

Article

Mapping Cultural Sentiments in Indonesian Digital Literature: An Annotated and Validated Multicultural Dataset

Suarni Syam Saguni  , Nensilianti *  and Ridwan 

Department of Indonesian Language and Letter, Universitas Negeri Makassar, Makassar 90224, Indonesia

* Correspondence: nensilianti@unm.ac.id

Received: 1 December 2025; **Revised:** 26 January 2026; **Accepted:** 27 January 2026; **Published:** 9 February 2026

Abstract: This study develops an annotated and validated multicultural sentiment dataset derived from Indonesian digital literature. The study integrates Cultural Sentiment Analysis (CSA) and Critical Discourse Analysis (CDA) to address a significant gap in existing research. No prior corpus has systematically combined cultural affective dimensions with linguistic and ethnographic validation, making this dataset crucial for mapping cultural value representations in a multicultural context. The corpus includes over 100 digital literary texts—short stories, online novels, and poems—sourced from platforms such as Wattpad, KBM App, and scholarly blogs, selected through purposive sampling to ensure diverse ethnic and thematic coverage. Annotation was carried out by trained annotators using a culturally grounded emotion lexicon, identifying sentiment polarity (positive, negative, and neutral), cultural values (social harmony, cooperation, spirituality, resistance, and adaptation), and linguistic indicators. Validation involved linguistic review for semantic accuracy and ethnographic verification through Focus Group Discussions with cultural experts from various ethnic groups. The resulting multi-layered dataset provides authentic, contextually grounded, and bias-mitigated representations of cultural sentiment in Indonesian digital literature. Beyond enriching digital humanities scholarship, it offers a reusable open resource for future research, automated sentiment analysis development, and data-driven policy formulation, all aimed at enhancing digital cultural literacy and intercultural understanding in Indonesia.

Keywords: Dataset; Cultural Sentiment; Digital Literature; Linguistic Validation; Ethnographic Verification

1. Introduction

The rapid development of digital technologies over the past two decades has profoundly transformed how cultural texts are produced, disseminated, and consumed. This transformation has not only reshaped literary practices but also repositioned literature as a critical arena for negotiating, contesting, and reimagining cultural values in the digital age [1, 2]. Globally, digital literary platforms function as participatory spaces that enable the production, circulation, and reinterpretation of narratives across linguistic, ethnic, and ideological boundaries [3, 4]. Within this broader shift, cultural sentiment—understood as the affective dimension embedded within narratives that reflects collective beliefs, identities, and ideologies—has become a crucial analytic entry point for understanding how societies construct and negotiate meaning in increasingly multicultural and interconnected digital ecosystems [5].

In Indonesia, this phenomenon is particularly salient given the country's complex sociocultural fabric and rapidly expanding digital landscape. Digital literature has evolved from a medium of aesthetic expression into a significant socio-discursive space where narratives of identity, power, resistance, and adaptation intersect [6–8]. Platforms such as Wattpad, KBM App, and Storial function as cultural microcosms that facilitate interaction across

ethnic, linguistic, and class boundaries, offering a dynamic arena for the articulation of cultural experiences and social imaginaries. According to the Indonesian Internet Service Providers Association (APJII), more than 80.66% of active internet users now access text-based content—including short stories, poetry, and online novels—indicating a profound shift in the role of literature from a purely aesthetic practice to a sociocultural instrument that shapes identity, represents cultural values, and mediates conflicts in a multicultural society [9].

Despite this exponential growth, scholarly engagement with digital literature in Indonesia remains limited in scope. Particularly in relation to its affective and cultural dimensions, existing studies relying on qualitative narrative or thematic approaches [10–12] often lack the capacity to capture large-scale emotional patterns or systematically map cultural sentiment [13,14]. This gap is significant in a socio-political context increasingly marked by identity polarization, misinformation, and cultural contestation [15]. In such an environment, understanding the affective structures underlying digital narratives is not merely an academic pursuit but a strategic necessity for fostering intercultural dialogue and mitigating societal fragmentation.

Cultural Sentiment Analysis (CSA) offers a promising methodological solution to these challenges. As a computational technique grounded in Natural Language Processing (NLP), CSA enables the large-scale detection, classification, and interpretation of sentiment expressions tied to cultural values [16,17]. When integrated with Critical Discourse Analysis (CDA), CSA extends beyond sentiment polarity to uncover deeper ideological structures, power relations, and identity formations that underlie textual production [2,18]. However, the application of this integrative approach within the context of Indonesian digital literature remains nascent. A bibliometric review of 27 Scopus-indexed journals revealed that only three studies have examined Indonesian digital literature using data-driven methods, and none explicitly integrate CSA and CDA [19]. This underscores a substantial research gap and an opportunity to develop innovative, interdisciplinary methodologies that bridge computational analysis and cultural theory.

