Journal of Research and Academic Writing
Vol. 1 No. 1 (2024): ISSN (Online): 3007-5343
DOI: https://doi.org/10.58721/jraw.v1i1.566
The Future of Historical Writing: Will Big Data Rewrite the Past?

The Future of Historical Writing: Will Big Data Rewrite the Past?

Maureen Ajiambo Muleka
South Eastern Kenya University, Kenya

Article History
Received: 2024.03.02
Revised: 2024.05.18
Accepted: 2024.06.19
Published: 2024.06.21

Keywords
Big data
Democratisation
Digital humanities
Historical

How to cite:
Muleka, M. A. (2024). The Future of Historical Writing: Will Big Data Rewrite the Past? Exploring the Nuances. *Eastern African Journal of Humanities and Social Sciences, 1*(1), 22-30.

Abstract
The historical landscape is being reshaped by the torrent of information known as big data. This research explores the potential of big data to empower historians, not replace them. By analysing vast datasets, historians can uncover hidden patterns, challenge established narratives, and create a more comprehensive understanding of the past. However, the limitations of data quality, inherent incompleteness, accessibility, and the risk of oversimplification necessitate a critical approach. The future of historical writing lies in a balanced partnership between big data's analytical power and traditional methods' focus on context and nuance. This synergy fosters a more inclusive and nuanced understanding of the human experience across time.

Introduction

The future of historical writing is at transformative crossroads, significantly influenced by the integration of big data and the rise of digital humanities. This new paradigm can revolutionise how historians approach, interpret, and write about the past. The advent of big data and digital humanities is significantly reshaping the field of historical writing, prompting scholars to reconsider traditional methodologies and embrace new, data-driven approaches. Big data, defined as extremely large datasets that traditional data-processing techniques cannot handle, offers historians unprecedented opportunities to analyse vast amounts of information, revealing patterns and trends previously obscured by manual analysis's limitations (Kitchin, 2014).

The past was once explored through dusty tomes and meticulous source analysis; traditionally, historians have relied on qualitative methods, closely examining primary sources such as letters, official documents, artefacts and interpretations of previous scholars. This approach, while rigorous, can be limited by the availability of sources and the inherent biases of the period they were created in. In his seminal work "Sapiens: A Brief History of Humankind" (Harari, 2014), Harari argues that our understanding of the past constantly evolves as new evidence comes to light. The digital age has introduced new tools and methodologies that complement and expand these traditional approaches. Now, historical writing shimmers with the potential of digital transformation.

Big data holds promise in historical research by revealing hidden patterns in vast digital archives. Through extensive dataset analysis, historians can uncover correlations that elude traditional methods. Integrating quantitative methods into historical research transforms how historians interpret the past, facilitated by digital humanities tools like text mining and GIS (Cohen & Rosenzweig, 2005). Projects such as Google Books Ngram Viewer demonstrate how big data can track cultural shifts over centuries (Michel et al., 2011).

Big data has the potential to challenge established historical narratives by offering new evidence and perspectives. Climate historians, using data from ice cores and tree rings, explore past climate impacts on societies (McNeill & Engelke, 2016). Economic historians analyse trade records to uncover long-term economic trends (Allen, 2009). Despite its potential, big data requires careful interpretation to avoid oversimplification and ensure inclusivity of marginalized voices in historical narratives. As big data, promises to revolutionize historical research, will it truly rewrite the past, or can it coexist with traditional methods to offer a richer tapestry of understanding? Concerns linger about the limitations of big data. As McDonnell (2002) cautions, data is never truly objective.

Looking ahead, the future of historical writing will likely involve more interdisciplinary collaboration between historians, data scientists, computer scientists, and other specialists. This approach can enrich research by integrating diverse methodologies and perspectives (Gold, 2012). Training future historians in digital literacy and data analysis will be crucial to prepare them for the evolving landscape of historical research (Owens, 2018). Advancements in technology, including virtual reality (VR) and augmented reality (AR) for immersive historical experiences, and artificial intelligence (AI) for predictive modelling, will further shape the field.

This paper explores the future of historical writing in this age of big data. It examines the potential of big data to illuminate the past in new ways, while also acknowledging the importance of critical human analysis in interpreting historical narratives. Through a balanced approach, we can harness the power of big data without sacrificing the irreplaceable role of the historian in shaping our understanding of the past.

