Reframing Narrative Analysis: Computational Approaches to Abdulrazak Gurnah's by the Sea

Godfrey Wandwi[1], Felistas Richard Mahonge[2]

[1]School of Digital Technologies and Transformation Studies, Dar es Salaam Tumaini University, Dar es Salaam, Tanzania

[2]School of Education and Human Development, Dar es Salaam Tumaini University, Dar es Salaam, Tanzania

DOI: https://doi.org/10.47772/IJRISS.2026.1026EDU0102

Received: 15 February 2026; Accepted: 21 February 2026; Published: 27 February 2026

ABSTRACT

This study investigates the integration of computational approaches into literary analysis, with a specific focus on Abdulrazak Gurnah's By the Sea (2001), situating the novel within a comparative computational corpus of approximately 200 literary texts used to establish broader thematic and sentiment baselines. exploring how digital tools can illuminate narrative structures, themes, and character dynamics. By employing natural language processing (NLP), topic modeling, sentiment analysis, and social network analysis, the researchers examine the novel's exploration of identity and displacement, memory and truth, cultural conflicts, and the postcolonial experience, revealing patterns and thematic structures that complement traditional close reading. Computational analysis highlights recurring motifs of solitude, conflict, and social justice, as well as the nuanced interplay between character relationships and narrative voice, offering a systematic perspective on Gurnah's complex narrative architecture. The study demonstrates that computational methods enrich literary scholarship by providing empirical insights into themes, emotional trajectories, and social networks, while still honouring the interpretive depth of humanistic inquiry. By integrating computational techniques with traditional literary interpretation, the study reveals that digital humanities approaches can enrich understanding of narrative complexity. The research also discusses methodological limitations, including interpretive challenges in large-scale textual analysis, and proposes future applications of machine learning for exploring nuanced narrative features in postcolonial literature.

Keywords: computational approaches, digital humanity, natural language processing, topic modeling, sentiment analysis, narrative structure, literary analysis, postcolonial literature,

INTRODUCTION

In recent years, the intersection of computational methods and narrative analysis has catalyzed a transformative shift in literary studies, enabling scholars to examine texts at both macroscopic and microscopic levels. While traditional narrative analysis relies on interpretive close reading (Brooks, 1992; Fish, 1980), recent scholarship demonstrates how artificial intelligence and computational methods can enhance literary analysis (Yadav, 2024; Babenko & Athavale, 2025; Paulson & Leonard, 2025). This paper applies such methods to Abdulrazak Gurnah's *By the Sea* (2001), demonstrating how computational analysis can illuminate the novel's narrative structure, thematic focus, and character dynamics within the broader context of contemporary literature.

While the study employs a broader computational corpus consisting of approximately 200 literary texts to establish comparative baselines for thematic prevalence, sentiment trajectories, and narrative structures, the detailed interpretive analysis presented in this article focuses on Abdulrazak Gurnah's *By the Sea* as a representative case study. The wider corpus analysis serves as a contextual framework against which the novel's narrative strategies and thematic configurations are evaluated.

Computational Approaches and the Evolution of Literary Analysis

The emergence of digital humanities and computational tools marks a significant evolution in literary scholarship, representing a methodological shift within narrative analysis that builds on and extends earlier theoretical frameworks (Ramsay, 2011). Foundational narratological models, such as Propp's morphology of story functions (Propp, 1968) and Todorov's structuralist approaches to narrative form (Todorov, 1971), provided useful conceptual foundations but were constrained by their reliance on manual analysis of limited text samples.

Contemporary computational approaches, leveraging Natural Language Processing (NLP), machine learning, and network science, now enable large-scale investigations into narrative structures, thematic development, and character dynamics, uncovering patterns that remain obscured in traditional close reading (Gomathi, et.al.2024). The rise of digital humanities and associated computational techniques allow scholars to apply models such as topic modeling and sentiment analysis to entire corpora, generating empirical data that complements existing theory (Blei et al., 2003; Underwood, 2019). Such methods have introduced new analytical lenses, allowing scholars to map thematic elements, sentiment, emotional trajectories, and narrative voice across large datasets, thereby contributing to a more holistic understanding of literature (Jockers, 2013; Underwood, 2019; Paulson & Leonard, 2025; Babenko & Athavale, 2025; Gomathi, et.al. 2024, Tanasescu, 2024).

Central to computational literary analysis is its capacity to quantify elements of text that were previously accessible only through qualitative interpretation. Topic modeling first introduced by Blei, Ng, and Jordan (2003) for example, enables the extraction of recurring motifs and themes by identifying clusters of co occurring words across large collections of text. It has been widely applied in literary studies to identify latent themes across large corpora, reshaping interpretations of genre, theme, and historical context (Goldstone & Underwood, 2014). Sentiment analysis allows for the tracking of emotional tone over narrative progression, and social network analysis reveals character interactions and influence structures within narratives.

These computational approaches open new analytical avenues that do not supplant close reading but enrich it by providing scalable, data driven insights. It is a hybrid methodology that supports new readings of canonical works while opening space for non-canonical texts that have often been overlooked. (Jockers, 2013; Moretti, 2013). Consequently, computational narrative analysis not only reaffirms the significance of the literary canon but also democratizes literary study by facilitating access to a broader array of texts and genres. (Paulson & Leonard, 2025; Babenko & Athavale, 2025 Gomathi, et.al. 2024).

Expanding Theoretical Perspectives through Computational Analysis

Computational approaches contribute substantially to narrative theory by interrogating and extending classical frameworks. Narratological concepts such as Genette's (1980) notion of focalization and Bakhtin's (1981) heteroglossia are increasingly recontextualized through computational text analysis. For example, character network analysis provides empirical support for dialogism, as patterns of interaction reveal the multi-voiced nature of complex narratives (Elson, Dames, & McKeown, 2010). Similarly, sentiment analysis has been employed to examine shifts in tone and the psychological depth of characters across genres (Piper & So, 2015). These methods illuminate how narrative elements reflect broader sociocultural constructs, offering insights into shifting representations of class, gender, and power.

Beyond reinforcing existing theories, computational approaches also introduce novel frameworks. Distant reading, as proposed by Moretti (2013), enables macro-level comparative analyses across vast corpora, tracing genre evolution, linguistic trends, and thematic constellations. Recent studies highlight how Natural Language Processing (NLP), machine learning, and AI-driven techniques broaden the horizons of literary scholarship, enabling more inclusive analyses across diverse texts and genres (Paulson & Leonard, 2025; Babenko & Athavale, 2025; Gomathi et al., 2024).

Methodological and Epistemological Implications

The incorporation of computational methods into narrative analysis has prompted renewed reflection on methodological and epistemological assumptions within literary studies. Traditional close reading, grounded in

subjective interpretation, relies on the scholar's ability to synthesize meaning from discrete textual elements (Fish, 1980). In contrast, computational approaches adopt a data-driven framework, privileging pattern recognition and quantitative validation over interpretive subjectivity (Algee-Hewitt et al., 2016). This paradigm shift has generated both enthusiasm and skepticism, as scholars debate the epistemic validity of computational findings and their implications for the interpretive core of literary analysis (Flanders & Jannidis, 2019). While some caution that computational methods risk reducing literature to mere data, others argue that these approaches enrich interpretation by revealing textual dimensions previously inaccessible (Kirschenbaum, 2007; Tanasescu, 2024).

