

Multilinguality in Action: Towards Linguistic Diversity and Inclusion in Digital Humanities

Alíz Horváth, Cosima Wagner, David Joseph Wrisley,
Joanne Bernardi, Aanya Chadha, Danielle Garcia,
Till Grallert, Sharanya Ghosh, Yuri Ishida,
Aleksandra Kaye, Ahac Meden, Kiyonori Nagasaki
Dylan Palmer, Hugo Scheithauer, Marta Świetlik,
Jeffrey Tharsen, Yifan Wang

Abstract The article addresses the multilingual landscape in Digital Humanities, focusing on understanding its practitioners. We adopt the concept of user profiles from UX design to help create visibility and empathy for the unique needs of multilingual scholars. In a DH2023 workshop, using a dataset of six user profiles, participants examined multilingual DH, exploring the complex interaction between language use, identity, inclusivity, and infrastructure. Only by including multilingual perspectives, we argue, can DH promote diverse knowledge systems towards more supportive infrastructures and a more inclusive scholarly community.

Keywords Multilinguality. Multilingual DH. Non-Latin scripts. User personas. Research infrastructures.

Summary 1 Introduction. – 2 Multilinguality to the Fore. – 3 A Workshop to Address the Question of User Profiles. – 4 Decoupling Language Culture from Infrastructure. – 5 User Profiles for Representation and Fostering Inclusivity in DH Research. – 6 A Wider Perspective on Literacies. – 7 Moving Forward. – 8 Conclusion.



Peer review

Submitted 2024-04-15
Accepted 2024-06-18
Published 2024-12-17

Open access

© 2024 Horváth et al. | © 4.0



Citation Horváth, A. et al. (2024). "Multilinguality in Action: Towards Linguistic Diversity and Inclusion in Digital Humanities". *magazén*, 5(2), 179-196.

1 Introduction

In an era marked by claims of creating global interconnected knowledge, the multilingual landscape of Digital Humanities (DH) presents both promise and challenge. The creation of user profiles, stemming from persona creation in UX design, has emerged as a useful strategy for understanding the diverse needs and obstacles faced by researchers in multilingual environments. Through discussing the intricacies of the experiences of multilingual digital humanists, it is easy to see that there is an interplay between individual language usage and infrastructural dynamics. At a recent writing sprint at the 2023 ADHO conference, scholars converged to explore the dynamics of multilingual DH. From this collaborative endeavour, a number of themes emerged providing invaluable insights into the complexities of navigating linguistic diversity in digital scholarship. As we assess the current state of multilingual DH, it becomes evident that innovative initiatives exist on the horizon. These initiatives not only seek to bring many more multilingual scholars to the table and to centre their voices, but also to engage in critical conversations about decolonizing knowledge spaces. We contend that engaging with multilingual DH necessitates making room for a wide spectrum of stakeholders who work in and across multiple languages. By creating space for diverse voices, we enrich the scholarly discourse and increase our chances for meaningful, inclusive knowledge communities. Ultimately, acknowledging the crucial role of infrastructures for digital scholarship means making space for discussion of the role of humanists in shaping inclusive infrastructures.

Alíz Horváth (Central European University); Cosima Wagner (Freie Universität Berlin, Germany); David Joseph Wrisley (New York University Abu Dhabi, United Arab Emirates); Joanne Bernardi (University of Rochester, USA); Aanya Chadha (Indian Institute of Technology Jodhpur, India); Danielle Garcia (University of Rochester, USA); Till Grallert (Humboldt-Universität zu Berlin, Germany); Sharanya Ghosh (Indian Institute of Technology Jodhpur, India); Yuri Ishida (Okayama University, Japan); Aleksandra Kaye (Max Planck Institute for the History of Science Berlin, Germany); Ahac Meden (ZRC SAZU, Slovenia); Kiyonori Nagasaki (International Institute for Digital Humanities, Japan); Dylan Palmer (University of Rochester, USA); Hugo Scheithauer (ALMAAnaCH, Centre Inria de Paris, France); Marta Świetlik (Institute of Literary Research PAS, Poland); Jeffrey Tharsen (University of Chicago, USA); Yifan Wang (International Institute for Digital Humanities, Japan).

The authors of this article would like to thank the two anonymous reviewers for their thoughtful engagement with our ideas and for their suggestion for revisions, as well as the ADHO 2023 Program Committee for including our workshop and writing sprint on the program.

2 Multilinguality to the Fore

In prior investigations, we explored the distinction between multilingualism and multilinguality, a differentiation underscored by linguists to differentiate between the broader effort to define multiple “official” languages and more personal, strategic utilization of multilingual resources by individuals (Horváth et al. 2024a; 2024b). Given this distinction, it is perfectly possible for someone to be “multilingual”, yet to function in languages outside of the official multilingualism of the territory in which they live and work. This distinction between the personal and the official is particularly salient in the realm of textual Digital Humanities, where researchers actively engage with a vast array of linguistic tools in their scholarly endeavours. What we call “scholarly multilinguality” is a domain that requires a significant amount of reflection, since it both has the potential to enrich our research production and to form a stumbling block for digital research.

