

# JASIST special issue on digital humanities (DH)

## INTRODUCTION

More than 15 years ago, *A Companion to Digital Humanities* marked out the area of digital humanities (DH) “as a discipline in its own right” (Schreibman et al., 2004, p. xxiii). In the years that followed, there is ample evidence that the DH domain, formed by the intersection of humanities disciplines and digital information technology, has undergone remarkable expansion. This growth is reflected in *A New Companion to Digital Humanities* (Schreibman et al., 2016). The extensively revised contents of the second edition were contributed by a global team of authors who are pioneers of innovative research in the field. Over this formative period, DH has become a widely recognized, impactful mode of scholarship and an institutional unit for collaborative, transdisciplinary, and computationally engaged research, teaching, and publication (Burdick et al., 2012; Svensson, 2010; Van Ruyskensvelde, 2014).

The field of DH has advanced tremendously over the last decade and continues to expand. Meanwhile, competing definitions and approaches of DH scholars continue to spark debate. “Complexity” was a theme of the DH2019 international conference, as it demonstrates the multifaceted connections within DH scholarship today (Alliance of Digital Humanities Organizations, 2019). Yet, while it is often assumed that the DH is in flux and not particularly fixed as an institutional or intellectual construct, there are also obviously touchstones within the DH field, most visibly in the relationship between traditional humanities disciplines and technological infrastructures. Thus, it is still meaningful to “bring together the humanistic and the digital through embracing a non-territorial and liminal zone” (Svensson, 2016, p. 477). This is the focus of this *JASIST* special issue, which mirrors the increasing attention on DH worldwide.

The goals for this issue are to concentrate attention on DH concepts, theories, methods, and analyses and to showcase the infrastructures of DH and its standard practices, the methodological and technological advancements in DH research, and critical matters that have shaped the development of DH. It also aims to map out a research agenda that identifies critical points of intersection and gaps in knowledge that require collaboration

between information science and technologies (IST) professionals and DH scholars, as well as the establishment of a common theoretical framework for understanding key issues and research questions.

## IN THIS DIGITAL HUMANITIES SPECIAL ISSUE

This special issue presents works that contribute to an understanding of DH’s multifaceted landscape and the role of IST in humanities research in the 21st century. The special issue contains 12 articles, which span various disciplines, countries, and regions, consonant with themes of collaboration and diversity in the field (Spiro, 2012), and reveal a growing recognition of DH as a global phenomenon with particular manifestations in local settings (Risam, 2016). These contributions are built on previous work exploring conceptual connections between DH and IST (Poole, 2017), as well as the role of information professionals in DH (Clement & Carter, 2017). The editors have organized this issue around three principal themes: (a) Landscapes of DH; (b) infrastructures of DH; and (c) methodological innovations, challenges, and new interests in DH. Below, we briefly introduce each section and summarize how each article contributes to the understanding of scholarship and practice in DH.

### Landscapes of DH

While critical discussions about the scope and extent of the field abound in both the humanities and information science literature, fewer works have presented empirical studies of the content and contours of DH. Among these works are studies of scholars on social media (Grandjean, 2014a, 2014b, 2015; Quan-Haase et al., 2015; Ross et al., 2011), at conferences (Weingart et al., 2020), and in scholarly communications, including coauthorship (Nyhan & Duke-Williams, 2014; Spiro, 2009), citational practices (Gao et al., 2017, 2018; Palermo, 2019; Su et al., 2020; Tang et al., 2017), and the early history of the field (Sula & Hill, 2019). These works have documented both

collaborative and interdisciplinary features of the field, while also charting areas of topical significance. Building on these previous works, the articles in this section take novel approaches to the landscape of the field and, in some cases, assemble new datasets to do so.

In “Digital humanities—A discipline in its own right? An analysis of the role and position of digital humanities in the academic landscape,” Jan Luhmann and Manual Burghardt take up the question of whether DH is a discipline—on par with others in the academic landscape—or whether it occupies some other position, such as an interdisciplinary field. Using new hierarchical clustering techniques in text analysis, they explore three decades of English-language DH journal articles and compare them to a large corpus of research articles spanning 15 disciplines. Luhmann and Burghardt find that DH is both a discipline and a bridge between other disciplines—notably, computational linguistics and information science, on the one hand, and humanistic disciplines, on the other. They conclude by suggesting further areas of study, including analysis of conferences and citations, as well as diachronic representations of the ever-developing landscape of DH.