Recent developments in culturally aware NLP emphasize the importance of bias-sensitive dataset construction, affective computing beyond polarity, and ethical accountability in annotated resources [20–23]. Studies in digital humanities also highlight the integration of computational modeling with critical cultural interpretation to address the textual ecologies of the Global South [24–26]. These advances inform the present framework by positioning cultural sentiment not only as a technical variable but as a socially situated and ideologically mediated construct.

The shift from traditional qualitative literary studies to data-driven methodologies also reflects broader epistemological transformations within the digital humanities [24,27,28]. Advances in machine learning, probabilistic topic modeling, and emotion lexicon development have enabled scholars to analyze cultural phenomena at scale, revealing macro-level patterns in value distribution, emotional dynamics, and ideological shifts over time. Such approaches not only enhance analytical precision but also provide empirical foundations for evidence-based policy design in cultural literacy and education [29,30]. Moreover, they align with contemporary theoretical discourses—such as postcolonialism and multiculturalism—that conceptualize culture as a site of ongoing negotiation, hybridity, and contestation [31,32].

Responding to this scholarly and societal need, the present study develops an annotated and validated multicultural sentiment dataset derived from Indonesian digital literature. The dataset, comprising over 100 texts including short stories, online novels, and poetry, was curated through purposive sampling to capture the diversity of ethnic perspectives and cultural themes represented in the digital sphere. Annotation was conducted using a culturally grounded emotion lexicon and underwent dual validation: linguistic verification ensured the accuracy of lexical and syntactic tagging, while ethnographic validation—conducted through focus group discussions with cultural experts from diverse ethnic backgrounds—ensured contextual and cultural authenticity. The resulting multi-level annotation schema captures sentiment polarity (positive, negative, neutral), types of cultural values (e.g., social harmony, cooperation, spirituality, cultural resistance, value adaptation), and key linguistic indicators [17,33].

This research contributes to the field of digital humanities and cultural analytics in two fundamental ways. First, it provides a robust methodological framework that integrates CSA and CDA, enabling a deeper understanding of the affective, ideological, and cultural layers of digital literary texts. Second, it offers a publicly available, annotated dataset that serves as a foundational resource for future research in computational cultural studies, automated sentiment detection systems, and evidence-based policy design aimed at enhancing digital cultural literacy. By bridging the methodological gap between computational analysis and critical theory, this study not only advances the scholarly discourse on digital literature but also offers practical tools for promoting intercultural un-

derstanding and strengthening cultural resilience in Indonesia's increasingly complex digital society.

2. Methods

2.1. Research Design

This study employs an interdisciplinary hybrid research design that integrates computational linguistics, digital humanities, and critical cultural analysis to systematically map the cultural sentiments embedded in Indonesian digital literature. The methodological framework follows the principles of Design-Based Research (DBR), a paradigm characterized by iterative, data-driven, and context-sensitive cycles aimed at solving complex real-world cultural and linguistic challenges [30, 34]. This approach allows for the development of a validated annotated dataset of cultural sentiment. In parallel, it supports a computational-interpretive model that captures both linguistic structures and sociocultural meanings.

The overall research workflow consists of five key stages. The first stage, contextual and theoretical framing, involves mapping sociocultural dynamics in digital literature and establishing theoretical foundations such as Cultural Sentiment Analysis (CSA), Critical Discourse Analysis (CDA), postcolonial theory, and multiculturalism. The second stage, corpus construction and preprocessing, focuses on selecting, cleaning, and preparing digital literary texts to ensure their compatibility with computational analysis. The third stage, annotation and validation, involves developing a multi-layered annotated dataset that captures sentiment polarity, cultural values, and linguistic indicators, followed by rigorous validation to ensure accuracy and cultural relevance. The fourth stage, computational modeling and CDA integration, applies NLP techniques, machine learning models, and discourse analysis to detect and interpret cultural sentiments at scale. Finally, the fifth stage, synthesis and interpretation, integrates the findings into a cultural analytical framework and translates them into practical insights for digital literacy development and cultural policy formulation.

2.2. Data Collection and Corpus Construction

The primary dataset consists of 100 Indonesian digital literary texts (short stories, online novels, and poems) published between 2018 and 2023, collected from leading platforms such as Wattpad, KBM App, and Storial, as well as curated literary blogs. A purposive sampling strategy was employed to ensure representativeness in terms of ethnic diversity, thematic scope, and cultural contexts [6, 12]. The texts cover 10 major ethnic groups, including Javanese, Bugis, Batak, Toraja, Minangkabau, and others, and are categorized across four key thematic domains: identity and belonging, cultural resistance, value adaptation, and spirituality.

