

Exploring and Rethinking Digital Humanities in Philosophical Perspective

Fenfen Song

Department of Chinese Language and Literature, Northwest Minzu University, Lanzhou, China

Abstract. In light of the accelerated global growth of digital humanities, an escalating number of inquiries and critiques have been directed towards this field. Consequently, it becomes imperative to scrutinize the rationality of digital humanities research through a philosophical lens and address prevalent criticisms. Studies underscore that positivist philosophy forms the underpinning for quantitative research within the humanities, while structuralist theory furnishes a mature framework and structural model for digital humanities research practices. Upon comprehensive examination of diverse critiques aimed at digital humanities, a consensus emerges. Scholars engaged in digital humanities research are advised to integrate quantitative and qualitative research methodologies. This integration should encompass the incorporation of diverse data sources, analytical techniques, and expressive forms. This approach facilitates the analysis and presentation of intricate, dynamic, and multifaceted phenomena within the humanities. Moreover, it is essential to merge information processing with knowledge creation to forestall the generation of spurious findings. This necessitates a concerted effort to fortify the quality assessment and management of data, as well as to identify and rectify errors and biases inherent in various datasets. Such measures collectively enhance the usability and accuracy of the data in the realm of digital humanities research.

1. Introduction

In the wake of the digitalization wave, digital humanities, often regarded as a novel academic research methodology or "a new academic paradigm" [1], has found widespread application across various branches of the humanities. Rooted in computational thinking, digital humanities seeks to leverage digital technology for literature data processing, analysis, and visualization. Its goal is to propel the high-tech evolution of traditional humanities academic research by employing a research methodology that is relatively precise, empirical, and reproducible. Gerhard Lauer, Chair of the Center for Digital Humanities at the University of Basel, Switzerland, aptly expresses the transformative impact of digitization: "If change is about changing lives, then digitization is one such change that has dramatically altered every aspect of our lives, and the wind is at our backs" [2].

As the momentum behind digital humanities research continues to build, it has encountered skepticism and opposition. This article endeavors to examine the rationale behind digital humanities research methods or paradigms from a philosophical perspective, aiming to reflect on current criticisms within the field. The article is structured into two parts. The first part delves into the ideas and methodologies of positivism and structuralism, exploring their impact on digital humanities. The second part offers reflections on the evolution of digital humanities research, providing insights drawn from four key criticisms: the

misunderstanding of quantitative research, reflections on structuralism, the pursuit of intellectual creativity, and the quest for research accuracy.

2. Philosophical Exploration of Digital Humanities

Digital humanities represents an interdisciplinary research paradigm that leverages computational tools and methodologies to enhance traditional humanities. This approach integrates the quantitative and mathematical research methods inherent in the natural and social sciences, while preserving the interpretative essence of the humanities.

Within the realm of digital humanities research, abstract symbolic systems, including mathematics, logic, and language, assume a pivotal role as both research tools and subjects of investigation. Consequently, the philosophical underpinnings of digital humanities can be scrutinized through two primary lenses. First, the positivist perspective prioritizes the objective, precise, and verifiable description and analysis of humanistic phenomena. Second, structuralism places emphasis on unveiling and elucidating the deep structure and intrinsic relationships within humanistic phenomena. This involves the analysis of meaning and culture through the utilization of symbols and language, as well as the development of theories and methodologies.

1836201408@qq.com

On a macro level, positivist thinking and practices establish a robust philosophical foundation for conducting quantitative research in the humanities. On a micro level, the theories and methods of structuralism furnish comprehensive theoretical frameworks and structural models that underpin the research endeavors within the field of digital humanities.

2.1. Digital Humanities and Positivism

Compared to traditional humanities research, a defining characteristic of digital humanities research lies in its transformation of the methodology from qualitative to quantitative, or a fusion of both. In "Digital Humanities and Literary Studies", Pang Yuhou traces the roots of quantitative research in the humanities back to the 1850s, citing examples such as mathematician Augustus De Morgan's utilization of quantitative methods to analyze word lengths and physicist Thomas Mendenhall's application of word-frequency statistics to assess stylistic differences between authors and works [3].

