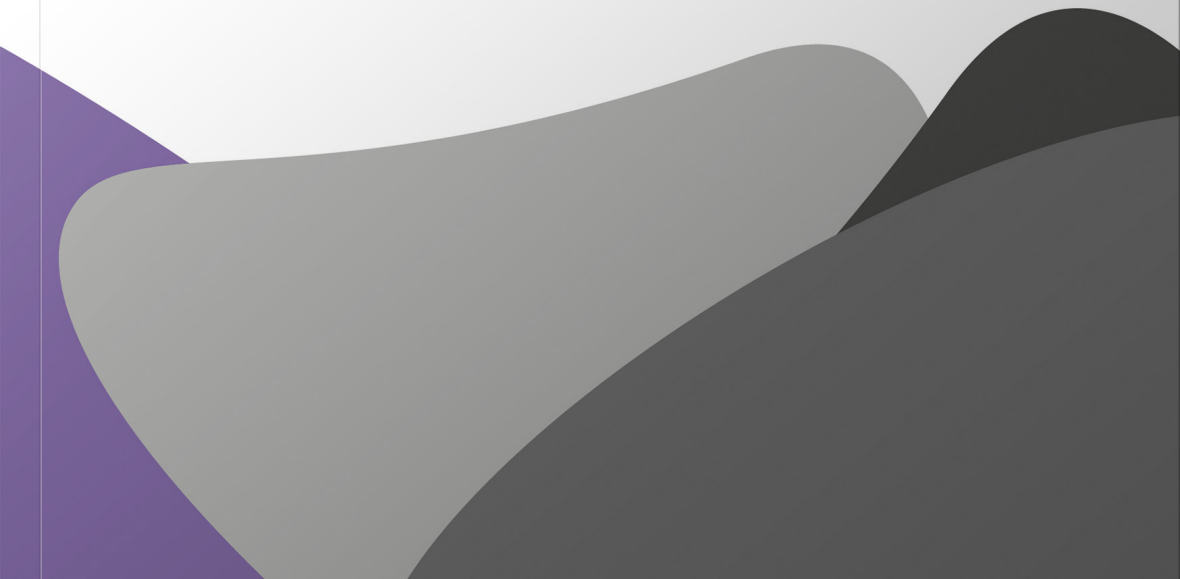


From Global to Local?

**Digitale Methoden in den Geisteswissenschaften im
deutschsprachigen Raum: Ein Triptychon**

*Ulrike Wuttke, Christopher Nunn, Christian Schröter,
Melanie Seltmann & Christian Wachter*



**Digitale Glokalisierung
Sogenannte ‚Kleine Fächer‘, Digital Humanities und die
Herausforderungen der Mehrsprachigkeit**

Elisabeth Burr
Frederik Elwert
Till Grallert
Natalie Kraneiß
Milanka Matić-Chalkitis
Jana-Katharina Mende
Jonas Müller-Laackman
Aleksej Tikhonov
Cristina Vertan
Cosima Wagner

**From Global to Local? Digitale Methoden in den
Geisteswissenschaften im deutschsprachigen Raum: Ein Triptychon**
Herausgegeben von Ulrike Wuttke, Christopher Nunn,
Christian Schröter, Melanie Seltsmann, Christian Wachter

Publiziert in Melusina Press (Universität Luxemburg), 2026
11, Porte des Sciences
L-4366 Esch-sur-Alzette
<https://www.melusinapress.lu>

Verlagsleitung:	Niels-Oliver Walkowski, Johannes Pause
Lektorat:	Carolyn Knaup, Niels-Oliver Walkowski, die Herausgeber:innen
Gestaltung:	Valentin Henning, Erik Seitz
Umschlagsbild:	Sara-Lee Nußbaum

Die digitale Version dieser Publikation steht unter <https://www.melusinapress.lu>
frei zur Verfügung.

Bibliografische Information der Nationalbibliothek Luxemburg:
Die Nationalbibliothek Luxemburg verzeichnet diese Publikation in der
Luxemburgischen Nationalbibliografie; detaillierte bibliografische Daten sind im
Internet über bnl.public.lu abrufbar.

Druck: Libri Plureos GmbH, Friedensallee 273, 22763 Hamburg.

DOI (Publication):	10.26298/1981-5722-dgsk
DOI (Container):	10.26298/1981-5722
ISBN (Web):	978-2-919815-72-2
ISBN (PDF):	978-2-919815-73-9
ISBN (Print):	978-2-919815-74-6

Das vorliegende Werk steht unter einer CC BY-SA 4.0 Lizenz. Informationen zu
dieser Lizenz finden Sie unter
<https://creativecommons.org/licenses/by-sa/4.0/deed.de>. Die in diesem Werk
enthaltenen Bilder und Ressourcen unterliegen der selben Lizenz, sofern sie
keiner anderen Quelle entnommen wurden oder mit einer anderen Lizenz
versehen sind.





2026

Digitale Glokalisierung: Sogenannte Kleine Fächer, Digital Humanities und die Herausforderungen der Mehrsprachigkeit

Elisabeth Burr, Frederik Elwert, Till Grallert, Natalie Kraneiß,
Milanka Matić-Chalkitis, Jana-Katharina Mende,
Jonas Müller-Laackman, Aleksej Tikhonov, Cristina Vertan &
Cosima Wagner

Einleitung

In den Digital Humanities (DH) ist die Spannung zwischen den globalen Entwicklungen und den Möglichkeiten der lokalen Teilhabe wohl nirgendwo größer als innerhalb der sogenannten Kleinen Fächer.¹ Laut BMBF bauen Kleine Fächer „Brücken zwischen Vergangenheit und Moderne“, sie halten „frühere Sprachen und untergegangene Kulturen präsent“ und „leisten [...] einen wesentlichen Beitrag zu unserem kulturellen Gedächtnis und Selbstverständnis – und das weit über den deutschen Tellerrand hinaus“.² Zudem wird angesichts der gesellschaftlichen Heraus-

¹ Die Arbeitsstelle Kleine Fächer versteht unter der Bezeichnung eine Disziplin, die über ein Selbstverständnis als eigenständiges Fach verfügt, eigene Studiengänge und Studienschwerpunkte sowie bis zu maximal drei eigene unbefristete Professuren pro Standort unterhält. Darüber hinaus muss das Fach über eigene Fachzeitschriften und Fachgesellschaften verfügen („Was ist ein kleines Fach?“, Kartierung, Portal Kleine Fächer, aufgerufen am 6. Dezember 2024, <https://www.kleinefaecher.de/kartierung/was-ist-ein-kleines-fach>). In diesem Text beziehen wir uns auf diejenigen Kleinen Fächer, die sich in Wort und Schrift sowie geistes- bzw. sozialwissenschaftlichen Methoden mit einer Weltregion befassen.

² „Kleine Fächer – Große Potenziale – BMBF“, Bundesministerium für Bildung und Forschung, BMBF, aufgerufen am 11. Juli 2024, https://web.archive.org/web/20230816105314/https://www.bmbf.de/bmbf/de/forschung/geistes-und-sozialwissenschaften/kleine-faecher/kleine-faecher_node.html.

forderungen unserer Zeit darauf verwiesen, dass diese „nicht an Ländergrenzen enden“ und daher „der Internationalisierung der Geistes- und Sozialwissenschaften zunehmende Bedeutung“ zukommt.³ Kleine Fächer mit regionalen Schwerpunkten (Regionalstudien, Area Studies) – von der Arabistik bis zur Ukrainistik – sind in besonderem Maße darauf angewiesen, Ressourcen der DH für und in verschiedenen Sprachen zu verwenden, sei es als multilinguale Korpora, Metadaten in nichtlateinischen Schriften oder als Regelwerke für die Codierung wie z. B. TEI. Innerhalb der DH-Community sind kleine Fächer wie auch insgesamt in den Wissenschaftsstrukturen unterrepräsentiert. *Digitale Globalisierung* bezieht sich daher auf die Situation derjenigen Fächer, die in einem transnationalen und globalen Zusammenhang individuelle Lösungen für konkrete Studien- und Forschungsstandorte finden müssen wie auch auf die Situation von mehrsprachigen DH generell. Als Lokales im Globalen schließt die Einbeziehung von Mehrsprachigkeit in den DH an die im Sammelband verhandelten Grundfragen an.

Aktivitäten zur Stärkung von Mehrsprachigkeit in den DH rücken in letzter Zeit immer stärker in den Vordergrund.⁴ Auch die Multilingual DH-AG im DHd-Verband will hierzu ihren Beitrag leisten.⁵ Jedoch kann die Feststellung einer anglophonen Dominanz in den DH weiterhin als korrekt gelten.⁶ In Bezug auf Europa allerdings hat sie sich von einer vorwiegend englischsprachigen Wissenschaft hin zu parallel nebeneinander existierenden nationalen bzw. nationalsprachlichen DH-Communities mit ver-

3 „Internationale Geistes- und Sozialwissenschaften,“ Forschung, Bundesministerium für Forschung, Technologie und Raumfahrt, BMFTR, aufgerufen am 6. Dezember 2024, <https://www.geistes-und-sozialwissenschaften-bmbf.de/de/Internationale-Geistes-und-Sozialwissenschaften-1696.html>.

4 Lorella Viola und Paul Spence, Hg., *Multilingual Digital Humanities* (London: Routledge, 2023), <https://doi.org/10.4324/9781003393696>; Pedro Nilsson-Fernández und Quinn Dombrowski, „Multilingual Digital Humanities,“ in *The Bloomsbury Handbook to the Digital Humanities*, hg. von James O’Sullivan (London: Bloomsbury, 2022), 83–92, <https://doi.org/10.5040/9781350232143.ch-8>; Alíz Horváth, „Enhancing Language Inclusivity in Digital Humanities: Towards Sensitivity and Multilingualism,“ *Modern Languages Open* 1: 1–21, <https://doi.org/10.3828/mlo.v0i0.382>.

5 „AG Multilingual DH,“ DHd: Digital Humanities im deutschsprachigen Raum, aufgerufen am 6. Dezember 2024, <https://digitalhumanities.de/ag-multilingual-dh/>. Dieser Beitrag beruht auf einem von der AG organisierten Panel zu „Zwischen Mehrsprachigkeit und Ressourcenlücke: Quo Vadis ‚Kleine Fächer‘ in den deutschsprachigen Digital Humanities?“ auf der DHd 2024.

bindenden Strukturen wie etwa DARIAH verschoben. Die in der DACH-Region traditionell Germanistik-nahen DH bilden dabei hauptsächlich Deutsch und Englisch als Fach- und Publikationssprache ab.⁷ Ein Hauptproblem der mehrsprachigen DH ist die implizite Vorannahme von Einsprachigkeit, die in der Linguistik als „monolingualer Habitus“⁸ bezeichnet wird, also die Annahme, dass Einsprachigkeit die Norm ist. Das zeigt sich z. B. in einsprachigen Metadaten, Tools, die andere Sprach- und Schriftformen nicht zulassen oder in der Bereitstellung einsprachiger Korpora durch Archive und Bibliotheken. Infolgedessen gibt es einen *resourcedness gap*⁹ bei der digitalen Präsenz dieser Disziplinen.

