FROM CLOSE TO DISTANT READING: DIGITAL HUMANITIES IN LITERARY STUDIES

Febi Puspitasari
Universitas Negeri Yogyakarta, Indonesia
Email: febi.puspitasari@uny.ac.id

Anwar Efendi
Universitas Negeri Yogyakarta, Indonesia
Email: anwar@uny.ac.id

Rony Kurniawan Pratama
Universitas Sebelas Maret, Indonesia
Email: ronykpratama@staff.uns.ac.id

Trida Chairu
Badan Riset dan Inovasi Nasional, Indonesia
Email: trid001@brin.go.id

Abstract
Digital humanities offers a novel approach to comprehensively understand, analyze, and interpret literary texts. This study aims to elucidate the various digital humanities methods and their implications for advancing literary studies. A systematic literature review with a descriptive-analytical approach was employed to examine digital humanities methods and analytical tools within Indonesian literature. Four primary approaches were identified: (a) digital text analysis, (b) data visualization, (c) sentiment analysis and text classification, and (d) geographic mapping. Software such as Voyant Tools and AntCont facilitate textual analysis, while Tableau and Gephi open new dimensions in visualizing narrative structures and interrelations between literary elements. Machine learning methods and natural language processing libraries enable more sophisticated sentiment analysis and text classification. In contrast, ArcGIS and Neatline expand the exploration of spatial dimensions in literary works. Digital humanities methods allow for the integration of quantitative and qualitative analyses, yielding novel insights into literary text analysis. The implementation of digital humanities methods should be studied rigidly and critically, always considering the cultural and historical contexts of Indonesian literature. Digital humanities methods should also be viewed as a complement, rather than a replacement, for established conventional approaches. The integration of digital humanities methods enriches the body of knowledge in Indonesian literary studies.

Keywords: Digital humanities, literature, digital text analysis, data visualization, sentiment analysis.

INTRODUCTION

The past decade has witnessed a transformative shift in literary studies, driven by the burgeoning discipline of digital humanities. The proliferation of digital tools and analytical methods have fundamentally reshaped the epistemological and methodological approaches to understanding, analyzing, and interpreting literary texts (Berry & Fagerjord, 2017). Digital humanities, rooted in ""humanities computing"", has undergone significant evolution over the past four decades, offering novel and more comprehensive perspectives on literary research (Schreibman et al., 2015).

In the realm of literary studies, digital humanities provide a research methodology that integrates quantitative and qualitative analysis. This approach enables "distant reading", complementing the traditional method of ''close reading'' (Moretti, 2013). Consequently, digital humanities should not be considered as a threat to established humanistic traditions but rather as a valuable supplement that enriches and expands the scope of literary studies (Jockers, 2013).

Integrating digital methodologies into literary studies offers efficiency in data processing and unlocks novel avenues of analysis that potentially revolutionize our understanding of the production, reproduction, reception, and distribution of literary works. This approach allows for a synthesis of meticulous close reading with comprehensive distant reading, creating a "hybrid" paradigm that bridges the limitation of each method. As a result, literary scholars will witness a "new landscape" in literary studies—a terrain where big data and algorithms synergize with traditional humanistic sensibilities, expanding the horizons of textual and cultural exploration.

The shift from close reading to distant reading reflects a significant methodological evolution in contemporary literary studies. Close reading, long the cornerstone of literary analysis, is recently complemented by distant reading, which enables large-scale analysis and the identification of researchers to analyze hundreds or even thousands of text simultaneously, revealing trends and patterns that might be obscured in traditional readings (Dobson, 2019; Moretti, 2013).

Nevertheless, applying digital humanities methods in literary studies also presents challenges and debates. Critics often express concerns about the loss of interpretive nuance as a characteristic of traditional literary studies (Dobson, 2019). Therefore, scholars emphasize the importance of an integrated approach that combines close and distant reading, merging the power of quantitative analysis with the interpretive sensitivity of close reading (Bode, 2018).

In the context of Indonesian literature, adopting digital humanities methods opens opportunities for broader and deeper explorations of the national literary corpus. This approach can potentially reveal prior unidentified thematic, stylistic, and narrative patterns and facilitate more comprehensive comparative analyses with regional and global literature (Riza, 2019).