Our research into user profiles revealed that researchers in Digital Humanities routinely leverage their varied linguistic resources in various aspects of their daily and professional activities, with differing degrees of success and ease. For example, an individual might use a second language (L2) for everyday communication within an academic setting or for more casual scholarly exchange and communication. This operational use of language contrasts with the use of what are commonly termed “research languages”. The latter refer to languages of specialization required for accessing archival materials, conducting source-based inquiries, or performing other scholarly tasks. Research languages are skills that researchers do not necessarily possess in extenso from the beginning of their PhD, but are rather developed over the course of a career. This use of research languages is a fascinating area of study, particularly as it often involves engagement with historical languages in different states of evolution or very different contextual features, where factors such as orthography, script, or abbreviation play critical roles in forms of professional reading. As the documentation, collection and even dissemination of textual data in one’s research languages becomes increasingly digital, a paradox presents itself. Although the promise of universal character encoding would seem to suggest that any textual data can be handled with ease, in reality, not all languages are on an equal footing. When discussing the multilingual capabilities of researchers in languages and literature, it is essential to acknowledge how they navigate and utilize these complex layers of language in their work, from note-taking and citation management to corpus creation, analysis, dissemination of research findings, and ultimately, the publication of their work in either digital or print formats, including how easy or seamless it is for them to combine these languages in the same workflow or publication.

This concept of multilinguality stands in contrast to multilingualism, which typically refers to a more static, agreed-upon set of languages used within a specific political geography. The limitations of traditional multilingualism are readily apparent to those living or working in a context where their essential linguistic tools derive from geographically or temporally distant sources. Moreover, specialists in historical humanities disciplines – such as history, philology, and literary history – readily understand that the linguistic resources necessary for modern or contemporary scholarship differ significantly from those needed to explore older states, or regionally variant forms, of language. The concept of scholarly multilingualism might then be introduced, referring to the expectation for scholars to disseminate research and teach in multiple languages, often influenced by national or regional expectations. While research infrastructures, like library catalogues or research platforms, may support this form of multilingualism to varying degrees they do so by performing a state function. And even then the globalization of the research publishing industry predominantly encourages a tilt towards English as the common language of scholarship.

Addressing the interplay between infrastructure and scholarly multilinguality, we observe that while infrastructures can both constrain and enable scholarly work, bringing together diverse users within a shared space, they do not inherently resolve the tension between these elements. Commercial infrastructures like Google Workspace and Microsoft 365 have made notable advancements in supporting a wide array of global languages with functionality such as machine translation, multilingual search functions, and transliteration tools. The suitability of these commercial tools for scholarly work warrants careful consideration, as they are usually more suited to general purpose uses by a wider public. Academic knowledge infrastructures are built for more specific uses and necessarily smaller audiences, and they do not receive the same amount of funding for system administration, research and development, let alone maintenance.

Expanding this dialogue, it becomes crucial to emphasize understanding and prioritizing the needs that will most significantly impact the scholarly community and finding ways to articulate the requirements of specific multilingual communities, not only from the perspective of legal compliance, but also as a question of language justice. By engaging in continuous dialogue with users and stakeholders, infrastructure providers can better identify and address the most pressing needs, fostering a more inclusive and effective environment for scholarly multilinguality. This approach does not merely aim to accommodate a wider array of linguistic resources, but also acknowledges the evolving nature of scholarly work and the diverse linguistic landscapes in which it operates. Through collaborative effort and

a commitment to language diversity and inclusivity, we can strive toward infrastructures that challenge conventional thinking through their ability to support and enrich the multilingual and multifaceted nature of contemporary scholarship. The question of multilingual infrastructures must not be left to tech companies to resolve, but as we have argued elsewhere (Horváth et al. 2024b, 4), digital humanists should have a place at the table to suggest innovative solutions.

3 A Workshop to Address the Question of User Profiles

To this end, we organized a workshop for the Alliance of Digital Humanities Organizations (ADHO) 2023 annual conference in Graz, Austria in order to explore the state of multilinguality in DH, based upon extensive discussion of six data-driven user profiles we had previously created and published (Horváth et al. 2023, dataset). These personas included a variety of knowledge actors in multilingual DH academic environments: an assistant professor, a full professor, a doctoral student, an academic technology specialist, a librarian, and an independent scholar. Such profiles were created to showcase various realities and challenges that people experience while navigating their multilinguality in global, digital contexts.