Rongqian Ma and Kai Li also consider questions of disciplinarity in “Digital humanities as a cross-disciplinary battleground: An examination of inscriptions in journal publications,” employing here a visual perspective rather than a textual one. Their study compares inscriptions (nonverbal artifacts and media, such as graphs, diagrams, and tables) in DH journal articles to science, technology, engineering, mathematics (STEM), and humanities domains. In doing so, the authors develop a picture of the distinct epistemic community of digital humanists. As one of the first comprehensive empirical analyses of inscription use in DH, the study bridges an important gap between the increasing interests in data visualization and data-driven research among DH scholars and other empirical studies of the field.

The final article in this section, “Digital Humanities in the iSchool,” explores the landscape of DH within a particular location: Information schools. John Walsh and 12 coauthors provide a snapshot of DH programs and initiatives in iSchools worldwide, including the perspectives of programs in Africa, Asia-Pacific, Europe, and North America. Topics covered in this review include educational models used for DH training in iSchools, analysis of DH courses and curricula as reflected in course descriptions and syllabi, cooperative endeavors between iSchools and other disciplines around research projects and educational programs, and DH careers, explored through analysis of relevant job advertisements. The authors comment upon their experiences managing DH programs within iSchools and provide guidance for those schools considering new programs in the area. Overall, the article helps

identify existing gaps in current educational offerings and points to potentials for future development.

## Infrastructures of DH

Scholars have long recognized infrastructure—the organizational and structural aspects of the field—as crucial to the support of DH. Projects are often seen as the “primary unit” of DH (Burdick et al., 2012), serving as the driving mechanism for organizing ideas, people, and resources in the field. Early commentators noted the role of infrastructures in supporting DH production and publishing (Pitti, 2004), while more recent work has focused on data curation and reuse: “capturing and preserving not only the data itself, but information about the methods by which it was produced” (Flanders & Muñoz, 2012). Central to questions of data are availability, interoperability, and workflows for processing data, as well as particular issues related to the cultural heritage datasets commonly found within DH. The articles in this section explore new methods and thinking around these datasets, especially through the lens of Linked Open Data (LOD), a long-standing interest within information science and Semantic Web research (Heath & Bizer, 2011). Like other works in this special issue, these articles point toward interdisciplinary collaborations as key to the success and sustainability of DH infrastructures.

Joyce Siqueira and Dalton Martins’ “Workflow models for aggregating cultural heritage data on the web: A systematic literature review” points out that integrating cultural data is not a trivial task, even though many cultural institutions have developed search interfaces that integrate their digital objects and facilitate data retrieval for lay users. This article presents a systematic literature review of data aggregation workflows, organized around nine stages: Harvesting, ingestion, mapping, indexing, storing, monitoring, enriching, displaying, and publishing LOD. The authors discuss various technologies and solutions associated with each stage—most of them are semi-automatic—but note that there is still no consensus regarding the stages, their nomenclatures, and technologies across the field, an area for further development.

In “Integrated interdisciplinary workflows for research on historical newspapers: Perspectives from humanities scholars, computer scientists, and librarians,” Sarah Oberbichler and seven coauthors consider the interdisciplinary opportunities and challenges of working with digital materials, namely, historical newspapers. The article provides insight into digital tools, methods, and hermeneutics in action, developing an integrated digital hermeneutics workflow that combines research approaches from computer science, the humanities, and

libraries. The authors contend that interdisciplinary collaborations can benefit from better understandings of the workflows and traditions of each discipline involved, but these initiatives must also find integrated approaches to exploit the full potential of digitized sources. Interdisciplinary research must build something between the disciplines involved while still respecting and understanding each other's expertise and expectations.

"Harmonizing and publishing heterogeneous premodern manuscript metadata as Linked Open Data," authored by Mikko Koho and 15 researchers from Finland, United States, United Kingdom, and France, presents work on the "Mapping Manuscript Migrations" project which brought together multiple metadata models and developed a shared approach to describe early manuscript material and their provenance metadata currently siloed in incompatible, heterogeneous databases. The article presents the major products of the project: A unified data model, a repeatable data transformation pipeline, a LOD knowledge graph, and a Semantic Web portal offering LOD service for DH research and application development. The authors emphasize the importance of engaging humanities researchers in the development of digital platforms and tools, in addition to considering how such tools might empower the researchers to conduct more complex analyses.

