Infrastructures, Cultures, Communities: Herausforderungen von Multilingualität in den Digital Humanities

Till Grallert

Jana Mende

Jonas Müller-Laackman

Cosima Wagner

2022-08-02

# Kurzzusammenfassung

Die Hegemonie der englischen Sprache und lateinischen Schrift in den digitalen Geisteswissenschaften wird bereits seit längerem als Ursache für kulturelle und technologische/digitale Asymmetrien in den Blick genommen. Insbesondere die Nutzung nicht-lateinischer Schriften in digitalen Umgebungen ist mit zur Verfügung stehender Software, Informationssystemen und Infrastrukturen bislang nur in sehr beschränktem Rahmen oder oft auch gar nicht möglich. Der von der AG Multlingual DH organsierte Workshop soll anhand von drei grundlegenden Themenbereichen multilinguale DH Schwerpunkte für die weitere AG Arbeit entwickeln. Dies soll in einem OpenSpace entlang der folgenden Fragen geschehen: Wie können die DH multilingualer werden bzw. die Community der multilingual DH practitioners sich stärker vernetzen? Welche Aufgaben erwachsen daraus für Wissenschaftsinfrastruktur-Verantwortliche wie Bibliotheken, Forschungs-IT Abteilungen etc. um multilinguale DH stärker zu unterstützen bzw. überhaupt erst zu ermöglichen? Welche Rolle kann die Dhd-AG MLDH hier einnehmen?

# Einreichung

## Einleitung

Der Workshop richtet sich an all diejenigen, die in ihrer Arbeit auf die Nutzung von multilingualen digitalen Tools (Plattformen, Software etc.) und insbesondere auf die Verwendung nicht-lateinischer Schriften in digitalen Umgebungen angewiesen sind. Die Hegemonie der englischen Sprache und lateinischen Schrift in dem stark wachsenden Feld der digitalen Geisteswissenschaften wird bereits seit längerem als Ursache für kulturelle und technologische/digitale Asymmetrien in den Blick genommen (Fiormonte 2017, 2021; Fiormonte, Chaudhuri, und Ricaurte 2022; Gil und Ortega 2016; Golumbia 2013; Liu 2012; Ortega 2014; Spence 2021). Insbesondere die Nutzung nicht-lateinischer Schriften (NLS) in digitalen Umgebungen ist mit zur Verfügung stehender Software, Informationssystemen und Infrastrukturen bislang nur in sehr beschränktem Rahmen oder oft auch gar nicht möglich (Asef und Wagner 2018; Dombrowski 2020a, 2020b; Ghorbaninejad, Gibson, und Wrisley 2022; Horvath 2022; Mahony 2018).

Seit der DHd 2022 widmet sich die AG Multilingual DH (<https://dig-hum.de/ag-multilingual-dh>) dem Anliegen der Unterstützung von Multilingualität in DH-Infrastrukturen, der Erarbeitung von Lösungsansätzen und Verbesserung technischer Anforderungen im Bereich der Multilingualität durch die Sichtbarmachung der Probleme im Austausch mit Entwickler\*innen sowie der kritischen Diskussion der deutschen und internationalen DH-Praxis im Hinblick auf (neo)koloniale Praktiken z.B. bezüglich Vokabularen, monolingualen Infrastrukturen, Methodologien und Themenfeldern. Wie können die DH multilingualer werden bzw. die Community der multilingual DH practitioners sich stärker vernetzen, welche Aufgaben erwachsen daraus für Wissenschaftsinfrastruktur-Verantwortliche wie Bibliotheken, Forschungs-IT Abteilungen etc. um multilinguale DH stärker zu unterstützen bzw. überhaupt erst zu ermöglichen? Welche Rolle kann die Dhd-AG MLDH hier einnehmen? Diese Fragen sollen im von der AG Multilingual DH organisierten Workshop anhand von drei grundlegenden Themenbereichen multilingualer DH (Definitionen von Multilingualität/Mehrsprachigkeit, Monolingualität; Multilingualität und Wissenschaftsinfrastrukturen; “user stories” aus der multilingualen Digital Humanities Praxis) diskutiert sowie Schwerpunkte für die weitere AG Arbeit entwickelt werden.