Umgekehrt resultiert dieser *gap* aber auch aus einer als einsprachig empfundenen Ausgangssituation. Der globalen Dominanz des Englischen entspricht in der DACH-Region die Dominanz von Deutsch, oft fungieren Deutsch und Englisch als akademische *linguae francae*. Gerade für philologisch aber auch für sozialwissenschaftlich arbeitende Kleine Fächer innerhalb der deutschen Wissenschaft (z. B. die Slavistik, die Sinologie, die Japanologie oder die Arabistik) entspricht das jedoch nicht der Fachrealität: hier werden sprachlich diverse Ressourcen in unterschiedlichen Schriften bearbeitet.

6 Isabel Galina Russell, „Geographical and Linguistic Diversity in the Digital Humanities,“ *Literary and Linguistic Computing* 29, Nr. 3 (1. September 2014): 307–16, <https://doi.org/10.1093/lilc/fqu005>.

7 Till Grallert et al., „Umgang mit Multilingualität im DACH und DHd Verband“ (Zenodo, 22. Mai 2023), <https://doi.org/10.5281/zenodo.7957187>; Christof Schöch „Ten Years of the DH(d) and its Disciplines,“ Internationales Wissenschaftsforum: Disciplinary Transformations? Humanities Impact on Reshaping the Digital Humanities, org. Christopher Nunn und Natalie Rauscher, Universität Heidelberg, 12.–13. Juni 2025, <https://dhtrier.quarto.pub/dhd10>; Fabian Cremer et al., „DHd Chronicles – Anreicherung und Analyse der Beiträge zu den Jahrestagungen der Digital Humanities im deutschsprachigen Raum 2014–2023“ (Zenodo, 21. Februar 2024), <https://doi.org/10.5281/zenodo.10698355>.

8 Ingrid Gogolin, *Der monolinguale Habitus der multilingualen Schule* (Münster, New York: Waxmann, 2008), <https://elibrary.utb.de/doi/book/10.31244/9783830970989>.

9 Gabriel Nicholas und Aliya Bhatia, „Lost in Translation: Large Language Models in Non-English Content Analysis,“ Center for Democracy & Technology, 23. Mai 2023, <https://cdt.org/insights/lost-in-translation-large-language-models-in-non-english-content-analysis/>.

Mehrsprachigkeit ist Teil von Diversität in den DH. Die sprachliche Vielfalt der Welt wird dabei im Digitalen nur bedingt abgedeckt: „Despite growing sensitivity to the issue on digital platforms, linguistic diversity is not well represented in digital environments overall and it is likely that only a few hundred languages are actively in use on web pages to a significant degree.“¹⁰ Was allgemein auf das Internet zutrifft, betrifft aber ebenso Infrastrukturen und Forschungsgegenstände der DH. Gerade hier sind die Kleinen Fächer schon allein für die eigene Zukunftsperspektive gezwungen, dieses Missverhältnis zu korrigieren. In den letzten Jahren wurden dafür Ansätze geschaffen, die entgegen den dominanten einsprachigen Entwicklungen im Wissenschaftsfeld Ressourcen für eine mehrsprachige und diverse Nutzer*innenschaft entwickeln.¹¹ Dabei betrifft die Mehrsprachigkeit und Förderung sprachlicher Diversität zahlreiche verschiedene Aufgabenfelder der DH: Auf der Ebene von Infrastrukturen, Bibliotheken und Archiven betrifft es insbesondere deren Kataloge, Erschließungssysteme, Regelwerke, Discovery-Systeme, Normdaten und Metadaten; hinsichtlich der Verfügbarkeit und Zugänglichkeit von Dokumenten, Texten und Korpora, die unterschiedliche Schriften, insbesondere nicht-lateinische Schriften, verwenden, betrifft es insbesondere Software und Tools.

Um Mehrsprachigkeit und Mehrschriftlichkeit als ‚globale‘ Aufgabe der DH in den Blick zu nehmen, lassen wir als Schreibkollektiv Stimmen aus der multilingualen DH-Praxis zu Wort kommen.

10 Paul Spence und Renata Faria Brandao, „Towards Language Sensitivity and Diversity in the Digital Humanities,“ *Digital Studies / Le Champ Numérique* 11, Nr. 9 (24. Dezember 2021): 1–29, <https://doi.org/10.16995/dscn.8098>.

11 Siehe z. B. „Multilingual Digital Humanities,“ GitHub, aufgerufen am 6. Dezember 2024, <https://github.com/multilingual-dh>.

Debatten zur Mehrsprachigkeit in den DH: Ein historischer Überblick

Elisabeth Burr

Internationale DH-Kongresse und -Verbände

Die *Humanities Computing Community* war immer schon mehrsprachig und beschäftigt sich von Beginn an mit unterschiedlichen Sprachen und Schriftsystemen. In *Computers in the Humanities (CHum)* und *Literary & Linguistic Computing (LLC)*¹² wurden von ihrer Gründung an auch Beiträge in mehreren Sprachen publiziert und in verschiedenen Sprachen veröffentlichte Bücher besprochen. Bei ihren Kolloquien und Konferenzen¹³ tat sich die Community allerdings lange schwer, ihre eigene Mehrsprachigkeit explizit anzuerkennen. Erst ein außerhalb der Verbände vollzogener Perspektivenwechsel und mühsames Lobbying bahnten der Mehrsprachigkeit auf Verbandsebene einen Weg.

Symposien und Konferenzen 1972–2003

In den Ankündigungen und Berichten zu den sieben zwischen 1970 und 1982 – ab 1974 von der Association for Literary and Linguistic Computing (ALLC) – im zweijährigen Rhythmus zu *Uses of the Computer in (Linguistic and) Literary Research* ausgerichteten (internationalen) Symposien findet sich kein Hinweis auf Sprache(n), ebenso wenig in den Calls for Papers (CfP) zu den fünf zwischen 1984 und 1988 jährlich stattfindenden ALLC-Konferenzen und den von ALLC und Association for Computers and the Humanities (ACH) 1990, 1991, 1992, 1993, 1995 und 1999 gemeinsam abgehaltenen Tagungen. Und das, obwohl der CfP der ALLC-Konferenzen 1984 (Löwen), 1987 (Göteborg) und 1988 (Jerusalem) mit Association for Literary and Linguistic Computing und Association de Littérature et de

12 Heute *Digital Scholarship in the Humanities (DSH)*.

13 Daten zu den Konferenzen finden sich in: Scott B. Weingart et al., „The Index of Digital Humanities Conferences Data,“ Carnegie Mellon University, 2022, <https://doi.org/10.1184/R1/12987959.v4>.

Linguistique Computationelles überschrieben ist und zu Löwen 1984, Norwich 1986 und Siegen 1990 nachweislich auch französisch- bzw. französisch- und deutschsprachige Beiträge angenommen wurden. Dagegen wurden mit dem Vermerk „the official language is English“ im CfP zur ALLC / ACH im norwegischen Bergen 1996 nicht englischsprachige Beiträge explizit ausgeschlossen.¹⁴ Bei den von ALLC und ACH gemeinsam in (auch) französischsprachigen Ländern ausgerichteten Tagungen Toronto 1989, Paris 1994 und Kingston 1997 waren (selbstverständlich) Englisch und Französisch Konferenzsprachen. Im Falle Kingstons gab es sogar einen zweisprachigen CfP. In den CfPs zu ALLC / ACH 1998, 2000 und 2002 sowie zu ACH / ALLC 2001 und 2003 wurden dann zwar auch französisch- und deutschsprachige Vorschläge akzeptiert, aber nur zusammen mit einer englischen Übersetzung „to facilitate the reviewing process“.¹⁵ Die Frage, welcher Sprachen die Reviewer*innen mächtig waren, wurde nicht einmal gestellt.

Perspektivenwechsel zur Mehrsprachigkeit

Unbemerkt blieb zunächst der Perspektivenwechsel, den die 1998 von Domenico Fiormonte gegründeten Seminare Computers, Literature and Philology (CLiP) mittels mehrsprachiger Webseiten, Präsentationen und Publikationen sowie der Schaffung des European Network of Excellence Computers & Humanities in Multilingual Europe (CHiME) vollzogen. Mit der Vorstellung von CLiP und CHiME bei der ALLC-Vorstandssitzung in Pisa 2002 galt die Frage *Mehrsprachigkeit und Mehrkulturalität* in den DH dann aber als gesetzt. 2003 in Athens (USA) wurde der als damals riskant betrachtete Beschluss gefasst, den Fokus der ALLC / ACH 2004 in Göteborg auf Mehrsprachigkeit und kulturelles Erbe zu richten und Deutsch, Eng-

¹⁴ ACH/ALLC, „Call for Papers 1996,“ *Humanist Discussion Group* 9, Nr. 155 (14. September 1995).

¹⁵ ACH/ALLC, „Call for Papers 1998,“ *Humanist Discussion Group* 11, Nr. 172 (15. Juli 1997); ACH/ALLC, „Call for Papers 2000,“ *Humanist Discussion Group* 13, Nr. 133 (12. August 1999); ACH/ALLC, „ACH-ALLC 2001 Call for Papers and Information for Speakers,“ aufgerufen am 12. Juni 2025, https://web.archive.org/web/20010629125629/http://www.nyu.edu/its/humanities/ach_allc2001/call.html; ACH/ALLC, „CALL FOR PAPERS and Information for Speakers,“ ALLC/ACH 2002, aufgerufen am 12. Juni 2025, <http://web.archive.org/web/20020615204323/http://www.uni-tuebingen.de/allcach2002?sflag=cfp>; ACH/ALLC, „Call for Papers 2003,“ *Humanist Discussion Group* 16, Nr. 227 (28. September 2002).

lisch, Französisch, Italienisch, Spanisch und die Landessprache Schwedisch als Konferenzsprachen zuzulassen. Diversifizierung der Themenbereiche und der durch Mitglieder der Community in sieben Sprachen übersetzte CfP zeugen von der befreienden Wirkung des Beschlusses. Nach der Konferenz rückten erstmals Fragen eines gegenüber unterschiedlichen akademischen Kulturen und Muttersprachen sensitiven *Reviewings* und bei der ACH / ALLC 2005 in Victoria *minority, indigenous and rare languages* sowie mehrsprachige Präsentationen in den Fokus. In Victoria wurde zudem das Standing Committee on Multi-Lingualism & Multi-Culturalism (MLMC) gegründet und Elisabeth Burr die Leitung übertragen.¹⁶

