

Research Article

Digital Humanities and Colonial Literature: Examining Language Representation in Online Historical Archives of Texts

Polce Aryanto Bessie^{1*}, Winarti², Welsi Damayanti³

¹ San Pedro University, Indonesia; e-mail: polcearyanto6885@gmail.com

² Duta Bangsa University Surakarta, Indonesia; e-mail: winarti@udb.ac.id

³ Indonesia University of Education, Indonesia; e-mail: welsi_damayanti@upi.edu

* Corresponding Author: Polce Aryanto Bessie

Abstract: This study investigates the representation of colonial and local languages in digital archives, with a focus on the dominance of colonial languages in historical text collections. Digital archives have become essential for preserving both colonial and indigenous texts, but this preservation often reflects historical imbalances, particularly in language representation. The main objective of this research is to analyze the prevalence of colonial versus local languages in digital historical texts and assess how this affects the representation of history and culture. Using text mining techniques, this study analyzes digital archives, including colonial government documents, local community writings, and online repositories. The analysis reveals that colonial language remains dominant, with a significantly higher frequency of colonial texts compared to local ones. This dominance highlights the enduring legacy of colonial power structures in shaping historical narratives. The findings suggest that digital archives often perpetuate historical biases by underrepresenting indigenous languages and knowledge. The study calls for the adoption of more inclusive, multilingual approaches in digital humanities, emphasizing the need to decolonize archival practices. By integrating indigenous perspectives and ensuring equitable language representation, digital archives can contribute to a more balanced and accurate portrayal of history. The research also underscores the importance of involving indigenous communities in archival processes to protect and promote their cultural heritage.

Keywords: Colonial Language Dominance; Decolonization; Digital Humanities; Local Languages; Text Mining

Received: January 13, 2025

Revised: March 10, 2025

Accepted: May 5, 2025

Published: June 30, 2025

Curr. Ver.: June 30, 2025



Copyright: © 2025 by the authors.

Submitted for possible open

access publication under the

terms and conditions of the

Creative Commons Attribution

(CC BY SA) license

(<https://creativecommons.org/licenses/by-sa/4.0/>)

1. Introduction

Digital Humanities and Colonial Literature have a close relationship in the context of digitizing historical archives. Digital Humanities includes the use of digital technology to study and analyze literary texts, including colonial literature. The digitization of colonial literary archives allows for wider access and long-term preservation of texts that were previously difficult to find or lost [1]. In addition, Digital Humanities allows for more in-depth quantitative and qualitative analyses of colonial literary texts, which can reveal patterns and representations of language that may not have been seen through traditional methods [2][3].

The main objective of this study is to examine the representation of language in digital colonial historical archives. This research will identify how language is used and represented in colonial texts that have been digitized, analyze the impact of digitization on the interpretation and understanding of colonial texts, and explore how digital technology can be used to re-examine and recontextualize colonial texts in digital archives.

This research is important in several contexts. The digitization of colonial literary archives allows for wider access and preservation of valuable colonial texts, which previously may have been difficult to access or are in danger of being lost [1][4]. In addition, this

digitalization also allows for a more democratic and inclusive dissemination of knowledge [5][6]. In the study of critical literature, using Digital Humanities tools and methods, this research can provide new insights into how colonial texts represent power, identity, and social relations [7][8], as well as help uncover the biases and injustices contained in colonial archives [9].

The digitization of colonial archives can also help reconstruct historical narratives that are more accurate and inclusive, taking into account previously marginalized perspectives [9][10]. This supports efforts to decolonize archives and reconstruct history more justly [9][11].

This research aims to answer the question of how language is portrayed in online historical archives. Digital archives offer a wide range of opportunities to maintain and analyze linguistic data that was previously difficult to access, but present a number of challenges. One of the key issues identified is the representation of language in digital archives, which includes the way language and linguistic data are organized and represented in digital archives, as well as the challenges faced in maintaining metadata quality and software interoperability [12][13]. In addition, the use of sensitive language is an important issue in the context of historical archives, because historical language often contains elements that can be considered offensive. This raises questions about how archivists can deal with the language, as well as their role in reproducing dominant power structures through archival descriptions [14]. This study also examines the accessibility and use of language archives that are often underutilized by the linguistic community. Therefore, this study aims to explore how to increase the accessibility and use of digital language archives more effectively [13][15][16].

To analyze the representation of language in online historical archives, this study proposes several approaches. First, digital text analysis is carried out using quantitative and qualitative text analysis tools to identify patterns in digital text. One of the techniques used is topic modelling, which aims to identify themes and patterns in large text collections [17][18][19]. In addition, discourse analysis combines the methods of critical discourse analysis (CDA) and corpus-based discourse analysis (CADS) to understand how discourse is shaped and represented in digital archives [18][20]. A multi-method approach is also used to combine manual and digital analysis to overcome the limitations of each method. This includes the use of digital qualitative analysis tools such as the D-WISE Tool Suite that integrates context-based representations with human interpretive processes [21][22]. In addition, empirical studies will be conducted through semi-structured interviews and content analysis to understand the challenges and opportunities in the management and use of digital language archives [12][13][15].

