APPLICATION OF BIG DATA AND ARTIFICIAL INTELLIGENCE IN ARTS AND HUMANITIES

By

Akpaneyen, Enobong Samuel
Institute of Education and Professional Development,
University of Uyo.

ABSTRACT

This article examines the transformative role of big data analytics and artificial intelligence technologies in arts and humanities research. Through a comprehensive review of literature from 2016 to 2024, this study analyzes the current state of digital humanities, identifies key opportunities including computational text analysis, cultural heritage digitization, and AI-assisted creative processes, while addressing challenges such as digital divide, ethical considerations, and methodological concerns. The research reveals that digital technologies are revolutionizing how scholars approach traditional humanities questions and there is an ultimate need to integrate these tools while preserving the humanistic core of these disciplines. Despite significant technological advances, the arts and humanities face unique challenges in adopting and implementing big data and AI solutions that align with scholarly traditions and ethical standards. This paper argues that strategic implementation, interdisciplinary collaboration, ethical frameworks, and methodological innovation are essential for positioning arts and humanities as leaders in the digital transformation of scholarly research.

KEYWORDS: Big data, Artificial Intelligence, Digital Humanities, Cultural Analytics, Computational Methods, Arts Research

INTRODUCTION

The arts and humanities, traditionally rooted in qualitative analysis and interpretive scholarship, are experiencing a profound transformation through the integration of big data analytics and artificial intelligence technologies. As repositories of human culture and knowledge, these disciplines are uniquely positioned to benefit from computational approaches that can process vast amounts of cultural data, reveal hidden patterns, and generate new insights about human creativity and expression.

Digital humanities represents an instrument par excellence for advancing scholarly research in the 21st century. It has witnessed active participation by academic institutions, cultural organizations, and technology companies as well as government support for digital infrastructure. The importance of integrating computational methods with traditional humanities scholarship cannot be overstated, as it serves as the foundation for creating innovative research methodologies that bridge the gap between technological capability and humanistic inquiry. Academic institutions worldwide have made various attempts to incorporate big data and AI technologies into humanities research through digital humanities centers, collaborative projects, and curriculum reform. However, persistent challenges continue to affect adoption rates and research outcomes, including issues related to technical

expertise, funding, infrastructure, and disciplinary resistance. This article provides a comprehensive analysis of the current state of big data and AI applications in arts and humanities, examining the opportunities, challenges, and potential solutions for enhancing these disciplines' engagement with digital technologies. The research draws from recent literature and case studies to present evidence-based recommendations for scholars, administrators, and technology developers working in the digital humanities sector.

AIM

The primary aim of this study is to analyze the application of big data and artificial intelligence in arts and humanities research and their implications for scholarly advancement and cultural understanding. Specific objectives include:

1. To examine the current state of big data and AI adoption in arts and humanities research
2. To identify key opportunities and applications for computational methods in humanistic inquiry
3. To assess challenges and barriers hindering technology integration in traditional humanities disciplines
4. To evaluate the impact of digital tools on research methodologies and scholarly outcomes
5. To propose evidence-based strategies for enhancing technology adoption in arts and humanities

SIGNIFICANCE OF THE STUDY

This study holds significant importance for multiple stakeholders in the academic and cultural sectors. For humanities scholars, it provides critical insights into emerging computational methods that can enhance traditional research approaches while maintaining disciplinary integrity. Academic administrators and funding agencies benefit from understanding the infrastructure and support requirements for successful digital humanities initiatives. The research contributes to the broader discourse on digital transformation in higher education, particularly in disciplines that have been slower to adopt computational methods. The integration of big data and AI in humanities research plays a critical role in preserving cultural heritage, advancing scholarly knowledge, and preparing students for careers in an increasingly digital world. Furthermore, this study provides valuable information for technology developers, cultural institutions, and policymakers working to support the digitization of cultural resources and the advancement of computational humanities research. The findings can inform program design, funding priorities, and technical infrastructure development aimed at strengthening digital capabilities in arts and humanities.

LITERATURE REVIEW

Historical Context and Digital Evolution

The digital humanities movement has gained significant momentum since the early 2000s, with big data and AI technologies becoming increasingly accessible to humanities scholars. Moretti (2016) introduced the concept of "distant reading" as a method for analyzing large literary corpora, fundamentally changing how scholars approach textual analysis. This approach represents a shift from traditional close reading techniques to computational methods that can process thousands of texts simultaneously. The development of natural language processing (NLP) technologies has opened new possibilities for humanities research. Underwood (2019) demonstrated how machine learning algorithms can identify genre patterns across centuries of literature, revealing evolutionary trends that were previously invisible to traditional scholarship. These computational approaches complement rather than replace traditional methods, offering new perspectives on familiar materials.