## Herausforderungen multilingualer DH

Digitalität wohnt ein großes egalitäres Heilsversprechen inne: der uneingeschränkte Zugang zum Wissen der Welt – für jede\*n, an jedem Ort, zu jeder Zeit. Erweitert für die DH versprechen wir uns die Beantwortung zentraler Forschungsfragen durch die Anwendung computationeller Methoden auf das digital(isiert)e Kulturerbe der Menschheit. Die Grundlage für den Glauben an ein solches Heilsversprechen ist die Verschleierung der Zugangsvoraussetzungen für die Teilhabe an dieser schönen neuen Welt. Metaphern in Interfaces und Zeichenkodierungen, d.h. der Möglichkeit geschriebene Sprache digital abzubilden, sind die Basis sämtlicher Mensch-Maschine Interaktionen und begründen einen fundamentalen *digital divide* zwischen hegemonialem, anglophonen Globalen Norden und Globalen Süden.

Vor diesem Hintergrund setzt sich der Workshop mit drei grundlegenden Themenbereichen dazu auseinander und behandelt allgemeine Fragen zur Mehrsprachigkeitsdefinitionen in den DH, darunter auch Fragen zum Verhältnis von Programmiersprachen und natürlichen Sprachen, das Verhältnis von Einsprachigkeit und Mehrsprachigkeit in Wissenschaftsinfrastrukturen sowie konkrete Probleme aus der mehrsprachigen, digitalen Praxis.

Die Beschäftigung mit Mehrsprachigkeit im Kontext der digitalen Geisteswissenschaften steht dabei unter dem Vorzeichen anglophoner Sprachdominanz, die sich auf verschiedenen Ebenen zeigt: Englisch ist die häufigste Arbeitssprache für (internationale) DH-Projekte, Journals, Konferenzen und Veranstaltungen (vgl. Fiormonte, Chaudhuri, und Ricaurte 2022, XIX). Wenn sich DH Gemeinschaften der anglophonen Anpassung entziehen, dann leiden sie oft darunter, international nicht wahrnehmbar und damit unsichtbar zu sein (vgl. Fiormonte, Chaudhuri, und Ricaurte 2022). Außerdem wirkt sich die anglophone Einsprachigkeit auf Projekte, Tools, Korpora, Metadaten und Infrastrukturen aus. Mehrsprachigkeitsdefinitionen in den DH beziehen sich dadurch erstens auf die Kritik (dominanter) Einsprachigkeit (vgl. Fiormonte 2021) und befassen sich mit Sprachenvielfalt, indem Tools, Korpora, Daten, Projekte in unterschiedlichen Sprachen bzw. auf unterschiedlichen Sprachen verfügbar gemacht werden. Eine besondere Positionen kommt dabei Sprachen mit geringen Sprecher\*innenzahlen sowie Sprachen ohne Schrifttradition zu. Mehrschriftigkeit ist ebenfalls ein Aspekt dieser Form von Mehrsprachigkeit. Mehrsprachigkeit und Mehrschriftigkeit gehen hier von einem Nebeneinander unterschiedlicher Sprachen aus. Zweitens befassen sich mehrsprachige DH-Ansätze mit inhärent mehrsprachigen Projekten, die mit mehreren Sprachen, über Sprachgrenzen hinweg oder zu gemischt-sprachlichen Phänomenen (fused lects) arbeiten. Hierzu gehören etwa Übersetzungen, die automatisierte Erkennung von Code-Switching oder Sprachvergleiche. Auch das Arbeiten mit verschiedenen historischen Sprachstufen fällt in diese Kategorie. Zuletzt umfassen Mehrsprachigkeitskonzepte in den DH auch das Verhältnis von Markup-Sprachen, Programmiersprachen, Beschreibungssprachen zu natürlichen Sprachen. Auch die individuelle Mehrsprachigkeit und Sprachkenntnisse der einzelnen Forschenden bzw. der Forschungsgruppen können in dem Zusammenhang reflektiert werden und sind auch häufig Ausgangspunkt in Diskussionen zum Mehrsprachigkeitsbegriff in den DH (vgl. Horvath 2021, 2).