Hindernisse und Erfolge

Das MLMC musste sich in den ersten zehn Jahren heftigen Diskussionen, Infragestellungen und Verdrehungen seiner Ziele nicht nur durch die anglophonen Teile der DH-Community stellen, sondern auch durch deren das Dogma des *English only* in der Wissenschaft verteidigende Mitglieder. Mithilfe von viel Überzeugungsarbeit und konkreten Aktionen wurde dann doch die Einsicht erwirkt, dass Konferenzen und die entsprechenden CfP prinzipiell so viele Sprachen wie möglich abdecken sollten. Wenngleich das mehr Arbeit bedeutet, so hat sich schließlich ein mehrsprachiger Pragmatismus auch durch die Unterstützung der Communities selbst durchgesetzt. So gibt es seit 2015 in der Regel fünf bis sechs Konferenz- und CfP-Sprachen, je nachdem ob zusätzlich die Sprache des gastgebenden Landes oder einer wichtigen Community berücksichtigt wird (z. B. Polnisch 2016, Portugiesisch 2018 und Algonquin 2020). Aus dem MLMC ging zudem auch GO::DH¹⁷ hervor, das 2013 zur ersten Special Interest Group (SIG) der Alliance of Digital Humanities Organizations (ADHO) wurde. Darüber hinaus wurde langfristig das Bewusstsein für die Bedeutung von Sprachen bei der Beschäftigung mit wissenschaftlichen Themen und der Verbreitung von Ergebnissen gefördert (siehe zum Beispiel die Zeitschrift *Digital Humanities Quarterly* (DHQ)¹⁸). Vor allem aber

¹⁶ Siehe auch Elisabeth Burr, „Discussion Paper: Internationalization, Multi-Lingual & Multi-Cultural Agenda“, *ADHO – Standing Committee on Multi-Lingualism & Multi-Culturalism*, Version 1 (4. Juli 2006), https://adho.org/wp-content/uploads/2022/07/MLMC_DiscussionPaper_2006.pdf.

¹⁷ „About,“ Global Outlook::Digital Humanities, GO::DH, aufgerufen am 12. Juni 2025, <http://www.globaloutlookdh.org/>.

ist dies sichtbar in der Öffentlichkeitsarbeit (siehe die mehrsprachigen CFP und inzwischen auch Webseiten von ADHO), in Fragen der Inklusion oder Exklusion (siehe die mehrsprachigen Präsentationen zum Beispiel bei der DH 2018) oder bei der Auseinandersetzung mit Software und Infrastruktur (siehe die SIG Multilingual DH und ihre Befassung mit u. a. *Tools and Tutorials*). Nicht zuletzt hat die Öffnung gegenüber Mehrsprachigkeit und Mehrkulturalität auch zur Diversifizierung der Themen und der Anerkennung von sprachlicher und kultureller Vielfalt geführt.

Lessons learned

Dieser kurze Abriss der Geschichte der für die DH konstitutiven Konferenzen und Initiativen belegt, dass die Community schon immer mehrsprachig war. Allerdings ist der Prozess der Anerkennung dieses Umstandes nicht abgeschlossen und das Englische dominiert weiterhin als scheinbar pragmatischster Zugang zu möglichst reibungsloser Kommunikation in relativ ressourcenarmen Feldern. Auffällig ist, dass die programmatischen Anstrengungen sich nicht in der gelebten Praxis widerspiegeln. Nach dem vorsichtigen Perspektivwechsel aus dem europäischen Kontext in den späten 1990er Jahren heraus wurden Fragen zu Mehrsprachigkeit und Mehrkulturalität zu den Leitthemen von Konferenzen. Praktische Experimente zur Umsetzung gelebter Vielsprachigkeit waren aber auf wenige Konferenzen beschränkt. Obwohl aktuelle ADHO-Konferenzen Beiträge in fünf bis sechs Sprachen zulassen, wird dieses Angebot praktisch kaum genutzt. Mehrsprachigkeit auf Verbandsebene bleibt also auch weiterhin ein wichtiges Aktionsfeld. Hierbei gilt es auch ein Bewusstsein dafür zu schaffen, dass das Schreiben in verschiedenen Sprachen die Korpusbildung und damit die Verfügbarkeit von Trainingsdaten für Sprachen fördert, die im Vergleich zum Englischen arm an Ressourcen sind.

¹⁸ „Home,“ Digital Humanities Quarterly, DHQ, aufgerufen am 12. Juni 2025, <https://www.digitalhumanities.org/dhq/>.

Multidimensionale Multilingualität: Eine Problemanzeige

Frederik Elwert

Von außen betrachtet scheinen sich die Probleme von Mehrsprachigkeit in den DH zu einem überschaubaren Bündel von Herausforderungen und mittlerweile grundsätzlich bekannten Lösungen zu kondensieren: Unicode, linksläufige Schriften wie Arabisch und Hebräisch, ungenügende Abdeckung von sprachverarbeitenden Modellen.¹⁹ In der Praxis ist Multilingualität aber ein Kristallisationspunkt für eine Kaskade von miteinander lose verbundenen, aber teils sehr unterschiedlich gelagerten Problemlagen im Rahmen der DH. Die folgenden Betrachtungen speisen sich aus meinen Erfahrungen als Digital Humanist in der Religionsforschung. Die angesprochenen Schwierigkeiten dürften aber beispielhaft für die Herausforderungen vieler Forschender stehen.

Am Centrum für Religionswissenschaftliche Studien (CERES) der Ruhr-Universität Bochum arbeiten Forschende verschiedener sogenannter Kleiner Fächer gemeinsam an religionsbezogenen Forschungsfragen. Jüngstes Unterfangen ist der Sonderforschungsbereich (SFB) 1475 – Metaphern der Religion.²⁰ Bei der Bandbreite disziplinärer Hintergründe und unterschiedlicher Quellsprachen ist Multilingualität kein isoliertes Problemfeld, sondern durchzieht jede Arbeit an gemeinsamen Methoden und Infrastrukturen. Multilingualität begegnet uns in verschiedenen Dimensionen: als Multilingualität in den *Gegenständen*, als Multilingualität in den *Repräsentationen* und als Multilingualität in den *Communities of Practice*.

¹⁹ Siehe etwa Alíz Horváth et al., „Towards Multilingually Enabled Digital Knowledge Infrastructures: A Qualitative Survey Analysis“, in *Multilingual Digital Humanities*, hg. von Lorella Viola und Paul Spence (London: Routledge, 2023), 197–212, <https://doi.org/10.4324/9781003393696-17>.

²⁰ Volkhard Krech et al., *Metaphors of Religion: A Conceptual Framework* (Bochum: SFB 1475, 2024), <https://doi.org/10.46586/mp.282>.

Multilingualität in den Gegenständen

Dieser Aspekt ist vermutlich derjenige, an den die Meisten als Erstes denken. In den Geisteswissenschaften, vor allem aber in den sogenannten Area Studies, befassen wir uns mit verschiedenen Sprachen und Schriftsystemen. Manche davon werden aktiv gesprochen (wie Arabisch), manche behandeln wir in früheren Sprachstufen, die sehr eigenen Regeln folgen (wie klassisches Chinesisch oder Mittelkoreanisch), und manche sind uns nur historisch überliefert (wie Sumerisch). Die Herausforderung besteht aber nicht nur darin, mit einer oder mehreren dieser Sprachen in digitalen Umgebungen zu operieren, sondern dass auch die Gegenstände selbst in vielen Fällen Gebrauch von mehreren Sprachen oder Schriftsystemen machen. So finden wir Texte, die Akkadisch und Sumerisch mischen, oder Schreibsysteme, die chinesische Schriftzeichen mit phonetischen Annotationen in anderen japanischen oder koreanischen Schriftsystemen kombinieren (die sogenannten Ruby-Annotationen).

Multilingualität in den Repräsentationen

Die Mehrsprachigkeit der Gegenstände potenziert sich noch einmal in den verschiedenen möglichen Repräsentationsformen. Selbst wenn man die Übersetzung als eine potenziell endlose Möglichkeit von Repräsentationen außer Acht lässt, haben wir es in der Praxis bei vielen Sprachen mit unterschiedlichen Formen der Transkription zu tun. Gerade die Repräsentation außereuropäischer Schriftsysteme mit lateinischen Buchstaben (sowie teils vielfältigen diakritischen Zeichen) ist in westlichen Forschungstraditionen tief verankert. Dies muss für die aktuelle Forschungspraxis ausgehandelt werden und ist über die Disziplinen auch nicht einheitlich. Während etwa die Sinologie häufig mit Chinesisch im originalen Skript arbeitet, ist in der Indologie die Transliteration die übliche Praxis. Und auch für die Arbeit mit historischen Datenbeständen, etwa Print-Editionen, ist Repräsentation eine Herausforderung. Während es zwar mittlerweile gute OCR-Modelle für Devanagari gibt,²¹ ist gerade die wissenschaftliche Umschrift mit ihren Diakritika für die Digitalisierung eine Herausforderung.

21 Etwa Nicole Merkel-Hilf, „Devanagari Gemischt M1A,“ Transkribus, aufgerufen am 13. Dezember 2024, <https://readcoop.eu/de/modelle/devanagari-mixed-m1a/>.

In den DH spielen zudem abstraktere Repräsentationen mittels Thesauri und Ontologien eine wichtige Rolle. Auch diese zumindest potenziell sprachunabhängigen Ordnungssysteme sind in der Praxis sprachlich²² oder in Bezug auf ihre inhaltliche Abdeckung²³ nicht unabhängig von den Sprach- und Denkstrukturen, die sie hervorgebracht haben.

Multilingualität in den Communities of Practice

Nicht nur die Gegenstände und ihre möglichen Repräsentationen sind von Mehrsprachigkeit geprägt, sondern auch die Communities, die sich mit diesen Gegenständen befassen. In welcher Sprache wird die Forschung über bestimmte Formen kultureller Überlieferung durchgeführt, publiziert und rezipiert? Wessen Publikationen sind sichtbar? Welche Öffentlichkeiten werden mit bestimmten Darstellungsformen adressiert? Die Arbeit über verschiedensprachige Gegenstände ist dabei noch nicht gleichbedeutend mit der Arbeit mit mehrsprachigen Forschungscommunities und Öffentlichkeiten.²⁴

Alte und neue Infrastrukturen

Digitale Infrastrukturen interferieren auf verschiedenen Ebenen mit diesen Gegenständen. Das Erbe von ASCII und das Problem konkurrierender und unvollständiger Zeichenkodierungen lebt bis heute fort, sei es in historischen Datensätzen oder in bestehenden Lücken in der Unicode-Kodierung.²⁵ Aber auch moderne Infrastrukturen sind häufig blind gegenüber den Anforderungen multilingualer Daten. So basieren viele Textanalyse- oder Annotationstools auf der Annahme, dass sich Texte als lineare Folge von Zeichenketten interpretieren und darstellen lassen können. Dies

²² Man denke etwa an die Gemeinsame Normdatei (GND).

