wilderness discourses,⁵⁶ the game is located in a northern Canadian wilderness⁵⁷. *The Long Dark*⁵⁸ is ideally suited for the present study because landscape here functions simultaneously as a socioculturally charged construct, as a playing field, and as a game element, which allows for a broad range of connective actions on the part of the players. Moreover, as a so-called independent title,⁵⁹ it is not subject to the mechanisms of the mass market for video games, so that its reception by gamers seems not easily predictable.

Despite the fact that *The Long Dark* does not require violent actions to survive, the game presents itself from a first-person perspective known from shooter games. While this is absolutely reasonable for the sake of impersonation and the game tries even to intensify this impression by both auditive response on even smaller bodily deficiencies and partially blurring the screen, when players are at the edge of falling unconscious by thirst, cold or injuries,

⁵⁵ Cf. Hinterland Studio 2017.

⁵⁶ Cf. Henderson 2001, p. 809.

⁵⁷ Bonner 2018 problematizes the notion of a wilderness in video games that appears to be visually constructed as pristine.

⁵⁸ Initially financed by the Canada Media Fund, subsequently realized via crowd-funding.

⁵⁹ Cf. Martin / Deuze 2009.

this might influence the way participants perceive the game in general. Beside of Figure 4 no other narrative is superimposed on the game, which in sum creates a perfect playground for individual sense-making of the landscape.

Large parts of the game map are covered by snowed woods and fields,⁶⁰ nested with a variety of build environments that provide resources and shelter. Roads and power lines connect the different parts of the map and can be used for navigation. ›Pleasant Valley‹ is surrounded by a scenic mountain range, whose edges can be used for orientation. As our key interest was to identify to what degree landscape was perceived as a scenic setting, as a set of sheer affordances or if there were more individual ways of spacing, The Long Dark provides all of these elements. Consistent with the game mechanics, the game even offers the experience of different weather conditions, daylight and night. Every action taken consumes a certain number of calories and even phases of sleeping come at the cost of calories. As opportunities for interaction exist both in the build and non-build environment in The Long Dark, landscape is usable as a field of affordances and not only evocating mood or atmosphere.

4. Results

All results were either obtained from our notes during the digital walk-alongs or the guided interviews. Information from notes was then related back to the video recordings we obtained from the participants after the sessions. Videos were used both for verification of observations and quotes. In the temporal context of these observations, videos were probed for other examples of interaction with landscape.

We focussed on the experience of landscape in the narrow sense and thus excluded a number of elements from our study.

- bodily environment: the systemic importance of the human body for starving and freezing in a survival game was reflected only by actions on the landscape
- built environment: only outside views of built environment were considered part of the landscape

Assigning image schemata, we further followed a prototypical approach identifying screenshots that reflected most the applicability of a scheme. We did not look deeper into the lifecycle of an active scheme from its early beginnings to its resolution.

Of course, our annotations reflect our experiences as players as well. We were able to reflect the following skill levels useful to understand our participants:

- very skilled player, who has played The Long Dark for some hours
- casual player with little experience in first-person perspective games, who has played The Long Dark only for study preparation purposes.

⁶⁰ Cf. Figure 5.

After having identified the prototypical key frames of image schemata we used the web tool **Recogito** to do explorative annotations of the scenes.⁶¹ Beside its primary use case to annotate maps, it can be used to analyse any 2D-graphics as well.



Fig. 5: Annotated key frame with possible anchor regions wood, bridge, road, snowed meadow, building and mountain range. Feature categories can be annotated in the data scheme of Recogito, but not visualized as labels. Blurred edges in the field of perception reflect that the avatar is hurt at the moment (bottom right symbol). The four icons on the bottom left show the state of the four main elements of the game mechanics: cold, tiredness, thirst and hunger. [Kremer et al. 2022]

In essence, our coding did a three-stage annotation of our key observations:

- Identify the key frame representing the scene of interaction
- Identify the possible anchor regions in each scene. For each anchor region we noted the scale as well loosely following Daniel Montello:⁶² S: nearby or haptic space; M: vista space; L: visible, but not instantly reachable space
- Identify the image schemata active at a specific anchor region according to the behaviour or self-attributed mindset of the participant.

In the sense of an explorative study, we now discuss the edge cases found in the individual spacings of landscape in detail.

4.1 Participant i

Due to limited gaming skills, participant *i* was not able to perceive landscape any other than a game setting. Interaction with the environment in general was a testing behaviour of real world induced hypotheses (Can I burn this? Can I light up a fire there? Will the ice break?). Navigation reacted strongly on salient features in the scene. During the whole session, participant *i* was completely challenged by complying to the game mechanics and showed no other types of performing landscape.

⁶¹ Cf. Figure 5.

⁶² Cf. Montello 1993.

Situation 1: Wayfinding	
active schemata	path, attraction
anchor region	M: bridge
	
Fig. 6: Bridge far back in the forest. [Kremer et al. 2022]	Fig. 7: Bridge appears behind an edge of rock. [Kremer et al. 2022]

Tab. 2: Situation 1: Wayfinding. [Kremer et al. 2022]

In this context, wayfinding tasks were observed. When looking for signs of civilization, elements far in the background were discovered and followed continuously. On first discovery as well as on arrival the element was named:

»Da kommt ne Brücke.« – »There’s a bridge back there.«⁶³

»Da ist die Brücke wieder.« – »There’s that bridge again.«⁶⁴

Situation 2: Ignored affordance	
active schemata	attraction
inactive schemata	path, enablement
anchor region	M: power lines, road, bridge

Tab. 3: Situation 2: Ignored affordance. [Kremer et al. 2022]

⁶³ Participant *i*, close playing session on 2021-3-26, originally stated in German, authors’ translation.

⁶⁴ Participant *i*, close playing session on 2021-3-26, originally stated in German, authors’ translation.

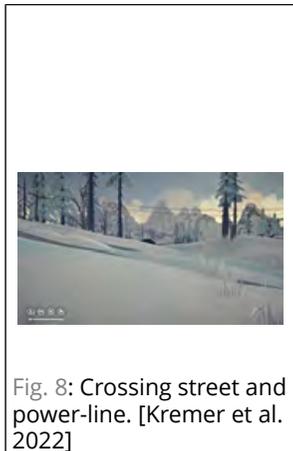


Fig. 8: Crossing street and power-line. [Kremer et al. 2022]

Tab. 3: Situation 2: Ignored affordance. [Kremer et al. 2022]

While following this route, any other affordances were ignored, including roads and power lines, which could have led to infrastructure elements much quicker.

Situation 3: Realism	
active schemata	blockage
anchor region	M: frozen lake
<p>Fig. 9: Frozen lake. [Kremer et al. 2022]</p>	

Tab. 4: Situation 3: Realism. [Kremer et al. 2022]

Shortly after having crossed the road and reaching a frozen lake, this was found to be some blockage:

»Mal gucken, ob ich da einbrech.« – »Let's see if I can break in there.«⁶⁵

⁶⁵ Participant *i*, close playing session on 2021-3-26, originally stated in German, authors' translation.

4.2 Participant b

Participant *b* showed a very strong preference for aesthetic qualities right from the beginning. He reacted to attractions of any kind, even if it was not clear if it would endanger him with respect to the game mechanics or cost at least time for backtracking. He also showed a strong collection behaviour which led to an overweight backpack before end of the short explorative session.

Situation 4: Landscape as attraction	
active schemata	attraction
anchor region	M: mountain range
	
Fig. 10: Mountain view. [Kremer et al. 2022]	Fig. 11: Repeated Mountain view. [Kremer et al. 2022]

Tab. 5: Situation 4: Landscape as attraction. [Kremer et al. 2022]

In the very first seconds of the game, the mountain range in the background gained his attention making up his plans:

»Natürlich, mich ziehen die Berge da hinten an.« – »Of course, I'm attracted to the mountains back there.«⁶⁶

Later on, he affirmed his intent to go to the mountains, and even when elements of civilization offering better shelter were already discovered, he only used them for preparing:

»Ich weiß natürlich nicht, ob das überlebenstechnisch sinnvoll wäre, aber ich würde hier in den Wald hinten reingehen, also die Berge, aber erst mal schau ich, was es hier so gibt.« – »I don't know if that would make sense survival-wise, of course, but I would go in the back of the woods here, so the mountains, but first I'll see what's around.«⁶⁷

⁶⁶ Participant *b*, close playing session on 2021-3-24, originally stated in German, authors' translation.

⁶⁷ Participant *b*, close playing session on 2021-3-24, originally stated in German, authors' translation.

Situation 5 / 6: Landscape as atmospheric aura	
active schemata	attraction
anchor region	S: footprint, L: wood
	
Fig. 12: Footprints in the snow. [Kremer et al. 2022]	Fig. 13: Snowy tree trunks. [Kremer et al. 2022]

Tab. 6: Situation 5 / 6: Landscape as atmospheric aura. [Kremer et al. 2022]

He paid attention to details of the game visualization itself, taking notice of footprints as special effects and showed signs of irritation to visualizations not consistent with real-world experiences. In detail, he complained about the tree trunks being snow covered from all directions which is in conflict with a major wind heading during a precipitation event.

Situation 7: Atmospheric indoor scenes	
active schemata	attraction, enablement
anchor region	S: grill
	
Fig. 14: Grill fire. [Kremer et al. 2022]	

Tab. 7: Situation 7: Atmospheric indoor scenes. [Kremer et al. 2022]

When he reached a barn building, he discovered a grill inside which imposed the strong affordance to him to light up a fire, evaluating his efforts with:

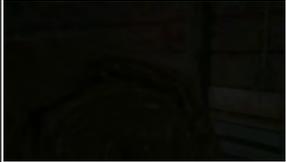
»Romantischer wird's nicht.« – »It doesn't get more romantic than that.«⁶⁸

After he realized:

»Verdammt, ich hab nichts zum Grillen dabei!« – »Damn, I didn't bring anything to grill!«⁶⁹

He looked up his bag and finally discovered that the game allowed him to grill a can of peaches.

Even if it does not precisely fit our understanding of landscape defined above and even if this scene doesn't lack a portion of self-stated irony, it shows how much situational pleasure can be obtained from simply staging explorative game play that does not relate to the core game mechanics at all.

Situation 8: Trap	
active schemata	blockage, container
anchor region	M: farm building
	
Fig. 15: Darkness. [Kremer et al. 2022]	

Tab. 8: Situation 8: Trap. [Kremer et al. 2022]

In the same barn, shortly after the grill session, he realized that his joy of grilling peaches led to a blockage situation, because he had no light source to find his way back to the door:

»Oh verdammt, ich find den Ausgang wahrscheinlich nicht mehr!« – »Oh damn, I probably won't be able to find the exit.«⁷⁰

⁶⁸ Participant *b*, close playing session on 2021-3-24, originally stated in German, authors' translation.

⁶⁹ Participant *b*, close playing session on 2021-3-24, originally stated in German, authors' translation.

⁷⁰ Participant *b*, close playing session on 2021-3-24, originally stated in German, authors' translation.

4.3 Participant n

Being the most experienced one amongst our participants, player *n* acted so controlled and calm that he had the opportunity to do some exploration tasks and even took his time to visualize his mindset by movement for the audience. Especially on that low level of difficulty he never got close to failing. Over time he grew more and more bored, especially unsatisfied with the amount of time it took to cross the landscape between different points of interaction. In his approach, landscape acted as a rich context for decision making and explorative tasks.

Situation 9: Contextualization on the move	
active schemata	path, enablement
anchor region	M: road
	
Fig. 16: Following a road. [Kremer et al. 2022]	Fig. 17: Context view from road. [Kremer et al. 2022]

Tab. 9: Situation 9: Contextualization on the move. [Kremer et al. 2022]

He never followed a path exclusively like in situation 2 (table 3). Situation 9 (table 9) shows an example of quick orientation glances he routinely integrated in the performance of a path. This behaviour gave him the opportunity to always think of further options.

Situation 10: Backtracking behaviour	
active schemata	blockage, path
anchor region	M: power line

Tab. 10: Situation 10: Backtracking behaviour. [Kremer et al. 2022]

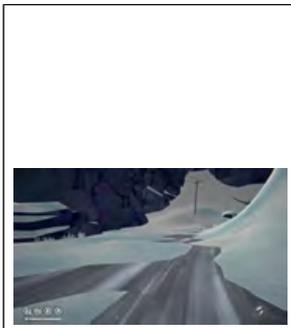


Fig. 18: End of power-line. [Kremer et al. 2022]

Tab. 10: Situation 10: Backtracking behaviour. [Kremer et al. 2022]

When he reached the end of a power line, but not the end of the road yet, he instantly realized that he must have reached a level boundary of ›Pleasant Valley‹, which caused him to backtrack immediately.

Situation 11: Shortcut	
active schemata	blockage, path
anchor region	M: tree trunk
<p>Fig. 19: Path blockage by tree trunk. [Kremer et al. 2022]</p>	

Tab. 11: Situation 11: Shortcut. [Kremer et al. 2022]

Returning from an exploration lap to the starting point, he once more complained about the time it took. After facing a tree trunk as a visual blockage, he started to walk right in the woods, what he hoped to be a shortcut:

»Wo läufst Du jetzt grad hin?« [observer] »Ich hab keine Ahnung!« – »Where are you going right now?« [observer] »I have no idea!«⁷¹

⁷¹ Participant *n*, close playing session on 2021-4-7, originally stated in German, authors' translation.

Interestingly, this decision point marked the beginning of him growing bored. Shortly thereafter, he tested if a steep slope would do harm to his avatar and after eventually experiencing some injuries, he fell out of impersonation and started to talk about his avatar as »you« and not as »I« further on.

Situation 12 / 13: Non-human actors	
active schemata	counterforce
anchor region	S: animal
	
Fig. 20: Wolf back the road. [Kremer et al. 2022]	Fig. 21: Nearby passing deer. [Kremer et al. 2022]

Tab. 12: Situation 12 / 13: Non-Human actors. [Kremer et al. 2022]

Animals are usually distant encounters on the easiest level of difficulty, if the player does not explicitly intend to fight or to hunt. A wolf far down the road did not scare participant *n*. The more he startled, when a deer overtook him in high speed, just to turn around in an erratic movement pattern and pass his other side:

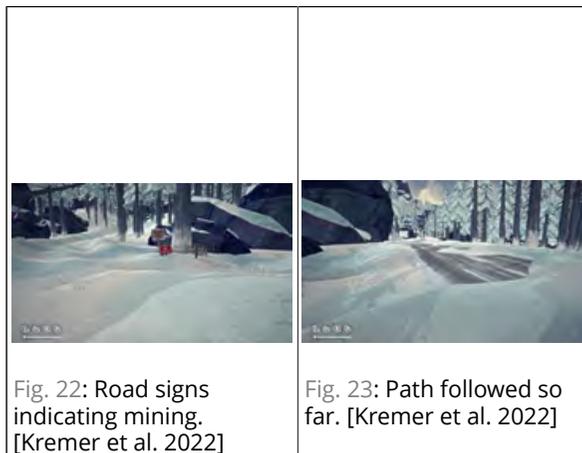
»Jesus! Der Hirsch hat mich erschreckt!« – »Jesus! The deer scared me!«⁷²

In this case it was less due to gameplay then to the unanticipated movement pattern.

Situation 14: Spatial decision making	
active schemata	path, enablement
anchor region	S: road signs, M: road, M: path

Tab. 13: Situation 14: Spatial decision making. [Kremer et al. 2022]

⁷² Participant *n*, close playing session on 2021-4-7, originally stated in German, authors' translation.



Tab. 13: Situation 14: Spatial decision making. [Kremer et al. 2022]

Similar to situation 9 (table 9), he stopped on his path when he reached signs and a path pointing to a mine up in the mountains. He oscillated a few times back and forth between the two views, clarifying his decision making with the words:

»Mine oder Zivilisation?« – »Mine or civilization?«⁷³

Situation 15: Heading calibration	
active schemata	path
anchor region	L: mountain
	
Fig. 24: Mountain view. [Kremer et al. 2022]	

Tab. 14: Situation 15: Heading calibration. [Kremer et al. 2022]

During his way up to the mine, he repeatedly raised his head to get feedback from a nearby mountain on his heading, because the path to the mine was not clearly visible in all places. This marked one of a few key frames that showed the usability of vertical elements of a landscape as landmarks.

⁷³ Participant *n*, close playing session on 2021-4-7, originally stated in German, authors' translation.

Situation 16: Landscape as atmospheric glance	
active schemata	attraction
anchor region	L: waterfall, L: mountain scene
	
Fig. 25: Panorama with waterfall. [Kremer et al. 2022]	

Tab. 15: Situation 16: Landscape as atmospheric glance. [Kremer et al. 2022]

After leaving the mine and before returning to the path downhill he stood for a few seconds perceiving an atmospheric mountain scene with a waterfall. Whereas this must not be overstated, he would not have done so, if nothing had attracted him.

4.4 Observed patterns and similarities

As expected, we were able to identify *landscape* both

1. as means of complying to game mechanics (find resources and survive) and
2. as a rich setting that invites exploration and interaction (grill some peaches).

Nevertheless, amongst the two more experienced players two very individual ways of performing gameplay on a landscape were observed:

- Landscape as aura: participant *b* made the plan to hide in the forest and try to reach the mountain from the very first glance on the mountain scene and regarded all possible shelters as only a possible source of equipment afterwards. He also reacted very strongly to the aesthetic quality of the visualization of footprints and the performative quality to grill food.
- Landscape as context: participant *n*, never happy with the game setting right from the beginning, systematically scanned the landscape for game elements providing the experience of joy and took any opportunity for explorative backtracking behaviour. When the relation of playtime and played time started to bother him and he started to get bored, he focused on the exploration of the game mechanic.

In essence, individual spacing in video games varies not only with gaming skills, but also with individual preferences of enacting landscape-like environments.

Methodically, all types of pre-chosen image schemata can be observed. Non-recurring patterns cover *container* and *counterforce*. Container as a cognitively very productive, hierarchy building pattern⁷⁴ was not observed, as exploration of the valley did not last long enough to either identify different regions of the ›Pleasant Valley‹ or neighbouring top level regions. Counterforces are not observed as the landscape does not present itself as immediately hostile in terms of aggressive animals or avalanches, but starts a silent process of freezing and starvation on a more subtle level. All other image schemata resample a rich variety of opportunities for interaction with the landscape and obstacles blocking them. Attractions can be both an iconic view or an object offering an affordance to act upon (Table 16).

Image schema	Number of scenes (n=16)
Container	1
Enablement	4
Blockage	4
Path	7
Attraction	6
Counterforce	1
Link	1

Tab. 16: Number of image schemata applied (multiple tagging). [Kremer et al. 2022]

A simple similarity measure can be applied. A scene is therefore represented by its vector of all image schemata. Each component is assigned to one of the ordinal values *active*, *inactive* or *non-existent*. Formally:

An image scene is a vector: $s = (\text{Container}, \text{Enablement}, \text{Blockage}, \text{Path}, \text{Attraction}, \text{Counterforce}, \text{Link})$ with $s \# \{\text{active}, \text{inactive}, \text{non-existent}\}$ ⁷

Similarity is then computed by a simple component distance measure on the defined ordinal scale divided by maximum value of 14 for normalization purposes. The derived index ranges from 0 (maximal similarity) to 1 (maximal dissimilarity). The measure is symmetric.

Taken from the scenes above, Figure 26 shows an example of two different scenes with maximal structural similarity (index: 0), both tagged with ›blockage‹. In contrast, Figure 27 shows two pictures of a road but in completely different tasks and perspectives (index: 0,36).

⁷⁴ Cf. Tversky 1993.

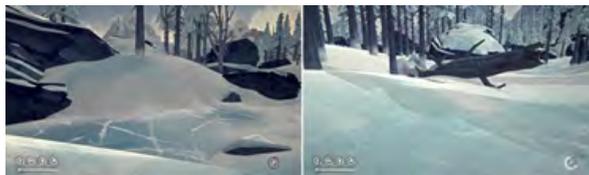


Fig. 26: Example of structural equivalent scene (blockage; participant i vs. participant n). [Kremer et al. 2022]

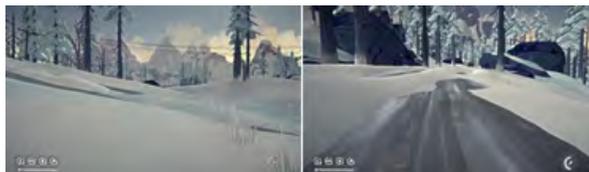


Fig. 27: Example of structural different scene (functions of road and power-line, participant i vs. participant n). [Kremer et al. 2022]

5. Discussion

All pre-identified image schemata are useful to provide information on individual spacings of landscape in our close playing walk-along sessions. Due to the exclusion of the systemic environment of the avatar body and the exclusion of build environment, tasks of landmark driven navigation (paths) are most productive. Because the experiment was laid out to observe early appropriation of a game unknown before to the players, a lot of explorative behaviour including reaction on attractions as affordances⁷⁵ or back-tracking⁷⁶ was taking place. From our limited sample, we can raise the hypothesis that skilled players are more likely to show curiosity or enjoying familiarization of a new game, as they have the capabilities to adapt quickly to game control and are trained to anticipate the core principles of any game mechanic. Observations made during the digital walkalongs reveal that the comparison to other known games are made on early appropriation in order to accommodate with the game. Of course, the sheer aesthetics of the visualization of landscape in video games can be observed several times while watching mountain scenes. As *The Long Dark* as a survivable game doesn't offer rich possibilities of interaction outside shelters and puts a focus on a proportional relation of played time and play time, not many other agents were perceived as counterforce.

Although our coding scheme is designed to annotate paronomies of anchor regions or taxonomies of image schemata, we were not able to observe such complex patterns. Our observations indicate that active image schemata are always bound to a set of higher-level intentions rendering other potential image schemata inactive. Only participant *n* performed a continuous active scan for context. Partially due to a lack of recognition capabilities with less experienced players, we know from computer and cognitive science that intentions are a powerful filter mechanism in general.⁷⁷ In detail, we saw:

1. Active intention sets effectively block affordances.⁷⁸

⁷⁵ Cf. Gibson 1982.

⁷⁶ Cf. Kremer et al. 2013, p. 3.

⁷⁷ Cf. Kiefer 2012.

⁷⁸ Cf. Kiefer 2012; situation 1, wayfinding (table 2): »There's something red back there!«.

2. Empty intention sets⁷⁹ react to almost any environmental affordance.⁸⁰
3. Higher level intentions stay active for a long time.⁸¹

In addition, very strong occurrences of the scheme ›attraction‹ like the farm building in scene 8 that reveal to provide major blockage schemata on entering simply by darkness can be classified as a trap.

We showed that similarity of (active) image schemata provides solid means to compare the visual relevance of scenes to different players. Other than close readings of sceneries that superimpose external attribution by the coder on the one hand and intention sets that focus completely on the player's mindset on the other, image schemata provide structural means to identify task oriented active and inactive visualities and thus act as a link between the other two approaches. Of course, similarity measures can be easily fine-grained to include the scale and feature type of anchor regions and tested on their specific effect sizes.

Our study setup proved to provide a solid ground for testing of our hypotheses as we used a widely controlled lab setup, a well-established method of observation (walk-alongs) and a well-understood annotation technique (image schemata). The game itself revealed to be the right choice as a playground for different sense-makings of landscape. Neither a dense narrative nor a common-sense context of fighting was superimposed to the game situation, which encouraged individual intentions. As opportunities for interaction exist both in build and non-build environment in *The Long Dark*, landscape was usable as a field of affordances and not only evocating mood or atmosphere.

Nevertheless, as our study setup was experimental, we can ex post recommend several opportunities for optimization. As the study setup was held in private via screencast, it was not possible to control a number of factors including hardware resources, incidence of light, brightness of monitor on site and effect of possible other persons present. Also, due to the experimental setting, we tried to focus widely on observation and only encouraged ongoing verbalization of perceptions, plans and (inter)action. With the coding scheme derived from the data, it should be easy now to ask in detail for active intentions to distinguish between not recognized game elements and those filtered by task orientation. To keep on not disturbing the participants in their choice making in the game, confronting the participants with the video material captured directly after the playing sessions for cross validation purposes would provide optimal results. This approach would accelerate the check for usable material at the same time. The skill level of the participants now relies on self-attribution and observation. A more structured approach would be to hold a questionnaire in front of the playing session or even on sample generation.

⁷⁹ For information on the computational approach of mobile intention recognition see Cf. Kiefer 2012; due to boredom: cf. Kiefer et al. 2014.

⁸⁰ E. g. situation 11, shortcut (table 11).

⁸¹ This is reoccurring in situation 4, landscape as attraction (table 5) with the wish to visit the mountains.

A variety of future research is fostered by our findings. As image schemata are a structural query ontology suitable for carrying different semantics, it should be easy to transfer the method of digital walk-alongs back to real world situations. Especially interesting is to capture not only video material of these walks, but also to record eye tracking data from the first-person perspective.⁸² This would also allow for verification of relevant features and measure focus times. Using such temporal structured data would also allow for an analysis of the temporal lifecycle of active image schemata, i. e. from their very beginning (discovering) to the end of relevance (leaving behind).

6. Conclusion

In our experimental approach to obtain more detailed information on individual sense-making of the geographic concept of landscape in video games, first, we used a phenomenological framework to look for individual differences in both perception and conception of landscape. We frame landscape as primarily visual accessible, (re-)constructed by socio-cultural assumptions in the very moment of watching. While on the move, landscapes can be understood as a stream of consciousness continuously reshaped and evaluated according to the intentionality and (silent) expectations of the player. Landscapes can be used to create a certain mood or atmosphere but can also offer a rich field of affordances, evocative spaces that hide their opportunities for interaction behind visual clues. Taking advantage of that, narratives in the game do not necessarily have to be told explicitly, but are produced by players on interaction with the game environment.

Second, we extended the established practice of close playing to accompanied digital walk-alongs, separating the roles of player and observer. In doing so, we mitigated priming biases and context switches and were able to strengthen our results in an experiment-like situation. Following the well-understood instrument of image schemata, we developed a simple coding scheme to annotate active visual elements of engagement and interaction in screencast videos obtained from playing sessions. Our expectation was to observe individual ways of sense-making of landscape in video game environments, varying by skill of player and type of enacting. We selected the game *The Long Dark*, a survival game in the Canadian wilderness that does neither superimpose a strong narrative nor a context of fighting used in many other games. This gave us the opportunity to watch different types of gameplay unfold.

In essence, we found that an individual sense-making beyond sheer affordances can only happen with a certain skill level in the familiarization process of games. Our participant not used to playing video games was continuously challenged by the task of managing the game mechanics and simply had no time for individual sense-making. One of our more experienced participants used landscape intensely as context for exploration behaviour, both to learn about evocative spaces and game mechanics. A third participant strongly reacted to the visual quality of landscape as an aura which significantly determined his intentions for the whole session. Making use of our structural annotation scheme of image schemata we proposed a

⁸² Cf. Kiefer et al. 2014.

simple similarity measure to compute the degree of congruence between different styles of interaction in comparable scenes. Interestingly, we were able to observe that intentions can act as an effective filter for visual-sense making, rendering different conceptions of similar views.

Next steps cover the opportunity to use our study setup, which proved to provide solid ground for observation, in real-world environments. Our current approach can be further enhanced by focusing more on the influence of active intentions on visual sense-making of landscapes. Technically, this could be assisted by using eye-tracking to identify specific areas of interest or focus regions, and in doing so, will enable the development of a more thorough understanding of the entire lifecycle of player-landscape interactions.

Bibliographic references

- Espen Aarseth / Solveig Marie Smedstad / Lise Sunnanå: A multidimensional typology of games. In: Proceedings of the Digital Games Research Conference. Level Up, Ed. by Marinka Copier / Joost Raessens. (DiGRA 2003 – »Level Up«, Utrecht, 4.11.–6.11.2003). Utrecht 2003, pp. 49–53. [\[online\]](#) [\[Nachweis im GBV\]](#)
- Espen Aarseth: A Narrative Theory of Games. In: Proceedings of the International Conference on the Foundations of Digital Games (FDG'12: International Conference on the Foundations of Digital Games, Raleigh, NC, 29.05.–01.06.2012). New York, NY 2012, pp. 129–133. [\[Nachweis im GBV\]](#)
- Ludotopia. Spaces, Places and Territories in Computer Games. Ed. by Espen Aarseth / Stephan Günzel. Bielefeld 2019. (= Edition Medienwissenschaft, 63) [\[Nachweis im GBV\]](#)
- Playful Participatory Practices. Theoretical and Methodological Reflections. Ed. by Pablo Abend / Benjamin Beil / Vanessa Ossa. Wiesbaden 2020. (= Perspektiven der Game Studies, 1) [\[Nachweis im GBV\]](#)
- John A. Agnew: Place and politics. The Geographical Mediation of State and Society. Boston, MA 1987. [\[Nachweis im GBV\]](#)
- James Ash / Lesley Anne Gallacher: Cultural Geography and Videogames. In: Geography Compass 5 (2011), i. 6, pp. 351–368. [\[Nachweis im GBV\]](#)
- Henri Bergson: Matière et Mémoire. Essai sur la relation du corps à l'esprit. 82 edition. Paris 1968. [\[Nachweis im GBV\]](#)
- Marc Bonner: Die gekerbte Wildnis. Inszenierungen vermeintlich unberührter Umwelt in digitalen Spielwelten. In: Paidia – Zeitschrift für Computerspielforschung (2018). Article from 28.02.2018. [\[online\]](#) [\[Nachweis im GBV\]](#)
- Tim Cresswell: On the move. Mobility in the modern Western world. New York, NY 2006. [\[Nachweis im GBV\]](#)
- David Crouch: Spacing, Performing, and Becoming. Tangles in the Mundane. In: Environment and Planning A (2003), i. 11, pp. 1945–1960. (= Economy and Space, 35) [\[Nachweis im GBV\]](#)
- Steven Crowell: Normativity and Phenomenology in Husserl and Heidegger. New York, NY 2013. [\[Nachweis im GBV\]](#)
- Monica Montserrat Degen / Gillian Rose: The Sensory Experiencing of Urban Design: The Role of Walking and Perceptual Memory. In: Urban Studies 49 (2012), i. 15, pp. 3271–3287. [\[Nachweis im GBV\]](#)
- Discord landing page. Ed. by Discord. 2022. [\[online\]](#)
- Sebastian Domsch: Space and Narrative in Computer Games. In: Ludotopia. Spaces, Places and Territories in Computer Games. Ed. by Espen Aarseth / Stephan Günzel. Bielefeld 2019, pp. 103–123. (= Edition Medienwissenschaft, 63) [\[Nachweis im GBV\]](#)
- Simon Egenfeldt-Nielsen / Jonas Heide Smith / Susana Pajares Tosca: Understanding Video Games. The Essential Introduction. 3rd edition. New York, NY 2015. [\[Nachweis im GBV\]](#)
- Nicolas Esposito: A Short and Simple Definition of What a Videogame Is. In: Proceedings of DiGRA 2005 Conference. Changing Views – Worlds in Play (DiGRA 2005 – Changing Views: Worlds in Play, Vancouver, 16.–20.6.2005) Vancouver 2005. [\[online\]](#)
- Daniel Martin Feige: Computerspiele. Eine Ästhetik. Berlin 2015. [\[Nachweis im GBV\]](#)
- James Jerome Gibson: Wahrnehmung und Umwelt. Der ökologische Ansatz in der visuellen Wahrnehmung. München 1982. [\[Nachweis im GBV\]](#)
- Stephan Günzel: The Space-Image. Interactivity and Spatiality of Computer Games. In: Conference Proceedings of the Philosophy of Computer Games. Ed. by Stephan Günzel / Michael Liebe / Dieter Mersch. (The philosophy of Computer Games, Potsdam, 08.–10.05.2008) Potsdam 2008, pp. 170–189. URN: [urn:nbn:de:kobv:517-opus-20072](#) [\[Nachweis im GBV\]](#)
- Stephan Günzel: What Do They Represent? Computer Games as Spatial Concepts. In: Ludotopia. Spaces, Places and Territories in Computer Games. Ed. by Espen Aarseth / Stephan Günzel. Bielefeld 2019, pp. 13–40. (= Edition Medienwissenschaft, 63) [\[Nachweis im GBV\]](#)
- Jennifer Henderson: Critical Canadiana. In: American Literary History 13 (2001), i. 4, pp. 789–813. [\[Nachweis im GBV\]](#)
- The Long Dark. Ed. by Hinterland Studio Inc. PC-Version 1.93 from 01.08.2017.
- Johan Huizinga / Hans Nachod / Andreas Flitner: Homo ludens. Vom Ursprung der Kultur im Spiel. Reinbek / Hamburg 1987. [\[Nachweis im GBV\]](#)
- Rudolf Inderst: Digitale Medien und Methoden. Rudolf Inderst über Close Playing. Open-Media-Studies-Blog. In: Open-Media-Studies-Blog der Zeitschrift für Medienwissenschaft. Blogpost from 18. Januar 2020. [\[online\]](#)
- Mark Johnson: The body in the mind: The bodily basis of meaning, imagination, and reason. Chicago, IL 1987. [\[Nachweis im GBV\]](#)
- Peter Kiefer: Mobile intention recognition. New York, NY 2012. [\[Nachweis im GBV\]](#)
- Peter Kiefer / Ioannis Giannopoulos / Dominik Kremer / Christoph Schlieder / Martin Raubal: Starting to Get Bored. An Outdoor Eye Tracking Study of Tourists Exploring a City Panorama. In: Proceedings of the Symposium on Eye Tracking Research and

- Applications (ETRA '14: Eye Tracking Research and Applications, Safety Harbor, FL, 26.03.–28.03.2014) New York, NY 2014, pp. 315–318. [[Nachweis im GBV](#)]
- Dominik Kremer / Christoph Schlieder / Barbara Feulner / Ulrike Ohl: Spatial Choices in an Educational Geogame. In: 5th International Conference on Games and Virtual Worlds for Serious Applications (VS#Games). Ed. by IEEE. (VS-Games: 5, Bournemouth, 11.–13.09.2013) Piscataway, NJ 2013. [[Nachweis im GBV](#)]
- Game research methods: an overview. Ed. by Petri Lankoski / Staffan Björk. Pittsburgh, PA 2015. [[Nachweis im GBV](#)]
- Martina Löw: Raumsoziologie. 9. Auflage. Frankfurt / Main 2017. [[Nachweis im GBV](#)]
- Kevin Lynch: The image of the city. Cambridge, MA 1960. [[Nachweis im GBV](#)]
- Chase Bowen Martin / Mark Deuze: The Independent Production of Culture: A Digital Games Case Study. In: Games and Culture 4 (2009), i. 3, pp. 276–295. [[Nachweis im GBV](#)]
- Daniel R. Montello: Scale and multiple psychologies of space. In: Spatial Information Theory. A Theoretical Basis for GIS. Ed. by Andrew U. Frank / Irene Campari. (European Conference COSIT'93, Marciana Marina, 19.–22.09.1993) Berlin et al. 1993, pp. 312–321. [[Nachweis im GBV](#)]
- Franco Moretti: Distant reading. 4th edition. London et al. 2013. [[Nachweis im GBV](#)]
- OBS Studio. Open Broadcaster Software. Ed. by OBS Project. 2022. [[online](#)]
- Recogito. Semantische Annotation ohne die spitzen Klammern. Ed. by Pelagios Network. Vienna 2022. [[online](#)]
- Martin Raubal / Max J. Egenhofer / Dieter Pfoser / Nectaria Tryfona: Structuring space with image schemata. Wayfinding in airports as a case study. In: Spatial Information Theory. A Theoretical Basis for GIS. Ed. by Stephen C. Hirtle / Andrew U. Frank. (International Conference COSIT '97, Laurel Highlands, PA, 15.10.–18.10.1997) Berlin et al. 1997, pp. 85–102. [[Nachweis im GBV](#)]
- Edward Relph: Place and placelessness. London 1976. (= Research in planning and design, 1) [[Nachweis im GBV](#)]
- Gillian Rose: Feminism and geography. The limits of geographical knowledge. Minneapolis, MN 1996. [[Nachweis im GBV](#)]
- Gillian Rose: Visual methodologies: an introduction to researching with visual materials. 4. edition. Los Angeles, CA et al. 2016. [[Nachweis im GBV](#)]
- Theodore R. Schatzki: The site of the social. A philosophical account of the constitution of social life and change. University Park, PA 2002. [[Nachweis im GBV](#)]
- Alfred Schütz / Thomas Luckmann: Strukturen der Lebenswelt. Konstanz 2003. [[Nachweis im GBV](#)]
- Joshua Glen Tanenbaum / Jim Bizzocchi: Close Reading Oblivion. Character Believability and Intelligent Personalization in Games. In: Loading: The Journal of the Canadian Game Studies Association 3 (2009), i. 4, pp. 1–20. PDF. [[online](#)]
- Yi-Fu Tuan: Space and place. The Perspective of Experience. Minneapolis, MN 1977. [[Nachweis im GBV](#)]
- Barbara Tversky: Cognitive maps, cognitive collages, and spatial mental models. In: Spatial Information Theory. A Theoretical Basis for GIS. Ed. by Andrew U. Frank / Irene Campari. (European Conference COSIT'93, Marciana Marina, 19.9.–22.9.1993) Berlin et al. 1993, pp. 14–24. [[Nachweis im GBV](#)]
- Tobias Unterhuber: »(Don't) go digging around«. Gone Home, das Spiel als Theorie, der Raum als Archiv und die 90er. In: Paidia – Zeitschrift für Computerspielforschung (2015). Article from 06.05.2015. [[online](#)]
- James Patrick Williams / Jonas Heide Smith: The players' realm. Studies on the culture of video games and gaming. Jefferson, NC 2007. [[Nachweis im GBV](#)]
- XHead / stmSantana: Steam Community. The Long Dark. Detailed Region Maps In: steamcommunity.com. Version small fix ver. 1.56 – 1.9x from 29.10.2019. [[online](#)]

List of Figures and Tables with Captions

- Abb. 1: Key question analysing visual material. [Rose 2016, p. 25]
- Abb. 2: Selective list of image schemata. [Johnson 1987, p. 126]
- Tab. 1: Participants and amount of analysed screencast. [Kremer et al. 2022]
- Abb. 3: Fan made map of ›Pleasant Valley«. [XHead / stmSantana 2019]
- Abb. 4: Main quest of The Long Dark. [Kremer et al. 2022]
- Abb. 5: Annotated key frame with possible anchor regions wood, bridge, road, snowed meadow, building and mountain range. Feature categories can be annotated in the data scheme of Recogito, but not visualized as labels. Blurred edges in the field of perception reflect that the avatar is hurt at the moment (bottom right symbol). The four icons on the bottom left show the state of the four main elements of the game mechanics: cold, tiredness, thirst and hunger. [Kremer et al. 2022]
- Tab. 2: Situation 1: Wayfinding. [Kremer et al. 2022]
- Abb. 6: Bridge far back in the forest. [Kremer et al. 2022]
- Abb. 7: Bridge appears behind an edge of rock. [Kremer et al. 2022]
- Tab. 3: Situation 2: Ignored affordance. [Kremer et al. 2022]
- Abb. 8: Crossing street and power-line. [Kremer et al. 2022]
- Tab. 4: Situation 3: Realism. [Kremer et al. 2022]
- Abb. 9: Frozen lake. [Kremer et al. 2022]
- Tab. 5: Situation 4: Landscape as attraction. [Kremer et al. 2022]
- Abb. 10: Mountain view. [Kremer et al. 2022]
- Abb. 11: Repeated Mountain view. [Kremer et al. 2022]
- Tab. 6: Situation 5 / 6: Landscape as atmospheric aura. [Kremer et al. 2022]
- Abb. 12: Footprints in the snow. [Kremer et al. 2022]
- Abb. 13: Snowy tree trunks. [Kremer et al. 2022]
- Tab. 7: Situation 7: Atmospheric indoor scenes. [Kremer et al. 2022]
- Abb. 14: Grill fire. [Kremer et al. 2022]
- Tab. 8: Situation 8: Trap. [Kremer et al. 2022]
- Abb. 15: Darkness. [Kremer et al. 2022]
- Tab. 9: Situation 9: Contextualization on the move. [Kremer et al. 2022]
- Abb. 16: Following a road. [Kremer et al. 2022]
- Abb. 17: Context view from road. [Kremer et al. 2022]
- Tab. 10: Situation 10: Backtracking behaviour. [Kremer et al. 2022]
- Abb. 18: End of power-line. [Kremer et al. 2022]
- Tab. 11: Situation 11: Shortcut. [Kremer et al. 2022]
- Abb. 19: Path blockage by tree trunk. [Kremer et al. 2022]
- Tab. 12: Situation 12 / 13: Non-Human actors. [Kremer et al. 2022]
- Abb. 20: Wolf back the road. [Kremer et al. 2022]
- Abb. 21: Nearby passing deer. [Kremer et al. 2022]
- Tab. 13: Situation 14: Spatial decision making. [Kremer et al. 2022]
- Abb. 22: Road signs indicating mining. [Kremer et al. 2022]
- Abb. 23: Path followed so far. [Kremer et al. 2022]
- Tab. 14: Situation 15: Heading calibration. [Kremer et al. 2022]
- Abb. 24: Mountain view. [Kremer et al. 2022]
- Tab. 15: Situation 16: Landscape as atmospheric glance. [Kremer et al. 2022]
- Abb. 25: Panorama with waterfall. [Kremer et al. 2022]
- Tab. 16: Number of image schemata applied (multiple tagging). [Kremer et al. 2022]
- Abb. 26: Example of structural equivalent scene (blockage; participant *i* vs. participant *n*). [Kremer et al. 2022]
- Abb. 27: Example of structural different scene (functions of road and power-line, participant *i* vs. participant *n*). [Kremer et al. 2022]

Beitrag aus:

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. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

SPARQLing Ogham – Irische Ogham-Steine als Linked Open Data

Autor*in:

Florian Thiery

Kontakt: mail@fthiery.de

Institution: Römisch-Germanisches Zentralmuseum Mainz

GND: [1169955746](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9) ORCID: [0000-0002-3246-3531](https://orcid.org/0000-0002-3246-3531)

DOI des Artikels:

[10.17175/sb005_010](https://doi.org/10.17175/sb005_010)

Nachweis im OPAC der Herzog August Bibliothek:

[1783635401](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0011-9)

Erstveröffentlichung:

22.06.2022

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise: 24.05.2022

GND-Verschlagwortung:

[Konzeptionelle Modellierung](#) | [Linked Data](#) | [Mittelalterliche Archäologie](#) | [Oghamschrift](#) | [Semantic Web](#) |

Zitierweise:

Florian Thiery: SPARQLing Ogham – Irische Ogham-Steine als Linked Open Data. 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/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_010](https://doi.org/10.17175/sb005_010).

Florian Thiery

SPARQLing Ogham – Irische Ogham-Steine als Linked Open Data

Abstracts

Als ›Ogham-Steine‹ werden mit der ›Ogham‹-Schrift beschriftete Monolithen bezeichnet, die in Irland und im westlichen Teil Britanniens zwischen dem 4. und 9. Jahrhundert entstanden sind. Sie sind in verschiedenen analogen Werken und Online-Datenbanken verzeichnet. Eine gemeinsame Auswertung der unterschiedlichen Quellen ist jedoch derzeit nicht möglich. Das Ogi-Ogham Projekt schafft eine auf Basis von CIDOC-CRM semantisch modellierte Datengrundlage als Linked Open Data (LOD) in Wikidata, sowie als selbst-moderierte LOD. Dieses Data Paper beschreibt die Daten-Publikation, welche im Rahmen des Fellow-Projekts Freies Wissen der Wikimedia Foundation erstellt wurde.

›Ogham stones‹ are monoliths inscribed with the ›Ogham‹ script that were placed in Ireland and the western part of Britain between the 4th and 9th centuries. They are published in various analogue works and online databases. However, a joint evaluation of the different sources is currently not possible. The Ogi-Ogham project creates a semantically CIDOC-CRM based modelled data basis as Linked Open Data (LOD) in Wikidata, as well as self-moderated LOD. This data paper describes the first data publication created as part of the Wikimedia Foundation's Open Science Fellows Program.