Die verschiedenen Mehrsprachigkeitskonzeptionen und -konstellationen und implizite sprachpolitische Einstellungen wie die Dominanz des Englischen und die Marginalisierung kleinerer Sprachen beeinflussen sich gegenseitig. Im zweiten Schwerpunkt des Workshops soll es daher im die Rolle von Akteur\*innen und Infrastrukturen gehen, die im Kontext globaler DH sprachpolitisch wirksam sind.

Der *digital divide* zeigt sich strukturell und institutionell auf verschiedenen Ebenen. So ist Unicode ein gigantischer Fortschritt gegenüber ASCII. Aber es ist auch ein in Kalifornien inkorporiertes Industriekonsortium ohne jegliche demokratische Legitimation, das zum einen allen Schriftsystemen die Buchstabenmetapher überstülpt und zum anderen der Idee von Nationalsprachen anhängt. Im Ergebnis gibt es für Suchalgorithmen 32 verschiedene Schreibweisen von "mekkanisch" in arabischen Buchstaben, die für arabisch-sprachige Menschen allerdings identisch sind. Dazu kommt die Verschleierung der realen Kosten der Digitalität, d.h. des notwendigen, kontinuierlichen Einsatzes menschlicher Arbeitszeit und natürlicher Resourcen, für Bereitstellung, Aufbau und Erhaltung digitaler Inhalte und Infrastrukturen. Ein berühmter Protestslogan, der sich von Baghdad aus im östlichen Mittelmeerraum verbreitet hat, lautet "Du bist so schön, wie eine zusätzliche Stunde Strom". Weite Teile des Globalen Südens haben weder kontinuierlichen Strom noch Internetzugänge. Ihr in der Folge des Kolonialismus sich häufig in den Sammlungen des Globalen Nordens befindliches Kulturerbe wird, wenn überhaupt, häufig ohne ihre Beteiligung digitalisiert. Kataloge und Interfaces liegen fast ausschließlich in den hegemonialen Sprachen des Globalen Südens vor, während Webbrowser basale Funktionalität von HTML5 für die Darstellung nichtlateinischer Sprachen nicht implementiert haben. So führt die technische Angabe, dass ein Text in Arabisch geschrieben ist (@xml:lang="ar"), nicht dazu, dass dieser korrekt rechtsbündig angezeigt wird. Diese Nichtbeachtung setzt sich durch die Bank fort. Können Algorithmen für OCR, NLP, NER, Stilometrie, Topic Modelling ohne zusätzlichen Aufwand für Sprachen und Schriften des Globalen Südens verwandt werden und stellen Interfaces für diese Algorithmen diese Schriften korrekt dar? Die Antwort ist fast immer “nein!”. Im Ergebnis sind die Gesellschaften des Globalen Südens mit einer permanenten Entfremdung, Notwendigkeiten der Übersetzung und einen zusätzlichen Aufwand konfrontiert, während für die Gesellschaften des Globalen Nordens die konstante Mediation durch sozio-technische Systeme durch die unmittelbare Vertrautheit ihrer Metaphern und Zeichensysteme und den vergleichweise ubiquitären Zugang zur notwendigen technischen Infrastruktur quasi unsichtbar wird.

Aus diesen strukturellen Bedingungen ergeben sich zugleich konkrete Probleme aus und in der Praxis von Area Studies Fächern im Globalen Norden (wissenschaftspolitisch in Deutschland i.d.R. als sogenannte "Kleine Fächer" marginalisiert, wenn auch aktuell mit neuen Förderlinien fokussiert, (vgl. BMBF o. J.)), für die die Arbeit mit Quellen in nicht-lateinischen Schriften sowie die Zusammenarbeit mit Forschenden und Lehrenden in den Regionen Arbeitsgrundlage des Faches ist und die auf Multilingualität und internationale Zusammenarbeit ermöglichende Infrastrukturen angewiesen sind (vgl. Kirmizialtin und Wrisley 2022). Eine weiterführende im Zusammenhang mit dem Themenschwerpunkt der DHd Konferenz 2023 stehende Frage ist hier zugleich, wie Open (Digital) Scholarship im Rahmen von transnationalen Kontexten gefördert werden kann (vgl. Grallert 2022; Wrisley 2019). Der Workshop öffnet einen Raum, anhand von “user stories” der multilingualen DHd-community und zusammen mit allen Teilnehmenden die genannten Herausforderungen zu fokussieren und zu erörtern.