The integration of digital humanities methods into Indonesian literary studies is not intended to replace the established approach of close reading but rather to complement and, even, enrich it. A "scalable reading" approach that combines macro-analysis (distant reading) with in-depth readings of critical works can serve as an effective model for contemporary Indonesian studies (Mueller, 2012). Therefore, it is undeniable that by adopting digital humanities methods, Indonesian literary studies have the potential to uncover new insight into the evolution of themes, styles, and narratives within the national literary corpus, revealing the complexities of intertextual relations, and the strategic position of Indonesian literature in global dialogues of literature and culture.

METHOD

The study employs a systematic literature review through a descriptive-analytical approach to examine digital humanities methods and tools within the context of Indonesian literature. Data sources include academic literature in various formats, such as journal articles, textbooks, monographs, conference proceedings, and trusted online sources such as academic repositories and research institution websites. Data is in the form of digital or printed texts discussing digital humanities and their applications in literary studies. The focus is on publication within the past decade to ensure the relevance and currency of the information, given the rapid development of text mining and digital analysis in humanities studies (Jockers & Underwood, 2015).

Data analysis followed several stages: (1) identification and categorization of methods and tools based on their type and function; (2) comparative evaluation to assess the strengths and limitations of each method; (3) synthesis of findings to identify trends and potential implementation in the context of Indonesian literature, and (4) critical interpretation of the implication of using these methods in Indonesian literary studies.

Within the scope of Indonesian literature, this study also considers a postcolonial perspective in digital humanities (Risam, 2018). It is crucial to ensure that the methods and the tools are both technically sound and culturally sensitive within the context of

Indonesian literature. Such an approach facilitates a comprehensive exploration of the methodological landscape of digital humanities in literary studies while being relevant and applicable in the Indonesian context. The study also considers the implications of using digital collection in literary studies to provide a novel lens through which to view literary history (Bode, 2018)

FINDINGS AND DISCUSSION
Digital a variety and tools of digital humanities analysis

The first research question pertains to the diversity of digital humanities approaches. Table 1 presents a detailed examination of the diverse digital humanities methods in literary research and their implications for their potential contributions to the field.

Table 1. Digital humanities approaches in literary studies

Methods	Tools	Primary Function	Potential Applications in Literary Studies
Analysis of Digital Texts	Voyant Tools	Identification of linguistic patterns, themes, and stylistic elements	Frequency analysis of words with specific characteristics to identify literary elements
	AntConc	Concordance and collocation analysis	Study of metaphors in literary studies
Data Visualization	Gephi	Mapping of inter-character relationships	Social Network analysis in literary works
	Tableau	Visualization of thematic trends	Evolution of themes in Indonesian literature.
Sentiment Analysis	Python (NLTK)	Sentiment analysis in text	Comparison of sentiment in literary works across generations
Text Classification	Scikit-learn	Categorization of genre or theme	Automated classification of short stories or novels based on theme
Geographic Mapping	ArcGIS	Analysis of spatial representation	Mapping of settings in novels or short stories
	Neatline	Visualization of character journeys	Routes of characters in novels or short stories

The paradigm shift from close reading to distant reading

The evolution of research methodologies within literary studies has undergone a significant transformation during the 20th and early 21st centuries. This paradigmatic shift reflects broader intellectual dynamics within the humanities and social studies. One of the most notable transitions is the movement from structuralism to post-structuralism (Culler, 2011). Based on Saussurean linguistics, structuralism conceptualized literary works as a system of signs susceptible to objective analysis. This approach emphasized the significance of internal structure within a text and the interrelationships among its constituent elements.

Post-structuralism, as a critical response to structuralism, shifted its focus from fixed structures to the fluidity of meaning and the interpretive subjectivity of the reader. Culler (2011) explains how thinkers such as Jacques Derrida and Roland Barthes challenged the structuralist assumption of fixed and determinate meaning. Post-structuralism, consequently, expanded the possibilities for plural interpretation and affirmed the readers'' agency in shaping meaning, a notion popularized by Barthes as the ''death of the author''.

Subsequent developments were signified by the emergence of new historicism and cultural studies (Greenblatt et al., 2009). New historicism emphasized the importance of historical context in literary interpretation, rejecting the isolated formalist approach.