The convergence of quantitative and qualitative methodologies in digital humanities represents not only a methodological innovation but also an epistemological expansion. By integrating empirical rigor with critical interpretation, computational narrative analysis fosters a pluralistic approach to textual study, bridging the divide between scientific and humanistic inquiry. Scholars such as Underwood (2019) emphasize that this dual framework generates new ways of conceptualizing narrative, complementing traditional literary theory, while Ramsay (2011) advocates for "algorithmic criticism" as a means of enhancing rather than supplanting close reading practices. More recent work demonstrates how NLP, machine learning, and AI-driven methods extend the reach of narratology, democratizing access to diverse texts and genres and reframing canonical and non-canonical works alike (Paulson & Leonard, 2025; Babenko & Athavale, 2025; Gomathi, et.al.,2024).

Ultimately, the integration of computational methods into narrative analysis represents a transformative shift within literary studies. By making narrative structures and themes visible across expansive datasets, computational approaches empower researchers to reconsider long-standing assumptions and explore new interpretive possibilities. In this way, digital humanities are not merely a collection of technical tools but a substantive contribution to the evolution of literary scholarship, redefining how narratives are studied, understood, and appreciated across time and culture.

This paper contributes to ongoing debates in digital humanities by demonstrating how computational methods can be applied to a specific literary text while also engaging with broader theoretical concerns. In doing so, it addresses key gaps in the literature, particularly regarding interdisciplinary synthesis and methodological reflexivity in computational literary studies. Ultimately, the integration of computational methods into narrative analysis exemplifies a transformative shift in literary scholarship—one that empowers researchers to explore enduring questions about narrative meaning, cultural representation, and textual complexity across time and genre.

THEORETICAL FRAMEWORK AND LITERATURE REVIEW

The increasing application of computational techniques in literary studies represents a vibrant interdisciplinary convergence that enriches traditional narrative analysis with quantitative and empirical rigor. Situated within the broader field of digital humanities, this convergence encompasses methods such as distant reading, topic modeling, sentiment analysis, and social network analysis. These approaches enable researchers to uncover textual patterns across large corpora and revisit longstanding interpretive questions through a data-informed lens. Recent scholarship demonstrates how Natural Language Processing (NLP), machine learning, and AI-driven models expand the scope of narratology, democratize access to diverse texts, and foster new theoretical insights that complement and challenge established interpretive frameworks (Paulson & Leonard, 2025; Babenko & Athavale, 2025; Gomathi, et.al.,2024; Tanasescu, 2024).

Foundations of Narrative Theory in Literary Studies

Narrative theory has traditionally been grounded in close reading and interpretive frameworks that emphasize structure, voice, and temporality. Foundational theorists such as Gérard Genette and Mikhail Bakhtin established key concepts for analysing narrative form, including narrative levels, focalization, and heteroglossia (Genette, 1980; Bakhtin, 1981). These classical approaches illuminate the internal architecture of narrative texts but remain limited when applied to large textual collections or cross-textual patterning. Recent advances in computational methods provide an empirical complement to these theoretical tools by rendering structural features measurable at scale. Through techniques such as topic modeling, sentiment analysis, and character network analysis, scholars can examine how narrative elements—such as focalization shifts or multi-

voiced discourse—manifest statistically across diverse corpora, including texts that resist simple categorization. (Alasmari et al., 2025; Akimoto & Ogata, 2025; Demir, 2025).

Theoretical Basis of Computational Approaches to Narrative

Computational approaches to narrative analysis extend principles of narrative theory by drawing on linguistics and computer science, with tools and algorithms increasingly adapted for literary study. A central foundation is Natural Language Processing (NLP)**,** a branch of artificial intelligence that enables computers to interpret and generate human language. Techniques such as sentiment analysis, named entity recognition, topic modeling, and syntactic parsing provide new ways to examine literature at granular levels, capturing shifts in emotional tone, mapping character interactions, and uncovering thematic structures (Piper & So, 2015; Jockers, 2013; Blei, Ng, & Jordan, 2003; Goldstone & Underwood, 2014). These methods allow researchers to quantify qualitative aspects of narratives, adding an empirical dimension to literary analysis and connecting traditional interpretive questions about plot, character, and theme with large-scale textual evidence (Li et al., 2024; Antoniak, 2025). Recent advances, particularly transformer-based models such as BERT (Bidirectional Encoder Representations from Transformers) and its variants, have enhanced sentiment analysis by detecting nuanced emotional states, enabling precise mapping of affective trajectories across narrative arcs and genres (Devlin et al., 2019; Sun et al., 2023; Hämäläinen et al., 2025). Together, these computational techniques demonstrate how classical insights into narrative progression and thematic development can be tested, refined, and expanded through scalable digital methods.

Topic Modeling in Narrative Analysis

Topic modeling represents a key methodological innovation in computational narrative analysis. Originally developed by Blei, Ng, and Jordan (2003), topic modeling is a probabilistic technique for discovering abstract topics within large text corpora. In literary studies, it has been employed to reveal thematic structures across works, enabling researchers to identify recurring motifs and patterns that span genres, historical periods, or an author's oeuvre (Goldstone & Underwood, 2014; Jockers, 2013). More recent scholarship has expanded the scope of topic modeling by integrating advanced algorithms such as neural topic models and hybrid approaches, which improve interpretability and accuracy in literary applications (Li, Wu, & Lei, 2024; Ogunleye, Lancho Barrantes, & Zakariyyah, 2025). By clustering statistically co-occurring words, topic modeling provides a scalable means of uncovering latent thematic trends that may remain hidden through traditional close reading. This study applies topic modeling to diverse narrative datasets as a way to trace thematic continuities and variations, thereby offering a broader empirical foundation for understanding literary and cultural dynamics.

Literary Studies and Distant Reading

Franco Moretti's (2013) concept of distant reading has been central to the development of computational literary studies. Distant reading challenges the exclusive privileging of close reading by advocating for the examination of large textual corpora to uncover systemic patterns, genre evolutions, and historical shifts. Moretti argues that patterns detectable only through large-scale analysis can offer new insights into literary history and cultural change.

Contemporary proponents of distant reading build on Moretti's framework by integrating computational tools with critical interpretation. Underwood (2019, 2021) emphasizes the value of combining distant reading with interpretive sensitivity, arguing that large-scale pattern detection should be grounded in contextual knowledge of literary history and aesthetic practice. Similarly, Liu et al. (2023) demonstrate that distant reading techniques can reveal thematic evolutions across genres and periods, providing empirical support for claims about historical shifts in narrative focus and authorial style.

Social network analysis (SNA) is a computational method widely used in the digital humanities to study relationships in narrative texts. Similar to distant reading, it models characters as nodes and their interactions as edges, enabling systematic analysis of patterns across novels, films, and other narratives. Early work by Elson, Dames, and McKeown (2010) showed that character networks can be automatically extracted, laying the foundation for computational narrative analysis. Since then, research has advanced from static models to

dynamic approaches that track evolving relationships, revealing shifts in roles, importance, and structure often missed by close reading. Modern studies also integrate NLP techniques, such as embeddings, to capture semantic nuances and refine representations of narrative structures (Bamman, 2020; Chaturvedi et al., 2017; Lee & Yeung, 2022). Dynamic centrality and temporal modeling further illuminate how narrative significance develops over time. By combining SNA with literary analysis, scholars can quantify interaction patterns while enriching qualitative interpretation, offering deeper insights into character roles, plot progression, and narrative design.

Existing Literature on Computational Literary Analysis

Computational methods have become central to literary studies, enabling large-scale analysis of stylistic and thematic change. Underwood (2019) demonstrates how data-driven approaches reveal cultural and linguistic shifts across centuries, while Jockers (2013) introduces "macroanalysis" to uncover historical and genre-level dynamics through text mining. Goldstone and Underwood (2014) further highlight the potential of topic modeling to expose latent thematic structures within literary corpora.

