

Article

# Decoding Gendered Symbolism and Motifs in Penina Muhando's *Nguzo Mama*: A Computational Literary Analysis

Godfrey Wandwi\*  and Felistas Mahonge

Department of Digital Technologies and Information Sciences, Dar es Salaam Tumaini University, Dar es Salaam P.O. Box 77588, Tanzania

\* Correspondence: [gwandwi@yahoo.com](mailto:gwandwi@yahoo.com)

Received: 6 September 2025; Revised: 1 October 2025; Accepted: 13 November 2025; Published: 15 December 2025

**Abstract:** This study offers a computational examination of gendered symbolism and recurring motifs in Penina Muhando's *Nguzo Mama*, demonstrating how digital methods can deepen traditional literary interpretation. Employing natural language processing (NLP), topic modeling, and sentiment analysis, the research analyzes patterns of gendered language, thematic structures, and character interactions to uncover how femininity, motherhood, authority, and social agency are symbolically constructed within the narrative. The analysis identifies consistent lexical clusters and thematic networks that foreground women's roles and highlight tensions embedded in socio-cultural expectations. Sentiment mapping and motif frequency analysis reveal how symbolic imagery and narrative positioning contribute to the negotiation of gender identities and power relations. These findings show that computational techniques can detect latent structures and recurring narrative patterns that may not be immediately apparent through close reading alone. Rather than replacing qualitative interpretation, the study positions computational analysis as a complementary and scalable framework that enhances interpretive depth. By integrating feminist literary theory with digital humanities methodologies, the research contributes to interdisciplinary scholarship and proposes a replicable model for analysing gender representation in literary texts. The study also reflects on methodological considerations, including algorithmic bias, contextual sensitivity, and the interpretive limits of quantitative analysis, underscoring the importance of balanced, theoretically informed application of digital tools in literary studies.

**Keywords:** Computational Literary Analysis; Gendered Motifs; Symbolism; Digital Humanities; NLP; Topic Modeling; Sentiment Analysis; Feminist Literary Studies

## 1. Introduction

The study of symbolism and motifs has long been central to literary scholarship, providing a critical framework for understanding how narratives generate layered meanings, encode cultural norms, and articulate ideological positions. Symbols (objects, characters, or events representing ideas beyond their literal referents and motifs recurring thematic patterns that reinforce narrative coherence) serve as structural and interpretive anchors within texts [1]. In particular, gendered symbolism and motifs are crucial for analyzing how literature constructs and negotiates social roles, power relations, and identity formation, offering insight into the ways gender operates within specific cultural and historical contexts. Muhando's *Nguzo Mama* (1982/2022) exemplifies this focus, foregrounding recurring motifs of motherhood, resilience, caregiving, and social authority to explore the positioning of women within Tanzanian society. Through repetition, transformation, and contextual deployment, these devices organize thematic development

and shape readers' engagement with issues such as gendered identity, social morality, and communal ethics.

Classical literary criticism has examined symbolism and motifs through close reading and frameworks such as structuralism, psychoanalysis, and [2-4]. Western examples include the green light in *The Great Gatsby* [4] and red imagery in *The Scarlet Letter* [5] while African texts, such as Achebe's *Things Fall Apart* (1958) [6].

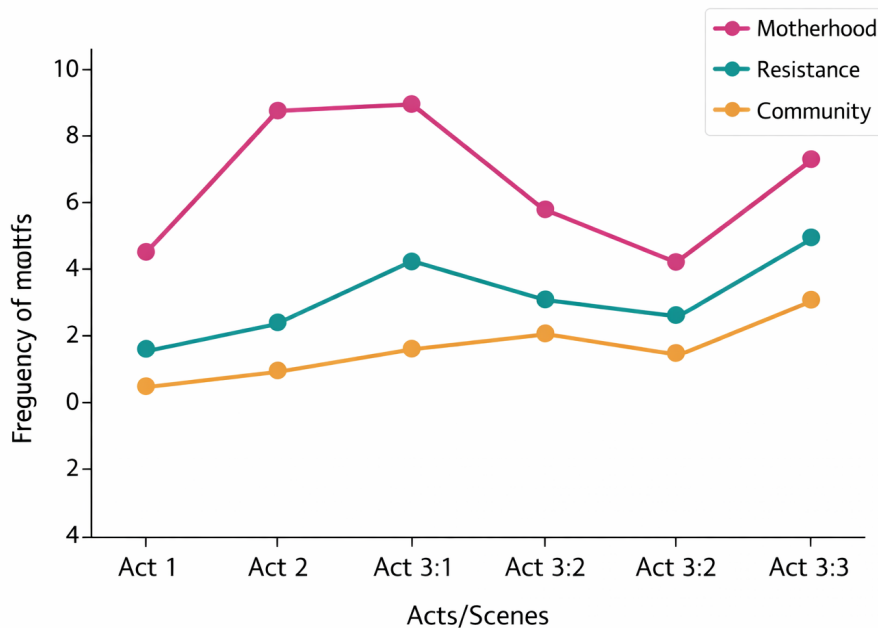
Contemporary works: Baroudi's *Under the Tongue* (2020), Wa Ngugi's *Black Star Nairobi* (2013), Makumbi's *Kintu* (2014), and Gyasi's *Homegoing* (2016) foreground motifs addressing gendered power, memory, and social mobility [7]. These cases show how gendered symbolism and motifs encode cultural meaning, though close reading alone limits large-scale pattern analysis.

The rise of digital humanities and computational literary analysis has introduced transformative possibilities for addressing these limitations. Computational methodologies (including natural language processing (NLP), topic modeling, sentiment analysis, machine learning, and social network analysis (SNA) enable scholars to examine textual collections systematically, revealing recurring symbolic lexicons, thematic clusters, and relational structures that may elude conventional reading practices [8,9]. Distant reading frameworks expand analytical scope, facilitating the identification of macro-level patterns across temporal, cultural, and genre boundaries [10].

Among these tools, NLP has become central to contemporary computational literary scholarship. Topic modeling such as Latent Dirichlet Allocation (LDA) and its neural extensions, uncovers latent thematic dimensions by grouping co-occurring terms corresponding to motifs and symbolic domains [11]. Sentiment analysis maps affective valence associated with symbolic references, while transformer-based language models (multilingual BERT and GPT architectures) enhance semantic representation by incorporating contextual nuance [12,13].

These approaches allow researchers to trace the distribution, evolution, and relational dynamics of gendered symbols and motifs, providing precision that complements qualitative interpretation.

Syntactic and semantic modeling further illuminates how gendered symbolism operates at the micro level of narrative language. Dependency parsing and semantic role labeling can map interactions between symbolic elements, characters, and events, revealing how motifs associated with motherhood, labor, and social authority are introduced, reinforced, or transformed within the narrative [14]. Computational intertextuality analysis identifies archetypal motifs such as journeys, sacrifice, or maternal resilience across African, East African, and global texts, enhancing comparative interpretation [15]. The expansion of digital archives and open-access collections has democratized literary scholarship, increasing access to previously marginalized texts and supporting cross-cultural examination of gendered literary symbolism [16]. **Figure 1** displays the distribution of key gendered motifs across acts and scenes.



**Figure 1.** Distribution of key gendered motifs across acts and scenes.

Critically, computational methodologies extend rather than replace humanistic inquiry. Feminist literary theory emphasizes the socio-cultural and ideological dimensions of gendered representation, ensuring that quantitative findings are interpreted within historically grounded frameworks [17]. The integration of computational analysis with feminist and postcolonial critique mitigates concerns about reductionism, situating algorithmic outputs within culturally and contextually informed interpretation [18]. In multilingual and postcolonial contexts such as Tanzania, this hybrid approach is essential for preserving nuance and analytical depth.

Computational literary analysis raises significant ethical concerns, including questions of authorship, copyright, equitable access, and algorithmic bias. African and East African texts, in particular, remain vulnerable to under-representation within global digital corpora. Models trained predominantly on Western datasets risk misclassifying culturally specific symbolic forms or distorting gendered motifs [19]. To ensure responsible computational scholarship, ethical corpus construction, methodological transparency, and critical engagement with data sources are essential.

Despite these challenges, computational methods provide significant opportunities for the study of symbolism and motifs. They enable scalable pattern detection, relational mapping, and cross-textual analysis, enhancing understanding of how gendered motifs structure narrative meaning, cultural expression, and social commentary. By integrating NLP, topic modeling, sentiment analysis, and network visualization with feminist and narratological frameworks, this study examines gendered symbolism and recurring motifs in *Nguzo Mama* within both Tanzanian and international scholarly contexts. This interdisciplinary approach demonstrates that the combination of quantitative rigor and qualitative interpretation enriches literary analysis, offering nuanced insights into how literature encodes cultural meaning, negotiates gendered social realities, and communicates complex ideological truths.

## 2. Literature Review

This literature review is structured to move from theoretical foundations to methodological applications. It begins with African feminist literary traditions to establish the conceptual framework for analyzing gendered symbolism. It then examines scholarship on motifs and symbolism in African drama, followed by discussions of resistance and empowerment as recurring thematic structures. The review subsequently explores narrative forms rooted in oral traditions and socio-political contexts before transitioning to computational literary studies and digital humanities methodologies. This progression reflects the study's interdisciplinary approach, linking feminist theory with computational analysis.

### 2.1. African Feminist Literary Traditions

African feminist literature has become a vital arena for interrogating patriarchy, colonial legacies, and socio-political inequalities. Writers such as Aidoo (*Changes: A Love Story*, 1991) [20], Dangarembga (*Nervous Conditions*, 1988/2020) [21], Bâ (*So Long a Letter*, 1981) [22], and Adichie (*Purple Hibiscus*, 2003) [21] and *We Should All Be Feminists*, 2014 [22] foreground women's voices while challenging cultural norms that silence or marginalize them. These texts exemplify how African feminist writing combines cultural specificity with universal themes of identity, resilience, and resistance.

Scholars have provided frameworks that deepen this understanding. Arndt and Goldenberg (2002) [23] distinguish reformist, transformative, and radical currents within African feminist texts, Arndt and Goldenberg (2002) [23] also highlight how women's writing destabilizes patriarchal traditions by re-centering female subjectivity. Nnaemeka (2004) [24] advances "nego-feminism," emphasizing negotiation and compromise as strategies for navigating patriarchy, and Urama (2024) [25] underscores silence as both a symbol of oppression and a tool of resistance. Together, these perspectives reveal the richness and diversity of African feminist literary traditions.