Greenblatt et al. (2009) demonstrated how literary works are products of specific social, political, and cultural conditions, and actively shape these historical realities.

The influence of postcolonial theory on literary studies introduced another critical dimension to the evolution of literary research methodologies (Said, 1993). Imperial and colonial perspectives tend to shape literary representations of non-Western cultures. The postcolonial analysis introduced concepts such as "Orientalism" and "writing back", allowing for a re-reading of the literary canon with a critical awareness of power dynamics and representation. This approach not only transformed the way to read colonial text but also offered different insights for prior marginalized voices within the global literacy discourse.

This evolution of research methods reflects a broader shift in understanding the nature of language, meaning, and the relationship between text and context. According to the internal structure of texts, literary studies have shifted toward a more comprehensive understanding of the interplay between texts, readers, and socio-historical contexts.

Following the preceding transformations, the shift from close to distant reading also represents a paradigmatic shift in contemporary literary studies. Moretti (2013) suggests that distant reading facilitates analysis on a more extensive scale, revealing patterns and trends which are probably discernible through traditional reading approaches. This paradigm allows for exploring literary corpora in a broader and more profound dimension.

However, this paradigmatic shift does not negate the significance of close reading. Bode (2018) elucidated that the integration of distant and close reading can yield a more holistic analysis. Therefore, distant reading functions as a heuristic tool to identify macro-level patterns, which can subsequently be elaborated through in-depth close reading.

Underwood (2019) emphasizes that paradigmatic shift also necessitates a reconceptualization of what constitutes ''reading'' in the digital age. He proposes the concept of distant reading not as an antithesis to close reading. Instead, the concepts are a spectrum of interpretive practices that enable researchers to traverse between micro and macro scales in literary analysis.

Integration of quantitative and qualitative analysis

The integration of quantitative and qualitative analysis in literary studies represents one of the most significant manifestations of the adoption of digital humanities methods. Data visualization tools such as *Gephi* enable visual representations of narrative structures or character networks,s visual representations of narrative structures or character networks, serving as a starting point for more in-depth qualitative analysis. Drucker (2017) underlines the importance of maintaining a critical perspective in interpreting humanistic data visualization, given that data in the humanities is generally interpretive and contextual.

This integration creates opportunities for more comprehensive exploration. For instance, visualizing character networks in Indonesian historical novels can reveal power dynamics or social relationships that reflect specific historical contexts. This aligns with Piper (2018) view that computational analysis can help uncover "hidden patterns" in literary texts that may be unseen during conventional reading.

The novel approach, digital humanities methods, constitutes a veritable revolution in analyzing and understanding literary texts. One of the most significant approaches is stylometry. Stylometry, which focuses on identifying and quantifying stylistic features, has proven invaluable in various applications, ranging from authorship attribution to the analysis of the evolution styles over time. These techniques allow researchers to identify a writer''s "linguistic fingerprint", aiding in solving authorship controversies and providing new insights into the development of an individual writer''s style or broader stylistic trends within a particular literary period.

Exploration of emotional and thematic dimension

Sentiment analysis and text classification using machine learning methods reveal new dimensions in understanding literary works' emotional and thematic aspect. These methods enable researchers to explore emotional nuances and thematic developments on a larger scale and more precisely. However, as Liu (2020) cautions, it is necessary to consider the potential biases inherent in algorithms and validate results by involving deep domain knowledge.

Sentiment analysis explores changes in mood or tone in poems from a specific period, such as comparing poems from the New Order and Reform era. It aligns with Kim & Klinger (2018) study demonstrating how computational sentiment analysis helps uncover emotional patterns in larger poetry corpora.

Furthermore, machine learning-based text classification can assist in identifying and analyzing dominant themes in Indonesian literary corpora. It can provide new insights into the evolution of national literature over time. However, as Underwood (2019) highlights, it is essential to understand that the generated thematic categories by machine learning algorithms may not always align with categories used in traditional literary analysis.

The challenge in exploring emotional and thematic dimensions using digital methods is the limitation in capturing the nuances and ambiguities that commonly characterize literary works. As Piper (2018) asserts, computational analysis tends to produce uniform representations of literary texts, potentially overlooking the complexity and depth of meaning through human reading.