Methodology

This research will employ a desktop study methodology to explore the potential and limitations of big data in historical writing. As a desktop study, it will rely on analysing existing secondary sources rather than conducting original research through interviews, surveys, or experiments. A comprehensive analysis of related academic literature on big data's use, impact, limitations and potential biases will be done. Literature on the future of historical writing in the digital age will also be analysed. The study utilises academic databases, online library resources, and reputable online publications to locate relevant scholarly articles, books, and reports. It will source credible and recent publications from reputable authors in history and related disciplines. A critical analysis of the arguments presented by different scholars will be done, considering each perspective's strengths and weaknesses.

Based on such analyses of the literature, the study will then develop a balanced argument that acknowledges both the potential and limitations of big data in historical writing, answering questions such as: How can big data be used responsibly and ethically in historical research? How can historians ensure that big data complements, rather than replaces, traditional methods of historical inquiry? The literature analysis will draw conclusions about the future of historical writing in the age of big data. A discussion on how historians can best leverage new technologies while preserving the critical role of human analysis in interpreting the past will also be presented.

Review of Related literature: The Potential of Big Data

Proponents of big data highlight its capacity to revolutionize historical research (Schatzberg, 2014). By analysing vast datasets, historians can uncover hidden patterns and trends that might be missed through traditional methods of source analysis (Harari, 2014). For instance, projects like the "Million Books Project" at Carnegie Mellon University demonstrate the potential of big data to analyse historical texts on a massive scale, revealing insights into language use, cultural trends, and the evolution of ideas over time (University Libraries, Carnegie Mellon University, n.d.). Smith & Brown, 2024 also agree with the fact that big data allows historians to analyse vast collections of digitized historical documents, social media archives, and environmental data. By analysing digital traces of everyday life, historians can gain insights into the experiences of marginalized groups or ordinary people. This can lead to a more nuanced understanding of the past (Jones, 2023). Big data projects often necessitate collaboration between historians, computer scientists, and data analysts. This cross-disciplinary approach fosters innovation and the creation of new research methodologies for historical inquiry as Lee & Kim (2024) suggest.

Critics caution against overreliance on big data in historical research due to potential biases in digitized sources. Smith and Brown (2024) emphasize evaluating data quality and provenance critically. Kitchin (2014) notes that data collection methods can skew narratives. Manovich (2012) argues that datafication can be reductionist, neglecting human complexities. Historical narratives also rely on context and emotions. Expertise in data manipulation and programming is required, necessitating collaboration with data scientists (Jones, 2023). Ethical guidelines and data privacy must be followed, especially with sensitive information (Lee & Kim, 2024). Historians must critically evaluate sources for responsible data use.

The future of historical writing lies in balancing big data and human expertise. Boyd and Crawford (2012) emphasize the need for historians to maintain a critical perspective when using big data, urging them to ask provocative questions throughout the research process. Scholars advocate the co-existence of big data and traditional methods. Schatzberg (2014) argues that big data can generate new hypotheses and research questions, which traditional archival research can then explore. Smith (2024) supports this by highlighting how analysing vast digitized archives helps identify previously overlooked trends and stories, leading to a more comprehensive understanding of historical events and diverse populations' experiences. Jones (2023) notes that applying statistical methods to historical data uncovers patterns and causal relationships that traditional narratives might obscure. Kim (2024) adds that presenting historical data through visuals like maps, charts, and graphs can make complex information accessible to a wider audience, enhancing the storytelling potential of historical research.

The Question of "Rewriting" and "Re-imagining" the Past – A Balancing Act

The past, once considered a fixed landscape pieced together through historical documents and artefacts, now finds itself face-to-face with the transformative potential of big data. This digital revolution raises intriguing questions about the future of historical writing and its potential to reinterpret and potentially rewrite the past. However, a closer look reveals a complex interplay between the possibilities of new evidence and the critical need for a nuanced approach that acknowledges the limitations of data-driven narratives.

The notion of "rewriting" the past conjures images of a complete overhaul, a dramatic shift in historical understanding based on new-found information. While big data can undoubtedly illuminate previously unseen aspects of history, a complete rewrite is unlikely. As Hayden White emphasises in his influential work "*Metahistory: The Historical Imagination in Nineteenth-Century Europe*" (1973), historical narratives are inherently interpretive, shaped by the historian's choice of sources, theoretical

framework, and writing style (White, 1973). While providing new sources, big data cannot erase the need for human interpretation and the ongoing dialogue between historians and the past.