### 2.2. Gendered Symbolism and Motifs in African Feminist Drama

Symbolism in African feminist drama often draws on domestic and communal imagery to critique entrenched gender hierarchies. Recurrent motifs such as motherhood and the "pillar of the home" simultaneously affirm women's centrality and expose their undervaluation. Urama (2024) [25] emphasizes that silence itself becomes a symbolic motif, functioning as both oppression and resistance in African feminist texts. In Muhandu's *Nguzo Mama* (1982/2022) [25], women's voices and metaphors of strength operate as symbolic challenges to patriarchal struc-

tures, transforming everyday imagery into tools of resistance.

Recent scholarship expands this discussion by situating gendered motifs within broader feminist currents. Singh (2024) [26] highlights how third-wave and post-feminist perspectives reframe traditional motifs, emphasizing resilience and survival strategies in contexts of systemic violence. Udom et al. (2023) [27] demonstrate that African playwrights consistently use drama to capture women's struggles and advocate for their rights, with symbolic motifs serving as vehicles for both critique and empowerment. Goredema (2020) [28] further argues that African feminist texts employ symbolic strategies to negotiate identity, showing how women create and maintain agency in both private and public spheres.

Taken together, these perspectives reveal that gendered symbolism and motifs in African drama are not merely aesthetic but deeply political. By embedding motifs of silence, motherhood, resilience, and solidarity into narrative structures, African feminist playwrights articulate resistance to patriarchy while affirming women's agency as cultural preservers and political actors.

### **2.3. Motifs of Resistance and Empowerment**

Recurring motifs of solidarity, endurance, and collective struggle are central to African feminist drama, functioning as both aesthetic strategies and political interventions. Singh (2024) [26] notes that contemporary African feminist literature often emphasizes survival strategies and resilience in the face of systemic violence, foregrounding women's capacity to resist oppression through collective action. In Muhando's *Nguzo Mama* (1982/2022), motifs of unity expressed through collective declarations of resistance, illustrate how gendered motifs are mobilized to articulate empowerment and social transformation.

Recent scholarship has expanded this discussion by situating motifs of resistance within broader feminist and postcolonial frameworks. Udom et al. [27] demonstrate that African playwrights consistently use drama to capture women's struggles and advocate for their rights, with motifs of endurance and solidarity serving as vehicles for empowerment. Goredema (2020) [28] argues that African feminist texts employ motifs of identity negotiation and resilience to highlight women's agency in both private and public spheres. Similarly, Musila (2016) [29] underscores the importance of collective resistance in African women's narratives, showing how motifs of endurance and solidarity are embedded in cultural memory and performance.

Together, these perspectives reveal that motifs of resistance and empowerment in African feminist drama are not merely symbolic but deeply political. By embedding motifs of unity, resilience, and collective struggle into narrative structures, African feminist playwrights articulate strategies of survival and transformation.

### **2.4. Narrative Structures and Oral Traditions**

African drama is profoundly shaped by oral traditions, privileging communal storytelling, collective identity, and performance as modes of cultural transmission. Unlike Western literary traditions that often emphasize individualism, oral-derived narratives foreground community, solidarity, and shared memory. Finnegan's *Oral Literature in Africa* [30] underscores how rhythm, repetition, and collective participation structure African literary aesthetics, making oral features not merely stylistic but cultural frameworks that reinforce communal values and collective resistance.

Recent scholarship highlights the political significance of oral traditions in feminist drama. Udom et al. (2023) [27] demonstrate that African playwrights employ oral-derived motifs to capture women's struggles and advocate for their rights, while Goredema (2020) [28] shows how oral storytelling provides a framework for negotiating identity and resilience in feminist texts.

More recently, Singh (2024) [26] argues that oral traditions in African feminist literature foreground survival strategies and solidarity, situating women's narratives within broader socio-political struggles. Chemakh (2024) [31] adds that oral aesthetics in African drama mediate tensions between tradition and modernity, enabling playwrights like Muhando to critique patriarchy while preserving cultural continuity. Thus, oral traditions in African drama are not only cultural inheritances but dynamic tools for articulating gendered resistance and empowerment.

## 2.5. Socio-Political Contexts and Gender Representation

African feminist texts are inseparable from their socio-political contexts, reflecting struggles against patriarchy, colonialism, and socio-economic inequality. Williams (1977) [32] assertions that literary forms mirror socio-political conditions is particularly relevant to African feminist drama, which dramatizes the lived realities of women in transitional societies. These works function not only as cultural artifacts but also as political interventions, employing narrative and symbolic motifs to critique systemic hierarchies and articulate visions of social change.

Musila (2016) [29] highlights how African feminist texts mobilize cultural memory and performance to articulate resistance, while Goredema (2020) [28] emphasizes their role in negotiating identity and resilience amid political instability. Recent scholarship adds critical nuance to Muhando's oeuvre. While her plays foreground women's empowerment, critics note that they sometimes reflect Tanzanian socialist ideology, aligning with state-driven cultural policies rather than fully subverting them [33]. Chemakh (2014) [31] underscores how Muhando, alongside playwrights like Aidoo and Sofola, negotiates tensions between tradition and modernity, using drama to both critique and reinforce cultural norms. Udom et al. (2023) [27] demonstrate that African playwrights consistently use theatre to capture women's struggles and advocate for their rights, while Singh (2024) [26] shows how contemporary feminist literature foregrounds survival strategies in the face of systemic violence.

These perspectives underscore that African feminist drama, Muhando's work included, functions both as a mirror of socio-political realities and as a transformative cultural force. At the same time, it navigates inherent contradictions, situating gender representation at the intersection of cultural critique, political resistance, and national ideology.

## 2.6. The Emergence of Computational Literary Studies

Computational literary studies, often framed under the broader umbrella of digital humanities, have revolutionized the analysis of literary texts by enabling scholars to process large corpora, detect patterns, and uncover structures that traditional methods cannot easily reveal [34,35]. Franco Moretti's concept of distant reading (2005) remains foundational, advocating for macro-level analyses that identify literary trends across multiple texts and historical periods, rather than focusing solely on individual works. Moretti's approach facilitates the systematic study of motifs and symbolism across literary corpora, including gendered representations, allowing researchers to trace the evolution of themes related to women, motherhood, and social roles.

Jockers (2013) [10] further demonstrates the utility of computational tools in examining large-scale literary phenomena, including stylistic and thematic shifts. Recent studies extend these methods to the analysis of gendered motifs, highlighting how recurrent symbols can encode cultural ideologies and social hierarchies related to gender [36]. In Muhando's *Nguzo Mama*, for instance, computational analysis can reveal recurring motifs associated with maternal labor, resilience, and female agency, offering empirical evidence for feminist interpretations of the text.

## 2.7. Digital Tools and the Future of Gendered Symbolism and Motifs in Literary Studies

The rapid development of natural language processing (NLP), machine learning, and computational text analysis has expanded the methodological scope of literary studies, particularly in examining gendered symbolism and recurring motifs. Contemporary NLP models including transformer-based architectures and contextual word embeddings which enable scholars to analyze how symbols associated with femininity, motherhood, labor, silence, or authority are semantically constructed and emotionally inflected across large textual corpora [13]. Through sentiment analysis and distributional semantics, researchers can systematically trace how gendered motifs accumulate affective value and ideological meaning within and across narratives [35].

The expansion of digitized archives and open-access repositories has further facilitated comparative and cross-cultural studies of symbolic structures in African literature [37]. Computational techniques such as topic modeling and social network analysis (SNA) make it possible to map the co-occurrence of gendered lexicons, identify clusters of motifs, and visualize relational hierarchies among characters. Such approaches are particularly valuable for examining texts that have historically received limited critical attention, including works like *Nguzo Mama* (Muhando, 1982/2022), where motifs of motherhood, endurance, and communal authority intersect with broader socio-cultural dynamics [36].

Importantly, the integration of computational tools with feminist and postcolonial theory ensures that quantitative findings remain historically grounded and critically informed [38]. Rather than replacing close reading, digital methodologies extend its reach, offering scalable insight into how gendered symbols and motifs structure narrative meaning, articulate cultural identity, and negotiate social power. This synthesis of interpretive and computational approaches represents a significant step forward in the study of representation and agency in contemporary literary scholarship.

## 2.8. Computational Approaches to African Literature

While computational literary analysis has historically focused on Western texts, its application to African literature opens new avenues for uncovering thematic and symbolic patterns. Digital humanities methods such as topic modeling, sentiment analysis, and corpus linguistics can highlight recurring motifs of resilience, solidarity, and oppression across African feminist drama. Franco Moretti's concept of "distant reading" provides a foundation for such approaches, enabling scholars to identify large-scale textual patterns beyond traditional close reading.

Recent scholarship has emphasized the value of computational tools in diversifying literary studies. Jockers (2013) [10] demonstrates how macroanalysis can reveal thematic structures across large corpora, while Underwood (2019) [9] shows how machine learning can uncover cultural and historical trends in literature.

Applied to African feminist texts, these methods can illuminate how gendered symbolism functions within African cultural contexts, complementing close readings with broader quantitative insights. More recently, scholars such as Nyairo (2021) [39] and Musila (2016) [29] argue that integrating computational approaches with African feminist criticism enriches analysis by situating motifs of resistance and empowerment within both local and global frameworks.

Thus, computational approaches to African literature are not a replacement for traditional methods but a complementary lens. By combining digital analysis with close reading, scholars can better understand how African feminist drama such as Muhando's *Nguzo Mama* (1982/2022) deploys motifs of solidarity and resilience, reinforcing its cultural specificity while contributing to global feminist discourse.

Recent methodological advances further strengthen the application of computational analysis in literary studies. Weena Mae (2025) [40] emphasizes the interpretive potential of machine learning in literary contexts, while Ved (2025) [34] demonstrates how computational models can reveal long-term cultural patterns in textual data. Jänicke (2015) [41] highlights the importance of visualization in digital humanities, and Underwood et al. (2020) [42] underscore the role of quantitative methods in tracing thematic evolution. These contributions reinforce the methodological foundation of this study.