The corpus construction followed a rigorous three-step process. First, selection and compilation involved filtering texts based on thematic relevance—such as identity, adaptation, and resistance—and ensuring cultural diversity. Second, metadata documentation enriched each text with contextual information, including author, platform, publication date, cultural theme, and genre, to support accurate annotation. Finally, preprocessing prepared the data for analysis through tokenization, lowercasing, punctuation removal, stopword filtering, and lemmatization using Python 3.10 with NLTK and spaCy, ensuring compatibility with NLP tools [16, 28].

Although the corpus size is limited to 100 texts, this number was deliberately chosen to balance cultural depth and annotation reliability in a pilot-scale, culturally intensive study. Unlike large-scale sentiment datasets that prioritize volume, this research emphasizes semantic richness and cultural representativeness. The purposive sampling strategy ensured coverage of ten major Indonesian ethnic groups and four dominant cultural themes, thereby maximizing internal diversity within a manageable corpus. Nevertheless, the authors acknowledge that platform-related bias may exist, as dominant platforms such as Wattpad and KBM App tend to favor urban, youth-oriented, and popular narrative styles, while literary blogs may reflect more academically mediated discourse. Consequently, the corpus should be interpreted as indicative rather than exhaustive of Indonesian digital literary production, and future extensions will incorporate community archives and regional platforms to enhance representational balance.

2.3. Annotation Framework

The annotation process constitutes the methodological backbone of this study and was designed to capture multi-layered cultural sentiments, summarized in **Table 1**. Manual annotation was conducted by a team of five trained annotators, each with a background in linguistics, literature, or cultural studies. Annotators underwent

two calibration rounds and a pilot annotation phase to refine annotation guidelines and ensure consistency.

Table 1. Annotation dimensions.

Layer	Description	Example
Sentiment Polarity	Label each sentence or clause as positive, negative, or neutral, reflecting affective orientation.	"Inter-tribal cooperation is our strength." → Positive
Cultural Value Category	Tags cultural themes such as social harmony, cooperation (<i>gotong royong</i>), spirituality, cultural resistance, and value adaptation.	"The fight against cultural discrimination" → Cultural Resistance
Linguistic & Semiotic Indicators	Identifies metaphors, idioms, lexical items, or narrative structures encoding cultural sentiment.	Words like customs, ancestors, and struggle

The annotation lexicon was developed from a 1850-term culturally grounded lexicon, derived from corpus keyword extraction, expert consultation, and ethnographic field interviews. Annotation reliability was assessed using Cohen's Kappa, achieving an agreement score of $\kappa = 0.84$, exceeding the 0.80 reliability benchmark recommended for computational annotation [35]. Given the culturally nuanced nature of categories such as spirituality and cultural resistance, several mechanisms were implemented to mitigate annotation subjectivity. These included iterative guideline refinement, double-blind annotation rounds, adjudication meetings for disagreement resolution, and ethnographic validation through FGDs to anchor labels in culturally situated interpretations. While the current framework was optimized for expert-driven annotation, its design supports scalability through semi-automatic pre-labeling using lexicon-based and transformer-assisted models, followed by targeted human verification. This structure enables controlled expansion to larger corpora and facilitates cross-cultural transfer by adapting the core sentiment dimensions to local cultural ontologies.

2.4. Validation Procedures

A dual validation framework was applied to ensure linguistic precision and cultural authenticity of the annotations. Linguistic validation involved computational linguists reviewing 20% of the annotated corpus to verify lexical, syntactic, and semantic accuracy, with misannotations corrected through consensus workshops. Ethnographic validation engaged cultural experts from seven ethnic groups in Focus Group Discussions (FGDs) to contextualize annotations, ensuring that cultural meanings were accurately represented, contextually grounded, and free from ethnocentric bias.

2.5. Computational Modeling and Cultural Sentiment Analysis

After annotation and validation, the dataset was processed through a multi-stage NLP and machine learning pipeline implemented in Python 3.10 using libraries such as scikit-learn, TensorFlow 2.12, and Hugging Face Transformers. The process included feature extraction to identify key linguistic and cultural features (n-grams, TF-IDF vectors, cultural lexicon frequencies, co-occurrence matrices), followed by model training using traditional classifiers (SVM, Logistic Regression) and deep learning models (BiLSTM for contextual analysis and BERT/IndoBERT for fine-tuned sentiment detection). Evaluation with a stratified 80/20 train-test split and 5-fold cross-validation showed that BiLSTM achieved an F1-score of 0.87 and accuracy of 0.89, exceeding the ≥ 0.85 benchmark. Finally, an interpretive layer using Critical Discourse Analysis (CDA) [18] contextualized the results, uncovering underlying ideologies, power relations, and socio-political narratives [24].

