

International Journal Research Publication Analysis

Page: 01-08

ARTIFICIAL INTELLIGENCE AND DIGITAL HUMANITIES: EMERGING PERSPECTIVES ON ENGLISH LITERARY RESEARCH

***Dr. A.K. Deepa Prem**

Associate Professor, Department of English, SKP Arts and Science College, Tiruvannamalai,
Tamil Nadu.

Article Received: 14 June 2026

Article Revised: 04 July 2026

Published on: 24 July 2026

***Corresponding Author: Dr. A.K. Deepa Prem**

Associate Professor, Department of English, SKP Arts and Science College,
Tiruvannamalai, Tamil Nadu.

DOI: <https://doi-doi.org/101555/ijrpa.8589>

ABSTRACT

The rapid advancement of Artificial Intelligence (AI) has begun to reshape the methodological landscape of the Humanities, opening new avenues within the interdisciplinary field of Digital Humanities (DH). This paper examines the emerging role of AI technologies including Natural Language Processing, machine learning classifiers, stylometric tools, and Large Language Models in English literary research. Employing a qualitative, literature-based analytical approach, the study surveys current applications of AI in distant reading, authorship attribution, sentiment and affect analysis, digital archiving, and literature pedagogy. It further maps the opportunities these tools present alongside persistent challenges of interpretive nuance, algorithmic bias, authenticity, and academic ethics. The paper argues for a balanced, hybrid model of literary scholarship in which AI serves as an augmentative instrument rather than a replacement for close reading and critical interpretation. Implications for curriculum design, research methodology, and future scholarship in English studies are discussed, and directions for further empirical research are proposed.

KEYWORDS: Artificial Intelligence; Digital Humanities; English Literature; Natural Language Processing; Distant Reading; Literary Research; Computational Criticism.

1. INTRODUCTION

Digital Humanities emerged in the late twentieth century as a field bridging computational methods and humanistic inquiry, and it has, over the past decade, been substantially reoriented by the rise of Artificial Intelligence. Where early DH scholarship relied on rule-

based text encoding and statistical word-counting, the current generation of AI systems encompassing machine learning, deep neural networks, and generative Large Language Models (LLMs) offers literary scholars tools capable of processing, classifying, and even generating text at a scale and speed unattainable through manual analysis alone.

For the discipline of English literary studies specifically, this shift raises both methodological opportunities and disciplinary anxieties. On one hand, AI-assisted approaches allow scholars to move fluidly between close reading of individual texts and distant reading of entire corpora, tracing stylistic, thematic, and generic patterns across centuries of literary production. On the other hand, questions persist about whether computational pattern-recognition can adequately capture the ambiguity, irony, and cultural specificity that define literary meaning.

This paper offers a structured review and conceptual mapping of how AI is currently being adopted within English literary research, synthesising recent scholarship to identify dominant application areas, associated opportunities, and unresolved challenges. It is intended as a foundation for scholars and postgraduate researchers seeking an entry point into this rapidly developing intersection.

2. Review of Literature

Foundational work in computational literary studies laid much of the groundwork for present-day AI applications. Moretti's (2013) concept of 'distant reading' proposed that literary history could be studied through the aggregate analysis of thousands of texts rather than close attention to a canonical few, a proposition later extended through corpus-scale stylistic modelling (Jockers, 2013) and algorithmically assisted criticism (Ramsay, 2011). Underwood (2019) demonstrated how supervised machine learning models could track the slow evolution of genre and literary style across two centuries of English fiction, offering empirical evidence for long-standing literary-historical claims.

More recent scholarship has turned to the specific affordances and risks of AI in text analysis. Yadav (2024) examined AI-assisted analysis of Shakespeare's Hamlet, finding that while AI tools efficiently surfaced thematic and stylistic patterns, they struggled to replicate the interpretive depth of human critical engagement with cultural and historical context. Parallel work on AI and digital humanities more broadly has highlighted concerns about bias, authenticity, and the ethics of algorithmic mediation in humanistic inquiry, alongside its evident gains in efficiency and scale.

A separate but related strand of research addresses the material infrastructure underlying AI-assisted literary study: the quality of digitised source texts. Hill and Hengchen (2019) demonstrated that inconsistencies in Optical Character Recognition (OCR) of historical texts can meaningfully distort downstream computational analysis, underscoring that AI-based literary research is only as reliable as the archival data feeding it. On the theoretical side, Hayles (2012) situates such tools within a broader account of human cognition as increasingly co-constituted with digital media a framing many DH scholars use to argue that AI does not simply assist literary interpretation but actively reshapes it.