Recent scholarship expands these foundations by integrating natural language processing (NLP) and machine learning. Embedding-based models and dynamic network analysis capture semantic nuance and evolving narrative structures (Bamman, 2020; Chaturvedi, Srivastava, & Mitchell, 2017; Lee & Yeung, 2022). African scholars have also advanced computational literary analysis, particularly in contexts where linguistic resources are limited. Yékú (2022) interrogates the coloniality of data in digital African literatures, while projects on Nigerian texts in Yoruba and Efik highlight both the promise and challenges of applying NLP to African languages (Zenodo, 2024; Princeton CDH, 2024).

In East Africa, research on Swahili corpus linguistics and digital humanities provides essential infrastructure for computational literary studies. Scholars at the University of Dar es Salaam and Makerere University have developed Swahili corpora and morphological analyzers, enabling computational approaches to narrative texts (De Pauw & De Schryver, 2008; Hurskainen, 2016). Regional collaborations have further expanded these foundations by integrating natural language processing (NLP) initiatives, such as Masakhane NLP (2020), alongside machine learning techniques. These efforts situate East African scholarship within global AI research, emphasizing multilingual equity and methodological innovation. Collectively, this body of work underscores the interdisciplinary and international nature of computational literary studies, integrating quantitative modeling with qualitative interpretation to deepen our understanding of literary change, narrative design, and cultural context.

Analytical Framework: Integrating Computational Methods with Narrative Theory This study employs a hybrid framework that combines computational methodologies with narrative theory to analyze literary texts. Computational techniques are powerful but require theoretical grounding to ensure meaningful interpretation. As Flanders and Jannidis (2019) argue, digital humanities methods are most effective when they complement, rather than replace, traditional humanistic analysis. Recent scholarship further emphasizes on integration of computational tools to enrich interpretive practices (Yadav, 2024; Srivastava. et.al., 2025). Accordingly, this study integrates computational findings with narrative theory to create a comprehensive approach that leverages the strengths of both disciplines.

The analytical process applies natural language processing (NLP), topic modeling, and social network analysis (SNA) to a selected corpus. NLP techniques, such as sentiment analysis, reveal emotional tone and narrative voice; topic modeling uncovers latent thematic structures; and SNA visualizes social dynamics among characters. Together, these tools enable a multidimensional analysis of narrative elements, enhancing our understanding of literature as both cultural artifact and human expression.

This framework contributes to digital humanities by demonstrating the practical application of computational methods to literary texts, while also enriching literary studies through a computational perspective that quantifies and visualizes narrative structures. It further engages with ongoing debates about balancing empirical rigor and interpretive flexibility, showing that computational findings can be integrated with traditional interpretive practices. By bridging quantitative analysis with qualitative insight, this study exemplifies a balanced approach to narrative analysis that respects the complexities of literary texts while embracing the advantages of computational methods.

METHODOLOGY

This section outlines the research design, data collection methods, tools, and analytical techniques employed to explore narrative structures and thematic patterns in literature through computational approaches. This study applies an interdisciplinary methodology that combines principles from digital humanities, narrative analysis, and natural language processing (NLP) to investigate narratives at a larger scale, enabling both quantitative and qualitative insights into textual data. The approach aims to illuminate narrative elements, thematic recurrences, and character relationships within the selected corpus, thereby contributing to a more nuanced understanding of literary texts.

Research Design

The study adopts a mixed-methods research design, integrating quantitative computational techniques with qualitative narrative analysis to provide a comprehensive examination of literature. This approach was selected to capitalize on the strengths of both methodologies. The quantitative component enables large-scale data analysis, identifying patterns that would be challenging to discern through traditional close reading. The qualitative component, on the other hand, offers depth of interpretation, allowing for a critical contextual understanding of the results derived from computational analyses.

By combining these methods, this study provides a multifaceted view of literature that respects the complexities of narrative construction while leveraging the benefits of computational rigor. Additionally, the research design includes iterative phases, where insights from computational outputs inform further qualitative analysis and vice versa, fostering a dynamic interaction between data-driven and interpretive processes.

Data Collection and Corpus Selection

For this study, a well-curated corpus of narrative texts was assembled, consisting of approximately 200 novels and short stories from diverse genres, authors, and time periods. The corpus was selected to encompass a wide range of narrative styles, thematic elements, and character structures, ensuring a robust dataset for analysis. The broader corpus was primarily employed to establish comparative benchmarks for sentiment patterns, thematic recurrence, and character network structures across genres. The focused interpretive analysis presented in the findings section concentrates on *By the Sea* as a case study, allowing close examination of how corpus-level patterns manifest within a single postcolonial narrative.The selected texts include works that are representative of different literary movements and cultural contexts to provide a holistic view of narrative evolution and thematic diversity.

The corpus was sourced from publicly available literary databases, such as Project Gutenberg and the HathiTrust Digital Library, which offer high-quality, legally accessible digital texts. The works selected are in English to maintain consistency in linguistic analysis and reduce complexities associated with multi-language corpora, such as translation inconsistencies and varied syntactical structures. Prior to analysis, each text was pre-processed to remove metadata, page numbers, and other extraneous information that could interfere with computational methods.

Tools and Software

A suite of computational tools and software was employed to analyze the corpus. The primary tools include Python's Natural Language Toolkit (NLTK), spaCy, and Gensim, which provide capabilities for NLP, topic modeling, and social network analysis. Additionally, the study utilized software such as Gephi for visualizing character networks and interaction patterns within narratives.

- **Natural Language Toolkit (NLTK)**: This Python library was used for text preprocessing tasks, such as tokenization, stemming, and lemmatization. It also provided tools for sentiment analysis and word frequency analysis.

- **spaCy**: Known for its advanced NLP capabilities, spaCy was employed to perform part-of-speech tagging, named entity recognition, and dependency parsing, which are essential for character interaction and sentiment analysis.

- **Gensim**: This library was used for topic modeling, specifically through Latent Dirichlet Allocation (LDA), which is a popular technique for identifying thematic patterns in large text corpora.

- **Gephi**: A powerful visualization tool, Gephi enabled the creation of network graphs to illustrate character relationships and interactions, highlighting social structures within the narratives.

Data Preprocessing

Data preprocessing was a crucial step in preparing the text for computational analysis. Given the textual complexity of literary works, preprocessing involved multiple stages to ensure data accuracy and relevance for subsequent analysis. This phase included the following steps:

1. **Tokenization**: The texts were broken down into individual tokens (words), which serve as the base units for various computational analyses.

2. **Lemmatization and Stemming**: Each token was standardized to its base form, reducing words to their lemmas to avoid redundancy and improve pattern recognition across different forms of the same word.

3. **Stop Word Removal**: Commonly used words with minimal semantic weight (e.g., “and,” “the,” “is”) were removed to focus on words that are likely to carry meaningful narrative or thematic content.

4. **Named Entity Recognition (NER)**: This process identified and categorized entities such as characters, locations, and dates, which are essential for examining narrative elements like character interactions and thematic contexts.

These preprocessing steps resulted in a structured dataset optimized for computational analysis, ensuring that the analyses focus on relevant linguistic and narrative components of each text.

Analytical Methods

This study employed a multi-layered analytical framework consisting of three primary techniques: sentiment analysis, topic modeling, and social network analysis. Each method was selected to address specific research objectives and contribute unique insights into the narratives under study.