With the fourteen participants of the workshop, our objective was to think through collectively the aspirations and difficulties of DH researchers working with non-Anglophone textual data, particularly text written in non-Latin and right-to-left scripts. Following a guided discussion, our main goal was a co-writing experience in the interest of knowledge sharing and community building. The idea of having a co-writing sprint at the ADHO conference was first and foremost one of knowledge sharing and collaborative learning. The authors of this paper have been researching and publishing on the topic of multilinguality in DH for some time now, but our perspectives are enhanced by the diversity of experiences that one finds at a global conference. The contributing group included individuals with significant experience in different geographical areas, with varied linguistic backgrounds, and holding positions in global academia in France, Germany, Hungary, India, Japan, Poland, Slovenia, the United States, and the United Arab Emirates. Admittedly, this variety of perspectives was not fully global, but nonetheless provided a rich starting point from which to enrich and refine existing debates about multilinguality in DH. All participants were invited to be a co-author of this article.

We believe in the importance of strengthening the agency of digital humanists in contributing to the design of workflows sensitive to key multilingual issues. However, as our workshop also confirmed, a relative lack of awareness of the complex issues that multilingual users face still persists in the broader DH community. These problems

are particularly pertinent to the limitations of infrastructural functionalities in the context of under-resourced languages and scripts, which have been significantly neglected.¹ Our stated goal was thus to “create space for dialogue and debate, as well as to introduce and develop further the creation of multilingual DH UX personas” to explore “how these debates can inform future plans for the multilingual DH community” as well as “how multilingual user personas can be recreated in local environments and globally” (Horváth et al. 2023).

The workshop consisted of a discussion of a pre-circulated paper which itself had been based on the creation and exploration of six user personas that we, the organizers, had previously created as a dataset (Horváth et al. 2023). Following the discussion, we held a collaborative writing sprint with all the participants. The result was an extensive collection of insights regarding multilingual aspects in DH and the potential roles of the user profiles within the larger DH community, far too many topics than can be reported on in a short length piece of writing. In the present article we have chosen to address a specific strain of insights which emerged from the writing sprint. More specifically, the participants elaborated on the need to challenge the monolingualist hegemony of digital systems, to use user profiles as a way of imagining and increasing more inclusive participation in DH research, and to expand our notion of knowledge infrastructures as social objects intersecting with a variety of literacies. These nuanced approaches to multilinguality underscored the roles and significance of language in research and provided potential means and contexts to refine the idea of user profiles in practical ways. Interestingly, the former aspect regarding the question of language eventually not only included discussions related to scholarly multilinguality and language-specific issues in the development of knowledge infrastructures, but also revealed hitherto largely overlooked, potential connections and confluences between DH and other fields such as identity, literacy and even language acquisition.

In what follows, we attempt to shed light on the multifaceted role of language in DH practice and claim that the question of language cannot be considered in isolation, nor in DH alone, but rather it should be discussed in relation to relevant fields which can in turn mutually influence each other.

1 Fiorimonte 2021; Ghorbaninejad, Gibson, Wrisley 2023; Columbia 2013; Grallert 2022; Horváth 2021; 2022; Horváth et al. 2024a; 2024b; Kirmizialtin, Wrisley 2022; Spence, Brandao 2021; Wrisley 2019.

4 Decoupling Language Culture from Infrastructure

Arguments have been forming around the importance of language equity in the existing scholarship, which can be witnessed, for example, in the publication of volumes such as *Global Language Justice*, exploring the emergent theme of language marginalization and the importance of linguistic diversity and language loss from a variety of perspectives: social justice, environmentalism, human rights, technological advances, and infrastructure development. As one of the editors, Lydia Liu mentioned in a recent interview, the purpose for the creation of this multifaceted edited volume was not to “advocate for the preservation of languages as fixed cultural artifacts” but to “ask how current circumstances might produce new formations and emergent lifeworlds” while considering language as an “open, experimental frontier” (Glasberg 2024).

Our pre-conference 2023 workshop discussion reflected the value of that argument; the participants expressed their need for multilingually enabled knowledge infrastructures to support their daily research and teaching practices. In a globally connected academe, they insisted that infrastructures should be accessible for as many people as possible. They criticized the assumption in application and interface design that localization can seamlessly bring digital environments to global language communities, on account of the fact that such apps and interfaces have not been designed with internationalization in mind. More thought should be given to the ways in which people interact with systems of thought in language, rather than assuming everything can flow – through translation from English or any other hegemonic language – to communities of a “target language”. Multilingual options should be standard for academic knowledge infrastructures which means bringing more multilingual minds to the design process; developing tools for English speakers that await translation for the world creates a new manifestation of the digital divide.

Another aspect of why language culture should be decoupled from knowledge infrastructures in a digitally connected world has to do with the abundance of digital sources of knowledge already found in non-Latin scripts, available, but perhaps siloed or undiscoverable, in digital knowledge spaces of non-Anglophone countries. In order to enable a culture of truly globally connected knowledge, libraries (and other GLAM infrastructures) in the “West” who have adopted English as one of their working languages, should strive to secure the interoperability of knowledge infrastructures, not only within their own language sphere, but also with non-Western, non-Anglophone countries and data in non-Latin scripts.