Geographic and spatial contextualization

Geographic mapping using tools such as ArcGIS enables the exploration of spatial dimensions in literary works, revealing new insights into representations of space and mobility. Presner et al. (2014) demonstrate how digital mapping can uncover geopolitical aspects within literary narratives, offering fresh perspectives on understanding the relationship between texts and their geographic contexts.

Mapping the geographic settings in travel novels or works of literature centered on locality can provide a deeper understanding of spatial representations and identity within national narratives. It aligns with the concept of "literary geography" proposed by Piatti et al. (2009), which emphasizes the importance of understanding spatial dimensions in the interpretation of literary texts.

Digital mapping contributes to the visualization and analysis of mobility patterns within Indonesian literary narratives. Such mapping reveals how the movement of characters within geographic space reflects broader social, cultural, and historical dynamics. Bodenhamer et al. (2015) state that digital mapping in the humanities is about representing physical space and about exploring the "conceptual spaces" shaped by narratives and interpretations.

However, digital mapping in a literary context also has limitations. As Cresswell (2014) cautions, cartographic representations tend to reduce the complexity of spatial experiences to static points on a map. Therefore, integrating digital mapping with in-depth qualitative analysis is necessary to comprehend the nuances and complexities of spatial representations in literary works.

Challenges and opportunities in the context of Indonesian literature

The adoption of digital humanities methods in Indonesian literary studies contributes to challenges and opportunities. Riza (2019) emphasizes the need to develop tools and methods sensitive to Indonesian literature's linguistic and cultural characteristics. For example, by developing sentiment analysis algorithms that are capable

of capturing the nuances of Indonesian and regional languages, which may not be accurately detected by algorithms designed for English.

Another substantial challenge lies in the availability of structured and comprehensive digital corpora. Vania et al. (2014) highlight the need for systematic efforts to digitize and index Indonesian literary works, especially classical ones or those written in regional languages. Such digitization and indexing are vital prerequisites for the effective and representative implementation of digital humanities methods.

Meanwhile, great opportunities are available due to the adoption of digital humanities methods in Indonesian literary studies. Digital techniques can help uncover patterns and trends in large literary corpora, which seems challenging to which seemed challenging for traditional methods. In the realm of Indonesian literature, it benefits a broader and more systematic understanding of the evolution of national literature, in terms of themes, styles, and structures.

Digital humanities methods also contribute to more extensive studies. As Bode (2018) demonstrates, computational analysis allows systematic comparisons between different literary corpora, which can provide new insights into the relationships and differences among various literary traditions.

Epistemological and ethical implication

Digital humanities methods trigger a multitude of ethical and epistemological questions. A prominent ethical issue relates to privacy and copyright in collecting and analyzing digital data. As Birnbaum et al. (2017) emphasize, the use of contemporary literary texts, particularly digital ones, raises questions about boundaries between fair use for research and potential copyright infringement.

From the epistemological standpoint, adopting digital methods in literary studies necessitated a reconceptualization of what constitutes ""knowledge"" within the context of literacy scholarship. As Ramsay (2011) explains, computational methods alter the way to access and analyze texts and the procedures for understanding and conceptualizing the objects of our study. Liu (2011) further highlights the significance of preserving ""humanity"" within the digital humanities. She argues that while digital methods offer powerful tools for analysis, human interpretation, and understanding remain at the core of humanities research, including literary studies.

Meanwhile, D'Ignazio & Klein (2020a) underscore the critical issue of representation and bias in digital data. They contend that the processes of collecting, processing, and analyzing digital data contain biases rooted in social structural inequalities. It reflects dominant groups' overrepresentation of works or the underrepresentation of marginal voices within digital corpora.

'D''Ignazio & Klein (2020) propose a more inclusive and reflexive methodology for data collection and analysis, including considering the socio-historical context of data, acknowledging the limitations and potential biases of datasets, and involving diverse perspectives in interpreting the results. In this context, it could mean actively seeking to include works by authors from diverse gender, ethnic, and socioeconomic backgrounds in digital corpora and developing linguistically and culturally sensitive analytical tools.