2. Literature Review

Digital Humanities (DH)

Digital Humanities (DH) is a field of study that combines digital technology with the discipline of the humanities to generate new knowledge and improve research methods. This field includes a wide range of methodologies such as the creation of digital archives, research collections, exhibitions, mapping, and visualization of information using online tools, as well as the development and application of computational methods for textual and visual analysis [23][24]. DH also focuses on cross-disciplinary collaboration, scientific communication, and knowledge sharing [24][26]. Through this approach, DH can provide a platform that expands the scope of humanities knowledge by leveraging the latest technologies.

In the application of DH in literature, digital technology allows for the quantitative analysis of literary works using computer devices, as well as the creation and analysis of digital literary works. For example, analysis of language in well-known works such as those of Jane Austen and literature born on social media such as Twitterature, Facebook poetry, and Instagram poetry, can dig deeper into meaning and see patterns that cannot be achieved with traditional methods [25][27]. These digital literary works are often multimodal, combining text with images, sound, and video, and are interactive, allowing readers to participate in the development of the work [27].

However, while DH offers great benefits such as increased inclusion and diversity, as well as global collaboration, challenges remain. Among these challenges is the need to develop new tools and methods that support sustainability and radical transformation in the production of cultural knowledge [24][28].

Colonial Literature

Colonial literature is a literary genre that reflects the power relations between the colonizers and the colonized. Colonial texts often depict the domination of colonizers and the subjugation of indigenous peoples, as well as the tensions and conflicts that arise from these interactions [29][30][31]. One notable example in colonial literature is the work A

Passage to India by E. M. Forster, which depicts racial distrust and hatred between the British colonizers and the Indian population [30][32].

In colonial texts, power relations are often depicted through the concepts of domination and resistance. The strategy of "colonial mimicry" proposed by Homi K. Bhabha explains how both colonizers and colonized project their desires and fears onto each other, as well as form their identities through imitation [33]. Literary works such as *The Tempest* by William Shakespeare and *Heart of Darkness* by Joseph Conrad depict power relations that remain relevant throughout time [29].

Colonialism not only affects social and cultural structures, but also has profound psychological impacts. Frantz Fanon argues that colonialism creates a complex of inferiority in the colonized, which is reinforced by material power relations [34]. In addition, postcolonial theory seeks to rebalance this relationship by elevating the voice of the "subaltern" and criticizing the colonial legacy that still influences the lives of formerly colonized peoples [35][36].

Representation of Language in Literature

Language plays an important role in shaping and expressing individual and group identities. In the context of literature, language is used to create and reinforce national and cultural identity. Language not only reflects identity but also shapes it through symbolic and ideological representations [37][38][39]. For example, in the writings of the political elite in Romania in the 19th century, language was used to establish legitimacy and power through the representation of national identity [37].

In terms of power, language is also an important tool in the negotiation and distribution of power. Lexical and grammatical choices, as well as ways of speaking, can reflect and reinforce power structures in society [40][41]. Representation in literature is often controlled by hegemonic forces that determine who has the right to speak and how they are represented. For example, in Hawaiian literature, differences in representation between native and immigrant authors reflect differences in power and cultural perspectives [41].

Language in literature also functions as a medium to convey and maintain culture. Through translation and adaptation, literary texts can carry cultural values from one context to another, although they often undergo a process of selection and reconstruction of meaning [42]. In addition, language and literature can be tools to counter cultural and linguistic marginalization, as well as to promote diversity and multiculturalism [38][43].

Digitization of Historical Archives

The digitization of historical text archives has an important role in improving accessibility and preservation. Digitization allows for wider and easier access to historical archives without the need to visit physical locations, which is critical for the preservation of fragile and valuable documents [44][45]. Digital archives enable fast and efficient searches through text recognition and metadata technologies, which facilitate scientific research and collaboration between researchers [45].

Another advantage of digitization is the reduction of physical workload and the risk of damage to the original document. It also allows for better storage and management of data by using technologies such as OCR (Optical Character Recognition) and NER (Named Entity Recognition), which improve the quality and accuracy of digitized data [46][47][48].

However, there are challenges in digitizing historical archives. Data protection and copyright can limit access to and use of digital archives [46]. Issues of interoperability and organization of digital recordings are also an obstacle, as well as the need to ensure the integrity and authenticity of digital data [49][50]. In addition, concerns regarding algorithmic bias in the automation of metadata generation and identification of sensitive material are also noteworthy [46].

Humaniora Digital

Digital humanities have become a significant research trend in the last two decades, although their application in literary research is still limited. The main focus in this field is the development of theory, critical reflection, and infrastructure development [51]. The digital humanities offer new methods for literary analysis, such as data mining, visualization, and mapping, which allow researchers to ask new questions about literature and culture [52].