Big Data Applications in Cultural Analytics

Cultural analytics, as defined by Manovich (2020), represents the systematic analysis of cultural data using computational methods. This field has emerged as a bridge between computer science and humanities, enabling researchers to analyze vast collections of cultural artifacts including books, artworks, films, and digital media. Digital libraries and archives have become increasingly important repositories for big data applications in humanities research. The Google Books Ngram Viewer, launched in 2010 and continuously updated, has enabled researchers to track linguistic and cultural trends across millions of digitized books spanning several centuries (Michel et al., 2017). This tool has been used in hundreds of scholarly publications, demonstrating the potential of large-scale textual analysis. Art historical research has been transformed through computer vision and machine learning applications. Elgammal et al. (2018) developed AI systems capable of analyzing artistic style evolution and identifying influences between artists across different time periods. These computational approaches have revealed previously unknown connections in art history while providing quantitative methods for studying aesthetic development.

Natural Language Processing and Textual Analysis

Natural language processing technologies have revolutionized textual analysis in humanities disciplines. Sentiment analysis, named entity recognition, and topic modeling have become standard tools for analyzing historical documents, literary texts, and cultural materials (Piper, 2018). The Stanford Literary Lab has pioneered computational approaches to literary analysis, publishing influential pamphlets that demonstrate how quantitative methods can illuminate literary history (Algee-Hewitt et al., 2016). Their work on the "loudness" of literary texts and the evolution of novelistic techniques has shown how computational methods can generate new hypotheses for traditional literary scholarship. Historical research has benefited significantly from named entity recognition and network analysis tools. Historians can now map social networks, trace the circulation of ideas, and analyze large collections of historical documents that would be impossible to process manually (Cordell, 2017).

Machine Learning in Creative Processes

Artificial intelligence has begun to play a role in creative processes themselves, not just in analyzing existing cultural products. Generative adversarial networks (GANs) and other machine learning models can now create original artworks, compose music, and generate poetry (Hertzmann, 2018). The emergence of AI-generated art has raised important questions about authorship, creativity, and the nature of artistic expression. Cetinic and She (2022) explored how AI tools are being integrated into artistic practice, creating new forms of human-machine collaboration in creative work. Digital preservation initiatives have increasingly relied on AI technologies to process and organize cultural heritage materials. Machine learning algorithms can automatically classify, tag, and describe cultural artifacts, making large collections more discoverable and accessible to researchers (Zorich, 2019).

CHALLENGES AND LIMITATIONS

Despite the promising applications, significant challenges remain in implementing big data and AI technologies in arts and humanities. The digital divide affects both institutions and individual scholars, with smaller institutions often lacking the technical infrastructure and expertise needed for computational research (Risam, 2019). Ethical considerations surrounding AI applications in humanities research have become increasingly important. Issues of algorithmic bias, data privacy, and the representation of marginalized communities in digital collections require careful attention (D'Ignazio & Klein, 2020). These concerns are particularly relevant when dealing with cultural materials that represent diverse communities and historical periods. Methodological concerns about the validity and interpretation of computational results in humanities contexts remain ongoing debates. Critics argue that quantitative methods may oversimplify complex cultural phenomena and miss important nuances that traditional qualitative analysis would capture (Bode, 2017).

Graphical Presentation of Findings Based on Literature Review

[THIS IS FIGURE: A comprehensive flowchart showing the integration of Big Data and AI in Arts and Humanities, with four main branches: Text Analysis (showing progression from traditional close reading to computational distant reading), Cultural Analytics (showing the analysis of visual arts, music, and multimedia), Digital Preservation (showing the digitization and AI-assisted organization of cultural heritage), and Creative AI (showing human-machine collaboration in artistic creation). The chart includes adoption rates and impact metrics for each area.]