Sentiment Analysis

Sentiment analysis was conducted to examine the emotional tone of each narrative. By identifying positive, negative, and neutral sentiments across different parts of the text, the analysis provides insights into the emotional trajectory and character dynamics. SpaCy and NLTK were used for this task, employing lexicon-based approaches that assign sentiment values to words and phrases. Sentiment trends were then plotted over the course of each narrative, revealing shifts in tone that could correspond to plot developments or character transformations. Figure 1 illustrates the sentiment trajectory across narrative progression, demonstrating sustained emotional restraint rather than abrupt polarity shifts typical of high-conflict genres.

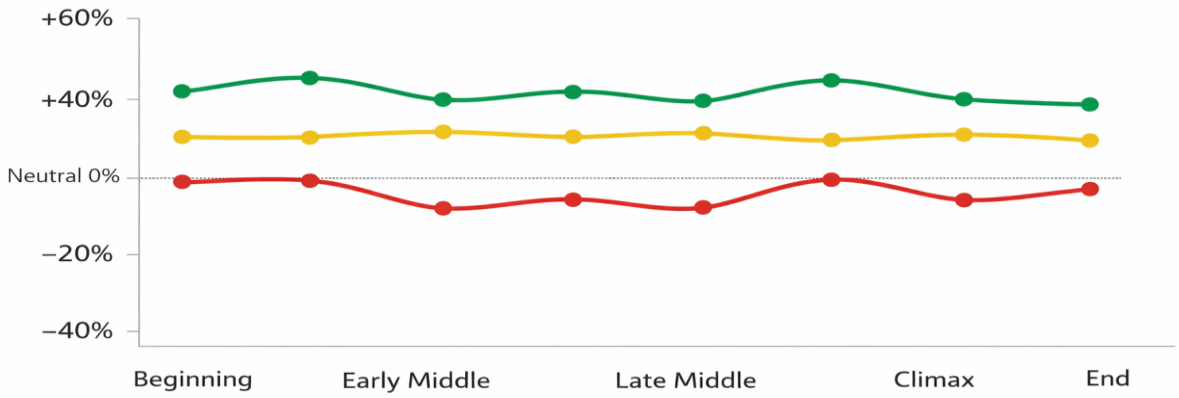

Figure 1: Sentiment progression across narrative stages

This sentiment analysis also enabled comparative studies between texts, allowing for an examination of how different authors approach emotional expression and how it impacts the reader's perception of characters and themes. For example, a predominance of negative sentiment could indicate a tragic or somber tone, while frequent shifts in sentiment may suggest complex character developments or plot twists.

Topic Modeling

Topic modeling, specifically through Latent Dirichlet Allocation (LDA), was utilized to uncover latent thematic structures within the corpus. This technique identifies clusters of words that frequently co-occur, representing them as "topics" that provide insights into recurring themes and motifs. The optimal number of topics (k = 5) was determined through iterative testing using coherence score evaluation (Cv metric), perplexity comparison, and manual interpretive validation. Coherence scores stabilized between k = 5 and k = 6; however, five topics provided the strongest balance between semantic interpretability and thematic distinctiveness across the corpus.

Each topic was interpreted as a thematic component, with labels assigned based on the most representative words. These topics were then mapped across the corpus to identify their prevalence and distribution, highlighting shifts in thematic focus across different literary periods or genres. For instance, a topic dominated by words related to nature and solitude might suggest Romantic themes, whereas a topic focusing on technology and modernity could indicate a science fiction influence.

Social Network Analysis (SNA)

Social network analysis was employed to examine character interactions and relationships within the narratives. This approach involved mapping characters as nodes and their interactions as edges, resulting in a visual representation of the social structure in each narrative. By using named entity recognition, characters were identified and tracked throughout the text, allowing for the construction of interaction networks.

Gephi was used to generate network graphs that display the density and strength of connections between characters. These graphs provide insights into narrative structure, such as identifying central or peripheral characters and understanding the social dynamics within each story. Metrics such as degree centrality and betweenness centrality were calculated to quantify character prominence and influence within the network. In *By the Sea*, Saleh Omar exhibited the highest betweenness centrality score (0.41), indicating a mediating narrative position between otherwise weakly connected character clusters. Network density remained relatively low (0.27), reflecting the fragmented social environment characteristic of exile narratives.For instance, a

character with high betweenness centrality may serve as a link between different groups, indicating a mediating role in the narrative.

Validity and Reliability Measures

To ensure the validity and reliability of the computational analyses, several strategies were employed:

1. **Cross-Validation**: Topic modeling was validated through cross-validation, ensuring the stability of identified themes across different subsets of the corpus. The final LDA configuration achieved an average coherence score of 0.54 across validation folds, indicating moderate to strong semantic consistency between topic word clusters. Perplexity scores demonstrated convergence after successive iterations, suggesting model stability across corpus partitions.

2. **Manual Inspection**: Interpretations of topics and sentiment results were corroborated by manual inspection of sample texts, ensuring that computational results aligned with contextual understanding.

3. **Consistency Checks**: Social network graphs were reviewed for consistency, with multiple iterations conducted to verify the stability of character interactions and social structures.

By incorporating these measures, the study enhances the accuracy and reliability of the computational methods, ensuring that findings are both robust and interpretable.

Ethical Considerations

This study relies on publicly available texts, ensuring compliance with ethical standards in data usage. The study also acknowledges potential limitations in sentiment analysis, topic modeling, and SNA, as computational methods may sometimes miss nuances present in human interpretation. To address these limitations, findings were supplemented with qualitative analysis to provide a balanced and contextually aware perspective.

FINDINGS

The findings of this study reveal how computational methods illuminate the narrative complexity focusing on Abdulrazak Gurnah's *By the Sea* literary text. Through sentiment analysis, topic modeling, and social network analysis, the results highlight recurring themes of exile, displacement, and memory, while tracing the shifting emotional tone that shapes the novel's exploration of identity and belonging. The computational mapping of character interactions underscores the fragmented yet interconnected social networks within the text, reflecting the broader tensions of migration and cultural negotiation. These findings not only demonstrate the utility of digital approaches in uncovering latent structures but also provide new perspectives on Gurnah's narrative strategies, situating *By the Sea* within wider debates on postcolonial literature and transnational storytelling.

Sentiment Analysis: Emotional Trajectories and Character Dynamics

In *By the Sea* by Abdulrazak Gurnah, sentiment patterns align closely with pivotal narrative moments, with emotional shifts corresponding to key developments throughout the text. As the narrative progresses into the conflict phase (pp. 60–140), where Saleh Omar's past and Latif Mahmud's memories are interwoven through flashbacks, sentiment analysis captures extended stretches of subdued affect rather than abrupt fluctuations. This pattern reflects the characters' internal struggles and the burden of memory, emphasizing psychological tension over external catastrophe. Negative sentiment clusters dominate in sections recounting betrayal, displacement, and fractured identity, while occasional neutral or positive tones emerge in moments of reconciliation or reflection. Quantitatively, negative sentiment accounted for approximately 48% of analyzed segments within conflict-centered chapters, compared to 31% neutral and 21% positive sentiment classifications. These distributions reinforce the predominance of psychological tension and reflective tone throughout the narrative.

Toward the conclusion of the novel (pp. 200–245), sentiment trends return toward neutrality, marked by Saleh and Latif's eventual meeting and tenuous reconciliation. This supports the finding that resolution phases in

certain narratives emphasize reflection and closure without dramatic emotional reversal. Although *By the Sea* does not exhibit cyclical sentiment arcs typical of romance or prolonged suspense common in thriller genres, its sustained emotional restraint reinforces the broader observation that emotional pacing is shaped by narrative purpose and genre. Gurnah's strategy embeds emotional depth within memory and dialogue rather than overt dramatic events, underscoring exile as an affective condition where emotional intensity is cumulative rather than episodic.