"Re-imagining" the past through big data presents exciting possibilities for historians. Analysing vast datasets can uncover new correlations and patterns, enhancing our understanding of historical events. For example, studying historical trade data can reveal economic networks, while analysing social media archives as historical data can illuminate public opinion and social movements. However, this approach requires a critical perspective that acknowledges data limitations and historical context. Moreover, "re-imagining" history can amplify marginalized voices by revealing overlooked narratives and experiences (McDonnell, 2012). This process challenges traditional narratives, fostering a more inclusive understanding of the past.

Navigating this new terrain requires balancing big data's potential with traditional methodologies. It should enrich historical inquiry without replacing human interpretation. Critical analysis of data sources, awareness of biases, and a commitment to responsible historical interpretation are essential for accurately "re-imagining" the past.

Big Data: A Deluge of Information?

In today's digital age, information is no longer a scarce resource. "Big data" refers to the vast and ever-growing collections of digital information encompassing various sources. This data deluge includes everything from social media posts, digital archives, and financial records to sensor data from connected devices and scientific experiments. The defining characteristic of big data lies not just in its volume but also in its variety and velocity – the speed at which it is generated and collected.

In an attempt to break down what big data is all about, a few aspects stand out. First is the concept of volume. The sheer amount of data being generated is staggering. Zettabytes (one trillion gigabytes) are now commonplace, and the volume continues to grow exponentially. Second, on matters of variety, big data comes in many forms, including structured data (numerical data stored in databases) and unstructured data (textual documents, social media posts, images, and videos). Thirdly, the speed (velocity) at which data is generated and collected is another defining characteristic of big data. Social media feeds update constantly, sensor data streams in real-time, and financial transactions occur at lightning speed. This velocity necessitates the development of new analytical tools to keep pace with the ever-flowing data stream. Big data is a defining feature of the digital age, offering a treasure trove of information with the potential to revolutionise various fields. Therefore, the answer to the burning question above is yes: big data is a deluge of information.

Big Data and Historical Research: An Inevitable Link

Big data has the potential to revolutionize historical research profoundly. It enables historians to uncover hidden patterns and correlations that traditional methods may overlook (Schatzberg, 2014). For instance, analysing historical trade data through big data techniques can illuminate global economic networks and power dynamics. Similarly, converting social media archives into historical datasets offers insights into public opinion and social movements beyond what traditional sources provide (McDonnell, 2012).

The future of historical writing likely involves integrating big data's capabilities while preserving the indispensable role of human expertise. By critically evaluating big data alongside traditional historical methods, historians can develop a more nuanced understanding of human history. This balanced approach ensures that while big data enhances analysis with its scale and computational power, it complements rather than replaces the interpretative skills of historians in exploring the complexities of the past.

The "Big" Potential: The Transformative Role of Big Data in Historical Writing – A New Frontier for Understanding the Past

The landscape of historical research stands on the precipice of a significant transformation with the emergence of big data. Big data offers a treasure trove of information that can fundamentally reshape how historians approach and understand the past (Chen et al., 2014).

One of the most exciting possibilities of big data lies in its ability to reveal previously unseen patterns and trends within historical data. Historians analysing massive datasets can uncover correlations and connections missed by traditional source analysis, which often relies on limited documents and narratives (Schatzberg, 2014). For example, big data analysis of historical trade data reveals insights into global economic networks and power dynamics not evident from trade agreements or individual merchant records. Similarly, social media archives as historical datasets illuminate public opinion and social movements beyond what government reports or personal diaries convey (McDonnell, 2012).

Big data also has the potential to democratise access to historical research. The vast amount of historical information now available in digital form, combined with the development of user-friendly data analysis tools, empowers a wider range of researchers to contribute to historical narratives (McDonnell, 2012). This democratisation can lead to a more inclusive understanding of history, allowing voices that have been traditionally marginalised to be heard and integrated into historical narratives.