# Workshopkonzeption

Der eintägige Workshop beginnt mit einer Q&A-Session, in der die Workshop-Veranstalter\*innen anhand eines visualisierten Charts zunächst die zentralen Themengebiete vorstellen, mit denen wir uns im Rahmen von Multilingualität in den Digital Humanities auseinandersetzen. Dabei werden zugleich von AG-Mitgliedern und anderen Interessierten eingereichte “user stories” vorgestellt, die die praktischen Herausforderungen und Desiderate des forschens und lehrens mit multilingualen Quellen verdeutlichen. In diesem Kontext ist es für alle Teilnehmer\*innen möglich, ein gemeinsames Verständnis für die sich stellenden Herausforderungen zu entwickeln, Fragen zu stellen und eine vertiefte inhaltliche Auseinandersetzung mit dem Thema zu üben. Hierbei wird genügend Raum und Zeit gegeben, dass alle Teilnehmer\*innen gleichermaßen dem weiteren Verlauf des Workshops folgen können.

Im Anschluss wird der Workshop als OpenSpace weitergeführt, in dem die Teilnehmer\*innen auf der Basis ihrer eigenen Wünsche und Vorstellungen das weitere Programm entwickeln. Hierbei folgen wir den zentralen Prinzipien der OpenSpace-Methode:

1. *Jede Person ist die richtige Person*: Unabhängig vom fachlichen Hintergrund kann jede\*r Teilnehmer\*in ihren kreativen Beitrag zu ihren Herausforderungen bzw. Lösung von Problemen mit Multilingualität leisten
2. *Was auch immer geschieht, ist das einzige was geschehen sollte*: Der kreative Freiraum bietet Teilnehmer\*innen mehr Möglichkeiten, flexibel und offen über die Themengebiete, die sie unter dem Thema multilinguale DH am meisten interessieren, zu diskutieren
3. *Es beginnt, wenn die Zeit reif ist*: Auch, wenn wir einen Zeitrahmen vorgeben, ist den Teilnehmer\*innen freigestellt, ihre Pausen und Arbeitsphasen selbstständig und nach den eigenen Bedürfnissen zu legen, um im richtigen Moment kreativen Austausch zu ermöglichen
4. *(Nicht) Vorbei ist (nicht) vorbei*: Ebenso bieten wir den Arbeitsgruppen die Freiheit, selbst zu bestimmen, wann sie zu einem Ende kommen und ihre abschließenden Ergebnisse formulieren wollen.

Es können selbstständig Themenwünsche eingebracht werden, die im Kontext zu den drei vorgestellten Themenbereichen stehen, oder auch neue Themenbereiche erschließen. Nach einer kurzen Findungsphase werden diese Themen dann in Arbeitsgruppen diskutiert und bearbeitet. Entsprechend dem OpenSpace-eigenen *Gesetz der zwei Füße* ist es jederzeit möglich und erwünscht, zwischen den Gruppen zu wechseln oder auch kurzfristig neue Situationen des kreativen Austauschs zu finden und zu schaffen.

Zuletzt entwickeln die Arbeitsgruppen in einem One-Minute-Paper gebündelt zentrale Positionen, Probleme und Fragestellungen, die dann im Plenum gemeinsam vorgestellt und diskutiert werden. One-Minute-Papers sind eine Methode zur Ergebnissicherung und sollen dazu beitragen, die diskutierten Themen und Ergebnisse auch über den Workshop hinaus produktiv weiterzuführen. Hieraus soll abschließend ein Chart entstehen, das sowohl die Ergebnisse des Workshops zusammenfasst, als auch einen visualisierter Wegweiser für die künftige Arbeit der DHd AG Multilingual DH darstellt und entsprechend publiziert wird.

## Ziel und Ergebnis

Eine visualisierte Präsentation in Form eines Charts, die im Workshop erarbeitete Herausforderungen und Probleme, sowie Lösungsansätze und Perspektiven zusammenfasst und sowohl als eigenständiges Werk publiziert wird, als auch einen Wegweiser für die künftige Arbeit der DHd AG Multilingual DH darstellt.