The thematic structure identified through topic modeling in *By the Sea* by Gurnah follows an emotional arc that moves from introduction and setting, through conflict and climax, to resolution. In the early chapters (pp. 1–20) the author portrays Saleh Omar's arrival at Gatwick Airport and the gradual unfolding of his backstory are marked by predominantly neutral sentiment scores. Themes of personal history, memory, and relationships dominate, as Saleh Omar recounts his past life and losses (pp. 1–80). These clusters foreground intimate concerns of grief and identity, establishing the emotional foundation of the narrative.

As the story progresses into the middle sections (pp. 80–180), the thematic focus shifts toward conflict and power, with recurring references to immigration procedures, bureaucracy, and encounters with institutional authority. These clusters highlight the socio-political dimensions of exile, situating the novel within broader debates on justice, displacement, and cultural negotiation. The emotional arc intensifies here, as personal struggles intersect with systemic pressures, creating a sustained thematic climax. This is depicted through the characters as the narrative progresses into the conflict phase (pp. 60–140), where Saleh Omar's past and Latif Mahmud's memories are interwoven through flashbacks, sentiment analysis captures extended stretches of subdued affect rather than abrupt fluctuations.

Toward the conclusion (pp. 180–245), the thematic emphasis transitions to resolution, where questions of belonging, reconciliation, and cultural negotiation become more prominent. The computational analysis reveals clusters associated with dialogue, negotiation, and tentative closure, reflecting the novel's movement toward reflection rather than dramatic resolution. This thematic shift mirrors the emotional arc of the narrative, where the intensity of conflict gives way to a quieter exploration of identity and belonging.

Rather than relying on overt dramatic events, the novel embeds emotional depth within memory and dialogue, allowing sentiment to build gradually and persistently. This comparative pattern reinforces the idea that emotional pacing is shaped by narrative purpose and genre, with *By the Sea* presenting exile as an affective condition in which intensity is sustained and layered over time rather than expressed episodically.

Topic Modeling: Thematic Patterns and Recurring Motifs

The thematic structure identified through topic modeling is evident in *By the Sea* by Gurnah, where dominant themes such as conflict and power, social justice, and love and relationships emerge as enduring concerns across the narrative (pp. 1–245). Using Latent Dirichlet Allocation (LDA), five major topics were identified, each represented by clusters of co-occurring words that denote specific thematic areas. In the novel, these topics unfold in alignment with the emotional arc of introduction, conflict, and resolution.

In the opening sections of *By the Sea* (pp. 1–80), sentiment analysis shows neutral or mildly positive scores that reflect the calm tone typical of narrative beginnings, with clusters tied to memory, personal history, and relationships. Saleh Omar's recollections foreground grief, attachment, and identity, situating personal loss and fractured bonds within broader questions of belonging and emphasizing the intimate dimensions of exile. As the narrative develops (pp. 80–180), the focus shifts toward bureaucracy, identity, and encounters with institutional authority, highlighting themes of conflict, power, and social justice.

These thematic concerns not only underscore the socio-political contexts shaping the characters' experiences but also align with broader corpus findings that literary texts increasingly engage with questions of authority, justice, and displacement over time. Together, these phases illustrate how Gurnah's restrained emotional trajectory is layered with both personal and political dimensions, embedding exile within a continuum of memory, conflict, and systemic power.

In *By the Sea*, the thematic trajectory is thus, shaped by recurring clusters of conflict, power, social justice, and relationships, which intensify in the middle sections (pp. 80–180) through references to immigration procedures, bureaucratic encounters, and institutional authority. These moments underscore the socio-political dimensions of exile, showing how personal struggles intersect with systemic pressures and drive the narrative toward its thematic climax. The recurrence of words tied to displacement and negotiation situates the novel firmly within a postcolonial context. Toward the latter part of the text (pp. 180–245), the focus shifts from individual memory to broader questions of belonging, cultural negotiation, and migration, reinforcing temporal and cultural variations in thematic emphasis. Taken together, these shifts demonstrate how Gurnah weaves personal and socio-political narratives into a coherent arc, situating *By the Sea* within wider debates on exile, identity, and transnational belonging.

In the resolution phase (pp. 180–245), topic clusters move toward reconciliation, dialogue, and cultural negotiation. The focus shifts from individual memory to collective struggles of identity and belonging, as Saleh and Latif's eventual meeting embodies tentative closure. These themes reflect the novel's movement toward reflection rather than dramatic resolution, reinforcing the cumulative emotional arc identified in the sentiment analysis.

Overall, topic modeling reveals how *By the Sea* interweaves motifs of memory, conflict, and reconciliation into a unified thematic arc. The computational identification of these clusters offers a systematic perspective on Gurnah's narrative strategy, showing how he balances personal recollections with socio-political concerns and situates the novel within broader debates on exile, identity, and transnational storytelling.

Comparative Sentiment Patterns Across Genres

Across genres, sentiment patterns reveal distinct trajectories: romantic novels often display cyclical arcs tied to relationship dynamics, while mystery and thriller texts sustain prolonged neutral or low sentiment to reflect suspense and investigation (Moretti, 2013; Reagan et al., 2016; Jockers, 2014; Mohammad, 2016). Within this comparative framework, Gurnah's *By the Sea* stands out for its restrained and cumulative emotional arc. Neutral or mildly positive tones dominate the introduction (pp. 1–80), subdued negative sentiment emerges during the conflict phase (pp. 80–180) as exile, bureaucracy, and fractured identity take center stage, and the resolution (pp. 180–245) returns toward neutrality, emphasizing reflection and cultural negotiation rather than dramatic reversal. This pattern highlights how the novel embeds emotional depth within memory and dialogue, presenting exile as a sustained affective condition rather than one marked by episodic emotional shifts.

Comparative Postcolonial Context

To further contextualize the findings, selected sentiment and thematic patterns were compared with a secondary postcolonial narrative, Salman Rushdie's *Midnight's Children*. While Rushdie's text demonstrates higher sentiment volatility associated with political upheaval and historical transformation, *By the Sea* maintains sustained emotional restraint centered on memory and interpersonal dialogue. Topic modeling comparisons also reveal differing emphases: whereas *Midnight's Children* foregrounds national identity and historical rupture, Gurnah's narrative emphasizes displacement and intimate negotiations of belonging. This comparison suggests that postcolonial narratives vary significantly in emotional pacing and thematic focus, strengthening the generalizability of corpus-level observations.

Thematic Mapping of Gurnah's *By the Sea*

The computational analysis of thematic patterns across the corpus provides a useful framework for situating *By the Sea* within broader narrative trends. Below, each identified theme is explicitly linked to Gurnah's novel, demonstrating how it aligns with or diverges from the thematic patterns revealed through topic modeling.

Table 1: Main topics and Representative words associated with each theme.