²³ Etwa der deutliche *western bias* des etablierten Getty Arts and Architecture Thesaurus (AAT) oder von IconClass. Siehe hierzu Alina Kühnl, „Iconclass: Ein Klassifizierungssystem für Kunst – und Mensch?“, *the ARTicle* (Blog), *Hypotheses*, 6. Oktober 2020, <https://doi.org/10.58079/uqtu>.

²⁴ Vgl. Spence und Faria Brandao, „Towards Language Sensitivity and Diversity in the Digital Humanities.“

²⁵ Antonio Perri, „Al di là della tecnologia, la scrittura. Il caso Unicode,“ *Annali - Università degli Studi Suor Orsola Benincasa*, Nr. 2 (2009), <https://universitypress.unisob.na.it/ojs/index.php/annali/article/view/1059>.

gilt aber nicht für komplexe Schriftsysteme wie Ruby-Annotationen. Auch scheinbar einfache Operationen wie die Tokenisierung sind z. B. für indische Sprachen eine Herausforderung, bei denen nicht nur die Wortgrenzen typografisch nicht angezeigt werden, sondern auch dasselbe Zeichen aufgrund phonologischer Verschmelzungen (*Sandhi*) zu zwei verschiedenen Wörtern gehören kann.

Was folgt daraus für die Arbeit im SFB 1475? Zunächst einmal sind viele etablierte Tools ‚von der Stange‘, etwa zur Textannotation, nicht ohne weiteres für die Bandbreite unserer Daten geeignet. Der Vielfalt an möglichen Konfigurationen kann dabei nicht allein durch zusätzliche Werkzeuge begegnet werden. In vielen Fällen ist auch eine Anpassbarkeit durch fallspezifische Regeln der Verarbeitung und Darstellung erforderlich. Es braucht also nicht nur ‚bessere‘ Werkzeuge, weil die Kriterien für ‚gut‘ sehr fallabhängig sind, es braucht in erster Linie flexiblere Tools. Als SFB sind wir in der komfortablen Situation, dass wir die Ressourcen haben, zur Not auch eigene Entwicklungen stemmen zu können. Aber auch mit einer hinreichend flexiblen Gesamtarchitektur erfordert es sehr viel Aufmerksamkeit für die jeweiligen Besonderheiten des Materials, ohne dabei die Anforderungen an eine gewisse Einheitlichkeit in der Verarbeitung aufzugeben. Dies ist nur in einem engen Austausch zwischen DH und den Fachwissenschaftler*innen möglich und erfordert eine geduldige Rekonstruktion der jeweiligen Spezifika der Sprache und des Materials, der etablierten fachspezifischen Forschungspraxis und der sich daraus ergebenden Herausforderungen für eine digitale Methodik.

Lessons learned

Es zeigt sich, dass Multilingualität ein multidimensionales Problem darstellt, das letztlich auf allen Ebenen der wissenschaftlichen Arbeit und vor allem Infrastruktur wirkt. Sprache wird zu einer ein- oder wenigdimensionalen Sache reduziert, die es analog zu einfacher Internationalisierung im Zweifel über Unicode oder Latinisierung, in Fragen der Community hilfsweise mit Englisch abzubilden gilt. Dabei wird übersehen, wie leerstellenbehaftet diese vermeintlichen Lösungen oft sind und inwieweit bestehende Ansätze und Toolkits weit davon entfernt sind, der Komplexität der multilingualen Realität gerecht zu werden.

Ohne Lehre keine DH-Forschung: Ein Problembericht aus der Islamwissenschaft und Arabistik

Natalie Kraneiß

Im Studium und bei der Promotion in den Fächern Arabistik und Islamwissenschaft in Deutschland zeigt sich in Hinblick auf DH und Mehrsprachigkeit ein doppeltes Problem: Zum einen besteht nur begrenzt Interesse bei Studierenden an den Methoden der DH. Zum anderen fehlt es an Angeboten, die diese Methoden unter Berücksichtigung der sprachlichen und historischen Besonderheiten des Fachs vermitteln. Das Desinteresse der Studierenden führe ich auf die geringe Bereitschaft der Forschenden zurück, das Potenzial der DH auszuloten und sich an der Diskussion über die Herausforderungen zu beteiligen. DH spielen in der Ausbildung bisher kaum eine Rolle, mit ersten Ausnahmen etwa in Hamburg und Berlin.²⁶

Disziplinübergreifende Angebote wie etwa das vom Service Center for DH angebotene Zertifikat DH an der Universität Münster sind offen für alle Studierenden und bieten eine Einführung in die DH. Allerdings wird das Angebot größtenteils von bestimmten Fachbereichen getragen und ist nicht auf außereuropäische Sprachen ausgelegt. Viele der vorgestellten Tools und Methoden funktionieren daher nicht oder nur eingeschränkt

²⁶ In Hamburg gibt Maxim Romanov seit 2022 im Rahmen des Studiums der Afrika-Asien-Studien Einführungskurse in R, die auch die spezifischen Bedarfe der am Studiengang beteiligten Fächer berücksichtigen (z. B. im SoSe 2022 Maxim G. Romanov, „DH in AAS - TA with R (2022S)“, Syllabus, 17. Juni 2022, https://eis1600.github.io/course_TA_R_2022S/ und im SoSe 2023 Maxim G. Romanov, „HDIS: Analysis of Prosopographical Data—with R“, Syllabus, 30. Oktober 2023, <https://eis1600.aai.uni-hamburg.de/hdis-pua/index.html>. Xenia Kudela gibt im WiSe 2024 an der Freien Universität Berlin einen Workshop „Introduction to Digital Research Methods“ für Studierende, die mit nicht-lateinischen Schriften arbeiten („Introduction to Digital Research Methods – M. Xenia Kudela“, Vorlesungsverzeichnis, Freie Universität Berlin, aufgerufen am 31. Mai 2024, https://www.geschkult.fu-berlin.de/e/semiarab/arabistik/seminar/termine/2024_11_KUDELA_DH-Workshop.html).

mit arabischem, persischem oder türkischem Material.²⁷ Und wie kann man Korpusanalyse üben, wenn es keine kuratierten Korpora gibt?

Eine besondere Herausforderung stellen die unterschiedlichen Schriftsysteme dar, mit denen wir arbeiten. Wer heute in der Islamwissenschaft und Arabistik digitale Methoden einsetzen möchte, muss nicht nur selbstständig das Potenzial erkennen, die passende Methode und ein geeignetes Werkzeug finden sowie die oft mühsame Einarbeitung durchlaufen, sondern stößt fast zwangsläufig auf technische Probleme, die sich aus der Verwendung anderer Schriftsysteme und fehlender Software ergeben. Zur Lösung dieser Probleme fehlt es Anfänger*innen an Know-how: Es gibt nur wenige Tutorials, diese sind oft von und für Informatiker*innen geschrieben, und es fehlen Anlaufstellen, an die man sich mit Problemen wenden kann. Während *trial and error* in den DH zum Alltag gehört, ist in Kleinen Fächern, wie der Islamwissenschaft, die mit nichtlateinischen Schriften arbeiten, eine äußerst hohe Frustrationstoleranz erforderlich. Wer sich eigenständig in digitale Methoden einarbeiten möchte, bleibt auf persönliche Kontakte und das Wohlwollen von Menschen angewiesen, die in ihrer Freizeit bei Problemen helfen.

Diese Herausforderungen sind sicher bekannt, aber der Unterschied zwischen Kleinen Fächern wie der Islamwissenschaft und Großen Fächern wie der Geschichtswissenschaft liegt darin, dass in letzteren bereits systematisch Materialien und Werkzeuge gesammelt und bereitgestellt werden, um dem wissenschaftlichen Nachwuchs den Einstieg zu erleichtern – bis hin zur Einrichtung eigener Lehrstühle. Auf solche Strukturen können Angehörige meines Fachs nicht hoffen.

27 Maxim Romanov, Leiter der Emmy Noether-Nachwuchsgruppe The Evolution of Islamic Societies (c. 600-1600), und sein Team wiesen erst kürzlich auf die zahlreichen Schwierigkeiten hin, mit denen Forschende bei der Anwendung von Named Entity Recognition auf klassische arabische Quellen konfrontiert sind und präsentierten erste Ergebnisse der Entwicklung und des Trainings von Deep-Learning-Modellen, die speziell für die Bewältigung verschiedener Aufgaben im Zusammenhang mit klassischen arabischen Texten erstellt wurden (Tariq Yousef et al., „The Evolution of Islamic Societies (c. 600-1600): From Algorithmic Analysis into Social History,“ in *Proceedings of the Ancient Language Processing Workshop: Associated with the 14th International Conference on Recent Advances in Natural Language Processing RANLP 2023*, hg. von Adam Anderson und Association for Computational Linguistics (Shoumen: Incoma, 2023), 160–69.

Lessons learned

Es steht zu befürchten, dass Kleine Fächer wie die Islamwissenschaft und Arabistik mit den Entwicklungen in den DH nicht Schritt halten werden können. Dies ist nicht nur ein Nachteil für Studierende und Wissenschaftler*innen der Fächer, die nicht die notwendigen Kompetenzen erwerben und weitergeben, um sich mit anderen Digital Humanists auszutauschen, sondern auch für die DH insgesamt. Denn diese verlieren die wertvollen Perspektiven von Wissenschaftler*innen, die sich mit unterrepräsentierten Menschen, Sprachen und Regionen befassen – sowohl in der Tech-Welt als auch in den Geisteswissenschaften.

Toolentwicklung vs. Anpassung allgemeiner Lösungen – Lehren aus zwei Kleinen Fächern

Cristina Vertan

In diesem Beitrag werden exemplarisch zwei Projekte aus jeweils einem Kleinen Fach vorgestellt. Die Projekte unterscheiden sich grundlegend nicht nur bezüglich der Sprache(n) der bearbeiteten Dokumente, sondern auch bezüglich des Forschungsziels sowie der bereits existierenden Sprachressourcen. Die betroffenen Kleinen Fächer sind: (1) Äthiopistik und (2) Altamerikanistik. Die Projekte und deren Lösungen werden anhand der folgenden Kriterien beschrieben: (A) Sprachtypologie, (B) Annotationszweck, (C) bereits existierende Ressourcen und (D) ausgewählte digitale Verfahren.

Sprachtypologie

Alle Projekte beschäftigen sich mit historischen Sprachen.