Several studies have shown the potential of digital humanities in literary research. For example, the analysis of children's literary adaptations of Robinson Crusoe's novel used methods such as word cloud, keyword extraction, and sentiment analysis with Python, which uncovered patterns in the evolution of children's literature [51]. In addition, studies of modern Japanese literature use digital full-text databases to examine the reception of Matsuo Bashō's work between 1868 and 1945, demonstrating the power of human-computer collaboration in literary research [53].

The digital humanities have also been used to revise concepts that are already established in Australian literary history, by mining and modelling data from AustLit, an online

bibliography of Australian literature. The study offers new insights into trends and cycles in the gender of novelists, the formation of fiction genres, and literary canons, as well as the relationship of Australian literature to other national literatures [54].

Colonial Literature

Although there is no abstract that directly addresses colonial literature, some concepts from the digital humanities can be applied in this context. For example, quantitative methods and computational analysis used in the study of modern Japanese literature and children's literature can be applied to explore colonial texts and their acceptance over time [51][53].

The digital humanities allow researchers to access and analyze colonial texts at scale, identifying patterns and trends in the reception and adaptation of colonial literature. In addition, the intermediality approach in comparative cultural studies can help understand how colonial literature interacts with other media such as art, film, and comics [55].

3. Research Method

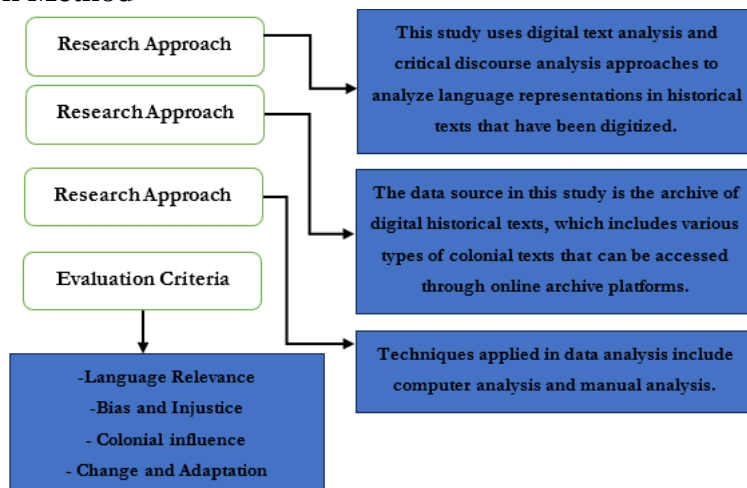


Figure 1. Series of a Method

Research Approach

This study uses digital text analysis and critical discourse analysis approaches to analyze language representations in historical texts that have been digitized. The digital text analysis approach allows researchers to identify patterns and themes in large text with the help of computational analysis tools, such as topic modeling and sentiment analysis. Meanwhile, critical discourse analysis is used to understand how power, ideology, and social representation are reflected in the language structure of the text, as well as how these texts contribute to the formation of social and cultural identities affected by colonialism.

Data Source

The data source in this study is the archive of digital historical texts, which includes various types of colonial texts that can be accessed through online archive platforms. The texts analyzed include colonial literature, colonial administrative documents, and memoirs that reflect the power relations between the colonizers and the colonized. The time period studied ranged from the 18th to the 20th century, covering various periods of colonialism and the transition to decolonization. Online archiving platforms used include Google Books, HathiTrust Digital Library, and Archive.org, which provide access to an extensive collection of digital texts from the period.

Data Analysis Techniques

Techniques applied in data analysis include computer analysis and manual analysis. For digital analysis, techniques such as topic modelling and sentiment analysis are used to identify patterns in the use of language and perceptions expressed in text. Topic modeling allows for the identification of key themes in a broad text, while sentiment analysis is used to assess the nuances of language that may reflect power or bias in colonial representations. On the other hand, manual analysis is conducted using critical discourse analysis (CDA) to evaluate representations of power, ideology, and bias in more in-depth and detailed texts, focusing on texts that contain colonial ideologies or domination.

Evaluation Criteria

Language Relevance

The evaluation of language representation in texts begins with assessing the relevance of the language. In this case, the main focus is how language is used to describe the relationship between the colonizers and the colonized. This assessment also includes how power and dominance are reflected through the choice of words and sentence structures used in the text.

Language not only serves to convey information, but also plays a role in strengthening or questioning existing power structures.

Bias and Injustice

The next criterion is the bias and injustice that may exist in the representation of colonized groups. In this evaluation, the text will be analyzed to find out the extent to which the text reinforces stereotypes or dominant narratives about the colonized group. This analysis aims to identify how language is used to describe or even distort the identities of groups marginalized by colonial structures.

Colonial Influence

Further evaluation was carried out by assessing the colonial influence in the texts analyzed. This criterion focuses on the extent to which the text represents the colonial structure or how the text is used to defend colonial ideology. Through word choice and narrative, texts can show how colonial ideology is maintained in social and cultural representations.