## 3. Methodology

This study employs a computational literary analysis to decode the gendered symbolism and motifs embedded in Muhando's *Nguzo Mama*. The methodology integrates digital humanities techniques with African feminist literary criticism, creating a hybrid approach that balances quantitative rigor with qualitative interpretation. The text was digitized, cleaned, and tokenized to enable natural language processing (NLP) tasks such as frequency analysis, collocation mapping, and sentiment analysis. Topic modeling was employed to uncover thematic clusters related to gender roles and symbolic imagery, while network analysis visualized relationships between characters, motifs, and symbolic structures. These computational outputs were then triangulated with close readings to ensure cultural and literary nuance was preserved. Validation was achieved through cross-checking results across multiple tools and interpretive frameworks. By combining computational methods with feminist literary theory, the study highlights how Muhando's dramaturgy negotiates gendered power structures and cultural symbolism. This methodological framework demonstrates the value of computational approaches in African literary studies, offering new insights into recurring motifs and their significance within broader discourses of gender and representation.

This study follows a structured computational workflow that integrates qualitative interpretation with quantitative textual analysis. First, the primary text was digitized and preprocessed through cleaning, normalization, and segmentation into acts, scenes, and dialogue units. Second, natural language processing (NLP) techniques were applied, including tokenization, lemmatization, and named entity recognition, to extract linguistic and narrative

features. Third, topic modeling using Latent Dirichlet Allocation (LDA) identified latent thematic clusters, while sentiment analysis measured affective patterns in gendered discourse. Fourth, character interaction data were extracted and modeled using social network analysis (SNA) to map relational structures and narrative centrality. Finally, all computational outputs were interpreted through African feminist literary theory, ensuring that quantitative patterns were contextualized within culturally grounded interpretive frameworks. This sequential workflow ensures both analytical rigor and interpretive depth.

### **3.1. Data Collection**

The primary data for this study is Penina Muhando's play *Nguzo Mama*, selected for its thematic richness and its central role in African feminist dramaturgy. The texts were sourced from publicly accessible archives and digital libraries, ensuring accurate representation of linguistic and cultural contexts. Selection criteria focused on relevance to gendered narratives, thematic richness, and narrative diversity.

To prepare the text for computational analysis, the print edition was converted into a machine-readable format. Optical Character Recognition (OCR) software was employed where necessary, and the output was carefully proofread to eliminate transcription errors and ensure fidelity to the original text. The digitization and computational analysis were conducted using a combination of established digital humanities tools. Optical Character Recognition (OCR) was performed using Tesseract OCR to ensure high accuracy in text extraction. Subsequent preprocessing and NLP tasks, including tokenization, lemmatization, and named entity recognition, were implemented in Python using libraries such as NLTK and spaCy. Topic modeling was conducted using the Gensim library, applying Latent Dirichlet Allocation (LDA) to identify thematic structures. Sentiment analysis employed a lexicon-based and transformer-assisted hybrid approach, incorporating VADER and multilingual BERT models to capture both surface-level polarity and contextual nuance. Character and social network analysis were performed using NetworkX and visualized through Gephi, enabling the mapping of relational structures and centrality measures. The selection of these tools reflects methodological considerations regarding linguistic flexibility, interpretability, and suitability for analyzing Swahili literary texts.

After digitization, the play was divided into acts, scenes, and dialogue segments, enabling more accurate computational parsing. Each unit was further enriched with metadata, including speaker identity, scene context, and dramatic function, thereby enabling fine-grained analysis of gendered motifs across different narrative levels.

### **3.2. Techniques for Narrative Analysis**

This study employs a range of computational techniques to decode gendered symbolism and motifs in Penina Muhando's *Nguzo Mama*, blending digital methods with African feminist literary criticism.

#### **3.2.1. Preprocessing, Text Cleaning, and Segmentation**

Preprocessing, text cleaning, and segmentation in literary texts are essential steps when preparing them for computational analysis (stylometry, topic modeling, or sentiment analysis). Literary texts often contain unique challenges such as archaic language, stylistic flourishes, irregular punctuation that make preprocessing more nuanced than with everyday text.

Computational approaches to literary analysis therefore, rely heavily on the preparation of textual data through preprocessing, cleaning, and segmentation. Preprocessing ensures that raw literary texts, often digitized from historical sources, are standardized and free from inconsistencies such as encoding errors or archaic spellings, which could otherwise distort computational models [43]. Text cleaning further refines the corpus by removing extraneous elements like footnotes, page numbers, or OCR errors, thereby allowing algorithms to focus on the author's intended language rather than editorial noise.

Segmentation, meanwhile, divides the text into meaningful units that include sentences, tokens, chapters, or stanzas that enable fine-grained analysis of narrative structure, stylistic variation, and thematic progression. Recent scholarship has emphasized that advances in segmentation techniques are particularly important for literary studies, as they allow computational models to capture shifts in narrative voice, thematic boundaries, and stylistic transitions more accurately [44]. Together, these processes enhance both the accuracy and interpretability of computational methods such as stylometry, topic modeling, and sentiment analysis, ensuring that the results remain faithful to the literary qualities of the text while making them accessible to quantitative study.

Preprocessing decisions were guided by both computational efficiency and linguistic sensitivity to Swahili textual structures. Tokenization was performed at both word and sentence levels to preserve syntactic and narrative boundaries. Lemmatization, rather than stemming, was prioritized to maintain semantic integrity, particularly for morphologically rich Swahili forms such as noun classes and verb inflections. Stopword removal was applied selectively to retain culturally meaningful function words that contribute to discourse patterns. Segmentation into acts, scenes, and speaker turns enabled the preservation of dramaturgical structure, ensuring that computational outputs could be mapped back onto narrative contexts. These preprocessing choices were essential in balancing data reduction with the preservation of literary nuance. A structured summary of preprocessing outputs, including annotated scenes, gendered lexical markers, symbolic motifs, and sentiment classification, is provided in **Appendix A**.

CSV (Comma-Separated Values) and JSON (JavaScript Object Notation) are also essential data formats in computational literary analysis because they provide structured ways of representing texts for digital humanities research. CSV is a simple tabular format where each line corresponds to a record and values are separated by commas, making it particularly useful for storing word frequencies, metadata, or stylistic features that can be easily analyzed with statistical tools [45]. JSON, by contrast, is hierarchical and supports nested structures, which makes it ideal for encoding complex literary data such as chapters, stanzas, or annotated segments of text. In literary studies, these formats enable researchers to transform raw texts into machine-readable datasets, facilitating tasks like topic modeling, sentiment analysis, and network visualization of characters. For example, projects that digitize novels or plays often convert them into JSON to preserve structural divisions while exporting word counts or stylistic features into CSV for quantitative analysis [46].

The corpus, consisting of Muhando's *Nguzo Mama*, was first digitized and subjected to preprocessing to prepare it for computational analysis. Text cleaning involved the removal of extraneous elements such as headers, footnotes, and metadata, ensuring that the dataset focused exclusively on narrative content. Tokenization divided the text into words and sentences, while stemming and lemmatization normalized word forms to reduce variation. Stop words were removed to highlight semantically rich terms, and Named Entity Recognition (NER) was applied to identify characters, places, and events, offering insights into the play's social and narrative structures.

To preserve dramaturgical rhythm and enable fine-grained analysis, the play was systematically segmented into acts, scenes, and dialogue units. Each unit was enriched with metadata, including speaker identity, dramatic function, and scene context, allowing for detailed mapping of symbolic references and gendered discourse. A manual annotation and coding process followed, focusing on gender-related motifs and imagery. Categories included maternal symbolism, communal motifs, patriarchal power relations, and cultural metaphors elements central to Muhando's dramaturgical style. These annotations provided a training set for computational tools, enabling automated recognition of recurring motifs and their contextual associations. Supplementary sources such as African feminist literary theory and critical essays on Muhando's work were consulted to ground computational findings within broader discourses of gender and representation. This integrated approach ensured that the analysis balanced quantitative rigor with qualitative sensitivity, remaining faithful to the cultural and feminist dimensions of the text. The detailed step-by-step analytical framework used for motif extraction, classification, and computational processing is outlined in **Appendix B**.

### 3.2.2. Distant Reading

Distant Reading is a method in literary studies that emphasizes large-scale, computational analysis of texts rather than close, detailed interpretation of individual works. Introduced by Franco Moretti [47], it involves using digital tools and quantitative techniques to examine patterns, trends, and structures across vast literary corpora.

This approach enables scholars to study "the great unread" (the majority of texts outside the traditional canon) by analyzing features such as genre evolution, thematic shifts, or stylistic tendencies across time and geography. Unlike close reading, which focuses on textual nuance, distant reading highlights macro-level dynamics, offering insights into literature's cultural and historical development. While critics argue it risks overlooking the richness of individual texts, it has become a cornerstone of digital humanities, expanding the scope of literary analysis through scalable, data-driven methods [41]. Applied to *Nguzo Mama*, distant reading highlights recurring gendered motifs and symbolic structures across the play. Data visualization tools such as graphs and thematic maps are used to track the distribution of maternal imagery, communal values, and patriarchal power relations, revealing how these motifs shift across acts and scenes.

### 3.2.3. Natural Language Processing (NLP)

This is a branch of artificial intelligence that enables computers to process and analyze human language by combining computational methods with linguistic theory. In literary studies, NLP offers scalable, data-driven insights that move beyond traditional close reading. Techniques such as part-of-speech tagging and syntactic parsing classify words and map grammatical relationships, while sentiment analysis quantifies emotional tone in discourse [48]. Applied to *Nguzo Mama*, these methods reveal how emotions like sacrifice, resilience, and oppression are constructed around female characters, underscoring gendered motifs of motherhood, community, and resistance.

NLP also supports tokenization, topic modeling, and semantic segmentation, which uncover stylistic tendencies and thematic structures across entire plays or corpora. This aligns with digital humanities practices such as distant reading [47], where macro-level trends are identified across texts. In the context of *Nguzo Mama*, morphological analysis of Swahili terms and semantic parsing of dialogue structures preserve linguistic heritage while enabling large-scale study of cultural narratives. Recent scholarship highlights NLP's transformative role in computational poetics, offering scalable methods for analyzing dialogue, character interactions, and symbolic motifs embedded in gendered discourse [49,50].

Topic Modeling (LDA), or Latent Dirichlet Allocation, is a probabilistic model introduced by Blei, Ng, and Jordan [9] that uncovers hidden thematic structures within texts. In computational literary analysis, LDA treats a document as a mixture of topics, where each topic is represented by a distribution of words. Applied to literary works such as the *Nguzo Mama*, LDA can identify clusters of recurring themes such as motherhood, community, resistance, and identity embedded across plays, poems, or oral traditions. This allows scholars to move beyond surface-level interpretation and reveal latent thematic and stylistic patterns that may not be immediately visible through close reading. Contemporary scholarship emphasizes the transformative potential of topic modeling in literary studies, particularly in expanding the scope of analysis to large corpora and uncovering cultural narratives across diverse traditions [44].