This was seen as even more relevant, as the entire technology stack, including the basic standards for protocols, interfaces, formats

etc., is biased towards English. When the developers only speak a single (Anglophone) language, their mental model of knowledge is often unconsciously coupled with their language culture.² It was concluded that having multilingual eyes in the development (or design) process helps the decoupling and enables more truly universal knowledge representation. Knowing that most people authoring the standards, writing the implementations, maintaining the infrastructures, or feeding the algorithms embody multiple multilingual personas, might enable changes towards a more diverse and inclusive digital environment.

5 User Profiles for Representation and Fostering Inclusivity in DH Research

Enhancing the visibility of these scholarly experiences is crucial to set the foundation of practical solutions. Venues like our workshop at DH2023 helped us understand the field better and encouraged us to identify common problems while shedding light on hitherto marginalized factors by creating an action space for more minds to come together.

In our communities we are all in silos. The user personas that we created show us that across communities similar issues can arise, issues that many scholars can relate to, revealing the ubiquitousness of problems that may have previously been considered to be “edge cases”. In this vein, such user personas can also create empathy for different kinds of people, bringing the lived experience of individuals of diverse backgrounds closer to the larger DH community and relevant stakeholders in a broader sense. Our data-driven user profiles can contribute to the amelioration of the situation in a DH context by highlighting the realities and challenges of multilingual users. In turn, more researchers might find themselves in the company of others, having tried to use different digital environments.

The six user profiles, which served as a starting point for the workshop, showcase data-driven, but fictional, individuals with diverse training and professional background, but all of them can be characterized as multilingual actors in DH. We have argued elsewhere for the need to differentiate between official multilingualisms and scholarly multilinguality (see Horváth et al. 2024a; 2024b) and understood the latter as a means to highlight the different contexts in and degrees to which scholars use various languages (such as a native speaker of German specializing in Chinese studies who works in Southeast Asia and predominantly publishes in English).

² Examples of publications on these topics include Fiormente 2021; Phillipson 1996; 2012; 2013.

To take this a step further, our workshop has shown the importance of also considering the extent to which different languages are used by DH practitioners for specific tasks. For example, further investigating how important each research language an individual uses is in the context of digital projects and scholarly communication, how certain languages appear in a user's professional and personal life, and how these spheres can become blurred in certain forms of research can shed light on the nuances of multilingual realities.

At the same time, it is important to realize that these user profiles also constitute a moving target. Generating user stories and showcasing scholarly realities, these persona profiles narrativize multilinguality. However, one point that emerged in our workshop discussion was related to the inherent change that these stories entail, showcasing how the careers of these users evolve and shift in light of and as a result of the situation of the current DH landscape. The profiles are not static characters, as it was highlighted during the workshop, therefore their authenticity could be further enhanced through continuous updates about their future career stages – this could enrich the ephemeral, snapshot-like state of their current, momentary position and tell the stories as multilingual journeys.

6 A Wider Perspective on Literacies

During our pre-conference workshop and writing spring at the 2023 ADHO conference, as we previously mentioned, we prepared a reading and circulated a set of published user profiles. We also had a set of questions for a guided discussion of these issues. One of the ways in which the participants steered the discussion was a pleasantly surprising new one for us, that of the thinking about multilingual users in DH and the question of evolving literacies. It was pointed out that the user profile dataset, while it encompassed researchers who work in different languages within specific academic or alt+ac positions, not enough emphasis had been placed on the degree to which they use their different languages for specific purposes. For instance, it would be insightful to compare the significance of different languages in digital research and scholarly communication, as well as the blurred lines between personal and professional language engagement within certain research contexts.

Participants felt a need to recognize that literacy is a process, rather than a static achievement assumed by the published dataset of personas. Languages are deeply intertwined with the history and culture of their communities, emphasizing the importance of language learning for understanding different perspectives. By integrating acquiring languages and literacies into multilingual profiles, perhaps including even the learning and acquisition of technical languages,

and recognizing that multilingual users' experiences with language resources evolve, we can paint a more inclusive picture of literacies as in a constant state of development. Ideally, a broader range of profiles, including non-academic or non-professional ones, should be incorporated to ensure the accessibility and usability of digital textual practices across diverse user groups.

Furthermore, the discussion delved into the dual nature of language usage, distinguishing between working *in* and working *on* specific languages and scripts. This duality underscores both the potential and the challenge for enhancing linguistic inclusivity in DH. Encouraging scholars to publish and produce research in languages other than English, supported by DH infrastructures, particularly for non-Latin scripts and marginalized languages. At the time of writing this article, such practices are still relegated to the margins of DH. Some participants even raised the question whether non-Latin scripts should have dedicated infrastructures. In both cases, the field's accessibility could extend beyond the Anglosphere, although the global interconnectedness of the field would find different expression depending on the vision that would be chosen.