The concept of quantitative research in the humanities finds its philosophical foundation in positivism, a viewpoint rooted in facts, observation, and experimentation. Originating in 17th to 18th century England, positivism emerged as a response to the speculative philosophies of knowledge. Advocates such as Francis Bacon, Isaac Newton, Edward Herbert, and David Hume championed experiments and observations as means to establish scientific laws and theories, rejecting metaphysics and theology grounded in reason or authority. Auguste Comte, a 19th-century French philosopher, formalized positivism through his work at the Institut de Philosophie Positive, dividing the development of human reason into three stages: theological, metaphysical, and empirical [4]. Comte emphasized the use of the scientific method to observe and study natural and social phenomena, seeking truth, usefulness, certainty, and precision [4].

John Stuart Mill, a British philosopher, further discussed the scientific research method within Comte's positivism. Mill advocated induction to explore causal relationships and introduced "Mill's Five Laws," emphasizing the elimination of potential interfering factors [5]. In the 1920s, the Vienna School expanded on positivism, giving rise to logical positivism. This branch focused on the logical analysis of knowledge, especially scientific language, defining criteria for proposition meaning through the "principle of verifiability". Logical positivists argued that only empirically verifiable statements possess meaning. Statements beyond empirical verification, such as metaphysics or ethics, were considered meaningless. They sought to analyze and reconstruct the language of science using symbolic logic and mathematics to eliminate linguistic ambiguity and confusion.

In essence, positivism promotes the acquisition of objective knowledge through empirical observation, determining regularity and causality. Logical positivism extends this by emphasizing the principle of verifiability, providing a more rigorous standard for quantitative

research. Additionally, it incorporates mathematical logic, aiming to express scientific thinking and conclusions with formalized language. Thus, the philosophical thinking and research practices of positivism establish a solid foundation for quantitative research in the humanities.

2.2. Digital Humanities and Structuralism

The impact of positivism on digital humanities is characterized by its fundamental and macroscopic influence, contrasting with the specific and microscopic influence of structuralism. Structuralism, originating in Europe in the early 20th century, emerged as an intellectual movement with the primary objective of deciphering the underlying code of human culture. Its core tenet posits that cultural phenomena are governed by hidden rules and patterns, discernible through symbolic representation and relational analysis. Initially rooted in linguistics, structuralism extended its purview to encompass anthropology, psychology, literature, and other disciplines, utilizing tools such as mathematics, logic, and language to analyze the structure and function of human phenomena. The theory and methodology of structuralism furnish a mature framework and structural model for the research practice of digital humanities.

Firstly, structuralism provides digital humanities with a structural model for analyzing humanistic phenomena. This approach contends that humanistic phenomena possess internal logic and laws, which can be represented through abstract symbols and relations, forming a structural model. Digital humanities leverages this model to code, label, classify, retrieve, and analyze digital texts, uncovering their structural features and semantic information.

Secondly, structuralism offers digital humanities an analytical method employing mathematics and logic. Drawing on mathematical and logical thinking, it transforms humanistic phenomena into quantifiable and operable data and symbols, enhancing the objectivity and precision of humanistic research. Digital humanities employs structuralism's analytical methods of mathematics and logic to quantify, calculate, reason, and prove, thereby augmenting the scientific rigor of humanities research.

Thirdly, structuralism provides a linguistic method of analysis for digital humanities. Originating in linguistics, it views language as a symbolic system and delves into its structure and function by analyzing phonemic, lexical, syntactic, and semantic levels. Digital humanities utilizes structuralist linguistics to conduct operations such as lexical segmentation, annotation, syntactic analysis, semantic analysis, and pragmatic analysis on digital texts, enriching the depth and comprehensiveness of humanities research.

Despite being distinct philosophical trends, positivism and structuralism exhibit connections and interactions in digital humanities. Both draw on mathematical and logical ways of thinking, employ technical means, and seek to uncover regularities and universality in humanities phenomena. Digital humanities serves as the successor and practitioner of both positivism and structuralism,

leveraging their concepts and methods while employing digital technology for large-scale, high-rate, and innovative computational research in the humanities.

3. Reflections on Criticism of Digital Humanities

The examination of the philosophical underpinnings of digital humanities provides a basis for a nuanced reflection on the critiques levied against this interdisciplinary field. Present criticism of digital humanities emanates from four principal areas: a misinterpretation of quantitative research, contemplation of structuralism, the pursuit of creativity within digital humanities, and the quest for precision in digital humanities.