Finally, concerns specific to generative AI have grown alongside the rise of LLMs. Huang et al. (2025) catalogue the tendency of LLMs toward 'hallucination' the generation of plausible but factually incorrect content a risk with direct implications for AI-assisted literary annotation and criticism, where fabricated interpretive claims could pass undetected. Institutional bodies have likewise begun responding to these developments: the Council of Europe (2024) issued a framework convention addressing the ethical and rights-based dimensions of AI adoption across public and academic life, while universities have begun introducing dedicated coursework, such as a graduate seminar on AI and Culture within English departments (St. Mary's University, 2025), reflecting the field's pedagogical adaptation.

3. Objectives of the Study

1. To examine the principal AI technologies currently applied in English literary research.
2. To map the specific applications of these technologies across literary sub-domains, including stylistics, authorship studies, and pedagogy.
3. To critically assess the opportunities and challenges AI presents for literary interpretation.
4. To propose a balanced framework for integrating AI tools within traditional literary scholarship.

4. RESEARCH METHODOLOGY

This study adopts a qualitative, descriptive-analytical research design based on a structured review of secondary sources. Peer-reviewed journal articles, conference proceedings, and institutional publications addressing AI and Digital Humanities, published predominantly between 2019 and 2025, were purposively selected and thematically analysed. Sources were drawn from academic databases and publisher repositories and were categorised according to

their primary area of application within English literary studies (see Figure 1). This method is appropriate for an emerging, conceptually unsettled field where the priority is synthesis and framework-building rather than statistical generalisation; it does not involve primary data collection from human subjects.

5. AI Applications in English Literary Research

5.1 Distant Reading and Stylometric Analysis

Building on Moretti's foundational proposition, contemporary stylometric tools use machine learning to quantify features such as sentence length, lexical density, and syntactic complexity across large corpora, enabling scholars to trace stylistic change across literary periods, movements, and individual authorial careers with a precision unavailable to unaided reading.

5.2 Authorship Attribution

Machine learning classifiers trained on stylistic 'fingerprints' have been applied to long-standing attribution debates, including disputed passages within the Shakespearean canon, by statistically comparing candidate texts against verified authorial baselines.

5.3 Sentiment and Affective Mapping

Sentiment analysis models allow researchers to chart the emotional trajectory of a narrative, offering quantitative support for qualitative claims about tonal shifts, tragic structure, or affective reader response in poetry and fiction.

5.4 Digital Archiving, OCR, and Corpus Building

AI-enhanced OCR and text-mining pipelines have expanded access to rare and fragile literary archives, though, as Hill and Hengchen (2019) caution, uncorrected recognition errors can introduce systematic noise into subsequent computational analysis.

5.5 AI in Literature Pedagogy

Generative AI tools are increasingly used in classroom settings as Socratic interlocutors, translation aids, and drafting partners, prompting English departments to design new coursework explicitly addressing AI's cultural and pedagogical implications.

6. AI Techniques and Their Applications

Table 1 summarises the principal AI techniques identified in the reviewed literature and their corresponding applications within English literary research.

Table 1: AI Techniques and Their Applications in English Literary Research.

AI Technique / Tool	Core Function	Application in English Literary Research
Natural Language Processing (NLP)	Parses syntax, semantics and discourse structure of text	Automated close reading, thematic tagging, discourse analysis of novels and drama
Stylometry / Distant Reading tools (e.g., Voyant, Stylo)	Quantifies stylistic features across large corpora	Tracking stylistic evolution, genre classification, canon-scale pattern detection
Machine Learning Classifiers	Learns patterns from labelled text data	Authorship attribution (e.g., disputed Shakespearean passages), genre prediction
Sentiment / Affect Analysis Models	Detects emotional tone and polarity in text	Mapping affective arcs in narrative fiction and poetry
Large Language Models (LLMs)	Generates and summarises text, answers queries on corpora	Draft annotation, translation assistance, Socratic tutoring in literature classrooms
OCR & Text-Mining Pipelines	Converts and cleans scanned archival material	Digitisation of rare manuscripts and periodicals for corpus-based study

7. Opportunities and Challenges

Table 2 synthesises the principal opportunities and challenges associated with AI adoption in literary scholarship, as identified across the reviewed literature.