Languages are deeply intertwined with the history and culture of their communities, emphasizing the importance of language learning for understanding different perspectives. By integrating language learning into multilingual profiles, we can highlight both professionals and students who choose to engage with languages beyond their immediate expertise, or who are required to do so in order to gain access to research conversations centred in specific geographic or linguistic communities. By centring the question of literacies in the multilingual DH user profiles, it may even be possible to encourage more monolingually leaning practitioners to do so.

7 Moving Forward

Handling the role and significance of language in digital scholarship constitutes an ongoing challenge, particularly in the context of non-Latin scripts. A growing number of initiatives, however, have been engaged in raising awareness of the relevant difficulties and language-based inequities and/or aim to integrate linguistic perspectives into their activities. In the final section of our paper, we will briefly provide an overview of some recent examples of different natures as case studies to offer practical inspirations moving forward.

Besides the above-mentioned, and much-needed, edited volume, *Global Language Justice*, which provides an interdisciplinary meditation on the realities and broader consequences of language loss, and another recently published volume, *Multilingual Digital Humanities* (edited by Lorella Viola and Paul Spence) with a more specific

focus on multilingual aspects related to technology, concerted efforts have recently started to take shape in the form of collaborative research and digital tool development projects, working groups and special interest groups, as well as on an infrastructural level. The year-long initiative, *New Languages for NLP: Building Linguistic Diversity in the Digital Humanities*³ (funded by the National Endowment for the Humanities), aimed to mitigate the limited scope of languages and scripts currently available in SpaCy by supporting the creation of NLP models for eleven languages ranging from Yiddish to Kanbun, Classical Arabic, Old Chinese, Yoruba, Kannada, Ottoman Turkish, Nineteenth-Century Russian, Tigrinya, Efik, and Quechua. Between 2021 and 2022, the Institute contributors, including junior and more senior researchers, graduate students, and research software engineers from around the world, have managed to make varying, but unequivocally significant, progress towards diversifying the landscape of Natural Language Processing.

In our DH2023 workshop, the importance of and need for language sensitivity on an organizational and infrastructural level also emerged as a key motif. Existing research infrastructures and DH organizations handle the question of language in multiple ways, depending on their area of focus. Without attempting to provide an exhaustive overview, here we will briefly discuss numerous examples to illustrate how communities of varying size and geographical embeddedness deal with multilingual perspectives.

First of all, on an infrastructural level, two of the key research infrastructures in the European area, DARIAH (Digital Research Infrastructure for the Arts and Humanities) and OPERAS (Open Scholarly Communication in the European Research Area for Social Sciences and Humanities) have both recognized the need to include multilingual perspectives in their respective initiatives. The DARIAH-sustained metablog, OpenMethods, for example, concentrates on making digital tools and methods more visible in a creative and openly accessible format by sharing relevant published contents (which includes the appropriate documentation of a digital tool or method) with critical introductions written by the editors or external (volunteer) contributors. While the introductions have been consistently produced in English as lingua franca, the platform showcases projects in a variety of languages, thereby making an effort to be linguistically inclusive.

In addition, early 2023 brought the establishment of the DARIAH working group dedicated to Multilingual DH by co-founders Author and Maroussia Bednarkiewicz. Through frequent conferences, workshops, and further public presentations, regular meetings, and the

3 See <https://newnlp.princeton.edu/> and a related book chapter by the organizers in Tasovac et al. 2024.

organization of trainings, the working group aims to strengthen the representation and visibility of the challenges and realities of under-resourced languages and scripts in DH and digital research infrastructures and to enhance digitally-enabled scholarship in such languages thus fostering the creation of a linguistically more inclusive DH environment. On the other hand, multilingual perspectives not only appear in the case of other working groups as well under the auspices of DARIAH and provide common ground for intersectional collaborations. For instance, together with the Research Data Management working group, the Multilingual DH working group has recently won a DARIAH grant to organize a multi-day workshop on “Creating, Managing, and Archiving Textual Corpora in Under-Resourced Languages” scheduled for late 2024 in Hamburg to collaboratively produce a practical workflow to help scholars handle their sources as basis for further analysis. Meanwhile, multilingual aspects have found their way to the activities of the Bibliodata working group, the members of which has recently published on open bibliographical workflows and the challenges of multilinguality (Malínek et al. 2024), as well as the Ethics and Legality in the Digital Arts and Humanities (ELDAH) working group, which has been working on a multilingual consent form wizard. These initiatives offer multiple paths, directions, and trajectories to consider to see how language can appear in the context of digital scholarship while providing possibilities for future collaborations.