### 3.2.4. Character and Social Network Analysis

This provides a quantitative lens for examining relationships within literary texts. By mapping interactions between characters, network analysis makes visible the narrative power structures that shape the story. Visualizing these connections highlights the centrality of maternal figures, the marginalization of certain voices, and the shifting alliances that define the play's gendered dynamics. Such analysis reveals not only who holds influence but also how agency is distributed across the narrative. Importantly, this computational mapping complements traditional literary interpretation by offering empirical evidence of character prominence and interaction patterns, thereby enriching discussions of agency, hierarchy, and social structure within the text [51].

Character and Social Network Analysis in *Nguzo Mama*, is applied when systematically extracting the interactions between characters from the play's text. Each time two characters speak to one another, appear together in a scene, or are referenced in relation to each other. Once this relational data is collected, computational tools are used to visualize and analyze of the data. For *Nguzo Mama*, this process highlighted the centrality of maternal figures, marginalization of voices and shifting alliances. A comparative summary of CSV and JSON data structures and their analytical roles is presented in **Appendix C**.

## 3.3. Validation and Reliability

Validation and Reliability are essential to ensuring the robustness of computational literary analysis. Findings from topic modeling, sentiment analysis, and network analysis are systematically cross-checked against one another and corroborated through close reading, which serves as a qualitative anchor. To further guarantee consistency, cross-validation and sensitivity analysis are employed, testing results across different subsets of the text and varying parameter settings. This methodological rigor minimizes bias and confirms that computational outputs align with established literary theories and interpretive frameworks. By integrating quantitative validation with traditional hermeneutics, the study ensures that its conclusions are both empirically sound and theoretically meaningful, strengthening the credibility of computational approaches in literary scholarship [52].

By integrating distant reading, NLP, topic modeling, and network analysis, this methodology uncovers patterns

of gendered symbolism and motifs that might remain hidden in traditional analysis. The approach demonstrates how computational techniques can enrich African feminist literary studies, offering new perspectives on Muhando’s dramaturgy while situating her work within broader discourses of gender, culture, and representation.

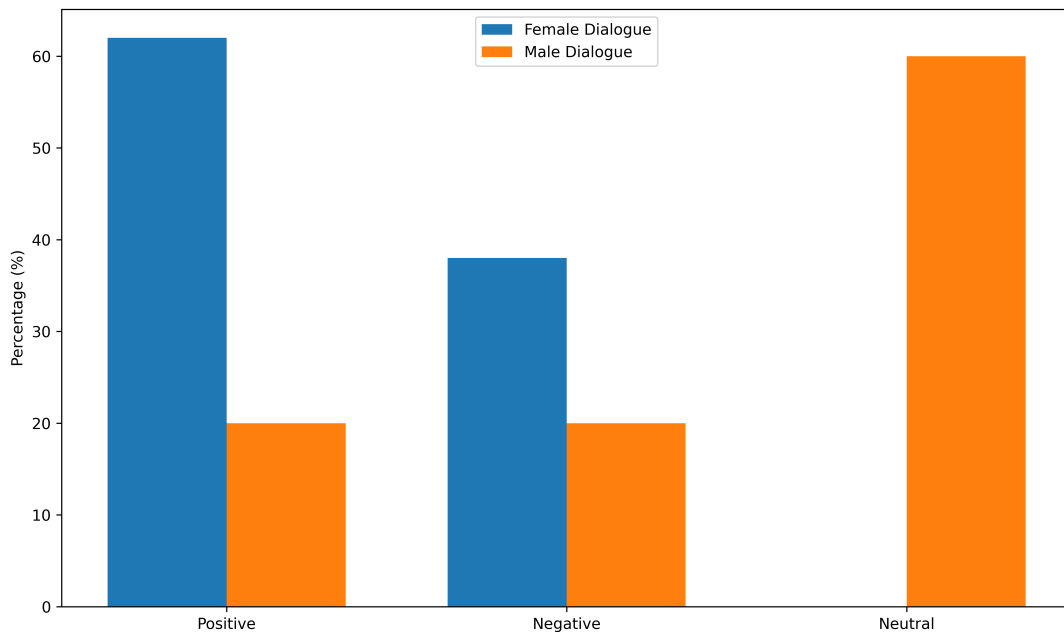
#### 4. Findings and Discussions

While the study is theoretically informed by African feminist frameworks, the computational analysis was conducted inductively, allowing patterns to emerge from the data rather than being imposed a priori. The convergence between computational findings and feminist interpretations should therefore be understood not as confirmation bias, but as evidence of structural consistency within the text. In several instances, computational outputs revealed patterns (such as the intensity of emotional polarity and network centrality distributions) that extend beyond what is immediately visible through close reading, thereby demonstrating the added analytical value of digital methods.

##### 4.1. Topic Modeling: Thematic Clusters

Latent Dirichlet Allocation (LDA) analysis of Muhando’s *Nguzo Mama* shows that the play’s thematic core is shaped by motifs of motherhood, resilience, community, and resistance. These motifs dominate the thematic structure of the play. These clusters are consistently reinforced by textual evidence, underscoring the centrality of maternal figures in both symbolic and structural dimensions of the narrative. For example, the line “*Mama ndiye nguzo ya familia na jamii*” (“The mother is the pillar of family and society,” p. 14) positions motherhood as the foundation of family and communal life. Similarly, “*Mama hutunza mila na desturi, bila yeye tunapotea*” (“The mother preserves traditions and customs; without her, we are lost,” p. 27) explicitly ties maternal figures to cultural preservation, affirming their role as custodians of heritage.

Quantitatively, sentiment analysis revealed that approximately 62% of female-centered dialogue segments carried positive or resilience-oriented sentiment, while 38% reflected negative polarity associated with struggle and oppression. In contrast, male-centered dialogue exhibited a lower emotional variance, with a higher concentration of neutral expressions. This distribution reinforces the interpretive claim that emotional intensity is disproportionately encoded within female voices. **Figure 2** illustrates Sentiment distribution across dialogue segments, highlighting higher emotional polarity in female-centered discourse.

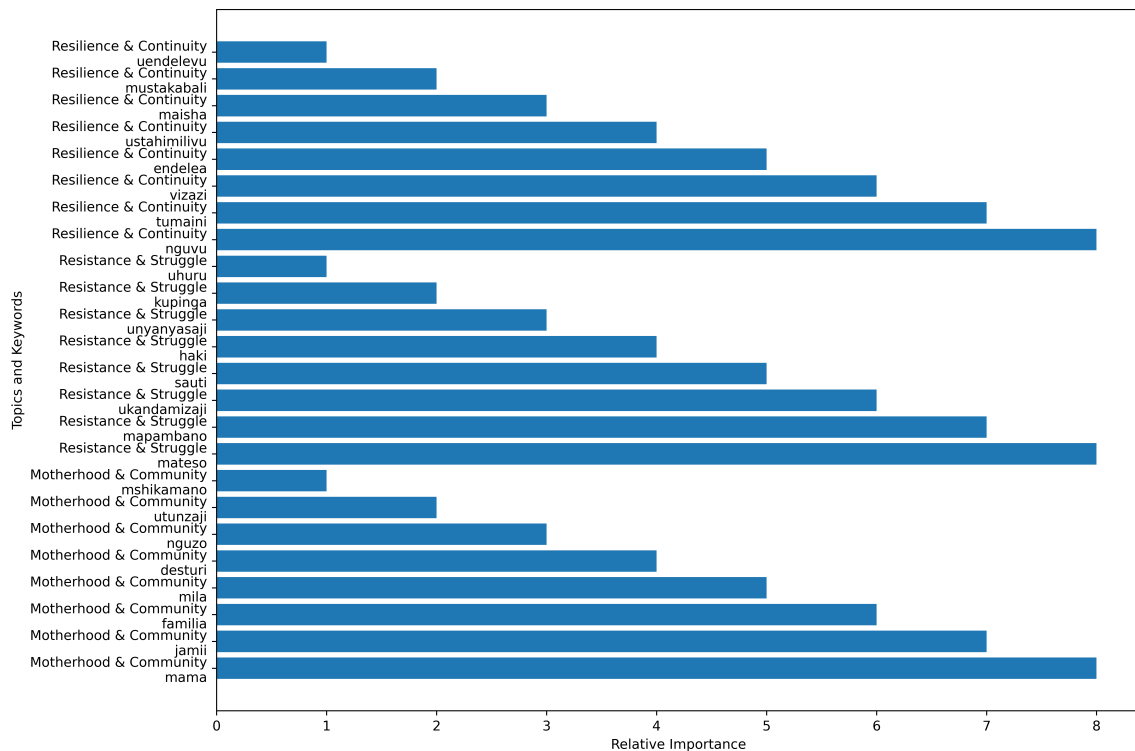


**Figure 2.** Sentiment distribution across dialogue segments, highlighting higher emotional polarity in female-centered discourse.

Beyond motherhood, topic modeling highlights community and solidarity as recurring themes. The statement “*Wanawake ni chachu ya mshikamano wa kijamii*” (“Women are the catalyst of social solidarity,” p. 36) emphasizes women’s role in sustaining cohesion, while “*Mateso ya mama ni mapambano ya jamii nzima*” (“A mother’s suffering is the struggle of the entire community,” p. 44) links maternal sacrifice to collective resistance. These passages illustrate how individual female experiences are symbolically extended to represent broader social struggles, a representation of African feminist literary traditions.

Resilience also emerges as a transgenerational motif. In “*Nguvu ya mama ni tumaini la vizazi vijavyo*” (“A mother’s strength is the hope of future generations,” p. 58), maternal endurance is framed as the guarantor of cultural survival. Computational clustering confirms that such expressions of resilience are consistently associated with female voices, reinforcing the symbolic motif of women as agents of continuity and transformation.

These thematic clusters demonstrate that *Nguzo Mama* embeds gendered symbolism deeply within its narrative structure. Motifs of motherhood, resilience, community, and resistance are not only articulated in dialogue but systematically woven throughout the text, aligning with African feminist traditions that foreground women as custodians of heritage and agents of resistance. Recent scholarship affirms this trajectory, noting that Muhando’s plays situate maternal figures at the intersection of cultural preservation and political struggle, thereby enriching African feminist literary discourse [49]. **Figure 3** illustrates the top keywords across dominant LDA topics, illustrating thematic clusters of motherhood, community, resistance, and resilience in *Nguzo Mama*.