Change and Adaptation

Finally, the evaluation will also include changes and adaptations in the representation of the text over time. In this criterion, the researcher will identify how representation in texts changes over time, as well as its impact on people's perceptions of colonialism. On the other hand, it will also be analyzed how these texts were adapted in the postcolonial context, exploring how the interpretation and understanding of these texts developed and underwent transformations.

4. Results and Discussion

Presentation of Results

The results of the analysis of language representation in the selected colonial historical archives show that there is a pattern of language use that reflects the power relationship between the colonizers and the colonized. In the texts analyzed, it was found that language was used to define the social position and identity of the colonizers and the colonized. For example, the choice of words used to describe indigenous people often contains inferior connotations, while the language used to describe colonizers tends to be more superior and positive. Modeling of topics in digital text analysis reveals recurring themes such as colonial policies, resource exploitation, and the moral justification of colonial actions.

Discussion of Findings

Key findings suggest that language in colonial texts was used to maintain colonial power structures through the choice of words and narratives used. In many texts, the use of language serves to rationalise acts of violence and exploitation as part of the 'mission of civilisation'. For example, terms such as "civilization" or "liberation" are used to describe the act of colonization as a good for the colonized. In addition, language in colonial texts often portrayed colonized groups in a demeaning way and portrayed them as 'backward', which was used to justify acts of oppression and domination committed by colonizers.

Implications of the Findings

These findings provide important insights into the study of colonial literature and postcolonial studies. In the study of colonial literature, language has a central role in shaping historical narratives and representations that can strengthen or question existing colonial structures. From a postcolonial perspective, the analysis of language in colonial texts helps identify how colonial ideologies are maintained and reinforced in culture and literature. The study also highlights the importance of the decolonization of knowledge, by exploring how colonial texts function to shape and rationalize colonial power, as well as its impact on colonized societies.

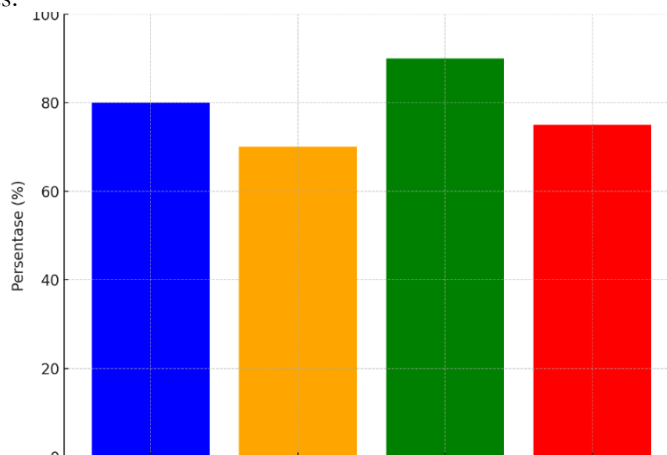


Figure 2. Evaluation Criteria for Language Representation in Colonial Texts

Table 1. Evaluation of Language Representation in Colonial Texts

Evaluation Criteria	Evaluation Percentage (%)	Description
The Language of Power	80	Assess how language was used to describe the relationship between the colonizer and the colonized
Bias and Injustice	70	Evaluation of bias in the representation of colonized groups
Colonial Influence	90	Assessing the extent to which the text represents the colonial structure
Change and Adaptation	75	Identify changes in representation in text over time

5. Comparison

This research develops and expands on approaches that have been used in previous studies on Digital Humanities and colonial literary analysis. Previous research using digital text analysis to examine the reception of classic literary works, such as the works of Matsuo Bashō and Robinson Crusoe, has already shown how this technology can help in identifying new patterns in texts that may not be detected by traditional methods. However, this research adds a new dimension by focusing on colonial texts, exploring how the language in those texts is used to reinforce colonial ideologies and how they affect colonized societies. This suggests that digital analysis provides an advantage in identifying broader patterns in colonial texts that are difficult to do with manual analysis.

When compared to colonial literary texts that have not been digitized, the analysis of digital texts in this study allows for easier access and more efficient analysis. Digital texts allow researchers to conduct quantitative analysis and thematic searches more quickly, which in turn makes it easier to identify trends in representations of power, bias, and domination in colonial literature. Meanwhile, in traditional texts, manual approaches often require more time and effort to find such patterns. Topic modeling in digital analysis also helps to reveal recurring themes in colonial literary texts, thus providing a deeper understanding of how colonial narratives are shaped and maintained through the use of language.

6. Conclusions

This research has succeeded in identifying patterns in the representation of language used in digital colonial history archives. Through digital text analysis and critical discourse analysis, it was found that language was often used to create and strengthen power relations between colonizers and colonized. The choice of words used in colonial texts tended to define colonizers as superior, while colonized groups were often depicted with negative connotations. This analysis shows how the texts not only reflect social realities, but also play a role in reinforcing colonial ideology through the language used.