(1) Beim Altäthiopischen (Ge'ez) handelt es sich um eine südsemitische Sprache. Bis ins 19. Jahrhundert war sie die bedeutendste Schriftsprache des christlichen Äthiopiens. Die reiche christlich-äthiopische Literatur ist zunächst von Übersetzungen – aus dem Griechischen und Arabischen – geprägt, bevor sich eine vielfältige indigene Literatur entwickelt hat. Ins-

besondere Texte, die nur im Altäthiopischen vollständig überliefert sind (z. B. das *Henochbuch*), erlangen dabei besondere Bedeutung.²⁸ Das Altäthiopische hat ein eigenes Silbenalphabet entwickelt, das bis heute in modernen Sprachen Äthiopiens und Eritreas Verwendung findet. Innerhalb der semitischen Sprachen fällt es durch die verwendete Schreibrichtung (von links nach rechts) auf. Dieses Silbenalphabet führt dazu, dass Morphemgrenzen in der Schrift nicht darstellbar sind, sodass beispielsweise ein einzelner Vokal als Bestandteil einer Silbe eine eigenständige Wortart darstellen kann und tokenisiert werden muss.²⁹

(2) Das Schriftsystem des Klassischen Maya zeichnet sich durch seinen ikonischen Charakter aus, weshalb es auch als Hieroglyphenschrift bezeichnet wird. Typologisch handelt es sich um ein logosyllabisches System, das durch zwei grundlegende Zeichenklassen, die Logo- und Syllabogramme, gekennzeichnet ist. Logogramme stehen für konkrete sprachliche Begriffe und verweisen mit wenigen Ausnahmen immer nur auf ein Denotat. Vokal- und Silbenzeichen repräsentieren einerseits Silbenkomponenten und dienen andererseits zur Schreibung lexikalischer und grammatischer Morpheme.³⁰

Annotationen und Projektziele

Im ersten Projekt TraCES – From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages³¹ lag

28 Cristina Vertan, Andreas Ellwardt und Susanne Hummerl, „Ein Mehrebenen-Tagging-Modell für die Annotation altäthiopischer Texte,“ in *DHd 2016 Modellierung – Vernetzung – Visualisierung. Die Digital Humanities als Fächerübergreifendes Forschungsparadigma* (Zenodo, 2. Mai 2017), <https://doi.org/10.5281/zenodo.4645397>.

29 Stephan Druskat und Cristina Vertan, „Nachnutzbarmachung von Forschungsdaten und Tools am Beispiel altäthiopischer Korpora,“ in *DHd 2018 Kritik der digitalen Vernunft* (Zenodo, 26. Februar 2018), <https://doi.org/10.5281/zenodo.4622354>.

30 Franziska Diehr et al., „Modellierung von Entzifferungshypothesen in einem digitalen Zeichenkatalog für die Maya-Schrift,“ in *Die Modellierung des Zweifels – Schlüsselideen und -konzepte zur graphbasierten Modellierung von Unsicherheiten*, hg. von Andreas Kuczera et al. (Wolfenbüttel: ZfdG, 2019), https://doi.org/10.17175/sb004_002.

31 „TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages,“ Press release, Universität Hamburg, auf-

der Fokus auf der diachronen Untersuchung der altäthiopischen Sprache, um die kulturellen und sprachlichen Einflüsse des klassischen Griechisch und Arabisch zu erforschen. Dafür war eine sehr tiefgehende linguistische Annotation der Sprache erforderlich. Der agglutinative Charakter der Sprache (z. B. Nomen und Präposition sind im selben Token) sowie andere Besonderheiten machten eine detaillierte morphologische Annotation im Originalskript unmöglich. Da einige Sprachentwicklungen auch im Schriftsystem beobachtet wurden, mussten sowohl das Originalskript als auch eine Transkription parallel dargestellt werden. Annotiert wurde auf mehreren Ebenen: Textstruktur, Morphologie, Entitäten (Orte, Namen usw.) und Edition.

Im zweiten Projekt, einer Textdatenbank und Wörterbuch des klassischen Maya³² ist die Annotation ein wesentliches Werkzeug zur Entzifferung der Sprache. Es gibt zehn Schritte der Transkription bzw. Transliteration sowie der linguistischen Annotation, die von der ikonischen Darstellung bis hin zur Übersetzung des Textes führen. Alle Schritte müssen die Erfassung mehrerer Arbeitshypothesen ermöglichen.

Sprachliche Ressourcen

Beide Projekte konnten nur auf wenige vorhandene Ressourcen zurückgreifen. Für TRaCES existierte das allgemein akzeptierte Wörterbuch der altäthiopischen Sprache zu Beginn des Projekts nur in gedruckter Form. Die Digitalisierung erfolgte parallel zur Entwicklung des Annotations-tools. Die Annotation der Texte diente auch der Entdeckung neuer Wörter und somit der Erweiterung des Wörterbuchs. Das zweite Projekt hat die Erstellung eines Wörterbuchs der Maya-Sprache als Endprodukt des Annotationstools zum Ziel. In TraCES sind bis auf wenige Ausnahmen für das Transkriptionsalphabet alle Zeichen in Unicode erfasst. Für die Maya-Schrift gibt es derzeit keine allgemein akzeptierte Unicode-Darstellung, weshalb für die ikonische Darstellung noch mit Bildern gearbeitet wird.

gerufen am 6. Dezember 2024, <https://www.aai.uni-hamburg.de/en/ethiostudies/research/traces.html>.

³² „Textdatenbank und Wörterbuch des Klassischen Maya,“ Institut für Archäologie und Kulturanthropologie (IAK), bonndoc, aufgerufen am 6. Dezember 2024, <https://bonndoc.ulb.uni-bonn.de/xmlui/handle/20.500.11811/1117>.

Ausgewählte digitale Verfahren

Zu Beginn jedes Projekts wurden existierende Annotationstools analysiert, jedoch konnte keines die Anforderungen des jeweiligen Fachs ohne erhebliche Informationsverluste erfüllen. In beiden Projekten wurde daher ein eigenes Annotationstool entwickelt. Die Entwicklung basierte auf einem generischen Modell (GraMO),³³ das zuerst die kleinste zu annotierende Einheit definiert und darauf basierend zwei Graphen modelliert: einen für die Annotation und einen für die Sprachdarstellung. Die annotierten Einheiten sind dann Graphknoten, deren Labels (der sprachliche Inhalt) auch während des Annotationsprozesses unter bestimmten Bedingungen geändert werden können. Obwohl das Modellierungskonzept dasselbe ist, unterscheiden sich die drei Implementierungen stark voneinander. Sie basieren auf den jeweiligen spezifischen Projektanforderungen. Die Komplexität der Annotation erforderte eine interne Speicherung in JSON, gefolgt von einem TEI/XML-Export, um die Daten einem breiteren Publikum zugänglich zu machen. Die Tool-Entwicklung stellte die Projekte hinsichtlich Zeit- und Personalressourcen vor erhebliche Herausforderungen. Da die jeweiligen Sprachen nicht vollständig erforscht sind, mussten die Modellierungs- und Implementierungsschritte mehrmals angepasst werden. So wurde die GraMO-Instanz für Maya zunächst als Prototyp implementiert. Während der Nutzung des Prototyps wurden Lücken oder Fehler im Modell entdeckt, die teilweise so gravierend waren, dass eine neue Implementierung erforderlich wurde. Alle Implementierungen wurden in JAVA realisiert, blieben auch mehrere Jahre nach der Entwicklung stabil und nutzbar und sind als Open Source verfügbar. Langfristig müssen sie in die derzeit in Entwicklung befindlichen Infrastrukturen integriert werden. Die Entwicklung einer eigenen Lösung erwies sich trotz des Aufwands in allen Fällen als unverzichtbar.

Lessons learned

Mehrsprachigkeit sowie wenig erforschte Sprachen sind zentraler Untersuchungsgegenstand in Kleinen Fächern. Die hier vorgestellten Beispiele

³³ Cristina Vertan, „Supporting Hermeneutic Interpretation of Historical Documents by Computational Methods,“ in *Proceedings of Computational Humanities Conference 2018* (Lausanne: University of Lausanne, 2018), 77–82, <https://ceur-ws.org/Vol-2314/>.

demonstrieren die Breite und den Umfang von historischen Lücken, die digital geschlossen werden sollen. Dadurch wird auch der Mehrwert solcher Projekte ersichtlich.

Die Lösungen wurden hier aus einem generischen Open Source Model entwickelt und dann für die jeweiligen Projekte intensiv spezifisch aufbereitet. Für eine mehrsprachige DH bedeutet das, auch mögliche Nutzungen für Kleine Fächer und wenig unterstützte Sprachen mitzudenken, um eine Adaption möglich zu machen.

Handschriftenerkennung automatisieren: Das MultiHTR-Projekt für Ukrainisch, Jiddisch und deutsche Kursive

Aleksej Tikhonov & Milanka Matić-Chalkitis

Das MultiHTR-Projekt³⁴ widmet sich in der zweiten Förderphase (2022–2024) der Entwicklung von auf maschinellem Lernen basierenden Modellen zur automatisierten Handschriftenerkennung oder *Handwritten Text Recognition* (HTR) unter Nutzung der Transkribus-Plattform.³⁵ Im Fokus stehen hierbei Sprachen, die entweder nicht auf lateinischen Schriftzeichen basieren oder durch komplexe Sonderformen des lateinischen Alphabets verschriftlicht werden wie Ukrainisch, Jiddisch und deutsche Kursive.

³⁴ „Projektbeschreibung,“ MultiHTR - Multilinguale Handschriftenerkennung, aufgerufen am 6. Dezember 2024, <https://www.multihtr.uni-freiburg.de/>; „MultiHTR,“ Projekt, DHd: Digital Humanities im deutschsprachigen Raum, aufgerufen am 24. August 2025, <https://digitalhumanities.de/projekt/projekt-2-ueberschrift/>.

³⁵ „Künstliche Intelligenz und Küchenrezepte: Ein Programm zur Handschriftenerkennung entziffert alte Briefe, Postkarten oder Tagebücher,“ Archiv der Hochschul- und Wissenschaftskommunikation (bis 08/2024), Universität Freiburg, 25. November 2020, <https://kommunikation.uni-freiburg.de/pm/online-magazin/erleben-und-mitmachen/kuenstliche-intelligenz-und-kuechenrezepte>.

Für das Ukrainische stellte sich von Beginn des Projekts an die Herausforderung, geeignete Handschriften aus dem 19. und 20. Jahrhundert für das Training der Algorithmen zu finden, während die russische Invasion in die Ukraine im Februar 2022 begann und bis heute andauert. Ein zentraler Aspekt in diesem Zusammenhang ist die russische Kriegsführung, die gezielt Kultur- und Bildungseinrichtungen zerstört. So wurde bereits in den ersten Tagen der Invasion das Iwankiwer Museum für Geschichte und Heimatgeschichte in der Kyjiwer Oblast zerstört. Weitere Zerstörungen betrafen unter anderem das Nationale Skovoroda Literatur- und Gedenkmuseum (Oblast Charkiw), Teile des Holocaust-Gedenkzentrums Babyn Jar sowie das Museum des russischen Dichters Sergej Esenin in Charkiw. Am 4. September 2024 bestätigte die UNESCO die Zerstörung oder Beschädigung von 438 kulturellen, religiösen und historischen Objekten in der Ukraine, darunter 32 Museen, 16 Bibliotheken und ein Archiv.³⁶

Trotz der bedrohlichen Lage für Menschenleben und der ständigen Gefahr der Zerstörung von Archivmaterialien konnten innerhalb weniger Monate Kooperationen mit verschiedenen ukrainischen Institutionen, darunter die Vernadskyi-Nationalbibliothek der Ukraine, das General Regionally Annotated Corpus of Ukrainian (GRAC)³⁷ und das Nationalmuseum des Holodomor-Genozids in Kyjiw etabliert werden. Die Zusammenarbeit mit den ukrainischen Partner*innen gestaltete sich reibungslos und hochprofessionell. Absprachen wurden auf Ukrainisch per E-Mail getroffen. Das Projekt setzte hierbei auf flache Hierarchien und effiziente Zusammenarbeit, um Arbeitsprozesse zu optimieren und Verzögerungen in Arbeitsabläufen zu minimieren.