1. Einleitung

Dieser Artikel ist eine wissenschaftliche Dokumentation (Data Paper) von Geo- und Ogham-Daten (vgl. Kapitel 3); er bezieht sich somit auf die in Zenodo¹ / GitHub² und Wikidata³ publizierten Daten. Die Daten wurden erhoben, um einen gemeinsamen und standardisierten Daten-Hub für irische Ogham-Steine zu schaffen, Verlinkungen untereinander herzustellen und diese semantisch zu beschreiben.

Informationen zu Ogham-Steinen liegen in vielen analogen Quellen sowie teilweise offenen Online-Datenbanken bzw. -Repositorien vor. Da allerdings kein gemeinsamer Standard der Identifizierung (z. B. Steine, Fundorte) besteht, sind diese Daten nicht gemeinsam auswertbar, kurz gesagt, sie sind nicht interoperabel. Hier setzt das **Ogi-Ogham Projekt** (OOP) mit den im Folgenden vorgestellten Daten an. Zunächst wird eine gemeinsame Datenbasis der unterschiedlichen Ogham-Repositorien als CSV-Dateien (für Geodaten GeoJSON) geschaffen, die in viele Softwaretools (Datenbanken, Auswertungstools wie ›QGIS‹ und ›R‹) niedrigschwellig integriert werden können. Zur standardisierten, semantischen Beschreibung der Ogham-Daten eignet sich insbesondere das Resource Description Framework (RDF)⁴ mit seinem Dialekt

¹ Vgl. Thiery 2021a.

² Vgl. Thiery 2021c. Das **GitHub Repository**, wird ständig weiterentwickelt.

³ Vgl. Thiery 2021d. Das WikiProject Irish Ogham Stones sammelt alle Wikidata-Items zum ›Irish Ogham Stones into the Wikimedia Universe (Q100530634)‹ Projekt; **SPARQL-Abfrage**, Wikidata 2021.

⁴ Vgl. Klyne et al. 2014.

›Turtle⁵, da die Daten nativ als Graph-Netzwerk strukturiert sind. Als Konsequenz ergeben sich aus den CSV-Dateien in Verbindung mit einer Ontologie sogenannte Linked Open Data (LOD)⁶ und damit die Anbindung dieser vorher unverknüpften Daten an die [Linked Open Data Cloud](#). In diesem Data Paper werden der archäologische Hintergrund der Ogham-Steine (Kapitel 2), die Beschreibung der Daten (Kapitel 3), die Relevanz der Daten (Kapitel 4) sowie ein Ausblick in die Zukunft der Ogham-Daten (Kapitel 5) thematisiert.

2. Hintergrund

Ogham-Steine sind eine wichtige Quelle für die archaische oder proto-irische Sprache und Gesellschaft. Die im 4.–9. Jahrhundert in Irland und im westlichen Teil Britanniens (Wales, Schottland, Cornwall, Devon und Isle of Man) errichteten Monolithen sind mit der sogenannten Ogham-Schrift beschrieben.⁷ Die Buchstaben des Ogham-Alphabets bestehen aus Strichen, die in unterschiedlichen Längen und Winkeln entlang von (oft natürlichen) Rillen der Steine von links unten nach oben und rechts wieder nach unten geschrieben wurden.⁸ Als Beispiel sei an dieser Stelle der CIIC-Stein Nr. 81 mit seiner Inschrift »C[A]SSITT[A]S MAQI MUCOI CALLITI«⁹ gezeigt (vgl. Abbildung 1).

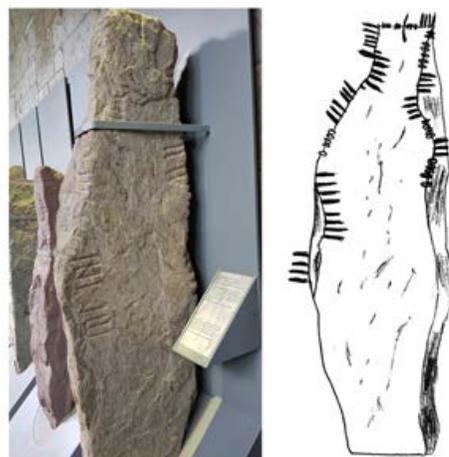


Abb. 1: links: Ogham-Stein CIIC 81 im Stone Corridor des University College Cork (UCC); rechts: Umzeichnung nach Macalister 1945. [beide Thiery 2022]

Als Standardwerk der Ogham-Inschriften gilt das *Corpus Inscriptionum Insularum Celticarum* (CIIC) von Macalister.¹⁰ Darin wird zum einen das heute weit verbreitete Nummerierungsschema vorgestellt und zum anderen zwei Arten von Wörtern beschrieben:

- formular words, die meist persönliche Beziehungen bezeichnen, und

⁵ Vgl. Beckett et al. 2014.

⁶ Vgl. Berners-Lee 2006.

⁷ Vgl. Ferguson 1864; Graves / Limerick 1876.

⁸ Vgl. MacManus 1997.

⁹ Vgl. Macalister 1945, Bd. 1, S. 83.

¹⁰ Vgl. Macalister 1945.

- nomenclature words, die durch irische Personennamen Details der frühgälischen Gesellschaft aufzeigen.

Die Namen werden auf den Steinen im Genus Genitiv angegeben, sodass ein Stein einer Person gewidmet zu sein scheint. Es bleibt jedoch unklar, ob die Steine zum Beispiel Grabmarkierungen waren oder Landbesitz bezeichneten.¹¹

Neben dem CIIC existieren weitere Quellen für das Studium der Ogham-Steine: Eine Reihe von Print-Publikationen wurde vor allem in den 1990er-Jahren als Online-Datenbanken publiziert. Für das OOP¹² und das von Wikimedia Deutschland geförderte **Fellow-Programm Freies Wissen** Projekt ›**Irische Ogham Steine im Wikimedia Universum**‹ sind neben dem CIIC, das Celtic Inscribed Stones Project (CISP)¹³ und das Ogham in 3D Projekt (O3D)¹⁴ besonders wichtige Ressourcen.



Abb. 2: 3D Scan des Ogham in 3D Project von CIIC 180. Emlagh East (IMLEACH DHÚN SÉANN), Co. Kerry. [Thiery 2022]

Das CISP-Projekt hat seine Datenbank im Jahr 2001 als Online-Datenbank bereitgestellt.¹⁵ Das O3D-Projekt bietet verschiedene Typen von offenen Daten zum Download an: eine steigende Anzahl von 3D-Scans von Ogham-Steinen (vgl. **Abbildung 2**) und EPIDOC-Dateien mit kommentierten Steininschriften. Da Ogham-Datensätze nicht dynamisch miteinander verknüpft sind, ist das Zusammenführen der Informationen sehr umständlich. Das OOP zielt darauf ab, Linked Open Data (in Wikidata und eine eigene Ontologie in einem selbst-gehosteten Triplestore) aus den unterschiedlichen Datenbanken zu erstellen.

¹¹ Vgl. MacManus 1997.

¹² Das **Ogi-Ogham Projekt** wurde 2019 vom **Research Squirrel Engineers Netzwerk** gegründet, vgl. Research Squirrel Engineers (Hg.) 2021.

¹³ Vgl. CISP 2001.

¹⁴ Vgl. Ogham in 3D project (Hg.) 2022.

¹⁵ Da die ursprüngliche CISP-Datenbank nicht offen verfügbar ist, war **Kris Lockyear**, UCL (Hg.) 2022, (wissenschaftlicher Mitarbeiter des Projekts) so freundlich, dem Ogi-Ogham Projekt eine Access-Datenbank zur Verfügung zu stellen und die Erlaubnis zu erteilen, diese Daten zu verwenden.

3. Datenbeschreibung

Die in diesem Data Paper beschriebenen Daten wurden in zwei Repositorien bzw. Knowledge Bases publiziert: Zenodo¹⁶ (über GitHub¹⁷) und Wikidata¹⁸. Die Daten sind ›lebendige Daten‹, d. h. sie werden seit Beginn des Projektes sukzessive erstellt bzw. kuratiert und werden sich auch nach dieser Publikation weiterentwickeln. GitHub verfügt über die Möglichkeit Releases zu erzeugen, die bestimmte Versionen zu bestimmten Zeitpunkten darstellen. In Zenodo werden diese Versionen so gespiegelt, dass diese Versionen jeweils mit einem eigenen Digital Object Identifier (DOI) verfügbar sind. In Wikidata ist eine Historien-Übersicht jedes Items vorhanden, sodass eine Änderungsverfolgung stattfinden kann.

Grundsätzlich können die Daten in zwei Kategorien eingeteilt werden: Geodaten (Ordner ›geodata‹, Prefix ›gs‹) und Ogham-Daten (Ordner ›ogham‹, Prefix ›og‹). Für Daten dieser Kategorien sind im Zenodo-Release ›unique lists‹ verfügbar, in denen jedes Datum eine ID enthält. Diese ID wird in der RDF¹⁹-Darstellung in einen Unique Identifier (URI) umgewandelt und stellt somit für das Semantic Web einen eindeutigen Identifikator dar. Darüber hinaus enthält das Zenodo-Release Kreuztabellen (Ordner ›crosstable‹; Prefix ›ct‹), welche durch die Angabe von Anfangs- und End-ID die Verknüpfungen zwischen den Daten darstellen. Zur Sicherstellung der Nachnutzung und Interoperabilität sind von der wissenschaftlichen Community akzeptierte und in zurzeit gängige Software importierbare Datentypen gewählt worden: CSV²⁰ (Ordner ›csv‹), GeoJSON²¹ (Ordner ›geojson‹, Prefix ›gs‹), sowie Turtle-RDF²² (Ordner ›rdf‹). Die GeoJSON-Daten enthalten eine Teilmenge der CSV-Daten, nämlich administrative Grenzen der Republik Irland und Nordirlands sowie Ogham-Fundorte. Die RDF-Daten basieren auf den CSV-Daten und wurden mithilfe von Python-Skripts (Ordner ›rdf/py‹) entsprechend der Web Ontology Language (OWL) basierenden Ontologie (Ordner ›rdf/ontology‹) in semantisch modellierte LOD transformiert und generiert. Die in Wikidata publizierten Daten sind durch ein Mapping Scheme repräsentiert (Ordner ›wikidata‹), welches mit Hilfe der Software [Open Refine](#) erstellt wurde. Die in Wikidata publizierten Daten basieren auf den CSV-Daten des Zenodo-Releases, welche mithilfe von Python-Skripts und der Software [Quick Statements](#) in das [Quick Statements-Format](#) überführt wurden. Die Datenpublikation in Wikidata hat bereits im Jahr 2019 begonnen und entwickelt sich mit Hilfe der Community laufend fort. [Beispiele](#) für Python-Skripts und Quick Statements-Dateien sind im GitHub [Fellow-Projekt Working-Repository](#) zu finden.

Die folgenden Unterkapitel erläutern die Herkunft, Umfang, Repräsentativität und Konsistenz der Daten (Kapitel 3.1), den Prozess der Datenerhebung, -aufbereitung und -anreicherung (Kapitel 3.2), die semantische Datenmodellierung als LOD und in Wikidata (Kapitel 3.3) sowie die Datenpublikation und Zugangsmöglichkeiten.

¹⁶ Vgl. Thiery 2021a.

¹⁷ Vgl. Thiery 2021c.

¹⁸ Vgl. Thiery 2021d.

¹⁹ Vgl. Klyne et al. 2014.

²⁰ Vgl. Shafranovich 2005.

²¹ Vgl. Butler et al. 2016.

²² Vgl. Beckett et al. 2014.

3.1 Herkunft, Umfang, Repräsentativität und Konsistenz

Dieses Kapitel beschreibt den Ursprung der Datenpublikation auf Zenodo und Wikidata. Macalister beschreibt im CIIC Ogham-Inschriften und Steine in Irland²³, in Wales²⁴, in England²⁵, auf der Isle of Man²⁶ und in Schottland²⁷. Die hier publizierten Ogham-Daten konzentrieren sich nur auf Vorkommen auf der irischen Insel, administrativ gesehen die Republik Irland und Nordirland als Teil des Vereinigten Königreiches von Großbritannien und Nordirland. Dementsprechend sind auch nur Geodaten dieses Bereiches veröffentlicht, welche den administrativen Einheiten der beiden Länder entsprechen und den staatlichen Vermessungsstellen (Ordnance Survey, OSi, OSNI): irische Insel, Staat, Provinz, County, Barony und die kleinste Einheit Townland. Die publizierten Ogham-Daten beziehen sich auf zwei große Standardwerke,²⁸ den CIIC²⁹ von R. A.S. Macalister, sowie den Guide to Ogham (GTO)³⁰ von Damian MacManus. Des Weiteren werden repräsentative Online-Datenbanken und -Repositorien als Quellen genutzt: Die Online-Datenbank des CISP³¹, sowie das Online-Repositorium des O3D.³² Innerhalb dieser Daten und Quellen gibt es sich aufeinander beziehende Referenzen, sodass eine Konsistenz gewährleistet ist. Darüber hinaus gibt es weitere Online-Repositorien, die zukünftig in diesen Datensatz integriert werden könnten: z. B. den Thesaurus Indogermanischer Text- und Sprachmaterialien (TITUS), sowie den [Historic Environment Viewer](#), der Einträge des National Monuments Service Sites and Monuments Record (SMR) und des National Inventory of Architectural Heritage beinhaltet. Die folgenden Unterkapitel geben einen Überblick über die Herkunft der Geo- und Ogham-Daten.

3.1.1 Geodaten

Geodaten werden aus drei verschiedenen Quellen bezogen: staatliche Vermessungsstellen der Republik Irland ([Ordnance Survey Ireland](#), OSi) und Nordirlands ([Ordnance Survey of Northern Ireland](#), OSNI), sowie Community-Daten aus [Open Street Map](#) über die Plattform [Irish Townlands](#). Die unterschiedlichen Datengeber stellen ihre Geodaten in verschiedenen Geodaten-Formaten (insbesondere als ESRI Shape-File) und unter verschiedenen Lizenzen zur Verfügung: [CC BY-SA 4.0](#), [Open Government Licence Version 3](#) (OGL v3), bzw. [Open Data Commons Open Database License](#) (ODbL). In Tabelle 1 sind die hier genutzten Layer, mit ihren Lizenzen aufgeführt:

²³ Vgl. Macalister 1945, Bd. 1, S. 1–305, CIIC 1-317.

²⁴ Vgl. Macalister 1945, Bd. 1, S. 306–434, CIIC 318-456.

²⁵ Vgl. Macalister 1945, Bd. 1, S. 435–479, CIIC 457-499.

²⁶ Vgl. Macalister 1945, Bd. 1, S. 479–483, CIIC 501-505.

²⁷ Vgl. Macalister 1945, Bd. 1, S. 484–501, CIIC 506-520.

²⁸ Vgl. Das [Ogham in 3D Projekt](#) gibt eine umfassende Literaturliste an, vgl. Ogham in 3D project (Hg.) 2022.

²⁹ Vgl. Macalister 1945.

³⁰ Vgl. MacManus 1997.

³¹ Vgl. CISP 2001.

³² Vgl. Ogham in 3D project (Hg.) 2022.

Institution	Name	Lizenz
OSi	Provinces – OSi National Statutory Boundaries – Generalised 20m	CC BY-SA 4.0
OSNI	OSNI Open Data – Largescale Boundaries – NI Outline	OOGL v3
OSi	Administrative Areas – OSi National Statutory Boundaries – Generalised 20m	CC BY 4.0
OSNI	OSNI Open Data – Largescale Boundaries – County Boundaries	OOGL v3
OSi	Administrative Areas – OSi National Statutory Boundaries – Generalised 20m	CC BY 4.0
OSM	townlands.ie OSM data	ODbL
OSi	Townlands – OSi National Statutory Boundaries – Generalised 20m	CC BY 4.0
OSNI	OSNI Open Data – 50K Boundaries – Townlands	OOGL v3

Tab. 1: Quell-Layer der Geodaten und Lizenzen. [Thiery 2022]

3.1.2. Ogham-Daten

Ogham-Daten wurden sowohl aus analoger Literatur als auch aus digitalen Quellen bezogen. Die wichtigsten analogen Quellen sind das CIIC³³ und GTO³⁴. Zu den wichtigsten digitalen Quellen (wie Online-Datenbanken, bzw. Repositorien und Access-Datenbanken) zählen die Online-Datenbank (Web-Interface) des CISP³⁵, sowie die Online-Datenbank (Web-Interface) und Repository (EPIDOC-Dateien) des O3D-Projekts³⁶.

³³ Vgl. Macalister 1945.

³⁴ Vgl. MacManus 1997.

³⁵ Vgl. CISP 2001. Darüber hinaus wurde dem Ogi-Ogham Project eine Access-Datenbank, die die Basis der CISP Online-Datenbank bildet von Dr. Kris Lockyear zur Verfügung gestellt, der dem Projekt die Nutzungs- und Publikationserlaubnis gab.

³⁶ Vgl. Ogham in 3D project (Hg.) 2022.

Das CIIC dient vor allem als Grundlage zur Modellierung der CIIC-Steine nach Macalister (vgl. Kapitel 3.2). Als Beispiel für die Art der Informationen in dieser Quelle sei die Beschreibung von CIIC 81 in der Barony of Kinalmeaky gezeigt:

»81.-Garranes (84). [...] On this townland there is an interesting group of earthworks which have recently been examined by excavation. C[A]SSITT[A]S MAQI MUCOI CALLITI. The present stone, which has been known since the sixties, appears to have come from a souterrain in the group; in Windele's time it was standing loosely in the ground beside one or two others, but it must have fallen shortly after that; for long it lay prostrate on one of the ditches, where I first saw it. It is now in the Museum of University College, Cork. An irregular pillar of clayslate, 5' 9" X 1' 7" X 0' 7": the inscription begins at 2' 6" from the bottom of one angle, runs over the top, and ends 3' 0" from the bottom of the adjacent angle. It is made with bold scores, punched and rubbed: a little chipped, so that some of the vowels are lost, but of the reading there can be no doubt. The notches of the I in the first word are accidentally grouped so as to make AUA, but the engraver's intention cannot be questioned. Preceding the initial C there are some scratches, apparently modern, but in any case, of no significance.«³⁷

Im GTO³⁸ werden Ogham-Worte als formula words³⁹ sowie nomenclature words⁴⁰ beschrieben. Bei nomenclature words ist zu beachten, dass

»[...] personal names in the Ogham corpus, almost all of which appear in the genitive case (see §6.25), can be classified broadly into three types: (A) compounded dithematic names (German Vollname), (B) uncompounded monothematic nouns and adjectives with or without suffix (German Kurz- and Kosename) and (C) uncompounded dithematic names of the structure inflected noun + dependent genitive or attribute adjective«⁴¹.

Die hier erfassten Daten für ›nomenclature words‹ beziehen sich vor allem auf MacManus' Kategorien (A) und (B). Beispiele für (A) sind »CUNA- (also CUNO-, CONU-, CVNO-)«⁴² oder »CAT(T)U- (also CATA-, CAT-, CATO)«⁴³, bzw. (B) »[...] 5 ALATTOS, 215 ALATTO (gen. Alta, OI allaid ›wild‹), 90 BAIT (Báeth, báeth ›foolish‹) [...] 277 BIR (Berr, berr ›short-haired‹)«⁴⁴. Prominente Beispiele für formula words sind »[...] MAQ(Q)I (MAQ, MAC, MACI, MACCI etc.) the most frequently occurring word on the inscriptions, is the gen sg. of PI *maqqas [...] the standard Irish word for ›son‹ [...]«⁴⁵ oder »[...] MUCOI (mocu, maccu) is exceptional among formula words in being confined in later Irish [...] to the stereotype sequences X mocu Y and mocu Y [...]«⁴⁶.

³⁷ Vgl. Macalister 1945, Bd. 1, S. 83–84, CIIC 81.

³⁸ Vgl. MacManus 1997.

³⁹ Vgl. MacManus 1997, S. 118–120, §6.27.

⁴⁰ Vgl. MacManus 1997, S. 101–114, §6.2–§6.22.

⁴¹ Vgl. MacManus 1997, S. 101, §6.2.

⁴² Vgl. MacManus 1997, S. 102, §6.3.

⁴³ Vgl. MacManus 1997, S. 102, §6.3.

⁴⁴ Vgl. MacManus 1997, S. 106, §6.11.

⁴⁵ Vgl. MacManus 1997, S. 119, §6.28.

⁴⁶ Vgl. MacManus 1997, S. 119, §6.28.

Die Datenbanken und Repositorien von CISP und O3D umfassen u. a. eine große Anzahl irischer Ogham-Steine mit weiteren Informationen zu Fundorten, Maßen, Inschriften und Transkriptionen verschiedenster Wissenschaftler*innen, aber auch zu auf den Inschriften vermerkten Personen. In der CISP-Datenbank sind diese Informationen im [HTML/XML-Format](#) verfügbar; im O3D-Repositorium sind diese Daten in den [EPIDOC-Dateien](#) codiert. Literaturverweise sind vor allem in der CISP-Datenbank durch Seitenangaben in Macalisters Publikation angegeben (jedoch nicht in der dem Autor derzeit zur Verfügung stehenden Access-Datenbank und nicht auf die Nummerierung nach CIIC), z. B. [GARES/1](#) = Macalister/1945:81.

Die Daten zu Fundorten stellen in allen Quellen eine Herausforderung dar. Das O3D-Projekt gibt, wenn möglich, GPS-Koordinaten zu ›found‹, ›original‹ bzw. ›last recorded‹ an, z. B. [CIIC 12](#). Die CISP-Datenbank bietet in der Online-Datenbank eine ›Irish Grid Reference‹ an, z. B. [Garranes](#) (W 472 629), diese ist jedoch manchmal unvollständig, z. B. [Glanmore](#) (Q 661 52), was eine valide Überführung in WGS84⁴⁷ beispielsweise mit dem [Irish Grid Reference Finder](#) erschwert. Auch die Überführung der Fundort-Informationen aus dem CIIC ist eine Herausforderung. Zwar kann davon ausgegangen werden, dass die erwähnten Orte ›rechts neben der Nummer‹⁴⁸ ein Townland darstellen, jedoch ist die Zuordnung auch mit Hilfe der Informationen zu Barony und County oft schwierig. Als erste Quelle zum Abgleich diente das [Irish Townlands](#) Repositorium. Jedoch waren Macalisters Angaben in manchen Fällen nicht vorhanden, unvollständig oder falsch (z. B. Tippfehler, falscher Ortsname). Dazu kommt, dass es zwischen 1995 und 2021 Gebietsreformen gab. Somit konnten einige Townlands nicht (direkt) identifiziert werden, bzw. erst durch Betrachtung und Plausibilitätsprüfung mit Hilfe anderer Repositorien. z. B. [Historic Environment Viewer](#) und [Placenames Database of Ireland](#) (Logainm), eine Annäherung erreicht werden.

3.2 Datenerhebung, Datenaufbereitung und Datenanreicherung

Für die publizierten Daten fanden eine Datenerhebung (Digitalisierung), eine Datenaufbereitung und eine Datenanreicherung statt. Diese Methoden müssen für Geodaten und Ogham-Daten voneinander getrennt betrachtet werden. Als Grundlage für die Bearbeitung diente eine postgres- / PostGIS-Datenbank, in die zunächst die zur Verfügung gestellte Access-Datenbank des CISP-Projekts integriert wurde. Shapefiles der Geodaten (OSi, OSNI, OSM) wurden mittels der Software QGIS ebenfalls transferiert. Alle weiteren Daten konnten über mehrere Arbeitsschritte und Scripting mit Python (vgl. Kapitel 3.2.2) als CSV in die Datenbank überführt werden. Die nachfolgenden Unterkapitel zeigen auf, wie einzelne Daten erhoben, aufbereitet und angereichert wurden.

⁴⁷ WGS 84 (World Geodetic System 1984) ist ein geodätisches Referenzsystem, das als einheitliche Grundlage für Positionsangaben auf der Erde genutzt wird.

⁴⁸ Z. B. die Überschrift »248. – Bawnaglanna«, Macalister 1945, Bd. 1, S. 243.

3.2.1 Geodaten

Grundlage der administrativen Grenzen, wie die Insel Irland, die administrativen Einheiten Republik Irland (RI) bzw. Nordirland (NI), Provinzen, Counties, Baronies, sowie Townlands bilden die in Kapitel 3.1.1 genannten Daten. Diese wurden mit Hilfe der Software QGIS in einen gemeinsamen ›Insel-Layer‹ aggregiert (insbesondere Geodaten der RI und aus NI) und in die PostGIS-Datenbank verschoben. Besonders zu erwähnen ist hierbei, dass die Daten der Insel sowie der Landesgrenzen aus den Geodaten der Provinzen aggregiert wurden. Für die Baronies Nordirlands wurden die Daten aus OSM benutzt, da hierzu keine Daten beim OSNI vorlagen. Zu den Daten der Townlands wurden fehlende Bereiche (die der Stadtbezirke Dublin, Cork, Galway und Waterford) digitalisiert und mit den Townland-Geodaten verschnitten. In der postgres-Datenbank wurden allen Entitäten bis auf die Townlands (da nicht vollständig in Wikidata enthalten) Wikidata-IDs zugewiesen (durch Daten aus Wikidata). Den Townlands wurde mit Hilfe von [Irish Townlands](#) eine OSM-ID sowie ein Logainm-ID zugewiesen, was auch hier eine weitere Interoperabilität gewährleistet.

3.2.2 Ogham-Daten

Für die CIIC basierten Ogham-Daten musste zunächst eine manuelle Datenerhebung (Retrodigitalisierung) für alle Entitäten (Steine, Sites, Inschriften) erfolgen. Durchgeführt wurde dies manuell in Spreadsheets durch Sophie C. Schmidt und Florian Thiery. Diese Daten bilden als CSV-Datei die Basis der PostgreSQL-Datenbank. In PostgreSQL wurden diese Daten verknüpft und als CSV zur Verfügung gestellt (vgl. Kapitel 3.4). Die Erstellung von Koordinaten der einzelnen Fundorte folgte dem Vorgehen aus Kapitel 3.1.2.

Ogham-Daten der CISP-Datenbank standen zu Beginn bereits als postgres-Datenbank zur Verfügung. Diese Datenbank wurde zunächst nach dem `SCRIPTTYP = ogams / ogham / Ogams / Ogham` `SCRIPTTYP = ogams / ogham / Ogams / Ogham` und `COUNTRY = Ireland` `COUNTRY = Ireland` gefiltert. Die resultierenden Steine konnten verknüpft und als CSV zur Verfügung gestellt werden (vgl. Kapitel 3.4). Die zugrundeliegende Access-Datenbank war jedoch nicht vollständig. Einige Informationen aus der Online-Datenbank fehlten⁴⁹ (z. B. GRID-Koordinaten, Referenzen). Diese wurden manuell digitalisiert und in die Datenbank integriert. Die nachfolgende Transformation der GRID-Koordinaten erfolgte nach dem in Kapitel 3.1.2 beschriebenen Workflow.

Daten des O3D-Projekts standen sowohl als EPIDOC-Files als auch als Informationen auf der Website zur Verfügung. Die EPIDOC-Files wurden mit Hilfe eines [Python-Scripts](#) geparsed und das resultierende CSV in die postgres-Datenbank integriert. Zudem wurden weitere

⁴⁹ Hinweis: Das [Handbuch zur CISP Datenbank](#), Lockyear (Hg.) 2000, gibt Daten-Felder an, welche nicht in der zur Verfügung gestellten Access-Datenbank verfügbar sind; zudem zeigt die CISP-Online-Datenbank in manchen Fällen detailliertere Informationen als die Access-Datenbank.

Informationen (z. B. Inschrift in Ogham-Script) manuell von der Website digitalisiert und als CSV in die Datenbank integriert. Nach der Aufbereitung konnten die Daten verknüpft und ebenfalls als CSV zur Verfügung gestellt werden (vgl. Kapitel 3.4).

Auf Basis der zuvor gezeigten Schritte für Ogham-Steine wurde als ›Stein-Daten-Container‹ ein sogenannter Squirrel-Stone⁵⁰ erstellt. Diese Squirrel-Identifikationsnummer für Ogham-Steine integriert Ogham-Steine der CIIC -, CISP - und O3D-Systeme.

Informationen zu Wörtern wurden aus GTO⁵¹ digitalisiert (vgl. Kapitel 3.1.2) und als CSV in die postgres-Datenbank importiert. Personen auf Ogham-Steinen konnten aus den Datensammlungen der CISP- und O3D-Projekte mit Hilfe eines Python-Skripts extrahiert und als CSV in die postgres-Datenbank importiert werden (vgl. Kapitel 3.1.2). Nach einer Datenaufbereitung und -verknüpfung sind diese Daten als CSV verfügbar (vgl. Kapitel 3.4).

In den Daten ist jeweils die Information der Ursprungsquelle hinterlegt. Alle Daten sind mittels Primär- und Fremdschlüssel miteinander verbunden, was zu (semantischen) Verknüpfungen in Kreuztabellen führt, die ebenfalls als CSV verfügbar sind. Die Daten wurden nach erfolgreicher Erhebung und Aufbereitung in Wikidata importiert (vgl. Kapitel 3.3.1). Dieser Schritt diente dazu, dass alle Daten mit Informationen aus Wikidata angereichert werden konnten. Darüber hinaus konnten Ogham-Daten auch mit Informationen der Geodaten (inkl. Wikidata-ID, OSM-ID, Logainm-ID) versehen werden. Durch einen Export in maschinenlesbare, semantisch modellierte Formate (vgl. Kapitel 3.3.1) mit Verknüpfungen zu bestehenden Ressourcen sind die hier publizierten Daten Teil der Linked Open Data Cloud, bzw. des Knowledge Graphs.

3.3 Semantische Datenmodellierung

Die aus den in Kapitel 3.2 dargestellten Schritten erzeugten CSV-Dateien ermöglichen nun eine semantische Modellierung mittels OWL als LOD, bzw. die Integration in den community-basierten ›Wikidata Knowledge Graph‹. Die folgenden Unterkapitel geben einen tieferen Einblick in die semantische Modellierung der Geo- und Ogham-Daten in den jeweiligen Repositorien bzw. Knowledge Bases.

⁵⁰ Als ›Squirrel-Stone‹ wird ein ›Stein-Daten-Container‹ bezeichnet, der eine ›Stein-Daten-Identifikationsnummer‹ besitzt. Er wurde nach der Forscher*innengruppe der ›Research Squirrel Engineers‹ benannt.

⁵¹ Vgl. MacManus 1997.

3.3.1 Linked Open Ogham Data

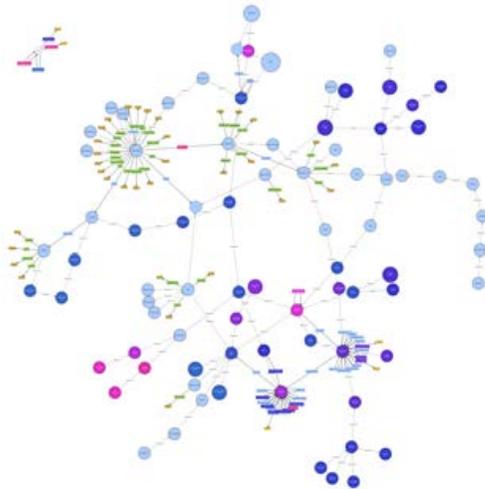


Abb. 3: Linked Ogham Ontology, vgl. Thiery 2021b, visualisiert mit WebVOWL 1.1.4. [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Grundlage für eine semantische Modellierung der CSV-Dateien ist eine Ontologie (Ordner ›ontology‹), eine formal geordnete Darstellung (Regelwerk) einer Menge von Begriffen (Klassen) und der zwischen ihnen bestehenden Beziehungen (Properties) in einem bestimmten Geltungsbereich. Diese wird mit Hilfe der Web Ontology Language⁵² maschinenlesbar dargestellt. Als Software zur Erstellung der Ontologie wird ›Protegé‹ genutzt. Die Ogham-Ontologie⁵³ (vgl. Abbildung 3) basiert unter anderem auf dem in der Archäologie als Standard angesehenen CIDOC Conceptual Reference Model (CIDOC CRM), Simple Knowledge Organization System (SKOS), Friend of a Friend (FOAF), GeoSPARQL und für Provenance Information auf PROV-O. Darunter sind Klassen zur Beschreibung der Geo- und Ogham-Daten-Entitäten eingehängt. Zur Beschreibung der Eigenschaften und Verknüpfungen (Crosstabes) sind zudem ›Object‹ und ›Data Properties‹ vorhanden. Tabelle 2 zeigt die genutzten Klassen sowie deren Bezüge (›Object Properties‹). Zur detaillierten Ansicht kann das OWL-File in Protegé geöffnet werden. Aus der Kombination zwischen Ausgangsdaten (CSV) und Ontologie (OWL) entstehen Linked Open Data (Ordner ›rdf‹).

Klasse (Elternelement)	Prefix	Relation
Word (crm:E55 Type)	OW	Y -[shows]->OW
Ogham Stone (crm:E22 Man-Made-Object)	Y	Y -[hasCollection]->Y
Ogham Site (pleaides:place)	OS	Y -[disclosedAt]->OS
Reading (crm:E25 Man-Made-Feature)	OR	OI -[identifiedAs]->OR

Tab. 2: Klassen, Object-Properties und Prefixes. [Thiery 2022]

⁵² Vgl. Hitzler et al. 2012.

⁵³ Vgl. Thiery 2021b.

Ogham Person (foaf:Person)	OP	Y -[shows]->OP
Geographic Location (pleaides:place)	OL	OS -[hasLocation]->OL
Inscription (crm:E25 Man-Made-Feature)	OI	Y -[carries]->OI
State/ Province / County / Barony / Townland (geosparql:Feature)	GSD	GSD -[within]->GSD

Tab. 2: Klassen, Object-Properties und Prefixes. [Thiery 2022]

3.3.2 Semantik in Wikidata

Der Community-Hub Wikidata folgt einem streng semantischen Prinzip. Wikidata gibt dabei eine Struktur an ›Properties‹ (Verknüpfungsmöglichkeiten) vor. Klassen und Instanzen können eigenständig erzeugt werden. Im Rahmen dieser Datenpublikation wurden die in Kapitel 3.2.1 beschriebenen semantischen Daten in einer verkürzten Version in Wikidata bereitgestellt. Daten in Wikidata unterliegen einer CC0-Lizenz, sodass aufgrund von Lizenzbeschränkungen nicht alle Geo- und Ogham-Daten publizierbar sind. In Wikidata konnten Items zu ›Ogham-Sites‹, ›Ogham-Stein-Concepts‹, sowie ›Ogham-Words‹ hinterlegt werden. Die minimale semantische Modellierung in Wikidata kann im Ordner ›wikidata‹ nachvollzogen werden. Die JSON-Dateien wurden mit Hilfe der Software ›Open Refine‹ erstellt.

3.4 Datenpublikation und Access-Points

Die in den Kapiteln 3.2 und 3.3 beschriebenen Daten können nun auch semantisch angereichert in verschiedenen Formaten zur Verfügung gestellt werden. Hierfür wurden im Rahmen des Projekts folgende zwei Möglichkeiten gewählt:

1. Publikation der Daten als CSV, GeoJSON und RDF über Zenodo / GitHub und
2. Publikation über Wikidata.

Die folgenden Unterkapitel zeigen auf, welche Daten jeweils publiziert wurden und welche Access-Points verfügbar sind.

3.4.1 Zenodo / GitHub

Die Schritte aus Kapitel 3.2 und 3.3 führen zu den in diesem Kapitel beschriebenen Daten, welche auf Zenodo / GitHub veröffentlicht sind. Basis für die GeoJSON- und RDF-Dateien sind immer die gleich benannten CSV-Dateien (vgl.). Die unterschiedlichen Endlizenzen ergeben

sich durch die verschiedenen Lizenzen der Quell-Dateien (vgl. [Tabelle 3](#); vgl. [Tabelle 1](#)). Die RDF-Daten als LOD sind über einen [SPARQL-Endpoint](#) des Mainzer Zentrum für Digitalität in den Geistes- und Kulturwissenschaften ([mainzed](#)) verfügbar.

Dateityp	Kategorie	Dateiname	Lizenz
csv	ogham	og_inscriptions.csv	CC BY 4.0
csv	ogham	og_loactions.csv	CC BY 4.0 / CC BY-NC-SA 3.0 Ireland
csv	ogham	og_persons.csv	CC BY 4.0
csv	ogham	og_reading.csv	CC BY 4.0 / CC BY-NC-SA 3.0 Ireland
csv	ogham	og_sites.csv	CC BY 4.0
csv	ogham	og_stones.csv	CC BY 4.0 / CC BY-NC-SA 3.0 Ireland
csv	ogham	og_stones_references.csv	CC BY 4.0
csv	ogham	og_words.csv	CC BY 4.0
csv	geodata	gs_baronies.csv	CC BY-SA 4.0
csv	geodata	gs_counties.csv	CC BY-SA 4.0 / OOGL v3
csv	geodata	gs_countries.csv	CC BY-SA 4.0 / OOGL v3
csv	geodata	gs_ireland_island.csv	CC BY-SA 4.0 / OOGL v3
csv	geodata	gs_provinces.csv	CC BY-SA 4.0 / OOGL v3
csv	geodata	gs_townlands_centroids.csv	CC BY-SA 4.0 / OOGL v3
csv	crosstables ⁵⁴	ct_itemX_itemY.csv	CC BY 4.0
geojson	---	gs_baronies.geojson	CC BY-SA 4.0
geojson	---	gs_counties.geojson	CC BY-SA 4.0 / OOGL v3
geojson	---	gs_countries.geojson	CC BY-SA 4.0 / OOGL v3
geojson	---	gs_ireland_island.geojson	CC BY-SA 4.0 / OOGL v3
geojson	---	gs_provinces.geojson	CC BY-SA 4.0 / OOGL v3

Tab. 3: Veröffentlichte CSV-, GeoJSON- und RDF-Dateien mit Lizenzangaben. [Thiery 2022]

⁵⁴ Aus dem Modellierungsprozess resultierten eine Vielzahl an Kreuztabellen. Aus Platzgründen sind in dieser Tabelle nicht alle aufgeführt (X / Y als Platzhalter).

geojson	---	gs_sites.geojson	CC BY 4.0
geojson	---	gs_townlands.geojson	CC BY-SA 4.0 / OOGL v3
rdf	ogham	og_inscriptions.ttl	CC BY 4.0
rdf	ogham	og_locations.ttl	CC BY 4.0 / CC BY-NC-SA 3.0 Ireland
rdf	ogham	og_persons.ttl	CC BY 4.0 / CC BY-NC-SA 3.0 Ireland
rdf	ogham	og_reading.ttl	CC BY 4.0 / CC BY-NC-SA 3.0 Ireland
rdf	ogham	og_sites.ttl	CC BY 4.0
rdf	ogham	og_stones.ttl	CC BY 4.0 / CC BY-NC-SA 3.0 Ireland
rdf	ogham	og_stones_references.ttl	CC BY 4.0
rdf	ogham	og_words.ttl	CC BY 4.0
rdf	geodata	gs_baronies.ttl	CC BY-SA 4.0
rdf	geodata	gs_counties.ttl	CC BY-SA 4.0 / OOGL v3
rdf	geodata	gs_countries.ttl	CC BY-SA 4.0 / OOGL v3
rdf	geodata	gs_ireland_island.ttl	CC BY-SA 4.0 / OOGL v3
rdf	geodata	gs_provinces.ttl	CC BY-SA 4.0 / OOGL v3
rdf	geodata	gs_townlands_centroids.ttl	CC BY-SA ⁵⁵ 4.0 / OOGL v3
rdf	crosstables	ct_item1_item2.ttl	CC BY 4.0

Tab. 3: Veröffentlichte CSV-, GeoJSON- und RDF-Dateien mit Lizenzangaben. [Thiery 2022]

⁵⁵ Aus dem Modellierungsprozess resultierten eine Vielzahl an RDF-Dateien mit Centroiden der Townlands. Aus Platzgründen sind in dieser Tabelle nicht alle aufgeführt (X als Platzhalter).

3.4.2 Wikidata



Abb. 4: Wikidata Ogham Sites (Q72617071) des Irische Ogham Steine im Wikimedia Universum Fellow-Projekt (Q100530634). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Die Schritte aus Kapitel 3.2 und 3.3 führen zu in Wikidata veröffentlichen Daten⁵⁶ (vgl. **Tabelle 4** und **Abbildung 4**) die durch die Community bearbeitet werden können. Basis für die Quick Statements-Dateien, welche im Import und zur Modifikation genutzt wurden, sind die aus Kapitel 3.4.1 erzeugten CSV-Dateien. Wikidata ist unter anderem über einen **SPARQL-Endpoint** erreichbar.

Wikidata Query	Items
https://w.wiki/3JCh	Alle Items des Fellow Projekts
https://w.wiki/3JGo	Ogham Sites
https://w.wiki/3JGp	Ogham Stones
https://w.wiki/3JGu	Ogham Words

Tab. 4: Wikidata SPARQL-Abfragen für Ogham Items. [Thiery 2022]

4. Geostatistische Analysen und Potentiale der Daten

Die vorgestellten Daten bieten ein hohes Forschungs- und Nutzungspotential. Das O3D-Projekt zeigt, dass das Thema aktuell wissenschaftlich bearbeitet wird und die Daten somit zu laufenden Forschungsprozessen beitragen können. Durch die Nutzung von Linked Open Data-Konzepten und -Technologien wird eine Interoperabilität zwischen unterschiedlichen Quellen für Ogham-Daten geschaffen. Durch (Retro-)Digitalisierung, Nutzung semantischer Technologien bzw. Modellierungen sowie einheitliche Standards (RDF), Verknüpfung der LOD und Veröffentlichung der CSV- und RDF-Daten, werden analoge Kataloge und nicht maschinenlesbare Ogham-Daten für die DH-Community und aktuelle Ogham-Forschung

⁵⁶ Zusammengefasst im Wikidata:WikiProject **Irish Ogham Stones**, Thiery 2021d.

nutzbar. So können darauf aufbauende statistische⁵⁷ und geostatistische Analysen, z. B. in QGIS mit dem **SPARQLing Unicorn QGIS Plugin**,⁵⁸ bisher rein deskriptive Auswertungen der CIIC- und CISP-Kataloge reproduzierbar machen und den Mehrwert semantischer Modellierung in der DH-Community aufzeigen.⁵⁹



Abb. 5: Kartierung der Squirrel-Steine (Q106602643), erstellt mit QGIS und dem Baselayer Esri.WorldShadedRelief (Tiles © Esri – Source: Esri), Maßstab: 1:5.000.000, EPSG:3857 (Stand: 15.05.2021). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]



Abb. 6: Kartierung der Squirrel-Steine (Q106602643), aggregiert auf die administrative Ebene der Counties, eingefärbt nach der Anzahl der Vorkommen im County, erstellt mit QGIS und dem Baselayer Esri.WorldShadedRelief (Tiles © Esri – Source: Esri), Maßstab: 1:5.000.000, EPSG:3857 (Stand: 15.05.2021). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

⁵⁷ Vgl. Schmidt 2021.

⁵⁸ Vgl. Thiery / Homburg 2021.

⁵⁹ Vgl. Homburg / Thiery 2020.



Abb. 7: Kartierung der Squirrel-Steine (Q106602643), aggregiert auf die administrative Ebene der Baronies, eingefärbt nach der Anzahl der Vorkommen im Barony, erstellt mit QGIS und dem Baselayer Esri.WorldShadedRelief (Tiles © Esri – Source: Esri), Maßstab: 1:5.000.000, EPSG:3857 (Stand: 15.05.2021). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Der Ordner ›example‹ beinhaltet ein QGIS-Projekt, das die im Datensatz enthaltenen Squirrel-Stones umfasst. Abbildung 5 zeigt die räumliche Verteilung der Ogham-Stein-Concepts nach der aggregierten Squirrel-Identifikationsnummer. In Abbildung 6 sind die Squirrel-Stones (graue Punkte) je County (grüne Farbskalierung nach Häufigkeit), in Abbildung 7 je Barony abgebildet. Es ist jedoch zu beobachten, dass die Counties ›Kerry⁶⁰ (153) und ›Cork⁶¹ (92) (Abbildung 6) – genauer die ›Barony Corkaguiny⁶² (67) (Abbildung 7) auf der ›Dingle-Halbinsel – die höchste Dichte an Ogham-Stein-Vorkommen aufweist. Hierbei ist klar zu erkennen, dass eine Aggregation auf verschiedene administrative Einheiten zu sehr unterschiedlichen Darstellungen führt.

