Research Journal of English Language and Literature (RJELAL)
A Peer Reviewed (Refereed) International Journal
Impact Factor 6.8992 (ICI) http://www.rjelal.com;
Email:editorrjelal@gmail.com; ISSN:2395-2636 (P); 2321-3108(O)

Vol.14.Issue 3. 2026 (July-Sept.)

REVIEW ARTICLE

INTERNATIONAL STANDARD SERIAL NUMBER INDIA
2395-2636 (Print):2321-3108 (online)

From Humanities Computing to Digital Humanities: The Emergence and Development of a Field

Syed Saqib Anwar[1] and Rahul Vipan[2]

[1] Department of English, University of Jammu, Jammu, Jammu and Kashmir, India
Email: saqibanwar15@gmail.com

[2] Department of English, University of Jammu, Jammu, Jammu and Kashmir, India
Email: Rahulmahi754@gmail.com

DOI: 10.33329/rjelal.14.3.198

Article info
Article Received: 16/07/2026
Article Accepted: 18/08/2026
Published online: 23/08/2026

Abstract

The paper outlines the evolution of Digital Humanities from Roberto Busa's pioneering computational work in the 1940s to the current scenario where it occupies the interdisciplinary state of the inclination. It explores the rise of data computing, the evolution of digital standards and institutions, the evolution of computational methods, and the rise of "digital humanities" as a larger discipline. The paper studies how Digital Humanities has emerged as a result of the interplay of technology, institutional support, and evolving conceptions of humanities scholarship.

Keywords: digital humanities, humanities computing, computational methods, digital archives, distant reading.

INTRODUCTION

Digital humanities is an emerging field in the humanities, but its roots go back more than seventy years to an attempt to use a machine to index the complete works of a medieval philosopher. Today, it encompasses text encoding, data visualization, digital archives, and computational text analysis. It was not developed at once. It grew incrementally from a small technical niche known as "humanities computing" to a discipline with journals, conferences, university centers, degree programs, and disciplinary debates. This history is important because the methods, institutions, and debates of the field are inseparable from the technological and institutional context in which they emerged.

The paper outlines that development in a number of stages. These encompass the early days of humanities computing, the establishment of common standards and institutions, the move of computing beyond textual studies, the emergence of the "Digital Humanities" term in the 2000s, along with the rise of large-scale quantitative analysis in the 2010s, and ongoing debates regarding the identity and surrounding issues in the field. These stages demonstrate that Digital

Humanities emerged from technology, institutional support, and evolving conceptions of the potential of humanities scholarship with computers.

Its history is not a simple or linear success story. In nearly every instance, there was a discussion about whether computational approaches were providing any new value to humanities research or merely new technical forms to old questions. This tension between methodological novelty and scepticism about real change continues throughout the field's history. It is one of the reasons why the boundaries of Digital Humanities are still up for discussion.

Origins

The beginning of the history of digital humanities is the Italian Jesuit Father Roberto Busa, who in 1949 asked IBM to help him create a full concordance of the works of Thomas Aquinas and other authors. The result was the *Index Thomisticus*, which later took decades in the making. The project was not completed for print until 1980, and it is often regarded as the symbolic birth of the field. As Hockey (2004) notes in her history of the discipline:

> The origins of humanities computing are quite well known, unlike many other interdisciplinary experiments. In 1949, an Italian Jesuit priest, Father Roberto Busa, started what even today is a monumental task: to make an index verborum of all the words in the works of St Thomas Aquinas and related authors, totalling some 11 million words of medieval Latin. Father Busa thought that perhaps a machine could help him, and, hearing of computers, he went to see Thomas J. Watson at IBM in the United States for help. (p. 3)

The significance of the project was not in its technical size, which was modest by later standards, but in the fact that it demonstrated that computational methods used in scientific and business applications could be applied to literary and textual scholarship. It also demonstrated that a humanist could lead a computational project, not just supply materials to technical specialists.