The OPERAS research infrastructure also aims to integrate the question of language into its projects, particularly through its Multilingualism Special Interest Group. OPERAS predominantly stands for fostering open scholarly communication in the humanities and social sciences, which explains the somewhat different manifestation of multilingualism here compared to the above-mentioned examples in the context of DARIAH. In the realm of scholarly publishing, the Multilingualism SIG focuses on remedying the dominance of English by supporting “researchers and institutions who want to continue publishing in their own language and to develop transnational scientific cooperation” (Multilingualism). An important highlight of their activities includes the development of a multilingual discovery tool and a pilot Translation Service, while also collaborating with external stakeholders to help “address language biases in assessment and supporting language skills of non-fluent speakers in concrete situations” (Multilingualism).

On the level of DH organizations, the Alliance of Digital Humanities Organizations (ADHO) assembles and connects a variety of local

or regional DH communities.⁴ The localized focus of these organizations seems to also imply language diversity or at least a focus on certain languages (often beyond English), depending on the dominant languages utilized in the respective areas and regions. However, the case of ADHO constituent organizations also showcases the importance of distinguishing between multilingualism and scholarly multilinguality. On the one hand, the majority of these organizations are connected to regions where English is not the official language, which can be considered a fundamental source of support towards ensuring the presence of certain languages that local scholars work on and/or in. On the other hand, the dominance and prioritization of specific language(s), for example at the annual conferences of these organizations, does not necessarily entail the equal representation of scholarly multilinguality, particularly with regard to the case of scholars working in a certain language environment while using sources in one or more other languages or scripts.

Here, a workshop on “Towards Multilingualism In Digital Humanities: Achievements, Failures And Good Practices In DH Projects With Non-latin Scripts”⁵ at the ADHO DH 2019 conference in Utrecht acted as a catalyst for strengthening the international collaboration and networking of a globally dispersed multilingual DH community. It resulted in a website and mailing list (<https://multilingualdh.org/en/>), as well as the idea to propose a special interest group (SIG) on multilingual DH within the ADHO association, which was finally enacted at the ADHO DH conference 2023 in Graz.

Furthermore, on a national level, working groups on multilingual DH have been launched such as in the German-speaking DHd association or as a Multilingual DH community interest group within the UK-Ireland DH association. The multilingual DH group in the DHd association addresses not only the hardships of multilingual DH practitioners who are constantly confronted with “incompatibilities or simply a lack of support and functionality for certain written languages” but also the “marginalization of multilingual DH both in the Digital

⁴ The ADHO constituent organizations as of March 2024 are: Australasian Association for Digital Humanities, the Association for Computers and the Humanities, centerNet, Canadian Society for Digital Humanities/Société canadienne des humanités numériques, Digital Humanities Alliance for Research and Teaching Innovations, Digital Humanities Association of Southern Africa, Digital Humanities im deutschsprachigen Raum (DHd), European Association for Digital Humanities, Humanistica: L'association francophone des humanités numériques/digitales, Japanese Association for Digital Humanities, Korean Association for Digital Humanities, Red de Humanidades Digitales, and the Taiwanese Association for Digital Humanities. For more information, see <https://adho.org/membership-individual/>.

⁵ See <https://multilingualdh.org/en/dh2019/>, an expanded version of this text has been published as Grallert et al. 2023a. On the topic of multilingualism in the DHd association see also Grallert et al. 2023b; 2023c.

Humanities community and in the ‘conservative sciences’”. Furthermore they connect their working group initiative to the criticism of a new “colonialism” in the digital transformation of scholarship where due to its economical/financial superiority,

the global North [...] shapes and determines the field of discourse and opportunities. This ‘colonialism’ and the marginalization of the digital science practice of the global South and the one-sided focus on topics of the global North represent a hierarchization of scientific practice that must be addressed, discussed, and practically challenged.⁶

The UK-Ireland Multilingual DH community interest group aims at fostering and expanding “multilingual digital research and practices within the UK and Ireland’s community of Digital Humanities researchers and related practitioners”, “more explicitly address[ing] the anglophone assumptions that dominate our digital research and practices” and to collaborate “with international initiatives to foster more language-inclusive research and digital infrastructures, and to raise awareness of the ways language intersects with other forms of exclusion in our field and in wider society”.⁷

Other community building efforts – here supported by a university library – include the establishment of a Multilingual DH Lab at the new Ada Lovelace Center for Digital Humanities at Freie Universität Berlin. While taking up on the idea to provide “a space for research on and experimentation with the use of non-Latin scripts in digital environments” the organizers of the lab also point to the aspect of digital monolingualism as “a barrier to global scholarly exchange in general (e.g. lack of visibility/accessibility of and access to research or sources in non-Anglophone languages and non-Latin scripts)”.⁸

This short overview of recent initiatives illustrates some of the ways digital humanists in the Global North have begun to foreground multilingualism (and multilinguality) in a variety of different domains such as NLP, scholarly publishing and translation, but it is important to underscore that they are thus far mostly accessible to researchers located in these countries. They also show how digital humanists have begun to insist on the importance of the topic for its capacity to “[disrupt] digital monolingualism” in the profession, not only because it concerns themselves in their daily research practices but also because of how it intersects with the general transformation

⁶ See DHd MLDH WG <https://m-l-d-h.github.io/DHd-AG/> (our translation).