## **Methodological innovations, challenges, and new interests in DH**

Much of DH scholarship reflects the application of new information technology tools, methods, and epistemological frameworks to the humanities. In this section (and elsewhere in this special issue), we present papers that report new research and development on these applications. The projects described here engage with various domains, languages, geospatial regions, resource types, and facilities, yet they all share common goals: Widening the scope of the humanities, opening access to sources, and presenting new forms of scholarly activity (Thomas III, 2016). In addition to presenting methodological innovations, the authors discuss the challenges faced, gaps found, and new questions identified through their research. They also reflect how humanities scholarship, traditions, and practices have shaped or have the potential to shape digital tools, platforms, and innovations.

Starting with the question, "A bridge too far for artificial intelligence?," Álvaro Pérez Pozo and five coauthors report their findings on "Automatic classification of stanzas in Spanish poetry." The authors point out that advances in artificial intelligence and machine learning have been primarily focused on prose text, leaving mostly

aside figurative or poetic expressions of language due to their rich semantics and syntactic complexity. Although machine learning has achieved such tasks as metrical annotation and syllabification, there has been little work on the classification of stanzas and the inner structures of verses on which poems are built. The authors find that stanza classification remains hard for computer systems, both those using classical machine learning and those using statistical language models. At present, neither can compete with traditional computational paradigms based on the knowledge of experts.

Machine learning algorithms, known as deep neural networks, have dominated the field of automatic text analysis and natural language processing in recent years. In "Text analysis using deep neural networks in digital humanities and information science," Omri Suissa, Avshalom Elmalech, and Maayan Zhitomirsky-Geffet present the practical challenges that DH and library and information science (LIS) experts may encounter when applying deep neural network models in their research. To assist DH/LIS experts in choosing appropriate approaches, the authors develop a decision model based on two strategies: The data availability strategy and the domain adaptation strategy. They argue that DH/LIS researchers can no longer see natural language processing and machine learning researchers as their "toolmakers," and it is time to introduce courses that foster understanding of the inner workings of machine learning, deep learning, and related algorithms into the DH and LIS curricula. DH/LIS experts need not become computer science experts; they simply need to be able to comprehend and adapt deep learning algorithms for their own needs.

Extending work on natural language processing, "Using parsed and annotated corpora to analyze parliamentarians' talk in Finland" highlights Finnish parliamentary records and interviews with former parliamentarians, which have been annotated with metadata of talk structure and named persons. Using natural language processing, Mykola Andrushchenko and 10 coauthors investigate how politicians talk about power, how ideological terms are used in political speech, and how to identify narratives in the corpus. The authors argue that DH investigations of such datasets cannot rely exclusively on computational humanities modeling but must accommodate a range of perspectives starting with simple searches and quantitative exploration and ending with modeling. The article underscores the need for more thorough discussions on how IST tools may alter the research questions posed by humanists.

Open research data repositories have been hailed as cornerstones promoting collaboration, interoperability, and large-scale sharing and reuse. Yet humanities scholars are often reluctant to use them, leading many scholars to ask why. Karin Hansson and Anna Dahlgren

investigate this gap between the possibilities and constraints of open research data in the humanities through a case study of five image data repositories. Their paper, “Open research data repositories: Practices, norms, and metadata for sharing images,” presents an analysis of what these platforms share in common, as well as the norms and practices embedded in this datafication of academic practice. It reveals an ontological gap between the metadata needs and practices of humanities researchers, as well as the limitations of copyright. For those in the humanities, these data-sharing tools are not yet sophisticated enough when it comes to sharing and reusing images.

In “Giving shape to large digital libraries through exploratory data analysis,” Peter Organisciak, Benjamin Schmidt, and Stephen Downie provide an additional perspective on data-sharing facilities, this time from the standpoint of data providers, specifically large, multi-institutional digital libraries. These cooperative enterprises allow scholars to study textual data across centuries of books, but the potential of such analysis can be constrained by burdens of scale, where statistical inference is technically complex and can be hampered by limitations on data access and lack of flexibility in the tools provided by the systems. The article provides detailed technical information about the HathiTrust+Bookworm tool, which allows multifaceted exploration of the multimillion work HathiTrust Digital Library. The authors reflect on the role of hypothesis-building as part of scholarly workflows, as well as the broader role of exploratory tools in approaching large datasets.