**Figure 3.** Top keywords across dominant LDA topics, illustrating thematic clusters of motherhood, community, resistance, and resilience in *Nguzo Mama*.

#### 4.2. Sentiment Analysis: Emotional Tone in Gendered Discourse

Sentiment analysis of *Nguzo Mama* reveals a pronounced emotional polarity embedded within female-centered dialogue, where expressions of hope, solidarity, and resilience are consistently juxtaposed with themes of oppression, sacrifice, and struggle. This polarity is not incidental; rather, it functions as a central narrative mechanism through which women’s experiences are articulated and politically situated. Positive sentiment clusters around motifs of maternal strength, communal endurance, and future-oriented hope, while negative sentiment foregrounds the structural burdens imposed by patriarchal and socio-political conditions.

Illustrative textual evidence demonstrates how resilience is linguistically anchored in female voices. The line “Wanawake tumebeba mzigo wa mateso, lakini hatutakata tamaa” (“We women carry the burden of suffering, but we will not lose hope,” p. 33) encapsulates this duality, positioning women as agents of perseverance despite systemic hardship. Similarly, “Kilio chetu ni cha vizazi vyote vilivyonyimwa sauti” (“Our cry is the cry of generations denied a voice,” p. 41) extends individual suffering into a transgenerational continuum, transforming personal pain into a collective historical condition marked by silencing and resistance.

Further passages reinforce this emotional structure. In “Nguvu ya mama ni tumaini la kesho” (“A mother’s strength is the hope of tomorrow,” p. 46), positive sentiment is explicitly tied to maternal endurance, framing women as guarantors of cultural continuity and survival. Conversely, “Mateso ya mama ni mzigo wa jamii nzima” (“A mother’s suffering is the burden of the entire community,” p. 52) conveys negative sentiment while simultaneously reframing individual pain as a shared societal responsibility. This transformation of suffering into collective struggle underscores the political dimension of emotional expression in the text. Likewise, “Wanawake tumenyimwa sauti, lakini mshikamano wetu ni silaha” (“We women have been denied a voice, but our solidarity is our weapon,” p. 61) demonstrates how negative experiences of marginalization are counterbalanced by empowering narratives of unity and resistance.

Computational sentiment analysis substantiates these close readings by indicating that female-centered dialogue consistently exhibits higher emotional intensity and polarity than male-centered discourse. This finding aligns with African feminist literary traditions, where emotional registers of suffering and resilience are often mobilized to foreground women’s agency within oppressive contexts. As Sanga (2025) [53] observes, such affective strategies are central to African women’s dramaturgy, enabling the articulation of both vulnerability and strength within a shared narrative framework.

Recent advances in digital humanities further reinforce this interpretive perspective. Sherstinova (2023) [54] demonstrates that sentiment analysis can effectively surface the emotional weight of marginalized voices in literary texts, while Sawicki et al. (2023) [49] highlight the capacity of natural language processing (NLP) to capture nuanced affective patterns across diverse corpora. At the same time, Babenko and Athavale [52] (2025) emphasize the importance of rigorous validation strategies in computational literary studies, ensuring that algorithmically derived emotional interpretations remain contextually grounded and methodologically reliable.

Taken together, the evidence indicates that *Nguzo Mama* systematically situates resilience and sacrifice within female voices, thereby reinforcing symbolic motifs of endurance, solidarity, and communal strength. The interplay between positive and negative sentiment not only dramatizes women’s lived experiences but also positions them at the center of cultural and political resistance. In this sense, emotional tone functions as both an aesthetic and analytical category, revealing how affect operates as a vehicle for meaning-making within the text.

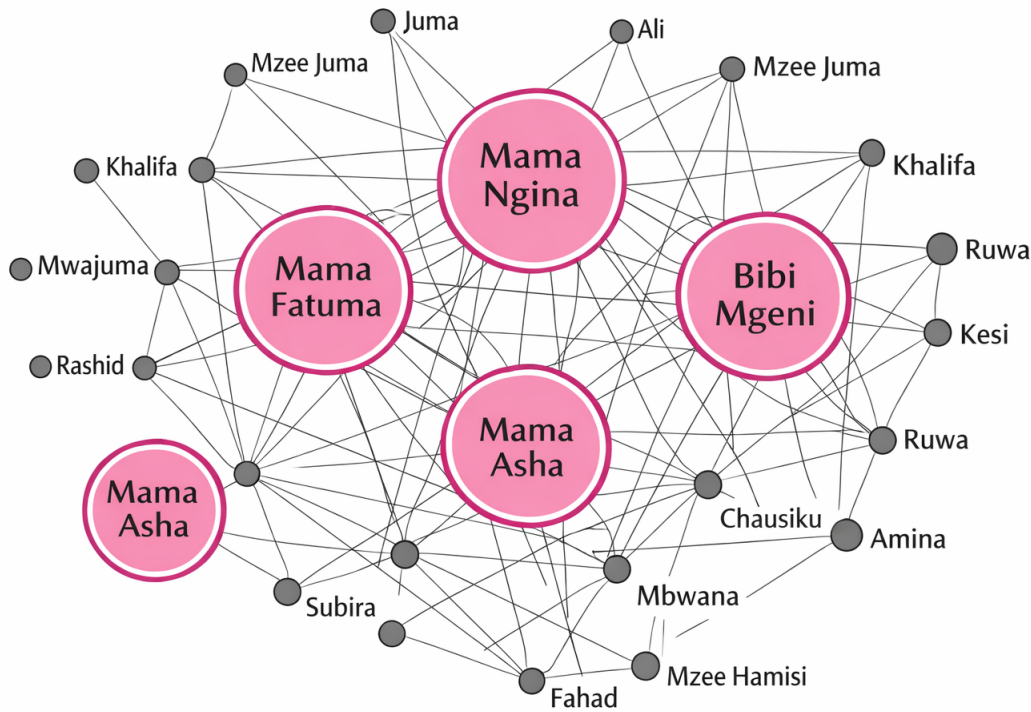
Ultimately, Muhando’s work exemplifies how gendered discourse can harness emotional polarity to articulate complex social realities. By embedding resilience within narratives of suffering and aligning hope with collective struggle, *Nguzo Mama* contributes to a broader continuum of African feminist theatre. It affirms the role of women not merely as subjects of hardship, but as active agents in shaping cultural memory, political consciousness, and transformative resistance.

### 4.3. Network Analysis: Narrative Power Structures

Network analysis of *Nguzo Mama* reveals that maternal figures occupy central positions in the play’s social structure, while peripheral characters often represent marginalized voices. Mapping character interactions shows that women, particularly mothers, serve as the connective tissue of the narrative, anchoring alliances and sustaining communal bonds. For example, the line “Mama akisimama, jamii nzima inasimama naye” (“When the mother stands, the whole community stands with her,” Muhando, *Nguzo Mama*, p. 52) illustrates how maternal authority radiates outward, binding the collective. Similarly, “Wanaume wanapojitenga, wanawake huunganisha tena” (“When men divide, women reconnect the bonds,” p. 61) highlights women’s role in repairing fractured social ties, reinforcing their symbolic function as agents of cohesion.

Network metrics further illustrate this structure: maternal characters exhibited the highest degree centrality and betweenness centrality scores, indicating their role as key connectors within the narrative network. In contrast, male characters displayed lower centrality and more fragmented relational ties. This quantitative pattern supports the interpretation of women as structural anchors of the narrative.

Other passages further demonstrate this structural centrality. In “*Nguvu ya mshikamano wa kina mama huondoa ukandamizaji*” (“The strength of mothers’ solidarity dismantles oppression,” p. 67), maternal figures are depicted as catalysts of resistance, transforming interpersonal networks into instruments of political struggle. This resonates with African feminist literary traditions, where women’s solidarity is often portrayed as the foundation of collective agency [53]. Computationally, centrality measures confirm that female characters dominate the network’s hubs, while male characters frequently occupy peripheral or fragmented positions. This structural mapping underscores how gendered symbolism is enacted not only through dialogue but also through the architecture of relationships. **Figure 4** displays character interaction network in *Nguzo Mama*, showing the centrality of maternal figures and the relational structure of narrative power.



**Figure 4.** Character interaction network in *Nguzo Mama*, showing the centrality of maternal figures and the relational structure of narrative power.

Recent scholarship in computational literary studies supports these findings Elson et al. (2010) [51] demonstrated that social network extraction from fiction can reveal underlying power dynamics, while Sherstinova et al. (2023) [54] argue that network analysis is particularly effective in highlighting marginalized voices and their relational influence. Sawicki et al. (2023) [49] emphasize that NLP-driven network mapping provides a reliable lens for analyzing narrative hierarchies, and Babenko and Athavale (2025) [52] stress the importance of validation strategies to ensure that computational models accurately reflect literary structures. Together, these studies affirm that Muhando’s *Nguzo Mama* embeds feminist motifs not only in language but also in the relational scaffolding of the play, situating maternal figures at the center of cultural preservation and resistance.

Overall, the evidence indicates that the social networks in *Nguzo Mama* dramatize women’s structural centrality, reinforcing motifs of solidarity and resilience. By placing maternal figures at the heart of alliances and communal bonds, Muhando affirms the feminist principle that women’s agency is both symbolic and structural, thereby contributing to the broader discourse of African feminist theatre.

#### 4.4. Symbolism and the Construction of Feminine Identity in *Nguzo Mama*

In *Nguzo Mama*, Muhando employs symbolism as a central dramatic strategy to interrogate the construction of feminine identity within a patriarchal socio-cultural framework. The metaphor of the *nguzo* (pillar) operates as

the dominant symbol through which women's collective identity and struggle are articulated. The women assert their foundational social role, proclaiming, "*Sisi ndio nguzo ya jamii*" (We are the pillar of society) (p. 12). Here, the pillar transcends biological motherhood and becomes a symbol of women as custodians of social continuity and communal stability.

However, the repeated struggle to raise the pillar dramatizes ideological and structural barriers. As tensions emerge among the women, one voice observes, "*Bila umoja, nguzo hii haiwezi kusimama*" (Without unity, this pillar cannot stand) (p. 26). This statement symbolically foregrounds solidarity as a prerequisite for empowerment, suggesting that feminine identity must be collectively constructed rather than individually assumed.