Busa's project was not the first. Parallel projects appeared during the 1950s and 1960s in areas such as authorship attribution and statistical stylistics. A well-known instance is the comparative word-frequency analysis that was applied to the authorship of the disputed *Federalist Papers*. All these projects had a similar approach to the treatment of texts as countable and comparable data. This method remained important in computational literary studies even after the tools changed.

Consolidation

After Busa's work, humanities computing became more established during the 1960s through the 1980s. The history by Hockey (2004) tells of a shift from early experiments with concordances and word counts to more organized infrastructure. Dedicated journals such as *Computers and the Humanities* and *Literary and Linguistic Computing* appeared. Professional organizations such as the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing also developed. Annual conferences facilitated the exchange of research and methodology among scholars in various locations. As Hockey (2004) puts it:

> The one word that would most likely describe this period would be "consolidation." There was an increased use of methodologies developed in the early period. There were more electronic texts being produced and more projects using the same applications were begun. The knowledge of what was possible had trickled out through normal scholarly channels of communication and more and more people had encountered computers in their daily lives and had

> begun to consider what computers might do for their research and teaching. (p. 7)

These institutions established a stable community that could educate new scholars and pass on technical knowledge. An important development was the establishment of common standards for the representation of texts in digital form. Prior to the development of these standards, individual projects would develop their own encoding systems. This hindered scholars from sharing data or building on the work of others. In the late 1980s, scholars such as Susan Hockey initiated the Text Encoding Initiative, which provided a common vocabulary for scholars to mark up manuscripts, editions, and archives. The international, multidisciplinary Text Encoding Initiative (TEI), founded in 1987, is "undoubtedly the single most important collaborative project in humanities computing to date" (McCarty, 2004, p. 1228). It enabled projects to collaborate and facilitated scholars to build upon digital texts produced at various institutions.

In the early 1990s, the advent of personal computers and early multimedia technologies enabled humanities computing to transcend the realm of plain-text analysis to images, sound, and hypertext. The field also started to view itself as a field with its own methods and questions, not just as computer science techniques applied to humanities materials. In 1992, the Institute for Advanced Technology in the Humanities was established at the University of Virginia, and other centers were founded, providing the field with institutional space.

Funding, Digitisation and Infrastructure Development

Institutional growth alone does not explain how a small practice became a field with dedicated academic departments. Large digitization projects and external funding were also important. They provided scholars with extensive source material in machine-readable format. The National Endowment for the Humanities (NEH) created its Office of Digital Humanities in 2008 in the United States. This provided the field with a specific source of funding. Similar initiatives developed in the United Kingdom through the Joint Information Systems Committee (JISC). Digital archives and tool-building projects were also supported by private foundations like the Andrew W. Mellon Foundation.

Projects for digitization outside the academy were also significant. In 2004, the Google Books project and the subsequent Hathi Trust digital library began to make millions of books searchable. Many materials could also be downloaded in large quantities for texts that are not under copyright. These collections were used for large-scale quantitative text analysis, one of the most visible methods of the Digital Humanities in the 2010s.

The founding generation did not have access to this amount of data. It facilitated the growth of Digital Humanities methods. As technology, funding, and digital collections grew, scholars were able to pose questions that were previously hard or impossible to investigate.

Expansion Beyond Text

The early projects were primarily concerned with texts, such as concordances, stylistic analysis, and textual editing. Computational techniques also began to be used in art history, archaeology, musicology, and history during the 1990s and early 2000s. Art historians and archaeologists were able to study visual materials through computational methods with the help of digital image processing. Historians used geographic information systems to map spatial and demographic data over time. Digital editing projects also evolved from simple transcription to detailed and linked scholarly editions, which could display various manuscript versions of the same work. This growth contributed to the subsequent claim that computational

approaches were no longer confined to texts and required a new, more general term: "Digital Humanities."