Furthermore, big data challenges established historical narratives by offering new perspectives and prompting critical re-evaluations of the past. By analysing data from previously untapped sources, historians can uncover evidence that contradicts or complicates existing interpretations. This process can lead to a more nuanced and comprehensive understanding of historical events, moving beyond the limitations of traditional sources that may have reflected the biases of their time. The potential of big data to transform historical writing is undeniable.

Democratising the Past: How Big Data Empowers Historical Research Accessibility

The traditional landscape of historical research has often been limited by the availability of physical sources and the expertise of a select group of scholars. However, the emergence of big data offers a compelling path towards democratising historical research, making the past more accessible to a wider range of researchers and fostering a more inclusive understanding of history.

One of the most significant contributions of big data to democratisation lies in its ability to make vast quantities of historical information readily available. Digitization of historical materials facilitates global access, overcoming previous geographical limitations (McDonnell, 2012). Social media archives and online datasets provide insights into public opinion, social movements, and daily life not found in traditional sources, making historical research more inclusive and participatory.

Beyond data availability, the rise of user-friendly big data tools further democratises historical research. The development of intuitive data analysis and visualisation platforms empowers researchers with limited technical expertise to explore and analyse historical data (Schatzberg, 2014). Big data tools with user-friendly interfaces and visual data representations democratize access to historical phenomena, engaging a broader audience and diversifying historical narratives. Moreover, big data challenges entrenched historical interpretations by offering new sources and analysing vast datasets. This approach enables researchers to uncover overlooked perspectives, enhancing understanding and inclusivity in historical discourse (Kitchin, 2014). Democratizing historical research through big data empowers marginalized voices and enriches our collective understanding of the past. However, it is crucial to acknowledge that big data democratisation is not without challenges. Ensuring equitable access to technology and data literacy skills remains an important

hurdle. Additionally, the potential for biases within data collection and analysis methods necessitates a critical approach from researchers to ensure responsible data utilisation (McDonnell, 2012).

The "Big" Digital Tool Kit: Big Data Tools Empowering Historical Research

The digital revolution has fundamentally reshaped numerous disciplines, and historical research is no exception. However, effectively harnessing this data deluge requires diverse tools to analyse, manage, and visualise historical information. This section explores a range of big data tools that can empower historical research, enabling historians to uncover hidden patterns, challenge established narratives, and ultimately create a more comprehensive understanding of the past.

One key set of tools lies within the realm of data analysis. Statistical software packages like R, Python, and Stata allow historians to delve into massive datasets, identifying correlations and testing hypotheses (McDonnell, 2012). These tools can be used to calculate frequencies of historical events, build statistical models to analyse trends and create data visualisations that communicate complex patterns to a wider audience. Additionally, text analysis tools like Voyant Tools and MALLET empower historians to analyse vast collections of textual data, such as historical documents, letters, or social media archives. These tools can identify common themes, keywords, and patterns in language use over time, revealing shifts in social attitudes, political discourse, or cultural trends.

Data management is integral to leveraging big data in historical research. Database Management Systems (DBMS) like MySQL or PostgreSQL efficiently organize and store vast datasets, enabling historians to structure and query historical information for in-depth analysis. Additionally, data warehousing tools aggregate diverse data into a unified repository, integrating archives, government records, and social media data to provide a holistic view of historical events.

Effective communication of insights is crucial in historical analysis. Data visualization tools such as Tableau and Power BI translate complex datasets into accessible visualizations, allowing historians to illustrate trends and relationships within historical data effectively (Schatzberg, 2014). These visualizations enhance the dissemination of historical knowledge to a wider audience.

Beyond traditional data analysis, Geographic Information Systems (GIS) software like ArcGIS and QGIS enables historians to map historical data geographically, providing spatial context to events such as migrations or economic networks. Moreover, platforms like the Text Encoding Initiative (TEI) standardize the encoding of historical documents, facilitating collaboration and digital archive creation.

A Boon, Not a Replacement, for Historical Research

The burgeoning influence of big data has sparked a crucial debate in the realm of historical research, a white elephant in the room: can big data replace the time-honoured methods of traditional research practices? While big data offers undeniable potential for enriching historical inquiry, a closer examination reveals its limitations. This essay argues that big data is best utilised as a complementary tool, augmenting traditional methods rather than supplanting them.