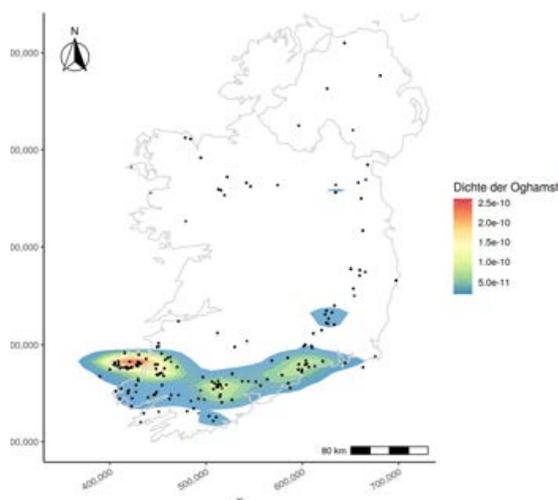


Abb. 8: Dichtekartierung aller Ogham Steine (Ogham Stone Concepts) des Research Squirrel Ogham Projects (Q106602643), erstellt mit ›R‹ (ggplot2, ggspatial) und dem Baselayer von naturalearthdata.com (Stand: 15.05.2021). [Schmidt 2021, CC BY 4.0 via Wikimedia Commons]

⁶⁰ Vgl. Logainm (Ciarraí / Kerry), Government of Ireland / Gaois (Hg.) 2022.

⁶¹ Vgl. Logainm (Corcaigh / Cork), Government of Ireland / Gaois (Hg.) 2022.

⁶² Vgl. Barony of Corkaguiny, Co. Kerry, OpenStreetMap Foundation (Hg.) 2021.

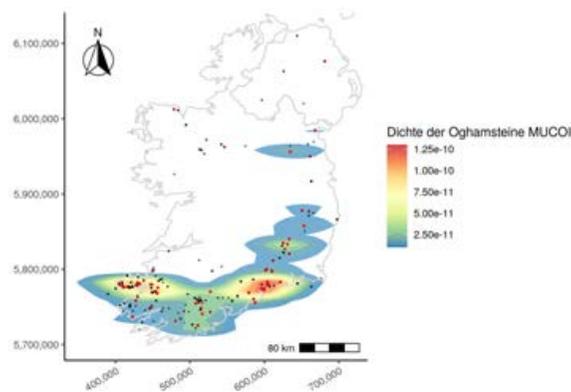


Abb. 9: Dichtekartierung des Wortes MUCOI auf den Ogham-Steinen (Ogham Stone Concepts) des Research Squirrel Ogham Projects (Q106602643), erstellt mit ›R‹ (ggplot2, ggspatial) und dem Baselayer von naturalearthdata.com (Stand: 15.05.2021). [Schmidt 2021, CC BY 4.0 via Wikimedia Commons]

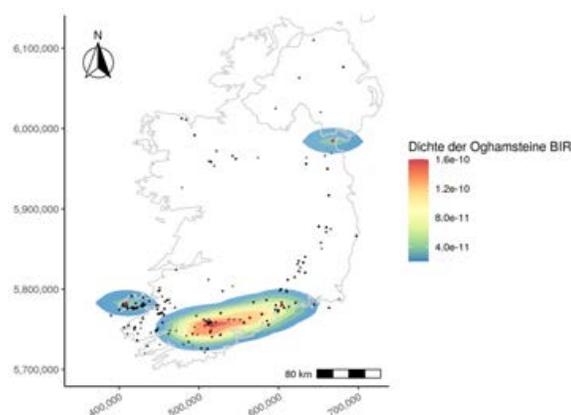


Abb. 10: Dichtekartierung des Wortes BIR auf den Ogham-Steinen (Ogham Stone Concepts) des Research Squirrel Ogham Projects (Q106602643), erstellt mit ›R‹ (ggplot2, ggspatial) und dem Baselayer von naturalearthdata.com (Stand: 15.05.2021). [Schmidt 2021, CC BY 4.0 via Wikimedia Commons]

Gleiches lässt sich auch durch geostatistische Analysen⁶³ mit Hilfe der Software ›R‹⁶⁴ beobachten. Die Dichtekartierung (vgl. Abbildung 8) zeigt deutlich einen Hotspot auf der ›Dingle-Halbinsel‹. Darüber hinaus lassen sich die räumlichen Verteilungen von einzelnen Wörtern auf den Steinen analysieren. Betrachtet man , so ist z. B. ersichtlich, dass das formula word ›MUCOI‹⁶⁵ (túath, tribe) besonders häufig auf der ›Dingle-Halbinsel‹ und im ›County Waterford‹⁶⁶ auftritt, das nomenclature word ›BIR‹⁶⁷ (short haired) jedoch vorrangig im ›County Cork‹ (vgl. Abbildung 10). Während eine Häufung auf der ›Dingle-Halbinsel‹ aufgrund der Gesamtverteilung der Steine nicht verwunderlich ist, lassen andere Verteilungsschwerpunkte eventuell Schlussfolgerungen auf regionale Präferenzen zu.

⁶³ Code und Daten publiziert unter Schmidt 2021.

⁶⁴ R Core Team, genutzt wurden folgende Pakete: tidy (Wickham / Girlich 2021), dplyr (Wickham et al. 2021), magrittr (Bache / Wickham 2020), sp (Pebesma / Bivand 2005; Bivand et al. 2013; r-spatial 2022), rgdal (Bivand et al. 2019), ggplot2 (Wickham 2016), ggspatial (Dunnington / Thorne 2018).

⁶⁵ Vgl. MacManus 1997, S. 119–120, §6.27.

⁶⁶ Vgl. Logainm (Port Láirge / Waterford), Government of Ireland / Gaois (Hg.) 2022.

⁶⁷ Vgl. MacManus 1997, S. 106, §6.11.

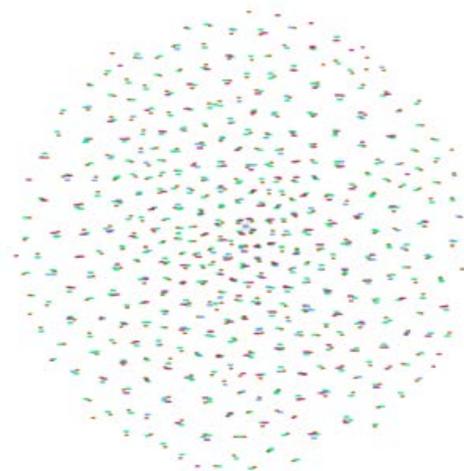


Abb. 11: Ogham-Stein-Verweise in verschiedener Literatur, visualisiert mit vis.js (grün: CIIC / Macalister, braun: CISP, lila: Ogham 3D, blau: Sonstige). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]" >Abb. 11: Ogham-Stein-Verweise in verschiedener Literatur, visualisiert mit vis.js (grün: CIIC / Macalister, braun: CISP, lila: Ogham 3D, blau: Sonstige). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

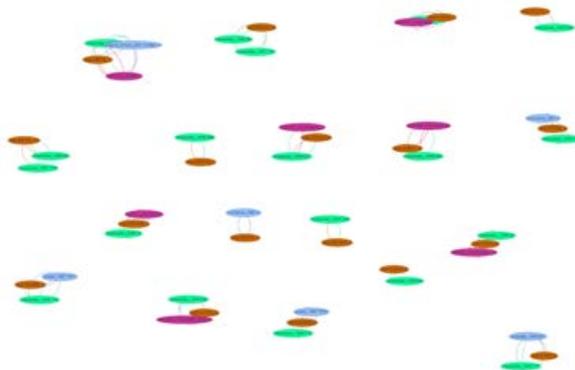


Abb. 12 Ogham-Stein-Verweise in verschiedener Literatur (Detailansicht), visualisiert mit vis.js (grün: CIIC / Macalister, braun: CISP, lila: Ogham 3D, blau: Sonstige). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Die hier beschriebenen Daten zu Ogham-Steinen besitzen Kreuz-Referenzen zu verschiedenen Werken mit anderen Nummerierungsschemata. Die CISP-Datenbank enthält viele dieser Referenzen. Dieses Referenz-Netzwerk wurde mit Informationen aus der postgres-Datenbank ergänzt (>og_stones_references.csv<) und in den Abbildungen 11 und 12 visualisiert.

5. Ausblick

Der publizierte Datensatz bietet bereits einen großen Datenschatz für die Erforschung des frühen Mittelalter Irlands an. Wie jedoch schon erwähnt, gibt es weitere Quellen, z. B. **TITUS**, welche ihn inhaltlich erweitern und validieren könnten. Diese sollten nach Möglichkeit zukünftig in den Datensatz integriert werden. Des Weiteren sind noch nicht alle Eigenschaften aus der CISP-Datenbank in CSV / RDF transformiert. In einer zukünftigen Version des Datensatzes wird sich somit der Korpus vergrößern und damit weitere (statistische) Analysen

ermöglichen. Durch die transparente, interoperable und standardisierte Modellierung mittels RDF und einer Ontologie sowie die Bereitstellung der Daten für die Community in Wikidata bieten sich viele zukünftige Möglichkeiten (vgl. Abbildungen 13).

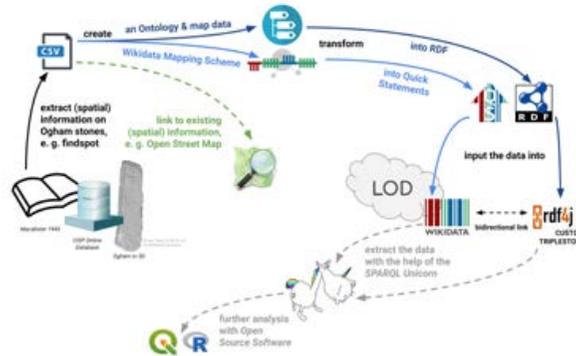


Abb. 13: Der hybride LOD-Ogham-Workflow. [Thiery et al. 2022, CC BY 4.0 via Wikimedia Commons]

In Wikidata können Citizen Scientists und die Wikidata-Community Daten zu Ogham-Steinen verbessern und erweitern. Durch einen bidirektionalen Link zwischen Wikidata und dem eigenen Ogham-Triplestore ist es perspektivisch möglich, aktualisierte Daten aus Wikidata zu übernehmen, um so den eigenen Korpus zu erweitern. Das Ogham-Portal ogham.link⁶⁸ stellt einen LOD-Hub zur Ogham Forschung dar, welche die zuvor beschriebenen Potentiale bündelt.

6. Danksagung

Als Mitglieder des Research-Squirrel-Engineers-Netzwerks haben Sophie C. Schmidt, Timo Homburg und Martina Trognitz die Entwicklung des OOP und dieses Artikels sehr unterstützt. Wir danken auch Dr. Kris Lockyear dafür, dass er dem OOP die CISP-Datenbank zur Verfügung gestellt und die Nutzung der Daten gefördert hat. Ein Dank geht auch an Wikimedia Deutschland, an das Fellow-Projekt Freies Wissen und an die Mentees Charlotte Oertel, Sophie C. Schmidt, sowie an meinen Mentor Jakob Voß, die maßgeblich am Gelingen dieses Projektes beteiligt waren.

⁶⁸ Research Squirrel Engineers (Hg.) 2021.

Bibliographische Angaben

- Roger Bivand: Applied spatial data analysis with R. 2. Auflage. New York, NY 2013. (= Use R!, 10). [\[Nachweis im GBV\]](#)
- Samuel Ferguson: Account of Ogham Inscriptions in the Cave at Rathcroghan, County of Roscommon. In: Proceedings of the Royal Irish Academy 9 (1864–1866), S. 160–170. [\[online\]](#) [\[Nachweis im GBV\]](#)
- Charles Graves / C. Limerick: The ogham alphabet. In: Hermathena 2 (1876), H. 4, S. 443–472. [\[Nachweis im GBV\]](#)
- Robert Alexander Stewart Macalister: Corpus inscriptionum insularum Celticarum. 2 Bde. Dublin 1945–1949. [\[Nachweis im GBV\]](#)
- Damian MacManus: A Guide to Ogam. Maynooth 1997. (= Maynooth monographs, 4) [\[Nachweis im GBV\]](#)
- Edzer Jan Pebesma / Roger Simon Bivand: Classes and methods for spatial data in R. In: R News 5 (2005), H. 2, S. 9–13. PDF. [\[online\]](#)

Onlinequellen

- Stefan Milton Bache / Hadley Wickham: magrittr: A Forward-Pipe Operator for R. Version 2.0.2. In: cran.r-project.org. 26.01.2022. [\[online\]](#)
- David Beckett / Tim Berners-Lee / Eric Prudhommeaux / Gavin Carothers: RDF 1.1. Turtle Terse RDF Triple Language. In: w3.org. W3C Recommendation vom 25.02.2014. [\[online\]](#)
- Tim Berners-Lee: Linked Data – Design Issues. In: w3.org. vom 27.07.2006. [\[online\]](#)
- Applied spatial data analysis with R. Hg. von r-spatial. 2022. HTML. [\[online\]](#)
- Roger Simon Bivand / Tim Keitt / Barry Rowlingson: rgdal: Bindings for the »Geospatial« Data Abstraction Library. Version 1.5-28 vom 15.12.2021. [\[online\]](#)
- Howard Butler / Martin Daly / Allan Doyle / Sean Gillies / Stefan Hagen / Tim Schaub: The GeoJSON Format. RFC 7946. In: datatracker.ietf.org. Internet Standards Track document vom August 2016. [\[online\]](#)
- Celtic Inscribed Stones Project (CISP) online database. Hg. von Celtic Inscribed Stones Project. In: ucl.ac.uk. 2001. [\[online\]](#)
- Logainm.ie: Ciarraí / Kerry. In: logainm.ie. Hg. von Government of Ireland / Gaois. 2022. [\[online\]](#)
- Logainm.ie: Corcaigh / Cork. In: logainm.ie. Hg. von Government of Ireland / Gaois. 2022. [\[online\]](#)
- Logainm.ie: Port Láirge / Waterford. In: logainm.ie. Hg. von Government of Ireland / Gaois. 2022. [\[online\]](#)
- Ogham in 3D. Bibliography. Hg. von Ogham in 3D project. In: ogham.celt.dias.ie. 28.01.2022. [\[online\]](#)
- Dewey Dunnington / Brent Thorne: ggspatial: Spatial Data Framework for ggplot2. In: cran.r-project.org. Version 1.0 vom 12.07.2018. [\[online\]](#)
- Pascal Hitzler / Markus Krötzsch / Bijan Parsia / Peter F. Patel-Schneider / Sebastian Rudolph: OWL 2 Web Ontology Language Primer (Second Edition). In: w3.org. W3C Recommendation vom 11.12.2012. [\[online\]](#)
- Timo Homburg / Florian Thiery: Linked Ogham Stones – Semantische Modellierung und prototypische Analyse irischer Ogham-Inschriften. In: zenodo.org. Version 1 vom 05.03.2020. DOI: [10.5281/zenodo.3697060](#)
- Graham Klyne / Jeremy J. Carroll / Brian McBride: RDF 1.1 Concepts and Abstract Syntax. In: w3.org. W3C Recommendation vom 25.02.2014. [\[online\]](#)
- Kris Lockyear: The Celtic Inscribed Stones Project database. A guide. Überarbeiteter Entwurf vom 18.10.2000. [\[online\]](#)
- Barony of Corkaguiny, Co. Kerry. Hg. von OpenStreetMap Foundation. In: townlands.ie. 15.03.2022. [\[online\]](#)
- The Research Squirrel Engineers Network. Hg. Research Squirrel Engineers. Mainz 2021. In: squirrel.link. [\[online\]](#)
- Ogham.Link. A Linked Open Data Hub for Ogham Research. Hg. von Research Squirrel Engineers. In: ogham.link. 2021. [\[online\]](#)
- The R Project for Statistical Computing. Getting Started. Hg. von The R Foundation. In: r-project.org. Wien 2020. [\[online\]](#)
- Sophie Charlotte Schmidt: ogi-ogham/oghamaps: v0.2. In: zenodo.org. Version 0.2 vom 15.01.2021. DOI: [10.5281/zenodo.4443014](#)
- Yakov Shafranovich: Common Format and MIME Type for Comma-Separated Values (CSV) Files. RFC 4180. In: datatracker.ietf.org. Memo vom Oktober 2005. [\[online\]](#)
- Florian Thiery (2021a): Ogham Data. In: zenodo.org. Version 1.0 vom 16.05.2021. DOI: [10.5281/zenodo.4765604](#)
- Florian Thiery (2021b): Linked Ogham Ontology. In: zenodo.org. Version 1.0.1 vom 16.05.2021. DOI: [10.5281/zenodo.4765568](#)

Florian Thiery (2021c): Ogham Data. In: github.com. Release v1.0 vom 16.05.2021. DOI: [10.17605/OSF.IO/KPQ2C](https://doi.org/10.17605/OSF.IO/KPQ2C) [online]

Florian Thiery (2021d): Wikidata: WikiProject Irish Ogham Stones. In: Wikidata. 20.05.2021. [online]

Florian Thiery: Sparqlunicorn/sparqlunicornGoesGIS. In: Github.com. 10.02.2022. [online]

Florian Thiery / Timo Homburg: SPARQLing Unicorn QGIS Plugin. In: zenodo.org. Version 0.12.1 vom 21.01.2021. DOI: [10.5281/zenodo.6039443](https://doi.org/10.5281/zenodo.6039443)

Kris Lockyear. Hg. von University College London. In: ucl.ac.uk. London 2022. [online]

Hadley Wickham: ggplot2: Elegant Graphics for Data Analysis. In: ggplot2.tidyverse.org. 2016. [online]

Hadley Wickham / Romain François / Lionel Henry / Kirill Müller: dplyr: A Grammar of Data Manipulation. In: cran.r-project.org. Version 1.0.6 vom 05.05.2021. [online]

Hadley Wickham / Maximilian Girlich: tidy: Tidy Messy Data. In: cran.r-project.org. Version 1.1.3. 03.03.2021. [online]

Ogi-Ogham Project. Hg. von Wikidata. In: wikidata.org. 26.04.2021. HTML. [online]

Wikidata Query Service. Hg. von Wikidata. In: query.wikidata.org. 2021. [online]

Abbildungs- und Tabellenverzeichnis

Alle Abbildungen sind unter freien Lizenzen (CC BY(-SA)) auf Wikidata Commons veröffentlicht und auch [hier](#) verfügbar.

Abb. 1: links: Ogham-Stein CIIC 81 im Stone Corridor des University College Cork (UCC); rechts: Umzeichnung nach Macalister 1945. [beide Thiery 2022]

Abb. 2: 3D Scan des Ogham in 3D Project von CIIC 180. Emlagh East (IMLEACH DHÚN SÉANN), Co. Kerry. [Thiery 2022]
Tab. 1: Quell-Layer der Geodaten und Lizenzen. [Thiery 2022]

Abb. 3: Linked Ogham Ontology, vgl. Thiery 2021b, visualisiert mit WebVOWL 1.1.4. [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Tab. 2: Klassen, Object-Properties und Prefixes. [Thiery 2022]

Tab. 3: Veröffentlichte CSV-, GeoJSON- und RDF-Dateien mit Lizenzangaben. [Thiery 2022]

Abb. 4: Wikidata Ogham Sites (Q72617071) des Irische Ogham Steine im Wikimedia Universum Fellow-Projekt (Q100530634). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Tab. 4: Wikidata SPARQL-Abfragen für Ogham Items. [Thiery 2022]

Abb. 5: Kartierung der Squirrel-Steine (Q106602643), erstellt mit QGIS und dem Baselayer Esri.WorldShadedRelief (Tiles © Esri — Source: Esri), Maßstab: 1:5.000.000, EPSG:3857 (Stand: 15.05.2021). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Abb. 6: Kartierung der Squirrel-Steine (Q106602643), aggregiert auf die administrative Ebene der Counties, eingefärbt nach der Anzahl der Vorkommen im County, erstellt mit QGIS und dem Baselayer Esri.WorldShadedRelief (Tiles © Esri — Source: Esri), Maßstab: 1:5.000.000, EPSG:3857 (Stand: 15.05.2021). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Abb. 7: Kartierung der Squirrel-Steine (Q106602643), aggregiert auf die administrative Ebene der Baronies, eingefärbt nach der Anzahl der Vorkommen im Barony, erstellt mit QGIS und dem Baselayer Esri.WorldShadedRelief (Tiles © Esri — Source: Esri), Maßstab: 1:5.000.000, EPSG:3857 (Stand: 15.05.2021). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Abb. 8: Dichtekartierung aller Ogham Steine (Ogham Stone Concepts) des Research Squirrel Ogham Projects (Q106602643), erstellt mit R (ggplot2, ggspatial) und dem Baselayer von naturalearthdata.com (Stand: 15.05.2021). [Schmidt 2021, CC BY 4.0 via Wikimedia Commons]

Abb. 9: Dichtekartierung des Wortes MUCOI auf den Ogham-Steinen (Ogham Stone Concepts) des Research Squirrel Ogham Projects (Q106602643), erstellt mit R (ggplot2, ggspatial) und dem Baselayer von naturalearthdata.com (Stand: 15.05.2021). [Schmidt 2021, CC BY 4.0 via Wikimedia Commons]

Abb. 10: Dichtekartierung des Wortes BIR auf den Ogham-Steinen (Ogham Stone Concepts) des Research Squirrel Ogham Projects (Q106602643), erstellt mit R (ggplot2, ggspatial) und dem Baselayer von naturalearthdata.com (Stand: 15.05.2021). [Schmidt 2021, CC BY 4.0 via Wikimedia Commons]

Abb. 11: Ogham-Stein-Verweise in verschiedener Literatur, visualisiert mit vis.js (grün: CIIC / Macalister, braun: CISP, lila: Ogham 3D, blau: Sonstige). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Abb. 12: Ogham-Stein-Verweise in verschiedener Literatur (Detailansicht), visualisiert mit vis.js (grün: CIIC / Macalister, braun: CISP, lila: Ogham 3D, blau: Sonstige). [Thiery 2022, CC BY 4.0 via Wikimedia Commons]

Abb. 13: Der hybride LOD-Ogham-Workflow. [Thiery et al. 2022, CC BY 4.0 via Wikimedia Commons]

Beitrag aus:

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. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Der MUSE Datensatz

Autor*in:

Johanna Barzen

Kontakt: johanna.barzen@iaas.uni-stuttgart.de

Institution: Universität Stuttgart, Institut für Architektur von Anwendungssystemen

GND: [1261754875](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0001-8397-7973](https://orcid.org/0000-0001-8397-7973)

Autor*in:

Fabian Bühler

Kontakt: fabian.buehler@iaas.uni-stuttgart.de

Institution: Universität Stuttgart, Institut für Architektur von Anwendungssystemen

GND: [1261770226](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0003-2834-4342](https://orcid.org/0000-0003-2834-4342)

Autor*in:

Frank Leymann

Kontakt: frank.leymann@iaas.uni-stuttgart.de

Institution: Universität Stuttgart, Institut für Architektur von Anwendungssystemen

GND: [172224780](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0002-9123-259X](https://orcid.org/0000-0002-9123-259X)

DOI des Artikels:

[10.17175/sb005_012](https://doi.org/10.17175/sb005_012)

Nachweis im OPAC der Herzog August Bibliothek:

[1783635630](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9)

Erstveröffentlichung:

01.09.2022

Lizenz:

Sofern nicht anders angegeben



Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise: 01.07.2022

GND-Verschlagwortung:

[Datenbank](#) | [Filmgenre](#) | [Filmwissenschaft](#) | [Kostüm](#) | [Ontologie](#) |

Zitierweise:

Johanna Barzen, Fabian Bühler, Frank Leymann: Der MUSE Datensatz. 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/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_012](https://doi.org/10.17175/sb005_012).

Johanna Barzen, Fabian Bühler, Frank Leymann

Der MUSE Datensatz

Abstracts

Das Filmkostüm ist ein vergleichsweise wenig untersuchtes diegetisches Gestaltungsmittel. Im Besonderen mangelt es an einem übergreifenden Konzept, Kostüme systematisch beschreibbar und damit, über Einzelfilmanalysen hinaus, analysierbar zu machen. Der MUSE Datensatz will diese Lücke schließen und stellt detaillierte, in eine umfassende Ontologie eingebettete Daten über Filmkostüme zur Verfügung. Die Kostümdateien enthalten sowohl Daten zu den Kostümen selber, zu den einzelnen Kleidungsstücken, den Attributen wie Farbe, Materialien, Zustand und deren Zusammenstellung, wie auch kontextualisierende Daten zu deren Vorkommen im Film.

Film costumes as diegetic design elements are only rudimentarily explored. There is especially a lack of an overarching concept to make costumes describable in a systematic way and thus analysable beyond individual film analyses. The MUSE dataset aims to fill this gap by providing detailed data on film costumes embedded in a comprehensive ontology. The costume data set contains both, data on the pieces of clothes that compose the costume, their arrangement, and their attributes such as colour, materials, and condition, as well as contextualizing data on the occurrence of the costumes in the film and the role wearing the costume.

1. Einleitung

Frei nach Gottfried Kellers Kleider machen Leute¹ unterstützen Kostüme visuell die Schaffung der Rolle von Schauspieler*innen über deren sichtbare Hülle.² Der Grundstein für das geltende Verständnis von filmisch vestimentärer Kommunikation ist die häufig bemühte Metapher von Kleidung als Sprache, die sich bereits in den Titeln vieler Werke zu Bekleidung, Mode oder Kostümen spiegelt.³ Wie allerdings eine solche vestimentäre Sprache über die Metapher hinaus aussehen kann, wird aufgrund der Komplexität der Aussagen, wie auch deren kontextuell und kulturell beeinflusste Interpretation, als schwierig angesehen.⁴ Um sich trotz der gegebenen Schwierigkeiten einer potenziellen Kostümsprache zu nähern, soll hier drei Annahmen gefolgt werden:

1. Dass jedes Kleidungsstück, seine Farben, Materialien, sein Zustand etc. von Kostümbildner*innen⁵ ausgewählt wird, um einen bestimmten Effekt zu erzielen, der

¹ Keller 2017.

² Devoucoux 2007, S. 29f.

³ Vgl. beispielsweise: Die Sprache der Mode, Barthes 1985; The Language of Clothes, Lurie 2000; Kleidersprache, Hoffmann 1985; Clothing as Communication, Holman 1980; Die Sprache der Kleidung, Petrascheck-Heim 1966; Fashion as Communication, Barnard 2002.

⁴ Vgl. Burger 2002, S. 21.

⁵ Dies geschieht unter Einbezug der in den Drehbüchern verfügbaren Informationen und der Berücksichtigung von Absprachen mit den anderen Departments im Filmstehungsprozess, wie beispielsweise der Regie, der Maske oder der Lichttechnik.

beispielsweise die Charakterisierung einer Rolle, deren Beziehung zu anderen Rollen oder zur Umwelt unterstützt.⁶

2. Dass Kostümbildner*innen hierbei auf Konventionen zurückgreifen, die sich etabliert haben, um bestimmte Stereotype, Zeit- und Ortsgegebenheiten etc. über Kostüme und deren Attribute zu kommunizieren.⁷

3. Dass diese etablierten Konventionen und Stilmittel, mittels derer über das Filmkostüm kommuniziert wird und welche Rezipient*innen über ihre Mediensozialisation zu verstehen gelernt haben, in Filmen reflektiert sind.

2. Hintergrund

Genau hier setzt das Projekt **MUSE** (was für ›Muster Suchen und Erkennen‹ steht) an. MUSE hat zum Ziel, das Wissen über diese Konventionen aus Filmen zu extrahieren und die signifikanten Elemente, welche zur Kommunikation beitragen, beispielsweise ein bestimmtes Stereotyp oder eine Charaktereigenschaft, als Kostümmuster darstellbar zu machen.⁸ Das Verständnis von Kostümmustern folgt hier dem von Alexander et al. (1977) geprägten Musterbegriff, nach welchem Muster als bewährte Lösungen von wiederkehrenden Problemen verstanden und dokumentiert werden.⁹ Muster als Dokumentationstechnik haben sich in den unterschiedlichsten Domänen etabliert,¹⁰ werden aber vorrangig in der Informatik angewendet.¹¹ In dem vorliegenden Anwendungsfall stellen Kostümmuster für Kostümbildner*innen bewährte Lösungen zur Verfügung, wie mit Hilfe von Kleidungsstücken, deren Zusammenstellung, Attributen oder Zustand bestimmte Eigenschaften, Stereotype oder Charakterzüge kommuniziert werden. Die Kostümmuster erfassen diese wiederkehrenden Elemente und deren kontextuelle Einbettung in einem für Muster üblichen Format und werden untereinander durch Relationen unterschiedlicher Semantik zu einer Mustersprache verbunden, hier einer Kostümmustersprache.¹²

Um Muster und deren Relationen zu identifizieren, werden in MUSE in großem Umfang Daten zu Filmkostümen erfasst, analysiert und interpretiert. Hierzu soll der qualitative Schritt der Auswertung und Interpretation der Daten mittels quantitativer Methoden unterstützt werden, um der Menge an Informationen habhaft zu werden. Das Vorgehen und die Werkzeugumgebung (siehe Abschnitt 3) von MUSE unterstützt dabei:

1. die Definition der potenziell relevanten Parameter zur umfassenden Beschreibung eines Kostüms,
2. die Auswahl des Untersuchungsgegenstandes in Form eines Filmkorpus,

⁶ Vgl. La Motte 2010, S. 69–73.

⁷ Vgl. Giannone 2005, S. 140–153; Devoucoux 2007, S. 15.

⁸ Vgl. Barzen et al. 2018; Barzen 2018.

⁹ Vgl. Alexander et al. 1977, S. ix–xvii.

¹⁰ Vgl. beispielsweise: Patterns for Pedagogy, Lilly 1996; Pattern Language for Creative Presentations, Iba 2014; A Pattern Language for Film Production, Fießer 2007.

¹¹ Vgl. beispielsweise: Design Patterns, Gamma et al. 1994; Pattern-Oriented Software Architecture, Buschmann et al. 1996; Enterprise Integration Patterns, Hohpe / Woolf 2004; Cloud Computing Patterns, Fehling et al. 2014.

¹² Vgl. Barzen / Leymann 2014 ; für mehr Details: Barzen 2018, S. 88–90.

3. die Erfassung der Daten mittels des MUSE Repositoriums,
4. die auf die Art und Struktur der Daten abgestimmte Analyse der Daten und
5. die Überführung der validierten Ergebnisse in Kostümmuster.¹³

Insbesondere die Schritte eins bis drei sollen im Folgenden ausgeführt werden.

3. Beschreibung der Daten

Seit dem Start des MUSE Projekts 2013 wurden und werden kontinuierlich Daten zu Filmkostümen erfasst, die den MUSE Datensatz bilden. Der erste Schritt hin zu einer systematischen Erfassung ist die Definition der Domäne, welche die grundlegende Struktur vorgibt, in der die Daten erfasst sind.

3.1 Die zugrundeliegende Struktur

Um die potenziell relevanten Parameter zur Beschreibung eines Kostüms zu definieren, wurde eine umfassende Ontologie aufgebaut.¹⁴ Der Begriff der Ontologie wird hier, anders als in der Philosophie,¹⁵ als Werkzeug der Wissensrepräsentation verstanden, welches klare Begriffsdefinitionen durch deren Merkmale sowie deren Einbettung in ein Ordnungsschema erlaubt.¹⁶ Als Basis der Ontologie systematisieren und kategorisieren einzelne Taxonomien in hierarchischer Struktur und einem hohen Detailgrad diejenigen Parameter, die potenziell Einfluss auf die Wirkung eines Kostümes haben können.¹⁷ Hier lassen sich Taxonomien der Basiselemente (Bezeichnungen eines Kleidungsstückes, wie ›Lange Hose‹, ›Reithose‹ oder ›Sporthose‹), Teilelemente (Bezeichnungen der kompositen Elemente von Kleidungsstücken wie ›Haifischkragen‹, ›Fledermausärmel‹, oder ›Reisverschluss‹), Design, Form, Zustand, Material, Materialeindruck (zur Spezifizierung der Materialien durch Eindrücke wie ›steif‹, ›fließend‹ oder ›flauschig‹), Farbe, Farbeindruck (zur Spezifizierung der Farben durch Eindrücke wie ›glänzend‹, ›stumpf‹ oder ›transparent‹), Funktion, Trageweise und Körpermodifikation finden. Neben diesen sind zudem Taxonomien der Kontextualisierung der Kostüme entstanden. Hierzu zählen die Taxonomien der Ortsbegebenheit, Stereotyp, Geschlecht, Rollenberuf, Alterseindruck, Spielzeit, Tageszeit, Familienstand, Charaktereigenschaft, Spielort und Genre. Allein die Taxonomie der Basiselemente hat 914 Konzepte, alle Taxonomien insgesamt kommen auf mehr als 3000 Konzepte. Um auch Relationen zwischen Basiselementen untereinander und deren Vorkommen in Bezug auf Körperregionen zu erfassen, erlauben speziell definierte Operatoren (z. B. ›darüber getragen‹, ›angesteckt‹, ›umgegürtet‹) die Basiselemente durch die Verwendung eines Subjekt-Prädikat-Objekt-Schemas in Bezug zu setzen.

¹³ Vgl. für mehr Details: Barzen et al. 2018, S. 5–16; Barzen 2018, S. 91–213.

¹⁴ Vgl. Barzen 2013.

¹⁵ Vgl. Hesse 2014, S. 306.

¹⁶ Vgl. Furrer 2014, S. 308–309.

¹⁷ Welche der potenziellen Parameter, deren Zusammenstellungen und Ausprägungen sich hier als relevant erweisen, gilt es mittels der Auswertung zu untersuchen. Daher ist hier z. B. mittels Bekleidungslexika, verschiedener Internetquellen und Expertengesprächen eine möglichst umfassende und detaillierte Auswahl als Ausgangswert der Erfassung zusammengestellt worden, vgl. Barzen 2013, S. 3–7.

Als initiale Untersuchungsgrundlage wurden 60 Filme ausgewählt, jeweils 20 aus den Genres Western, High-School-Comedy und Märchen.¹⁹ Da das MUSE Projekt auf das Identifizieren etablierter Konventionen und damit auf wiederkehrende Elemente in den Kostümen abzielt, wurden Genres mit relativ festem Figureninventar ausgewählt. Zur Auswahl der 20 Filme innerhalb der Genres sind nach vorab aufgestellten Kriterien diejenigen Filme bestimmt worden, bei welchen potenziell ein großer Einfluss auf die Mediensozialisation angenommen werden kann: Der Argumentation folgend, dass die Filme mit großem Publikum die Entstehung von Konventionen begünstigen, sind pro Genre, basierend auf Rankings von nationalen und internationalen Filmdatenbanken sowie Boxoffice-Zahlen²⁰, diejenigen Filme ausgewählt worden, die im Vergleich am häufigsten auf den oberen Plätzen zu finden waren. Schwierigkeiten, die sich in diesem Zusammenhang ergeben, sind die Unschärfe der Genre-Begriffe oder die Mehrsprachigkeit, die durch den Einbezug der internationalen Datenbanken gegeben ist, aber auch der Umgang mit animierten oder schwarz-weiß gedrehten Filmen, um nur einige zu nennen. Indem über mehrere Quellen hinweg die Anzahl der Nennungen als ausschlaggebend gewertet wurde, sollte ein möglichst repräsentativer Querschnitt zusammengestellt werden.²¹ Für das Genre Märchen beispielsweise wurden die fünf folgenden Quellen herangezogen:

1. [moviepilot](#)
2. [kino](#)
3. [imdb](#)
4. [boxofficemojo a](#) (Genre: Live Action Fantasy)
5. [boxofficemojo b](#) (Genre: Romantic Fantasy)²²

Tabelle 1 gibt eine Übersicht über die 20 untersuchten Filme für das Genre Märchen in alphabetischer Reihenfolge sowie deren Nennungen in den Quellen.

Filmkorpus: Märchen									
Titel	Originaltitel	Regisseur	Produktionsland	Produktionsjahr	moviepilot.de	kino.de	imdb.com	boxofficemojo.com a	boxofficemojo.com b
Beady	Beady	David Berza	USA	2011	x	x	x	x	2
Brothers Grimm	The Brothers Grimm	Terry Gilliam	CZ, USA, UK	2005	x	x	x	x	3
Cinderella	Cinderella	Kenneth Branagh	USA, UK	2015	x	x	x	x	5
Der Sturzwanderer	Standard	Matthew Vaughn	USA, UK	2007	x	x	x	x	2
Die Tag des Falken	Larkbenke	Richard Donner	USA	1985	x	x	x	x	3
Die Braut des Prinzen	The Princess Bride	Rob Reiner	USA	1987	x	x	x	x	3
Die Chroniken von Namia - Der König von Namia	The Chronicles of Namia: The Lion, the Witch and the Wardrobe	Andrew Adamson	UK	1988	x	x	x	x	2
Die fantastische Welt von Oz	Oz the Great and Powerful	Sam Raimi	USA	2013	x	x	x	x	2
Die Schöne und das Biest	Beauty and the Beast	Bill Condon	USA	2017	x	x	x	x	3
Die Schöne und das Biest	La Belle et le Bête	F	2014	x	x	x	x	2	
Die unendliche Geschichte	The Never Ending Story	Wolfgang Petersen	D, USA	1984	x	x	x	x	2
Das Handluder für Achenbach	Tri otvsky pro Popelku	Vladimir Vainok	CSSR, DDR	1973	x	x	x	x	2
Elle - Verflucht & unsterblich	Elle Enchanted	Timothy O'Flaherty	USA, UK, IRL	2004	x	x	x	x	2
Into the Woods	Into the Woods	Rob Marshall	USA	2014	x	x	x	x	2
Legend	Legend	Kidley Scott	USA, UK	1985	x	x	x	x	3
Maleficent	Maleficent	Robert Stromberg	USA	2014	x	x	x	x	2
Peter Pan	Peter Pan	Paul John Hooper	USA	2003	x	x	x	x	2
Snow White and the Huntsman	Snow White and the Huntsman	Rupert Sanders	USA	2012	x	x	x	x	2
Spiegeln Spiegel	Mirror Mirror	Tarun Singh	USA	2012	x	x	x	x	2
Verwünscht	Enchanted	Kevin Lima	USA	2007	x	x	x	x	2

Tab. 1: Das Korpus der 20 untersuchten Märchenfilme mit Nennung der Quellen. [privat]

¹⁹ Wobei wir uns im vorliegenden Fall als initiales Korpus auf den westlichen Unterhaltungsfilm fokussieren. Dies kann nur als Startpunkt verstanden werden, dessen Aussagekraft durch Erweiterung des Korpus auf Filme aus anderen Genres und Kulturkreisen in weiterführenden Arbeiten bereichert werden kann.

²⁰ Welche potenziell mit der Anzahl von Besucher*innen korrelieren.

²¹ Vgl. für mehr Details: Barzen 2018, S. 128–136.

²² Die Filme wurden den 11. Mai 2016 ausgewählt, die verwendeten Listen wurden teilweise aktualisiert oder sind heute nicht mehr abrufbar.

3.3 Die Erfassung der Daten mittels MUSE Repository

Das Interface des **MUSE Repositorys** unterstützt bei der Erfassung der Daten (vgl. Abbildung 2) und führt durch den Prozess, in welchem detaillierte Informationen zu jedem Film, den vorkommenden Rollen und den jeweiligen Kostümen aufgenommen werden. Dies beinhaltet für den Film Informationen zu Filmtitel, Originaltitel, Regisseur*in, Kostümbildner*in, Erscheinungsjahr, Dauer, Produktionsorten, Genrezugehörigkeiten, Farbkonzepten und Stil. Da gerade bei den Nebenrollen und Statist*innen davon ausgegangen werden kann, dass hier die Kostümbildner*innen auf Konventionen zurückgreifen, um ein schnelles Kommunizieren dieser Rolle zu gewährleisten, wird jede Rolle mit Sprechanteil erfasst und alle Kostüme beschrieben. Pro Rolle werden allgemeine Informationen wie Rollenname, Darsteller*innen-Name, Rollenberuf, Geschlecht, dominanter Alterseindruck, dominante Charaktereigenschaft, Familienstand, Rollenrelevanz (was sich auf Haupt-, Nebenrolle oder Statist bezieht) und Stereotypen erfasst. Bei jedem Kostüm werden sowohl Daten über den Kontext (Wie lange ist es zu sehen? Wo ist es zu sehen? Was passiert in der Szene? etc.), allgemeine Informationen über das Kostüm (Was ist die dominante Farbe? Was ist die dominante Funktion oder der dominante Zustand des Kostüms? etc.), wie auch sehr detaillierte Beschreibungen (Was ist die Kragenfarbe und Kragenform? Welche Verschlussarten und Verzierungen liegen vor? etc.) erhoben. Präzise Informationen über die Dauer, die ein Kostüm im Bild ist,²³ erlauben beispielsweise in der Auswertung Gewichtungen der Kostüme vorzunehmen. Der Annahme folgend, dass ein Kostüm, das über den gesamten Film getragen wird, einen größeren Effekt haben könnte, als eines, das nur für ein paar Sekunden zu sehen ist, kann man dieses Kostüm für die Auswertung dementsprechend gewichten.²⁴

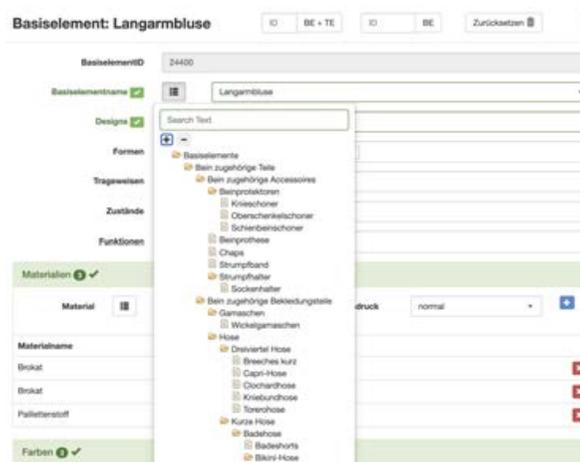


Abb. 2: Der Screenshot verdeutlicht exemplarisch wie das MUSE Repository durch die Erfassung eines Basiselementes führt. Zu sehen ist hier ein Ausschnitt der Maske zur Erfassung eines Basiselementes mit aufgeklappter Taxonomie der Basiselemente. [privat]

²³ Wobei über den Timecode erfasst wird, wann das Kostüm zu sehen ist, so dass eine genaue Angabe, wie lange das Kostüm im Bild zu sehen war, über das Aufsummieren der Sekunden möglich ist.

²⁴ Vgl. für mehr Details: Barzen 2018, S. 136–170.

Da die Kostüme bis in die Teilelemente eines Kleidungsstücks, sowie deren jeweilige Ausprägungen in Material, Farbe, und Form erfasst werden, ist der Prozess zeitaufwendig. In der bisher sieben Jahre andauernden Erfassung waren zehn studentische Hilfskräfte²⁵ des Instituts für Architektur von Anwendungssystemen der Universität Stuttgart beteiligt, von denen drei aktuell²⁶ noch an der Datenerfassung arbeiten. Zudem waren zwei weitere technische Hilfskräfte an der Implementierung und Wartung des MUSE Repositoriums beteiligt.²⁷ Die Filme werden in deutscher Sprache und gleicher Filmfassung auf dem **VLC Player** abgespielt und bearbeitet, um Timecodes etc. synchron zu halten. Trotz der Formularstruktur, der vorwiegend vordefinierten und auswählbaren Werte innerhalb des Formulars und der Möglichkeit des optischen Markierens fehlender Werte, lassen sich Fehler nicht ganz vermeiden. Um eine möglichst optimale Datenqualität zu gewährleisten, wird im Vier-Augen-Prinzip vorgegangen. D. h. jeder Film wird zweifach komplett bearbeitet, wobei beim ersten Mal die Daten angelegt werden und diese beim zweiten Mal überprüft und, wenn nötig, korrigiert werden. So werden einerseits Flüchtigkeitsfehler behoben und andererseits Absprachen im Team einheitlich umgesetzt. Diese beziehen sich beispielsweise darauf, wie mit bestimmten kritischen oder mehrdeutigen Fällen umgegangen wird, z. B. bei nicht eindeutigen Funktions- oder Materialzuweisungen oder was als ›dominant‹ angesehen wird. Aber auch Absprachen darüber, dass, wenn möglich, immer die tiefste Ebene der Taxonomie als das spezifischste Element ausgewählt wird oder bei der Erfassung der Timecodes immer dann ein neuer Timecode angelegt wird, wenn das Kostüm länger als zwei Sekunden aus dem Bild ist, fallen darunter.