3.1. Quantitative Research and Digital Humanities Criticism

Positivism serves as a robust foundation for the integration of quantitative research methodologies in the humanities. However, the application of quantitative research in this domain has elicited a viewpoint that seeks to capture the qualitative essence of the humanities, thereby sparking criticism of digital humanities. Scholars, including literary expert Stephen Mahey, have contested the practice of "remote reading" in digital humanities, drawing a comparison to the "close reading" methodology of the New Criticism. In his work "Literature Is not Data: Against Digital Humanities", Mahey argues that "literature is not data. Literature is the opposite of data." "The algorithmic analysis of novels and of newspaper articles is necessarily at the limit of reductionism. The process of turning literature into data removes distinction itself. It removes taste. It removes all the refinement from criticism. It removes the history of the reception of works" [6]. According to Mahey, the process of transforming literature into data erodes critical distinctions, taste, refinement, and the historical reception of works.

The underlying criticism stems from an ideological equation of quantitative research with positivism, positioning it in opposition to qualitative research, which is perceived as aligned with humanistic research paradigms. This perspective has traditionally led to the confinement of quantitative and qualitative research methods to specific disciplinary areas. Addressing this, Xie Lizhong, a professor at Peking University's Department of Sociology, challenges the binary thinking by asserting in "Reconsidering the Relationship Between Quantitative and Qualitative Research in the Field of Social Studies" that positivism is not synonymous with quantitative research, nor is humanism exclusively tied to qualitative research. Both positivist and humanist research paradigms can incorporate either quantitative or qualitative research modes [7].

Recognizing that quantitative and qualitative research represent distinct modes, not rigid paradigms, it is imperative to avoid simplistic limitations, such as confining quantitative research to natural and social sciences and qualitative research to the humanities. Instead, the selection of research modes should align with the

research's purpose, object, problem, and methodology. It is crucial to underscore that the humanities' inclination towards qualitative statements, reasoning, and interpretation does not preclude the application of quantitative research methodologies. However, an exclusive reliance on quantitative research, neglecting qualitative aspects, can have detrimental consequences, including a lack of research depth, innovation, and the emergence of what is termed as "technophilia."

The recommended approach is the integration of both quantitative and qualitative research methodologies in digital humanities research. By combining data analysis with humanistic interpretation, this approach optimally leverages the strengths of big data research while preserving the artistic and emotional characteristics inherent in the humanities. This nuanced integration ensures a holistic exploration of the subject matter, steering clear of reductionist tendencies and promoting a more comprehensive understanding of the complexities within digital humanities research.

3.2. Structuralism and Digital Humanities Criticism

An additional facet of criticism directs attention towards the technical and methodological challenges inherent in digital humanities research. This critique encompasses two primary concerns. Firstly, there are notable deficiencies in the selection, processing, analysis, and presentation of data, leading to outcomes that are deemed unreliable, inaccurate, insufficient, or irrelevant. Secondly, digital humanities research is faulted for quantifying humanistic knowledge as a calculable, manipulatable, and comparable entity, thus failing to capture the intricate complexity and specificity inherent in the humanities.

Nan Z. Da, in "The Computational Case against Computational Literary Studies," highlights the technical, logical, and conceptual deficiencies within computational literary studies. These include challenges in processing data attributes, utilizing statistical methods, and selecting appropriate tools. Da contends that the intricate nature of literature surpasses the capabilities of statistical tools, underscoring the inadequacies in the application of these methodologies within digital humanities research [8].

This critique on a technical and methodological level is deeply rooted in a philosophical reflection on structuralism. While structuralism has introduced novel perspectives and methodologies to the humanities, it has faced challenges and refutations from post-structuralist, deconstructivist, and postmodernist schools of thought. These opposing currents argue that structuralism neglects the history, diversity, and uncertainty inherent in humanities phenomena. They assert that structuralism excessively emphasizes the stability and universality of structures, ignores the creativity and uniqueness of the subject, and reduces humanities phenomena to an abstract, static, and formal system. Consequently, structuralism is deemed inadequate in adequately explaining the complexity and dynamism within humanities phenomena.

In response, scholars who embrace digital humanities research are urged not to hastily dismiss criticisms

concerning technical and methodological shortcomings. Such critiques offer valuable insights, prompting digital humanists to recognize the limitations of their field. This acknowledgment steers them away from excessive reliance on algorithms and the generation of technology-centric theories. Simultaneously, these critiques serve as inspiration for digital humanities researchers to diversify their approaches by introducing additional data sources, analytical methods, and expressive modalities. The objective is to demonstrate the intricate complexity, dynamism, and plurality inherent in humanities phenomena.