Significant efforts to define the field also occurred during this period. In the late 1990s, Willard McCarty's publications attempted to provide a more solid intellectual foundation for humanities computing than merely a technical service, as he believed that computational methods transform scholarship, not just its speed or convenience. As he put it:

> Computing the humanities is not about speeding up conventional scholarship, or making its performance more efficient or accurate, although all those improvements can occur, because what we mean by scholarship itself changes in the process, or as Roberto Busa SJ put it in 1980, "computers are not intended to save time and work, but to perform work of an analytical nature that would not otherwise be possible." (McCarty, 1998)

This idea became a key theme in subsequent debates on Digital Humanities. The computer was not just a quicker way to conduct traditional research. It also enables scholars to pose new questions and to look at materials in new ways.

Humanities Computing to Digital Humanities

The name "Digital Humanities" became widely used in the mid-2000s, especially after the publication of *A Companion to Digital Humanities* in 2004, edited by Susan Schreibman, Ray Siemens, and John Unsworth. The editors contended that the new name demonstrated a broader range. It encompassed digital approaches in a wide range of humanities fields, not just text-based. As they explain:

> This collection represents a pivotal moment in the field of Digital Humanities: for the first time, a broad spectrum of theorists and practitioners… have been gathered together to think about Digital Humanities as a field in its own right, and to consider how it relates to the areas of traditional humanities scholarship. (Schreibman et al., 2004, p. xxiii)

The shift in terminology thus was linked to a shift in the scope of the field. Literature, history, art, cultural studies, archives, and other fields might be considered Digital Humanities.

Kirschenbaum (2010) writes about the change from the point of view of English departments:

> The terminological change from "humanities computing" to "digital humanities" has been attributed to John Unsworth, Susan Schreibman, and Ray Siemens who, as editors of the anthology *A Companion to Digital Humanities* (2004), tried to prevent the field from being viewed as "mere digitization." [20] Consequently, the hybrid term has created an overlap between fields like rhetoric and composition, which use "the methods of contemporary humanities in studying digital objects," [20] and digital humanities, which uses "digital technology in studying traditional humanities objects".[20] The use of computational systems and the study of computational media within the humanities, arts and social sciences more generally has been termed the 'computational turn'. (pp. 56–57)

The new name coincided with a number of developments, including the mass digitization of library collections (Google Books, HathiTrust), growing institutional support (National Endowment for the Humanities' Office of Digital Humanities), and events like THATCamp. These advances united humanists and technologists outside of the traditional academic conference.

This fast growth also prompted scholars to make more direct definitions of the field. *The Digital Humanities Manifesto*, which was circulated and revised by a group of scholars starting in 2008, attempted to describe what was different about Digital Humanities from previous computational scholarship and from traditional humanities. Its authors describe the field in broad terms: "Digital Humanities is not a unified field but an array of convergent practices that explore a universe in which: a print is no longer the exclusive or the normative medium in which knowledge is produced and/or disseminated" (*Digital Humanities Manifesto 2.0*, 2008–2009, p. 2).

These claims contributed to the expansion of the Digital Humanities from a niche discipline to a broader one. In this timeframe, digital humanities centers, masters and doctoral programs, and faculty positions also grew.

The 2010s: Distant Reading and Large-Scale Quantitative Analysis

The 2010s witnessed the development of one of the most controversial approaches in the field: large-scale quantitative analysis of literary corpora, with the availability of large digital collections and institutional funding. This is a method that is closely linked to Franco Moretti and the Stanford Literary Lab, which he co-founded in 2010. Moretti's term for this approach, "distant reading," first appeared in his 2000 essay "Conjectures on World Literature." The method became easier to practice when large machine-readable collections and stronger computing infrastructure became available. Instead of studying only a small number of canonical novels, scholars could study thousands of novels as part of literary history. This shows a larger pattern in the field: new methods sometimes had to wait for the technology and institutions needed to support them.

The popularity of distant reading also changed how people viewed Digital Humanities. The field became more identified with data visualization, network graphs, and statistics, and less with digital editing, archival work, and multimedia scholarship. This shift in public opinion also had an impact on subsequent discussions of the nature and goals of Digital Humanities.