3.4 Aktueller Stand der Datenerfassung

Nach aktuellem Stand sind 55 der 60 initial ausgewählten Filme bereits komplett erfasst, wobei alle Filme der Genres Western und High-School-Comedy auch bereits final korrigiert sind. Insgesamt sind damit 5.245 Kostüme mit 28.822 Basiselementen, die wiederum 63.773 Teilelemente aufweisen, in dem Datensatz zu finden. Das entspricht durchschnittlich ca. 95 Kostümen pro Film mit 5,6 Basiselementen pro Kostüm und 12,5 Teilelementen pro Basiselement. Dabei sind beispielsweise 159.526 Mal bestimmte Farben und 180.274 Mal bestimmte Materialien als Attribute zugeordnet worden.

3.5 Verfügbarkeit

Um eine umfassende und für die Nachnutzung der Daten optimierte Veröffentlichung des Datensatzes zu ermöglichen, wird das **MUSE Datenset** diese der Öffentlichkeit über **DaRUS**, das Datenrepositorium der Universität Stuttgart, zur Verfügung gestellt.²⁸ Eine Teilmenge der

²⁵ Anika Trischler; Chantelle Brand; Kim Kopec; Lea Intelmann; Leandra Brand; Lisa Beckh; Lisa Schepelmann; Maria Meyers; Regina Ababkov; Sophia Kaiser.

²⁶ Im April 2021.

²⁷ Fabian Bühler; Michael Falkenthal.

²⁸ Vgl. Barzen et al. 2021a.

Daten wurde bereits im Dezember 2018 über GitHub veröffentlicht,²⁹ indem der Code des **MUSE Repositoriums** zusammen mit einem Teil der Daten dort öffentlich zugänglich gemacht wurde. Nun hat sich in der Zwischenzeit der Datensatz signifikant vergrößert, die Richtlinien bei GitHub haben sich dahingehend geändert, dass Dateien ab einer bestimmten Größe nicht mehr kostenfrei zur Verfügung gestellt werden können, und auch die Anforderungen an die Veröffentlichungen von Forschungsdaten und die dafür bereitgestellten Infrastrukturen haben sich verschärft. DaRUS begegnet allen drei Gegebenheiten: es ermöglicht das Archivieren, Teilen und Veröffentlichen von Forschungsdaten und unterstützt explizit die FAIR-Prinzipien, wobei:

1. die Daten mit einer DOI versehen werden, was sie innerhalb der Community sichtbar und langfristig nutzbar macht,
2. die Datensätze mittels Metadaten so beschrieben werden, dass sie leicht auffindbar und teilbar sind und
3. mittels einer API der Zugriff auf die Daten automatisierbar ist.

Um die Daten einer möglichst breiten Interessens- und Nutzer*innengemeinschaft zur Verfügung zu stellen, sind diese in mehreren Formaten zugänglich. Die Daten können in eine MySQL- oder MariaDB-Datenbank importiert oder direkt in Form einer SQLite-Datenbank verwendet werden. Die beiden Formate wurden gewählt, um Informationsverluste durch das Konvertieren der Daten auszuschließen. Bei einem Export in das CSV-Format, würden beispielsweise Informationen darüber, wie Datensätze über mehrere Tabellen hinweg verbunden sind, verloren gehen. Um den Zugang so einfach wie möglich zu gestalten, ist zudem eine vorgefertigte SQLite-Datenbank Teil des Datensatzes. Diese kann z. B. mit dem **DB Browser for SQLite** geöffnet werden. So können bei Bedarf auch einzelne Tabellen gezielt als CSV exportiert werden.

Der Datensatz wurde zusätzlich in mehrere Teile aufgeteilt, welche nach Bedarf in die Datenbank importiert werden können. Der für eine Auswertung der Kostümdateien notwendige Teil – basierend auf deren textuellen Beschreibungen der Kostüme – umfasst etwa 50 MB.³⁰ Um in der Auswertung und deren Interpretation auch visuelle Repräsentationen der Kostüme zu ermöglichen, sind Illustrationen der Basiselemente in einer eigenen Datei mit einer Größe von etwa 500 MB gespeichert. Da bei der Erfassung der Kostüme zur internen Kontrolle der Daten Screenshots zu den Kostümen und deren Details hinzugefügt werden, welche dem Urheberrecht unterliegen, kann nicht die vollständige Datenbank von MUSE öffentlich gemacht werden. Der Teildatensatz mit den Screenshots ist deshalb auf DaRUS nicht für die Öffentlichkeit freigegeben. In den Vollversionen inklusive der Screenshots umfasst der Datensatz mehr als 12 GB. Da noch aktiv an dem Datensatz gearbeitet wird, beinhaltet die

²⁹ Hierbei ist nur ein Teil der Kostüme aus den Genres Western und High-School-Comedy, meist noch ohne die Korrektur, veröffentlicht worden.

³⁰ Der Datensatz enthält zwei Filme ohne inhaltlich relevante Daten (›1.setr‹ und ›1.yx‹), die zum Testen des MUSE Repositoriums genutzt werden, und fünf Versionen der gleichen kurzen Sequenz aus dem Film *Bad teacher* (als ›Bad teacher Test 1-5‹ betitelt). Letztere wurden zur Abschätzung genutzt inwieweit sich die erfassten Daten unterschiedlicher Erfasser*innen unterscheiden um potenzielles ›Rauschen‹ in den Daten zu ermitteln.

Referenz auf den veröffentlichten Datensatz in DaRUS die Version ›Eins‹ des Datensets, welche aktualisiert wird, sobald sich signifikante Änderungen ergeben haben, wobei hierbei die Versionen hochgezählt werden.

4. Relevanz der Daten

Die Relevanz der Daten innerhalb des Projekts MUSE ist durch die intrinsische Forschungsfrage des Projektes definiert und zielt auf das Schärfen des Verständnisses filmisch vestimentärer Kommunikation. Hier im Besondern auf das Identifizieren von Konventionen und Stilmitteln, welche sich – in einem wohldefinierten Rahmen –, entwickelt haben, um mittels des Kostüms bestimmte Eigenschaften, Stereotypen oder Zugehörigkeiten zu kommunizieren. Über die immanenten Forschungsinteressen von MUSE hinaus, bietet der Datensatz für weiterführende Untersuchungen aber durchaus Potenzial. Beide Aspekte werden die nachfolgenden Abschnitte beleuchten.

4.1 Auswertung innerhalb von MUSE

Der MUSE Datensatz wird im Zuge des MUSE Projektes mithilfe unterschiedlicher semi-automatisierter Analysemethoden untersucht, um die signifikanten Elemente zu identifizieren, welche als häufig angewandte und damit bewährte Lösung in Kostümmuster überführt werden können. Um innerhalb der Daten Hinweise auf Kostümmuster als Bausteine einer möglichen Kostümsprache zu identifizieren, werden aktuell Analyseverfahren sowohl aus dem Bereich der Data Analytics³¹ als auch dem des Maschinellen Lernens³² angewandt.

Das generelle Vorgehen der Auswertung folgt den häufig iterativ ausgeführten Analyseschritten:

1. Datenbereinigung, welche beispielsweise Formatkonvertierung oder Behandlung fehlender Werte beinhaltet.
2. Datenaufbereitung, welche sich im vorliegenden Fall z. B. mit der Überführung von kategorialen in numerische Daten befasst.
3. Auswahl des Algorithmus, was sowohl die Auswahl des Verfahrens (Clustering, Klassifikation etc.), als auch die Auswahl eines darin zur Verfügung stehenden Algorithmus, dessen Implementierung (bei Clustering z. B. OPTICS, k-Means, Maximum Cut etc.) und dessen Hyperparameter beinhaltet.
4. Ausführen des gewählten Algorithmus.
5. Visualisierung der Ergebnisse.
6. Erstellung von Hypothesen.
7. Anschließende Validierung der Hypothesen.

³¹ Vgl. Falkenthal et al. 2017.

³² Vgl. Barzen 2021.

Wenn eine solche Hypothese gefunden ist, kann diese auf Musterkandidaten hinweisen, so dass noch die Schritte:

8. des Identifizierens der Musterkandidaten
9. der Interpretation und Überführung in ein Muster

und

10. die Erstellung der Relationen zwischen Mustern

durchgeführt werden können.³³

Aktuell werden u. a. verschiedene Clusteralgorithmen erprobt, um diejenigen Kostüme zu gruppieren, welche ähnliche Attribute aufweisen und somit auf Kostümmuster hinweisen können (vgl. [Abbildung 3](#)). Hierbei sind willkürlich zu Anschauungszwecken 100 Kostüme aus dem Datensatz ausgewählt worden, wobei diese über ihre dominante Farbe und den dominanten Alterseindruck gruppiert werden sollen. Eine große Herausforderung bei der Anwendung verschiedener Analysealgorithmen ist, dass viele Algorithmen numerische Daten voraussetzen, der MUSE Datensatz aber zu einem großen Teil aus kategorialen Daten besteht. Um im zweiten Arbeitsschritt, der Datenaufbereitung, die Problematik eines sehr hochdimensionalen Merkmalsraums bei der Überführung der Daten in numerische zu vermeiden,³⁴ was in der Auswertung häufig die Anwendung von Feature-Reduction-Verfahren nötig machen würde, kann hier die Besonderheit des MUSE Datensatzes genutzt werden, um den in einer Baumstruktur vorliegenden Daten Ähnlichkeitswerte zuzuordnen. D. h. für die in [Abbildung 3](#) ausgewählten Kostüme können jedem untersuchten Kostüm zu jedem anderen Kostüm über die Attribute dominante Farbe und den dominanten Alterseindruck mittels des Ähnlichkeitsmaßes nach Wu und Palmer³⁵ mengenwertige Ähnlichkeitswerte zugeordnet werden, welche auf den Abständen der ausgewählten Werte innerhalb der Taxonomien beruhen. Hellblau und dunkelblau sind damit deutlich näher in der Taxonomie und bekommen so auch einen höheren Ähnlichkeitswert als hellblau zu schwarz.³⁶ Somit wird das semantische Wissen, welches in der Ontologie enthalten ist, genutzt, um die Kostüme zueinander über ihre Attribute in Beziehung zu setzen. Ähnlichkeitswerte können dann in Distanzwerte überführt und als solche mittels Einbettung in einem metrischen Merkmalsraum abgebildet werden. Dabei spiegeln die Distanzen der Datenpunkte zueinander die Ähnlichkeitswerte der Kostüme zueinander wider. Hierauf können nun unterschiedliche Clusteralgorithmen, wie beispielsweise eine Implementierung des Clusteralgorithmus k-Means³⁷, angewendet werden (vgl. [Abbildung 3](#)).

³³ Vgl. für mehr Details: Barzen 2021, S. 11–16.

³⁴ Wie sie beispielsweise durch das One-Hot-Encoding, als eines der gängigen Verfahren um kategoriale Daten in einen Merkmalsraum abzubilden, entstehen: Hierbei wird jedem Attribut ein Merkmalsvektor zugeordnet.

³⁵ Vgl. Wu / Palmer 1994.

³⁶ Vgl. Barzen et. al 2021, S. 40–46.

³⁷ Vgl. Arthur / Vassilvitskii 2007.

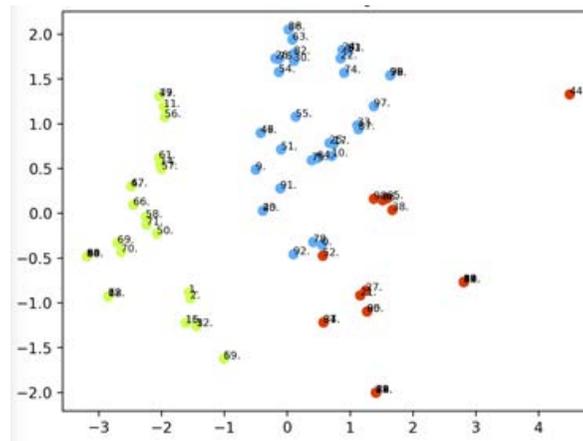


Abb. 3: Exemplarisches Ergebnis des Clusteralgorithmus k-Means mit drei identifizierten Clustern. Hierzu wurden die Kostüme auf Basis ihre Distanzwerte zueinander vorab mittels der Multidimensionalen Skalierung (MDS) in einen zweidimensionalen metrischen Merkmalsraum eingebettet. Die Zugehörigkeit von einem Kostüm zu einem Cluster wird über den Farbcode angezeigt. [privat]

Abbildung 4 zeigt einen Ausschnitt der in den gefundenen Clustern gruppierten Kostüme (in Abbildung 3 ist etwa Cluster 1 mit grünen Punkten markiert und Cluster 2 mit roten Punkten). Wie Abbildung 4 verdeutlicht, lässt sich für jugendliche Rollen viel mehr Farbvarianz in der dominanten Farbe der Kostüme verzeichnen, als bei Kostümen für ältere Charaktere, so dass bei den Rollen mit Alterseindrücken in den 30er- bis 60er-Jahren verstärkt Schwarz, vereinzelt auch Weißtöne, zu finden sind. Dies ist ohne Frage ein sehr simplifiziertes Beispiel, bei dem es gilt, weiter zu untersuchen, ob dieses Ergebnis beispielsweise über Genre, Gender, etc. hinweg stabil bleibt, genauso wie es für deutlich größere Datenmengen – die der MUSE Datensatz zur Verfügung stellt – zu verifizieren wäre.

Würde sich diese Hypothese nun durch weiterführende und umfassendere Untersuchungen als stichhaltig erweisen, könnte sie in ein Kostümmuster überführt werden. Wie Abbildung 4 zudem veranschaulicht, sind in der Auswertung die Kostüme jeweils mit dem MUSE Repositoryum verlinkt (mit Direktlinks zu den Informationen zum Film, der Rolle und dem Kostüm), was den nötigen interpretatorischen Schritt in der Überführung von Musterkandidaten in ein Muster unterstützt. Auch wenn die Auswertung es ermöglicht, Hinweise auf Musterkandidaten zu identifizieren, sowie Nachvollziehbarkeit und Reproduzierbarkeit der getroffenen Aussagen über vestimentäre Kommunikation zur Verfügung zu stellen, ist die eigentliche Überführung in das Musterformat eine Abstraktionsleistung, die auf Auslegung, Interpretation und Übersetzung beruht. Hierbei kann das Heranziehen der detaillierten Kostümbeschreibung im MUSE Repositoryum wiederum unterstützend genutzt werden. Die identifizierten Muster, mit ihren Relationen untereinander,³⁸ können dann in einem Muster-Repositoryum, wie beispielsweise dem Pattern Atlas,³⁹ abgelegt werden und so für weiterführende Anwendungen in Theorie und Praxis zur Verfügung gestellt werden.

³⁸ Vgl. für identifizierte Kostümmuster: Barzen 2018, S. 192–213.

³⁹ Vgl. Leymann / Barzen 2021.

Cluster 1: grüne Punkte						Cluster 2: rote Punkte					
ID	Referenz Film	Referenz Rollen	Referenz Kostuem	Dominant Color	Dominant Age Impression	ID	Referenz Film	Referenz Rollen	Referenz Kostuem	Dominant Color	Dominant Age Impression
13	Link	Link	Link	[Blautöne]	[Jugendlicher]	19	Link	Link	Link	[Schwarz]	[30ern]
14	Link	Link	Link	[Oliv]	[Jugendlicher]	21	Link	Link	Link	[Schwarz]	[60ern]
15	Link	Link	Link	[Antrazit]	[Jugendlicher]	23	Link	Link	Link	[Schwarz]	[40ern]
16	Link	Link	Link	[Antrazit]	[Jugendlicher]	27	Link	Link	Link	[Schwarz]	[20ern]
18	Link	Link	Link	[Dunkelblau]	[Jugendlicher]	28	Link	Link	Link	[Schwarz]	[40ern]
49	Link	Link	Link	[Mittelbraun]	[Jugendlicher]	29	Link	Link	Link	[Schwarz]	[30ern]
50	Link	Link	Link	[Hellblau]	[Kind]	31	Link	Link	Link	[Creme]	[40ern]
56	Link	Link	Link	[Dunkelbraun]	[Jugendlicher]	32	Link	Link	Link	[Schwarz]	[30ern]
57	Link	Link	Link	[Dunkelgrün]	[Jugendlicher]	34	Link	Link	Link	[Schwarz]	[40ern]
58	Link	Link	Link	[Zitronengelb]	[Jugendlicher]	35	Link	Link	Link	[Schwarz]	[30ern]
59	Link	Link	Link	[Schwarz]	[Jugendlicher]	37	Link	Link	Link	[Creme]	[40ern]
60	Link	Link	Link	[Blautöne]	[Jugendlicher]	38	Link	Link	Link	[Weißtöne]	[30ern]
61	Link	Link	Link	[Grüntöne]	[Jugendlicher]	39	Link	Link	Link	[Schwarz]	[30ern]
62	Link	Link	Link	[Dunkelblau]	[Jugendlicher]	40	Link	Link	Link	[Schwarz]	[30ern]
66	Link	Link	Link	[Rottöne]	[Jugendlicher]	41	Link	Link	Link	[Schwarz]	[30ern]
67	Link	Link	Link	[Blaurot]	[Jugendlicher]	42	Link	Link	Link	[Schwarz]	[40ern]
68	Link	Link	Link	[Blautöne]	[Jugendlicher]	43	Link	Link	Link	[Schwarz]	[30ern]

Abb. 4: Ausschnitt der Auflistung der Kostüme von zwei der Cluster aus Abbildung 3, wobei Cluster 1 in Abb. 3 mit grünen Punkten und Cluster 2 mit roten Punkten markiert ist. [privat]

Der Schritt der Analyse kann aber nicht nur bei der Erstellung der Muster helfen, sondern auch bei der Zuordnung von neuen Kostümen zu schon identifizierten Clustern, so dass beispielsweise die kontinuierlich weiter erfassten Kostüme direkt dem richtigen Cluster zugeordnet werden können. Zur Anwendung kommen hier verschiedene Klassifikationsalgorithmen, wie beispielsweise Support Vector Machines (SVM) oder neuronalen Netzen. Um die Präzision der Zuordnung zu erhöhen, werden hierbei nicht nur Algorithmen eingesetzt, welche auf dem klassischen Computer ausgeführt werden, sondern auch solche, die bereits auf Quantencomputern ausgeführt werden können. Gerade in der Klassifizierung konnten hier deutlich genauere Ergebnisse erzielt werden als mit der klassisch implementierten Variante.⁴⁰

4.2 Nachnutzungspotenzial über MUSE hinaus

Ohne DOI und Metadaten sowie außerhalb der MUSE Werkzeugumgebung war der aufwendig erstellte umfassende Datensatz bisher kaum für weiterführende Anwendungen auffindbar und nutzbar, so dass das inhaltliche Potenzial von Interessierten wenig genutzt werden konnte. Dies gilt sowohl für weiterführende Fragestellungen auf den eigentlichen Daten, aber auch in der Verbindung der MUSE Daten mit anderen Datensätzen. Dem intrinsischen Forschungsinteresse des Projektes MUSE geschuldet, fokussiert sich die Auswertung aktuell stark auf das stereotype Element von Kostümen, auf die Konventionen, bei der über die zu beobachtende Wiederholung argumentierend Aussagen über eine mögliche Kostümsprache getroffen werden können. Vielversprechend wäre nun, von den identifizierten Konventionen ausgehend, zu untersuchen, inwieweit und warum bei bestimmten Filmen oder Rollen von diesen abgewichen wird, wie diese aufgegriffen oder eventuell bewusst kontrastiert werden. Aber auch Untersuchungen, inwieweit bestimmte Kostümbildner*innen oder Regisseur*innen eigene Variationen einer Kostümsprache entwickelt haben, könnten mittels der Daten nähere Betrachtung erfahren. Aktuell ist das Projekt auf das Kostüm als isolierte Entität fokussiert, wobei es als filmisches Gestaltungsmittel stark mit anderen filmischen Gestaltungsmitteln interagiert. Um dieses Zusammenspiel zu untersuchen, ist die Verknüpfung des MUSE

⁴⁰ Vgl. Barzen 2021, S. 44f.

Datensatzes mit anderen Datensätzen vielversprechend. So wäre z. B. eine Kombination mit Daten zu Filmräumen äußerst spannend, um hier Kostüme in den Raum, dessen Farbigkeit und Struktur, einzubetten. Dies wiederum kann durch Verschmelzung oder Kontrastierung eine wichtige Komponente in der Wahrnehmung des Kostüms darstellen, was gleichzeitig aber auch Synergieeffekte für das Verständnis über Filmräume ermöglichen kann. Gleiches gilt für eine Kombination der auditiven Komponente des Kostüms, wie das Klappern von Absätzen oder das Rauschen schwerer Seidenröcke, mit Daten zu Filmgeräuschen oder Filmmusik, was das Verständnis über Kostüme weiter schärfen könnte.⁴¹ Darüber hinaus wäre auch eine Verbindung der Daten mit den in den Drehbüchern vorhandenen Kostümbeschreibungen und den vorgenommenen Abweichungen von Interesse. Jenseits des Mediums Film wäre auch eine Kombination von textuellen Beschreibungen von Kostümen in Theaterstücken oder Kleidung in der Literatur vielversprechend.

Über den fachlichen Mehrwert hinaus ist zudem die strukturelle Besonderheit des Datensatzes bereichernd: So können, wie in Abschnitt 4.1 beschrieben, durch die zugrundeliegenden taxonomischen Strukturen der textuellen Beschreibungen die jeweiligen Kostüme untereinander mit Ähnlichkeitswerten versehen werden. Dies ermöglicht, das in der Ontologie abgebildete semantische Wissen nutzbringend anzuwenden. Durch die Überführung der kategorialen in numerische Daten wird zudem eine deutlich breitere Anwendung von potenziellen Analysealgorithmen möglich.⁴² Dass das vorgestellte Vorgehen von MUSE nicht auf Kostüme beschränkt ist, sondern auch in ganz anderen Bereichen eingesetzt werden kann, wird durch MUSE4Musik verdeutlicht.⁴³ MUSE4Musik beschäftigt sich mit der Emotionalität symphonischer Musik des 19. Jahrhunderts. Hierbei werden, der in Abschnitt 3.1 beschriebenen Strukturierung folgend, detaillierte Daten über Musikausschnitte erfasst, um diese auf wiederkehrende Elemente zu untersuchen, welche potenziell ähnliche Emotionen transportieren. Sowohl das MUSE als auch das MUSE4Music Repositorium wurden jeweils aufbauend auf einer vorab erstellten und dem Anwendungsfall angepassten Ontologie, aufwendig und zeitintensiv implementiert. Um weitere Anwendungen des Vorgehens in unterschiedlichen Domänen effizienter zu ermöglichen, haben wir **MUSE4Anything** entwickelt.⁴⁴ MUSE4Anything erlaubt, die Ontologien und darauf aufbauend generierte Eingabeformulare für die Daten direkt in einem Werkzeug zu erstellen und es so für die strukturierte Datenerfassung in unterschiedlichen Domänen anzupassen. Darüber hinaus kann somit auch der Erkenntnisgewinn, der anhand des MUSE Datensatzes über die Auswertung der Daten erzielt werden konnte, für andere Anwendungsfälle, beispielsweise für anderer filmischer Gestaltungsmittel, potenziell nutzbringend eingesetzt werden.

5. Ausblick

⁴¹ Vgl. Falkenthal et al. 2015, S. 56f.

⁴² Vgl. Barzen et al. 2021b, S. 40–47.

⁴³ Vgl. Barzen et al. 2017.

⁴⁴ Vgl. Bühler 2021.

Wie in den vorherigen Abschnitten beschrieben, ist das MUSE Projekt nicht abgeschlossen. So werden aktuell noch die letzten Daten des initialen Filmkorpus erfasst, wie auch die Auswertungsumgebung von MUSE um weitere Filtermöglichkeiten und Analysealgorithmen erweitert, um die Daten von verschiedenen Seiten zu beleuchten und das Potenzial des Datensatzes hinsichtlich der gegebenen Fragestellung voll auszuschöpfen. Auch wenn erste Muster und deren Relationen zueinander bereits identifiziert sind, gilt es, weitere Muster als Bausteine einer Kostümsprache über das Musterrepositorium Pattern Atlas zur Verfügung zu stellen, um ein besseres – und verifizierbares – Verständnis der Kommunikation durch Kostüme in Filmen zu unterstützen.

Darüber hinaus sind die Daten, die darunterliegende Ontologie, wie auch das MUSE Repositorium und die Werkzeugumgebung zur Auswertung der Daten Open-Source- implementiert. Somit stehen sie Interessierten zur Verifikation oder zur Adressierung eigener, weiterführender Forschungsfragen zur Verfügung. Durch die Publikation über DaRUS sind die Daten mittels DOI und Metadaten so aufbereitet, dass ein dauerhafter und intuitiv gestalteter Zugang gewährleistet ist. Darüber hinaus wird mit MUSE4Anything ein Werkzeug zur Verfügung gestellt, welches es erleichtern soll, das MUSE Vorgehen, wie in Abschnitt 3 beschrieben, ohne Implementierungsaufwand auch in anderen Domänen einzusetzen. Damit sind sowohl die vorgestellten Daten an sich, wie auch deren struktureller Mehrwert durch einfache Übertragbarkeit des Vorgehens für andere Anwendungsfälle nutzbar.

Bibliografische Angaben

Christopher Alexander / Sara Ishikawa / Murry Silverstein / Max Jacobson / Ingrid Fiksdahl-King / Shlomo Angel: A Pattern Language: Towns, Buildings, Constructions. New York, NY 1977. [\[Nachweis im GBV\]](#)

David Arthur / Sergei Vassilvitskii: k-means++: The Advantages of Careful Seeding. In: SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms (2007), S. 1027–1035. [\[Nachweis im GBV\]](#)

Malcom Barnard: Fashion as Communication. London u. a. 2002. [\[Nachweis im GBV\]](#)

Roland Barthes: Die Sprache der Mode [Système de la Mode 1967]. Frankfurt / Main 1985. [\[Nachweis im GBV\]](#)

Johanna Barzen: From Digital Humanities to Quantum Humanities: Potentials and Applications. In: Quantum Computing in the Arts and Humanities: An Introduction to Core Concepts, Theory and Applications. Hg. Eduardo R. Miranda. Preprint 2021. DOI: [10.48550/arXiv.2103.11825](https://doi.org/10.48550/arXiv.2103.11825)

Johanna Barzen: Taxonomien kostümrelevanter Parameter: Annäherung an eine Ontologisierung der Domäne des Filmkostüms. Stuttgart 2013. (= Technischer Bericht, 2013/04) [\[online\]](#)

Johanna Barzen: Wenn Kostüme sprechen – Musterforschung in den Digital Humanities am Beispiel vestimentärer Kommunikation im Film. Köln 2018. [\[online\]](#)

Johanna Barzen / Frank Leymann: Costume Languages As Pattern Languages. In: Proceedings of Pursuit of Pattern Languages for Societal Change (PURPLSOC) Preparatory Workshop. Krems 2014, S.88–117. [\[Nachweis im GBV\]](#)

Johanna Barzen / Fabian Bühler / Frank Leymann (2021a): MUSE Datenset. DaRUS, Version 1 2021. DOI: [10.18419/darus-1805](https://doi.org/10.18419/darus-1805)

Johanna Barzen / Frank Leymann / Michael Falkenthal / Daniel Vietz / Benjamin Weder / Karoline Wild (2021b): Relevance of Near-Term Quantum Computing in the Cloud: A Humanities Perspective. In: Cloud Computing and Services Science. CLOSER 2020. Hg. Donald Ferguson / Claus Pahl / Markus Helfert. In: Communications in Computer and Information Science 1399 (2021), S. 25–58. [\[Nachweis im GBV\]](#)

Johanna Barzen / Michael Falkenthal / Frank Leymann: Wenn Kostüme sprechen könnten: MUSE – Ein musterbasierter Ansatz an die vestimentäre Kommunikation im Film. In: Digital Humanities. Perspektiven der Praxis. Hg. Peggy Bockwinkel / Beatrice Nickel / Gabriel Viehhauser. Berlin 2018, S.223–241. [\[Nachweis im GBV\]](#)

Johanna Barzen / Uwe Breitenbücher / Linus Eusterbrock / Michael Falkenthal / Frank Hentschel / Frank Leymann: The vision for MUSE4Music. Applying the MUSE method in musicology. In: Computer Science – Research and Development 32 (2017), H. 3–4, S. 323–328. [\[Nachweis im GBV\]](#)

Fabian Bühler: MUSE4Anything. Stuttgart 2021. DOI: [10.18419/opus-11410](https://doi.org/10.18419/opus-11410)

Rosa Burger: Contemporary Costume Design. Dress Codes und weibliche Stereotype im Hollywood-Film. Wien 2002. [\[Nachweis im GBV\]](#)

Frank Buschmann / Regine Meunier / Hans Rohnert / Peter Sommerlad / Michael Stal: Pattern-Oriented Software Architecture. In: A System of Patterns 1. Chichester u. a. 1996. [\[Nachweis im GBV\]](#)

Daniel Devoucoux: Mode im Film. Zur Kulturanthropologie zweier Medien. Bielefeld 2007. [\[Nachweis im GBV\]](#)

Michael Falkenthal / Johanna Barzen / Uwe Breitenbücher / Christoph Fehling / Frank Leymann / Aristotelis Hadjakos / Frank Hentschel / Heizo Schulze: Leveraging Pattern Applications via Pattern Refinement. In: Proceedings of the International Conference on Pursuit of Pattern Languages for Societal Change (PURPLSOC). Berlin 2015, S. 38–61. [\[Nachweis im GBV\]](#)

Michael Falkenthal / Johanna Barzen / Uwe Breitenbücher / Sascha Brüggmann / Daniel Joos / Frank Leymann / Michael Wurster: Pattern Research in the Digital Humanities: How Data Mining Techniques Support the Identification of Costume Patterns. In: Computer Science – Research and Development 32 (2017), H. 3–4, S. 311–321. [\[Nachweis im GBV\]](#)

Christoph Fehling / Frank Leymann / Ralf Retter / Walter Schupeck / Peter Arbitter: Cloud Computing Patterns Fundamentals to Design, Build, and Manage Cloud Applications. Wien 2014. [\[Nachweis im GBV\]](#)

Andreas Fießler: A Pattern Language for Film Production. In: Proceedings of the 12th European Conference on Pattern Languages of Programs (EuroPLoP'07). Irsee 2007, S. A1-1–A1-A16. [\[Nachweis im GBV\]](#)

Frank J. Furrer: Eine kurze Geschichte der Ontologie. Von der Philosophie zur modernen Informatik. In: Informatik Spektrum. Organ der Gesellschaft für Informatik e.V. und mit ihr assoziierter Organisationen 37 (2014), H. 4, S. 308–317. [\[Nachweis im GBV\]](#)

Erich Gamma / Richard Helm / Ralph Johnson / John Vlissides: Design Patterns: Elements of Reusable Object-Oriented Software. Boston u.a. 1994. [\[Nachweis im GBV\]](#)

Antonella Giannone: Kleidung als Zeichen. Ihre Funktionen im Alltag und ihre Rolle im Film westlicher Gesellschaften. Eine kultursemiotische Abhandlung. Berlin 2005. [\[Nachweis im GBV\]](#)

Muster-Suchen-und-Erkennen/muse-for-anything. Hg. von Michael Wurster / Fabian Bühler. In: GitHub.com. Hg. von buehlefs. 2022. [\[online\]](#)

Wolfgang Hesse: Ontologie und Weltbezug – vom philosophischen Weltverständnis zum Konstrukt der Informatik. In: Informatik Spektrum. Organ der Gesellschaft für Informatik e.V. und mit ihr assoziierter Organisationen 37 (2014), H. 4, S. 298–307. [[Nachweis im GBV](#)]

Hans-Joachim Hoffmann: Kleidersprache. Eine Psychologie der Illusionen in der Kleidung, Mode und Maskerade. Frankfurt/M. 1985. [[Nachweis im GBV](#)]

Gregor Hohpe / Bobby Woolf: Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions. Boston 2004. [[Nachweis im GBV](#)]

Rebecca H. Holman: Clothing as Communication: an Empirical Investigation. In: Advances in Consumer Research. Hg. Jerry C. Olson / Ann Abor. 7 (1980), S. 372–377. [[Nachweis im GBV](#)]

Takashi Iba: Presentation Patterns. A Pattern Language for Creative Presentations. Yokohama 2014. [[Nachweis im GBV](#)]

Gottfried Keller: Kleider machen Leute. 4. Auflage. Frankfurt / Main 2017. [[Nachweis im GBV](#)]

Richard La Motte: Costume Design 101. The business and art of creating costumes for film and television. Saline, MI 2010. [[Nachweis im GBV](#)]

Frank Leymann / Johanna Barzen: Pattern Atlas. In: Next-Gen Digital Services. A Retrospective and Roadmap for Service Computing of the Future. Hg. Marco Aiello / Athman Bougettaya / Damian Tamburri / Willem-Jan van den Heuvel. Cham 2021, S. 67–76. (= Lecture Notes in Computer Science, 12521) [[Nachweis im GBV](#)]

Susan Lilly: Patterns for Pedagogy. Object Magazine 5 (1996), H. 8, S.93–96. [[Nachweis im GBV](#)]

Alison Lurie: The Language of Clothes. Featuring a new introduction about fashion today. New York, NY 2000. [[Nachweis im GBV](#)]

Ingeborg Petrascheck-Heim: Die Sprache der Kleidung. Wesen und Wandel von Tracht, Mode, Kostüm und Uniform. Wien 1966. [[Nachweis im GBV](#)]

Zhibiao Wu / Martha Palmer: Verb Semantics and Lexical Selection. In: ACL '94 Proceedings of the 32nd annual meeting on Association for Computational Linguistics (1994), S. 133–138. DOI: [10.3115/981732.981751](#)

Abbildungs- und Tabellenverzeichnis

Abb. 1: Exemplarisch generierter Kostüm-Graph für das erste Kostüm der Stiefmutter aus Cinderella (Kenneth Branagh 2015). TC 00:09:13–00:09:47, 00:09:55–00:10:30. [privat]

Tab. 1: Das Korpus der 20 untersuchten Märchenfilme mit Nennung der Quellen. [privat]

Abb. 2: Der Screenshot verdeutlicht exemplarisch wie das MUSE-Repositorium durch die Erfassung eines Basiselementes führt. Zu sehen ist hier ein Ausschnitt der Maske zur Erfassung eines Basiselements mit aufgeklappter Taxonomie der Basiselemente. [privat]

Abb. 3: Exemplarisches Ergebnis des Clusteralgorithmus k-Means mit drei identifizierten Clustern. Hierzu wurden die Kostüme auf Basis ihre Distanzwerte zueinander vorab mittels der Multidimensionalen Skalierung (MDS) in einen zweidimensionalen metrischen Merkmalsraum eingebettet. Die Zugehörigkeit von einem Kostüm zu einem Cluster wird über den Farbcode angezeigt. [privat]

Abb. 4: Ausschnitt der Auflistung der Kostüme von zwei der Cluster aus *Abbildung 3*, wobei Cluster 1 in *Abb. 3* mit grünen Punkten und Cluster 2 mit roten Punkten markiert ist. [privat]

Beitrag aus:

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. Teilband 2 / Sonderband 5 der ZfdG: DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Epistemologie digitaler Experimentalsysteme am Beispiel von Zeitungsportalen. Methodologische und praktische Herausforderungen, Probleme und Perspektiven

Autor*in:

Erik Koenen

Kontakt: ekoenen@uni-bremen.de

Institution: Universität Bremen, Zentrum für Medien-, Kommunikations- und Informationsforschung

GND: [173572502](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-73572502) ORCID: [0000-0001-9824-385X](https://orcid.org/0000-0001-9824-385X)

DOI des Artikels:

[10.17175/sb005_013](https://doi.org/10.17175/sb005_013)

Nachweis im OPAC der Herzog August Bibliothek:

[1783635886](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-1783635886)

Erstveröffentlichung:

20.09.2022

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autor*innen.

Letzte Überprüfung aller Verweise: 07.09.2022

GND-Verschlagwortung:

[Erkenntnistheorie](#) | [Experiment](#) | [Portal \(Internet\)](#) | [Zeitung](#) |

Zitierweise:

Erik Koenen: Epistemologie digitaler Experimentalsysteme am Beispiel von Zeitungsportalen. Methodologische und praktische Herausforderungen, Probleme und Perspektiven. 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/sb005_001](https://doi.org/10.17175/sb005_001) PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb005_013](https://doi.org/10.17175/sb005_013).

Erik Koenen

Epistemologie digitaler Experimentalsysteme am Beispiel von Zeitungsportalen. Methodologische und praktische Herausforderungen, Probleme und Perspektiven

Abstracts

Mit der Digitalisierung historischer Zeitungen und deren Einspeisung in digitale Zeitungssammlungen und -portale hat sich eine ebenso neuartige wie spezifische Forschungsinfrastruktur für Zeitungen herausgebildet. Gleichwohl werden digitalisierte Zeitungen und Zeitungsportale als Datenform bzw. digitale Infrastruktur bislang noch eher selten thematisiert. Vor diesem Hintergrund entwirft und verfolgt der Beitrag mithilfe des Konzepts digitaler Experimentalsysteme eine epistemologische Perspektive, die die Eigenarten, Effekte und Folgen digitaler Forschungskontexte für die Zeitungsforschung methodologisch hinterfragt sowie die sich hieraus ergebenden experimentellen Spielräume aufzeigt. Methodologische und praktische Herausforderungen, Probleme und Potenziale werden exemplarisch für Zeitungsportale als (1.) digitale Archive bzw. Sammlungen, (2.) ›reading machines‹ und (3.) ›virtual research environments‹ herausgearbeitet und reflektiert.

With the digitization of historical newspapers and their embedding into digital newspaper collections and portals a novel as well as specific research infrastructure for newspapers has emerged. Nevertheless, digitized newspapers and newspaper portals as a form of data and digital infrastructure are still rarely discussed. In light of this, the paper uses the concept of digital experimental systems to develop and pursue an epistemological perspective that methodologically questions the characteristics, effects and implications of digital research contexts in newspaper research, and that makes the resulting experimental scope visible. Methodological and practical challenges, problems and potentials will be elaborated and reflected exemplarily for newspaper portals as (1.) digital archives and collections, (2.) ›reading machines‹ and (3.) ›virtual research environments‹.

1. Einleitung

In den Digitalisierungsstrategien historischer Medien spielen Zeitungen eine herausragende Rolle.¹ Mit der kontinuierlichen Digitalisierung historischer Zeitungen und deren Einspeisung in spezifische Zeitungsportale entstehen seit etwa zwei Jahrzehnten weltweit riesige digitale Sammlungs- und Quellenkorpora. Gleichwohl werden digitalisierte Zeitungen und Zeitungsportale als Datenform bzw. digitale Forschungsinfrastruktur noch eher selten im Kontext der digitalen Geistes- und Geschichtswissenschaften thematisiert. Folgt man Madleen Podewski, dann sind Zeitungen ebenso sehr wie Zeitschriften ›bislang kein starkes Forschungsobjekt‹² der Digital Humanities. In gewisser Weise ist diese derzeitige Situation paradox: Während mittlerweile mehrere tausend Zeitungstitel und Millionen von Zeitungsseiten digitalisiert vorliegen und das korrespondierend zu beobachtende merklich gestiegene Interesse, Zeitungen zum Forschungsgegenstand bzw. zur Forschungsquelle zu

¹ Koenen 2018.

² Podewski 2018, S. 2.

machen, zweifellos eine Folge dessen ist, dass Zeitungen als massenhafte und sperrige Medien in ihrer digitalen Form einfach leichter zu haben, zu handhaben, zu lesen und zu sichten sind, wird die Eigenart und Nutzung von Zeitungen als digitale Objekte kaum methodologisch oder quellenkritisch hinterfragt und reflektiert.

Vielmehr spielt in der Forschungspraxis mit digitalisierten Zeitungen die objekt- und quellenkritische Frage »der Transformation ins digitale Format [...] offenbar keine nennenswerte Rolle«. ³ Astrid Blome hat deshalb gefordert, die Forschung mit digitalisierten Zeitungen durch eine »notwendige Erweiterung der Quellenkritik« zu fundieren, »die quellenkritisch Quantität und Relevanz, Selektion und Vermittlung des digitalen Materials« mit in den Blick nimmt. ⁴ Bob Nicholson spricht sogar von einer notwendigen exklusiven »digital methodology« für digitalisierte Zeitungen und Zeitungsportale: »In order to develop robust methodological responses to digital archives, we need to understand how newspapers are changed by digitisation.« ⁵ Damit weist Nicholson auf die wesentlichen forschungspraktischen Folgen der Digitalisierung für die Zeitungsforschung hin. Zeitungen werden im Zuge ihrer Digitalisierung zu *Metaquellen*: »A set of structured information, passed on the computer and processed by it.« ⁶ Und um sie in dieser digitalen Form zu erschließen, werden u. a. Datenformate, Portalinterfaces, Programmierschnittstellen oder Suchtechnologien und hiermit verbundene informatorisch-technische Logiken, Prinzipien, Prozesse, Routinen und Strukturen unmittelbar für die Forschung relevant. Vor diesem Hintergrund interessiert sich dieser Beitrag für eine epistemologische Perspektive, die diese neuen »technisch-medialen Voraussetzungen« ⁷ in eine Erkenntnis- und Quellenpraxis der Zeitungsforschung integriert, welche die signifikante Digitalität von Forschungsobjekt und Forschungskontext ernst nimmt und sowohl methodologisch transparent macht wie methodisch produktiv operationalisiert.

2. Herausforderungen der Konzeption digitaler Methodologien

Die Herausforderungen der Konzeption digitaler Methodologien unter den technisch-medialen Bedingungen von Digitalität und Digitalisierung sind komplex. ⁸ Im Hinblick auf den fortwährenden technologischen Wandel von digitalen Forschungsinfrastrukturen stellt sich hierbei insbesondere die grundsätzliche Frage der Nachhaltigkeit, ⁹ d. h., inwiefern solche methodologischen Reflexionen mehr als nur Momentaufnahmen sein können, die vielleicht den neuesten Entwicklungsstand einfangen, aber trotzdem prinzipiell dem technologischen Fortschritt hinterherlaufen. Zudem sind digitale Forschungsinfrastrukturen vielschichtige sozio-technische »data assemblages« ¹⁰, die nicht bloß Forschungsinteressen dienen, sondern ebenso informatorische, institutionelle, technische, rechtliche und unternehmerische Konditionen der Datengenerierung wie des Umgangs mit den vorgehaltenen Daten und Digitalisaten implizieren.

³ Podewski 2018, S. 3.

⁴ Blome 2018, S. 1, 33.

⁵ Nicholson 2013, S. 64.

⁶ Genet 1994, hier übersetzt von Fickers 2016.

⁷ Heinen 2011.

⁸ Vgl. Stalder 2021.

⁹ Vgl. DARIAH-DE 2015, S. 128–130.

¹⁰ Kitchin 2014, S. 22–26.

Entsprechend sollten digitale Methodologien genauso elastisch wie robust konzipiert sein. Sie sind aus einer epistemologischen Position zu entwerfen, die fragt, welche spezifisch neuen Denkstile und Forschungspraktiken sich bei der informatorisch-technisch gerahmten Auseinandersetzung mit Daten als digitalen Objekten und Quellen entfalten.