## Zielgruppe

Der Workshop wird von Mitgliedern der AG Multilingual DH ausgerichtet und wendet sich an alle, die sich mit mehrsprachigen DH-Projekten beschäftigen oder sich generell für Mehrsprachigkeit in den DH interessieren. Spezielles Vorwissen ist nicht notwendig.

## Organisator\*innen / Beitragende

Die im Sinne des CfP Beitragenden dieses Workshops sind die einreichenden Organisator\*innen in ihrer Funktion als Co-convenors der AG Multilingual DH: Till Grallert (HU Berlin, [till.grallert@ub.hu-berlin.de](mailto:till.grallert@ub.hu-berlin.de)), Jana Mende (Universität Halle-Wittenberg, [jana-katharina.mende@germanistik.uni-halle.de](mailto:jana-katharina.mende@germanistik.uni-halle.de)), Jonas Müller-Laackman (Universität Hamburg, [jonas.mueller-laackman@fu-berlin.de](mailto:jonas.mueller-laackman@fu-berlin.de)) und Cosima Wagner (FU Berlin, [cosima.wagner@fu-berlin.de](mailto:cosima.wagner@fu-berlin.de)). Die Forschungsinteressen konvergieren um Multilinguale DH, v.a. im Bereich nicht-lateinische Schriften mit Schwerpunkt Arabisch, Chinesisch, Japanisch und Koreanisch, Qualitative und quantitative Ansätze mehrsprachiger Literaturgeschichte, Literatur und Mehrsprachigkeit, digitale Literaturwissenschaft, Postkoloniale DH und Dekolonisierung in Forschungsinfrastrukturen und -praxis, kritische Algorithmen-Studien und Wissenschaftsinfrastrukturen, Kritik epistemischer Gewalt, Gender- und Metawissenschaften/Kritische DH, DH Pädagogik.

## Ablauf

9:00 - 9:15 Uhr Vorstellungsrunde  
9:15 - 10:00 Uhr Impuls und Q&A  
10:00 - 10:15 Uhr Kaffeepause  
10:15 - 10:45 Uhr OpenSpace Themenfindung  
10:45 - 15:15 Uhr OpenSpace, freie Einteilung der Arbeitszeit, Pausen nach Bedarf  
15:15 - 15:30 Uhr Kaffeepause  
15:30 - 16:30 Uhr Plenum  
16:30 - 17:00 Uhr Formulierung der Ergebnisse und Abschluss

## Zahl der möglichen Teilnehmerinnen und Teilnehmer

20

## Benötigte technische Ausstattung

Pinnwände, Beamer

# Notizen

## Formalia

* 1200 Wörter, max. 1500
* Deadline: 2023-08-03
* Kurzzusammenfassung der Einreichung mit ca. 100–150 Wörtern
* Kontaktdaten aller Beitragenden sowie einen Absatz zu deren Forschungsinteressen
* die Zahl der möglichen Teilnehmerinnen und Teilnehmer
* Angaben zu einer etwa benötigten technischen Ausstattung
* KEIN extra call für Workshop!

# Literatur

Asef, Esther, und Cosima Wagner. 2018. „Workshop-Bericht „Nicht-lateinische Schriften in multilingualen Umgebungen: Forschungsdaten und Digital Humanities in den Regionalstudien““. DHd Blog. 24. Oktober 2018. <https://dhd-blog.org/?p=10669>.

BMBF. o. J. „Kleine Fächer – Große Potenziale - BMBF“. Bundesministerium für Bildung und Forschung - BMBF. Zugegriffen 1. August 2022. <https://www.bmbf.de/bmbf/de/forschung/geistes-und-sozialwissenschaften/kleine-faecher/kleine-faecher_node.html>.

Dombrowski, Quinn. 2020a. „Preparing Non-English Texts for Computational Analysis“. *Modern Languages Open*, 1, August, 45. <https://doi.org/10.3828/mlo.v0i0.294>.

———. 2020b. „What’s a ‚Word‘: Multilingual DH and the English Default“. Online, Oktober 15. <https://quinndombrowski.com/blog/2020/10/15/whats-word-multilingual-dh-and-english-default/undefined>.