⁷ See <https://digitalhumanities-uk-ie.org/community-interest-groups/multilingual-dh/>.

⁸ See <https://www.ada.fu-berlin.de/en/ada-labs/multilingual-dh/index.html>.

of scholarship in the humanities and social sciences within the digital space. From the perspective of a globally connected culture of thinking, the digital has introduced more barriers across languages than it has removed. Furthermore, the initiatives, while starting points, link practices in Digital Humanities and multilingualism to large-scale projects predominantly in the Global North. More dialogue with the societies outside of the Global North needs to take place. More consideration needs to be given from an equity perspective as well to ensure that research done there can also benefit from the advocacy for multilingual perspectives (Ghorbaninejad, Gibson, Wrisley 2023). Advocating for the place of multilingual perspectives requires us first to establish their visibility in multiple aspects of our professional life: conferences, transnational organizations, peer review and the evaluation and funding of research. While these initiatives are essential for laying the foundation for a multilingual future in DH, they are first steps.⁹

8 Conclusion

In recent years, much has been made of decolonizing cultural practices, particularly in the domains of educational curricula and GLAM collections. These debates bring to the fore not only questions of representation and inclusion, but also the inherent privilege of European and Western knowledge systems and the power dynamics of education and cultural institutions. The realm of digital infrastructure and tools has absorbed and reproduced much of this power and privilege. When users engaging with multilingual resources – be it various data forms in distinct languages or textual data mixing different languages – approach those responsible for developing, maintaining, or adapting scholarly infrastructure, a common refrain is heard: it is not feasible to accommodate every unique user need or use case. This response reflects institutional power to balance the capabilities of infrastructure and larger institutional priorities in technology investment and management.

In this article, using a range of methods from participative persona creation process analysis, collective ideation and co-writing sprint in a workshop setting, as well as community-driven data collection and comparative analysis, we hope to have made a case for seeing multilinguality as a strength, rather than a weakness. It is a subject that should be taken on at levels higher than the individual institution.

⁹ For a first attempt at sketching concrete steps in the many different communities of the larger knowledge ecosystem of Digital Humanities, see our article: “Who are the Users in Multilingual Digital Humanities?”.

We have also made the case for infrastructure as a site of decolonial reflection and action for Digital Humanities. As we suggested in the previous section, there is space for community engagement and ethical stewardship in the domain of multilingual knowledge practices. In the limited words available to us here we can only hope to raise awareness of the question and to call actors within the community to imagine change and implementation of frameworks encouraging greater linguistic diversity within our academic communities and the larger societal spheres that they represent. Ultimately, our call is for a change in mentality, away from infrastructure as a set of constraining functions which exclude some members of our community, to one in which we constantly keep in mind what multilingual functionality on the web can be and how (knowledge) infrastructure developers and managers can contribute to decolonizing digital spaces and amplifying underrepresented voices and cultures, improving the quality of infrastructure but also empowering individuals to engage with and celebrate linguistic diversity in meaningful ways.