Der Krieg Russlands gegen die Ukraine stellt zwar die größte Herausforderung für die wissenschaftliche Arbeit dar, jedoch bestehen auch weitere relevante Hindernisse wie die Verwendung des kyrillischen Alphabets. Von den etwa 160 öffentlichen Transkribus-Modellen verwenden derzeit 14 Modelle (ca. 8,75 %) kyrillische Alphabete. Dies erschwert zum einen

³⁶ „Damaged Cultural Sites in Ukraine Verified by UNESCO,“ News, UNESCO, 7. Februar 2024, <https://www.unesco.org/en/articles/damaged-cultural-sites-ukraine-verified-unesco>.

³⁷ Maria Shvedova et al., „GRAC: General Regionally Annotated Corpus of Ukrainian,“ GRAC, uacorpus, aufgerufen am 25. Juni 2025, <https://uacorpus.org/>.

die Suche nach geeigneten Kooperationspartner*innen innerhalb und außerhalb der Transkribus-Community, zum anderen haben die Trainingssysteme bislang weniger Erfahrung mit diesen Alphabeten im Vergleich zu den 35 Modellen für Deutsch oder den 22 für Niederländisch. Dies bietet jedoch neue Möglichkeiten für Forschende, die sich mit der Entwicklung und Veröffentlichung kyrillischer Modelle befassen. Das Transliterationsmodell *Ukrainian generic handwriting*³⁸ wurde mit Handschriften in einem Umfang von rund 100.000 Wortformen trainiert und erreichte einen Fehlerquotienten von 4,2 %.

Für Jiddisch, das auf dem hebräischen Alphabet basiert, ist die Expertise in der Transkribus-Community noch geringer. Das MultiHTR-Projekt konnte jedoch von dieser kleinen Community profitieren und Partnerschaften mit Ruthie Abeliovich und Sinai Rusinek vom Dybbuk-Projekt³⁹ (Universität Tel Aviv) sowie mit Astrid Lembke (Germanistische Mediävistik, Universität Mannheim) etablieren. Die Trainingsprozesse stießen auf verschiedene Schwierigkeiten, insbesondere die Übertragung handgeschriebener hebräischer Grapheme in das lateinische Alphabet. Diese Latinisierung des Jiddischen verfolgte zwei primäre Zwecke: (A) die Unterstützung von Jiddisch-Lernenden und (B) die Demokratisierung des Wissens, indem auch Personen ohne Kenntnisse des hebräischen Alphabets Zugang zu jiddischen Texten erhalten.

Eine weitere Herausforderung stellte die Umstellung der Leserichtung von rechts-nach-links (RTL) auf links-nach-rechts (LTR) dar. Auf gemeinsame Initiative des MultiHTR-Projektteams mit Forschenden aus weiteren Projekten führte der Transkribus-Support eine entsprechende Funktion ein, die nun auch für andere RTL-Sprachen wie Osmanisch-Türkisch oder Arabisch genutzt werden kann. Die größte Schwierigkeit bestand jedoch darin, dass es keine allgemeingültige Konvention für die Übertragung vom hebräischen ins lateinische Alphabet gibt. Zwar existieren verschiedene ausgearbeitete Systeme, doch treten hier ebenfalls Komplika-

38 „About this Model,“ *Ukrainian Generic Handwriting*, Transkribus, aufgerufen am 6. Dezember 2024, <https://www.transkribus.org/model/ukrainian-generic-handwriting>.

39 „The DYBBUK Project: Yiddish Popular Theatre 1880–1920,“ *DYBBUK*, aufgerufen am 6. Dezember 2024, <https://www.dybbuk.co>.

tionen auf, da es in einigen Fällen keine Eins-zu-eins-Entsprechung zwischen den hebräischen und lateinischen Graphemen gibt. So kann das Aleph <א> je nach Kontext als [a], [o] oder als Zwischenvokal gelesen werden; gelegentlich bleibt es auch stumm. Das Projektteam entschied sich, für dieses Graphem das Zeichen <ə> zu verwenden, um auf die Flexibilität in der Aussprache hinzuweisen. Darüber hinaus wurde beschlossen, die Modelle⁴⁰ mit einem Fehlerquotienten zwischen 4,73 % und 5 % nicht als Transliterationstools, sondern als Transkriptionshilfen zu definieren, da eine exakte eins-zu-eins-Übertragung des Jiddischen ins lateinische Alphabet nicht gewährleistet werden kann. Sie erleichtern jedoch denjenigen, die das hebräische Alphabet nicht beherrschen, den Zugang.

Zu den trainierten Modellen⁴¹ für verschiedene deutsche Kurzschrift- bzw. Stenografiesysteme⁴², eine der größten Herausforderungen für aktuelle HTR-Technologien, gehören einerseits größere generische Modelle, die sowohl auf natürlichen als auch teilweise auf synthetischen Trainingsdaten basieren, d. h. elektronisch vorliegenden Langschrift-Texten überführt in das entsprechende Kurzschriftsystem, sowie kleinere, spezifische Modelle, die auf bestimmte Handschriften spezialisiert sind. So wurde anhand mehrerer Tagebücher ein auf die Handschrift von Michael Kardinal von Faulhaber aus dem Projekt Kritische Online-Edition der Tagebücher Michael Kardinal von Faulhabers (1911–1952)⁴³ spezialisiertes

40 „About this Model,“ Transcription Aid Model 1 for Handwritten Yiddish (Hebrew to Latin), Transkribus, aufgerufen am 6. Dezember 2024, <https://www.transkribus.org/model/yiddish-to-latin-0.01>.

41 „About this Model,“ Faulhaber, Transkribus, aufgerufen am 13. Dezember 2024, <https://www.transkribus.org/model/faulhaber>; „About this Model,“ Gabelsberger Natural, Transkribus, aufgerufen am 6. Dezember 2024, <https://www.transkribus.org/model/gabelsberger-natural>; „About this Model,“ Stolze-Schrey Combined, Transkribus, aufgerufen am 6. Dezember 2024, https://www.transkribus.org/model/stolze-schrey_combined; „About this Model,“ DEK German Combined, Transkribus, aufgerufen am 6. Dezember 2024, <https://www.transkribus.org/model/dek-german-combined>.

42 Stenografie ist ein Schriftsystem, das mithilfe von verkürzten Wortformen und Sonderzeichen eine besonders schnelle handschriftliche Aufzeichnung von Sprache ermöglicht und wurde vor dem Computerzeitalter oft für das Anfertigen von Gerichtsprotokollen, Unterlagen von Geschäftsführungen oder auch in privaten Unterlagen benutzt (Fritz Kaselitz, *Kritische Würdigung der deutschen Kurzschrift-Systeme von Gabelsberger, Stolze und Arends* (C. W. Born, 1875).

43 „Kritische Online-Edition der Tagebücher Michael Kardinal von Faulhabers (1911–1952),“ Faulhaber Edition, aufgerufen am 6. Dezember 2024, <https://www.>

Modell für die Gabelsberger-Kurzschrift⁴⁴ trainiert, das im August 2024 veröffentlicht wurde. Die Besonderheit dieses Modells ist, dass es sowohl die Lang- als auch die Kurzschrift transkribiert, die verschiedenen positionierten Grapheme innerhalb des gesamten Schreibraumes erkennt und kontextabhängig zwischen nahezu identisch aussehenden Lexemen unterscheidet.

Um Trainingsdaten aus anderen Projekten erweitert, wurden kleinere spezifische Modelle zu größeren generischen Modellen weiterentwickelt, darunter ein generisches Gabelsberger-Modell, welches in der Lage ist, diverse in Gabelsberger-Kurzschrift verfasste Handschriften automatisch transkribieren zu lassen.

Mithilfe von synthetischen Trainingsdaten, elektronisch vorliegenden Langschrift-Texten überführt in das entsprechende Kurzchriftsystem, wurden ebenfalls Modelle für das Stolze-Schrey-Kurzchriftsystem⁴⁵ und die Deutsche Einheitskurzschrift (DEK)⁴⁶ trainiert.

faulhaber-edition.de/; „Kritische Online-Edition der Tagebücher von Michael Kardinal von Faulhaber (1911-1952),“ Institut für Zeitgeschichte, aufgerufen am 6. Dezember 2024, <https://www.ifz-muenchen.de/forschung/ea/forschung/kritische-online-edition-der-tagebuecher-von-michael-kardinal-von-faulhaber-1911-1952>.

44 Die Gabelsberger Kurzschrift ist eine im 19. Jahrhundert von Franz Xaver Gabelsberger entwickelte, handschriftliche Stenografiemethode (Tim Lethen, „Gabelsberger Kurzschrift,“ in *Gespräche, Vorträge, Séancen: Kurt Gödels Wiener Protokolle 1937/38* (Cham: Springer International Publishing, 2021), 11–16, https://doi.org/10.1007/978-3-030-67606-3_2).

45 Das Stolze-Schrey-Kurzchriftsystem ist eine im 19. Jahrhundert entwickelte Stenografiemethode, die Elemente der Systeme von Ferdinand Schrey und Wilhelm Stolze kombiniert, um eine besonders flüssige und leicht erlernbare Kurzschrift zu ermöglichen (Georg Amsel, *Kurzschrift: Lehrbuch der Vereinfachten Deutschen Stenographie (Einigungs-System Stolze-Schrey); nebst Schlüssel, Lesestücken und einem Anhang* (Berlin: De Gruyter, 1907)).

46 Die Deutsche Einheitskurzschrift (DEK) ist ein 1924 eingeführtes, standardisiertes Stenografiesystem, das Elemente verschiedener deutscher Kurzchriftsysteme vereint und in drei Ausbaustufen eine effiziente, einheitliche Aufzeichnung gesprochener Sprache ermöglicht (Rudolf Willimsky, „Schreibtechnik,“ in *Sekretärinnen Handbuch*, hg. von Brigitte Diez, Wiesbaden: Gabler Verlag, 1982: 113–75, https://doi.org/10.1007/978-3-663-05217-3_4).