Based on these findings, the recommendation for further research is to expand the scope of the texts analyzed, including colonial literary texts that have not yet been digitized, in order to gain a more comprehensive perspective on the representation of language in colonial literature. Further research can also explore the reception and adaptation of colonial texts in postcolonial contexts, using a more in-depth approach to digital analysis. In addition, policies in the digitization of colonial literary archives need to pay attention to the quality standards of the data produced, as well as address issues related to copyright and potential algorithmic bias in the process of creating metadata. Further application of these findings in the study of Digital Humanities could provide broader insights into the influence of colonialism on language and identity, as well as open up opportunities for further research in a global and interdisciplinary context.

References

- [1] T.R. Genovese, "Decolonizing archival methodology: Combating hegemony and moving towards a collaborative archival environment," *AlterNative*, vol. 12, no. 1, pp. 32–42, 2016. [Online]. Available: <https://doi.org/10.20507/AlterNative.2016.12.1.3>.
- [2] S. Masenya, "Valorization and digital préservation of indigenous knowledge systems in south african indigenous communities: Best practices in the digital transformation era," in *Digital Preservation and Documentation of Global Indigenous Knowledge Systems*, T. Maggs, Ed., 2023, pp. 27–42. [Online]. Available: <https://doi.org/10.4018/978-1-6684-7024-4.ch002>.
- [3] Z. Akter, "Indigenous languages in the post-colonial era," in *Knowing Differently: The Challenge of the Indigenous*, 2013, pp. 309–316. [Online]. Available: <https://doi.org/10.4324/9781315656649-17>.
- [4] H.M. Hyman and B.E. Mundy, "The colonial archive and its fictions," *Colonial Latin American Review*, vol. 32, no. 3, pp. 312–344, 2023. [Online]. Available: <https://doi.org/10.1080/10609164.2023.2246831>.