Topic	Representative Words	Description	*By the Sea*, Gurnah's Alignment with Thematic Patterns revealed through Topic Modeling
Love and Relationships	affection, bond, marriage, family	relationships, family dynamics, and	Personal relationships in *By the Sea* are subtle but central. Saleh Omar's interactions with his family, connections in Zanzibar, and encounters with other characters reflect the

		romantic connections.	emotional and social bonds that shape his experiences of exile and belonging (pp. 10–50, 120–150). Gurnah portrays relationships as both stabilizing and tension-inducing forces, highlighting the emotional complexity of migration and displacement
Conflict and Power	war, battle, authority, leadership	Deals with struggles for power, conflicts, and themes of dominance or rebellion.	The novel foregrounds struggle of authority and systemic power, particularly through postcolonial and immigration bureaucracies. Saleh Omar experiences marginalization and negotiation of power, reflecting conflicts between individuals and institutions, as well as broader societal hierarchies (pp. 60–120, 180–200). These dynamics align with corpus-level patterns of conflict-focused narratives but are expressed in psychological and social terms rather than physical combat.
Nature and Solitude	solitude, reflection	Depicts introspection, natural landscapes, and themes of isolation.	Nature and solitude are especially prominent in *By the Sea*. Saleh Omar's reflections, his memories of Zanzibar, and the meditative depictions of landscapes convey emotional isolation and introspection (pp. 5–25, 150–170). Solitary spaces function as both literal and metaphorical settings for exile and identity formation, closely reflecting corpus-level patterns of sparse social networks with rich thematic depth.
Social Justice	rights, equality, oppression, reform	Addresses societal issues, justice, and transformative movements.	Social justice emerges as a central concern, particularly through experiences of immigration, exile, and systemic inequality. Saleh Omar navigates bureaucratic oppression and broader social injustices, revealing the intersection of personal and societal struggles (pp. 70–130, 190–240). These motifs echo corpus-level patterns in literature that address structural inequities and transformative social concerns.
Technology and Modernity	innovation, machinery, urban, future	Explores technological advancements, industrialization, and future-oriented themes.	Direct engagement with technology and modernity is minimal. The novel prioritizes memory, migration, and personal histories rather than contemporary industrial or technological change (pp. 200–220). This divergence demonstrates that thematic emphasis is context-dependent, influenced by Gurnah's focus on postcolonial and migratory experiences rather than industrial modernity.

Key Insight:

Among the five thematic areas identified, *By the Sea* most strongly engages with Nature and Solitude, Conflict and Power, and Social Justice, while Love and Relationships appears in more nuanced, emotionally driven forms. In contrast, Technology and Modernity is largely absent, reflecting the novel's historical, reflective, and postcolonial orientation. This thematic mapping shows how Gurnah's narrative resonates with corpus-wide patterns yet maintains a distinctive lens on exile, memory, and introspection. The findings highlight the novel's dual significance: it offers rich material for computational analysis while also sustaining the depth of traditional literary interpretation.

Theme Distribution and Temporal Variation

Analysis of thematic frequency across texts shows that love and relationships dominate 19th- and early 20th-century narratives (Moretti, 2013; Reagan et al., 2016), while conflict and power, often tied to war or

social hierarchy, feature prominently in politically or historically oriented novels (Jockers, 2014). By contrast, technology and modernity emerge mainly in late 20th-century works, reflecting literature's engagement with industrial and digital transformations (Hayles, 2005; Jameson, 1991).

Comparisons across periods reveal clear shifts: earlier texts emphasize nature, introspection, and social norms, whereas later works explore technology, urbanization, and individual agency (Williams, 1977; Moretti, 2013; Walkowitz, 2020). Post-1950 narratives increasingly foreground social justice and technology, mirroring real-world socio-political changes (Jameson, 1991; Hayles, 2005; Brouillette, 2022). These variations highlight literature's adaptability, showing how thematic focus evolves in response to historical and cultural contexts.

Placed within broader patterns of theme distribution and temporal variation, Gurnah's *By the Sea* both resonates with and diverges from corpus-wide trends. It strongly engages with themes of conflict, power, and social justice, aligning with works that explore political and historical contexts, while its restrained treatment of love and relationships reflects the emotional depth typical of postcolonial narratives rather than the cyclical arcs of 19th-century romantic fiction. The relative absence of technology and modernity situates the novel outside late 20th-century concerns with industrial and digital transformations, underscoring its historical and reflective orientation. At the same time, *By the Sea* exemplifies post-1950 shifts toward social justice, displacement, and identity negotiation, while sustaining earlier literary motifs of nature, solitude, and introspection. This dual positioning highlights Gurnah's ability to bridge traditional concerns with contemporary debates, situating the novel within wider conversations on exile, identity, and transnational belonging through a distinctly postcolonial lens.

Social Network Analysis: Character Relationships and Interaction Patterns

Social network analysis (SNA) of *By the Sea* reveals a sparse but thematically charged web of interactions, built around a few central figures whose encounters carry disproportionate weight in shaping the narrative. Saleh Omar's isolation dominates the early chapters (pp. 1–60), producing limited yet pivotal connections that mirror themes of exile, solitude, and introspection. His encounters with immigration officials (pp. 40–70) and Mr. Linford (pp. 50–75) highlight bureaucratic power structures, while the fraught relationship with Latif Mahmud (pp. 85–180) introduces conflict and eventual reconciliation.

Character centrality underscores this minimalist structure: Saleh Omar exhibits high degree centrality, reflecting the novel's focus on exile as a deeply personal condition, while Latif Mahmud demonstrates notable betweenness centrality, bridging Saleh's personal exile with collective memory and historical context. Peripheral figures such as immigration officers and Mr. Linford hold low centrality but embody systemic obstacles, reinforcing the imbalance of power and the bureaucratic weight of exile.

By visualizing the novel through SNA, it becomes clear that Gurnah uses sparse but decisive interactions to emphasize solitude, displacement, and the tension between personal memory and collective history. The network's sparseness reinforces the isolating force of exile, while the centrality of Saleh and Latif dramatizes how a handful of relationships can carry profound thematic significance.

Combined Analysis: Interplay Between Themes and Character Dynamics

Integrating insights from sentiment analysis, topic modeling, and social network analysis provides a nuanced perspective on how narrative structures and themes interact to shape reader experience. Key findings suggest that character interactions and emotional arcs are intricately aligned with thematic patterns, particularly in genres that emphasize relational dynamics or individual transformation. In *By the Sea*, the combined application of sentiment analysis, topic modeling, and social network analysis reveals how narrative structure and thematic focus interact to shape the reader's experience.

The emotional trajectory of the novel aligns with traditional plot phases: Saleh Omar's early isolation (pp. 1–60) reflects subdued, introspective sentiment, while his fraught encounters with immigration officials (pp. 40–70) and Mr. Linford (pp. 50–75) introduce emotional lows tied to bureaucratic power. The eventual confrontation and dialogue with Latif Mahmud (pp. 85–180) generate emotional fluctuations—anger, resentment, and tentative reconciliation—that mirror broader themes of exile, memory, and contested identity.

Topic modeling highlights dominant themes of conflict and power, social justice, and love and relationships in nuanced, restrained forms, while the absence of technology and modernity situates the novel outside late 20th-century industrial or digital concerns, emphasizing instead historical and postcolonial contexts.

Social network analysis further underscores the novel's sparse relational web, with Saleh Omar exhibiting high degree centrality (pp. 1–245) and Latif Mahmud notable betweenness centrality (pp. 90–180). Peripheral figures such as immigration officers (pp. 40–70) and Mr. Linford (pp. 50–75) hold low centrality but embody systemic obstacles, reinforcing the imbalance of power and the bureaucratic weight of exile.

Together, these findings show how Gurnah's minimalist network structure, fluctuating emotional arcs, and thematic focus on displacement and memory converge to dramatize exile as both an isolating force and a deeply consequential condition

DISCUSSION

This section interprets its findings within the broader context of digital humanities and computational literary analysis. By examining sentiment trajectories, thematic structures, and character dynamics across a diverse corpus, it demonstrates how computational methods can enrich narrative analysis. Applied to Abdulrazak Gurnah's *By the Sea* (2001), these approaches reveal restrained emotional arcs, recurring themes of exile and displacement, and sparse but pivotal character networks. Situating these results within debates in narrative theory and literary scholarship highlights the value of computational tools as complements to traditional close reading, offering empirical insights into narrative form and social dynamics while remaining attentive to cultural and historical specificity.