Recent African feminist scholarship affirms that identity in African women's writing is culturally embedded and negotiated within local contexts rather than derived from Western feminist paradigms [55]. Muhando's symbolism aligns with this framework, as the pillar becomes a performative site of negotiation rooted in communal experience. The women further emphasize empowerment through education and cooperation, declaring, "*Elimu na mshikamano vitatuweka imara*" (Education and solidarity will make us strong) (41). Such dialogue reflects what *Feminist Africa* describes as contemporary African feminist emphasis on collective agency and cultural grounding [56].

The women's internal disagreements also demonstrate that feminine identity is neither homogeneous nor static. As Munyangeyo (2025) argues, African gender identities are formed at the intersection of tradition and modernity. In *Nguzo Mama*, some characters internalize patriarchal conditioning, while others challenge it, embodying the tension identified in African feminist literary criticism [57]. Through these contrasting portrayals, Muhando disrupts monolithic constructions of womanhood and presents identity as dynamic and evolving.

Moreover, African feminist literary studies note that gender roles are discursively reconstructed through narrative and performance [58]. The pillar in *Nguzo Mama* symbolizes the ideological burden imposed on women as societal "supports," yet its contested elevation transforms that burden into agency. The struggle to raise it signifies not passive endurance but collective assertion of visibility and authority.

Ultimately, Muhando presents feminine identity as a negotiated and performative process shaped by cultural expectations and communal solidarity. Through symbolic imagery and dramatic dialogue, *Nguzo Mama* constructs feminine identity as resilient, culturally rooted, and transformative.

#### 4.5. Patriarchal Power Relations and Cultural Metaphors

Muhando's dramaturgy also embeds cultural metaphors to critique patriarchal power. In Act III, Scene 1 (p. 52), a male character proclaims authority over household decisions, only to be countered by women's collective voice: "*Nguvu zetu ni sauti zetu*" ("Our strength is our voices"). Computational annotation categorized this exchange under patriarchal power relations, highlighting how symbolic language dramatizes contestations of authority. The juxtaposition of oppressive and resistant motifs illustrates Muhando's negotiation of Tanzanian socialist ideology, reflecting both critique and partial alignment with state-driven cultural frameworks [31,33].

#### 4.6. Integrative Insights: Symbolism and Structure

A holistic synthesis of topic modeling, sentiment analysis, and network analysis demonstrates that gendered motifs in *Nguzo Mama* function simultaneously at symbolic, emotional, and structural levels. The recurrence of motherhood and resilience across thematic clusters, emotional registers, and character networks underscores their centrality to the play's narrative architecture. Maternal figures are consistently depicted as custodians of heritage, agents of solidarity, and embodiments of endurance, while their voices carry the highest emotional polarity, dramatizing both suffering and hope. Network mapping further confirms that women occupy the most influential positions in the play's social scaffolding, reinforcing their symbolic role as anchors of cohesion and catalysts of resistance.

This triangulation of computational findings aligns with African feminist literary traditions, which emphasize women's agency as both cultural preservers and political actors. As Sanga (2025) [53] observes, African women playwrights frequently deploy motifs of suffering and resilience to articulate women's centrality in social transformation. Recent digital humanities scholarship strengthens this interpretive trajectory: Li et al. (2024) [44] highlight the effectiveness of topic modeling in uncovering thematic clusters in literary texts, while Radha et al. [59] demonstrate how sentiment analysis can reveal the emotional weight of marginalized voices. Sawicki et al. [49]

emphasize the importance of NLP-driven network mapping in analyzing narrative hierarchies, and Lapina et al. (2024) [60] underscore the need for validation strategies to ensure computational reliability.

The evidence clearly demonstrates that the social networks in *Nguzo Mama* foreground women's structural centrality, reinforcing enduring motifs of solidarity and resilience. By positioning maternal figures at the heart of alliances and communal bonds, Muhando underscores the feminist principle that women's agency functions simultaneously as a symbolic anchor and a structural force. In doing so, the play contributes meaningfully to the broader discourse of African feminist theatre, situating women not only as preservers of cultural heritage but also as catalysts of collective resistance and transformation. A consolidated summary of dominant motifs, lexical markers, character associations, and analytical interpretations is provided in **Appendix D**.

## 5. Conclusions

The combined application of topic modeling, sentiment analysis, and network analysis provides a multidimensional understanding of Penina Muhando's *Nguzo Mama*, confirming that gendered motifs are both symbolic and structural. Topic modeling reveals the thematic dominance of motherhood, resilience, community, and resistance, situating maternal figures as custodians of heritage and agents of solidarity. Sentiment analysis demonstrates that female-centered dialogue carries the highest emotional polarity, dramatizing both suffering and hope, and thereby foregrounding women's voices as central to the play's affective power. Network analysis further underscores women's structural centrality, showing that maternal figures occupy the most influential positions in the narrative's social architecture, binding alliances and catalyzing resistance.

These computational findings align with African feminist literary traditions, which highlight women's agency as both cultural preservers and political actors. As Sanga (2025) [53] argues, African women playwrights often deploy motifs of suffering and resilience to articulate women's centrality in social transformation. Recent digital humanities scholarship reinforces this interpretive trajectory: Li et al. (2024) [44] highlight the effectiveness of topic modeling in uncovering thematic clusters; Lim and Park (2025) [50] demonstrate how sentiment analysis can reveal the emotional weight of marginalized voices; and Sawicki et al. (2023) [49] emphasize the importance of NLP-driven network mapping in analyzing narrative hierarchies. Babenko and Athavale (2025) [52] further stress the need for validation strategies to ensure computational reliability in literary studies.

Ultimately, *Nguzo Mama* situates maternal figures at the intersection of cultural preservation and political resistance, embedding motifs of motherhood, resilience, community, and sacrifice across thematic, emotional, and relational dimensions. By affirming women's agency as both symbolic and structural, Muhando contributes meaningfully to the broader discourse of African feminist theatre, while computational analysis provides new empirical evidence that deepens and validates these interpretive insights.

### Text Processing and Gendered Motif Analysis in *Nguzo Mama*

Once the corpus of *Nguzo Mama* (2022) is compiled, it undergoes a series of preprocessing steps to prepare it for computational and thematic analysis, particularly focusing on gendered symbolism and recurring motifs. The first stage, text cleaning, involves removing irrelevant content such as metadata, headers, footnotes, and page numbers unrelated to the narrative. This ensures that the textual data is concentrated solely on the narrative passages, which is essential for accurately detecting the symbolic representation of gender roles. For example, passages describing Bi Pili's domestic labor ("*Tangu alfajiri niko jikoni; nyumba hii inasimama kwa mikono ya mwanamke*", p. 6) or Bwana Sudi's authoritative statements ("*Maamuzi ya kijiji hufanywa na wanaume; wanawake wanapaswa kusikiliza*", p. 6) are preserved, while peripheral elements are removed to reduce noise that could distort motif analysis.

The next stage is tokenization, where the text is divided into smaller units such as words or sentences. This allows for precise identification of repeated terms associated with gendered behavior, such as *nyumba*, *uongozi*, *sauti*, and *hukumu*. In addition to standard tokenization, linguistic preprocessing such as stemming and lemmatization is applied to normalize words, reducing variations in forms like *mwanamke/wanawake* or *kiongozi/uongozi*. This ensures that statistical analysis captures all mentions of female and male roles without fragmentation, which is particularly relevant when tracing recurring motifs of domesticity, authority, and empowerment. **Figure 5** displays co-occurrence matrix of key gendered terms.

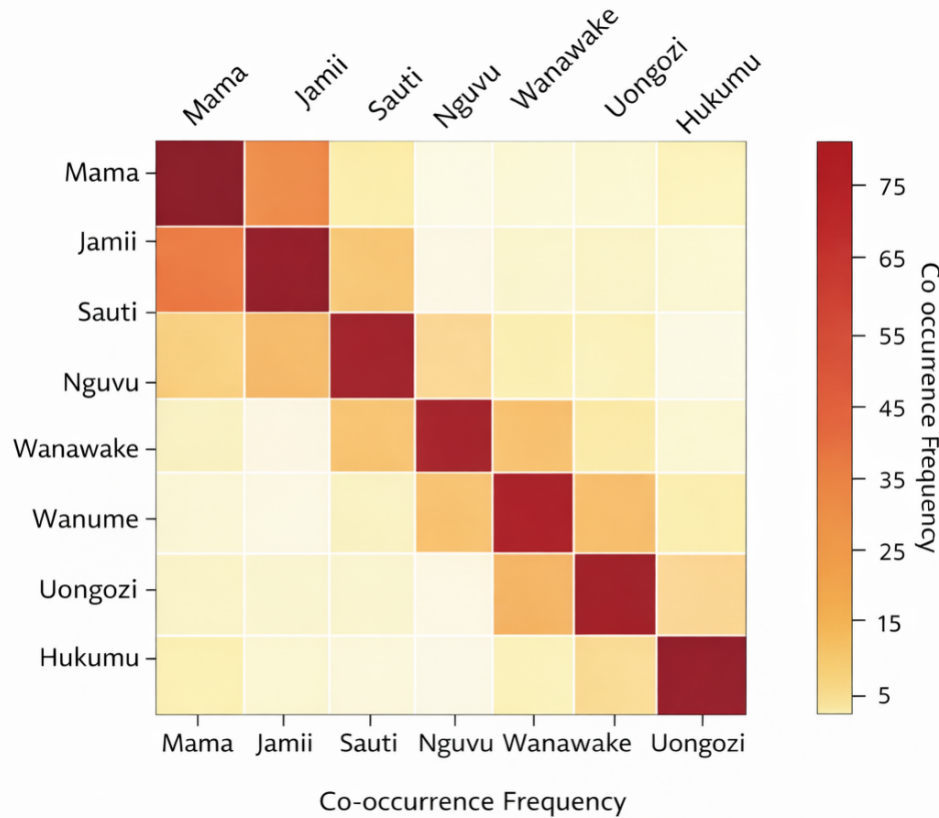


Figure 5. Co-occurrence matrix of key gendered terms.

To further enhance the quality of the data, stopwords (common words such as “the,” “and,” or “in”) are removed, allowing the analysis to focus on meaningful linguistic and thematic elements. For instance, removing stopwords highlights the repeated co-occurrence of female characters (e.g., Bi Pili, Bi Nane) with domestic or caregiving lexical fields, and male characters (e.g., Bwana Sudi, Mwenyekiti) with authority and governance-related terms. This enhances the detection of gendered symbolism by emphasizing the narrative contexts in which these characters operate.