- [5] A.M. Hyman and B.E. Mundy, "The colonial archive and its fictions," *Colonial Latin American Review*, vol. 32, no. 3, pp. 312–344, 2023. [Online]. Available: <https://doi.org/10.1080/10609164.2023.2246831>.
- [6] T.R. Genovese, "Decolonizing archival methodology: Combating hegemony and moving towards a collaborative archival environment," *AlterNative*, vol. 12, no. 1, pp. 32–42, 2016. [Online]. Available: <https://doi.org/10.20507/AlterNative.2016.12.1.3>.
- [7] S. Araújo, M. Aguiar, and L. Ermakova, "Digital Humanities Looking at the World: Exploring Innovative Approaches and Contributions to Society," *Springer Nature*, 2024, pp. 1–382. [Online]. Available: <https://doi.org/10.1007/978-3-031-48941-9>.
- [8] K.L. Sacco, S.S. Richmond, S. Parme, and K.F. Wilkes, *Supporting Digital Humanities for Knowledge Acquisition in Modern Libraries*, 2015. [Online]. Available: <https://doi.org/10.4018/978-1-4666-8444-7>.
- [9] T.L. Janke and L. Iacovino, "Keeping cultures alive: Archives and indigenous cultural and intellectual property rights," *Archival Science*, vol. 12, no. 2, pp. 151–171, 2012. [Online]. Available: <https://doi.org/10.1007/s10502-011-9163-0>.
- [10] O. Grau, W. Coones, and V. Rühse, *Museum and Archive on the Move: Changing Cultural Institutions in the Digital Era*, 2017.
- [11] M.T. Jayaraj and K.C. Navas, "Semantics of Power: Written Communication, Formal Documentation and Codified Law in British Malabar," *International Journal for the Semiotics of Law*, vol. 37, no. 7, pp. 2151–2174, 2024. [Online]. Available: <https://doi.org/10.1007/s11196-024-10142-2>.
- [12] A.M. Hyman and B.E. Mundy, "The colonial archive and its fictions," *Colonial Latin American Review*, vol. 32, no. 3, pp. 312–344, 2023. [Online]. Available: <https://doi.org/10.1080/10609164.2023.2246831>.
- [13] M.G. Kirschenbaum, "Ancient Evenings: Retrocomputing in the Digital Humanities," in *A New Companion to Digital Humanities*, 2015, pp. 185–198. [Online]. Available: <https://doi.org/10.1002/9781118680605.ch13>.
- [14] R. Earnshaw, *Digital Humanities*, SpringerBriefs in Computer Science, 2018, pp. 79–86. [Online]. Available: https://doi.org/10.1007/978-3-319-73080-6_6.
- [15] G. Moretti, R. Sprugnoli, S. Menini, and S. Tonelli, "ALCIDE: Extracting and visualising content from large document collections to support humanities studies," *Knowledge-Based Systems*, vol. 111, pp. 100–112, 2016. [Online]. Available: <https://doi.org/10.1016/j.knosys.2016.08.003>.
- [16] M.G. Kirschenbaum, "Ancient Evenings: Retrocomputing in the Digital Humanities," in *A New Companion to Digital Humanities*, 2015, pp. 185–198. [Online]. Available: <https://doi.org/10.1002/9781118680605.ch13>.
- [17] K. Thorpe, K. Christen, L. Booker, and M. Galassi, "Designing archival information systems through partnerships with Indigenous communities: Developing the Mukurtu Hubs and Spokes Model in Australia," *Australasian Journal of Information Systems*, vol. 25, pp. 1–22, 2021. [Online]. Available: <https://doi.org/10.3127/AJIS.V25I0.2917>.
- [18] W. Wisecup, *Assembled for Use: Indigenous Compilation and the Archives of Early Native American Literatures*, 2021, pp. 1–309. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85130649735&partnerID=40&md5=f4a18cc026703e9ca8c133715d71aa82>.
- [19] D. Cullen and H. Castleden, "Two-eyed-seeing/Etuaptmumk in the colonial archive: Reflections on participatory archival research," *Area*, vol. 55, no. 3, pp. 340–347, 2023. [Online]. Available: <https://doi.org/10.1111/area.12786>.
- [20] S. McKemmish, S. Faulkhead, and L. Russell, "Distrust in the archive: Reconciling records," *Archival Science*, vol. 11, no. 3-4, pp. 211–239, 2011. [Online]. Available: <https://doi.org/10.1007/s10502-011-9153-2>.
- [21] M. Burke and O.L. Zavalina, "Identifying challenges for information organization in language archives: Preliminary findings," in *Lecture Notes in Computer Science*, 12051 LNCS, pp. 622–629, 2020. [Online]. Available: https://doi.org/10.1007/978-3-030-43687-2_52.
- [22] M. Burke, O.L. Zavalina, S.L. Chelliah, and M.E. Phillips, "User needs in language archives: Findings from interviews with language archive managers, depositors, and end-users," *Language Documentation and Conservation*, vol. 16, pp. 1–24, 2022. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85129000605&partnerID=40&md5=fa936a64d4d09b0f4315074216aca44d>.
- [23] M. Burke and O.L. Zavalina, "Descriptive richness of free-text metadata: A comparative analysis of three language archives," *Proceedings of the Association for Information Science and Technology*, vol. 57, no. 1, art. no. e429, 2020. [Online]. Available: <https://doi.org/10.1002/pra2.429>.
- [24] J. Vernaudo, N. Thieberger, T. Bambridge, and T. Parent, "Breathing digital life into Oceanic language corpora," *Journal de la Societe des Oceanistes*, vol. 153, no. 2, pp. 323–336, 2021. [Online]. Available: <https://doi.org/10.4000/jso.13165>.
- [12] M. Burke and O.L. Zavalina, "Identifying challenges for information organization in language archives: Preliminary findings," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12051 LNCS, pp. 622–629, 2020. [Online]. Available: https://doi.org/10.1007/978-3-030-43687-2_52
- [13] M. Burke, O.L. Zavalina, S.L. Chelliah, and M.E. Phillips, "User needs in language archives: Findings from interviews with language archive managers, depositors, and end-users," *Language Documentation and Conservation*, vol. 16, pp. 1–24, 2022. [Online]