To further enhance model interpretability, an error analysis was conducted on misclassified instances. The majority of false negatives occurred in metaphorical and implicit affect expressions, particularly those embedding cultural symbolism (e.g., ancestral metaphors, spiritual idioms), which are less explicitly represented in pre-trained embeddings. False positives were frequently associated with polysemous cultural lexemes whose sentiment orientation depends strongly on narrative context. These patterns indicate that although BiLSTM and BERT-based models achieve high overall performance, they remain sensitive to culturally implicit language use. Moreover, pre-trained transformer models, primarily optimized on general-domain Indonesian corpora, may underrepresent minority dialectal forms and localized cultural expressions. This bias underscores the need for culturally enriched fine-tuning data and motivates the present dataset as a corrective resource for low-resource and culturally specific NLP.

2.6. Triangulation and Synthesis

To enhance analytical robustness, a triangulation approach integrates computational, linguistic, and ethnographic findings, enabling the identification of dominant and minority cultural sentiments, temporal and thematic shifts across genres and publication periods, and relational dynamics between language, identity, and emotion. The triangulated results informed the development of digital cultural literacy modules and evidence-based policy recommendations, offering actionable insights for education, technology design, and intercultural communication.

2.7. Expected Outcomes

The proposed methodological framework is designed to achieve four main outcomes: a publicly accessible annotated and validated dataset of cultural sentiments in Indonesian digital literature; a high-accuracy machine learning model for sentiment classification and narrative interpretation; a computational-critical analytical framework integrating CSA and CDA for advanced cultural analysis; and policy and educational tools that foster intercultural dialogue, promote cultural literacy, and strengthen social cohesion in digital environments.

3. Results

3.1. Development of the Annotated Multicultural Cultural Sentiment Dataset

3.1.1. Multicultural Cultural Sentiment Dataset: Empirical Insights and Analytical Implications

The construction of a multicultural cultural sentiment dataset offers a nuanced cartography of cultural expression as reflected in contemporary Indonesian digital literature. Comprising over 100 curated literary texts sourced from diverse digital platforms—including Wattpad, KBM, and community-based literary blogs—the dataset encapsulates a wide range of ethnic representations, sociocultural themes, and historical narratives. Exploratory data analysis reveals not only a high degree of narrative variation but also a rich spectrum of emotional intensity. By employing both quantitative and qualitative analytical lenses, this study captures the dynamic movement of cultural sentiment—ranging from affirmation and solidarity to critique and ambivalence, which are closely tied to the cultural values depicted and the ideological frameworks in which the texts are situated.

The dataset was annotated across three distinct analytical dimensions that collectively enhance the granularity of sentiment interpretation. First, the sentiment polarity layer indicates that 45% of the narratives are characterized by negative sentiment, 32% by positive sentiment, and 23% by ambivalent or mixed sentiments. Negative sentiment is predominantly linked to narratives of cultural conflict, marginalization, and resistance to hegemonic norms. In contrast, positive sentiment typically appears in texts emphasizing cooperation (*gotong royong*), social cohesion, and spiritual reverence. Ambivalent sentiment emerges in narratives that portray cultural dilemmas or processes of negotiation, particularly those that highlight tensions between tradition and modernity.

Second, the cultural values annotation identifies five dominant thematic categories—spirituality, *gotong royong*, cultural resistance, value adaptation, and social harmony—each of which is visualized in frequency distribution diagrams. Third, the linguistic dimension demonstrates that culturally charged phrases—such as “saving a life,” “treated unfairly,” or “sacred kris”—simultaneously function as emotional cues and markers of cultural meaning, thus enriching the affective and symbolic depth of the texts.

These empirical findings are further substantiated through cross-tabulation of sentiment polarity and cultural value categories, illustrated via heatmap visualizations. The analysis reveals that values such as social harmony and *gotong royong* are strongly associated with positive sentiment. In contrast, values related to adaptation and identity formation frequently correspond with ambivalent sentiment, reflecting the complex emotional negotiations involved in cultural transformation. Meanwhile, cultural resistance and social injustice are predominantly linked to negative sentiment, underscoring the function of digital literary narratives as spaces of critique, dissent, and ideological resistance. These patterns collectively suggest that cultural sentiments within the corpus are not randomly distributed, but instead follow discernible thematic, ideological, and emotional trajectories. Accordingly, Indonesian digital literature emerges as an active discursive site for the articulation of cultural tensions, emotional reflection, and social negotiation.