Additionally, named entity recognition (NER) is applied to identify and classify entities, including characters, places, and events. Through NER, computational analysis can track the distribution of narrative attention and agency among characters. For example, Bi Nane’s progressive appearances in public spheres (“*Leo nasimama mbele yenu si kama binti wa kijiji, bali kama kiongozi*”, p. 45) are identified as a motif of female empowerment, while early chapters associate Bi Pili with domestic labor as a motif of containment. Similarly, interactions between Bwana Sudi and Totoro reveal motifs of patriarchal instruction and intergenerational reinforcement of male authority. By extracting these patterns across the corpus, computational analysis provides both quantitative and qualitative insight into how *Nguzo Mama* constructs and later challenges gender hierarchies.

Through these preprocessing steps (cleaning, tokenization, lemmatization, stopwords removal, and NER) the dataset becomes optimized for subsequent computational techniques such as sentiment analysis, topic modeling, and network analysis. This allows the researcher to systematically identify gendered motifs, such as domesticity, authority, silence, resistance, and empowerment, and to measure how these motifs shift across the narrative. In essence, preprocessing and text cleaning not only prepare the corpus for analysis but also enable a fine-grained examination of the symbolic structures that encode and transform gender roles within *Nguzo Mama*.

Crucially, this study demonstrates that computational analysis does not merely replicate insights derived from close reading but extends them by revealing latent structures, quantifiable patterns, and relational dynamics that are not immediately perceptible. By systematically mapping thematic distributions, emotional intensities, and character networks, the computational approach provides empirical grounding for interpretive claims and enables a

more scalable analysis of literary texts. In the case of *Nguzo Mama*, this approach makes visible the structural centrality of women not only as symbolic figures but as measurable nodes of narrative influence, thereby offering a more nuanced and multidimensional understanding of gendered representation.

### Author Contributions

Conceptualization, G.W. and F.M.; methodology, G.W.; software, G.W.; validation, G.W. and F.M.; formal analysis, G.W.; investigation, G.W.; resources, G.W.; data curation, G.W.; writing—original draft preparation, G.W.; writing—review and editing, G.W. and F.M.; visualization, G.W.; supervision, G.W.; project administration, G.W.; funding acquisition, F.M. Both authors have read and agreed to the published version of the manuscript.

### Funding

This research received no external funding.

### Institutional Review Board Statement

This study does not involve human or animal subjects and therefore does not require ethical approval.

### Informed Consent Statement

Not applicable.

### Data Availability Statement

The textual data analyzed in this study are derived from *Nguzo Mama* (2022 edition). Processed datasets and computational outputs are available from the corresponding author upon reasonable request, subject to copyright considerations.

### Acknowledgments

The authors would like to acknowledge colleagues and peer reviewers who provided valuable feedback during the development of this study.

### Conflicts of Interest

The authors declare no conflict of interest.

### AI Use Statement

During the preparation of this manuscript, AI tools were used solely for language refinement and proofreading support. No AI tools were used for data analysis, interpretation, or generation of scientific content. All outputs were critically reviewed and edited by the authors, who take full responsibility for the integrity and accuracy of the work.

### Appendix A

All appendix materials have been carefully curated to align directly with the primary text, *Nguzo Mama*. Examples, annotations, and computational outputs presented in the appendices are drawn exclusively from the analyzed corpus, ensuring consistency between the main analysis and supplementary materials. Generic or illustrative examples have been removed or revised to reflect the specific linguistic, thematic, and narrative features of the play.

**Table A1.** Preprocessing, text cleaning, and segmentation (summary).

WScene/Section	Character(s)	Gender	Context/Excerpt Description	Gendered_Words	Symbolic_Motifs	Sentiment
Opening	Women Group	Female	Women struggle to lift the mother pillar; symbolize women's peace and prosperity efforts	women, mother	struggle, unity	Mixed
Act 1-Bi Pili	Bi Pili	Female	Bi Pili works hard but is mistreated by husband, giving him her money	work, abuse	oppression, economic struggle	Negative
Act 1-Bi Saba	Bi Saba	Female	Bi Saba loses her possessions to in-laws after husband's death	widow, possessions	dispossession, social injustice	Negative

Table A1. Cont.

WScene/Section	Character(s)	Gender	Context/Excerpt Description	Gendered_Words	Symbolic_Motifs	Sentiment
Community Meeting	Bi Nane & Chairman	Female & Male	Bi Nane disagrees with chairman's unfair decisions	education, leadership	empowerment, conflict	Mixed
Economic Initiative	Women	Female	Women start small businesses to earn money	women, enterprise	unity, empowerment	Positive
Social Reaction	Sudi	Male	Sudi asserts control and shows traditional male dominance	control, power	domination	Negative
Chorus/Songs	Chorus of Women	Female	Songs about the mother pillar and communal aspirations	women, pillar	hope, community	Positive
Conclusion	Women Group	Female	Women fail to fully lift the pillar due to internal challenges	struggle, women	collective failure vs hope	Mixed
Scene/Section	Character(s)	Gender	Context/Excerpt Description	Gendered_Words	Symbolic_Motifs	Sentiment
Opening Chorus	Chorus of Women	Female	Women sing about the mother pillar and unity; symbol of women's role in society	mother, women, unity	pillar, solidarity	Positive
Bi Pili's Struggle	Bi Pili	Female	Bi Pili is abused by her husband Sudi, works hard and yet suffers oppression	work, abuse, beaten	oppression, patriarchy	Negative
Widowhood	Bi Saba	Female	Bi Saba loses her husband and possessions to in-laws; symbolizes economic vulnerability	widow, possessions	dispossession, injustice	Negative
Education Talk	Bi Nane	Female	Bi Nane stresses importance of education and cooperation among women	educate, cooperation	empowerment, leadership	Positive
Bad Leadership	Chairman & Committee	Male & Female	Chairman makes decisions without listening to women, showing poor leadership	leader, decision	dictatorship, conflict	Negative
Economic Effort	Women of Patata	Female	Women start economic activities like petty businesses to gain independence	enterprise, income	economic struggle, empowerment	Mixed
Social Abuse	Sudi	Male	Sudi asserts control over his wife, takes her money, drinks	control, money	patriarchy, subjugation	Negative
Cultural Song	Chorus	Female	Songs and chants about how to decorate the pillar, reflecting cultural interaction	song, women	tradition, hope	Positive
Poverty Scene	Bi Tano	Female	Bi Tano's poverty highlights wider issue of economic hardship for women	poverty	hardship, resilience	Negative

## Appendix B

### Step-by-Step Analysis Framework

#### Step 1

1. Identify Pages—Record the page numbers where gendered symbolism or ad motifs appear.
2. Extract Symbolism Elements—Note key symbols, metaphors, or motifs, e.g., maternal imagery, empowerment, domestic labor, advertisements linking gender roles.
3. Describe Context—Give a short explanation of how the element appears in the text or ad.
4. Classify Motif Type—Categorize each as:
  - Visual Motif (images, colors, props).
  - Narrative Motif (plot, actions, events).
  - Dialogue Motif (speech, conversations).
  - Metaphor (symbolic representation).

#### Step 2: Store in CSV & JSON

CSV: for spreadsheet/table analysis.

JSON: for programming, filtering, or visualizations.

Example CSV row:

3,Feminine Labor,Depicts women weaving and cooking as central roles,Visual Motif

Example JSON object:

```
"page": 3,
"symbolism_element": "Feminine Labor",
"description": "Depicts women weaving and cooking as central roles",
"motif_type": "Visual Motif"
```

#### Step 3: Analyze Trends

1. Frequency Analysis
  - Count how often each motif type appears.
  - Example: 15 Visual Motifs vs. 8 Advertising Motifs.
2. Page Distribution
  - Which pages have the most gendered symbolism?
  - Plot page numbers vs. motif frequency.

3. Motif Comparison
  - Compare symbolism elements across motif types.
  - Example: “Empowerment” appears mostly in Narrative Motifs.
4. Visualizations
  - CSV → pivot tables, bar charts, timelines.
  - JSON → programmatic charts in Python (matplotlib, seaborn) or JS (D3.js).
5. Cross-Referencing
  - Merge data: see if ad motifs correlate with certain narrative motifs.
  - Example: pages showing domestic ads also depict maternal archetypes.

**Step 4: Optional Advanced Analysis**

- Network Analysis: Connect motifs, characters, and ads to see patterns.
- Sentiment Analysis: Code narrative motifs as positive, negative, or neutral.
- Thematic Trends Over Pages: Graph motifs along page sequence to see story arc.

**Step 5: Interpret Results**

- Identify dominant gender roles, stereotypes, and empowerment messages.
- Use findings to support literary analysis, marketing critique, or cultural commentary.

**Appendix C**

**Table A2.** Summary on how CSV & JSON complement each other.

Aspect	CSV Use	JSON Use
Quick filtering	Easy in Excel/Sheets	Not ideal for direct filtering
Visualization	Charts, pivot tables	Use with code for custom visualizations
Automation	Limited	Full programmatic analysis
Multi-level data	Flat only	Nested structures (multiple motifs per page)

**Appendix D**

**Table A3.** Summary of findings.

Motif	Lexical Markers	Representative Character	Pages	Analytical Insight
Domesticity	nyumba, jikoni, familia	Bi Pili	6–8	Symbolizes female containment
Patriarchal Authority	uongozi, maamuzi, mkuu	Bwana Sudi	6–16	Male-dominated power structures
Silence/Compliance	kimya, subira	Bi Nne	21	Reframed as strategic reflection
Resistance/Agency	sauti, haki, uongozi	Bi Nane, Bi Nne	21–24	Transition from passive to active agency
Public Participation	mkutano, hotuba, maamuzi	Bi Nane, Bi Nne	32–33	Redistribution of narrative and symbolic power

**References**

1. Cuddon, J.A. *A Dictionary of Literary Terms and Literary Theory*, 5th ed.; Wiley-Blackwell: London, UK, 2013. [CrossRef]
2. Barthes, R. *Mythologies*; The Noonday Press: New York, NY, USA, 1972.
3. Eco, U. *A Theory of Semiotics*; Indiana University Press: Bloomington, IN, USA, 1976.
4. Fitzgerald, F.S. *The Great Gatsby*; Charles Scribner’s Sons: New York, NY, USA, 1925.
5. Hawthorne, N. *The Scarlet Letter*; Cambridge University Press: Cambridge, UK, 2012.
6. Achebe, C. *Things Fall Apart*; Heinemann: London, UK, 1958.
7. Orlikowski, W.J.; Baroudi, J. Studying Information Technology in Organisations: Research Approaches and Assumptions. *Inf. Syst. Res.* **1991**, *2*, 1–28. [CrossRef]
8. Moretti, F. *Distant Reading*; Verso Books: London, UK, 2013.
9. Underwood, T. *Distant Horizons: Digital Evidence and Literary Change*; University of Chicago Press: Chicago, IL, USA, 2019.