Current Directions

Since these debates became prominent, Digital Humanities has continued to change rather than become a fixed discipline. Machine learning and natural language processing have become increasingly important in computational text analysis. These technologies, however, raise questions about interpretation, bias in training data, and the relationship between statistical patterns and humanistic meaning. One of the prime examples of this development is the integration of gender and intersectional feminist perspectives, which challenge the historical assumption that data, code, and algorithms are inherently neutral or objective agents.

Contemporary Digital Humanities increasingly reject the techno-heroic narrative that data can speak for itself. Instead, current scholarship applies an intersectional framework to expose how data collection and quantitative modeling are inextricably bound to societal power dynamics. This critical direction is heavily anchored by the principles of *Data Feminism* (D'Ignazio & Klein, 2020), a framework established by Catherine D'Ignazio and Lauren F. Klein, who argued that "Data is a double-edged sword" (p. 14). It is weaponized "by those in power to consolidate their control over places and things, as well as people" (D'Ignazio & Klein, 2020, pp. 14–17).

The current scholarship delves into issues where scholars challenge "objective" data. Domains like Data Feminism, a framework within the Digital Humanities, critiques how historical archives, when fed into machine-learning workflows and search algorithms, perpetuate historic forms of structural

misogyny and marginalization. In response, contemporary directions focus on radical archival recovery. Rather than accepting the "epistemic flattening" caused by standardized automated categorization, Digital Humanities projects are leveraging specialized digital environments such as the landmark *Orlando Project*, a digital scholarly database that documents the lives and writings of women authors writing in English from the earliest periods to the present. Unlike a simple digital archive, it combines biography, literary history, and textual relationships using structured digital encoding, published by Cambridge University Press (The Orlando Project, 2022), or feminist-driven linked open data initiatives to uncover, preserve, and center the historically erased voices of women and marginalized creators

Meanwhile, Digital Humanities centers around the globe have been growing in their public-facing and community-based digital archives. Some projects are concerned with the preservation of minority languages, oral history, and materials from communities that were historically underrepresented in print-based scholarly archives. These developments are in response to concerns of access and representation. Technology, funding, and discussions of what questions and materials should be addressed continue to influence the field.

CONCLUSION

Digital Humanities is undergoing a critical epistemological shift. Digital humanities has gone through a number of phases, from Roberto Busa's concordance project in 1949 to the current international organization of the field. It started with groundbreaking text-analysis projects and evolved into common standards, journals, professional societies, funding mechanisms, and digital archives. Then the use of computational methods spread to other fields such as art history, archaeology, musicology, and others. The field expanded quickly in the 2000s under the banner of "Digital Humanities." Distant reading and large-scale quantitative analysis became more prominent in the 2010s.

The paper demonstrates that Digital Humanities is not a static collection of methods that have been passed down unaltered from Busa's project. It is a practice that is influenced by technology, institutional funding, and the evolving questions that scholars pose. Ongoing debates about the nature and scope of Digital Humanities demonstrate that the field continues to evolve through the work of its scholars

References

Busa, R. A. (2004). Foreword: Perspectives on the digital humanities. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. xvi–xxi). Blackwell.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press. https://doi.org/10.7551/mitpress/11805.001.0001

Digital humanities manifesto 2.0. (2008–2009).

Hockey, S. (2004). The history of humanities computing. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. 3–19). Blackwell.

Kirschenbaum, M. G. (2010). What is digital humanities and what's it doing in English departments? *ADE Bulletin, 150*, 55–61.

Liu, A. (2012). Where is cultural criticism in the digital humanities? In M. K. Gold (Ed.), *Debates in the digital humanities*. University of Minnesota Press.

McCarty, W. (1998). *What is humanities computing? Toward a definition of the field*. Centre for Computing in the Humanities, King's College London.

Moretti, F. (2000). Conjectures on world literature. *New Left Review, 1*, 54–68.

Schreibman, S., Siemens, R., & Unsworth, J. (2004). The digital humanities and humanities computing: An introduction. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A companion to digital humanities* (pp. xxiii–xxvii). Blackwell.

The Orlando Project. (2022). *Orlando: Women's writing in the British Isles from the beginnings to the present*. Cambridge University Press. https://orlando.cambridge.org/