The final article in this section and issue, “Revisiting the digital humanities through the lens of Indigenous studies—or how to question the cultural blindness of our technologies and practices” interrogates technoscientific traditions from the perspective of Indigenous studies. Reflecting on work with the Samí people of Norway, Sweden, Finland, and northwestern Russia, Copp  lie Cocq proposes several interventions to research practices, specifically around data harvesting, categorization, and sharing. This article—one of the first to bring Indigenous studies to the journal—calls into question common assumptions around values, perspectives, and ethics that characterize fieldwork and research subjects. As such, it contributes to calls for critical approaches from and within DH and provides an occasion to revisit our digital research practices more broadly.


## SUMMARY

In conclusion, the dominant narrative of the emergent DH field holds that information technologies have


transformed humanities through the application of new tools, methods, and epistemological frameworks. Less frequently does one see the argument that the humanities disciplines have shaped the products and applications of information technology. The papers in this special issue provide research findings and discussions that represent both aspects of this phenomenon of co-shaping and co-constitution within the field. These works reflect how digital tools and platforms have shaped or have the potential to influence humanities inquiry; how humanities scholars work, conduct research, and disseminate their work and the associated implications; and conversely, how humanities scholarship, traditions, and practices have shaped or have the potential to shape digital tools, platforms, and innovations. The editors of this issue hope that readers will benefit from these rich perspectives on the ecology of DH.

## ACKNOWLEDGMENTS

This special issue is a result of work by five guest coeditors. We thank the anonymous reviewers, who are experts in some combination of information science and technologies and humanities. We thank Javed Mostafa, former JASIST editor-in-chief, for inviting this special issue.

Marcia Lei Zeng, Guest Editor<sup>1</sup> 

Chris Alen Sula, Guest Editor<sup>2</sup>

Karen F. Gracy, Guest Editor<sup>1</sup> 

Eero Hyv  nen, Guest Editor<sup>3</sup>

V  nia Mara Alves Lima, Guest Editor<sup>4</sup>

<sup>1</sup>Kent State University, Kent, Ohio, USA

<sup>2</sup>Pratt Institute, New York, New York, USA

<sup>3</sup>Aalto University and University of Helsinki (HELDIG), Espoo, Finland

<sup>4</sup>University of Sao Paulo, Sao Paulo, Brazil

## ORCID

Marcia Lei Zeng  <https://orcid.org/0000-0003-0151-5156>

Karen F. Gracy  <https://orcid.org/0000-0002-8399-4746>

## REFERENCES

- Alliance of Digital Humanities Organizations. (2019). *Digital humanities 2019: “Complexities”*. Call for papers. <http://staticweb.hum.uu.nl/dh2019/dh2019.adho.org/call-for-papers/cfp-english/index.html>
- Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., & Schnapp, J. (2012). *Digital Humanities*. MIT Press.
- Clement, T. E., & Carter, D. (2017). Connecting theory and practice in digital humanities information work. *Journal of the Association for Information Science and Technology*, 68(6), 1385–1396. <https://doi.org/10.1002/asi.23732>
- Flanders, J., & Mu  oz, T. (2012). *An introduction to humanities data curation. DH curation guide: A community resource guide to data curation in the digital humanities*. <http://guide.dhcurator.org/intro>