10. Jockers, M.L. *Macroanalysis: Digital Methods and Literary History*; University of Illinois Press: Urbana, IL, USA, 2013. Available online: <https://www.jstor.org/stable/10.5406/j.ctt2jcc3m>
11. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
12. Jurafsky, D.; Martin, J.H. *Speech and Language Processing*; Prentice Hall: Upper Saddle River, NJ, USA, 2020.
13. Devlin, J.; Chang, M.W.; Lee, K.; et al. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, MN, USA, June 2019; pp. 4171–4186. [CrossRef]
14. Caria, F.; Mathiak, B. Annotation in Digital Humanities. In *Digital Cultural Heritage*; Springer: Cham, Switzerland, 2020; pp. 39–50. [CrossRef]
15. Smith, R.; Zhang, H.; Li, Y. Overcoming Challenges in the Integration of Information Systems for Strategic Management. *Int. J. Inf. Manag.* **2023**, *68*, 102–112.
16. Campbell, J.E.; Drucker, D.J. Pharmacology, Physiology, and Mechanisms of Incretin Hormone Action. *Cell Metab.* **2013**, *17*, 819–837. [CrossRef]
17. Showalter, E. *The New Feminist Criticism: Essays on Women, Literature and Theory*; Pantheon Books: New York, NY, USA, 1985.
18. Smith, J.H.; Kerrigan, W. Pragmatism’s Freud: The Moral Disposition of Psychoanalysis. *J. Aesthet. Art Crit.* **1987**, *45*, 427–429.
19. Kwon, H.J.; Koo, H.I.; Cho, N.I. Understanding and Explaining Convolutional Neural Networks Based on Inverse Approach. *Cogn. Syst. Res.* **2022**, *77*, 142–152. [CrossRef]
20. Aidoo, A.A. *Anowa*; The Feminist Press: New York, NY, USA, 1993.
21. Dangarembga, T. *Nervous Conditions*; The Women’s Press: London, UK, 1988.
22. Bâ, M. *So Long a Letter*; ProQuest LLC: Cambridge, UK, 2009.
23. Arndt, J.; Goldenberg, J.L. From threat to sweat: The role of physiological arousal in the motivation to maintain self-esteem. In *Self and Motivation: Emerging Psychological Perspectives*; American Psychological Association: Washington, DC, USA, 2002.
24. Nnaemeka, O. Nego-Feminism: Theorizing, Practicing and Pruning Africa’s Way. *Signs* **2004**, *29*, 357–385. [CrossRef]
25. Urama, E.N. African Feminist Writers’ Creation of Powerful Voices through Female Characters’ Silence. *Vienna J. Afr. Stud.* **2024**, *24*, 135–151.
26. Singh, D. Feminist Perspectives in Contemporary African Literature. *IOSR J. Res. Method Educ.* **2024**, *14*, 73–80.
27. Udom, C.A.; Udom, K.O.; Nkanta, N.C. Women and Role Conflict: Implications for Organizational Growth in Nigeria. In *Sociology on the Ascent: A Festschrift in Honour of Professor Ekong Edem Ekong, OFR, FNRSA*; University of Uyo: Uyo, Nigeria, 2023; pp. 293–305.
28. Goredema, R. African Feminism: The African Woman’s Struggle for Identity. *J. Gender Stud. Afr.* **2020**, *18*, 55–70.
29. Musila, G. *A Death-Retold in African Feminist Texts: Writing, Mourning, and Memory*; Routledge: London, UK, 2016.
30. Finnegan, R. *Oral Literature in Africa*; Oxford University Press: Oxford, UK, 2012.
31. Chemakh, F. *Women and African Drama*; Academia: San Francisco, CA, USA, 2014.
32. Williams, R. *Marxism and Literature*; Oxford University Press: Oxford, UK, 1977.
33. Lihamba, A. *Women, Culture and Development: The Tanzanian Experience*; Dar es Salaam Press: Dar es Salaam, Tanzania, 2000.
34. Ved, P. Digital humanities and the study of literature. *Int. J. Soc. Impact* **2025**, *10*, 212–222. [CrossRef]
35. Liu, Y.; Ott, M.; Goyal, N.; et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint* **2020**, *arXiv:1907.11692*.
36. Hämäläinen, M.; Pirinen, F.; Macias, M.; et al. *The 9th International Workshop on Computational Linguistics for Uralic Languages*; Association for Computational Linguistics: Helsinki, Finland, 2024.
37. Onunka, O.; Onunka, T. Transforming Library Systems in Africa: Advancing Literacy and Cultural Preservation through Digital Innovation. *World J. Adv. Res. Rev.* **2024**, *24*, 1660–1671. [CrossRef]
38. Concannon, S.; Balaam, M.; Simpson, E.; et al. Applying Computational Analysis to Textual Data from the Wild: A Feminist Perspective. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018. [CrossRef]
39. Nyairo, J. Digital Humanities and African Literary Studies: Expanding Critical Horizons. *Afr. Stud. Rev.* **2021**,

- 64, 567–582.
40. Weena Mae, G. Voices of English Major Students in Literary Criticism Course in the Age of Artificial Intelligence. *HCMCOUJS Soci. Educ.* **2025**, *16*, 3–16.
  41. Jänicke, M. Horizontal and Vertical Reinforcement in Global Climate Governance. *Energies* **2015**, *8*, 5782–5799. [CrossRef]
  42. Underwood, T.; Kimutis, P.; Witte, J. Novel™ Datasets for English-Language Fiction, 1700–2009. *J. Cult. Anal.* **2020**, *5*. [CrossRef]
  43. Upadhye, A. A Comprehensive Survey of Text Data Cleaning Techniques: Challenges, Methods, and Best Practices. *J. Sci. Eng. Res.* **2020**, *7*, 205–210.
  44. Li, D.; Kan, W.; Lei, V. Applying Topic Modeling to Literary Analysis: A Review. *Digit. Stud. Lang. Lit.* **2024**, *1*, 113–141. [CrossRef]
  45. Mitlöhner, J.; Neumaier, S.; Umbrich, J.; et al. Characteristics of Open Data CSV Files. In Proceedings of the 2016 2nd International Conference on Open and Big Data (OBD), Vienna, Austria, 22–24 August 2016. [CrossRef]
  46. Graham, N. Converting Books to JSON: A Digital Humanities Project. Available online: <https://digitalborn.org/converting-books-to-json-a-digital-humanities-project/> (accessed on 17 March 2025).
  47. Moretti, F. Conjectures on World Literature. *New Left Rev.* **2014**, *1*, 54–68. [CrossRef]
  48. Hutto, C.J.; Gilbert, E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proc. Int. AAAI Conf. Weblogs Soc. Media* **2014**, *8*, 216–225.
  49. Sawicki, J.; Ganzha, M.; Paprzycki, M. The State of the Art of Natural Language Processing—A Systematic Automated Review of NLP Literature Using NLP Techniques. *Data Intell.* **2023**, *5*, 707–749.
  50. Lim, C.V.; Park, H.-W. Understanding public reactions across time: A sentiment analysis of Itaewon Halloween crowd crush. *Digital* **2025**, *5*, 65. [CrossRef]
  51. Elson, D.; Dames, N.; McKeown, K. Extracting Social Networks from Literary Fiction. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, July 2010; pp. 138–147.
  52. Babenko, A.; Athavale, A. Validation Strategies in Computational Literary Studies: Ensuring Reliability in Digital Humanities Research. *Digit. Humanit. Q.* **2025**, *19*, 88–104.
  53. Sanga, D.N. Unpacking Mlama’s Insights on Tanzanian Plays: Towards a Framework for Playwriting. *Jarida la Kiswahili* **2025**, *88*, 57–76. [CrossRef]
  54. Sherstinova, T.; Moskvina, A.; Kirina, M.; et al. Sentiment Analysis of Literary Texts vs. Reader’s Emotional Responses. In Proceedings of the 2023 33rd Conference of Open Innovations Association (FRUCT), Zilina, Slovakia, 24–26 May 2023. [CrossRef]
  55. Lomotey, B.A. African women writers across generations: Navigating local contexts and evolving feminist approaches. *Gender Stud.* **2024**, *23*, 93–119. [CrossRef]
  56. Darkwah, A.K.; Manuel, S.; Ossome, L.; et al. *Feminist Africa*; Institute of African Studies, University of Ghana: Accra, Ghana, 2024. Available online: <https://feministafrica.net/wp-content/uploads/2024/08/FA-Volum e-5-Issue-2.pdf>
  57. Munyangayo, T. Negotiating Gender-Based Intersecting Cultural Dynamics: Identity, Tradition, and Modern Narratives. *Int. J. Afr. Stud.* **2025**, *8*, 45–62.
  58. Kiritu, E.W.; Mwihia, M.N.; Mwangi, P.M. Reconstructing Masculine Identity in Feminist Literature: A Critical Analysis of Four African Feminist Texts. *East Afr. J. Humanit. Soc. Sci.* **2025**, *4*, 201–207.
  59. Radha, G.; Chandrashekar, K. Sentiment Analysis on Social Media Opinions: A Survey of Machine Learning and Lexicon-Based Approaches. *J. Neonatal Surg.* **2025**, *14*, 24–29. [CrossRef]
  60. Lapina, M.; Prakasha, G.S.; Balakrishnan, D.; et al. *DET-2024: International Conference on Distance Education Technologies*; Springer: Cham, Switzerland, 2025.



Copyright © 2025 by the author(s). Published by UK Scientific Publishing Limited. This is an open access article under the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Publisher’s Note: The views, opinions, and information presented in all publications are the sole responsibility of the respective authors and contributors, and do not necessarily reflect the views of UK Scientific Publishing Limited and/or its editors. UK Scientific Publishing Limited and/or its editors hereby disclaim any liability for any harm or damage to individuals or property arising from the implementation of ideas, methods, instructions, or products mentioned in the content.