Fiormonte, Domenico. 2017. „Digital Humanities and the Geopolitics of Knowledge“. *Digital Studies / Le Champ Numérique* 7 (1). <https://doi.org/10.16995/dscn.274>.

———. 2021. „Taxation Against Overrepresentation? The Consequences of Monolingualism for Digital Humanities“. In *Alternative Historiographies of the Digital Humanities*, herausgegeben von Dorothy Kim und Adeline Koh, 333–76. Earth: punctum books. <https://doi.org/10.53288/0274.1.00>.

Fiormonte, Domenico, Sukanta Chaudhuri, und Paola Ricaurte, Hrsg. 2022. „Introduction“. In *Global Debates in the Digital Humanities*, ix–xxxiii. Debates in the Digital Humanities. Minneapolis: University of Minnesota Press. <https://www.upress.umn.edu/book-division/books/global-debates-in-the-digital-humanities>.

Ghorbaninejad, Masoud, Nathan P. Gibson, und David Joseph Wrisley. 2022. „RTL“. In *Debates in the Digital Humanities 2022*. Debates in the Digital Humanities.

Gil, Alex, und Élika Ortega. 2016. „Global Outlooks in Digital Humanities: Multilingual Practices and Minimal Computing“. In *Doing Digital Humanities*, herausgegeben von Richard Lane, Siemens, Raymond, und Constance Crompton. Routledge.

Golumbia, David. 2013. „Postcolonial Studies, Digital Humanities, and the Politics of Language“. Uncomputing. 31. Mai 2013. <https://www.uncomputing.org/?p=241>.

Grallert, Till. 2022. „Open Arabic Periodical Editions: A Framework for Bootstrapped Scholarly Editions Outside the Global North“. *Digital Humanities Quarterly* 16 (2).

Horvath, Aliz. 2021. „Enhancing Language Inclusivity in Digital Humanities: Towards Sensitivity and Multilingualism: Includes Interviews with Erzsébet Tóth-Czifra and Cosima Wagner“. *Modern Languages Open*, Nr. 1 (Dezember): 26. <https://doi.org/10.3828/mlo.v0i0.382>.

———. 2022. „Digital Brush Talk: Challenges and Potential Connections in East Asian Digital Research“. In *Global Debates in the Digital Humanities*, herausgegeben von Domenico Fiormonte, Sukanta Chaudhuri, und Paola Ricaurte, 127–40. Debates in the Digital Humanities. Minneapolis: University of Minnesota Press. <https://www.upress.umn.edu/book-division/books/global-debates-in-the-digital-humanities>.

Kirmizialtin, Suphan, und David Joseph Wrisley. 2022. „Automated Transcription of Non-Latin Script Periodicals: A Case Study in the Ottoman Turkish Print Archive“. *Digital Humanities Quarterly* 16 (2).

Liu, Alan. 2012. „Where Is Cultural Criticism in the Digital Humanities?“ In *Debates in the Digital Humanities*, herausgegeben von Matthew K. Gold, 490–509. Debates in the Digital Humanities 1. Minneapolis: University of Minnesota Press. <https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfbd1e/section/896742e7-5218-42c5-89b0-0c3c75682a2f#ch29>.

Mahony, Simon. 2018. „Cultural Diversity and the Digital Humanities“. *Fudan Journal of the Humanities and Social Sciences* 11 (3): 371–88. <https://doi.org/10.1007/s40647-018-0216-0>.

Ortega, Élika. 2014. „Multilingualism in DH“. Disrupting the Digital Humanities | 2015 MLA Position Papers. 31. Dezember 2014. <https://web.archive.org/web/20210424073656/https://www.disruptingdh.com/multilingualism-in-dh/>.

Spence, Paul. 2021. „Disrupting Digital Monolingualism: A report on multilingualism in digital theory and practice“. London: Language Acts & Worldmaking project. <https://doi.org/10.5281/zenodo.5743283>.

Wrisley, David Joseph. 2019. „Enacting Open Scholarship in Transnational Contexts“. *Pop! Public. Open. Participatory*, Nr. 1 (Oktober). <https://doi.org/20191217170056885>.