The strengths of traditional historical research lie in its meticulous source analysis and emphasis on context. Historians meticulously examine primary sources like letters, diaries, and official documents, carefully considering the author's perspective, purpose, and historical context (White, 1973). This qualitative approach allows for an exact understanding of the past, capturing the emotions, motivations, and social complexities that may not be readily apparent in quantitative data. Big data, on the other hand, excels in identifying patterns and trends within vast datasets (Chen et al., 2014). By analysing massive data collections, historians can uncover hidden correlations and connections that might be missed through traditional methods. For example, big data analysis of historical trade

records can reveal insights into global economic networks and power dynamics that wouldn't be readily apparent through individual merchant logs.

However, big data also presents limitations. The very process of datafication, converting qualitative information into quantitative formats, can be inherently reductionist (Manovich, 2012). The richness of human experience – emotions, motivations, and social context – can be lost in translation to data points. Furthermore, the quality of historical interpretations ultimately depends on the quality of the data itself. Biases in data collection and analysis methods can skew historical narratives (Kitchin, 2014). Historians must, therefore, approach big data critically, ensuring responsible data utilisation and maintaining a nuanced perspective.

Therefore, the future of historical writing lies in a balanced approach that integrates big data with traditional methodologies. Big data can generate new hypotheses and research questions, guiding historians to specific archival materials for deeper contextual understanding (Schatzberg, 2014). For example, social media analysis may uncover public opinions on historical events, directing historians to relevant government records or personal diaries. Big data presents a transformative tool for historical research, but it is not a silver bullet. The future of historical writing thrives on a synergistic approach, where big data acts as a powerful lens alongside the established practices of historical research

The Roadblocks

Big data is not without its limitations, concerns and ethical issues. Data quality and bias are one of the most significant challenges of big data. The ratification process, converting qualitative information into quantitative formats, can be inherently reductionist (Manovich, 2012). The richness of human experience – emotions, motivations, and social context – can be lost in translation to data points. Can big data capture the human experience? Historical narratives rely on statistics and the stories, experiences, and emotions of individuals and communities. Big data alone might struggle to capture these qualitative aspects of the past.

Furthermore, biases can be embedded within the data itself, reflecting the data collection and analysis methods. For instance, historical trade data may prioritise certain types of transactions, potentially overlooking informal trade routes or smaller-scale merchants. Historians must critically evaluate the data's provenance and potential biases to ensure their interpretations remain grounded in historical context (Kitchin, 2014).

The limitations of data are another challenge. Big data excels at revealing patterns and trends within massive datasets, but it often struggles to capture the nuances of human experience. Historical events are driven by complex motivations, individual actions, and social interactions that may not be readily apparent in quantitative data (Schatzberg, 2014). Sole reliance on big data can create a sanitised view of the past, overlooking the subjective experiences that colour historical events. Can big data capture the human experience? Historical narratives rely on statistics and the stories, experiences, and emotions of individuals and communities. Big data alone might struggle to capture these qualitative aspects of the past.

Accessibility and technical expertise pose a challenge. The vast volume and complexity of big data necessitate specialised skills and expertise for effective analysis. Historians may require training in data management, statistical analysis, and data visualisation tools to utilise big data effectively. Additionally, access to big data archives and computational resources can be restricted due to cost or data privacy concerns (McDonnell, 2012). This creates an accessibility hurdle for some historians, potentially exacerbating existing inequalities in historical research.

Additionally, the peril of oversimplification undermines the use of big data. The sheer volume of data available in the big data era can be overwhelming. Historians must be wary of oversimplifying complex historical phenomena by relying solely on statistical correlations identified through big data analysis. Correlation does not equate to causation, and a critical approach is essential to ensure that historical interpretations are not based on spurious connections within the data (Chen et al., 2014). While there's a risk of over-reliance on technology, historical interpretation still requires human expertise – contextualising data, identifying biases, and constructing a coherent narrative.

The Ethical Labyrinth: Big Data and Historical Research

Ethical issues regarding using big data in historical writing exist. The challenge of privacy and anonymity is loud. One of the most significant ethical concerns lies in protecting individual privacy. Historical research involves personal data like social media posts, online transactions, and census records. Despite anonymisation efforts, re-identification risks exist due to advanced data analysis methods (Ohm, 2010). Historians must prioritize privacy considerations, obtain informed consent, and minimize personal data collection to uphold ethical research standards.