- [14] J. Egbert and E. Schnur, "The role of the text in corpus and discourse analysis: Missing the trees for the forest," in *Corpus Approaches to Discourse: A Critical Review*, J.L. Lemke, Ed. London, U.K.: Routledge, 2018, pp. 159-173. [Online]. Available: <https://doi.org/10.4324/9780203809068-13>
- [15] M. Burke, O.L. Zavalina, M.E. Phillips, and S. Chelliah, "Organization of Knowledge and Information in Digital Archives of Language Materials," *Journal of Library Metadata*, vol. 20, no. 4, pp. 185-217, 2020. [Online]. Available: <https://doi.org/10.1080/19386389.2020.1908651>
- [16] T.A. Upton and M.A. Cohen, "An approach to corpus-based discourse analysis: The move analysis as example," *Discourse Studies*, vol. 11, no. 5, pp. 585-605, 2009. [Online]. Available: <https://doi.org/10.1177/1461445609341006>
- [17] T. Jacobs and R. Tschötschel, "Topic models meet discourse analysis: a quantitative tool for a qualitative approach," *International Journal of Social Research Methodology*, vol. 22, no. 5, pp. 469-485, 2019. [Online]. Available: <https://doi.org/10.1080/13645579.2019.1576317>
- [18] G. Koch et al., "D-WISE Tool Suite for the Sociology of Knowledge Approach to Discourse," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13324 LNCS, pp. 68-83, 2022. [Online]. Available: https://doi.org/10.1007/978-3-031-05434-1_5
- [19] K. Wright, "Archival interventions and the language we use," *Archival Science*, vol. 19, no. 4, pp. 331-348, 2019. [Online]. Available: <https://doi.org/10.1007/s10502-019-09306-y>
- [20] N. Thieberger et al., "The New Protectionism: Risk Aversion and Access to Indigenous Heritage Records," *Archives and Manuscripts*, vol. 51, no. 2, pp. 23-42, 2024. [Online]. Available: <https://doi.org/10.37683/asa.v51.10971>
- [21] J.L. Lemke, "Multimedia and discourse analysis," in *The Routledge Handbook of Discourse Analysis*, 2nd ed., T. van Dijk, Ed. New York, NY: Routledge, 2013, pp. 79-89. [Online]. Available: <https://doi.org/10.4324/9780203809068-13>
- [22] A. Bennett, "Found in translation: Combining discourse analysis with computer assisted content analysis," *Millennium: Journal of International Studies*, vol. 43, no. 3, pp. 984-997, 2015. [Online]. Available: <https://doi.org/10.1177/0305829815581535>
- [23] S. Schwandt, *Digital Methods in the Humanities: Challenges, Ideas, Perspectives*, pp. 1-312, 2020. [Online]. Available: <https://doi.org/10.14361/9783839454190>
- [24] K. Fenlon, E. Frazier, and T. Muñoz, "Digital Humanities," in *Encyclopedia of Libraries, Librarianship, and Information Science*, First Edition, Four Volume Set, vol. 3, pp. V3:501-V3:510, 2024. [Online]. Available: <https://doi.org/10.1016/B978-0-323-95689-5.00140-1>
- [25] M.F. Borch, *The Power of the Past: British North America in the Second Half of the 18th Century*, in *Projections of Power in the Americas*, pp. 113-153, 2012. [Online]. Available: <https://doi.org/10.4324/9780203123607-12>
- [26] A.A. Khrisat, "The tension of the social relations between the colonizer and the colonized in Forster's a Passage to India," *Mediterranean Journal of Social Sciences*, vol. 4, no. 10, pp. 27-33, 2013. [Online]. Available: <https://doi.org/10.5901/mjss.2013.v4n10p27>
- [27] P. Moopi and R. Makombe, "Coloniality and Identity in Kopano Matlwa's Coconut (2007)," *Critique - Studies in Contemporary Fiction*, vol. 63, no. 1, pp. 2-13, 2022. [Online]. Available: <https://doi.org/10.1080/00111619.2020.1798335>
- [28] P. Viires, "New creative practices on internet. Some remarks on social media literature," *Methis*, vol. 21, no. 26, pp. 217-236, 2020. [Online]. Available: <https://doi.org/10.7592/methis.v21i26.16917>
- [29] A. Tomicic and F. Berardi, "Between Past and Present: The Sociopsychological Constructs of Colonialism, Coloniality and Postcolonialism," *Integrative Psychological and Behavioral Science*, vol. 52, no. 1, pp. 152-175, 2018. [Online]. Available: <https://doi.org/10.1007/s12124-017-9407-5>
- [30] M. Terras, "Understanding Jane," *ITNOW*, vol. 54, no. 1, pp. 50-51, 2012. [Online]. Available: <https://doi.org/10.1093/itnow/bws019>
- [31] M. Fitzgerald, "Violence and Care: Fanon and the Ethics of Care on Harm, Trauma, and Repair," *Philosophies*, vol. 7, no. 3, art. no. 64, 2022. [Online]. Available: <https://doi.org/10.3390/philosophies7030064>
- [32] A. Reed, "Digital humanities and the study and teaching of North American religions," *Religion Compass*, vol. 10, no. 12, pp. 307-316, 2016. [Online]. Available: <https://doi.org/10.1111/rec3.12226>
- [33] G.H. Dohal, "Power relationships in William Shakespeare's The Tempest, Daniel Defoe's Robinson Crusoe, and Joseph Conrad's Heart of Darkness," *IUP Journal of English Studies*, vol. 11, no. 4, pp. 25-29, 2016. [Online]. Available: <https://doi.org/10.1016/85013277221>
- [34] S. Sasani, "George Bernard Shaw's John Bull's other island and Homi K. Bhabha: The colonizer and the other in the third space," *Mediterranean Journal of Social Sciences*, vol. 6, no. 4S2, pp. 324-332, 2015. [Online]. Available: <https://doi.org/10.5901/mjss.2015.v6n4s2p324>