Lessons learned

Bei unzureichender Menge an zur Verfügung stehenden Trainingsdaten, gerade für wenig erforschte, komplexe oder nicht mehr aktiv verwendete Schriftsysteme, ermöglicht die Verwendung von synthetisch erzeugten Trainingsdaten als Ausgangsbasis erst die Entwicklung von HTR-Modellen. Hervorzuheben hierbei ist, dass das Modelltraining für die deutschen Kurzschriftsysteme als Pionierarbeit in diesem Projekt stattgefunden hat, was bis dato zahlreiche Anfragen sowohl aus der wissenschaftlichen als auch aus der nichtwissenschaftlichen Community an das Projekt bestätigen. Obwohl einzelne Transkriptionsmodelle in diesen komplexen Fällen die typische Fehlerrate von um die 5 % überschreiten, stellen sie eine wichtige Grundlage für die Erschließung von Kurzchriftdokumenten vor allem für das nicht akademische Publikum dar.

Schluss / Ausblick

Digital Humanists arbeiten, „obwohl sie einer globalen Gemeinschaft angehören, in nationalen und lokalen Kontexten“.⁴⁷ Unser Beitrag hat deutlich gemacht, welche Herausforderungen hinsichtlich der Komplexität der multilingualen Arbeitsrealität, des Mangels an Ressourcen – sowohl hinsichtlich Korpora-Bildung als auch personeller Ressourcen beim Wissenschafts- und Wissenschaftsinfrastrukturpersonal – diejenigen Digital Humanists bewältigen müssen, die routinemäßig mit nicht-anglophonen Quellen arbeiten und ihre Forschungsergebnisse auch in anderen Sprachen als Englisch veröffentlichen bzw. kommunizieren.⁴⁸ Hauptanliegen

47 Ulrike Wuttke et al., „Call for Abstracts: From Global to Local? Digitale Methoden in den Geisteswissenschaften im deutschsprachigen Raum: ein Triptychon (Deadline 14.05.2024)“, *DHd-Blog* (Blog), 2. April 2024, <https://dhd-blog.org/?p=20783>.

48 Ein weiterführender Ansatz, mit dem Herausforderungen von multilingualen Digital Humanists zuletzt beschrieben wurden, ist die Erstellung von User Personas. Mittels User Personas können Bedarfe bezüglich technischer Konfigurationen und User Interfaces deutlich gemacht werden. Für Multilingual DH User Personas siehe Alíz Horváth et al., „Centring Multilingual Users: Thinking Through UX Personas in the DH Community“, *DH Unbound 2022, Selected Papers* 13, Nr. 3 (2024): 1–30, <https://doi.org/10.16995/dscn.9608>; Alíz Horváth et al., *Six User Personas for the Multilingual DH Community*, V1.0 (Zenodo, 9. April 2023), <https://doi.org/10.5281/zenodo.7811800>.

ist es zu vermitteln, dass ein umfassenderes Verständnis der wissenschaftlichen Multilingualität im Digitalen dringend erforderlich ist, damit die DH ihr volles Potenzial im Hinblick auf sprachliche Vielfalt und Inklusion ausschöpfen können. Darüber hinaus ist eine solche Awareness essenziell für die Entkolonialisierung des digitalen Raums im Bereich der DH.

Wir betrachten die Etablierung von entsprechend formulierten, um multilinguale Aspekte ergänzten Pflichtenheften, als den nächsten notwendigen Schritt zur Berücksichtigung von Mehrsprachigkeit für Infrastrukturen, Fachgegenstände und Arbeitspraktiken in den DH. Hierzu muss ‚multilinguale Expertise‘ bereits im Erstellungsprozess von Pflichtenheften eingeholt und ausreichend berücksichtigt werden. Am Ende erweitert eine Unterstützung weitgehender Mehrsprachigkeit potenziell auch bisher monolingual geführte Forschungskontexte durch eine größere Awareness gegenüber Schrift und Sprache. Leitfragen bei der Entwicklung solcher Pflichtenhefte könnten hierbei sein:

- Welche Bedarfe gibt es aus den Kleinen Fächern an die DH-Community, um ein sprachliches Gleichgewicht herzustellen?
- Wie muss ein Pflichtenheft für „multilingually enhanced knowledge infrastructures“⁴⁹ aussehen, das *multilingual* und *multiscriptural* DH stärkt?
- Welche Infrastruktur braucht eine mehrsprachige DH?
- Wie kann die Präsenz und der Austausch mit den Kleinen Fächern in den DH (auch auf Konferenzen) gestärkt werden?

Gute Praktiken des mehrsprachigen digitalen Arbeitens werden aktuell von der DHd-AG Multilingual DH gemeinsam mit den *multilingual DH*-Arbeitsgruppen der ADHO und von DARIAH im Rahmen eines Manifests für Multilingualität erarbeitet und dann publiziert.⁵⁰ Damit wird eine Grundlage für reflektiertes multilinguales Arbeiten in den DH geschaffen.

⁴⁹ Siehe Spence und Faria Brandao, „Towards Language Sensitivity and Diversity in the Digital Humanities,“ 1–29.

⁵⁰ Siehe das Programm der DH 2025: Till Grallert et al., „Manifesto for Multilingual Digital Humanities, Workshop SIG,“ Workshop am 15. Juli 2025, Digital Humanities Conference 2025, https://www.conftool.pro/dh2025/index.php?page=browseSessions&form_session=302&presentations=show.

Bibliographie

ACH/ALLC. „Call for Papers 1996.“ *Humanist Discussion Group* 9, Nr. 155 (14. September 1995).

ACH/ALLC. „Call for Papers 1998.“ *Humanist Discussion Group* 11, Nr. 172 (15. Juli 1997).

ACH/ALLC. „Call for Papers 2000.“ *Humanist Discussion Group* 13, Nr. 133 (12. August 1999).

ACH/ALLC. „ACH-ALLC 2001 Call for Papers and Information for Speakers.“ Aufgerufen am 12. Juni 2025. https://web.archive.org/web/20010629125629/http://www.nyu.edu/its/humanities/ach_allc2001/call.html.

ACH/ALLC. „CALL FOR PAPERS and Information for Speakers.“ ALLC/ACH 2002. Aufgerufen am 12. Juni 2025. <http://web.archive.org/web/20020615204323/http://www.uni-tuebingen.de/allcach2002?sflag=cfp>.

ACH/ALLC. „Call for Papers 2003.“ *Humanist Discussion Group* 16, Nr. 227 (28. September 2002).

Amsel, Georg. *Kurzschrift: Lehrbuch der Vereinfachten Deutschen Stenographie (Einigungs-System Stolze-Schrey); nebst Schlüssel, Lese-stücken und einem Anhang*. Berlin: De Gruyter, 1907.

BMBF. „Kleine Fächer – Große Potenziale – BMBF.“ Bundesministerium für Bildung und Forschung. Aufgerufen am 11. Juli 2024. https://web.archive.org/web/20230816105314/https://www.bmbf.de/bmbf/de/forschung/geistes-und-sozialwissenschaften/kleine-faecher/kleine-faecher_node.html.

BMFTR. „Internationale Geistes- und Sozialwissenschaften.“ Forschung, Bundesministerium für Forschung, Technologie und Raumfahrt. Aufgerufen am 6. Dezember 2024. <https://www.geistes-und-sozialwissenschaften-bmbf.de/de/Internationale-Geistes-und-Sozialwissenschaften-1696.html>.

bonndoc. „Textdatenbank und Wörterbuch des Klassischen Maya.“ Institut für Archäologie und Kulturanthropologie (IAK). Aufgerufen am 6. Dezember 2024. <https://bonndoc.ulb.uni-bonn.de/xmlui/handle/20.500.11811/1117>.

Burr, Elisabeth. „Discussion Paper: Internationalization, Multi-Lingual & Multi-Cultural Agenda.“ *ADHO – Standing Committee on Multi-Lingualism & Multi-Culturalism*, Version 1 (4. Juli 2006). https://adho.org/wp-content/uploads/2022/07/MLMC_DiscussionPaper_2006.pdf.

Cremer, Fabian, André Blessing, Patrick Helling, Ulrike Henny-Krahmer, Kerstin Jung und Nils Reiter. „Dhd Chronicles – Anreicherung und Analyse der Beiträge zu den Jahrestagungen der Digital Humanities im deutschsprachigen Raum 2014–2023.“ Zenodo, 21. Februar 2024. <https://doi.org/10.5281/zenodo.10698355>.

DHd: Digital Humanities im deutschsprachigen Raum. „AG Multilingual DH.“ Aufgerufen am 6. Dezember 2024. <https://digitalhumanities.de/ag-multilingual-dh/>.

DHd: Digital Humanities im deutschsprachigen Raum. „MultiHTR.“ Projekt. Aufgerufen am 24. August 2025. <https://digitalhumanities.de/projekt/projekt-2-ueberschrift/>.

DHQ. „Home.“ Digital Humanities Quarterly. Aufgerufen am 12. Juni 2025. <https://www.digitalhumanities.org/dhq/>.

Diehr, Franziska, Sven Gronemeyer, Christian Prager, Elisabeth Wagner, Katja Diederichs, Maximilian Brodhun und Uwe Sikora. „Modellierung von Entzifferungshypothesen in einem digitalen Zeichenkatalog für die Maya-Schrift.“ In *Die Modellierung des Zweifels – Schlüsselideen und -konzepte zur graphbasierten Modellierung von Unsicherheiten*, herausgegeben von Andreas Kuczera, Thorsten Wübbena und Thomas Kollatz. Wolfenbüttel: ZfdG, 2019. https://doi.org/10.17175/sb004_002.

Druskat, Stephan und Cristina Vertan. „Nachnutzbarmachung von Forschungsdaten und Tools am Beispiel altäthiopischer Korpora.“ In *DHd 2018 Kritik der digitalen Vernunft*. Zenodo, 26. Februar 2018. <https://doi.org/10.5281/zenodo.4622354>.

DYBBUK. „The DYBBUK Project: Yiddish Popular Theatre 1880–1920.“ Aufgerufen am 6. Dezember 2024. <https://www.dybbuk.co>.

Faulhaber Edition. „Kritische Online-Edition der Tagebücher Michael Kardinal von Faulhabers (1911–1952).“ Aufgerufen am 6. Dezember 2024. <https://www.faulhaber-edition.de/>.

Freie Universität Berlin. „Introduction to Digital Research Methods – M. Xenia Kudela.“ Vorlesungsverzeichnis. Aufgerufen am 31. Mai 2024. https://www.geschkult.fu-berlin.de/e/semiarab/arabistik/seminar/termine/2024_11_KUDELA_DH-Workshop.html.

Galina Russell, Isabel. „Geographical and Linguistic Diversity in the Digital Humanities.“ *Literary and Linguistic Computing* 29, Nr. 3 (1. September 2014): 307–16. <https://doi.org/10.1093/llc/fqu005>.

GitHub. „Multilingual Digital Humanities.“ Aufgerufen am 6. Dezember 2024. <https://github.com/multilingual-dh>.