As discussed by Birnbaum et al. (2017), privacy and copyright issues in digital research complement the ethical complexities of the digital humanities. The digitization and computational analysis of literary works raise questions about the boundaries of fair use in academic contexts, especially for copyrighted works. Birnbaum et al. (2017) underscore the importance of developing clear ethical and legal protocols for accessing and using digital materials for literary research.

These ethical challenges show the demand for a more holistic and reflexive approach to the digital humanities. Digital literary researchers should cultivate a critical awareness of their methodologies' ethical implications and collaborate with ethics, law, and information technology experts in ethics, law, and information technology. Therefore,

the academic community should engage in ongoing dialogue about ethical standards in digital research, ensuring that methodological advancement does not compromise ethical integrity or social justice.

Integration with traditional approaches

The integration of digital humanities with traditional approaches is another crucial aspect. Mueller (2012) proposes a "distant reading" approach that combines macro and micro analysis. It involves using digital text analysis to identify broad patterns in large corpora, followed by close reading of main works that represent these patterns.

This integrative approach aligns with the concept of "macroanalysis" proposed by Jockers (2013), emphasizing the importance of combining large-scale computational analysis with a close reading of contextual interpretation. It involves digital tools to identify trends and patterns in large Indonesian literary corpora, which can then be elaborated upon through in-depth analysis of representative works. Such integration also implies challenges. As Bode (2018) cautions, there is a risk that the result of the computational analysis may be biased towards the researcher's subjective interpretation. A critical and reflective attitude, therefore, is necessary in the integration of digital methods with traditional approaches.

Underwood (2019) underlines that the integration of digital methods with traditional approaches is not merely about analytical techniques but also about conceptualizing the object of the study. She argues that digital methods can facilitate an understanding of literature not merely as a collection of individual texts but also as a system that evolves and changes over time.

Developing digital literacy in literary studies

The adoption of digital humanities methods in Indonesian literary studies implies a need for the development of digital literacy among literacy researchers. It includes technical skills in using digital tools and a critical understanding of the potential and limitations of these methods. Mandell (2019) emphasizes the importance of "algorithmic literacy" for digital humanities researchers. It involves a basic understanding of how algorithms and computational models work, including an awareness of the potential biases and limitations in computational analysis.

Berry & Fagerjord (2017) propose the concept of "critical digital literacy" which encompassess technical skills and an understanding of the social, cultural, and ethical implications of using digital technology in humanities research. This is becoming increasingly important in the Indonesian context, where the adoption of digital technology into literary studies is a relatively recent phenomenon.

Developing digital literacy also involves the ability to integrate digital methods with traditional approaches in literary studies. As Jockers (2013) explains, literary researchers need to develop a methodological "bilingualism", the ability to move sinuously between quantitative and qualitative analysis, and between close and distant reading. Challenges in developing digital literacies in Indonesia are not only related to individual skills but also to infrastructure and access. As highlighted by Riza (2019), limited access to hardware, software, and training can be significant barriers to the adoption of digital methods among Indonesian literary researchers.

CONCLUSION

This research clarifies the potential of digital humanities methods in expanding and deepening the analysis of Indonesian literature. Four primary approaches were identified: Digital Text Analysis, Data Visualization, Sentiment analysis and Text Classification, and Geographic Mapping, each offering unique perspectives in literary analysis. Software such as VOcant Tools and AntConc facilitate textual analysis. Meanwhile, Tableau and Gephi

create new dimensions in visualizing narrative structures and relationships among literary elements. Machine learning and NLP libraries enable more sophisticated sentiment analysis and text classification, while ArccGIS and Neatline expand the exploration of spatial dimensions within literary works.

The integration of quantitative and qualitative analysis, made several methods and tools possible to yield novel insights into the analysis of literary texts. However, digital humanities methods should be studied rigidly and critically while considering Indonesian literature's cultural and historical context. The balance between methodological innovation and culturally sensitive viewpoints is vital in adopting digital humanities approaches. The approaches are considered complements rather than replacements for established conventional approaches. Thus, integrating digital humanities methods enriches the body of knowledge in Indonesian literary studies—despite the fact digital humanities inevitably raise ethical questions, particularly regarding the integrity and validity of research findings.