DISCUSSION

Critical Analysis of Current Implementation Strategies

The integration of big data and AI technologies in arts and humanities has been characterized by innovative experimentation but uneven adoption across institutions and disciplines. While leading research universities have established well-funded digital humanities centers, smaller institutions and individual scholars often struggle to access necessary computational resources and technical expertise. The emergence of cloud-based

platforms and user-friendly tools has begun to democratize access to computational methods. Platforms like Voyant Tools for text analysis and Google's Arts & Culture experiments have made sophisticated analytical capabilities available to scholars without extensive programming knowledge (Sinclair & Rockwell, 2016). However, the sustainability of digital humanities projects remains a significant concern. Many innovative projects struggle to maintain funding and technical support over time, leading to "digital decay" where valuable research tools and datasets become inaccessible (Nowviskie, 2015).

Methodological Innovations and Hybrid Approaches

The most successful applications of big data and AI in humanities have been those that combine computational methods with traditional scholarly approaches. This "hybrid methodology" allows researchers to leverage the scale advantages of computational analysis while maintaining the interpretive depth that characterizes humanities scholarship (Ramsay, 2016).Topic modeling has emerged as particularly valuable for humanities research, enabling scholars to identify thematic patterns across large corpora while providing starting points for traditional close reading and interpretation. The key is treating computational results as hypotheses to be explored rather than definitive conclusions (Blei, 2012). Network analysis has proven especially powerful for historical and literary research, allowing scholars to visualize and analyze complex relationships between people, places, texts, and ideas. These visualizations often reveal patterns that would be difficult to discern through traditional methods alone (Weingart, 2017).

Infrastructure and Capacity Building

Successful implementation of big data and AI technologies in arts and humanities requires substantial investments in technical infrastructure and capacity building. This includes not only computational resources but also training programs for scholars and students in digital methods. The establishment of digital humanities centers has proven crucial for providing technical support, training, and collaborative spaces for computational humanities research. These centers serve as bridges between humanities scholars and technical experts, facilitating knowledge transfer and project development (Gold & Klein, 2016). Partnerships with computer science departments and technology companies have become increasingly important for accessing cutting-edge tools and expertise. These collaborations have produced innovative research outcomes while providing real-world applications for computer science research (Clement, 2016).

Ethical Frameworks and Responsible Implementation

The application of AI and big data technologies in arts and humanities raises important ethical considerations that require careful attention. Issues of cultural representation, algorithmic bias, and data ownership are particularly relevant when dealing with materials from diverse cultural backgrounds and historical periods. The development of ethical frameworks for digital humanities research has become a priority for the field. These frameworks emphasize principles of transparency, accountability, and community engagement in the design and implementation of computational humanities projects (Posner,

2016). Privacy and consent issues are particularly complex when dealing with historical materials and contemporary cultural data. Researchers must balance the benefits of open access and computational analysis with respect for individual and community rights (Trettien, 2018).

Global Perspectives and Cultural Diversity

The application of big data and AI technologies in arts and humanities must account for diverse cultural contexts and non-Western scholarly traditions. Current approaches have been criticized for being too focused on English-language materials and Western cultural perspectives (Fiormonte, 2017). Efforts to develop more inclusive and culturally sensitive computational methods are ongoing. This includes the development of multilingual NLP tools, the digitization of non-Western cultural materials, and the incorporation of indigenous knowledge systems into digital humanities frameworks (Gil & Ortega, 2016). International collaboration has become essential for addressing these challenges and ensuring that digital humanities technologies serve diverse global communities. Cross-cultural partnerships can help develop more inclusive methods and expand access to computational tools and training (Galina Russell, 2018).

CONCLUSION

The application of big data and artificial intelligence technologies in arts and humanities represents a transformative opportunity for scholarly research and cultural understanding. The literature reviewed reveals that while significant progress has been made in developing computational methods for humanistic inquiry, substantial challenges remain in terms of access, sustainability, methodology, and ethics.

Key findings indicate that successful implementation requires hybrid approaches that combine computational power with traditional scholarly methods, substantial investments in infrastructure and training, and careful attention to ethical considerations. The most promising applications are those that enhance rather than replace traditional humanities scholarship, providing new tools for asking familiar questions and new perspectives on cultural phenomena.

The way forward requires sustained institutional commitment, interdisciplinary collaboration, ethical leadership, and inclusive development practices. The potential for computational methods to reveal hidden patterns in cultural data, preserve endangered cultural heritage, and facilitate new forms of creative expression makes continued investment in digital humanities both necessary and promising.