Nach Ludwik Flecks *Theorie des Erkennens* prägen nicht allein Denkmotive, Erkenntnisinteressen, Methoden oder Schulen wissenschaftliche Denkstile, sondern diese sind immer auch mit »technischen Termini« und dem genutzten »wissenschaftlichen Gerät« verknüpft.¹¹ »Indem wir inmitten von Geräten und Einrichtungen leben, die sich aus dem heutigen wissenschaftlichen Denkstil herleiten, empfangen wir ständig »objektive« Anstöße, so und nicht anders zu denken«,¹² so Fleck – was dann im Umkehrschluss heißt, dass *neue* Geräte, Techniken und Termini konventionelle Denkstile und Forschungspraktiken epistemologisch herausfordern und irritieren. Dass mit Computern, Daten und digitalen Forschungsinfrastrukturen *neue* Arrangements digitaler Denk- und Forschungswerkzeuge entstehen, um beispielsweise mit den Herausforderungen von *big data* in der historischen Forschung umzugehen, haben Shawn Graham, Ian Milligan und Scott Weingart mit dem Begriff des *Historian's Macroscope* illustriert, das für sie gleichermaßen ein »tool« und eine »perspective« meint: »What is more, a »macroscope« – a tool for looking at the very big – deliberately suggests a scientist's workbench, where the investigator moves between different tools for exploring different scales, keeping notes in a lab notebook.«¹³ Auch Daten, digitale Technologien und Software lenken also spezifisch neue Denk- und Forschungsstile »automatisch in die Bahnen«¹⁴ und ermöglichen und eröffnen Forschenden damit zugleich, so der Hinweis von Lev Manovich, vollkommen neue Forschungshorizonte: »Software is the interface to our imagination and the world – a universal language through which the world speaks, and a universal engine on which the world runs.«¹⁵

Unter diesen Voraussetzungen ist die Grundannahme dieses Beitrags, dass sich mit Daten, digitalen Technologien und Software neue Ensembles, Praktiken und Situationen geisteswissenschaftlicher Wissensordnung und -produktion herausbilden. Das heißt, mit dem digitalen Wandel der Infrastrukturen, Instrumente, Techniken und Quellen der *Fabrikation von Erkenntnis*¹⁶ verschiebt sich der epistemologische Rahmen von Erkenntnisgewinnung:

»To meet contemporary challenges, the humanities have made an »infrastructural turn« to build a physical place that, through the architecture itself, brings new features to the field: innovation, experimentation, hands-on practices, and collaboration.«¹⁷

Urszula Pawlicka-Deger hat diesen tiefgreifenden konzeptionellen Wandel in der geisteswissenschaftlichen Erkenntnispraxis, mit dem man in den Digital Humanities auf die Herausforderung reagiert, um »new research practices and methods related to the utilization

¹¹ Fleck 2011, S. 296f.

¹² Fleck 2011, S. 297.

¹³ Graham et al. 2016, S. XVI.

¹⁴ Fleck 2011, S. 297.

¹⁵ Manovich 2014, S. 80.

¹⁶ Knorr Cetina 1984.

¹⁷ Pawlicka-Deger 2020, S. 23.

of technology and digital tools« zu entwickeln und zu institutionalisieren, als »laboratory turn« bzw. »laboratorization« bezeichnet: »The lab becomes conceptualized as a way of thinking entailing new social practices and new research modes [...].«¹⁸

Von den von Pawlicka-Deger vorgestellten geisteswissenschaftlichen Labormodellen interessiert für die Ausleuchtung, Beschreibung und methodologische Reflexion von Zeitungsportalen als digitale Infrastrukturen vornehmlich das Modell des »virtual lab« bzw. des »virtual research environment«, das sie als »virtual space for facilitating digital research processes, providing databases, tools, and services, and promoting collaboration across the university, academic institutions, and cultural organizations« charakterisiert.¹⁹ Gleich, ob Zeitungsportale bloß der archivalisch-bibliothekarischen Logik des digitalen Archivs bzw. der digitalen Sammlung folgen oder ob sie explizit als virtuelle Forschungsumgebung konzipiert sind, ist ihnen die »virtual situatedness, defined as the digital, internet-based workspace with an infrastructure, connection, and operation that affect the work and research communication«²⁰ gemeinsam. Der Blick auf Zeitungsportale durch die konzeptionelle Brille des *virtual lab* erlaubt es uns folglich, die epistemologischen und forschungspraktischen Eigenarten, Folgen und Potenziale dieser originären *virtual situatedness* von Forschungsobjekt und Forschungskontext herauszuarbeiten. Für die weitere Operationalisierung der Laborperspektive wird erneut auf Hans-Jörg Rheinbergers *Epistemologie des modernen Experimentierens in Experimentalsystemen* zurückgegriffen,²¹ weil sich mit deren Fokus auf die generellen Bedingungen, Dinge und Praktiken der experimentellen Genese und Produktion von Wissen ebenfalls die epistemologisch, methodisch und technisch vielfach verschränkten Prozesse »serieller Wissensgenerierung«²² auf der Grundlage von Datenobjekten in digitalen Forschungskontexten einfangen lassen.²³ Im folgenden Abschnitt 3 wird Rheinbergers Konzeption des Experimentalsystems vorgestellt und als Arrangement digitaler Forschung interpretiert. Hieran anschließend werden digitalisierte Zeitungen als dynamische technisch-epistemische Assemblagen beschrieben, die hoch variable Praktiken der Quellendarstellung und -manipulation ermöglichen (Abschnitt 4). Schließlich wird kritisch diskutiert, inwieweit verschiedene Zeitungstypen als Experimentalsysteme der Zeitungsforschung solche Praktiken initiieren, stabilisieren und strukturieren (Abschnitt 5).

3. Digitale Forschungskontexte als ›Experimentalsysteme‹

In den Digital Humanities findet geisteswissenschaftliche Forschung in einer komplexen »medientechnischen Landschaft«²⁴ aus Daten, digitalen Technologien und Software statt. Aus einer grundlegenden epistemologischen Perspektive spiegelt sich in der Erkenntniserzeugung in solchen informatorisch-technisch gerahmten und mediatisierten Forschungskontexten das erkenntnisproduktive Oszillieren zwischen ›epistemischen Dingen‹ und ›technischen Dingen‹

¹⁸ Pawlicka-Deger 2020, S. 63f.

¹⁹ Pawlicka-Deger 2020, S. 60.

²⁰ Pawlicka-Deger 2020, S. 58.

²¹ Rheinberger 2001, S. 7; Hall 2014.

²² Link 2019, S. 30.

²³ Vgl. Koenen 2020; Koenen 2021, S. 123–128.

²⁴ Rheinberger 2021, S. 18.

der Wissensproduktion in Experimentalsystemen, wie es der Wissenschaftsforscher Hans-Jörg Rheinberger eingehend herausgearbeitet hat.²⁵ Folgt man diesem Gedanken, so lassen sich auch datenbasierte und in digitale Technologien und Software eingebettete und von diesen initiierte und vernetzte Erkenntnispraktiken und -prozesse als Experimentalsysteme konzipieren, in denen Epistemologien und Techniken des Erkennens und Erkenntnis stetig wechselseitig aufeinander verweisen.

Experimentalsysteme zeichnen sich für Rheinberger durch ihren originären *Werkstattcharakter* aus: Experimentalsysteme als »eigentliche Arbeitseinheiten« der Forschung sind ebenso »materielle Arrangements« wie »Versuchsfelder«, in denen »Wissensobjekte und die technischen Bedingungen ihrer Hervorbringung unauflösbar miteinander verknüpft« sind.²⁶ Im Gegensatz zum formalisierten, hypothesentestenden Experiment sind Experimentalsysteme erkenntnisoffen. Als »Generator von Überraschungen« sind sie induktiv voreingestellt explorativen und heuristischen Erkenntnispraktiken verpflichtet und müssen somit, »sollen sie produktiv bleiben, so organisiert sein, dass die Erzeugung von Differenzen zur reproduktiven Triebkraft der ganzen Experimentalmaschinerie wird.«²⁷

Mit Rheinbergers Experimental-Epistemologie wird so der Erkenntnis- und Forschungsprozess als wechselseitig wirkender Prozess sichtbar, in dem »Technisches und Epistemisches sich gegenseitig bedingen und vorantreiben.«²⁸ Dabei bilden sich beim Agieren der*des Forschenden *in* und deren*des Interagieren *mit* Experimentalsystemen quasi »von unten«²⁹ »die Objekte des wissenschaftlichen Interesses als epistemische Dinge«³⁰ heraus: »Man muß sich durch eine komplexe Experimentallandschaft hindurchschlagen, bis sich wissenschaftlich relevante einfache Dinge abzeichnen.«³¹ Forschen als »Prozess der Erkenntnisgewinnung« ist für Rheinberger somit »Arbeit am Wissen«³² und die epistemischen Dinge sind »die Dinge, denen die Anstrengung des Wissens gilt«³³, die also die Forschenden irritieren und ihre wissenschaftliche Neugier wecken. Hervorgebracht werden die epistemischen Dinge in »Austauschbeziehung«³⁴ mit »stabilen Umgebungen, die man als Experimentalbedingungen oder als technische Dinge bezeichnen kann«,³⁵ und die die Dinge des wissenschaftlichen Interesses und ihre »Sichtbarmachung«³⁶ »in doppelter Hinsicht« determinieren: »Sie bilden ihre Umgebung und lassen sie so erst als solche hervortreten, sie begrenzen sie aber auch und schränken sie ein.«³⁷

²⁵ Vgl. Rheinberger 2001.

²⁶ Rheinberger 2001, S. 7f., 21.

²⁷ Rheinberger 2001, S. 9.

²⁸ Rheinberger 2021, S. 19.

²⁹ Rheinberger 2001, S. 9.

³⁰ Hall 2014, S. 228.

³¹ Rheinberger 2001, S. 22f.

³² Rheinberger 2021, S. 231.

³³ Rheinberger 2001, S. 24.

³⁴ Rheinberger 2001, S. 27.

³⁵ Rheinberger 2001, S. 25.

³⁶ Hall 2014, S. 229.

³⁷ Rheinberger 2001, S. 26.

So gesehen liegen der eigentliche Mehrwert, ja der epistemologische Witz digitaler Technologien und Software verstanden als Experimentalsysteme eben gerade *nicht* im automatisierten schnellen Be- und Umrechnen von Daten, sondern in ihrer experimentell-interaktiven Aneignung zur Irritation konventioneller Erkenntnispraktiken des Sinnverstehens und der dabei erfolgenden Konstruktion von Daten als epistemische Dinge. In »Experimentalsystemen«, so Rheinberger, »steht die Interaktion des Experimentators mit seinem Material im Mittelpunkt. [...] Im Zuge der Interaktion mit den Dingen, mit denen man in einem Experiment arbeitet, erwachen diese Materialien selbst zum Leben.«³⁸ Ganz in diesem Sinne hat Arndt Niebisch festgestellt, dass sich geisteswissenschaftliche Forschung mit digitalen Methoden und Techniken gerade nicht auf die »computergestützte Auswertung von Daten« beschränkt, sondern eigentlich das »Experimentieren und Konstruieren von Werkzeugen, die diese Auswertung übernehmen«, im Mittelpunkt steht – verbunden mit »einer Diskussion, wie diese Daten zu interpretieren und zu verwenden sind.«³⁹ In dieser Weise inszeniert etwa das Kollektiv des Stanford Literary Lab rund um Franco Moretti seine digitale Forschung nicht bloß als experimentelle Laborarbeit, sondern sie konzедieren mit ihren literaturwissenschaftlichen Experimentalreihen vielmehr, dass bei diesen »neuen Methoden der Prozess eine fast genauso große Rolle spielt wie die Ergebnisse.«⁴⁰ Ebenso verorten sich die Forscher*innen des Text-Mining-Projekts ePol (Postdemokratie und Neoliberalismus) in einem »quasi-experimentellen Setting« und kennzeichnen den »Austausch über Erfahrungen mit der praktischen Anwendung computerunterstützter Analyseverfahren« und über »das Verhältnis von informationswissenschaftlicher Auswertungskapazität und menschlicher Interpretationsleistung« als zentrales methodologisches Moment ihrer Forschungen.⁴¹ In der Logik des Experimentalsystems geht es im »Spiel der Hervorbringung von Neuem«⁴² in digitalen Forschungskontexten also letztlich um erkenntnistheoretisch *und* technisch auszutrierende Praktiken der »Ausweitung der Aufmerksamkeit«⁴³ bzw. »experimentellen Geist«⁴⁴, mit dem »die Flexibilität des Metamediums Computer möglichst ausgereizt wird«, sodass »immer wieder andere Daten immer wieder anders verrechnet und formatiert« werden, »um völlig unterschiedliche Rekontextualisierungen der Daten zu erzeugen und damit unterschiedliche Interpretationsmöglichkeiten zu eröffnen.«⁴⁵ Im Kontext von informatorisch-technischen Praktiken und Prozessen interaktiv-experimenteller Datenverarbeitung werden so aus Daten eigenartige, flexible epistemische Dinge, wie am Beispiel *digitalisierte Zeitungen* im Folgenden gezeigt wird.

³⁸ Rheinberger 2018, S. 119.

³⁹ Niebisch 2018, S. 12.

⁴⁰ Moretti et al. 2017, S. 17.

⁴¹ Lemke / Wiedemann 2016, S. V.

⁴² Rheinberger 2001, S. 27.

⁴³ Schruhl 2019, S. 6.

⁴⁴ Rheinberger 2018.

⁴⁵ Bubenhofer / Rothenhäusler 2016, S. 64.

4. Digitalisierte Zeitungen als technisch-epistemische Assemblagen

Mit der digitalen Bereitstellung und Verfügbarkeit von Zeitungen hat sich ebenso für die Zeitungsforschung eine heterogene, vielfältige und kaum mehr zu überblickende medientechnische Landschaft herausgebildet. So stellt Huub Wijfjes kritisch zur Vielfalt der Zeitungsportale fest: »[A]ll kinds of specialised – regional, local, thematic – collections pop up in the online world. Each of these collections can make use of specific interfaces, standards and/or tariffs for accessibility and use.«⁴⁶ Zudem ist man mittlerweile mit mehreren »interface generations«⁴⁷ von Zeitungsportalen konfrontiert. Neben den als digitale Archive bzw. Sammlungen konzipierten Angeboten finden sich inzwischen Plattformen wie [Gale Digital Scholars Lab](#), [Impresso](#) oder [NewsEye.EU](#), die explizit als virtuelle Forschungsumgebungen entwickelt werden und digitale Forschungsinstrumente und Methoden integrieren.

Vor diesem Hintergrund gestalten sich digitalisierte Zeitungen, *definiert, fixiert und umgrenzt* durch die »technischen und instrumentellen Bedingungen«⁴⁸ der jeweiligen digitalen Umgebung, als äußerst diverse, dynamische und flexible digitale Objekte, die auf ein weites Spektrum an Datenniveaus, digitalen Formatierungen und Visualisierungen verweisen. Zeitungen werden schon im Prozess der Zeitungsdigitalisierung nicht einfach nur digitalisiert, sondern sie werden als digitale Objekte konstruiert und als solche wiederum für unterschiedlichste Zwecke modelliert und bereitgestellt. Das heißt, die variablen institutionellen und technischen Bedingungen und Ziele ihrer Prozessierung in einer digitalen Umgebung beeinflussen, in welcher Gestalt und Ordnung das epistemische Ding *digitalisierte Zeitung* letztlich auftritt. Digitalisierte Zeitungen existieren daher nicht aufgrund ihrer bloßen Digitalisierung, sondern erst infolge ihrer jeweiligen komplexen digitalen Verarbeitung. Bob Nicholson hat diese ebenso voraussetzungsreichen wie vielschichtigen Prozesse der Digitalisierung und digitalen Edition von Zeitungen entsprechend ausdrücklich als Medienwechsel bzw. *Remediatisierung* gekennzeichnet: Zeitungen werden im Zuge der Transformation ins Digitale »remediated« and not just reproduced«⁴⁹ und zu »complex objects determined by multiple layers of processing and datafication«⁵⁰. Sie werden zu Einsen und Nullen *verdatet*, werden zu Dateien, Dokumenten und digitalen Korpora, kurz: zu »in Form gebrachter Information«⁵¹ als Grundlage und Voraussetzung für maschinelle Lesbarkeit, Verarbeitung und Visualisierung. Mithilfe von Algorithmen, Code, Dateiformaten und Interfaces als technischen Dingen werden digitalisierte Zeitungen somit nicht allein darstellbar und sichtbar, sondern sie werden von diesen zugleich als originäre epistemische Dinge in Form gebracht, präsentiert, stabilisiert und strukturiert. Folgt man Rheinberger, so entfalten Instrumente »ihre wissenschaftliche Wirkungsmächtigkeit nur als Ermöglichungsbedingungen bestimmter Darstellungsformen, als konstituierende Elemente

⁴⁶ Wijfjes 2017, S. 14.

⁴⁷ Ehrmann et al. 2019, S. 10.

⁴⁸ Rheinberger 2001, S. 29.

⁴⁹ Nicholson 2013, S. 64.

⁵⁰ Ehrmann et al. 2019, S. 2.

⁵¹ Fickers 2016.

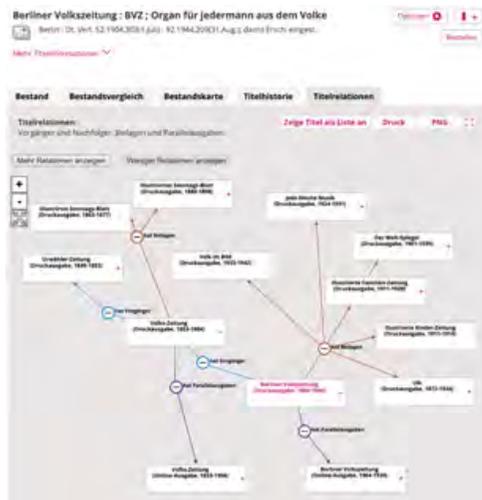


Abb. 2: Visualisierung von Titelrelationen in der ZDB am Beispiel der Berliner Volkszeitung. [Zeitschriftendatenbank 2022]

Genauso sind die mittels Leseinterfaces in Zeitungsportalen präsentierten Image-Digitalisate und Volltexte von Zeitungen technisch-epistemische Assemblagen mit wesentlichen methodologischen Folgen für die Forschungspraxis. Obwohl die in vielen Zeitungsportalen ein- und umgesetzten Reader in Design und Funktionalität auf den ersten Blick scheinbar nur analoge Zeitungen als »materielle Leseobjekte hinsichtlich ihrer typographischen Dispositive und Navigation (sequentielles Blättern) imitieren«, sind diese digital konstruierten Zeitungen weit mehr als lediglich »digitale Klone analoger Vorbilder«. ⁵⁸ Vielmehr überdeckt die von den Portalen suggerierte Anzeige und Nutzung eines digitalen Duplikats die quellenkritische Frage nach der originären Medialität digitalisierter Zeitungen, wie sie aus ihrer Remediation als digitale Objekte resultiert. So nivelliert die gleichförmige Anzeige der Viewer u. a. Format, Layout und Materialität der Originalvorlagen ebenso wie sie die vielschichtige Digitalität der Digitalisate verschleiert. ⁵⁹ Denn hinter der Oberfläche der Viewer wird das Zeitungsdigitalisat ja informatorisch-technisch aus unterschiedlichen digitalen Formaten und Schichten (Images, Metadaten, Volltexte) kompiliert und ergibt so erst die hochauflösende, volltexterschlossene, zoombare digitalisierte Zeitung wie sie beispielsweise im [Europeana Newspapers Project](#) modelliert wurde (vgl. Abbildung 3). Digitalisierte Zeitungen werden zudem in Portalen in einer neuen digitalen epistemischen Ordnung organisiert, die die konventionelle lineare Recherche- und Sichtungslogik von Zeitungen nach dem *Top-Down-Prinzip* ergänzt. Neben der digital simulierten analogen Druckordnung nach einzelnen Ausgaben und Jahrgängen werden in digitalen Sammlungen und im Volltext erschlossene Zeitungen zusätzlich *Bottom-Up*, d. h. vollkommen quer zur Ordnung der ursprünglichen Zeitungsoriginale durchsuch- und flexibel nutzbar (vgl. Abbildung 4). ⁶⁰

⁵⁸ Kuhn / Hagenhoff 2017, S. 35.

⁵⁹ Vgl. Blome 2018, S. 12.

⁶⁰ Vgl. Nicholson 2013, S. 66f.



Abb. 3: Kompilation, Modellierung und Ordnung digitalisierter Zeitungen im Europeana Newspapers Portal. [Fedasenka et al. 2014]

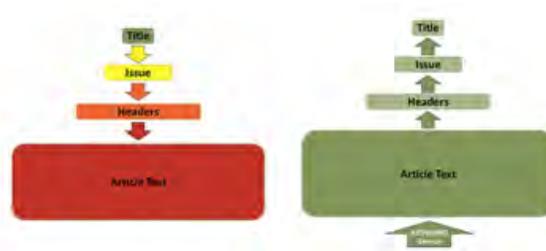


Abb. 4: Konventionelle Recherche- und Sichtungslogik von Zeitungen nach dem Top-Down-Prinzip und Bottom-Up-Volltextsuche in Zeitungsportalen. [Nicholson 2013, S. 66f.]

5. Zeitungsportale als ›Experimentalsysteme‹ der Zeitungsforschung

Zeitungsportale geben der Erzeugung und Konstruktion von digitalisierten Zeitungen als technisch-epistemische Assemblagen mithin mehr als einen bloßen informatorisch-technischen Rahmen. Zeitungsportale sind vielmehr interaktionsabhängige und infolgedessen offene digitale »Räume der Darstellung«⁶¹ und »Hervorbringung von neuen Wissensspuren«⁶², die Forschende in komplexe Prozesse der Formung und Ko-Konstruktion von Zeitungen als digitale Objekte involvieren. Sicher, »Daten entstehen in Prozessen«⁶³ – doch diese Prozesse determinieren nicht die Gestalt und Ordnung der Daten, sondern in diese Prozesse ist die*der Forschende selbst performativ mit eingebunden, insofern sie*er diese Daten mit Modifikationen der informatorisch-technischen Parameter selbst beeinflussen, formen und manipulieren kann.

Folglich definieren und konstituieren Zeitungsportale für Forschende ebenso eigentümliche wie höchst heterogene »epistemische Räume«⁶⁴ für die datenbasierte experimentelle Interaktion mit digitalisierten Zeitungssammlungen und den hiermit verbundenen »Prozess

⁶¹ Rheinberger 2001, S. 109.

⁶² Rheinberger 2001, S. 115.

⁶³ Couldry / Hepp 2021, S. 85.

⁶⁴ Rheinberger 2018, S. 125.

des Formens«⁶⁵ von digitalisierten Zeitungen – als »der dynamischen, relationalen und fragilen Natur von Daten«⁶⁶ – entsprechende epistemische Dinge. Zwar markieren und setzen die konkrete technische Architektur, Ausstattung und Umsetzung des jeweiligen Portals »den Horizont und die Grenzen«⁶⁷ potenziell möglicher Interaktionen, trotzdem liegt es in der Hand der*des Forschenden mit ihrer*seiner experimentellen Kreativität, diese Grenzen jeweils auszuloten, auszureizen oder sogar zu überschreiten.

Somit sind Zeitungsportale viel mehr als einfache »gateways«⁶⁸ der datenhaltenden- und -gebenden Institutionen für den Zugang zu digitalisierten Zeitungssammlungen und Zeitungen. Vielmehr gestalten sie mit ihren *user interfaces*, deren Design und Funktionsspektrum für Datenoperationen, und mit den jeweiligen institutionellen, rechtlichen und wirtschaftlichen Konditionen der Portalpolitik aktiv die »user workflows by offering different selections of tools and features for searching and exploring content«⁶⁹ mit und präkonfigurieren so die Praktiken, Prozesse und Situationen zwischen Daten, Portal und User. Vor diesem Hintergrund kommt Paul Gooding in seinen *user studies* zur Nutzung von Zeitungsportalen zu dem Schluss, dass gerade die von den Anwender*innen viel bemängelte *usability* oftmals die Annehmlichkeiten und Vorteile, digitalisierte Zeitungen jetzt »broadly, deeply and quickly«⁷⁰ zu nutzen, sowie die uns speziell interessierenden experimentellen Spielräume und Potenziale der Portale von vornherein entscheidend einschränkt und limitiert:

»Users of large-scale digitised collections find difficulties in using search functionality to return accurate results [...]. They also face the enforcement of a particular type of search and browsing behaviour, created by inflexible interfaces, which force all users into a narrow range of interactions.«⁷¹

Weiterhin kritisiert Astrid Blome, dass digitalisierte Zeitungen in einer enorm heterogenen Portallandschaft in unterschiedlichsten »Darstellungsformen« und »Erschließungstiefen« vorgehalten werden, was nicht nur »dem wissenschaftlichen Bedürfnis nach einer validen Datenerhebung auf einheitlicher Grundlage« entgegensteht, sondern von den Forschenden zudem »flexible« und »kontinuierlich und individuell zu modifizierende Suchstrategien« erfordert.⁷² Eine umfassende Evaluation der Funktionalitäten und Features einschlägiger Zeitungsportale, die jüngst eine Forscher*innengruppe um Maud Ehrmann unternommen hat, führt dies eindrücklich vor Augen (vgl. Abbildung 5). Aus forschungspraktischer Sicht resümiert Huub Wijfjes die sich hieraus ergebenden erheblichen Herausforderungen für die Zeitungsforschung bei der Nutzung von Zeitungsportalen recht treffend:

⁶⁵ Rheinberger 2018, S. 119.

⁶⁶ Fickers 2020.

⁶⁷ Rheinberger 2001, S. 25.

⁶⁸ Ehrmann et al. 2019, S. 2.

⁶⁹ Ehrmann et al. 2019, S. 2.

⁷⁰ Blome 2018, S. 12.

⁷¹ Gooding 2017, S. 182.

⁷² Blome 2018, S. 33.

»[A] researcher who wants to work with complete newspaper data needs to be able to organise, improvise and negotiate. [...] Last but not least, a researcher needs to realise that good preparation is more than half of the work; it is almost all of the work.«⁷³

Für den produktiven Umgang mit solchen aus der Vielfalt und Vielgestaltigkeit der Zeitungsportale resultierenden Problemen scheint es somit hilfreich, entlang der informatorisch-technischen Portalevolution und der vorrangigen Portallogik (Archivieren, Forschen, Recherchieren und Suchen) verschiedene Portaltypen zu beschreiben, die sich systematisch in Interfacedesign und -konzeption sowie Funktionalität, Interaktivität und Usability unterscheiden.⁷⁴ In der Forschungspraxis sind es ja gerade die jeweiligen Portaleigenschaften, die unmittelbar Erkenntnispraktiken und -prozesse stimulieren und materiell, performativ und technisch »die experimentellen Formen der Praxis, durch die sich diese Erkenntnisprozesse ausbilden und in denen sie sich vollziehen«⁷⁵, vorstrukturieren. Im Hinblick auf solche portaldifferenten experimentellen Anregungen, Effekte, Praktiken und Potenziale werden nun im Folgenden drei Portaltypen genauer betrachtet und charakterisiert: Digitale Archive bzw. Sammlungen, Zeitungsportale als *reading machines* und *virtual research environments*.

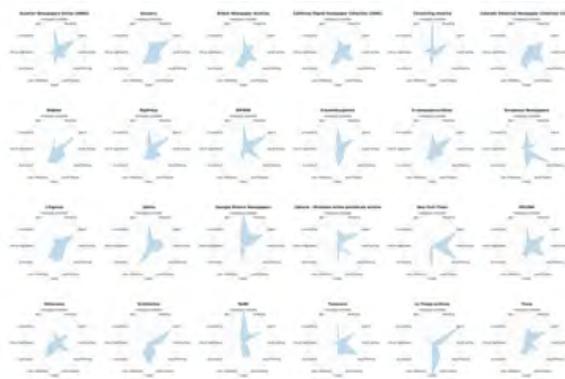


Abb. 5: Evaluation der Funktionalitäten und Features einschlägiger Zeitungsportale. [Ehrmann et al. 2019, S. 9]

5.1 Digitale Zeitungsarchive und -sammlungen: Experimentelle Effekte ›kleiner Archive‹

Digitale Zeitungsarchive und -sammlungen fokussieren mit ihrem Design und ihrer Funktionalität auf den »access«⁷⁶ sowie basale digitale Interfaces und Instrumente zum vorrangigen Zweck der Präsentation, Recherche und Sichtung von digitalisierten Zeitungen als Faksimiles in Form von Image-Digitalisaten. Lediglich einfache Filter wie Kalender oder Titelfilter erleichtern Auswahl und Zugang. Den Anwender*innen präsentieren die für das konkrete digitale Objekt Zeitung unspezifisch konzipierten Leseinterfaces dieses Portaltypus

⁷³ Wijfjes 2017, S. 20.

⁷⁴ Ehrmann et al. 2019; Pfanzelter et al. 2021.

⁷⁵ Rheinberger 2021, S. 8.

⁷⁶ Ehrmann et al. 2019, S. 10.

Zeitungen als »materielle Leseobjekte in Form digitaler Klone analoger Vorbilder«, d. h., sie *imitieren* quasi Zeitungen als »materielle Leseobjekte hinsichtlich ihrer typographischen Dispositive und ihrer Navigation (sequentielles Blättern)«. ⁷⁷

Insbesondere die deutsche Portallandschaft folgt von wenigen Ausnahmen abgesehen diesem Portalprinzip des digitalen Lesesaals, das den kostengünstigsten, technisch einfachsten und schnellsten Weg darstellt, Zeitungen digital bereitzustellen und online verfügbar zu machen. Dennoch impliziert diese recht simple Darstellungs-, Lese- und Speicherform für digitalisierte Zeitungen einen interessanten epistemischen Effekt. Madleen Podewski hat darauf hingewiesen, dass solche Portale primär nach der Logik »kleiner Archive« operieren und deren Wissensordnung analog zu »konkret-materiellen, periodisch aufeinander folgenden Heften geschieht, die sich einzeln in die Hand nehmen und durchblättern lassen«. ⁷⁸ Damit komprimieren sie das massenhaft vorliegende und sehr sperrige Materialobjekt Zeitung und konservieren es digital als genuines »physical, typographical, and information artifact« ⁷⁹, das erkenntnisperspektivisch wesentliche Elemente und Relationen der »semantisch-materiell-visuellen Mehrdimensionalität« ⁸⁰ von Zeitungen aufzeigt: »visuelle in der Seiten- und Heftgestaltung, periodisch rhythmisierte in der Erscheinungsweise, numerisch oder alphabetisch strukturierte in Rubrizierungsverfahren, elementare und komplexe Bedeutungsbeziehungen zwischen den jeweiligen Bild- und Texteinheiten.« ⁸¹ Zeitungsportale als »kleine Archive« lenken auf diese Weise die epistemische Aufmerksamkeit abseits der Inhalte vielmehr auf die journalistischen und redaktionellen »Aufbereitungsroutinen für die wechselnden Inhalte« ⁸², auf Aufmachung, Design, Layout, Rubriken und verschiedene weitere Momente periodischer und visueller »mediumspezifischer Ordnung« ⁸³ wie Erscheinungsrhythmus, Format, Illustrierung oder Umfang – also Aspekte, die Kevin G. Barnhurst und John Nerone in der Frage nach der »form« bzw. dem »look« von Zeitungen zusammengefasst haben:

»Form includes the things that are traditionally labeled layout and design and typography; but it also includes habits of illustration, genres of reportage, and schemes of departmentalization. Form is everything a newspaper does to present the look of the news. [...] Put another way, the form includes the way the medium imagines itself to be and to act. In its physical arrangement, structure, and format, a newspaper reiterates an ideal for itself.« ⁸⁴

Problematisch ist, dass in »kleinen Archiven« die möglichen Datenoperationen oft auf den einfachen Dateixport der Image-Digitalisate einzelner Ausgaben, Nummern oder Seiten limitiert sowie für die als Image-Digitalisat remedialisierte digitalisierte Zeitung zudem selten hinreichende und valide Metadaten zur Form des digitalisierten Objekts hinterlegt sind. Entsprechend ist der Einsatz visueller Analyseinstrumente und -perspektiven für

⁷⁷ Kuhn / Hagenhoff 2017, S. 35.

⁷⁸ Podewski 2018, S. 6.

⁷⁹ Liddle 2012, S. 231.

⁸⁰ Podewski 2018, S. 9.

⁸¹ Podewski 2018, S. 7.

⁸² Bucher 2016, S. 285.

⁸³ Podewski 2018, S. 6.

⁸⁴ Barnhurst / Nerone 2001, S. 3.

die über diesen Portaltyp verfügbaren Zeitungssammlungen keineswegs ausgelotet. Eine methodische Option, um solche Sammlungen vermehrt auf visuellen Wegen zu erkunden, ist die Implementierung von Prozeduren und Prozessen aus dem Bereich der von Lev Manovich etablierten »media visualization«, mit der korpusbasiert große Mengen an Bildern und Bilddaten automatisiert verarbeitet und in Form von Plots visualisiert werden⁸⁵: »Conceptually media visualization is based on three operations: zooming out to see the whole collection (image montage), temporal and spatial sampling, and remapping (rearranging the samples of media in new configurations).«⁸⁶ »This allows us«, so Manovich über die Erkenntnispotenziale dieser Methode, »to use the full capacities of our perception and cognition to compare these images and notice all kinds of patterns and details that manual text tags or numerical features may not capture.«⁸⁷ Kleinere Experimente, die im Kontext von Manovichs Cultural Analytics Labs bereits mit einzelnen großen Image-Korpora von Zeitschriften und Zeitungen unternommen wurden, geben vielversprechende Einblicke in das enorm anregende Potenzial dieses visuellen »distant viewing«⁸⁸ für die Zeitungsforschung. Die Abbildungen 6 und Abbildung 7 illustrieren dies beispielhaft anhand von Plots der Titelseiten des *Hawaiian Star* und des *Time* Magazins. Die mithilfe eines *Imageplotters* berechneten und erstellten Visualisierungen geben auf verschiedenen Ebenen Einsichten in einen reichhaltigen Formwandel der Periodika. Die großformatigen Übersichtsplots, die lediglich auf der einfachen Berechnung von Bildeigenschaften wie Farbton, Graustufen oder Helligkeit beruhen, irritieren auf den ersten Blick mit der ganzen Vielfalt dieses Wandels in Aufmachung, Drucktechnik, Layout und visueller Kultur. Erkundet man sie im Weiteren genauer, greift einzelne Momente dieses Wandels heraus oder zoomt in sie hinein, so lassen sich explorativ und intuitiv in den visuellen Mustern und Schichten unterschiedliche Phasen, Perioden und Strukturen dieses Wandels entdecken, denen man in vertiefenden Forschungen weiter nachgehen kann.

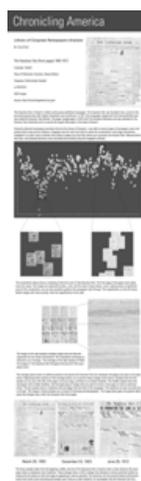


Abb. 6: The Hawaiian Star Front Pages 1893–1912. [Manovich / Kiani 2012]

⁸⁵ Manovich 2020, S. 223–244.

⁸⁶ Manovich 2020, S. 244.

⁸⁷ Manovich 2020, S. 222.

⁸⁸ Podewski 2018, S. 12.



Abb. 7: Time Magazine Covers 1923–2009. [Manovich / Douglass 2016]

5.2 Zeitungen im Volltext: Zeitungsportale als ›reading machines‹

Zeitungsportale, die digitalisierte Zeitungen volltexterschlossen interaktiv durchsuchbar machen, lassen sich mit Stephen Ramsay als »reading machines« klassifizieren, die mit ihren digitalen Recherche- und Suchfunktionen genuin neue Praktiken und Perspektiven der Textaneignung, -erschließung und -sichtung entfalten und erlauben.⁸⁹ Franco Moretti hat diesen informatorisch-technisch basierten, maschinellen Umgang mit digital(isiert)en Texten schon vor Ramsay pointiert mit dem Konzept des *distant reading* markiert.⁹⁰ Unter dem provokanten Motto »we know how to read texts, now let's learn how not to read them«⁹¹ hat Moretti eine digitale Textpraxis vorgezeichnet, die sich nicht mehr direkt, exklusiv und gründlich mit einigen wenigen Texten auseinandersetzt, sondern die die Verarbeitung vieler Texte zum Ziel hat. Im Gegensatz zum hermeneutisch-verstehenden, textnahen *close reading* werden hierfür beim *distant reading* die Leseprozesse in unterschiedlicher Weise an Computer und Software delegiert, was Forschenden hilft, sehr große Mengen Text sehr schnell *zu lesen*, und zugleich eine neue »spezifische Form«⁹² der Erkenntnis von Texten ermöglicht: »Umriss, Beziehungen und Strukturen werden so deutlich, Formen, letztlich Modelle.«⁹³

Nach dem Start der ersten Zeitungsportale setzte in der Zeitungsforschung sogleich eine eingehende und weitsichtige Diskussion über diese nachhaltigen epistemischen Effekte und Folgen für in digitalen Sammlungen *und* im Volltext erschlossene digitalisierte Zeitungen ein. In einem der frühen Diskussionsbeiträge hat Bob Nicholson die »ability to search large archives

⁸⁹ Vgl. Ramsay 2011.

⁹⁰ Vgl. Moretti 2000; Moretti 2016.

⁹¹ Moretti 2000, S. 57.

⁹² Moretti 2009, S. 7.

⁹³ Moretti 2009, S. 7.

of print culture for individual words or phrases« sogar als »central feature« für einen »digital turn« in der Zeitungsforschung herausgestellt,⁹⁴ das den Umgang mit Zeitungen geradezu revolutioniert:

»Classically [...], when we want to find something in a conventional newspaper we tend to approach it using a top-down approach. [...] Thanks to OCR technology and keyword search engines, we can now access this bottom level of text directly [...]. The search for information in newspapers need no longer be driven solely by top-down reading, but by the generation of new keywords.«⁹⁵

Nicht zuletzt hat die Erschließung und Verfügbarkeit von Zeitungen als digital integrierter, interaktiv durchsuchbarer Volltext wesentliche epistemische Folgen für den Quellenwert von Zeitungen. In Zeitungen werden Ausgabe für Ausgabe, Jahrgang für Jahrgang, Seite für Seite über sehr lange Zeiträume aktuell und unmittelbar gesellschaftliche Diskurse, Ereignisse und Entwicklungen in ihrer ganzen Breite und Vielfalt abgelagert, gespeichert und überliefert. Vor diesem Hintergrund operieren volltexterschließende Zeitungsportale, so Nicholson, nun gewissermaßen als »living dictionary, one which allowed to access the minutiae of everyday cultural discourse and track the complex ways in which words and phrases circulated, how their meanings evolved and how their usage changed on an almost daily basis«.⁹⁶ Neben dem offensichtlichen Komfort einer effektiven und schnellen Recherche und Sichtung eröffnen sie für Forschende somit ganz neue experimentelle Möglichkeiten und Perspektiven für die diachrone und synchrone Erforschung der medialen Evolution, Semantik, Trends und Strukturen von gesellschaftlichen und öffentlichen Diskurse: »They can make comparisons over time, between papers and between different media forms, to isolate innovations and specificities and to identify similarities, continuities and borrowings.«⁹⁷ So lässt sich beispielsweise in einzelnen bzw. vergleichend zwischen verschiedenen Zeitungen (eines Landes, einer Region, weltweit) der Intensität und den Mustern journalistischer Berichterstattung, den Konjunkturen, Zitationen und Zirkulationen von Nachrichten oder dem medialen »life-cycle«⁹⁸ von Medienereignissen oder Themenkarrieren nachspüren. In *Abbildung 8* ist z. B. das Interface der *U.S. News Map* zu sehen, das mittels einer Stichwortsuche in den Volltexten von *Chronicling America* die Häufigkeit von Begriffen und Themen in geografischer Verteilung und im zeitlichen Trend darstellt. Und *Abbildung 9* zeigt das Projekt *Viral Text*, das unter Verwendung der Ressourcen von *Chronicling America* und *Making of America* Netzwerke des Übernehmens, Teilens und Zirkulierens von Nachrichten und Texten zwischen Zeitschriften und Zeitungen sichtbar macht.

⁹⁴ Nicholson 2013, S. 62.

⁹⁵ Nicholson 2013, S. 66f.

⁹⁶ Bingham 2010, S. 70f.

⁹⁷ Bingham 2010, S. 228.

⁹⁸ Bingham 2010, S. 229.



Abb. 8: U.S. News Map. [Toon 2016]

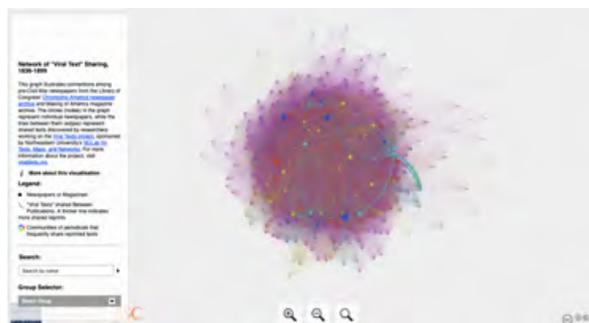


Abb. 9: Network of Viral Text Sharing 1836–1899. [Cordell / Smith 2017]

Trotz der schon frühen Einsichten in die immensen Potenziale volltexterschließender Zeitungsportale seitens der Zeitungsforschung werden diese bislang leider nur in wenigen und zudem kommerziellen Ausnahmen wie [Gale](#), [ProQuest](#) oder [RetroNews](#) ausgeschöpft. Stattdessen ist für sehr viele der zumeist öffentlich geförderten Zeitungsportale nach wie vor »a gap between growing user expectations and current interface capacities«⁹⁹ festzustellen, der sich vornehmlich hinsichtlich kaum ausgeprägter und portalübergreifend höchst unterschiedlicher Möglichkeiten der Ergebnisfilterung und -selektion sowie oftmals teils komplizierter, teils restriktiver Optionen des Exports und der Weiterverwertung von Suchresultaten offenbart. Für den gelegentlichen Gebrauch mag diese hinreichen, für Forschende, die auf den Skalierungsebenen von *big data* und *long data* mit Zeitungsvolltexten forschen wollen, heißt dies praktisch, von Zeitungsportal zu Zeitungsportal individuelle Forschungsworkflows zu konzipieren, um überhaupt Volltexte zu generieren, und für die weitere Verarbeitung der Volltexte auf den Einsatz und die Nutzung externer Tools zurückzugreifen.¹⁰⁰

5.3 Zeitungsportale als ›virtual research environments‹

Zeitungsportale als *virtual research environments* sind nun Lösungen für das integrierte Forschen mit digitalisierten Zeitungen, wobei hier gegenwärtig noch vorzugsweise der Volltext im Vordergrund steht. Als explizit virtuelle Forschungsumgebungen vereinen solche Portale »data services, digital materials, tools, and an environment for research practices and collaboration«¹⁰¹ und beschränken sich folglich nicht mehr nur auf die interaktive

⁹⁹ Pfanzelter et al. 2021, S. 2.

¹⁰⁰ Vgl. Koenen 2018, S. 550–553; Pfanzelter et al. 2021, S. 5–22.

¹⁰¹ Pawlicka-Deger 2020, S. 61.

Durchsuchbarkeit von und den Zugang zu digitalisierten Zeitungen, sondern koppeln diese Funktionalitäten mit weiteren Features zur Filterung, Kuratierung und Speicherung von Recherche- und Suchresultaten sowie spezifischen Instrumenten und Methoden für *Text Mining*. Mit diesem Typus von Portalen ergeben sich für die Zeitungsforschung somit »more creative and nuanced approaches«¹⁰², was die Möglichkeiten der Forschung mit digitalisierten Zeitungen merklich in Richtung einer dateninteraktiven, experimentellen Forschung verschiebt. Das heißt, die in solche Portale implementierten vielseitigen Funktionen, Methoden und Tools begleiten und erleichtern nicht bloß die Forschung mit informatorisch-technischen Mitteln. Sie operieren vielmehr genuin in der Logik eines experimentell-wissensproduzierenden Forschungsstils, dem ein »spielerischer Modus heuristischen Tastens«¹⁰³ eigen ist, der digitale Forschung eigentlich erst kreativ macht und so den digitalen Forschungsprozess permanent produktiv vorantreibt.