REFERENCES

Berry, D. M., & Fagerjord, A. (2017). *Digital humanities: Knowledge and critique in a digital age*. Polity Press

Birnbaum, D. J., Bonde, S., & Kestemont, M. (2017). The digital middle ages: An introduction. *Speculum*, *92*(S1), S1–S38. https://doi.org/10.1086/694236

Bode, K. (2018). *A world of fiction: Digital collections and the future of literary history*. University of Michigan Press. https://doi.org/10.3998/mpub.8784777

Bodenhamer, D. J., Corrigan, J., & Harris, T. M. (2015). *Deep Maps and Spatial Narratives* (D. J. Bodenhamer, J. Corrigan, & T. M. Harris, Eds.). Indiana University Press. https://doi.org/10.2307/j.ctt1zxxzr2

Cresswell, T. (2014). *Place: An introduction*. John Wiley & Sons

Culler, J. (2011). *Literary theory: A very short introduction*. Oxford University Press

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

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

Dobson, J. E. (2019). *Critical Digital Humanities*. University of Illinois Press. https://doi.org/10.5406/j.ctvfjd0mf

Drucker, J. (2017). Why distant reading isn't. *PMLA/Publications of the Modern Language Association of America*, *132*(3), 628–635. https://doi.org/10.1632/pmla.2017.132.3.628

Greenblatt, S., Županov, I., Meyer-Kalkus, R., Paul, H., Nyíri, P., & Pannewick, F. (2009). *Cultural Mobility*. Cambridge University Press. https://doi.org/10.1017/CBO9780511804663

Jockers, M. L. (2013). *Macroanalysis: Digital methods and literary history*. University of Illinois Press.

Jockers, M. L., & Underwood, T. (2015). Text-Mining the Humanities. In *A New Companion to Digital Humanities* (pp. 291–306). Wiley. https://doi.org/10.1002/9781118680605.ch20

Kim, E., & Klinger, R. (2018). A Survey on Sentiment and Emotion Analysis for Computational Literary Studies. *Zeitschrift Für Digitale Geisteswissenschaften*. https://doi.org/10.17175/2019_008

Liu, A. (2011). Friending the past: The sense of history and social computing. *New Literary History*, *42*(1), 1–30. https://doi.org/10.1353/nlh.2011.0004

Liu, A. (2020). Toward a Diversity Stack: Digital Humanities and Diversity as Technical Problem. *PMLA/Publications of the Modern Language Association of America*, *135*(1), 130–151. https://doi.org/10.1632/pmla.2020.135.1.130

Mandell, L. (2019). *Breaking the book: Print humanities in the digital age*. John Wiley & Sons

Moretti, F. (2013). *Distant reading*. Verso

Mueller, M. (2012). *Scalable reading*. https://scalablereading.northwestern.edu/

Piatti, B., Bär, H. R., Reuschel, A.-K., Hurni, L., & Cartwright, W. (2009). Mapping literature: Towards a geography of fiction. In *Cartography and art* (pp. 177–192). Springer. https://doi.org/10.1007/978-3-540-68569-2_15

Piper, A. (2018). *Enumerations: Data and literary study*. University of Chicago Press

Presner, T., Shepard, D., & Kawano, Y. (2014). *HyperCities: Thick mapping in the digital humanities*. Harvard University Press

Ramsay, S. (2011). *Reading machines: Toward an algorithmic criticism*. University of Illinois Press

Riza, H. (2019). Indigenous language processing in Indonesia. In K. V. Subbarao & P. Bhaskararao (Eds.), *The languages of South and Southeast Asia: Problems and solutions* (pp. 187–206). Cambridge Scholars Publishing

Said, E. W. (1993). *Culture and imperialism*. Vintage

Schreibman, S., Siemens, R., & Unsworth, J. (2015). *A New Companion to Digital Humanities* (S. Schreibman, R. Siemens, & J. Unsworth, Eds.). Wiley. https://doi.org/10.1002/9781118680605

Underwood, T. (2019). *Distant horizons: Digital evidence and literary change*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226612973.001.0001

Vania, C., Ibrahim, M., & Adriani, M. (2014). Sentiment lexicon generation for an under-resourced language. *International Journal on Electrical Engineering and Informatics*, *6*(4), 469–480