Bibliography

- Alliance of Digital Humanities Associations (s.d.). “Membership”.
<https://adho.org/membership-individual>
- Digitalhumanities-Uk-Ie.org (s.d.). “Multilingual DH in the UK and Ireland | UK-Ireland Digital Humanities Association”.
<https://digitalhumanities-uk-ie.org/community-interest-groups/multilingual-dh>
- Fiormonte, D. (2021). “Taxation against Overrepresentation? The Consequences of Monolingualism for Digital Humanities”. Kim, D.; Koh, A. (eds), *Alternative Historiographies of the Digital Humanities*. Santa Barbara: Punctum Books, 333-76.
- Ghorbaninejad, M.; Gibson, N.P.; Wrisley, D.J. (2023). “Right-to-Left (RTL) Text: Digital Humanists plus Half a Billion Users”. Gold, M.K.; Klein, L. (eds), *Debates in the Digital Humanities 2023*. Minneapolis: University of Minnesota Press, 47-73.
- Glasberg, E. (2024). “What Does Global Language Loss and Justice Look Like?”.
<https://news.columbia.edu/news/what-does-global-language-loss-and-justice-look>
- Golumbia, D. (2013). “Postcolonial Studies, Digital Humanities, and the Politics of Language”. *Uncomputing*, May 31.
<https://web.archive.org/web/20160623210924/http://www.uncomputing.org/?p=241>
- Grallert, T. (2022). “Open Arabic Periodical Editions: A Framework for Bootstrapped Scholarly Editions Outside the Global North”. *Digital Humanities Quarterly*, 16(2).
<https://web.archive.org/web/20231018192515/http://www.digitalhumanities.org/dhq/vol/16/2/000593/000593.html>
- Grallert, T. et al. (2023a). “Die AG Multilingual DH im DHD stellt sich vor”. *Zenodo*, March 8.
<https://doi.org/10.5281/zenodo.7744530>
- Grallert, T. et al. (2023b). “Dealing with Multilingualism in the DACH and DHD Association”. *Zenodo*, May 24.
<https://doi.org/10.5281/zenodo.7965774>
- Grallert, T. et al. (2023c). “Umgang mit Multilingualität im DACH und DHD Verband”. *Zenodo*, May 22.
<https://doi.org/10.5281/zenodo.7957187>
- Horváth, A. et al. (2023). “Six User Personas for the Multilingual DH Community (v1.0)”. Dataset related to “Towards Multilingually Enabled Digital Knowledge Infrastructures: Discussing the Role of the Digital Humanist”, discussion at DH Unbound 2022 [virtual conference] (May 17-19). *Zenodo*, April 9.
<http://doi.org/10.5281/zenodo.7811800>
- Horváth, A. et al. (2024a). “Towards Multilingually-Enabled Digital Knowledge Infrastructures: A Qualitative Survey Analysis”. Viola, L.; Spence, P. (eds), *Multilingual Digital Humanities*. London: Routledge, 197-212.
- Horváth, A. et al. (2024b). “Centring Multilingual Users: Thinking through UX Personas in the DH Community”. *Digital Studies/Le champ numérique*, 13(3), 1-30.
<https://doi.org/10.16995/dscn.9608>
- Horváth, A.; Wagner, C.; Wrisley, D.J. (2024). “Who Are the Users in Multilingual Digital Humanities?”. *Digital Scholarship in the Humanities*, under review.
- Horváth, A. (2021). “Enhancing Language Inclusivity in Digital Humanities: Towards Sensitivity and Multilingualism”. *Modern Languages Open*, 1(26), 26.
<http://doi.org/10.3828/mlo.v0i0.382>
- Horváth, A. (2022). “Digital Brush Talk: Challenges and Potential Connections in East Asian Digital Research”. Fiormonte, D.; Chaudhuri, S.; Ricaurte, P. (eds), *Global Debates in the Digital Humanities*. Minneapolis: University of Minnesota Press, 127-40.

- Kirmizialtin, S.; Wrisley, D.J. (2022). "Automated Transcription of Non-Latin Script Periodicals: A Case Study in the Ottoman Turkish Print Archive". *Digital Humanities Quarterly*, 16(2).
<https://web.archive.org/web/20230928110700/http://www.digitalhumanities.org/dhq/vol/16/2/000577/000577.html>
- Liu, L.H.; Rao, A.; Silverman, C.A. (eds) (2024). *Global Language Justice*. New York: Columbia University Press.
- Malínek, V. et al. (2024). "Open Bibliographical Data Workflows and the Multilinguality Challenge". *Journal of Open Humanities Data*, 10, 27, 1-14.
<http://doi.org/10.5334/johd.190>
- M-l-d-h.github.io (s.d.). "DHd AG Multilingual DH".
<https://m-l-d-h.github.io/DHd-AG>
- Multilingualdh.org (s.d.). "DH 2019 Workshop I Multilingual DH".
<https://multilingualdh.org/en/dh2019>
- Multilingualism (2022). "Multilingualism – OPERAS".
<https://operas-eu.org/special-interest-groups/multilingualism-2/>
- Phillipson, R. (1996). "Linguistic Imperialism: African Perspectives". *ELT Journal*, 50(2), 160-7.
<https://doi.org/10.1093/elt/50.2.160>
- Phillipson, R. (2012). "Linguistic Imperialism". Chapelle, C. (ed.), *The Encyclopedia of Applied Linguistics*. Malden, MA: Wiley-Blackwell, 1-7.
- Phillipson, R. (2013). *Linguistic Imperialism Continued*. Abingdon: Routledge.
- Princeton University (s.d.). "New Languages for NLP".
<https://newnlp.princeton.edu>
- Spence, P. (2021). *Disrupting Digital Monolingualism: A Report on Multilingualism in Digital Theory and Practice*. London: Language Acts and Worldmaking Project.
<http://doi.org/10.5281/zenodo.5743283>
- Tasovac, T. et al. (2024). "Bridging the Gap Between Digital Humanities and Natural Language Processing: A Pedagogical Imperative for Humanistic NLP". Viola, L.; Spence, P. (eds), *Multilingual Digital Humanities*. Abingdon: Routledge, 114-26.
- Wrisley, D.J. (2019). "Enacting Open Scholarship in Transnational Contexts". *POP! Public. Open. Participatory*, 1.
<https://web.archive.org/web/20231122183458/https://popjournal.ca/issue01/Wrisley>