- Gao, J., Duke-Williams, O., Mahony, S., Bold, M. R., & Nyhan J. (2017). The intellectual structure of digital humanities: An author co-citation analysis. *DH 2017 Conference*, Montréal. <https://dh2017.adho.org/abstracts/083/083.pdf>
- Gao, J., Nyhan, J., Duke-Williams, O., & Mahony, S. (2018). Visualising the digital humanities community: A comparison study between citation network and social network. *DH 2018 Conference*, Mexico City. <https://dh2018.adho.org/en/visualising-the-digital-humanities-community-a-comparison-study-between-citation-network-and-social-network>
- Grandjean, M. (2014a). The digital humanities network on Twitter (#DH2014). *Martin Grandjean*, July 14. <http://www.martingrandjean.ch/dataviz-digital-humanities-twitter-dh2014>
- Grandjean, M. (2014b). The digital humanities network on Twitter: Following or being followed? *Martin Grandjean*, September 8. <http://www.martingrandjean.ch/digital-humanities-network-twitter-following>.
- Grandjean, M. (2015). Digital humanities on Twitter, a small-world? *Martin Grandjean*, July 2. <http://www.martingrandjean.ch/digital-humanities-on-twitter>
- Heath, T., & Bizer, C. (2011). *Linked data: Evolving the web into a global data space*. Morgan & Claypool. <https://doi.org/10.2200/S00334ED1V01Y201102WBE001>
- Nyhan, J., & Duke-Williams, O. (2014). Joint and multi-authored publication patterns in the digital humanities. *Literary and Linguistic Computing*, 29(3), 387–399. <https://doi.org/10.1093/lc/fqu018>
- Palermo, G. (2019). Visualizing citations in *digital humanities Quarterly's* Biblio project. *ACH 2019 Conference*, Pittsburgh.
- Pitti, D. V. (2004). Designing sustainable projects and publications. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. 469–487). Blackwell.
- Poole, A. H. (2017). The conceptual ecology of digital humanities. *Journal of Documentation*, 73(1), 91–122. <https://doi.org/10.1108/JD-05-2016-0065>
- Quan-Haase, A., Martin, K., & McCay-Peet, L. (2015). Networks of digital humanities scholars: The informational and social uses and gratifications of Twitter. *Big Data & Society*, 2(1). <https://doi.org/10.1177/2053951715589417>
- Risam, R. (2016). Navigating the global digital humanities: Insights from Black feminism. In M. K. Gold & L. F. Klein (Eds.), *Debates in the digital humanities 2016* (Ch.29). University of Minnesota Press.
- Ross, C., Terras, M., Warwick, C., & Welsh, A. (2011). Enabled backchannel: Conference Twitter use by digital humanists. *Journal of Documentation*, 67(2), 214–237. <https://doi.org/10.1108/00220411111109449>
- Schreibman, S., Siemens, R., & Unsworth, J. (Eds.). (2004). *A companion to digital humanities*. Blackwell.
- Schreibman, S., Siemens, R., & Unsworth, J. (Eds.). (2016). *A new companion to digital humanities*. John Wiley & Sons.
- Spiro, L. (2009). Collaborative authorship in the humanities. *Digital Scholarship in the Humanities*. <https://digitalscholarship.wordpress.com/2009/04/21/collaborative-authorship-in-the-humanities>.
- Spiro, L. (2012). “This is why we fight”: Defining the values of the digital humanities. In M. K. Gold (Ed.), *Debates in the digital humanities* (Ch.3). University of Minnesota Press.
- Su, F., Zhang, Y., & Immel, Z. (2020). Digital humanities research: Interdisciplinary collaborations, themes and implications to library and information science. *Journal of Documentation*, 77(1), 143–161. <https://doi.org/10.1108/JD-05-2020-0072>
- Sula, C. A., & Hill, H. V. (2019). The early history of digital humanities: An analysis of *Computers and the Humanities* (1966–2004) and *Literary and Linguistic Computing* (1986–2004). *Digital Scholarship in the Humanities*, 34(Suppl. 1), i190–i206. <https://doi.org/10.1093/lc/fqz072>
- Svensson, P. (2010). The landscape of digital humanities. *Digital Humanities Quarterly*, 4(1).
- Svensson, P. (2016). Sorting out the digital humanities. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A new companion to digital humanities* (pp. 476–492). John Wiley & Sons.
- Tang, M. C., Cheng, Y. J., & Chen, K. H. (2017). A longitudinal study of intellectual cohesion in digital humanities using bibliometric analyses. *Scientometrics*, 113(2), 985–1008. <https://doi.org/10.1007/s11192-017-2496-6>
- Thomas, W. G., III (2016). The promise of the digital humanities and the contested nature of digital scholarship. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A new companion to digital humanities* (pp. 524–537). John Wiley & Sons.
- Van Ruyskensvelde, S. (2014). Towards a history of e-education? Exploring the possibilities of digital humanities for the history of education. *Paedagogica Historica*, 50(6), 861–870. <https://doi.org/10.1080/00309230.2014.955511>
- Weingart, S. B., Eichmann-Kalwara, N., Lincoln, M., Chidsey, C., Barlett, H. W., Dombrowski, Q., Jhunhunwala, A., Kwakye, G., Lee, H., & Nazareth, S. (2020). *The index of digital humanities conferences*. Carnegie Mellon University. <https://doi.org/10.34666/k1de-j489>