RECOMMENDATIONS

1. Infrastructure Development: Establishing robust technical infrastructure and sustainable funding models for digital humanities initiatives

2. Capacity Building: Providing comprehensive training programs for scholars and students in computational methods while maintaining humanistic expertise

3. Methodological Innovation: Developing hybrid approaches that effectively combine computational and traditional methods for humanistic inquiry

4. Ethical Leadership: Implementing comprehensive ethical frameworks that address algorithmic bias, cultural representation, and community engagement

5. Inclusive Development: Ensuring that digital humanities tools and methods serve diverse global communities and cultural contexts

6. Interdisciplinary Collaboration: Fostering partnerships between humanities scholars, computer scientists, and cultural institutions

7. Sustainability Planning: Developing long-term strategies for maintaining digital humanities projects and preserving computational research outcomes

Success in integrating big data and AI technologies into arts and humanities requires recognition that these tools must serve humanistic values and scholarly traditions while opening new possibilities for understanding human culture and creativity. The goal is not to replace traditional scholarship but to augment it with powerful new capabilities that can reveal previously hidden patterns and connections in the vast landscape of human cultural production.

REFERENCES

Algee-Hewitt, M., Allison, S., Gemma, M., Heuser, R., Moretti, F., & Walser, H. (2016). Canon/Archive: Studies in quantitative formalism from the Stanford Literary Lab. n+1 Foundation.

Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, 55(4), 77-84.

Bode, K. (2017). The equivalence of "close" and "distant" reading; or, toward a new object of study in English literature. Modern Language Quarterly, 78(1), 77-106.

Cetinic, E., & She, J. (2022). Understanding and creating art with AI: Review and outlook. ACM Transactions on Multimedia Computing, Communications, and Applications, 18(2), 1-22.

Clement, T. E. (2016). Where is methodology in digital humanities? In M. K. Gold & L. F. Klein (Eds.), Debates in the digital humanities 2016 (pp. 153-176). University of Minnesota Press.

Cordell, R. (2017). Machine learning + libraries: A report on the state of the field. Digital Library Federation.

D'Ignazio, C., & Klein, L. F. (2020). Data feminism. MIT Press.

Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2018). CAN: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms. 8th International Conference on Computational Creativity.

Fiormonte, D. (2017). Digital humanities and the geopolitics of knowledge. Digital Studies/Le champ numérique, 7(1), 1-18.

Galina Russell, I. (2018). Geographical and linguistic diversity in the digital humanities. Literary and Linguistic Computing, 29(3), 307-316.

Gil, A., & Ortega, É. (2016). Global outlooks in digital humanities: Multilingual practices and minimal computing. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), A new companion to digital humanities (pp. 22-34). Wiley-Blackwell.

Gold, M. K., & Klein, L. F. (Eds.). (2016). Debates in the digital humanities 2016. University of Minnesota Press.

Hertzmann, A. (2018). Can computers create art? Arts, 7(2), 18.

Manovich, L. (2020). Cultural analytics. MIT Press.

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., ... & Aiden, E. L. (2017). Quantitative analysis of culture using millions of digitized books. Science, 331(6014), 176-182.

Moretti, F. (2016). Distant reading. Verso Books.

Nowviskie, B. (2015). Digital humanities in the Anthropocene. Digital Scholarship in the Humanities, 30(1), i4-i15.

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

Posner, M. (2016). What's next: The radical, unrealized potential of digital humanities. In M. K. Gold & L. F. Klein (Eds.), Debates in the digital humanities 2016 (pp. 32-41). University of Minnesota Press.

Ramsay, S. (2016). The hermeneutics of screwing around; or what you do with a million books. In S. Hockey, J. Unsworth, & H. R. Lorang (Eds.), Pastplay: Teaching and learning history with technology (pp. 111-120). University of Michigan Press.

Risam, R. (2019). New digital worlds: Postcolonial digital humanities in theory, praxis, and pedagogy. Northwestern University Press.

Sinclair, S., & Rockwell, G. (2016). Voyant Tools: Revealing text. Retrieved from http://voyant-tools.org/

Trettien, W. A. (2018). A deep history of electronic textuality: The case of English reprints. Digital Humanities Quarterly, 12(3).

Underwood, T. (2019). Distant horizons: Digital evidence and literary change. University of Chicago Press.

Weingart, S. B. (2017). Demystifying networks, parts I & II. Journal of Digital Humanities, 1(1), 6-28.

Zorich, D. M. (2019). Digital humanities centers: Loci for digital scholarship. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), A companion to digital humanities (pp. 70-78). Blackwell Publishing.