In essence, digital humanities should be positioned as both the inheritor and practitioner of structuralism, acknowledging its contributions, and as the critic and transducer of structuralism, seeking to overcome its limitations. This nuanced approach ensures that digital humanities continues to evolve, incorporating diverse methodologies and addressing the complexities intrinsic to the study of humanities phenomena.

3.3. Digital Humanities and the Creation of Knowledge

Certain scholars have critiqued digital humanities research for its perceived overemphasis on data and technical tools, suggesting an immersion in the gratification of information acquisition at the expense of knowledge creation. Notably, Brennan, Timothy, in "The Digital-Humanities Bust: After a Decade of Investment and Hype, What Has the Field Accomplished? Not Much", contends that "the field commits a fundamental error by equating an abundance of information with increased knowledge"[9]. Brennan's perspective, however, is deemed by some as overlooking the innovative and diverse nature of digital humanities. His generalization is seen as dismissing the substantial value and contribution that digital humanities can make to scholarly pursuits.

While acknowledging the potential of digital humanities to generate new knowledge through statistical analysis, it is underscored that the field should transcend mere information acquisition. A call is made for digital humanities to prioritize the creation and dissemination of knowledge to forestall the generation of invalid research outcomes. Two primary categories of invalid research results are identified. The first pertains to an over-reliance on data analysis and presentation in actual humanities research, sidelining in-depth interpretation of humanities phenomena. The second involves the articulation of self-evident conclusions. While the former concern has been diligently addressed by many digital humanities researchers, the latter poses a critical question regarding the rationale behind using digital humanities research methods to validate established conclusions.

This consideration prompts reflection on the necessity of certain propositions and the potential hindrance posed by repetitive research to innovative development. The imperative for digital humanities researchers is to engage in continuous exploration and critical thinking in practice. While the field holds the capacity to validate existing conclusions, careful consideration must be given to the

selection of propositions deemed essential for such validation. The ongoing discourse encourages the digital humanities community to navigate a path that not only avoids the pitfalls of invalid research but also fosters innovation and meaningful advancement within the discipline.

3.4. Digital Humanities and Data

The final category of critique emanates from within the realm of digital humanities research and focuses predominantly on the quality of data. For instance, Zheng Yongxiao, a researcher at the Institute of Literature of the Chinese Academy of Social Sciences, underscores issues of proofreading and errors in documents digitized from ancient books, particularly within the context of classical Chinese literature. He laments, "It is indeed regrettable that the database of ancient books, developed at considerable expense, is utilized merely as a tool for searching, querying, and saving typing time"[10]. This perspective reflects the meticulous concerns of rigorous digital humanities researchers regarding the lack of data quality.

As the pace and scale of literature digitization have surged, there has been a concomitant increase in data errors, giving rise to digital humanities research with precarious foundations. These challenges are not confined solely to technical aspects but extend to scholarly dimensions as well. Errors in data can stem from various stages such as collection, cleaning, conversion, storage, and computation, impacting the completeness, accuracy, consistency, and credibility of the data. Consequently, these technical and scholarly issues diminish the efficacy and value of data analysis and data mining.

Several scholarly works, including Guo Zhimao and Zhou Aoying's "A Summary of Data Quality and Data Cleaning Research" and Guo Azhang, Liu Xiuyuan et al.'s "Research on Key Problems of Data Quality in Large Industrial Data Environment," have delved into improving data quality. Their findings underscore the necessity for digital humanities researchers to diligently assess, manage, and rectify errors and biases in diverse datasets, ultimately enhancing the usability and accuracy of the data. This reciprocal relationship between data and big data research fosters mutual advancement.

Albert Einstein's aphorism, "The important thing is not to stop questioning" [11], encapsulates the ethos of addressing these challenges. Critics, after rigorous analysis, present genuine problems that have evolved into new focal points for the advancement of digital humanities. Consequently, these challenges serve as a compelling impetus, driving the development of the field. Digital humanities researchers are urged to engage in thoughtful reflection and implement practical solutions to overcome these hurdles. By doing so, they can significantly contribute to the robust growth of digital humanities research and facilitate the high-tech transformation of traditional humanities academic research.

4. Summary

In summary, the philosophical tenets of positivism and structuralism are intricately linked to the field of digital humanities. Positivism, on a macro level, establishes a robust philosophical foundation, facilitating quantitative research in the humanities. Conversely, on a micro level, structuralism furnishes mature theoretical frameworks and structural models that underpin the nuanced research practices within digital humanities.