GO::DH. „About.“ Global Outlook::Digital Humanities. Aufgerufen am 12. Juni 2025. <http://www.globaloutlookdh.org/>.

- Gogolin, Ingrid. *Der monolinguale Habitus der multilingualen Schule*. Internationale Hochschulschriften, Band 101. Münster, New York: Waxmann, 2008. <https://elibrary.utb.de/doi/book/10.31244/9783830970989>.
- Grallert, Till, Merve Tekgürler, Alíz Horváth, Jana-Katharina Mende, Jonas Müller-Laackmann [sic], Paul Joseph Spence und David Joseph Wrisley. „Manifesto for Multilingual Digital Humanities, Workshop SIG.“ Workshop am 15. Juli 2025, Digital Humanities Conference 2025. https://www.conftool.pro/dh2025/index.php?page=browseSessions&form_session=302&presentations=show.
- Grallert, Till, Xenia Monika Kudela, Eliese-Sophia Lincke, Colinda Lindermann, Jana-Katharina Mende, Jonas Müller-Laackman und Larissa Schmid. „Umgang mit Multilingualität im DACH und DHd Verband.“ Zenodo, 22. Mai 2023. <https://doi.org/10.5281/zenodo.7957187>.
- Horváth, Alíz. „Enhancing Language Inclusivity in Digital Humanities: Towards Sensitivity and Multilingualism,“ *Modern Languages Open* 1: 1–21, <https://doi.org/10.3828/mlo.v0i0.382>.
- Horváth, Alíz, Cornelis Lit, Cosima Wagner und David Joseph Wrisley. *Six User Personas for the Multilingual DH Community*. V1.0. 9. Zenodo, April 2023. <https://doi.org/10.5281/zenodo.7811800>.
- Horváth, Alíz, Cornelis van Lit, Cosima Wagner und David Joseph Wrisley. „Towards Multilingually Enabled Digital Knowledge Infrastructures: A Qualitative Survey Analysis.“ In *Multilingual Digital Humanities*, herausgegeben von Lorella Viola und Paul Spence, 197–212. London: Routledge, 2023. <https://doi.org/10.4324/9781003393696-17>.
- Horváth, Alíz, Cornelis Lit, Cosima Wagner, und David Joseph Wrisley. „Centring Multilingual Users: Thinking Through UX Personas in the DH Community.“ *DH Unbound 2022, Selected Papers* 13, Nr. 3 (2024): 1–30. <https://doi.org/10.16995/dscn.9608>.

Institut für Zeitgeschichte. „Kritische Online-Edition der Tagebücher von Michael Kardinal von Faulhaber (1911-1952).“ Aufgerufen am 6. Dezember 2024. <https://www.ifz-muenchen.de/forschung/ea/forschung/kritische-online-edition-der-tagebuecher-von-michael-kardinal-von-faulhaber-1911-1952>.

Kaselitz, Fritz. *Kritische Würdigung der deutschen Kurzschrift- Systeme von Gabelsberger, Stolze und Arends*. C. W. Born, 1875.

Krech, Volkhard, Tim Karis und Frederik Elwert. *Metaphors of Religion: A Conceptual Framework*. Metaphor Papers 1. Bochum: SFB 1475, 2024. <https://doi.org/10.46586/mp.282>.

Kühnl, Alina. „Iconclass: Ein Klassifizierungssystem für Kunst – und Mensch?“ *the ARTicle* (Blog). *Hypotheses*. 6. Oktober 2020. <https://doi.org/10.58079/uqtu>.

Lethen, Tim. „Gabelsberger Kurzschrift.“ In *Gespräche, Vorträge, Séancen: Kurt Gödels Wiener Protokolle 1937/38*. Veröffentlichungen des Instituts Wiener Kreis 31, 11–16. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-67606-3_2.

Merkel-Hilf, Nicole. „Devanagari Gemischt M1A.“ Transkribus. Aufgerufen am 13. Dezember 2024. <https://readcoop.eu/de/modelle/devanagari-mixed-m1a/>.

MultiHTR - Multilinguale Handschriftenerkennung. „Projektbeschreibung.“ Aufgerufen am 6. Dezember 2024. <https://www.multihtr.uni-freiburg.de/>.

Nicholas, Gabriel und Aliya Bhatia. „Lost in Translation: Large Language Models in Non-English Content Analysis.“ Center for Democracy & Technology. 23. Mai 2023. <https://cdt.org/insights/lost-in-translation-large-language-models-in-non-english-content-analysis/>.

- Nilsson-Fernández, Pedro und Quinn Dombrowski. „Multilingual Digital Humanities.“ In *The Bloomsbury Handbook to the Digital Humanities*, herausgegeben von James O’Sullivan, 83–92. London: Bloomsbury, 2022. <https://doi.org/10.5040/9781350232143.ch-8>.
- Perri, Antonio. „Al di là della tecnologia, la scrittura. Il caso Unicode.“ *Annali - Università degli Studi Suor Orsola Benincasa*, Nr. 2 (2009). <https://universitypress.unisob.na.it/ojs/index.php/annali/article/view/1059>.
- Portal Kleine Fächer. „Was ist ein kleines Fach?“ Kartierung. Aufgerufen am 6. Dezember 2024. <https://www.kleinefaecher.de/kartierung/was-ist-ein-kleines-fach>.
- Romanov, Maxim G. „DH in AAS - TA with R (2022S).“ Syllabus. 17. Juni 2022. https://eis1600.github.io/course_TA_R_2022S/.
- Romanov, Maxim G. „HDIS: Analysis of Prosopographical Data—with R.“ Syllabus. 30. Oktober 2023. <https://eis1600.aai.uni-hamburg.de/hdis-pua/index.html>.
- Schöch, Christof. „Ten Years of the DH(d) and its Disciplines.“ Internationales Wissenschaftsforum: Disciplinary Transformations? Humanities Impact on Reshaping the Digital Humanities, org. Christopher Nunn und Natalie Rauscher. Universität Heidelberg, 12.–13. Juni 2025, <https://dhtrier.quarto.pub/dhd10>.
- Shvedova, Maria, Ruprecht von Waldenfels, Sergey Yarygin, Andriy Rysin, Vasylyl Starko und Tymofij Nikolajenko. „GRAC: General Regionally Annotated Corpus of Ukrainian.“ GRAC, uacorp.us. Aufgerufen am 25. Juni 2025. <https://uacorp.us.org/>.
- Spence, Paul und Renata Faria Brandao. „Towards Language Sensitivity and Diversity in the Digital Humanities.“ *Digital Studies / Le Champ Numérique* 11, Nr. 9 (24. Dezember 2021): 1–29. <https://doi.org/10.16995/dscn.8098>.

Transkribus. „About this Model.“ DEK German Combined. Aufgerufen am 6. Dezember 2024. <https://www.transkribus.org/model/dek-german-combined>.

Transkribus. „About this Model.“ Faulhaber. Aufgerufen am 13. Dezember 2024. <https://www.transkribus.org/model/faulhaber>.

Transkribus. „About this Model.“ Gabelsberger Natural. Aufgerufen am 6. Dezember 2024. <https://www.transkribus.org/model/gabelsberger-natural>.

Transkribus. „About this Model.“ Stolze-Schrey Combined. Aufgerufen am 6. Dezember 2024. https://www.transkribus.org/model/stolze-schrey_combined.

Transkribus. „About this Model.“ Transcription Aid Model 1 for Handwritten Yiddish (Hebrew to Latin). Aufgerufen am 6. Dezember 2024. <https://www.transkribus.org/model/yiddish-to-latin-0.01>.

Transkribus. „About this Model.“ Ukrainian Generic Handwriting. Aufgerufen am 6. Dezember 2024. <https://www.transkribus.org/model/ukrainian-generic-handwriting>.

UNESCO. „Damaged Cultural Sites in Ukraine Verified by UNESCO.“ News. 7. Februar 2024. <https://www.unesco.org/en/articles/damaged-cultural-sites-ukraine-verified-unesco>.

Universität Freiburg. „Künstliche Intelligenz und Küchenrezepte: Ein Programm zur Handschriftenerkennung entziffert alte Briefe, Postkarten oder Tagebücher.“ Archiv der Hochschul- und Wissenschaftskommunikation (bis 08/2024). 25. November 2020. <https://kommunikation.uni-freiburg.de/pm/online-magazin/erleben-und-mitmachen/kuenstliche-intelligenz-und-kuechenrezepte>.

Universität Hamburg. „TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages.“ Press release. Aufgerufen am 6. Dezember 2024. <https://www.aai.uni-hamburg.de/en/ethiostudies/research/traces.html>.

- Vertan, Cristina. „Supporting Hermeneutic Interpretation of Historical Documents by Computational Methods.“ In *Proceedings of Computational Humanities Conference 2018*, 77–82. Lausanne: University of Lausanne 2018. <https://ceur-ws.org/Vol-2314/>.
- Vertan, Cristina, Andreas Ellwardt und Susanne Hummerl. „Ein Mehrebenen-Tagging-Modell für die Annotation altäthiopischer Texte.“ In *DHd 2016 Modellierung – Vernetzung – Visualisierung. Die Digital Humanities als Fächerübergreifendes Forschungsparadigma*. Zenodo, 2. Mai 2017. <https://doi.org/10.5281/zenodo.4645397>.
- Viola, Lorella und Paul Spence, Hg. *Multilingual Digital Humanities*. London: Routledge, 2023. <https://doi.org/10.4324/9781003393696>.
- Weingart, Scott B., Nickoal Eichmann-Kalwara und Matthew Lincoln. „The Index of Digital Humanities Conferences Data.“ Carnegie Mellon University, 22. September 2020. <https://doi.org/10.1184/R1/12987959.v4>.
- Willimsky, Rudolf. „Schreibtechnik.“ In *Sekretärinnen Handbuch*, herausgegeben von Brigitte Diez, 113–75. Wiesbaden: Gabler Verlag, 1982. https://doi.org/10.1007/978-3-663-05217-3_4.
- Wuttke, Ulrike, Melanie Seltsmann, Christian Schröter (geb. Vater), Anne Baillot, Christian Wachter und Christopher Nunn. „Call for Abstracts: From Global to Local? Digitale Methoden in den Geisteswissenschaften im deutschsprachigen Raum: ein Triptychon.“ *DHd-Blog* (Blog), 2. April 2024, aufgerufen am 6. Dezember 2024. <https://dhd-blog.org/?p=20783>.
- Yousef, Tariq, Lisa Mischer, Hamid Reza Hakimi und Maxim Romanov. „The Evolution of Islamic Societies (c. 600-1600): From Algorithmic Analysis into Social History.“ In *Proceedings of the Ancient Language Processing Workshop: Associated with the 14th International Conference on Recent Advances in Natural Language Processing RANLP 2023*, herausgegeben von Adam Anderson und Association for Computational Linguistics, 160–69. Shoumen: Incoma, 2023.