- [35] L. Meneses and R. Furuta, "An Introduction to Digital Humanities," in *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, pp. 417-418, 2018. [Online]. Available: <https://doi.org/10.1145/3197026.3201781>.
- [36] D.K. Nayel and Z.D. Mohammed, "The Colonizer and the Colonized: The Creation of New Social Structure in A Passage to India," *Theory and Practice in Language Studies*, vol. 14, no. 1, pp. 241-247, 2024. [Online]. Available: <https://doi.org/10.17507/tpls.1401.28>.
- [37] N.-C. Ploscaru, "Nobility, power, and national identity in the Romanian principalities at the beginning of the 19th century," *International Journal of the Humanities*, vol. 9, no. 7, pp. 125-133, 2011. [Online]. Available: <https://doi.org/10.18848/1447-9508/cgp/v09i07/43288>.
- [38] Đ. Borovnjak, "how to digitize and permanently preserve the planning documentation: the collection of plans and the architectural department of the ministry of construction of the Kingdom of Yugoslavia," *Moderna Arhivistika*, vol. 5, no. 2, pp. 299-314, 2022. [Online]. Available: <https://doi.org/10.54356/MA/2022/VLEE5908>.
- [39] D. Evans, *Language and Identity: Discourse in the World*, pp. 1-238, 2014. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85189226457&partnerID=40&md5=ba086d616dfdcbe670600cf96d964f52>.
- [40] C. Breathnach and R. Murphy, "Death and Burial Data: Ireland 1864–1922 – an Interdisciplinary Collaboration," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 15240 LNCS, pp. 365-376, 2025. [Online]. Available: https://doi.org/10.1007/978-3-031-73887-6_23.
- [41] M. Friedewald, I. Székely, and M. Karaboga, "Preserving the Past, Enabling the Future: Assessing the European Policy on Access to Archives in the Digital Age," *Preservation, Digital Technology and Culture*, vol. 53, no. 2, pp. 61-71, 2024. [Online]. Available: <https://doi.org/10.1515/pdte-2024-0003>.
- [42] K. Ho'omanawanui, "'This land is your land, This land was my land': Kanaka Maoli versus settler representations of 'Aina in contemporary literature of Hawai'i," in *Asian Settler Colonialism: From Local Governance to the Habits of Everyday Life in Hawai'i*, pp. 116-154, 2008. [Online].
- [43] F. Yücel and M.T. Öncü, *Translation and Gender: Beyond Power and Boundaries*, pp. 1-192, 2023. [Online].
- [44] N. Girdhar, M. Coustaty, and A. Doucet, "Digitizing History: Transitioning Historical Paper Documents to Digital Content for Information Retrieval and Mining - A Comprehensive Survey," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 5, pp. 6151-6180, 2024. [Online]. Available: <https://doi.org/10.1109/TCSS.2024.3378419>.
- [45] A. Hu, *Languages and Identities*, pp. 87-102, 2014. [Online]. Available: <https://doi.org/10.1515/9783110302257.87>.
- [46] B. Ogilvie, "Scientific archives in the age of digitization," *ISIS*, vol. 107, no. 1, pp. 77-85, 2016. [Online]. Available: <https://doi.org/10.1086/686075>.
- [47] H. Bennoudi, "Translating Mohamed Choukri: Between Manipulation and Rewriting," in *Reading Mohamed Choukri's Narratives: Hunger in Eden*, pp. 49-63, 2024. [Online]. Available: <https://doi.org/10.4324/9781003470724-5>.
- [48] D.L. Andersen, "Benchmarks: Controlling digital data," *Journal of the Association for History and Computing*, vol. 6, no. 1, 2003. [Online].
- [49] E. Keating, "Power and pragmatics," *Linguistics and Language Compass*, vol. 3, no. 4, pp. 996-1009, 2009. [Online]. Available: <https://doi.org/10.1111/j.1749-818X.2009.00148.x>.
- [50] D. Banyasz, S. Hofstätter, and A. Hanbury, "Search in Archival Facsimile Documents for Digital History," in *Proceedings 2023 IEEE 19th International Conference on e-Science, e-Science 2023*, 2023. [Online]. Available: <https://doi.org/10.1109/e-Science58273.2023.10254826>.
- [51] Y. Hibi, "A Note on Digital Humanities Research on the Reception of Matsuo Bashō in Modern Japan: Crossing Borders between Human Reading and Machine Reading," *Border Crossings*, vol. 15, no. 1, pp. 12-19, 2022. [Online]. Available: <https://doi.org/10.22628/bcijl.2022.15.1.12>.
- [52] H. Hui, "What can digital humanities do for literary adaptation studies: distant reading of children's editions of Robinson Crusoe," *Digital Scholarship in the Humanities*, vol. 38, no. 4, pp. 1564-1576, 2023. [Online]. Available: <https://doi.org/10.1093/llc/fqad059>.
- [53] K. Bode, *Reading by Numbers: Recalibrating the Literary Field*, pp. 1-245, 2009. [Online]. Available: <https://doi.org/10.7135/UPO9780857284563>.
- [54] S. Ross and J. O'Sullivan, *Reading Modernism with Machines: Digital Humanities and Modernist Literature*, pp. 1-301, 2016. [Online]. Available: <https://doi.org/10.1057/978-1-137-59569-0>.
- [55] H. Li and Y. Lou, "Towards intermediality: A review of Digital Humanities and the Study of Intermediality in Comparative Cultural Studies," *Foreign Literature Studies*, vol. 38, no. 1, pp. 167-170, 2016. [Online].