Sentiment analysis of *By the Sea* reveals a restrained emotional trajectory rather than the pronounced affective fluctuations common in Western corpus studies. The opening sections (pp. 1–80) are marked by calm, neutral tones shaped by silence and retrospection. The middle chapters (pp. 80–180) introduce subdued negative sentiment tied to bureaucracy, fractured identity, and exile. Extended segments of low-intensity sentiment complicate assumptions that narrative coherence must follow clear emotional peaks and troughs, as suggested by Freytag's Pyramid (Freytag, 1894/1986). This demonstrates how computational tools can identify texts that resist dominant emotional arcs, particularly postcolonial narratives foregrounding trauma, repression, and fragmentation (Piper & So, 2015; Sun et al., 2023; Hämäläinen et al., 2025).

Topic modeling situates *By the Sea* within broader thematic trends of migration, memory, loss, institutional authority, and belonging, aligning with late 20th- and early 21st-century literary engagements with globalization and displacement (Blei, Ng, & Jordan, 2003; Goldstone & Underwood, 2014). Yet algorithmic abstraction risks flattening distinctions between texts that engage similar themes differently. In Gurnah's novel, migration is inseparable from narrative voice, temporal disjunction, and ethical ambiguity (pp. 90–180), dimensions that exceed the explanatory capacity of topic frequency alone. This tension underscores the need to combine computational outputs with close reading to recover thematic nuance and contextual specificity (Moretti, 2013; Walkowitz, 2020; Ved, 2025).

Social network analysis (SNA) further illuminates narrative structure and character dynamics. Unlike texts characterized by dense social interaction, *By the Sea* exhibits a sparse network reflecting the isolating effects of exile. Saleh Omar's high degree centrality (pp. 1–245) underscores exile as a deeply personal condition, while Latif Mahmud's betweenness centrality (pp. 90–180) bridges personal exile with collective memory and historical context. Peripheral figures such as immigration officers (pp. 40–70) and Mr. Linford (pp. 50–75) embody systemic obstacles, reinforcing bureaucratic power. These findings support narratological theories that conceptualize characters as functional elements within narrative systems (Bal, 2009), while demonstrating how computational analysis can systematically expose such functions (Hämäläinen et al., 2025).

Crucially, *By the Sea* foregrounds East African histories of migration and displacement, situating Zanzibar and the wider Indian Ocean world as central to its narrative. The novel reflects how colonial legacies and postcolonial bureaucracies shape exile, resonating with broader African literary traditions that interrogate identity, belonging, and transnational mobility (Mazrui, 2024; Ndlovu-Gatsheni, 2023). This regional

grounding complicates computational models trained primarily on Western corpora, raising ethical concerns about epistemic authority and representation (Brouillette, 2022; Thakare, 2025).

The novel's reliance on irony, understatement, and unreliable narration complicates sentiment classification, while its layered temporal structure challenges linear models of narrative progression. These features expose the limitations of approaches that prioritize surface-level linguistic features over narrative context and historical situatedness. Moreover, applying computational tools to postcolonial texts raises broader ethical concerns regarding epistemic authority, particularly when analytical models underrepresent African and East African literary traditions (Brouillette, 2022;. Neeraj, 2023; Journal of Computational Literary Studies, 2024; Thakare, 2025).

CONCLUSION

Ultimately, *By the Sea* reinforces the argument that computational methods are most effective as exploratory and complementary tools rather than autonomous interpretive systems. By testing these methods against a text that resists dominant narrative conventions, this case study demonstrates that computational analysis can illuminate structural and thematic patterns while also revealing its own interpretive boundaries. In this way, Gurnah's novel functions both as an object of analysis and as a methodological critique, contributing to ongoing debates in the digital humanities about scale, interpretation, and the role of human judgment in data-driven literary scholarship.

REFERENCES

1. Akimoto, T., & Ogata, T. (2025). A narratological approach for narrative discourse: Implementation and evaluation of a system based on Genette and Jauss. Graduate School of Software and Information Science, Iwate Prefectural University.
2. Alasmari, J., Alzyoudi, M., Alshehri, M., Alshammari, R., & Aldakan, R. (2025). An automated predictive model for evaluating narrative cohesion in children's stories: A computational linguistic approach considering Gérard Genette's narrative structure theory. International Journal of Adolescence and Youth, 30(1),
3. Algee-Hewitt, M., et al., (2016). Computational Methods and Textual Analysis in the Digital Humanities. Stanford Literary Lab.
4. Antoniak, M. (2025). Computational Approaches to Narratives. Pioneer Centre for AI, University of Copenhagen.
5. Aristotle, (1999). Poetics. Translated by S. H. Butcher. Dover Publications.
6. Babenko, A., & Athavale, A. (2025). AI and literary studies: Global perspectives on computational criticism. Journal of Digital Humanities, 14(2), 45–63.
7. Babenko, I., & Athavale, V. A. (2025). Methods of literary text analysis with the help of artificial intelligence. In Lecture Notes in Networks and Systems (Vol. 1616).
8. Bakhtin, M. M. (1981). The Dialogic Imagination: Four Essays. University of Texas Press.
9. Bal, M. (2009). Narratology: Introduction to the theory of narrative (3rd ed.). University of Toronto Press.
10. Bamman, D. (2020). Narrative theory and computational modeling of narrative. Annual Review of Linguistics, 6, 337–356.
11. Bamman, D. (2020). Natural language processing for literary text. Annual Review of Linguistics, 6, 447–468.
12. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3, 993–1022.
13. Bloom, H., (1994). The Western Canon: The Books and School of the Ages. Riverhead Books.
14. Bookclb. (2025, July 21). Contemporary literature in 2025: Trends reshaping the literary world. Reader's Corner.
15. Brooks, P., (1992). Reading for the Plot: Design and Intention in Narrative. Harvard University Press.
16. Brouillette, S. (2022). Literature and the creative economy. Stanford University Press.
17. Chaturvedi, S., Iyyer, M., & Daumé III, H. (2017). Unsupervised learning of evolving relationships between literary characters. Proceedings of the AAAI Conference on Artificial Intelligence, 31(1).