Ein Beispiel für ein solches *virtual lab* ist *DiaCollo*¹⁰⁴, »ein Browser-basiertes Werkzeug zur Analyse und Visualisierung von Worthäufigkeiten und Kollokationen im zeitlichen Verlauf«¹⁰⁵, mit dem einzelne sorgfältig kuratierte Zeitungsvolltextkorpora mit digitalen *lexikometrischen* Textwerkzeugen ausgewertet und untersucht werden können (vgl. Abbildung 10). Neben den niedrigen *Einstiegshürden* eignet sich diese explorative Forschungsschnittstelle für einfache »interaktive Experimente« vor allem wegen der »Echtzeitdarstellung der Ergebnisse« und der Rückkopplung jedes Experiments an die Fundstellen im Korpus, wodurch »ein stetiger Wechsel zwischen Distant und Close Reading möglich wird.«¹⁰⁶ Mit einer Experimentierreihe zum medialen Sprachgebrauch der DDR unter Verwendung des DDR-Zeitungskorpus wurde die Funktionsweise dieser Schnittstelle insbesondere für eine »digital informierte Historische Semantik«¹⁰⁷ und die Frage nach diskursiven »Thematisierungsereignissen und -konjunkturen«¹⁰⁸ sehr intensiv getestet. So wurde gezeigt, dass deren experimentelles Potenzial vornehmlich darin liegt, herauszufinden, »wann eine Sache in Gestalt eines Wortes bzw. eines Begriffs zum Thema und damit höchst wahrscheinlich »problematisiert« wurde« und wie sich langfristig die Verwendung im gesellschaftlichen Diskurs etabliert, routinisiert und weiterentwickelt¹⁰⁹: »Was sich effizienter und zugleich konkreter als bisher beschreiben lässt, sind Konjunkturen und Ritualisierungen hochfrequenter Begriffe und Wortzusammenhänge [...]. Hinzu tritt die Möglichkeit, den Wandel semantischer Felder und Begriffsnetze zu visualisieren.«¹¹⁰

¹⁰² Nicholson 2013, S. 71.

¹⁰³ Fickers 2019.

¹⁰⁴ Vgl. Jurish 2015.

¹⁰⁵ Burckhardt et al. 2019, S. 179.

¹⁰⁶ Burckhardt et al. 2019, S. 183, 194.

¹⁰⁷ Burckhardt et al. 2019, S. 178.

¹⁰⁸ Sarasin 2012, S. 163.

¹⁰⁹ Sarasin 2012, S. 163.

¹¹⁰ Burckhardt et al. 2019, S. 178.



Abb. 10: DiaCollo – Die Zeit-Korpus: Kollokationen von »Klima« im zeitlichen Verlauf. [Koenen 2022]

Im Vergleich zu DiaCollo als lediglich interaktiv nutzbarer Forschungsschnittstelle zu vorkuratierten Korpora gehen die beiden jüngsten virtual labs unter den Zeitungsportalen *Impresso* (vgl. Abbildung 11) und *NewsEye.EU* (vgl. Abbildung 12) noch einen großen Schritt weiter. Trotz vieler Unterschiede im Design, den Funktionalitäten und der Umsetzung haben beide zum Ziel, die Datenbereitstellung von digitalisierten Zeitungen und den gesamten Forschungsworkflow mit digitalisierten Zeitungen samt spezieller Forschungsinstrumente in einem Interface informatorisch-technisch und methodologisch transparent zu integrieren und Forschenden hiermit sowohl kollaboratives Forschen wie personalisierte Forschungsworkspaces für ein möglichst freies und vielseitiges Experimentieren zu ermöglichen. So wirbt das Projekt *Impresso* für sein Portal: »The *Impresso* app is a full-fledged, in production newspaper interface with powerful search, filter and discovery functionalities based on semantic enrichments together with experimental contrastive views.«

¹¹¹ Die Flexibilität solcher Portalinterfaces ist jedoch nicht bloß praktisch, sie hat zugleich weitreichende epistemologische und methodologische Konsequenzen. Denn mit der flexibel justierbaren »configurability« ¹¹² der Interfaces werden Forschende nun kontinuierlich interaktiv in den dynamischen »Prozess des Formens« ¹¹³ von digitalisierten Zeitungen als epistemische Objekte eingebunden – eine *configurability*, der sich Forschende »selbstreferentiell« ¹¹⁴ bewusst sein müssen, weil sie diese in eine erkenntnisproduktive »Mikrodynamik« ¹¹⁵ aus Schleifen, »Subroutinen«, »Verkettungen« und »Verzweigungen«, ¹¹⁶ einbindet und stetig herausfordert, den Forschungsworkflow in Gang zu halten, zu kontrollieren, zu regulieren und zu steuern: »[H]ow can we adjust our measures, tools, data set and so on, to serve our research design and our methodological ambitions better?« ¹¹⁷ In diesem Kontext sind u. a. nicht zuletzt die Gefahren und Verheißungen der angebotenen, leicht nutzbaren digitalen Textmethoden immer unter der Maßgabe kritisch zu reflektieren, dass es eben keine »best method for automated text analysis« ¹¹⁸ gibt und somit der Weg zu den Resultaten und die Resultate selbst stets zu validieren sind.

¹¹¹ Romanello et al. 2020.

¹¹² Marres 2017, S. 113.

¹¹³ Scheuermann 2016, S. 61.

¹¹⁴ Rheinberger 2018, S. 119.

¹¹⁵ Rheinberger 2018, S. 148.

¹¹⁶ Rheinberger 2018, S. 148f.

¹¹⁷ Marres 2017, S. 113.

¹¹⁸ Grimmer / Stewart 2013, S. 269.

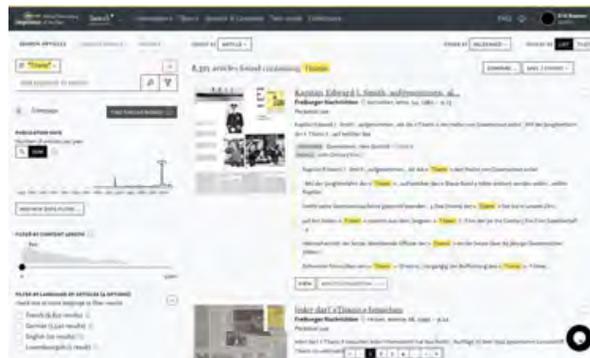


Abb. 11: Impresso-Interface. [Koenen 2022]

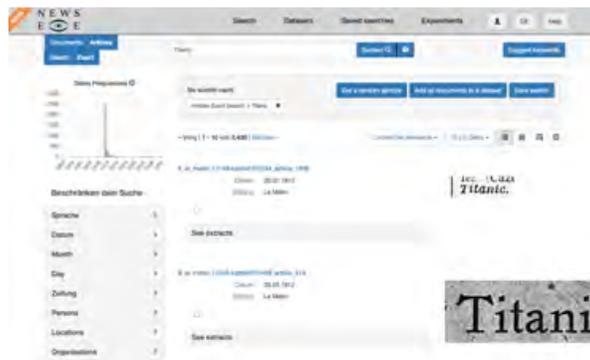


Abb. 12: NewsEye.EU-Interface. [Koenen 2022]

In dieser Weise haben sich beide Portale einem ko-kreativen Entwicklungsansatz und prinzipieller Entwicklungsoffenheit verschrieben. Hierbei ist ein wesentliches Moment, Interfacedesign, Funktionalität und Usability in der experimentellen Praxis ebenso von den Forschenden selbst zu evaluieren zu lassen. Mit dieser grundlegenden User*innen-sensiblen Orientierung an den Erfahrungen und Interessen von Forschenden liegt das schöpferische Potenzial dieser Portale für die künftige Zeitungsforschung insofern insbesondere in der kontinuierlichen und reflexiven Anpassung, Pflege und Weiterentwicklung nachhaltiger digitaler Forschungsinfrastrukturen und -umgebungen mit dem Ziel einer »integrative practice«¹¹⁹ des Experimentierens und Forschens mit digitalisierten Zeitungen.

6. Resümee und Schlussfolgerungen

Mit digitalisierten Zeitungen als digitalen Forschungsobjekten und Zeitungsportalen als digitalen Forschungskontexten ergeben sich gegenwärtig und für die Zukunft immer mehr neue Möglichkeiten einer innovativen und kreativen Forschung mit Zeitungen, deren epistemologische und methodologische Effekte, Folgen und Potenziale jedoch systematisch ebenso fachwissenschaftlich wie informatorisch-technisch bedacht und kritisch hinterfragt werden müssen. Vom Standpunkt der hier verfolgten Experimental-Epistemologie Hans-Jörg Rheinbergers konstituieren Zeitungsportale für eine solche »creative research«¹²⁰

¹¹⁹ Oberbichler et al. 2020.

¹²⁰ Niebisch 2018, S. 12.

gleichermaßen eigentümliche wie höchst heterogene, laborartige »epistemische Räume«¹²¹ für die experimentelle Interaktion mit digitalisierten Zeitungen und den hiermit verbundenen »Prozess des Formens«¹²² von digitalisierten Zeitungen als – »der dynamischen, relationalen und fragilen Natur von Daten«¹²³ – entsprechende epistemische Objekte. Das heißt ganz praktisch: Forschende gehen nicht mehr bloß mit Fragen und Forschungsinteressen an Zeitungen als Quellen heran, sondern sie treten mit der Nutzung von digitalisierten Zeitungen in Zeitungsportalen in »ein neues Handlungsfeld, in dem Datenproduzenten, Datenhüter und Datennutzer auf je eigene Weise in die Konstruktion des epistemischen Objekts ›Quelle‹ eingebunden sind.«¹²⁴ Vor allem durch ihre spezifische »semantisch-materiell-visuelle Mehrdimensionalität«¹²⁵, die zudem »massenhaft, variantenreich und periodisch prozessiert wird«¹²⁶, sind Zeitungen in Kontexten des Digitalen freilich bemerkenswert herausforderungsreiche Objekte, die hohe Anforderungen an ihre informatorische Erzeugung, Konstruktion und Reflexion als originäre technisch-epistemische Assemblagen stellen.

Unter dieser Voraussetzung fokussieren die einzelnen vorgestellten Typen von Zeitungsportalen bis dato indes lediglich einzelne Aspekte dieser Mehrdimensionalität und Prozessierung und markieren und rahmen so in der Folge für Forschende epistemisch deutlich begrenzte unterschiedliche experimentelle Kontexte für die Forschungspraxis mit digitalisierten Zeitungen. Das heißt, die epistemologischen Potenziale der einzelnen Portaltypen werden bislang kaum ausgeschöpft. Übergreifend ist hierbei besonders die »Sprach- und Textfixierung«¹²⁷ der Portale auffallend, die sich genauso in der dominanten Konzentration der Zeitungsdigitalisierung auf die Inhalts- und Textdimension der Zeitungen wie der vorherrschenden Konzeption von Interfaces und Portalen »zugunsten der reinen Textinformation und Textsuche«¹²⁸ in digitalisierten Zeitungen spiegelt. So wie mit der neuesten, als *virtual research environments* konzipierten Generation von Zeitungsportalen schon ein recht komfortabel skalierbares Experimentieren und Forschen auf verschiedenen Darstellungs- und Erschließungsebenen zwischen *close reading* und *distant reading* mit digitalisierten Zeitungen möglich ist, sollten nun gleichermaßen Optionen für ein *distant viewing* von digitalisierten Zeitungen etabliert und entwickelt werden.

Zeitungsportale als *Experimentalsysteme* stetig fortzuentwickeln, bleibt somit nicht bloß eine wichtige Herausforderung, um rein informatorisch und technisch nachhaltige und stabile Forschungsinfrastrukturen und -umgebungen für Forschende bereitzustellen und zu erhalten. Vielmehr gibt es genauso im Sinne der erkenntnisproduktiven Funktion der Portale für die Forschung noch reichlich Möglichkeiten der Verbesserung von Funktionalität, Interaktivität und Usability, sodass »newspaper interfaces are to be used as reliable gateways to in-depth analysis of these extensive archives of cultural heritage.«¹²⁹ Hierzu können Forschende selbst beitragen,

¹²¹ Rheinberger 2018, S. 125.

¹²² Rheinberger 2018, S. 119.

¹²³ Fickers 2020.

¹²⁴ Fickers 2016.

¹²⁵ Podewski 2018, S. 9.

¹²⁶ Podewski 2018, S. 15.

¹²⁷ Podewski 2018, S. 12.

¹²⁸ Blome 2018, S. 12.

¹²⁹ Pfanzelter et al. 2021, S. 12.

indem sie eine intensivere Diskussion über digitalisierte Zeitungen als exklusive epistemische Objekte anregen sowie ihre praktischen Erfahrungen und Erwartungen als User*innen hinsichtlich der Grenzen und Potenziale von digitalisierten Zeitungen und Zeitungsportalen kritisch reflektieren. Solches Praxis- und Reflexionswissen macht wiederum nicht nur die Nutzung von digitalisierten Zeitungen und Zeitungsportalen methodologisch transparent, sondern ist für Forschende zugleich Voraussetzung, um als *responsible users* gemeinsam mit den verantwortlichen Initiativen und Institutionen in vermehrt ko-kreativen Prozessen der zukünftigen Weiterentwicklung von Zeitungsportalen zusammenzuwirken.¹³⁰

¹³⁰ Vgl. Jensen 2021, S. 259.

Bibliografische Angaben

Kevin G. Barnhurst / John Nerone: The form of news. A history. New York u. a. 2001. [[Nachweis im GVK](#)]

Adrian Bingham: The digitization of newspapers archives. Opportunities and challenges for historians. In: Twentieth Century British History 21 (2010), H. 2, S. 225–231. DOI: [10.1093/tcbh/hwq007](#)

Astrid Blome: Zeitungen. In: Clio Guide – Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften. Hg. von Laura Busse / Wilfried Enderle / Rüdiger Hohls / Thomas Meyer / Jens Prellwitz / Annette Schuhmann. 2., erweiterte und aktualisierte Auflage. Berlin 2018, S. B.6-1–D.6-36. (= Historisches Forum, 23) DOI: [10.18452/19244](#)

Noah Bubenhofer / Klaus Rothenhäusler: »Korporatheken«. Die digitale und verdatete Bibliothek. In: Zeitschrift für Bibliothekskultur 4 (2016), H. 2, S. 60–71. DOI: [10.12685/027.7-4-2-154](#)

Hans-Jürgen Bucher: Ein »Pictorial Turn« im 19. Jahrhundert. Überlegungen zu einer multimodalen Mediengeschichte am Beispiel der Illustrierten Zeitungen. In: Historische Perspektiven auf den Iconic Turn. Die Entwicklung der öffentlichen visuellen Kommunikation. Hg. von Stephanie Geise / Thomas Birkner / Klaus Arnold / Maria Löblich / Katharina Lobinger. Köln 2016, S. 280–317. [[Nachweis im GVK](#)]

Daniel Burckhardt / Alexander Geyken / Achim Saupe / Thomas Werneke: Distant Reading in der Zeitgeschichte. Möglichkeiten und Grenzen einer computergestützten Historischen Semantik am Beispiel der DDR-Presse. In: Zeithistorische Forschungen / Studies in Contemporary History 16 (2019), H. 1, S. 177–196. DOI: [10.14765/zfz.dok-1345](#)

Ryan Cordell / David Smith: Viral texts. Mapping networks of reprinting in 19th-century newspapers and magazines. 2017. [[online](#)]

Nick Couldry / Andreas Hepp: Datafizierung. Wie digitale Medien und ihre Infrastrukturen unsere Praktiken, unser Wissen und unsere soziale Welt verändern. In: Mittelweg 36 30 (2021), H. 1, S. 85–101. [[Nachweis im GVK](#)]

Handbuch Digital Humanities. Anwendungen, Forschungsdaten und Projekte. Hg. von DARIAH-DE. Göttingen 2015. [[online](#)] [[Nachweis im GVK](#)]

Maud Ehrmann / Estelle Bunout / Marten Düring: Historical newspaper user interfaces. A review. In: IFLA WLIC 2019: Libraries – dialogue for change. Hg. Von IFLA. (WLIC: 85, Athen, 20.–30.08.2019), Den Haag 2019. [[online](#)]

Alena Fedasenko / Markus Muhr / Elizabeth Joss / Anastasia Gasia / Alastair Gunning: The Europeana newspapers content browser. Search and browse Europe's historical newspapers. In: slideshare.net/EuropeanLibrary/europeana-newspapers-poster. Hg. von European Library. Präsentation vom 17.02.2014. PDF. [[online](#)]

Andreas Fickers: Digitale Metaquellen und doppelte Reflexivität. In: H#Soz-Kult-Forum: Historische Grundwissenschaften und die digitale Herausforderung. Blogbeitrag vom 26.01.2016. [[online](#)]

Andreas Fickers: Zwischen Fremdheit und Vertrautheit – für eine digitale Hermeneutik in der Geschichtswissenschaft. In: Deutsches Historisches Institut Paris – Podcast. Podcast vom 25.04.2019. [[online](#)]

Andreas Fickers: Update für die Hermeneutik. Geschichtswissenschaft auf dem Weg zur digitalen Forensik? In: Zeithistorische Forschungen / Studies in Contemporary History 17 (2020), H. 1, S. 157–168. DOI: [10.14765/zfz.dok-1765](#)

Ludwik Fleck: Das Problem einer Theorie des Erkennens. In: Ludwik Fleck: Denkstile und Tatsachen. Gesammelte Schriften und Zeugnisse. Hg. von Sylwia Werner / Claus Zittel. Berlin 2011, S. 260–309. [[Nachweis im GVK](#)]

Jean-Philippe Genet: Source, Métasources, Texte, Histoire. In: Storia & Multimedia. Hg. von Francesca Bocchi / Peter Denley. Bologna 1994, S. 3–17. [[Nachweis im GVK](#)]

Paul Gooding: Historic newspapers in the digital age. Search all about it!. London u. a. 2017. [[Nachweis im GVK](#)]

Shawn Graham / Ian Milligan / Scott Weingart: Exploring big historical data. The historian's macroscope. London 2016. [[Nachweis im GVK](#)]

Justin Grimmer / Brandon M. Stewart: Text as data. The promise and pitfalls of automatic content analysis methods for political texts. In: Political Analysis 21 (2013), H. 3, S. 267–297. DOI: [10.1093/pan/mps028](#)

Kevin Hall: Hans-Jörg Rheinberger. Experimentalsysteme und epistemische Dinge. In: Schlüsselwerke der Science & Technology Studies. Hg. von Diana Lengersdorf / Matthias Wieser. Wiesbaden 2014, S. 221–233. [[Nachweis im GVK](#)]

Armin Heinen: Mediaspektion der Historiographie. Zur Geschichte der Geschichtswissenschaft aus medien- und technikgeschichtlicher Perspektive. In: Zeitenblicke 10 (2011), H. 1. Artikel vom 09.08.2011. URN: [urn:nbn:de:0009-9-30184](#).

Jessica Hubrich: Visualisierung von Titelzusammenhängen. Titelhistorie und Netzwerkgraph im neuen ZDB-Katalog. In: Offen(siv)e Bibliotheken. Neue Zugänge, neue Strukturen, neue Chancen. Hg. von Bruno Bauer / Andreas Ferus / Josef Pauser. (Österreichischer Bibliothekartag: 32, Wien, 15.–18.09.2015) Graz/Feldkirch 2016, S. 269–282. [[Nachweis im GVK](#)]

Jessica Hubrich / Hans-Jörg Lieder: Zeitungssuche interaktiv. Der neue ZDB-Katalog. In: o-bib 1 (2014), H. 1, S. 305–311. DOI: [10.5282/o-bib/2014H1S305-311](#)

Helle Strandgaard Jensen: Digital archival literacy for (all) historians. In: *Media History* 27 (2021), H. 2, S. 251–265. [[Nachweis im GVK](#)]

Bryan Jurish: DiaCollo – On the trail of diachronic collocations. In: CLARIN. Annual Conference 2015. Book of Abstracts. (Clarín 2015, Wrocław, 14.–16.10.2015), Wrocław 2015. PDF. [[online](#)]

Erik Koenen: Digitale Perspektiven in der Kommunikations- und Mediengeschichte. Erkenntnispotentiale und Forschungsszenarien für die historische Presseforschung. In: *Publizistik* 63 (2018), H. 4, S. 535–556. [[Nachweis im GVK](#)]

Erik Koenen: Experimentalsysteme und epistemische Dinge. Anregungen für eine Geschichte von Materialität und Praxis der Kommunikationswissenschaft. In: *Medien & Zeit* 35 (2020), H. 3, S. 30–34. DOI: [10.1007/s11616-018-0459-4](https://doi.org/10.1007/s11616-018-0459-4)

Erik Koenen: Forschungssoftware für die Kommunikations- und Mediengeschichte. Epistemologische Herausforderungen und Perspektiven. In: *Medien & Kommunikationswissenschaft / Themenhefte* 69 (2021), H. 1, S. 117–135. DOI: [10.5771/1615-634X-2021-1-117](https://doi.org/10.5771/1615-634X-2021-1-117)

Karin Knorr Cetina: Die Fabrikation von Erkenntnis. Zur Anthropologie der Wissenschaft. Frankfurt / Main 1984. [[Nachweis im GVK](#)]

Axel Kuhn / Svenja Hagenhoff: Kommunikative statt objektzentrierte Gestaltung. Zur Notwendigkeit veränderter Lesekonzepte und Leseforschung für digitale Lesemedien. In: *Lesen X.0. Rezeptionsprozesse in der digitalen Gegenwart*. Hg. von Sebastian Böck / Julian Ingelmann / Kai Matuszkiewicz / Friederike Schruhl. Göttingen 2017, S. 27–45. [[Nachweis im GVK](#)]

Text Mining in den Sozialwissenschaften. Grundlagen und Anwendungen zwischen qualitativer und quantitativer Diskursanalyse. Hg. von Matthias Lemke / Gregor Wiedemann. Wiesbaden 2016. [[Nachweis im GVK](#)]

Dallas Liddle: Reflections on 20,000 Victorian newspapers. »Distant reading« the Times using the times digital archive. In: *Journal of Victorian Culture* 17 (2012), H. 2, S. 230–237. [[Nachweis im GVK](#)]

Fabian Link: Neuere Entwicklungen in der Wissenschaftsgeschichte. Die mindestens drei Wissenschaftskulturen und ihre Historiographie. In: *Neue Politische Literatur* 64 (2019), H. 1, S. 27–53. DOI: [10.1007/s42520-019-0058-1](https://doi.org/10.1007/s42520-019-0058-1)

Rob Kitchin: The data revolution. Big data, open data, data infrastructures and their consequences. Los Angeles u. a. 2014. [[Nachweis im GVK](#)]

Lev Manovich: Software is the message. In: *Journal of Visual Culture* 13 (2014), H. 1, S. 79–81. [[Nachweis im GVK](#)]

Lev Manovich: Cultural analytics. Cambridge, MA u. a. 2020. [[Nachweis im GVK](#)]

Lev Manovich / Cyrus Kiani: Visualizing newspapers history: The Hawaiian Star, 5930 front pages, 1893–1912. In: *Software Studies Initiative*. Blogbeitrag vom 15.03.2012. [[online](#)]

Lev Manovich / Jeremy Douglass: 4535 Time magazine covers, 1923–2009. In: *Cultural Analytics Lab*. Blogbeitrag vom 16.04.2016. [[online](#)]

Noortje Marres: Digital sociology. The reinvention of social research. Cambridge u. a. 2017. [[Nachweis im GVK](#)]

Franco Moretti: Conjectures in world literatures. In: *New Left Review*, 1 (2000), S. 54–68. [[Nachweis im GVK](#)]

Franco Moretti: Kurven, Karten, Stammbäume. Abstrakte Modelle für die Literaturgeschichte. Frankfurt / Main 2009. [[Nachweis im GVK](#)]

Franco Moretti: Distant Reading. Konstanz 2016. [[Nachweis im GVK](#)]

Franco Moretti / Mark Algee-Hewitt / Sarah Allison / Marissa Gemma / Ryan Heuser / Matthew Jockers / Dominique Pestre / Erik Steiner / Amir Tevel / Hannah Walser / Michael Witmore / Irena Yamboliev: *Literatur im Labor*. Konstanz 2017. [[Nachweis im GVK](#)]

James Mussell: Beyond the »great index«. Digital resources and actual copies. In: *Journalism and the periodical press in nineteenth-century Britain*. Hg. von Joanne Shattock. Cambridge 2017, S. 17–30. [[Nachweis im GVK](#)]

News Eye. Anmelden. HTML. In: platform.newseye.eu. 2022. [[online](#)]

Bob Nicholson: The digital turn. Exploring the methodological possibilities of digital newspaper archives. In: *Media History* 19 (2013), H. 1, S. 59–73. DOI: [10.1080/13688804.2012.752963](https://doi.org/10.1080/13688804.2012.752963)

Arndt Niebisch: Agilität, Versionierung und Open Source. Softwareentwicklung und Praktiken der Geisteswissenschaften. In: *Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden*. Hg. von Martin Huber / Sybille Krämer. Wolfenbüttel 2018 (= *Zeitschrift für digitale Geisteswissenschaften / Sonderbände*, 3) DOI: [10.17175/sb003_009](https://doi.org/10.17175/sb003_009)

Sarah Oberbichler / Eva Pfanzelter / Stefan Hechl / Jani Marjanen: Doing historical research with digital newspapers. Perspectives of DH scholars. In: *EuropeanaTech Insights* (2020), H. 16: Newspapers. Blogbeitrag vom 16.10.2020. [[online](#)]

Urszula Pawlicka-Deger: The laboratory turn. Exploring discourses, landscapes, and models of humanities labs. In: *Digital Humanities Quarterly* 14 (2020), H. 3. [[online](#)]

Eva Pfanzelter / Sarah Oberbichler / Jani Marjanen / Pierre-Carl Langlais / Stefan Hechl: Digital interfaces of historical newspapers. Opportunities, restrictions and recommendations. In: *Journal of Data Mining and Digital Humanities, Histoinformatics* (2021). DOI: [10.46298/jdmhdh.6121](https://doi.org/10.46298/jdmhdh.6121)

Madleen Podewski: »Kleine Archive« in den Digital Humanities. Überlegungen zum Forschungsobjekt »Zeitschrift«. In: Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden. Hg. von Martin Huber / Sybille Krämer. Wolfenbüttel 2018 (=Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 3). DOI: [10.17175/sb003_010](https://doi.org/10.17175/sb003_010)

Stephen Ramsay: Reading machines. Toward an algorithmic criticism. Urbana, IL u. a. 2011. [[Nachweis im GVK](#)]

Hans-Jörg Rheinberger: Experimentalsysteme und epistemische Dinge. Eine Geschichte der Proteinsynthese im Reagenzglas. Göttingen 2001. [[Nachweis im GVK](#)]

Hans-Jörg Rheinberger: Experimenteller Geist. In: Experimentalität. Hans-Jörg Rheinberger im Gespräch über Labor, Atelier und Archiv. Berlin 2018, S. 118–159. [[Nachweis im GVK](#)]

Hans-Jörg Rheinberger: Spalt und Fuge. Eine Phänomenologie des Experiments. Berlin 2021. [[Nachweis im GVK](#)]

Matteo Romanello / Maud Ehrmann / Simon Clematide / Daniele Guido: The Impresso system architecture in a nutshell. In: EuropeanaTech Insights (2020), H. 16: Newspapers. Blogbeitrag vom 16.10.2020. [[online](#)]

Yann Ryan / Luke McKernan: Converting the British library's catalogue of British and Irish newspapers into a public domain dataset. Processes and applications. In: Journal of Open Humanities Data 7 (2021), H. 1, S. 1–5. DOI: [10.5334/johd.23](https://doi.org/10.5334/johd.23)

Philipp Sarasin: Sozialgeschichte vs. Foucault im Google Books Ngram Viewer. Ein alter Streitfall in einem neuen Tool. In: Wozu noch Sozialgeschichte? Eine Disziplin im Umbruch. Hg. von Pascal Maeder / Barbara Lüthi / Thomas Mergel. Göttingen 2012, S. 151–174. [[Nachweis im GVK](#)]

Leif Scheuermann: Die Abgrenzung der digitalen Geisteswissenschaften. In: Digital Classics Online 2 (2016), H. 2, S. 58–57. DOI: [10.11588/dco.2016.1.22746](https://doi.org/10.11588/dco.2016.1.22746)

Friederike Schruhl: Objektumgangsnormen in der Literaturwissenschaft. In: Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden. Hg. von Martin Huber / Sybille Krämer. Wolfenbüttel 2018. (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 3) DOI: [10.17175/sb003_012](https://doi.org/10.17175/sb003_012)

Felix Stalder: Was ist Digitalität? In: Was ist Digitalität? Philosophische und pädagogische Perspektiven. Hg. von Uta Hauck-Thum / Jörg Noller. Berlin 2021, S. 3–7. [[Nachweis im GVK](#)]

John Toon: What going viral looked like 120 years ago. In: Georgia Tech Research Horizons – Research News. Blogbeitrag vom 07.03.2016. [[online](#)]

Network of »Viral Text« Sharing, 1836–1899. In: networks.viraltxts.org. Hg. von University of Oxford et al. 2022. [[online](#)]

Olivia Vane: Press Picker: Visualising formats and title name changes in the British library's newspaper holdings. In: Living with Machines. Blogbeitrag vom 05.08.2020. [[online](#)]

Huub Wijfjes: Digital humanities and media history. A challenge for historical newspaper research. In: Tijdschrift voor Mediageschiedenis 20 (2017), H. 1, 4–24. DOI: [10.18146/2213-7653.2017.277](https://doi.org/10.18146/2213-7653.2017.277)

Berliner Volkszeitung. BVZ. Organ für jedermann aus dem Volke. Hg. von ZDB. In: zdb-katalog.de. 2022. [[online](#)]

Abbildungsverzeichnis

Abb. 1: Press Picker der British Library. [Vane 2020]

Abb. 2: Visualisierung von Titelrelationen in der ZDB am Beispiel der Berliner Volkszeitung. [Zeitschriftendatenbank 2022]

Abb. 3: Kompilation, Modellierung und Ordnung digitalisierter Zeitungen im Europeana Newspapers Portal. [Fedasenka et al. 2014]

Abb. 4: Konventionelle Recherche- und Sichtungslgik von Zeitungen nach dem Top-Down-Prinzip und Bottom-Up-Volltextsuche in Zeitungsportalen. [Nicholson 2013, S. 66f.]

Abb. 5: Evaluation der Funktionalitäten und Features einschlägiger Zeitungsportale. [Ehrmann et al. 2019, S. 9]

Abb. 6: The Hawaiian Star Front Pages 1893–1912. [Manovich / Kiani 2012]

Abb. 7: Time Magazine Covers 1923–2009. [Manovich / Douglass 2016]

Abb. 8: U.S. News Map. [Toon 2016]

Abb. 9: Network of Viral Text Sharing 1836–1899. [Cordell / Smith 2017]

Abb. 10: DiaCollo – Die Zeit-Korpus: Kollokationen von »Klima« im zeitlichen Verlauf. [Koenen 2022]

Abb. 11: Impresso-Interface. [Koenen 2022]

Abb. 12: NewsEye.EU-Interface. [Koenen 2022]

Beitrag aus:

Fabrikation von Erkenntnis – Experimente in den Digital Humanities (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5). 2022. DOI: [10.17175/sb005](https://doi.org/10.17175/sb005)

Titel:

Vom Topos des Scheiterns als konstituierender Kraft: Ein Essay über Erkenntnisprozesse in den Digital Humanities

Autor*in:

Tessa Gengnagel

Kontakt: tessa.gengnagel@uni-koeln.de

Institution: Universität zu Köln | Cologne Center for eHumanities (CCeH)

GND: [1248533763](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64864-p0011-9) ORCID: [0000-0001-8820-5112](https://orcid.org/0000-0001-8820-5112)

DOI des Artikels:

[10.17175/sb005_011_v2](https://doi.org/10.17175/sb005_011_v2)

Nachweis im OPAC der Herzog August Bibliothek:

[1902458362](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64864-p0011-9)

Erstveröffentlichung:

14.07.2022

Version 2.0:

19.09.2024

Lizenz:

Sofern nicht anders angegeben 

Letzte Überprüfung aller Verweise: 09.09.2024

Format:

PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:

[Diskursanalyse](#) | [Experiment](#) | [Paradigma](#) | [Scheitern](#) | [Wissenschaftstheorie](#)

Empfohlene Zitierweise:

Tessa Gengnagel: Vom Topos des Scheiterns als konstituierender Kraft: Ein Essay über Erkenntnisprozesse in den Digital Humanities. In: Manuel Burghardt / Lisa Dieckmann / Steyer Timo / Peer Trilcke / Niels-Oliver Walkowski / Joëlle Weis / Ulrike Wuttke (Hg.): Fabrikation von Erkenntnis – Experimente in den Digital Humanities (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 5). Wolfenbüttel 2022. 14.07.2022. Version 2.0 vom 19.09.2024. HTML / XML / PDF. DOI: [10.17175/sb005_011_v2](https://doi.org/10.17175/sb005_011_v2)

Änderungen in Version 2.0 (19.09.2024):

Korrekturen, Ergänzungen und Änderungen in den Abstracts, Text und Bibliografie gemäß der Gutachten.

Tessa Gengnagel

Vom Topos des Scheiterns als konstituierender Kraft: Ein Essay über Erkenntnisprozesse in den Digital Humanities

Abstracts

Das *Experimentieren* weckt in den Diskursen der Digital Humanities oftmals Assoziationen von Versuch und Ungewissheit. Eine darüber hinaus wichtige Komponente dieses rhetorischen Stilmittels wird in deutschsprachigen Publikationen allerdings weithin übersehen: und zwar seine Verbindung zu einer »importance of failure«. Der Beitrag hat primär zum Ziel, diesen Topos des Scheiterns vor dem Hintergrund erkenntnistheoretischer Überlegungen zu besprechen, die in dem Forschungsfeld gängig sind und sich um die *Möglichkeiten* der Digital Humanities drehen. Abschließend wird der Blick auf den Deutungsrahmen gelenkt, innerhalb dessen Infrastrukturen als ›Laboratorien‹ und Methoden als ›Experimente‹ verstanden werden; eine Einordnung, hinter der sich die Notwendigkeit für eine Wertbestimmung verbirgt.

The language of *experimentation* favours tendencies within Digital Humanities discourses that emphasize aspects of innovation and uncertainty. There is, however, another component that is often overlooked in the German context: and that is the association with an »importance of failure«. This essay examines whether the dimensions of failure mentioned within the field are rooted in epistemological concerns. To that end, it first (and foremost) reviews scholarship concerned with the *possibilities* of Digital Humanities knowledge production before concluding with a look at the way in which the framing of infrastructures as *laboratories* and methods as *experiments* may obscure the need for value assessment.

Einleitung

Wenn man von *Experimenten* oder einem *experimentellen Charakter* in den Digital Humanities spricht, dann wählt man nicht nur die Sprache des Versuchs und des Spiels,¹ sondern auch die Sprache des Erfolgs und Misserfolgs.² Das eine provoziert die Assoziation des anderen; bisweilen auch als bloßes Versprechen, das sich auf eine unbestimmte Zukunft richtet und Wertfindungsfragen der Gegenwart mit Hinweis auf einen noch ausstehenden Reifungs- und Konsolidierungsprozess auszublenden sucht.³ Dass Misserfolg in den Digital Humanities, zumindest in der englischsprachigen Forschungsliteratur, als Baustein auf dem Weg zum Erfolg gedeutet, für diesen teilweise gar als unabdingbar dargestellt wird, hat niemand so deutlich formuliert wie John Unsworth, der seinen Aufsatz zur Bedeutung des Scheiterns in den Digital Humanities – 1997, *avant la lettre* – mit dem denkwürdigen Satz einleitet: »If an electronic scholarly project can't fail and doesn't produce new ignorance, then it isn't worth a damn«.⁴

[1]

¹ Lisa Spiro zufolge gehört der Begriff *Experiment* zu einer »constellation of terms such as curiosity, play, exploration, and do-it-yourself« (Spiro 2012, S. 30.)

² Entsprechend ist der Begriff auch bei Spiro verknüpft, wenn sie feststellt, dass zwar nicht alle Experimente von Erfolg gekrönt seien, die DH-Community aber »the value of failure in pursuit of innovation« (Spiro 2012, S. 29) zu schätzen wisse.

³ Für den Zugang zum Thema wurde die Form des Essays gewählt, um die Perspektivität des Diskurses zu betonen und den Findungsprozess in den Vordergrund zu stellen, der sich darin widerspiegelt. Gleichzeitig wird kein Anspruch auf eine Vollständigkeit der Betrachtung oder abschließende Behandlung der umrissenen Selbstfindungsfragen erhoben. Zuletzt eignet sich die essayistische Form für eine Erörterung des kontroversen Charakters des Diskurses, ohne sich diesen explizit zu eigen zu machen (wie dies eine Polemik tun würde) oder auf jegliche Mutmaßungen zu verzichten (wie dies für eine wissenschaftlichen Abhandlung geboten wäre). Vielmehr lädt der vorliegende Text ein, die vorgebrachten Argumente abzuwägen, weiterzudenken und ihnen ggf. zu widersprechen.

⁴ Unsworth 1997.

Lisa Spiro hat den Zusammenhang zwischen den Unsworth'schen Vorstellungen des Scheiterns und einer definitiv eher das Vage favorisierenden Experimentrhetorik in den Digital Humanities bereits 2012 festgestellt⁵ – man kann allerdings nicht behaupten, dass dieser Diskurs in den deutschsprachigen Digital Humanities breit rezipiert worden wäre. Auf den Zusammenhang von Experimenten in den Digital Humanities und Scheitern als Chance hat Mareike König hingewiesen;⁶ ansonsten ist es bisher überwiegend anglophonen, insbesondere angloamerikanischen, Wortmeldungen überlassen gewesen, entsprechende Überlegungen anzustellen. Zu nennen wären beispielsweise Shawn Grahams *Failing Gloriously and Other Essays*⁷ oder der Eintrag *Failure* von Brian Croxall und Quinn Warnick in dem *Digital Pedagogy in the Humanities*-Projekt,⁸ aber auch Beiträge von Quinn Dombrowski, Max Kemman und Bethany Nowwiskie, die im Fall von Dombrowski und Nowwiskie ursprünglich als Konferenzvorträge gehalten wurden und somit als Spiegelung forschungsfeldspezifischer Dialoge gelten können.⁹ All diesen Auseinandersetzungen ist gemein, dass sie mit verschiedenen Ebenen des *Scheiterns* hadern, was sich wiederum an verschiedenen Verständnissen dessen, was die Digital Humanities sind und sollen, festmachen lässt. Gerade dieses »produktive Unbehagen«¹⁰ – um einen Terminus von Julia Flanders zu entwenden – könnte für ein Verständnis der Digital Humanities aufschlussreich sein, weil sich darin eben jene Wachstumsbeschwerden ausdrücken, die zwischen Versuch und Spiel nach einer Ernsthaftigkeit des Anliegens fahnden; schließlich drängt sich bei Überhöhungen des Misserfolgs die Frage auf: Was macht eigentlich *Erfolg* aus? Interessanterweise findet sich zu dieser Frage kein ähnlich dezidierter Diskurs in diesem Feld. Oder zielt die *Failure*-Debatte darauf ab? Wenn, dann nur im Umkehrschluss – dabei sind es erst Kriterien des Gelingens, die uns befähigen, ein Misslingen zu begründen statt es bloß zu konstatieren.¹¹ Insofern sich eine diskursive Nähe zwischen Experiment und Scheitern feststellen lässt, also zwischen einer Ungewissheit auf der einen und einer Gewissheit auf der anderen Seite, oder auch: zwischen einer Betonung von gemeinhin akzeptierten Formen der (Natur-)Wissenschaftlichkeit einerseits und einem Bekenntnis zu gemeinhin selten besprochenen Realitäten des wissenschaftlichen Arbeitens andererseits, lohnt es sich, diesen Stimmungsbildern in der Forschungsliteratur nachzugehen und ihr Verhältnis zu den digitalen Geisteswissenschaften zu bestimmen.

[2]

1. Failure

Zunächst einmal lässt sich zusammenfassen, dass sich die Idee eines greifbaren Scheiterns, das in der Nichterfüllung bestimmter Erwartungen oder einer unter anderen Umständen konkret realisierbaren Zielvorstellung wurzelt, nach einer Synopse der vorhandenen Diskussionsbeiträge in unterschiedliche Problemfelder aufschlüsseln lässt:

[3]

1. ein Versagen von Technologien, das heißt eine Diskrepanz zwischen der ihnen zugeordneten Aufgabe und den tatsächlich resultierenden Ergebnissen
2. ein menschliches Versagen, das heißt ein Unvermögen oder mangelnde Erfahrung im Umgang mit und in der Anwendung von Technologien

⁵ Vgl. Spiro 2012, S. 28–30.

⁶ Vgl. König 2015.

⁷ Vgl. Graham 2019.

⁸ Vgl. Croxall / Warnick 2015.

⁹ Vgl. Dombrowski 2019; Kemman 2019; Nowwiskie 2012.

¹⁰ Übersetzung der Autorin, vgl. Flanders 2009, Abs. 12.

¹¹ An dieser Stelle sei darauf hingewiesen, dass es Grund zu der Annahme gibt, die Failure-Rhetorik unmittelbar in der Nähe von wirtschaftlich beeinflussten Innovationsbemühungen zu verorten, die schon seit Jahrzehnten im Gründungsmythos der Digital Humanities verankert sind: So berichtet Roberto Busa, Pionier der ersten Stunde, in seinen *Annalen des Humanities Computing*, wie er 1949 im Warteraum von IBM (kurz vor seinem Schicksalstreffen mit dem Geschäftsführer) ein Plakat entdeckt habe mit einem Motivationspruch: »The difficult we do right away; the impossible takes a little longer.« (IBM always loved slogans). I took it with me into Mr. Watson's office. Sitting in front of him and sensing the tremendous power of his mind, I was inspired to say: »It is not right to say »no« before you have tried.« I took out the poster and showed him his own slogan. He agreed that IBM would cooperate with my project.« (Busa 1980, S. 84.) Dieser Aspekt der kommerziellen Verflechtungen und Werbelogiken wird in diesem Essay später noch einmal kurz anklingen, bedürfte allerdings einer separaten Aufarbeitung, insbesondere im Rahmen der Rolle von IBM, jenes Konzerns, der in viele Bereiche der Digital Humanities in ihren Anfangsjahren hineingewirkt hat (für die digitale Kunstgeschichte ist dies bereits von Pratschke 2018 angedeutet worden).

3. arbeitspraktisches Versagen, das heißt die gemäß zeitkritischer Horizonte aus diversen Gründen fehlende oder unzufriedenstellende Umsetzung selbst-deklarerter *Outcomes, Deliverables* und *Milestones*, insbesondere im Hinblick auf projektdominierte DH-Arbeitsumgebungen¹²
4. ein intellektuelles Versagen, das heißt eine Oberflächlichkeit in den Forschungsprämissen oder in der Interpretation der Forschungsergebnisse.

Sowohl arbeitspraktisches als auch intellektuelles Versagen könnten als Formen eines menschlichen Versagens eingestuft werden und es liegt nahe, dass die Kategorien, die die Forschungsliteratur an dieser Stelle ausbildet, mehrheitlich ineinandergreifen. Das gilt sowohl für die vier Versagensstufen bei Brian Croxall und Quinn Warnick¹³ als auch für die Taxonomie von Quinn Dombrowski, die einen umfassenderen Ansatz verfolgt und neben technologischen Aspekten insbesondere kommunikative und karrieristische Fallstricke betont.¹⁴ Shawn Graham bedient sich sowohl bei Dombrowski als auch bei Croxall / Warnick, um das eigene Lebenswerk einzuordnen.¹⁵ Max Kemman fragt, ob das *Scheitern* deswegen Einzug in die DH-Gedankenwelt gefunden habe, weil die DH primär als Form der Softwareentwicklung und nicht als Forschung verstanden würden.¹⁶ Bethany Nowwiskie knüpft ihre Ausführungen ganz konkret an die Entstehungsgeschichte des ›Scholars' Lab‹ an der University of Virginia, das als Nukleus einer ganzen Generation von DH-Wissenschaftler*innen in der Nachfolge John Unsworths und Jerome McGanns gelten kann und in dieser Tradition Fortschrittsnarrative verinnerlicht hat,¹⁷ welche sich bei Nowwiskie in Überlegungen zu »experimentellen Frühphasen«¹⁸, einer »Hermeneutik des Machens [oder] der Fabrikation«¹⁹ und einer »spielerischen Entdeckerlust«²⁰ äußern.