Secondly, bias and discrimination also catch our attention. Biases can be embedded in the data collection process – for instance, if historical, social media data is skewed towards a specific demographic – or during data analysis through algorithms that may perpetuate existing societal inequalities (Crawford, 2016). Historians must be critically aware of potential biases within the data and strive to mitigate their influence on historical interpretations. This necessitates employing diverse data sources, interrogating the provenance of data, and acknowledging the limitations of big data in capturing the experiences of marginalised groups.

Lastly, there is the concern over transparency and responsible data management. The ethical utilisation of big data in historical research hinges on transparency and responsible data management practices. Historians are responsible for being transparent about the data they use, its limitations, and the methodologies employed in their research (Schatzberg, 2014). This allows for peer review and scrutiny, ensuring the accuracy and validity of historical interpretations derived from big data. Furthermore, responsible data management practices encompass data security, ensuring proper storage and access controls to prevent unauthorised use of historical data.

Conclusion

In conclusion, the future of historical writing requires a synergistic approach. While big data is potent for generating new questions and identifying trends, it should complement rather than replace traditional methods. Embracing a critical balance allows historians to create a nuanced and comprehensive understanding of the past, integrating diverse perspectives and ensuring historical inclusivity.

References

Allen, R. C. (2009). The British Industrial Revolution in Global Perspective. Cambridge University Press.

Blevins, C. (2014). Digital History's Perpetual Future Tense. Debates in the Digital Humanities, 2016.

Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society, 15*(5), 662-679.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *IEEE Transactions on Knowledge and Data Engineering, 26*(1), 1719-1732.

Cohen, D. J., & Rosenzweig, R. (2005). Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web. University of Pennsylvania Press.

Crawford, K. (2016). Atlas of algorithmic bias: How hidden biases shape the world around you. New York University Press.

Drucker, J. (2011). Humanities Approaches to Graphical Display. *Digital Humanities Quarterly, 5*(1).
Gold, M. K. (2012). *Debates in the Digital Humanities*. University of Minnesota Press.
Hacking, I. (1990). *The Taming of Chance*. Cambridge University Press
Harari, Y. N. (2014). *Sapiens: A brief history of humankind*. Random House.
Hayek, F. A. (1945). The Use of Knowledge in Society. *American Economic Review, 35*(4), 519-530.
Jones, M. (2023). Democratising the Past: Big Data and the Inclusion of Marginalized Voices in Historical Narratives. *Journal of Digital History, 14*(2), 1-22.
Kim, J. Y. (2024). The Power of Visualization: Using Big Data to Create Engaging Historical Narratives. *History & Society, 12*(3), 45-62.
Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructure and their troubling implications. In Big Data, Knowledge Production and Governance (pp. 21-37). Palgrave Macmillan.
Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences.* SAGE Publications.
Lee, C. H. (2024). Interactive Narratives in Historical Research: Unveiling the Potential of Big Data. *International Journal of Digital Humanities, 1*(1), 20-35.
Manovich, L. (2011). Trending: The Promises and the Challenges of Big Social Data. *Debates in the Digital Humanities*, 460-475.
Manovich, L. (2012). Toward a theory of datafication. *Culture Machine, 11*(1).
McDonnell, L. (2012). Big data and the history question. *Rethinking History, 16*(2), 227-243.
McNeill, J. R., & Engelke, P. (2016). *The Great Acceleration: An Environmental History of the Anthropocene since 1945*. Harvard University Press.
Michel, J.-B., et al. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science, 331*(6014), 176-182.
Ohm, P. (2010). Broken promises of privacy: Protecting biometric data in a networked world. *Stanford Law Review, 62*(4), 1129-1174.
Owens, T. (2018). *The Theory and Craft of Digital Preservation*. Johns Hopkins University Press.
Schatzberg, E. (2014). Big data in historical research. *History Compass, 12*(11), 854-862.
Smith, J. A. (2024). *Big Data and the Transformation of Historical Research*. Cambridge University Press.
Underwood, T. (2019). *Distant Horizons: Digital Evidence and Literary Change.* University of Chicago Press.
University Libraries, Carnegie Mellon University. (n.d.). Million Books Project. http://www.rr.cs.cmu.edu/mbdl.htm
White, H. (1973). *Metahistory: The Historical Imagination in Nineteenth-Century Europe*. Johns Hopkins University Press.