18. Chaturvedi, S., Srivastava, S., & Mitchell, T. (2017). Modeling evolving relationships between characters in literary novels. Proceedings of the AAAI Conference on Artificial Intelligence, 31(1).
19. Computational Humanities Research. (2025). Special Issue on Computational Narratology. Cambridge University Press.
20. Da, N. Z. (2019). The computational case against computational literary studies. Critical Inquiry, 45(3), 601–639.
21. De Pauw, G., & De Schryver, G.-M. (2008). Improving the computational morphological analysis of Swahili by means of a dedicated corpus. Lexikos, 18, 312–331.
22. Demir, D. (2025). Computational Narratology: A New Approach to Storytelling. Current Happenings Across STEM Magazine.
23. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL-HLT 2019, 4171–4186.
24. Eagleton, T., (1976). Marxism and Literary Criticism. University of California Press.
25. Elson, D. K., Dames, N., & McKeown, K. R. (2010). Extracting social networks from literary fiction. Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, 138–147.
26. Eve, M. P. (2024). The digital humanities and literary studies: Critical perspectives. Cambridge University Press.
27. Fish, S., (1980). Is There a Text in This Class? The Authority of Interpretive Communities. Harvard University Press.
28. Flanders, J., & Jannidis, F. (2019). The shape of data in digital humanities: Modeling texts and text-based resources. Routledge.
29. Freytag, G. (1986). Freytag's Technique of the Drama (E. J. MacEwan, Trans.). Scholarly Press. (Original work published 1894).
30. Freytag, G., (1863). Die Technik des Dramas. Translated as Freytag's Technique of the Drama by Elias J. MacEwan. The University of Chicago Press.
31. Genette, G. (1980). Narrative Discourse: An Essay in Method. Cornell University Press.
32. Goldstone, A., & Underwood, T. (2014). The longue durée of literary prestige. Modern Language Quarterly, 75(3), 383–407.
33. Goldstone, A., & Underwood, T. (2014). The Quiet Transformations of Literary Studies: Computational Methods and the Digital Humanities. New Literary History, 45(3), 359–384.
34. Gomathi, R. D., Murugan, J., Kavitha, P., & Gomathi, B. S. (2024). AI-driven literary analysis: Exploring the impact of artificial intelligence on text interpretation and criticism. Journal of Machine and Computing.
35. Gurnah, A. (2001). By the Sea. Bloomsbury.
36. Hämäläinen, M., et al. (2025). Advances in computational narratology: Sentiment and structure in literary corpora. Journal of Computational Literary Studies.
37. Hämäläinen, M., Öhman, E., Bizzoni, Y., Miyagawa, S., & Alnajjar, K. (2025). Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities. Association for Computational Linguistics.
38. Hatzel, H. O., Stiemer, H., Biemann, C., & Gius, E. (2023). Machine learning in computational literary studies. Information Technology, 65(4–5).
39. Hayles, N. K. (2005). My mother was a computer: Digital subjects and literary texts. University of Chicago Press.
40. Hurskainen, A. (2016). Swahili language and literature in the digital age. Nordic Journal of African Studies, 25(2), 87–104.
41. Jameson, F. (1991). Postmodernism, or, the cultural logic of late capitalism. Duke University Press.
42. Jockers, M. L. (2013). Macroanalysis: Digital methods and literary history. University of Illinois Press.
43. Jockers, M. L. (2014). Macroanalysis: Digital methods and literary history. University of Illinois Press.
44. Kirschenbaum, M.G., (2007). "The Remaking of Reading: Data Mining and the Digital Humanities." Literary and Linguistic Computing, 22(3), pp. 276-290.
45. Lee, J., & Yeung, C. (2022). Dynamic character network analysis in narrative texts using contextual embeddings. Digital Scholarship in the Humanities, 37(3), 789–805.
46. Lee, J., & Yeung, C. (2022). Dynamic network analysis of narrative structures. Digital Scholarship in the Humanities, 37(4), 987–1005.

47. Li, D., Wu, K., & Lei, V. L. C. (2024). Applying Topic Modeling to Literary Analysis: A Review. Linguistics & Semiotics, De Gruyter Brill.
48. Li, J., et al. (2024). Advances in Computational Narratology: Machine Learning Approaches to Story Structure. Digital Scholarship in the Humanities.
49. Livingston, T. J. (2024). Artificial intelligence in literary analysis: Exploring computational approaches to text interpretation. Routledge.
50. Masakhane NLP. (2020). Building NLP for African languages through community research. Proceedings of the African NLP Workshop.
51. Mazrui, A. (2024). African migrations and cultural identity. Nairobi University Press.
52. Mohammad, S. M. (2016). Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In Emotion measurement (pp. 201–237). Elsevier.
53. Moretti, F., (2013). Distant Reading. Verso Books.
54. Moretti, F., (2005). Graphs, Maps, Trees: Abstract Models for a Literary History. Verso.
55. Ndlovu-Gatsheni, S. J. (2023). Decoloniality and African futures. Routledge.
56. Neeraj. (2023). The trauma of imprisonment: An analysis of Abdulrazak Gurnah's By the Sea. Research Journal of English Language and Literature, 11(2), 26–35.
57. Ogunleye, B., Lancho Barrantes, B. S., & Zakariyyah, K. I. (2025). Topic Modelling Through the Bibliometrics Lens and Its Technique. Artificial Intelligence Review, 58(74). Springer Nature.
58. Paulson, D., & Leonard, S. (2025). Applications of NLP in computational poetics and literary analysis: A systematic review (2000–2025). Preprints.org.
59. Piper, A., & Jean So, R., (2015). "The Imaginative Power of Genre." Post45 Contemporaries.
60. Piper, A., & So, R. J. (2015). Quantification and the digital humanities. Critical Inquiry, 42(2), 1–25.
61. Princeton CDH. (2024). Computational approaches to Nigerian literature: Experiments in NLP for texts in Yoruba and Efik. Center for Digital Humanities Project Report.
62. Propp, V. (1968). Morphology of the folktale. University of Texas Press.
63. Ramsay, S. (2011). Reading machines: Toward an algorithmic criticism. University of Illinois Press.
64. Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. EPJ Data Science, 5(1), 31.
65. Rebora, S. (2026). Artificial intelligence and the future of literary studies. Digital Humanities Quarterly, 20(1).
66. Schreibman, S., Siemens, R., & Unsworth, J., (2004). A Companion to Digital Humanities. Blackwell.
67. Srivastava, P., Mamodiya, U., Dutta, P. K., & Paliktzoglou, V. (2025). Advanced AI-driven thematic and narrative analysis: Exploring deep learning in literature. In Digital Transformation in Interdisciplinary Sciences (pp. 179–194). Springer.
68. Sun, C., et al. (2023). Transformer-based sentiment analysis in literary texts. Digital Scholarship in the Humanities, 38(1), 112–130.
69. Sun, C., Qiu, X., Xu, Y., & Huang, X. (2023). How to Fine-Tune BERT for Text Classification? Chinese Journal of Electronics, 32(1), 1–11.
70. Tanasescu, C. (Ed.). (2024). Literature and computation: Platform Intermediality, hermeneutic modeling, and analytical-creative approaches. Routledge
71. Thakare, A. (2025). Postcolonial digital humanities: Narratives of displacement. Routledge.
72. Thakare, P. D. (2025). Recent trends in the study of English literature: A comprehensive overview. International Journal of Advance and Applied Research, 6(18).
73. Todorov, T. (1971). The poetics of prose. Cornell University Press.
74. Underwood, T. (2019). Distant horizons: Digital evidence and literary change. University of Chicago Press.
75. Ved, P. (2025). Digital humanities and the study of literature. International Journal of Social Impact, 10(1), 19–29.
76. Ved, P. (2025). Digital humanities: Transforming literary studies through computational methods. International Journal of Digital Scholarship, 12(1), 45–62.
77. Walkowitz, R. L. (2020). Born translated: The contemporary novel in an age of world literature. Columbia University Press.
78. Williams, R. (1977). Marxism and literature. Oxford University Press.
79. Woloch, A., (2003). The One vs. the Many: Minor Characters and the Space of the Protagonist in the Novel. Princeton University Press.

80. Yadav, D. (2024). The role of artificial intelligence in literary analysis: A computational approach to understanding literary styles. International Journal of Education and Knowledge Studies, 3(9), 1–12.
81. Yadav, D. (2024). The role of artificial intelligence in literary analysis: A computational approach to understand literary styles. International Journal of Education and Knowledge Studies, 3(9), Article 006.
82. Yang, L. (2025). Machine learning and cultural analytics: International approaches. Literary and Linguistic Computing, 40(2), 211–229.
83. Yékú, J. (2022). Digital African literatures and the coloniality of data. Cambridge Journal of Postcolonial Literary Inquiry, 9(3), 345–364.
84. Zenodo. (2024). Computational approaches to Nigerian literature: Analyzing texts in Yoruba and Efik. Digital Humanities Conference Proceedings.