[4]

Dieser kurze Überblick soll lediglich als Einstieg dienen, denn es gilt einigen Fragen vertieft nachzugehen. Die Verknüpfung von Vorstellungen des *Scheiterns* mit methodischen Anleihen an *Experimente* und einer institutionellen Koppelung von DH-Forschung an sogenannte *Labs*²¹ legt nahe, dass dieses Dreieck aus Machbarkeit, Machen und Wirk- oder Werkstätte einen zumindest nicht unbedeutenden Anteil an dem Selbstverständnis des Feldes hat. Es erscheint daher sinnvoll, diese Phänomene nicht isoliert zu betrachten, sondern als Kontinuum zu verstehen, das an einer verdeckten Sinnfrage laboriert; einer Sinnfrage, die da eben nicht lautet: Was *sind* die Digital Humanities? Sondern: Was *können* die Digital Humanities? Mehr noch: Was *sollten* sie können? Und schließlich: Was können sie *nicht*?

[5]

Die Beantwortung solcher Fragen bewegt sich zwangsläufig in einem Spannungsfeld verschiedener Erwartungshaltungen und entsprechend differenziert muss der Blick auf mögliche Hinweisgeber sein. Wenn wir die Möglichkeiten der Wissensfindung in den Digital Humanities von den Möglichkeiten der Wissensfindung in den Geisteswissenschaften her denken, um anschließend eruieren zu können, inwiefern über diese hinausgegangen werden kann, respektive inwiefern die Digital Humanities hinter ihnen

[6]

¹² Fragen der projektdominierten DH-Arbeitspraxis und ihrer infrastrukturellen Bedingungen sind insbesondere in den letzten Jahren zunehmend in den Fokus von Sammelbänden gerückt, siehe hierzu Publikationen wie Cremer et al. (Hg.) 2024, Nyhan et al. (Hg.) 2023 und Pawlicka-Deger / Thomson (Hg.) 2023.

¹³ Croxall / Warnick unterscheiden folgende Kategorien: (1) »technological failure«, (2) »human failure«, (3) »failure as artifact« und (4) »failure as epistemology« (Croxall / Warnick 2015).

¹⁴ Neben dem (1) »technical failure« sind das (2) »disciplinary failure«, (3) »career planning failure«, (4) »communication failure«, (5) »strategic failure«, (6) »timing failure«, (7) »failure to probe assumptions«, (8) »failure to acknowledge change«, (9) »failure to forge a shared vision« und (10) »failure to do right by others« (Dombrowski 2019).

¹⁵ Vgl. Graham 2019, S. 12.

¹⁶ Vgl. Kemman 2019.

¹⁷ Die Rolle, die das English Department an der University of Virginia in der Verstetigung der angloamerikanischen Digital Humanities gespielt hat, war bereits Gegenstand einer Kontroverse, als deren Auslöser eine unsachlich-politisierende, wenn nicht gar polemische, Retrospektive angesehen werden kann; vgl. Allington et al. 2016 und die zeitnahe Replik Spahr et al. 2016.

¹⁸ »The early, experimental days of humanities computing at UVa taught me to see openness to failure and *openness itself* as our best paths to learning in DH« (Nowwiskie 2012). Hervorhebungen im Original.

¹⁹ »We are seeking evidence in our intellectual labor together, of the emergence of a new, non-discursive hermeneutic of ›making‹ in the digital humanities« (Nowwiskie 2012). Es sei darauf hingewiesen, dass *making* in diesen Diskursen die Konnotation des Tüftelns, Werkelns und Bastelns hat und mit ›Machen‹ nicht ganz adäquat zu übersetzen ist.

²⁰ »[W]ho taught me to learn by breaking, warping, deforming, loving, and above all by *playing with things* in a lighthearted way— with objects of our shared cultural heritage, to be sure—but also by playing with and within the *institutional structures* that shape and circumscribe or enable our work« (Nowwiskie 2012). Hervorhebungen im Original.

²¹ Vgl. Pawlicka-Deger 2020.

zurückbleiben, führt uns das in wissenschaftstheoretische Überlegungen des 19. und 20. Jahrhunderts. Wenn es sich bei dem dräuenden Schatten eines Scheiterns (oder der Befreiungsmacht desselben) hingegen um ein wissenschaftssoziologisches Phänomen handelt, das Selbstvergewisserungsprozesse narrativ einhegen soll, dann wird man nicht umhinkommen, vermeintliche oder reale Unzulänglichkeiten dezidiert als Frage der Umsetzung zu diskutieren und von dem theoretischen Diskurs um Erkenntnismöglichkeiten explizit abzugrenzen. Zuletzt hat das Thema eine außerakademische, gesellschaftliche Dimension: Wenn wir die Digital Humanities und ihren *Erfolg* oder *Misserfolg* als Fach, intellektuelle Bewegung oder lose verbandelte Interessensgemeinschaft in einen größeren zeitgeschichtlichen Kontext stellen – was für den wissenschaftlichen Austausch mangels belastbarer Aussagen jenseits der eigenen Beobachtung kaum geeignet, aber gerade deswegen von übergeordneter struktureller Bedeutung ist – müssen wir von eben jenen Fremdbildern Kenntnis nehmen, deren bejahende oder warnende Haltung sich entweder in abgeschwächter Form unmittelbar in den Digital Humanities wiederfindet oder auf die sie, wenngleich unterbewusst, in ihren Selbstdarstellungen reagieren.²² Dass solche Wahrnehmungen auch von Wissenschaftler*innen anderer Disziplinen in die gesellschaftliche Breite getragen werden, zeigt sich an Hans Ulrich Gumbrechts Jeremiade über den Zustand der Geisteswissenschaften, in der er von einer »intellektuellen Überbewertung der elektronischen Instrumente«²³ spricht und ausführt:

»Ihren Anspruch auf rigorosen Ernst und auf Spezialisierung versuchen die Geisteswissenschaften mittlerweile mit Programmen der Selbstdigitalisierung aufrechtzuerhalten, die den eigentlich kaum zu vermeidenden Schritt, sich elektronische Technologien als Instrumente zunutze zu machen, in den Status einer intellektuellen Revolution hochjubeln. Dank entsprechend eingestellten Suchprogrammen braucht man heute nur noch Sekunden – und nicht mehr Jahre –, um den Gebrauch eines Begriffs während eines Jahrzehnts oder die Eigenheiten im Satzbau eines bestimmten Autors identifizieren und dokumentieren zu können. Aber erst hier setzt dann weiterhin die einzig relevante Frage nach existenziell oder gar politisch relevanten Folgerungen an.«²⁴

[7]

Unabhängig davon, ob man diese Meinung für exemplarisch hält oder nicht, ist sie Teil des Rahmens, in den die Digital Humanities und ihre Rhetorik eingebettet sind. Bevor wir auf eben jene Rhetorik und das damit verbundene Selbstverständnis der Digital Humanities zu sprechen kommen, soll es im Nachfolgenden nun zunächst um übergeordnete wissenschaftstheoretische Perspektiven gehen, um daraus Impulse für die – gegenwärtig noch ausstehende – Einordnung des DH-Experimentbegriffs²⁵ sowie die – gegenwärtig noch nicht abschließend reflektierte – Prävalenz einer darüber hinausgehenden sprachlichen DH-Laborisierung abzuleiten.²⁶ Dass dieses Thema eine gewisse Dringlichkeit aufweist, liegt dabei nicht so sehr an einem allgemeinen Erkenntnisinteresse, sondern an den Verdrängungseffekten, die sich aus der Dominanz bestimmter Wortbilder ergeben können und anderweitig assoziierte Erklärungsansätze aus dem konsensorientierten Deutungsrahmen fallen lassen. In anderen Worten: Computergestützte Verfahren spielen in den Geisteswissenschaften ebenso eine Rolle wie die allgemeine Digitalisierung von Lebenswelten; doch was sagt uns das über Arbeitspraktiken und Wissenschaftsansprüche?

[8]

²² Ein Beispiel für eine bewusste Reaktion auf die kritische Fremdwahrnehmung der Digital Humanities im US-amerikanischen Kontext, die zugleich Argumente der Kritiker aufgreift und in Teilen bestätigt, ist Weed / Rooney (Hg.) 2014.

²³ Gumbrecht 2019. Vgl. für die Debatte an sich auch die Replik Kablitz 2019.

²⁴ Gumbrecht 2019.

²⁵ In Publikationen wie Sayers 2017 sucht man vergebens nach einer plausiblen Definition dessen, was dort sogar titelgebend ist. Zum Experimentbegriff in den Digital Humanities, siehe auch den Beitrag von Sarah Lang in diesem Band, vgl. Lang 2022.

²⁶ Es sei darauf hingewiesen, dass am 24. März 2021 im Rahmen der vDHd-Tagung »Experimente« ein Workshop zu dem Thema »(Keine) Experimente in den Digital Humanities?« stattgefunden hat, organisiert von Jonathan D. Geiger, Max Grüntgens und Dominik Kasper von der Akademie der Wissenschaften und der Literatur Mainz. Hierbei wurde dieses Desiderat thematisiert, vgl. vDHd (Hg.) 2021. In Bezug auf die sprachliche Laborisierung hat Pawlicka-Deger 2020 die umfangreichste Betrachtung vorgelegt, die jedoch weder in besonderem Maße auf politisch-ökonomische Sachzwänge noch auf die gesellschaftlich-wissenschaftliche Legitimationskrise eingeht, die die Humanities in den USA – woher der Trend der institutionellen »Laborisierung« in den Digital Humanities stammt – spätestens seit den *Science Wars* in den 1990ern verschärft begleitet und in der Abwägung solcher vermeintlich legitimierender Tendenzen nicht völlig außer Acht gelassen werden sollte; vgl. hierzu eine Aussage zum Ausgang der *Science Wars*, die nahelegt, warum manche *humanists* sich (unterbewusst) bemüßigt gefühlt haben mögen, anderen Wissenschaftler*innen – zumindest deklaratorisch – in die Labore zu folgen: »They [scientists] then returned to their labs, with many concluding that the humanities was really irrelevant to their research« (Pennock 2019, S. 210f.).

2. Denktradition: Zwei Kulturen

Es wäre vermessen, an dieser Stelle eine Zusammenfassung der gesamten Literatur zu Methodik und Gegenstand der Geisteswissenschaften leisten zu wollen; zu stark müsste man hierfür Begriffe und Positionen verkürzen, die im deutschsprachigen Raum mit Philosophen und Gelehrten wie Dilthey, Schleiermacher, Husserl, Heidegger und Gadamer verknüpft sind. Daher ist es zunächst einmal wichtig, jene epistemologischen Schwerpunkte zu identifizieren, die für eine Einordnung der Digital Humanities besonders relevant erscheinen. Der Rekurs auf die Geisteswissenschaften ergibt sich ganz prinzipiell aus der Tatsache, dass in den Digital Humanities oft ein geisteswissenschaftliches Erkenntnisinteresse vorausgesetzt wird, mutmaßlich beeinflusst durch die Tatsache, dass viele Forscher*innen, die sich mit diesem Feld identifizieren und in Fragen der Wissenschaftstheorie zu Wort melden, ursprünglich in einem traditionellen geisteswissenschaftlichen Fach beheimatet waren oder es, ungeachtet ihres Interesses an dem wie auch immer gearteten Einsatz von Computern und Technologien in den Geisteswissenschaften, weiterhin sind.²⁷ Es steht zu vermuten, dass man eine Betrachtung der Thematik auch aus einer anderen Sicht, etwa aus der Sicht der Informatik, konzipieren könnte; dies ist und bleibt zwar ein Desiderat (dessen sich die Computational Humanities zunehmend annehmen²⁸), allerdings würde es im vorliegenden Fall nicht dabei behilflich sein, die mehrheitlich bestehenden DH-Diskurse und die ihnen zugrundeliegenden Denktraditionen nachzuvollziehen.

[9]

In diesem Sinne sollten wir zuallererst Grundannahmen formulieren, die selbstverständlich erscheinen mögen, es im internationalen Vergleich allerdings nicht immer sind. Hierzu zählen beispielsweise die Annahmen, dass

[10]

1. es einen überordneten Wissenschaftsbegriff gibt, der
2. die Geisteswissenschaften einschließt und
3. eine Unterscheidung zwischen Wissenschaften einerseits mit ihrem Untersuchungsgegenstand und andererseits mit ihrem Methodenrepertoire, nicht aber ideologisch begründet.

Während die Antwort auf die Frage, ob denn die Geisteswissenschaften Wissenschaften seien, bei dem deutschen Begriff zumindest vordergründig inbegriffen ist (ohne dass wir damit näher spezifiziert hätten, was *Wissenschaft* ausmacht), wird es mit Blick auf den englischsprachigen Diskurs etwas komplizierter; und dieser Blick ist alleine deswegen unvermeidbar, weil Englisch als *lingua franca* der internationalen Digital Humanities gilt.²⁹ Um es deutlich festzuhalten: Es gibt im Englischen keinen Wissenschaftsbegriff, der sowohl die Geistes- als auch die Naturwissenschaften umfassen würde; *humanities* sind keine *sciences* und *humanists* keine *scientists*.³⁰ Wer meint, dies wäre für die Identitätsfindung der Digital Humanities irrelevant, irrt. Im Gegenteil: Aus dieser Grundkonstellation leitet sich ein nicht unerheblicher Anteil definitorischer Schwierigkeiten ab, da sich die Digital Humanities, grob gesprochen, nicht nur zu den Geisteswissenschaften verhalten und von diesen abheben oder auch nicht, sondern als Schnittstelle oder Mittler – so eine gängige Vorstellung ihrer Verortung³¹ – zwangsläufig von den Polen abhängig sind, zwischen denen sie angeblich pendeln. Eine Positionsbestimmung kann ohne Bestimmung des Rahmens, in dem sie stattfindet, nicht gelingen.

[11]

Bestes Beispiel für die Art und Weise, wie die Wissenschaftskonzeption im englischsprachigen Raum in die Digital Humanities hineinwirkt, ist ein Aufsatz von Paul S. Rosenbloom, der nach seiner Erstveröffentlichung in *Digital Humanities Quarterly (DHQ)* Eingang in den kuratierten und weithin beachteten Band *Defining Digital Humanities* gefunden hat.³² Rosenbloom beschäftigt sich mit der Frage, wie man die *humanities* als Teil

[12]

²⁷ Obwohl die Untersuchung von Melissa Terras zu der disziplinären Verortung von Teilnehmer*innen der ACH / ALLC-Konferenz bereits etwas älter ist, ist sie trotzdem noch interessant; ob sich hieraus weiterhin Rückschlüsse auf rezente Entwicklungen ziehen lassen, bedürfte einer neueren und auch umfassenderen Studie; vgl. Terras 2006.

²⁸ Vgl. bspw. Offert / Bell 2020.

²⁹ Vgl. Crane 2015; Fiormonte 2021.

³⁰ Vgl. Bod et al. 2016, S. 4.

³¹ Vgl. bspw. Rehbein 2020, S. 1427.

³² Vgl. Rosenbloom 2013.

von *science* begreifen könne³³ – mehr noch, als Untermenge der »social sciences«³⁴. Sein Hauptargument zielt darauf, dass »any enterprise that *tends to increase our understanding of the world over time*«³⁵ als »essentially scientific«³⁶ anzusehen sei. Zwar erkennt er an, dass es in anderen Ländern bereits andere Wissenschaftsverständnisse gibt als jenes, das *science* im Englischen sehr eng auslegt, und er verweist sogar darauf, dass man sich für einen umfassenden Wissenschaftsbegriff im Deutschen bedienen könne³⁷ – die Lektüre mutet aber trotzdem befremdlich an. Existente Überlegungen aus dem Bereich der Wissenschaftsphilosophie werden nicht näher rezipiert. So würde man annehmen, dass für die Theorie der Geisteswissenschaften im englischsprachigen Raum Schriften wie Georg Henrik von Wrights *Explanation and Understanding* von Interesse wären;³⁸ in anglophonen DH-Diskursen sucht man solche Referenzen allerdings vergebens, was mutmaßlich nicht so sehr auf ein spezifisches Versäumnis der Digital Humanities hindeutet als auf eine wenig ausgeprägte Auseinandersetzung mit solcher Theoriegeschichte in den Humanities an sich. Bei Rosenbloom finden Popper, Kuhn und Feyerabend als Vertreter allgemeiner Wissenschaftstheorien kurze Erwähnung.³⁹

Wenn man nun also fragt, ob denn die Geisteswissenschaften Wissenschaften seien, dann ist das eine andere Frage als die Frage, ob denn die *humanities sciences* seien, und dieser einfache (oder vielmehr höchst komplexe) Umstand suggeriert, dass die deutschsprachigen Digital Humanities ihre begriffliche Gleichsetzung von *humanities* und Geisteswissenschaften und daran anschließend von Digital Humanities und digitalen Geisteswissenschaften überdenken sollten; aus den Nuancen erwachsen Differenzen. Das lässt sich auch daran ablesen, in welchen divergenten Bahnen die englisch- und deutschsprachigen Diskurse zur Theorie der digitalen Geisteswissenschaften verlaufen.

[13]

In anglophonen Publikationen wird bei der Standortbestimmung der Digital Humanities oft auf C. P. Snows *two cultures* rekurriert,⁴⁰ die er 1959 in einer Rede kontrastiert und somit eine langlebige Erzählung geschaffen respektive perpetuiert hat, die Natur- und Geisteswissenschaften als fundamental verschiedenartig begreift.⁴¹ Die Rede selbst war als Angriff auf eine Schiefelage in der Prestigekultur an der Universität Cambridge konzipiert und macht sich den Wettbewerbsgedanken des Kalten Krieges zu eigen; entsprechend muss die Frage erlaubt sein, warum sie, in der Regel ohne nähere Beleuchtung ihrer Entstehungsumstände oder eigentlichen Argumente, weiterhin als primärer Referenzpunkt in gegenwärtigen Diskursen dient, so etwa, wenn Amy E. Earhart in ihren Ausführungen zu den *Digital Humanities as Laboratory* unter Verweis auf Snow davon spricht, dass es schon lange eine Trennlinie zwischen den *humanities* und den *sciences* gäbe, die sich zunehmend in einem Ungleichgewicht an Forschungsförderung äußern und entsprechend auf geisteswissenschaftlicher Seite Ressentiments Vorschub leisten würde.⁴² Die Existenz zweier »Wissenschaftskulturen«, von Snow einst behauptet, wird so in den Stand einer Tatsache erhoben. Dass es bei dieser Unterscheidung weniger um wissenschaftstheoretische als um wissenschaftssoziologische

[14]

³³ Vgl. Rosenbloom 2013, S. 220.

³⁴ Rosenbloom 2013, S. 232.

³⁵ Rosenbloom 2013, S. 221. Hervorhebung im Original.

³⁶ Rosenbloom 2013, S. 221.

³⁷ »[W]hether the generic is called philosophy, or science, or even *Wissenschaft* – a German word for science that includes not only those academic disciplines typically labeled as science in English but also other areas of academic study, such as the humanities« (Rosenbloom 2013, S. 223).

³⁸ Vgl. von Wright 1971.

³⁹ Vgl. Rosenbloom 2013, S. 221–223.

⁴⁰ Vgl. bspw. Porsdam 2013.

⁴¹ Vgl. Snow 1993.

⁴²

»As many working within digital humanities laud the laboratory model, there remains deep suspicion of bringing a science model to humanities work. The split between science and the humanities is longstanding, á [sic!] la C. P. Snow, and resistance to utilizing anything from science is increasing as university funding and prestige is increasingly seen, by some humanists, to correlate to STEM areas rather than humanities fields«

Aspekte geht, ist nicht nur aus der Art der Verweise ersichtlich, die sich auf sie beziehen, sondern aus dem programmatischen Text selbst, der Formalwissenschaften ausklammert oder zumindest nicht als distinkt zu Natur- und Geisteswissenschaften ausweist. C. P. Snows Beobachtungen erschöpfen sich, um ein Beispiel zu nennen, in Aussagen wie jener, dass den *scientists* die Zukunft in den Knochen stecken würde, während die ›traditionelle Kultur‹ (gemeint sind vor allen Dingen die Altphilologien) davon provoziert jedem Gedanken an die Zukunft entsagen würden: »[i]f the scientists have the future in their bones, then the traditional culture responds by wishing the future did not exist.«⁴³ Wäre nun die Zukunftsaffinität ein entscheidendes Kriterium, könnte man durchaus konstatieren, dass die Digital Humanities zumindest in technologischer Hinsicht eher am Puls der Zeit zu sein scheinen als ›traditionell‹ arbeitende Geisteswissenschaften. Den Puls der Zeit auf Technologie zu reduzieren, würde aber kaum den geisteswissenschaftlichen Ansprüchen gerecht werden, deren Blick in die Vergangenheit von einem Bedürfnis nach Wissensvermehrung geleitet wird, die in ihrer letzten kumulativen Konsequenz nie hinter das zurücktreten kann, was bereits einmal gewusst gewesen; so zumindest der wissenschaftliche Leitgedanke. Unabhängig von dem Wahrheitsgehalt solcher Charakterisierungen verwundert es nicht, dass das unbestimmte Bild von zwei ›Kulturen‹ wirkmächtiger zu sein scheint als das, was damit konkret gemeint sein mag.

In diesem Zusammenhang darf nicht unerwähnt bleiben, dass die Vorstellung zweier Kulturen dazu geführt hat, dass die Digital Humanities sich bisweilen als ›dritte Kultur‹ oder ›dritter Weg‹ positionieren.⁴⁴ Schon Jean-Claude Gardin, ein Pionier der Archäoinformatik, hat im Kontext von Expert*innensystemen und unter den Eindrücken semiologischer Hinwendungen zu den Literaturwissenschaften über einen dritten (Aus-)Weg nachgedacht, den er in der Archäologie in Ansätzen bereits verwirklicht sah, nicht in Opposition zur Hermeneutik, sondern in Verwandtschaft zu ihr.⁴⁵

[15]

3. Denktradition: Die ungenauen Wissenschaften

Die Frage der Hermeneutik erinnert indes an eine andere Art von Dualismus, die uns näher an wissenschaftstheoretische Überlegungen zur Unterscheidung zwischen den Wissenschaften und damit auch zu Fragen wissenschaftlichen *Erfolgs* und *Misserfolgs* heranführt. Gemeint ist die Aufteilung in genaue und ungenaue Wissenschaften, die Gerhard Lauer unter Bezugnahme auf Jacob Grimm diskutiert hat.⁴⁶ Der Rückgriff auf Jacob Grimm ist in der Tat wichtig, weil er in die (früh-)neuzeitliche Entstehungsphase der Konzeptionen und Kategorisierungen verweist, auf denen das heutige Wissenschaftssystem im deutschsprachigen Raum weiterhin zum größten Teil fußt.

[16]

Lauers Verweis auf Jacob Grimms 1846 veröffentlichte Rede *Über den Werth der ungenauen Wissenschaften*⁴⁷ ist dabei nicht nur deshalb interessant, weil sich der Eindruck aufdrängt, dass wissenschaftstheoretische Themen besonders dann aufgegriffen werden, wenn es eine Rede zu halten gilt, wie sich bereits bei der *Failure*-Thematik feststellen ließ.⁴⁸ Die Einbindung der Grimm'schen (wenn auch nicht originär Grimm'schen)⁴⁹ Unterscheidung zwischen genauen und ungenauen Wissenschaften ist von Interesse, weil Lauer sie auf die Polemiken abbildet, die insbesondere im angloamerikanischen Raum Kritik an den Digital Humanities üben und Erkenntnisprozesse in den *humanities* vor allen Dingen in der Kraft des menschlichen Geistes

[17]

⁴³ Snow 1993, S. 11.

⁴⁴ Vgl. bspw. Hunter 2014.

⁴⁵ Vgl. Gardin 1990, S. 27f. Zu Jean-Claude Gardin siehe auch Buzzetti 2019.

⁴⁶ Vgl. Lauer 2020.

⁴⁷ Vgl. Grimm 2016.

⁴⁸ Unsworth 1997, Nowwiskie 2012, Dombrowski 2019 und Lauer 2020 basieren auf mündlichen Vorträgen und man könnte diese Beobachtung noch um weitere bedeutende Texte wie Wilhelm Windelbands Straßburger Rektoratsrede (1894) ergänzen, wobei es sich natürlich um Zufall handeln mag – andererseits leuchtet es ein, dass gerade Vorträge dazu einladen, in unverbindlichem Rahmen grundsätzlicher zu werden; das nur als wissenschaftssoziologische Anmerkung am Rande.

⁴⁹ Es soll darauf hingewiesen sein, dass Grimm sich hierbei auf eine französische Tradition beruft, die es vertieft zu recherchieren gälte, wollte man dieser Wissenschaftsgeschichte weiter auf den Grund gehen. Lauer nennt einige frühneuzeitliche Vorläufer von ›exakten‹ Methoden in den Geisteswissenschaften, geht aber nicht näher auf die zeitgenössischen wissenschaftstheoretischen Diskurse ein; vgl. Lauer 2020, S. 159–161.

situiert wissen wollen, also jeglichen Versuch der Externalisierung a priori ablehnen.⁵⁰ Lauer schränkt beständig ein, dass Kriterien von vermeintlicher Objektivität und Exaktheit keine natürliche Trennlinie zwischen Natur- und Geisteswissenschaften bilden, vielmehr geht es ihm darum, zu beweisen, dass die (deutschsprachigen) Geisteswissenschaften ihren Anspruch auf Wissenschaftlichkeit schon seit Grimms Tagen aus ihrer Vorliebe für nachvollziehbare Schritte, Abgleiche, Einordnungen, man möchte sagen: Formen der historisch-kulturwissenschaftlichen Informationsverarbeitung ziehen, allerdings verliert er sich dabei in den Widersprüchen, die er aufzulösen sucht; nicht zuletzt deshalb, weil Fragen der Systematik des Vorgehens, der Gründlichkeit, der Überprüfbarkeit, der Ausdrücklichkeit (im Sinne ihrer Explizitmachung) mit dem Kriterium der Genauigkeit gleichgesetzt und mitunter verwechselt werden.⁵¹ Das Plädoyer, das Jacob Grimm zugunsten der ungenauen Wissenschaften hält, wird mit Blick auf das rückwirkende Postulat seines entgegen seines eigenen Bekenntnisses »strengen Methodenideals«⁵² zu einer Form der Außendarstellung deklariert und damit ausgehebelt, was an Rens Bods Argumente erinnert, der betont, die Theorie der Geisteswissenschaften im ausgehenden 19. Jahrhundert sei losgelöst von ihrer Praxis gewesen, was ihn in seinen Überlegungen zu Fragen der Modellierung in den (digitalen) Geisteswissenschaften aus der Verantwortung entlässt, sich mit eben jenen Theorien auseinanderzusetzen.⁵³ Zu Grimm sei noch gesagt, dass sein Impetus in seinem Loblied auf die Ambivalenz der Geisteswissenschaften patriotisch geprägt war und eben nicht oder nicht nur auf ihre Methoden zielte, die ihm editionswissenschaftlich natürlich genauso nahe waren wie einem sonst oft als Vertreter einer Verwissenschaftlichung der Geisteswissenschaften angeführten Karl Lachmann,⁵⁴ sondern auch (und in besonderem Maße) auf die Konsequenz ihrer Ergebnisse, das heißt in seinem Fall auf ihre nationalgeschichtliche Bedeutung, die – da der Untersuchungsgegenstand aus der Kultur erwachsen und nur in der Kultur zu verstehen – »uns näher zu Herzen [sic!]«⁵⁵ stünde. Man muss

⁵⁰ Lauer bezieht sich unter anderem auf die plakative Artikelreihe *The Digital Humanities War*, die 2019 in *The Chronicle of Higher Education* zwei Aufsätze demonstrativ gegeneinander antreten ließ; zudem wurden unter dem Titel der Reihe auch ältere Beiträge subsumiert. Diese Art von Beiträgen, bei denen es sich streng genommen eher um *op-eds* handelt, um Meinungskommentare, ist gerade in dieser Form durchaus typisch in einem polarisierten und bewusst polarisierenden angloamerikanischen Publikationsumfeld. Für die widerstreitenden Artikel in diesem Fall, vgl. Underwood 2019 und Da 2019.

⁵¹ Eine solche Verwechslung liegt im Fall der Stematologie vor, von der Lauer behauptet, dass Dilthey sie »aus seiner Theorie [der Geisteswissenschaften] definitorisch ausgeschlossen« (Lauer 2020, S. 163) habe, was natürlich keineswegs zutrifft, da eine Theorie, die auf das Verständnis von Texten fokussiert ist, deren Vorhandensein voraussetzt und auch die Verdienste der Editionswissenschaft zu schätzen weiß; es ist aber nicht so, als hätten selbst Philologen (und Dilthey war kein Philologe) die Stematologie im 19. Jahrhundert als geisteswissenschaftliche Methode ausführlich beschrieben oder reflektiert (eingeschränkt: mit Ausnahme von französischen Philologen wie Gaston Paris), ungeachtet der Tatsache, dass sie sie praktiziert haben, sonst hätte man im deutschsprachigen Raum nicht auf die schmalen und doch erstmals in dieser Form expliziten methodischen Ausführungen von Paul Maas warten müssen (vgl. Maas 1960). Zur Exaktheit der Stematologie sei gesagt, dass sie zwar regelbasiert und damit äußerst fundiert ist, in ihrer Priorisierung der Rekonstruktion eines unverfälschten »Originals« aber ideologisch ein Kind ihrer Entstehungszeit war (was Paolo Chiesa gar als »eugenic vision« (Chiesa 2020, S. 83) bezeichnet hat). Die der Stematologie zugrundeliegende Logik ist auch anderweitig scharf kritisiert worden, unter anderem von dem Paris-Schüler Bédier (vgl. Palumbo 2020, S. 91–97), und natürlich ist folgender Ausspruch von Maas wohlbekannt, den es in Hinblick auf eine »Exaktheit« der Stematologie stets vor Augen zu halten gilt: »Wie in der chemischen Formel die Anordnung der Atome für jedes Molekül einer Verbindung eindeutig und unveränderlich festgelegt ist, so im Stemma das Abhängigkeitsverhältnis der Zeugen für jede Stelle des Textes – wenn jungfräuliche Überlieferung vorliegt. Gegen die Kontamination ist kein Kraut gewachsen« (Maas 1960, S. 30).

⁵² Lauer 2020, S. 157. Die Rede ist hier sowohl von Jacob Grimm als auch von Wilhelm Grimm. Siehe zu der Frage der Methodenstrenge auch Anm. 51.

⁵³

»While the humanities were supposed to search for the unique, the sciences would deal with the general. This vision turned out to be extremely influential as it gave the humanities a powerful identity enabling them to differentiate and emancipate themselves from the other disciplines. This constitutive separation between the humanities and sciences, however, did not correspond to actual practice in the humanities before the nineteenth century, as we have already seen. [...] When Dilthey's and Windelband's visions were gaining ground – from the early twentieth century onwards – modelling practices in the humanities continued«

(Bod 2018, S. 85).

⁵⁴ Vgl. Lauer 2020, S. 156. Während Jacob Grimms Errungenschaften im Bereich der Linguistik bis dato Anerkennung finden, ist ihm in philologischer Hinsicht allerdings nachgesagt worden, er habe im Gegensatz zu Lachmann für eine »wilde Philologie« gestanden, vgl. Wyss 1979. Tatsächlich ist es so, dass Jacob Grimm und Karl Lachmann zu zwei Polen stilisiert worden sind, deren Legitimation durchaus zu hinterfragen ist, die in diesem Zusammenhang aber dennoch erwähnt werden sollten. Insbesondere zu Lebzeiten galt Lachmann im Gegensatz zu Grimm als Vertreter von Methodenstrenge, während Jacob Grimm im Gegensatz zu Lachmann »intuitives Nachfühlen« (Wolf 2012, S. 93) versinnbildlichte. Für mehr Informationen zu dem »bipolaren« Narrativ über beide Wissenschaftler, vgl. Wolf 2012, S. 93–98.

⁵⁵ Grimm 2016, S. 63.

sich dieser Argumentation nicht anschließen, um zu erkennen, dass die Frage nach einer *Exaktheit* in den Geisteswissenschaften keine Frage von Zahlen oder Zählungen ist, sondern vielmehr zur Disposition stellt, was sich daraus ableitet.

Anders formuliert: Was macht eine Methode in den Geisteswissenschaften zu einer genauen Methode? Der Einsatz von Mess- und Maßeinheiten? Kann eine genaue Methode in den Geisteswissenschaften genaue Ergebnisse, aber ungenaue Erkenntnisse, hervorbringen und wäre ein solcher Fall dann Grund, von einem *Scheitern* zu sprechen? Kann im umgekehrten Fall eine ungenaue Methode eine genaue Erkenntnis hervorbringen und wie ließe sich in beiden Fällen nicht die Methode, sondern die *Interpretation* der jeweiligen Ergebnisse nachvollziehen? Lauer sagt selbst, dass »auch das Zählen [...] das Interpretieren [braucht]«⁵⁶, nur bleibt offen, was diesen letzten Schritt in den Geisteswissenschaften genau ausmacht; dabei ist das Interpretieren, das *Verstehen*, genau jene erkenntnistheoretische Chiffre, an der sich die Theorie der Geisteswissenschaften seit jeher abarbeitet.

[18]

Wer Verstehen sagt, muss auch Dilthey sagen. Lauer wirft den Kritiker*innen der Digital Humanities – bei ihm überwiegend bezogen auf Alison Louise Kennedy und Nan Z. Da⁵⁷ – vor, in dieser Tradition zu stehen: »Tatsächlich übernehmen Kennedy, Da und andere in den gegenwärtigen Debatten um die Möglichkeiten der Digital Humanities für ihre Behauptungen nur ältere Thesen vor allem von Wilhelm Dilthey.«⁵⁸ Freilich erwähnen weder Kennedy noch Da Dilthey mit auch nur einer einzigen Silbe, was ebenso für andere bekannte Kritiker*innen der Digital Humanities gilt.⁵⁹ Es ist unklar, wer sonst mit den »anderen« gemeint sein könnte. Inwiefern die genannten Autor*innen Kenntnis von Dilthey und seinen Thesen haben, lässt sich ihren Schriften aufgrund mangelnder Erwähnung desselben nicht entnehmen; auch die Unterscheidung von *inexact* und *exact sciences* ist in den anglophonen DH-Diskursen bisher nicht besprochen worden. Man kann zwar einige der in den dortigen Texten vage mitschwingenden Empfindungen auf hiesige Debatten übertragen und unterstellen, dass sie aus einer verwandten Geisteshaltung heraus entstehen, aber es wäre wünschenswert, hier, soweit möglich, genau vorzugehen und zu definieren, welche Geisteshaltungen auf welche Art und Weise spezifisch verwandt sein sollen. Insbesondere bei Dilthey drängt sich der Eindruck auf, dass er, wie man in der englischen Umgangssprache zu sagen pflegt, mietfrei in den Köpfen seiner Kritiker*innen lebt. Das ließ sich schon bei Rens Bod beobachten, dessen Interesse hauptsächlich, wenn nicht gar ausschließlich, Formen der Mustererkennung gilt und der Dilthey und die Hermeneutik als vorgebliches Sinnbild für eine eher der Wahrsagerei denn der Wissenschaft nahestehende Methode in seiner Geschichte der Geisteswissenschaften mit einer betont beiläufigen Erwähnung straft.⁶⁰ Joris van Zundert hat die Oberflächlichkeit einer solchen Charakterisierung zurecht deutlich kritisiert⁶¹ und Andreas Fickers hat in seinem Gelehrtenstreit mit Bod darüber hinaus Argumente angebracht, die es in ihrer Konsequenz, wenn auch nicht in ihrer rhetorischen Schärfe, eher zu beachten gälte als eine imaginierte Dilthey-Tradition der Teesatzleserei;⁶² namentlich die Notwendigkeit für eine digitale Hermeneutik, also eine Quellenkritik, die überhaupt erst zum Verständnis digital-kultureller Artefakte – und das schließt Forschungsergebnisse in den Digital Humanities ein – befähigt.⁶³

[19]

⁵⁶ Lauer 2020, S. 165.

⁵⁷ Vgl. Lea 2016; Da 2019.

⁵⁸ Lauer 2020, S. 163. Und hernach wiederholt: »Dilthey und seine [Nachfolger].«

⁵⁹ Vgl. Fish 2012a; Fish 2012b; Kirsch 2014; Hafner 2016; Fish 2018; Klaue 2018; Görner 2019.

⁶⁰ Vgl. Bod 2013, S. 333f.

⁶¹ »In his recent history of the humanities, Rens Bod dedicates a mere two pages to the concept and history of hermeneutics [...]. He disposes of the »method« as being based on guesswork and premonitions. This dismissal might be cast aside as anecdotal were it not for Bod's position as professor of computational and digital humanities« (Zundert 2016, S. 340f).

⁶²

»Bod's research agenda for the humanities 2.0 reminds me of a positivist manifesto from the nineteenth century, translated into the digital jargon of the twentieth first [sic!] century and driven by the idea (better: ideology) that digital technologies will finally offer the tools to detect and uncover the (so far hidden) logical foundations of all

Es steht außer Frage, dass man über statistische Kenntnisse verfügen muss, wenn man statistische Methoden anwendet – doch wie verhält es sich beispielsweise mit der Kompetenz, Informationen angemessen zu visualisieren und zu kommunizieren oder auf Rezipient*innenseite entsprechend zu prozessieren? Wie verhält es sich um das Bewusstsein, wie etwas entstanden, wo seine Provenienz, wann das Ende seines Lebenszyklus? Wenn wir annehmen, dass die Geisteswissenschaften evidenzorientiert sind (und das waren sie schon bei Dilthey, denn wo keine Überlieferung, da keine Verstehensgrundlage) – wenn wir also annehmen, dass die Geisteswissenschaften evidenzorientiert sind und wenn wir weiterhin annehmen, dass die Digital Humanities an der Evidenz rühren und nicht etwa an der Fähigkeit des Erkennens an und für sich – wenn wir also annehmen, dass die Digital Humanities unseren Zugang zu bestehender Evidenz verändern, das heißt unseren Blick auf diese Evidenz, und wenn wir weiterhin annehmen, dass die Digital Humanities darüber hinaus den Bestand an Evidenz erweitern, dann ergibt sich daraus zunächst einmal kein fundamentaler Eingriff in Erkenntnisprozesse, die einen Verstand voraussetzen und sich an einen anderen Verstand richten.⁶⁴ Diltheys psychologische Überlegungen waren gerade in seinem Spätwerk eben keine Vereinzelnungsargumente, sondern zielten auf ein Bewusstsein von Zeithorizonten, innerhalb derer jedes denkende und fühlende Subjekt seine Aussagen über die eigene Erlebniswelt zu veräußern hat. Letztlich laufen die erkenntnistheoretischen Diskurse in den Digital Humanities an solchen Stellen auseinander, weil sie die Begründung eines geisteswissenschaftlichen Argumentes traditionell mit Konzepten wie Divination und Einfühlung⁶⁵ assoziieren, was eine Vereinbarkeit mit Konzepten der Digital Humanities in weite Ferne zu rücken lassen scheint – und das, obschon die Fähigkeit des Perspektivwechsels lediglich einen Akt der Plausibilisierung darstellt, der für das Verständnis der Handlungen und Gedanken menschlicher Akteur*innen, so es darum geht, tatsächlich unabdingbar ist, während das eigentliche geisteswissenschaftliche Argument auch traditionell in der fragmentarischen Überlieferung von Kulturerbe wurzelt, deren Lacunae durch Inferenz überbrückt werden müssen. Folglich setzt die Interpretation eine Kenntnis des Untersuchungsgegenstandes und eine Kenntnis von Methoden voraus, die auf diesen Untersuchungsgegenstand angewandt werden, und wenn auf diesen Untersuchungsgegenstand keine Methoden angewandt werden außer die der Hermeneutik, dann liegt der Argumentation dennoch eine Ausgangsevidenz zugrunde. Dilthey nennt diese Evidenz die »Reste menschlichen Daseins«⁶⁶ und schreibt darüber hinaus:

»Mit der Auslegung der auf uns gekommenen Reste ist innerlich und notwendig die Kritik derselben verbunden. [...] Auslegung und Kritik haben im geschichtlichen Verlauf immer neue Hilfsmittel zur Lösung ihrer Aufgabe entwickelt, wie die naturwissenschaftliche Forschung immer neue Verfeinerungen des Experiments.«⁶⁷

In diesem Sinne erscheint es naheliegend, Forschung in den Digital Humanities als Schicht zu betrachten, die sich zwischen Ausgangsevidenz und Interpretation derselben legt, ohne das eine oder andere zu ersetzen, sondern um – im besten Falle – das eine, das andere oder beides zu ergänzen. Aus diesem Grund ist es auch irreführend, die Digital Humanities als *tool box* zu betrachten, die sich über ihre Methoden definiert; stattdessen müssen sie über ihr Verhältnis zu »the research objects they study and the research questions

human activities [...]. Without denying the possibility of structural causalities, Dilthey emphasised the individuality of perception, imagination and reasoning in order to develop a critical approach to history paying attention to both structural forces and individual agency«

(Fickers 2013, S. 159f.).

⁶³ Vgl. für seine Kritik an Bod Fickers 2013; für seine Einordnung von Ranke, Droysen und Dilthey in DH-Diskurse Fickers 2012; und für sein Plädoyer für eine digitale Hermeneutik Fickers 2020.

⁶⁴ Vgl. in diesem Kontext auch Piper 2020.

⁶⁵ Weil es in diesen Diskussionen traditionell oft zur Sprache kommt, sei hierzu angemerkt, dass Dilthey gemeinhin für eine Einfühlungssemantik vereinnahmt worden ist, obwohl er selbst so gut wie nie von Einfühlung gesprochen hat und man sich hierfür eher Husserl zuwenden müsste (was mit seiner Arbeit zur Intersubjektivität einhergeht, von dieser aber zu unterscheiden ist), vgl. Makkreel 1996 und Gallagher 2019.

⁶⁶ Dilthey 1965, S. 217.

⁶⁷ Dilthey 1965, S. 217.

they aim to answer«⁶⁸ bestimmt werden, wie Michael Piotrowski und Mateusz Fafinski bereits ausgeführt haben. Zum Aspekt der Erkenntnis sollte zuletzt noch angemerkt werden, dass das Verständnis eines kulturellen Erzeugnisses, das aus einem Geist entstanden und dessen Geist mutmaßlich nur in einem Geist zu erahnen (siehe August Boeckhs *Erkenntnis des Erkannten*⁶⁹), etwas anderes ist als eine Erkenntnis *über* ein kulturelles Erzeugnis oder seine Entstehung oder seine Einbettung in einen größeren historischen Kontext. Insofern die Geisteswissenschaften für die Untersuchung ihrer weitreichenden Fragestellungen alle Evidenz einbeziehen, auch eine Evidenz, die sie selbst auf Grundlage der Überlieferung generieren und deren Bewertung sie nur vornehmen können, wenn sie in der Lage sind, diesen Vorgang nachzuvollziehen, braucht es die Differenzierung von Evidenzen für ein Argument anhand methodenkritischer und auch methodenspezifischer Kriterien; was jedoch nicht in das Verstehen eingreift, das der Beweisführung Auftrieb gibt.