As the inheritor and practitioner of these philosophical ideas, Digital Humanities employs digital technology to initiate large-scale, high-rate, and innovative computational research endeavors in the humanities. However, certain misconceptions regarding positivism and quantitative research, coupled with reflections on structuralism, have led to criticisms of digital humanities. Detractors either oppose the application of quantitative research in the humanities or scrutinize technical and methodological limitations within digital humanities research. Scholars have pointed out perceived deficiencies in creativity, knowledge depth, and research accuracy.

Addressing these concerns, we propose the following considerations: First, digital humanities research should ideally integrate quantitative and qualitative approaches, combining data analysis with the interpretation of humanistic nuances to leverage the strengths of big data research while preserving the qualitative essence of the humanities. Second, in the face of the complexity, dynamism, and plurality of humanities phenomena, researchers should acknowledge the limitations of digital humanities and avoid overreliance on algorithms. It is imperative to introduce diverse data sources, analytical methods, and expressions to comprehensively analyze and present the intricate, dynamic, and pluralistic nature of humanities phenomena. Third, amidst the wealth of data resources and advanced digital technology, researchers should not solely indulge in information acquisition but should also merge information processing with knowledge creation to prevent the generation of invalid research outcomes. Fourth, recognizing that data accuracy significantly impacts research reliability, digital humanities researchers must rigorously assess and manage data quality, identifying and rectifying errors and biases to enhance the usability and accuracy of data.

In conclusion, as an emerging academic research method or paradigm, digital humanities faces theoretical, technical, and methodological challenges. These include biases arising from technological and professional divides, imbalances between technological and theoretical development, and conflicts between humanities subjects and digital objects. Criticisms and reflections from diverse disciplines and perspectives prompt valuable discussions on the concept, connotation, method, value, and impact of digital humanities. While these critiques may be pointed, they contribute to steering the development of digital humanities towards a more scientific, rational, and robust trajectory. This, in turn, advances the construction of theoretical systems and methodologies within digital humanities, broadens the scope of research, and uncovers the potential for its further development, thereby making a

substantial contribution to the advancement of the humanities.

From current research findings, digital humanities not only offers new paradigms and perspectives for the humanities but also enriches data sources and problem domains, enhancing the scientificity and innovation of the humanities. Moreover, it provides novel cultural resources and services for society, fostering cultural innovation, tolerance, and societal progress. Thus, an optimistic and open-minded approach from humanities researchers is essential to embrace and support the cumulative research efforts within digital humanities.

References

- 1 Wang, N. (2023). Digital humanities as a paradigm of humanities scholarship. *Journal of Shanghai Jiao Tong University (Philosophy and Social Science Edition)*, 31(12), 1-12.
- 2 Lauer, G. (2018). Cultural measurement in the digital humanities. *Chinese Social Science Digest*, (12), 55-56. (Translated by Pang Nana)
- 3 Pang, Y. (2024). Digital humanities and literary studies. *Journal of Guangdong University of Foreign Studies*, 35(02), 104-113.
- 4 Comte, A. (2011). *Discours Sur L'esprit Positif. Une oeuvre du domaine public*.
- 5 Mill, J. S. (2002). *A system of logic, ratiocinative and inductive* (7th ed.). Longmans, Green, Reader and Dyer.
- 6 Marche, S. (2012). Literature Is not Data: Against Digital Humanities. Retrieved January 4, 2024, from <https://lareviewofbooks.org/essay/literature-is-not-data-against-digital-humanities>.
- 7 Xie, L. (2019). Reconsidering the relationship between quantitative and qualitative research in the field of social studies. *Hebei Journal*, 39(02), 160-170.
- 8 Nan, Z. D. (2019). The Computational Case against Computational Literary Studies. *Critical Inquiry*, 45(03), 601-639.
- 9 Brennan, T. (2017). The Digital-Humanities Bust: After a Decade of Investment and Hype, What Has the Field Accomplished? Not Much. *The Chronicle of Higher Education*, pp. B4.
- 10 Zheng, Y. (2019). Inheritance and transcendence: Ruminations on the future development of digital bibliography—Anecdotal discussion on the inspiration of Japanese literature digitization for China. *Chinese Comparative Literature*, (04), 5.
- 11 Einstein, A. (2011). *The Ultimate Quotable Einstein* (A. Calaprice, Ed.). Princeton University Press, p.425.