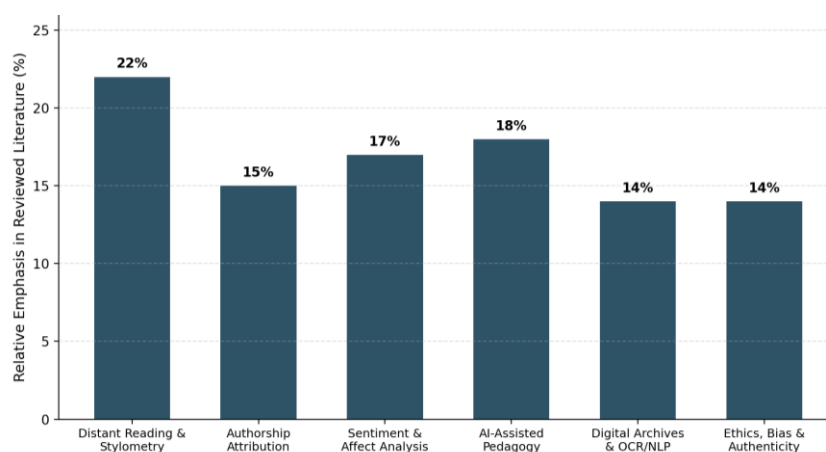
Table 2: Opportunities and Challenges of AI in English Literary Research.

Opportunities	Challenges
Scalable analysis of large literary corpora beyond manual capacity	Risk of reducing interpretive nuance to surface-level pattern matching
Support for comparative and cross-cultural literary studies	Bias embedded in training data affecting canon representation

Preservation and searchability of fragile archival texts	Questions of authenticity and authorship in AI-assisted writing
New pedagogical tools for student engagement with texts	Over-reliance risking decline in independent close-reading skills
Interdisciplinary bridges between literary theory and computational methods	Ethical and copyright concerns around training data and generated text

8. Thematic Distribution of Reviewed Literature

Figure 1 presents an illustrative classification of the literature surveyed in this study, organised by primary thematic focus. This classification is conceptual, reflecting the authors' categorisation of the reviewed sources for expository purposes rather than a formal bibliometric count.



9. DISCUSSION

The reviewed literature converges on a central tension: AI demonstrably extends the scale and efficiency of literary analysis, yet its outputs require sustained critical oversight from trained literary scholars. Studies applying AI to canonical texts consistently report strong performance in identifying surface-level thematic and stylistic patterns, but comparatively limited capacity to engage with irony, allusion, and culturally embedded meaning dimensions that remain central to literary interpretation. This suggests that AI is best conceptualised not as a substitute for critical reading but as a complementary instrument that can generate hypotheses, surface patterns, and manage archival scale, which human scholars then interpret, contextualise, and critique.

The risk of algorithmic 'hallucination' identified in generative AI research is particularly consequential for literary scholarship, where fabricated interpretive claims or misattributed

quotations could enter critical discourse without adequate verification. This reinforces the case for treating AI-generated analysis as a draft or provocation requiring scholarly verification, rather than as a finished critical product a principle equally relevant to the composition of academic papers using AI assistance, including work of the present kind.

10. CONCLUSION

Artificial Intelligence is reshaping the methodological toolkit available to English literary scholars, offering unprecedented scale in stylometric, archival, and affective analysis while simultaneously raising unresolved questions about interpretive depth, bias, and authenticity. This paper has mapped the principal areas of AI application in literary research and argued for a hybrid scholarly model in which computational tools augment, rather than replace, close reading and critical judgement. Future research should pursue empirical, corpus-based studies that directly compare AI-assisted and traditional critical readings of specific literary texts, alongside institutional guidelines for the ethical disclosure of AI use in literary scholarship and publishing.

REFERENCES

1. Council of Europe. (2024). The Council of Europe framework convention on artificial intelligence and explanatory report, human rights, democracy and the rule of law (Council of Europe Treaty Series No. 225).
2. Hayles, N. K. (2012). *How we think: Digital media and contemporary technogenesis*. University of Chicago Press.
3. Hill, M. J., & Hengchen, S. (2019). Quantifying the impact of dirty OCR on historical text analysis: Eighteenth Century Collections Online as a case study. *Digital Scholarship in the Humanities*, 34(4), 825–843.
4. Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 43(2).
5. Jockers, M. L. (2013). *Macroanalysis: Digital methods and literary history*. University of Illinois Press.
6. Moretti, F. (2013). *Distant reading*. Verso Books.
7. Ramsay, S. (2011). *Reading machines: Toward an algorithmic criticism*. University of Illinois Press.

8. St. Mary's University. (2025). Novel English class explores AI and the digital humanities. <https://www.stmarytx.edu/2025/digital-humanities/>
9. Underwood, T. (2019). Distant horizons: Digital evidence and literary change. University of Chicago Press.
10. Yadav, D. (2024). The role of artificial intelligence in literary analysis. International Journal of English, Knowledge and Skills, 3(9). <https://ijeks.com>