Um auf die Frage des Scheiterns zurückzukommen: Forschung scheitert nicht, sie lernt dazu. Formen der Weiterentwicklung, die auch bei Dilthey anklingen und von ihm gutgeheißen werden, bedingen sich aus anderen Formen der Weiterentwicklung, wie man sie etwa im technologischen Bereich beobachten kann. Gegebenenfalls ließe sich argumentieren, dass Forschung zwar nicht als gescheitert, aber als verfehlt oder fehlgeleitet gelten muss, wenn sie sich durch Unkenntnis der Evidenz, Unkenntnis der Methoden oder Unkenntnis ihrer eigenen Ergebnisse auszeichnet; das würde dann allerdings für jede Art von Forschung gelten, ungeachtet ihres disziplinären (und damit auch immer wissenschaftshistorischen) Zuschnitts.⁷⁰

[23]

4. Experimente, Laboratorien, Projekte

Warum aber sprechen die Digital Humanities nun so oft von ihrem eigenen Experimentcharakter? Es ließe sich spekulieren, dass damit eine bewusste Grenzüberschreitung intendiert ist, hin zu den Naturwissenschaften, die vermeintlich objektiver, wissenschaftlicher, empirischer.⁷¹ Aus geisteswissenschaftlicher Sicht kommen Zitate wie dieses in den Sinn: »One cannot put the Roman Empire in a test tube, add a dash of Christianity, and watch to see whether it rises or falls.«⁷²

[24]

Dazu sei gesagt, dass es immer auf die Variablen ankommt oder besser gesagt auf die Komplexität der involvierten Entitäten, ob man sie dahingehend modellieren kann, dass sich daraus extrafaktische oder kontrafaktische Szenarien ableiten lassen. Im Wissenschaftsrepertoire reiht sich das Experiment unter anderem neben die Beobachtung von Evidenz, die Modellierung von Evidenz, die Manipulation von Evidenz und die Simulation von Evidenz (die letzteren beiden als Sonderformen der Modellierung). Der sogenannte ›Untergang‹ des Römischen Reiches eignet sich nach derzeitigem Kenntnisstand nicht für solche Gedankenspiele, da es sich dabei bereits um die deutende Beschreibung eines multikausalen Vorgangs handelt, der sich nicht an einzelnen Ereignissen festmachen lässt und dessen Bewertungsgrundlage entsprechend zu polyvalent für solche Vereinfachungen ist. Gleichzeitig handelt es sich um eine vereinfachte, stark romantisierende Vorstellung, man müsse in den Naturwissenschaften bloß Reagenzien vermischen und dann zusehen, was passiert. Selbst wenn dem so wäre: Was reagiert in den Digital Humanities womit?

[25]

⁶⁸ Piotrowski / Fafinski 2020, S. 178.

⁶⁹ Vgl. hierzu Rodi 1990, S. 70–88. Diese Maxime muss in Verbindung zu der »Anerkennung des Prinzips der produktiven Unergründlichkeit und ihres methodischen Korrelats, der approximativ-unendlichen Verstehensaufgabe« (Rodi 1990, S. 97) gesehen werden.

⁷⁰ John Unsworth, der die »importance of failure« so sehr betont hat, begreift die Formulierung eigener Forschungsziele (und Kriterien für Erfolg und Misserfolg eines Projektes) als Voraussetzung für eine Anwendung von Poppers Falsifikationsprinzip auf digital-geisteswissenschaftliche Fragestellungen und Projekte, was noch erwähnt sein soll, um seine Beweggründe einzuordnen; vgl. Unsworth 1997.

⁷¹ Die Geschichtswissenschaften haben sich der Empirie freilich nie verwehrt, um nur ein Beispiel zu nennen. Ein Ansatz, der im Gegensatz zu Windelband (mit seiner Unterscheidung zwischen nomothetischen und idiographischen Wissenschaften) Gesetzmäßigkeiten in der historischen Forschung betont und sie als Teil der empirischen Wissenschaften definiert hat, findet sich bereits bei Hempel 1942.

⁷² Moore 2017, S. 4.

Ist es nicht ebenso naiv, anzunehmen, man könne Daten in Kontakt mit bestimmten Verarbeitungsschritten bringen und hätte damit die Hürde eines wissenschaftlichen Experimentes genommen? Zusehen, was passiert?

An dieser Stelle sei angemerkt, dass es womöglich zum eigenen Schaden der Digital Humanities ist, mit welcher Persistenz sich die Experimenterrhetorik bereits festgesetzt hat, ohne eine wissenschaftstheoretische Auseinandersetzung mit dieser sprachlichen Praxis zu zeitigen. Zumindest in Teilen scheint sich diese Praxis aus den Innovationsansprüchen des Feldes zu speisen, bei denen das *Experiment* als Antonym zu ›etablierten Methoden und Verfahren‹ agiert und sich damit stilistisch in die Verwendung des Experimentbegriffs in den Kulturwissenschaften einreicht:

[26]

»Die Kunst-, Kultur- und Literaturwissenschaften dagegen selektieren in ihren Verwendungen des Experimentbegriffs in der Mehrzahl der Fälle nur einen seiner semantischen Aspekte: den eines einmaligen Aktes des Ausprobierens neuartiger (künstlerischer) Techniken. Sie konturieren ›Experimentieren‹ als sowohl innovativen als auch singulären (nicht-reproduzierbaren) Akt der Erfindung, Entdeckung oder Schöpfung. Damit scheinen sie die Prägungen des naturwissenschaftlichen Experimentierverständnisses zu reduzieren. Doch blenden sie die naturwissenschaftlich dominierte Definitionsmacht des Begriffs durchaus nicht aus, die als expliziter oder impliziter Referenzpunkt stets erhalten bleibt.«⁷³

[27]

Nun ergeben sich dadurch zwei Schwierigkeiten: Zum einen überdeckt die starke Betonung eines Neuheitswertes ältere Forschungsliteratur und Forschungslogiken, die eine längere Tradition haben, als ihre *experimentelle* Anwendung suggerieren würde. Hierzu zählt insbesondere, wie auch bei Lauer sinnvoll dargelegt, der Einsatz von quantitativen Methoden in den Geisteswissenschaften, der prinzipiell weder besonders neu noch besonders kritikwürdig ist.⁷⁴ Zum anderen wird die Verwendung des Experimentbegriffs in den Digital Humanities durch die Schaffung von *Labs* als Arbeitsorten potenziert und stärker in die Nähe einer (Natur-)Wissenschaftlichkeit gerückt, die weder gegeben noch über eine allgemeine Legitimierungs- und Autorisierungsfunktion hinaus mutmaßlich intendiert ist. Keine Einrichtung, die in Zusammenhang mit den Digital Humanities als *Lab* bezeichnet wird, muss hierfür bestimmte Kriterien erfüllen, was die Vergleichbarkeit einschränkt und nahelegt, dass die assoziative Kraft des erfolgsdruckbefreiten Ausprobierens, händischen Arbeitens und kollaborativen gemeinschaftlichen Entdeckens im Vordergrund stehen soll. Hierzu passt etwa, dass dezidiert von dem Amalgam des »collaboratory«⁷⁵ gesprochen wird und man mitunter das *Lab* selbst zum *Experiment* deklariert.⁷⁶

[28]

In Anlehnung an die Ausführungen von Nowwiskie, die, wie anfangs erwähnt, genau diese Aspekte in Bezug auf das ›Scholars' Lab‹ an der University of Virginia hervorgehoben hat,⁷⁷ kann man sich an die Frühphase der analytischen Chemie erinnert fühlen, als Laboratorien im Entstehen begriffen waren, die im Laufe der Zeit dazu übergingen, ganz bestimmten Zwecken zu dienen, ganz bestimmte Bedingungen herzustellen und ganz bestimmte Forschung überhaupt erst zu ermöglichen, physisch, vor Ort, in der Beobachtung bestimmter

[29]

⁷³ Berg 2009, S. 53.

⁷⁴ Vgl. hierzu in Bezug auf die Literaturwissenschaft auch Bernhart 2018. Für die quantitative Geschichtswissenschaft des 20. Jahrhunderts sei stellvertretend auf die französische Annales-Schule und die Bielefelder Schule verwiesen, vgl. hierzu Buchner et al. 2020. Publikationen in den Digital Humanities wie McGillivray et al. 2018 gehen nicht auf solche Vorläufer ein.

⁷⁵ Vgl. Siemens / Siemens 2012. Man beachte das Resümee: »In the end, the experiment could be judged to be a failure as measured by many common benchmarks. Despite attempts to fuse what the lab felt to be the most desirable features of the single-researcher directed ›collaboratory‹ and that of the multiple-researcher directed ›co-laboratory‹, the lab became neither« (Siemens / Siemens 2012, S. 374).

⁷⁶ Vgl. Caton et al. 2017.

⁷⁷ Vgl. Nowwiskie 2012.

beeinflusster und unbeeinflusster Prozesse.⁷⁸ Dabei gilt es allerdings das zu bedenken, was Justus von Liebig in Hinblick auf den Zustand der analytischen Chemie zu Beginn des 19. Jahrhunderts so treffend beschrieben hat:

»Chemische Laboratorien, in welchen Unterricht in der Analyse erteilt wurde, bestanden damals nirgendwo; was man so nannte, waren eher Küchen, angefüllt mit allerlei Oefen und Geräthen zur Ausführung metallurgischer oder pharmaceutischer Prozesse. Niemand verstand eigentlich die Analyse zu lehren.«⁷⁹ [30]

Es ist zum gegenwärtigen Zeitpunkt eine offene Frage, ob sich die Laboratorien in den Digital Humanities an den Wortschöpfungen des Kreativsektors orientieren oder einen ähnlichen wissenschaftlichen Wandlungsprozess durchlaufen werden. Zumindest impliziert ihre Existenz Rahmenbedingungen, innerhalb derer Parameter für ein überwachtes, replizierbares, über sich selbst hinausdeutendes und in ein externalisiertes Ganzes hineindeutendes Schaffen festgelegt werden können. Dieser Punkt deutet eine weitere offene Frage an, nämlich die Frage, wie es um die Reproduzierbarkeit in den Digital Humanities bestellt ist, wobei dabei zwischen einer Nachvollziehbarkeit des Verfahrens und einer Wiederholbarkeit des Ergebnisses unterschieden werden sollte. Eine Erörterung dieses Aspektes würde an dieser Stelle zu weit führen, er soll aber erwähnt sein, da sich im Zuge der gesamtwissenschaftlichen »Replikationskrise«⁸⁰ durchaus der Gedanke auftut, man sollte Möglichkeiten des wissenschaftlichen *Scheiterns* womöglich auch in dieser Hinsicht thematisieren. Läge das Problem dann aber in der Wissenschaftstheorie oder nicht vielmehr in den Bedingungen, unter denen Wissenschaft stattfindet? [31]

Das führt uns in unserer Betrachtung verschiedener Dimensionen des *Scheiterns* zuletzt noch einmal in realpolitische Zusammenhänge. Wenn es um das projektbasierte Arbeiten in den Digital Humanities geht und wir zugleich anerkennen, dass das Festschreiben von Zielvorgaben Projekte in die missliche Lage bringt, ambitioniert zu sein und dennoch *produktorientiert*, so verwundert es nicht, dass sich aus dieser Quelle der großen Versprechungen die feuilletonistische Kritik eines Urs Hafner speist, der im Fall der Schweiz eine negative Bilanz dortiger DH-Großprojekte gezogen hat.⁸¹ Es wäre wahrscheinlich sinnvoll, aus dem Bewusstsein um die Versagensgründe, die sicherlich von Fall zu Fall zu unterscheiden und nicht immer mit Pauschalurteilen des *Scheiterns* treffend beschrieben sind, wie Tobias Hodel betont,⁸² eine Fachkultur der Aufarbeitung zu entwickeln, die sich offen und ehrlich mit pragmatischen Aspekten der Projektdurchführung in den Digital Humanities auseinandersetzt. Dass ein *Scheitern* teilweise gar intendiert ist oder zumindest in Kauf genommen wird, hebt Lisa Spiro mit einem Verweis hervor, der noch in eine andere Richtung deutet: [32]

»Not all experiments succeed as originally imagined, but the digital humanities community recognizes the value of failure in the pursuit of innovation. »[T]o encourage innovations in the digital humanities,« the National Endowment for the Humanities offers »Digital Humanities Start-Up Grants,« which »are modeled, in part, on the »high risk/high reward« paradigm often used by funding agencies in the sciences« (National Endowment for the Humanities). Failure is accepted as a useful result in the digital humanities, since it indicates that the experiment was likely high risk and means that we collectively learn from failure rather than reproducing it (assuming that the failure is documented).«⁸³ [33]

⁷⁸ Neben den *laboratory studies* in der Nachfolge Karin Knorr-Cetinas (vgl. Knorr-Cetina 1984), die gegenwärtige naturwissenschaftliche Arbeitspraktiken kartographiert, lohnt ein Blick in die Wissenschaftsgeschichte, die unter anderem auch den Aspekt der Kollaboration, den die Digital Humanities so sehr betonen, unter anderen Vorzeichen aufarbeitet (und entsprechende »unsichtbare« Arbeitskräfte ließen sich auch in der Geschichte der Geisteswissenschaften finden); vgl. Hentschel 2008.

⁷⁹ Liebig 1890, S. 822.

⁸⁰ Diese Replikationskrise ist auch in der Öffentlichkeit diskutiert worden, vgl. Wagner 2020.

⁸¹ Vgl. Hafner 2021.

⁸² »Dennoch möchte Hodel nicht von Scheitern sprechen. Die Projekte seien für die digitalen Geisteswissenschaften wichtige Ausbildungsstationen gewesen und hätten den »State of the Art« festgelegt« (Hafner 2021).

⁸³ Spiro 2012, S. 29.

In dieser Sicht steht nicht der Erkenntnisgewinn im Mittelpunkt, sondern die vermeintliche Fallhöhe. Max Kemman rückt diese Form des *Scheiterns* in die Nähe von Geschäftsmodellen und Vermarktungsstrategien aus der Gründer*innenszene, wo unter Vermeidung von fundierter Forschung und entsprechender Vorbereitung von Vorhaben Formen des *prototyping* und andere teils bloß als Nebelkerzen kurz aufleuchtende Aktivitäten favorisiert werden.⁸⁴ Mangels Fachkenntnis lohnt es sich an dieser Stelle nicht, näher auf solche Überlegungen einzugehen; es sollte genügen, darauf zu verweisen, dass die Digital Humanities universitär verankert sind und den Richtlinien der Forschungsförderung entsprechen müssen, wenn sie gefördert werden wollen; inwiefern sich verschiedene gesellschaftliche, wissenschaftspolitische und sozioökonomische Tendenzen gegenseitig bedingen, sei dahingestellt. Da Vorstellungen des *Scheiterns* in den Digital Humanities allerdings eng mit dieser Fragestellung verknüpft zu sein scheinen, ist es erstens notwendig, diesen Umstand zu benennen, und zweitens, den Dialog differenziert zu führen.

[34]

5. Fragen

Es gäbe noch viel zu sagen, doch die Zeichenzahl ist begrenzt. Daher in aller gebotenen Kürze: Wenn Quinn Dombrowski von Kommunikationsfehlern spricht,⁸⁵ dann handelt es sich dabei nicht unmittelbar um spezifisch digitalgeisteswissenschaftliche Problematiken. Es mag sein, dass das Ausmaß solcher Herausforderungen in interdisziplinären Zusammenhängen eine neue Qualität annimmt; und die begriffliche Präsenz von *collaboration*, sowohl in den ›Lab‹-Diskursen als auch in DH-Diskursen allgemein, kann hier als Fingerzeig dienen.⁸⁶ Was sich daraus als Handlungsempfehlung ableiten soll, ist momentan noch unklar. Überhaupt deutet eine Durchsicht der Forschungsliteratur an, dass in vielen Punkten Diskussionsbedarf besteht, der nicht immer oder nicht einmal hauptsächlich in der Forschungsliteratur selbst gedeckt werden kann, auch wenn er nicht spurlos an ihr vorübergeht. Daher kann dieser Themenkomplex um Formen des Scheiterns, Formen von Experimenten und Formen von Laboratorien keinesfalls als abgeschlossen erachtet werden. Vielmehr gilt es, folgenden Desideraten weitere Aufmerksamkeit zu schenken:

[35]

1. In den Digital Humanities vermischen sich verschiedene Fachkulturen, Denktraditionen, auch sprachlich und damit ideengeschichtlich vormals abgegrenzte Diskurse. Frage: Wie lässt sich ein wissenschaftstheoretischer Dialog aufbauen, wenn nicht in der direkten Konfrontation dieser Umstände?
2. Die Bandbreite an *Experimenten* und was mit dem Begriff gemeint sein kann, ist in den Digital Humanities nur vage umrissen. Frage: Wann verliert ein DH-Experiment den Status eines Experiments?
3. Erfolg und Misserfolg sind keine geeigneten Richtgrößen, um Forschung zu evaluieren. Man muss aber den Wert dessen, was man tut, erkennen und bewerten können. Man muss die ›Analyse lehren‹. Nur dann lassen sich Ergebnisse einordnen. Frage: Wer lehrt das Verstehen?

Da die Fragen rhetorisch klingen mögen, hier noch einige kurze Anmerkungen für ihre ausstehende Beantwortung:

[36]

Zu 1.: Dialog setzt verschiedene Gesprächspartner voraus. Vertreter*innen einer Kulturkritik an den Digital Humanities wie Domenico Fiormonte haben dabei unter anderem den Monolingualismus des Faches im Blick, der zumindest in seiner internationalen Ausrichtung Wissensaustausch fördert, aber auch behindert (am offenkundigsten, weil die sprachliche Kommunikation und Vermittlung auf akademischem Niveau sehr voraussetzungsreich ist und viele Fachtraditionen insbesondere in den Geisteswissenschaften eng mit ihren Landessprachen und -geschichten verwoben sind, von denen sie sich nicht immer in derselben

[37]

⁸⁴ Vgl. Kemman 2019 und Hall 2013. Eine gewisse Überhöhung des Scheiterns zeigt sich auch in folgender Publikation zur Startup-Kultur, in der fälschlicherweise (zumindest ohne Quellenangabe) Wilhelm Busch ein landläufiger Sinnspruch zugeschrieben und in das übliche Narrativ des ›Scheiterns als Chance‹ eingebettet wird: »Durch Fehler wird man klug, darum ist einer nicht genug.« Und passend von den Autoren selbst auf derselben Seite: »Scheitern muss als Chance aufgefasst werden« (Bogott et al. 2017, S. 166).

⁸⁵ Vgl. Dombrowski 2019.

⁸⁶ Vgl. Spiro 2012, S. 25f.; Siemens / Siemens 2012; Hunter 2014; aber auch bspw. Deegan / McCarty (Hg.) 2012.

Diskussionstiefe qua Übersetzung lösen können oder wollen).⁸⁷ Für die deutschsprachigen digitalen Geisteswissenschaften sollte es allerdings zweitrangig sein, ob und inwiefern theoretische Schriften wie jene von Dilthey in anglophonen DH-Publikationen rezipiert werden oder nicht; sie sollten nur darüber die Wissenschaftsgeschichte der deutschsprachigen Geisteswissenschaften und die damit verbundenen Diskurse selbst nicht vergessen. Perspektivisch wäre ein internationaler Austausch über solche Traditionen aus verschiedenen nationalen Kontexten wünschenswert.

Zu 2.: Die Antwort hängt davon ab, was man unter einem Experiment versteht. Wenn ein Experiment in den Digital Humanities darin besteht, Methoden aus anderen Wissenschaften auf geisteswissenschaftliche Untersuchungsgegenstände anzuwenden, vorausgesetzt, das Vorgehen wäre in der gegebenen Kombination unerprobt und hätte noch keinen gesicherten Nutzen, so würde ein solches Verfahren in dem Moment seinen experimentellen Status verlieren, in dem der Nutzen erbracht wäre. Wenn ein Experiment in den Digital Humanities hingegen darin besteht, neue Formen der Zusammenarbeit zu erproben, neue Formen der Wissenschaftskommunikation einzuführen, neue Formen von Arbeitsorten zu schaffen, dann wäre dieses wissenschaftskulturelle Experiment erst mit einer Etablierung entsprechender Strukturen und Formate beendet oder in dem Moment bereits wieder von einer sich verändernden Praxis überholt und erneut zum Experiment geworden. Noch ein Aspekt ist wichtig: Experimente müssen nicht aufhören, Experimente zu sein; ganz im Gegenteil ist das Experiment selbst als wissenschaftliche Methode beständig und folgt seiner eigenen Logik. Wenn wir uns also vorstellen, es würde neben diesen anderen Definitionen eines Experimentes auch solche Experimente in den Digital Humanities geben, die gezielt darauf ausgerichtet sind, immer nur bestimmte Daten mit anderen Daten oder bestimmten Verarbeitungsschritten zusammenzubringen und dabei zuzusehen, was passiert, dann würde diese Art des Experimentes immer Experiment bleiben. Hierbei gälte es *black boxes* zu vermeiden, da ein Experiment, wenn es auf die Beobachtung zielt, eine Beobachtung ermöglichen muss. Weitere Formen von Experimenten in den Digital Humanities sind vorstellbar; nur sollte Klarheit darüber herrschen, was mit dem Begriff jeweils gemeint ist.

[38]

Zu 3.: Diese Frage ist in der Tat die entscheidende. Sie muss – über das Offensichtliche, also die universitäre Lehre, hinaus – zunächst offen bleiben.

[39]

Es sei noch ein Schlusswort erlaubt:

[40]

Ein Antrieb für Forschung ist die Neugier (oder die Unkenntnis). Doch worauf zielt die Vorstellung eines wiederholten Versagens? Auf das letzte Gelingen, zumindest, wenn man voraussetzt, dass jeder eintretende Lerneffekt einem Zweck dient, nämlich der Verbesserung und Verfeinerung des Vorgehens. Wissenschaftsgeschichte ist auch immer eine Geschichte, die von Fehlschlägen geprägt ist⁸⁸ – das gilt im Übrigen für alle Wissenschaften inklusive der Geisteswissenschaften, wenn wir Thesen und Theorien dazu zählen, die mittlerweile als widerlegt oder überholt gelten; in dem Fall würden wir normalerweise aber wohl kaum von Fehlschlägen sprechen, weil es gemeinhin als selbstverständlich gilt, dass nur in der Auseinandersetzung, in der Fortentwicklung wissenschaftlicher Standards, das heißt in einer zunehmenden Tiefe und Breite der Auseinandersetzung (angesichts eines stetig anwachsenden Korpus an Forschungsliteratur), dass sich also nur in der Argumentation für und wider eine These und damit einhergehend einer Rückversicherung anhand der Quellenlage eine Erkenntnis herauskristallisiert, die es wiederum zu widerlegen gilt. Bewerten kann man das nur, wenn man Methode, Zielstellung und das Verhältnis beider zum untersuchten Material, respektive Datenmaterial, durchdrungen hat. In jedem Fall braucht es nicht nur Mensch oder Maschine, sondern den Wunsch zu wissen; so gut es geht zu wissen; sonst bleibt jedes Bemühen, um auf Unsworths ›new ignorance‹ zurückzukommen, nur Teil von jener Kraft, die stets das Wissenschaftliche will und stets das Nichtwissen schafft.

[41]

⁸⁷ Vgl. Fiormente 2021.

⁸⁸ Vgl. Pennock 2019, S. 157–159.

Bibliografische Angaben

- Daniel Allington / Sarah Brouillette / David Golumbia: Neoliberal Tools (and Archives). A Political History of Digital Humanities. In: Los Angeles Review of Books. Artikel vom 01.05.2016. [\[online\]](#)
- Gunhild Berg: Zur Konjunktur des Begriffs ›Experiment‹ in den Natur-, Sozial- und Geisteswissenschaften. In: Wissenschaftsgeschichte als Begriffsgeschichte. Terminologische Umbrüche im Entstehungsprozess der modernen Wissenschaften. Hg. von Michael Eggers / Matthias Rothe. Bielefeld 2009, S. 51–82. DOI: [10.14361/9783839411841-002](#) [\[Nachweis im GVK\]](#)
- Toni Bernhart: Quantitative Literaturwissenschaft. Ein Fach mit langer Tradition? In: Quantitative Ansätze in Literatur- und Geisteswissenschaften. Systematische und historische Perspektiven. Hg. von Toni Bernhart / Marcus Willand / Sandra Richter / Andrea Albrecht. Berlin et al. 2018, S. 207–220. DOI: [10.1515/9783110523300-009](#) [\[Nachweis im GVK\]](#)
- Rens Bod: A New History of the Humanities. The Search for Principles and Patterns from Antiquity to the Present. Oxford 2013. [\[Nachweis im GVK\]](#)
- Rens Bod / Julia Kursell / Jaap Maat / Thijs Weststeijn: A New Field. History of Humanities. In: History of Humanities 1 (2016), H. 1, S. 1–8. DOI: [10.1086/685056](#) [\[Nachweis im GVK\]](#)
- Rens Bod: Modelling in the Humanities. Linking Patterns to Principles. In: Historical Social Research / Supplement 31 (2018), S. 78–95. DOI: [10.12759/hsr.suppl.31.2018.78-95](#) [\[Nachweis im GVK\]](#)
- Nicole Bogott / Stefan Rippler / Branko Woischwill: Im Startup die Welt gestalten. Wie Jobs in der Gründerszene funktionieren. Wiesbaden 2017. [\[Nachweis im GVK\]](#)
- Michael Buchner / Tobias A. Jopp / Mark Spoerer / Lino Wehrheim: Zur Konjunktur des Zählens – oder wie man Quantifizierung quantifiziert. Eine empirische Analyse der Anwendung quantitativer Methoden in der deutschen Geschichtswissenschaft. In: Historische Zeitschrift 310 (2020), H. 3, S. 580–621. [\[Nachweis im GVK\]](#)
- Roberto Busa: The Annals of Humanities Computing. The Index Thomisticus. In: Computers and the Humanities 14 (1980), H. 2, S. 83–90. DOI: [10.1007/BF02403798](#)
- Dino Buzzetti: The Origins of Humanities Computing and the Digital Humanities Turn. In: Humanist Studies & the Digital Age 6 (2019), H. 1, S. 32–58. DOI: [10.5399/uo/hsda.6.1.3](#) [\[Nachweis im GVK\]](#)
- Paul Caton / Ginestra Ferraro / Luis Figueira / Elliott Hall / Neil Jakeman / Pam Mellen / Anna-Maria Sichani / James Smithies / Miguel Vieira / Tim Watts / Carina Westling: Mechanizing the Humanities? King's Digital Lab as Critical Experiment. In: Digital Humanities 2017. Conference Abstracts. Hg. von Rhian Lewis / Cecily Raynor / Dominic Forest / Michael Sinatra / Stéfán Sinclair. (DH 2017, Montréal, 08.–11.07.2017) Montréal 2017, S. 184–186. PDF. [\[online\]](#)
- Paolo Chiesa: Principles and Practice. In: Handbook of Stemmatology. Hg. von Philipp Roelli. Berlin 2020, S. 74–87. DOI: [10.1515/9783110684384](#) [\[Nachweis im GVK\]](#)
- Gregory Crane: The Big Humanities, National Identity and the Digital Humanities in Germany. In: Digital Humanities Universität Leipzig. Blogbeitrag vom 08.06.2015. Revision vom 20.07.2015. Google-Doc. [\[online\]](#)
- Projektmanagement und Digital Humanities. Zur klugen Gestaltung der Zusammenarbeit. Hg. von Fabian Cremer / Swantje Dogunke / Anna Maria Neubert / Thorsten Wübbena. Bielefeld 2024. DOI: [10.1515/9783839469675](#)
- Brian Croxall / Quinn Warnick: Failure. In: Digital Pedagogy in the Humanities. Concepts, Models, and Experiments. Hg. von Rebecca Frost Davis / Matthew K. Gold / Katherine D. Harris / Jentery Sayers. 2015. [\[online\]](#)
- Nan Da: The Digital Humanities Debacle. In: The Chronicle Review 65 (2019), Ausgabe 29. Onlineartikel vom 27.03.2019. [\[online\]](#) Druckversion vom 05.04.2019. [\[Nachweis im GVK\]](#)
- Collaborative Research in the Digital Humanities. Hg. von Marilyn Deegan / Willard McCarty. London 2012. [\[Nachweis im GVK\]](#)
- Wilhelm Dilthey: Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften. 4. unveränderte Auflage. Ursprünglich 1910. Stuttgart 1965. (= Gesammelte Schriften, 7) [\[Nachweis im GVK\]](#)
- Quinn Dombrowski. Towards a Taxonomy of Failure. In: Quinn Dombrowski. Hg. von Quinn Dombrowski. Blogbeitrag vom 30.01.2019. [\[online\]](#)
- Amy E. Earhart: The Digital Humanities as a Laboratory. In: Between Humanities and the Digital. Hg. von Patrik Svensson / David Theo Goldberg. Cambridge, MA u. a. 2015, S. 391–400. [\[Nachweis im GVK\]](#)
- Andreas Fickers: Veins filled with the Diluted Sap of Rationality. A Critical Reply to Rens Bod. In: BMGN – Low Countries Historical Review 128 (2013), H. 4, S. 155–163. DOI: [10.18352/bmgn-lchr.9347](#) [\[Nachweis im GVK\]](#)
- Andreas Fickers: Towards a New Digital Historicism? Doing History in the Age of Abundance. In: Journal of European Television History and Culture 1 (2012), H. 1, S. 19–26. DOI: [10.18146/2213-0969.2012.jethc004/](#) [\[Nachweis im GVK\]](#)
- Andreas Fickers: Update für die Hermeneutik. Geschichtswissenschaft auf dem Weg zur digitalen Forensik? In: Zeithistorische Forschungen 17 (2020), H. 1, S. 157–168. DOI: [10.14765/zzf.dok-1765](#) [\[Nachweis im GVK\]](#)
- Domenico Fiormonte: Taxation against Overrepresentation? The Consequences of Monolingualism for Digital Humanities. In: Alternative Historiographies of the Digital Humanities. Hg. von Dorothy Kim / Adeline Koh. Santa Barbara 2021, S. 333–376. [\[online\]](#)
- Stanley Fish (2012a): The Digital Humanities and the Transcending of Mortality. In: The New York Times. Post vom 09.01.2012. [\[online\]](#)
- Stanley Fish (2012b): Mind Your P's and B's. The Digital Humanities and Interpretation. In: The New York Times. Post vom 23.01.2012. [\[online\]](#)
- Stanley Fish: Stop Trying to Sell the Humanities. In: The Chronicle Review 64 (2018), Ausgabe 38. Onlineartikel vom 17.06.2018. [\[online\]](#) Druckversion vom 22.06.2018.
- Julia Flanders: The Productive Unease of 21st-Century Digital Scholarship. In: Digital Humanities Quarterly 3 (2009), H. 3. [\[online\]](#)
- Shaun Gallagher: Dilthey and Empathy. In: Interpreting Dilthey. Critical Essays. Hg. von Eric S. Nelson. Cambridge 2019, S. 145–158. DOI: [10.1017/9781316459447.008](#) [\[Nachweis im GVK\]](#)
- Jean-Claude Gardin: L'interprétation dans les humanités. Réflexions sur la troisième voie / Interpretation in the Humanities. Some Thoughts on the Third Way. In: Interpretation in the Humanities. Perspectives from Artificial Intelligence. Hg. von John Richard Ennals / Jean-Claude Gardin. London 1990, S. 22–59. [\[Nachweis im GVK\]](#)
- (Keine) Experimente in den Digital Humanities? Hg. von vDhd. In: vDhd 2021 Experimente. Blogbeitrag vom 28.01.2021. [\[online\]](#)
- Rüdiger Görner: Wenn die Saat des Digitalen aufgeht. In: Frankfurter Allgemeine Zeitung. Beitrag vom 30.06.2019. [\[online\]](#)

Shawn Graham: *Failing Gloriously and Other Essays*. Grand Forks 2019. DOI: [10.31356/dpb015](https://doi.org/10.31356/dpb015)

Jacob Grimm: Über den Werth der ungenauen Wissenschaften. In: *Texte zur Theorie der Geisteswissenschaften*. Hg. von Athena Panteos und Tim Rojek. Basierend auf der Druckausgabe Hildesheim et al. 1966. Ursprünglich Rede auf dem Frankfurter Germanistentag, 1846. Stuttgart 2016, S. 58–63. [[Nachweis im GVK](#)]

Hans Ulrich Gumbrecht: Wer würde denn die Geisteswissenschaften vermissen? In: *Neue Zürcher Zeitung*. Beitrag vom 29.10.2019. [[online](#)]

Urs Hafner: Der Irrtum der Zeitmaschinen. In: *Neue Zürcher Zeitung*. Beitrag vom 27.05.2016. [[online](#)] [[Nachweis im GVK](#)]

Urs Hafner: Leichen, Totgesagte, Paralytisierte – wenn Geisteswissenschaftler auf Informatik treffen. In: *Neue Zürcher Zeitung*. Beitrag vom 04.01.2021. [[online](#)]

Erika Hall: How the ›Failure‹ Culture of Startups Is Killing Innovation. In: *Wired*. Beitrag vom 11.09.2013. [[online](#)]

Carl Gustav Hempel: The Function of General Laws in History. In: *The Journal of Philosophy* 39 (1942), H. 2, S. 35–48. [[Nachweis im GVK](#)]

Unsichtbare Hände. Zur Rolle von Laborassistenten, Mechanikern, Zeichnern u. a. Amanuenses in der physikalischen Forschungs- und Entwicklungsarbeit. Hg. von Klaus Hentschel. Diepholz 2008. [[Nachweis im GVK](#)]

Andrea Hunter: Digital Humanities as Third Culture. In: *Mediekultur. Journal of Media and Communication Research* 30 (2014), H. 57, S. 18–33. DOI: [10.7146/mediekultur.v30i57.16318](https://doi.org/10.7146/mediekultur.v30i57.16318) [[Nachweis im GVK](#)]

Andreas Kablitz: Apokalyptiker und Alimenterte. In: *Frankfurter Allgemeine Zeitung*. Beitrag vom 09.11.2019. [[online](#)]

Max Kemman: DH Failures vs Findings. In: Max Kemman. Hg. von Max Kemman. Blogbeitrag vom 28.02.2019. [[online](#)]

Adam Kirsch: Technology Is Taking Over English Departments. The False Promise of the Digital Humanities. In: *The New Republic*. Beitrag vom 02.05.2014. [[online](#)]

Magnus Klaue: Auf der Suche nach dem verlorenen Objekt. In: *Frankfurter Allgemeine Zeitung*. Beitrag vom 12.05.2018.

Karin Knorr-Cetina: Die Fabrikation von Erkenntnis. Zur Anthropologie der Wissenschaft. Frankfurt / Main 1984. [[Nachweis im GVK](#)]

Mareike König: Jenseits der Metaphorik. Experimente in den Digital Humanities #dhiha6. In: *digitale:geschichte*. Hg. von Digitale Humanities Universität Wien. Blogbeitrag vom 07.06.2015. [[online](#)]

Gerhard Lauer: Über den Wert der exakten Geisteswissenschaften. In: *Geisteswissenschaft – was bleibt? Zwischen Theorie, Tradition und Transformation*. Hg. von Hans Joas / Jörg Noller. Freiburg et al. 2020, S. 152–173. (= *Geist und Geisteswissenschaft*, 5) [[Nachweis im GVK](#)]

Sarah Lang: Experiments in the digital laboratory. What the Computational Humanities can learn about their definition and terminology from the history of Science. 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. Teilband 1. Esch-sur-Alzette 2022. DOI: [10.26298/melusina.8f8w-y749-eitd](https://doi.org/10.26298/melusina.8f8w-y749-eitd)

Richard Lea: ›It's Like Hitting a Painting with a Fish‹. Can Computer Analysis Tell Us Anything New About Literature? In: *The Guardian* / Books. Beitrag vom 15.09.2016. [[online](#)]

Georg von Liebig: Justus von Liebig. Eigenhändige biographische Aufzeichnungen. In: *Berichte der deutschen chemischen Gesellschaft* 23 (1890), H. 3, S. 817–828. [[Nachweis im GVK](#)]

Paul Maas: *Textkritik*. 4. Auflage. Ursprünglich 1927. Leipzig 1960. [[Nachweis im GVK](#)]

Rudolf Adam Makkreel: How is Empathy Related to Understanding? In: *Issues in Husserl's II*. Hg. von Thomas Nenon / Lester Embree. Dordrecht 1996, S. 199–212. (= *Contributions to Phenomenology*, 24) [[Nachweis im GVK](#)]

Barbara McGillivray / Jon Wilson / Tobias Blanke: Towards a Quantitative Research Framework for Historical Disciplines. In: *Proceedings of the Workshop on Computational Methods in the Humanities*. Hg. von Michael Piotrowski (COMHUM 2018, Lausanne, 04.–05.06.2018). Lausanne 2018, S. 53–58. PDF. [[online](#)] [[Nachweis im GVK](#)]

Barrington Moore, Jr.: *Soviet Politics – The Dilemma of Power. The Role of Ideas in Social Change*. Ursprünglich 1950. London 2017. Onlineausgabe. [[Nachweis im GVK](#)]

Bethany Nowwiskie: Too Small to Fail. In: Bethany Nowwiskie. Hg. von Bethany Nowwiskie. Blogbeitrag vom 13.10.2012. [[online](#)]

On Making in the Digital Humanities. Hg. von Julianne Nyhan / Geoffrey Rockwell / Stéfan Sinclair / Alexandra Ortolja-Baird. London 2023. DOI: [10.14324/111.9781800084209](https://doi.org/10.14324/111.9781800084209)

Fabian Offert / Peter Bell: Generative Digital Humanities. In: *Proceedings of the Workshop on Computational Humanities Research*. Hg. von Folgert Karsdorp / Barbara McGillivray / Adina Nerghes / Melvin Wevers. (CHR: 1, Amsterdam, 18.–20.11.2020). Amsterdam 2020, S. 202–212. [[online](#)]

Giovanni Palumbo: Criticism and Controversy. In: *Handbook of Stematology*. Hg. von Philipp Roelli. Berlin 2020, S. 88–108. [[Nachweis im GVK](#)]

Urszula Pawlicka-Deger: The Laboratory Turn. Exploring Discourses, Landscapes, and Models of Humanities Labs. In: *Digital Humanities Quarterly* 14 (2020), H. 3. [[online](#)]

Digital Humanities and Laboratories. Perspectives on Knowledge, Infrastructure and Culture. Hg. von Urszula Pawlicka-Deger / Christopher Thomson. London 2023. DOI: [10.4324/9781003185932](https://doi.org/10.4324/9781003185932)

Robert Tatnall Pennock: *An Instinct for Truth: Curiosity and the Moral Character of Science*. Cambridge, MA 2019. [[Nachweis im GVK](#)]

Michael Piotrowski / Mateusz Fafinski: Nothing New Under the Sun? Computational Humanities and the Methodology of History. In: *Proceedings of the Workshop on Computational Humanities Research*. Hg. von Folgert Karsdorp / Barbara McGillivray / Adina Nerghes / Melvin Wevers (CHR: 1, Amsterdam, 18.–20.11.2020). Amsterdam 2020, S. 171–181. [[online](#)]

Andrew Piper: *Can We Be Wrong? The Problem of Textual Evidence in a Time of Data*. Cambridge 2020. DOI: [10.1017/9781108922036](https://doi.org/10.1017/9781108922036)

Helle Porsdam: Digital Humanities. On Finding the Proper Balance between Qualitative and Quantitative Ways of Doing Research in the Humanities. In: *Digital Humanities Quarterly* 7 (2013), H. 3. [[online](#)]

Margarete Pratschke: Geschichte und Kritik digitaler Kunst- und Bildgeschichte. In: *Computing Art Reader. Einführung in die digitale Kunstgeschichte*. Hg. von Piotr Kuroczyński / Peter Bell / Lisa Dieckmann. Heidelberg 2018, S. 20–37. DOI: [10.11588/arthistoricum.413.c5767](https://doi.org/10.11588/arthistoricum.413.c5767)

Malte Rehbein: Digitalisierung in den Geisteswissenschaften (Digital Humanities). In: *Handbuch Industrie 4.0. Recht, Technik, Gesellschaft*. Hg. von Walter Frenz. Berlin 2020, S. 1425–1431. DOI: [10.1007/978-3-662-58474-3_73](https://doi.org/10.1007/978-3-662-58474-3_73) [[Nachweis im GVK](#)]

Frithjof Rodi: Erkenntnis des Erkannten. Zur Hermeneutik des 19. und 20. Jahrhunderts. Frankfurt / Main 1990. [[Nachweis im GVK](#)]

Paul Simon Rosenbloom: Towards a Conceptual Framework for the Digital Humanities. In: *Defining Digital Humanities. A Reader*. Hg. von Melissa Terras / Julianne Nyhan / Edward Vanhoutte. Farnham, Surrey 2013, S. 219–233. Ursprünglich erschienen in *Digital Humanities Quarterly* 6 (2012), H. 2, [[online](#)]

Making Things and Drawing Boundaries. Experiments in the Digital Humanities. Hg. von Jentery Sayers. Minneapolis 2017. DOI: [10.5749/9781452963778](https://doi.org/10.5749/9781452963778) [[Nachweis im GVK](#)]

Lynne Siemens / Raymond Siemens: Notes from the Collaboratory. An Informal Study of an Academic DH Lab in Transition. In: Digital Humanities 2012. Conference Abstracts. Hg. von Jan Christoph Meister / Katrin Schönert / Bastian Lomsché / Wilhelm Schernus / Lena Schüch / Meike Stegkemper / Benjamin W. Bohl / Daniel Röwenstrunk. (DH: 24, Hamburg, 16.–22.07.2012) Hamburg 2012, S. 373–375. [[Nachweis im GVK](#)]

Charles Percy Snow: The Rede Lecture (1959). In: The Two Cultures. Einleitung von Stefan Collini. Cambridge 1993, S. 1–52. [[Nachweis im GVK](#)]

Juliana Spahr / Richard So / Andrew Piper: Beyond Resistance. Towards a Future History of Digital Humanities. In: Los Angeles Review of Books. Artikel vom 11.05.2016. [[online](#)]

Lisa Spiro: »This Is Why We Fight«. Defining the Values of the Digital Humanities. In: Debates in the Digital Humanities. Hg. von Matthew K. Gold. Minneapolis 2012, S. 16–35. [[online](#)]

Melissa Terras: Disciplined. Using Educational Studies to Analyse »Humanities Computing«. In: Literary and Linguistic Computing 21 (2006), H. 2, S. 229–246. DOI: [10.1093/lilc/fql022](https://doi.org/10.1093/lilc/fql022) [[Nachweis im GVK](#)]

Ted Underwood: Dear Humanists. Fear Not the Digital Revolution. In: The Chronicle Review 65 (2019), Ausgabe 29. Onlineartikel vom 27.03.2019. [[online](#)]
Druckversion vom 05.04.2019.

John Unsworth: Documenting the Reinvention of Text. The Importance of Failure. In: Journal of Electronic Publishing 3 (1997), H. 2. DOI: [10.3998/3336451.0003.201](https://doi.org/10.3998/3336451.0003.201) [[Nachweis im GVK](#)]

Gerald Wagner: Wissenschaftliche Irrtümer in Serie. In: Frankfurter Allgemeine Zeitung. Beitrag vom 12.04.2020.

In the Shadows of the Digital Humanities. Hg. von Elizabeth Weed / Ellen Rooney. Durham 2014. (= Differences, 25.1) [[Nachweis im GVK](#)]

Wilhelm Windelband: Geschichte und Naturwissenschaft. Rede. 3., unveränderte Auflage. Straßburg 1904. [[Nachweis im GVK](#)]

Johanna Wolf: Kontinuität und Wandel der Philologien. Textarchäologische Studien zur Entstehung der Romanischen Philologie im 19. Jahrhundert. Tübingen 2012. [[Nachweis im GVK](#)]

Georg Henrik von Wright: Explanation and Understanding. London 1971. [[Nachweis im GVK](#)]

Ulrich Wyss: Die wilde Philologie. Jacob Grimm und der Historismus. München 1979. [[Nachweis im GVK](#)]

Joris van Zundert: Screwmeneutics and Hermeneumericals. The Computationality of Hermeneutics. In: A New Companion to Digital Humanities. Hg. von Susan Schreibman / Ray Siemens / John Unsworth. Hoboken 2016, S. 331–347. [[Nachweis im GVK](#)]