3.1.2. Distribution of Sentiment Polarity

The distribution of sentiment polarity serves as a foundational analytical layer in understanding how Indonesian digital literature encodes emotional expressions of cultural identity. Sentiment polarity refers to the affective orientation of a text—whether positive (affirmative and constructive), negative (critical and oppositional), or ambivalent (ambiguous and reflective). Within the context of multicultural digital literature, these polarity patterns not only signal the emotional tone of the narratives but also reflect broader socio-political dynamics, ideological conflicts, and processes of cultural transformation.

A mixed-method analysis of polarity distribution enables the identification of a collective emotional landscape shaped by the representation of cultural values (see **Figure 1**). This includes tracing how emotional tones correlate with themes of identity, power asymmetries, and modes of social coexistence. To illustrate these affective orientations more precisely, the following diagram presents the proportional distribution of sentiment polarity within the annotated corpus. This visualization provides an empirical baseline for more granular analysis of the relationship between emotion, cultural values, and narrative structure—offering critical insights into how Indonesian digital literature functions as a conduit for cultural discourse and affective meaning-making.

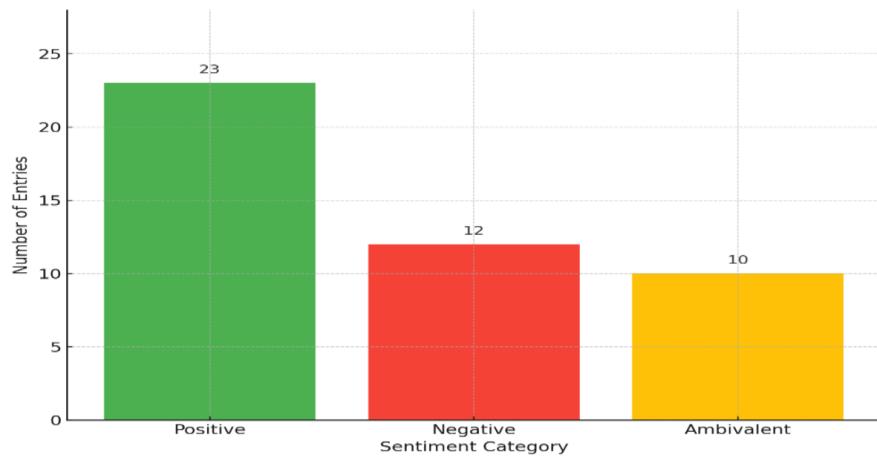


Figure 1. Distribution of sentiment categories in the narrative corpus.

The visualization of sentiment polarity distribution reveals that approximately 45% of the annotated narratives express negative sentiment, predominantly rooted in themes such as cultural conflict, discrimination, structural inequality, and rejection of dominant cultural norms. These texts frequently employ rhetorical strategies including symbolic opposition, conflict-driven dialogue, and representations of intergroup tensions, which serve to foreground critique, resistance, and contestation.

In contrast, 32% of the texts convey positive sentiment, commonly found in narratives that emphasize solidarity, spirituality, mutual cooperation (*gotong royong*), and social harmony. Such narratives tend to idealize traditional values, positioning them as cohesive forces essential to fostering communal bonds and cultural continuity.

Meanwhile, 23% of the texts exhibit ambivalent sentiment, characterized by a reflective or exploratory tone. These narratives often depict value dilemmas or negotiations of cultural meaning, particularly in the context of sociocultural transformation, modernity, or globalization.

Collectively, these three sentiment categories form a comprehensive spectrum of cultural emotion, illustrating that digital literary narratives in Indonesia do not merely serve an entertainment function but also operate as dynamic platforms for articulating evolving social attitudes, cultural identities, and emotional landscapes in contemporary society.

3.1.3. Distribution of Cultural Values in the Corpus

The distribution of cultural values within this digital literary corpus reflects the diverse ways in which Indonesian society constructs, interprets, and expresses cultural meanings through narrative. Through a rigorous and systematic annotation process, each text was classified according to its embedded cultural values, including spiri-

tuality, cooperation (*gotong royong*), social harmony, cultural resistance, and value adaptation (see **Figure 2**). This distributional analysis not only quantifies the frequency of each cultural value but also uncovers prominent thematic orientations shaping the discourse of multicultural digital literature. Consequently, the resulting cultural value map provides a critical foundation for understanding how local and universal values are negotiated, sustained, or contested within broader processes of social transformation.

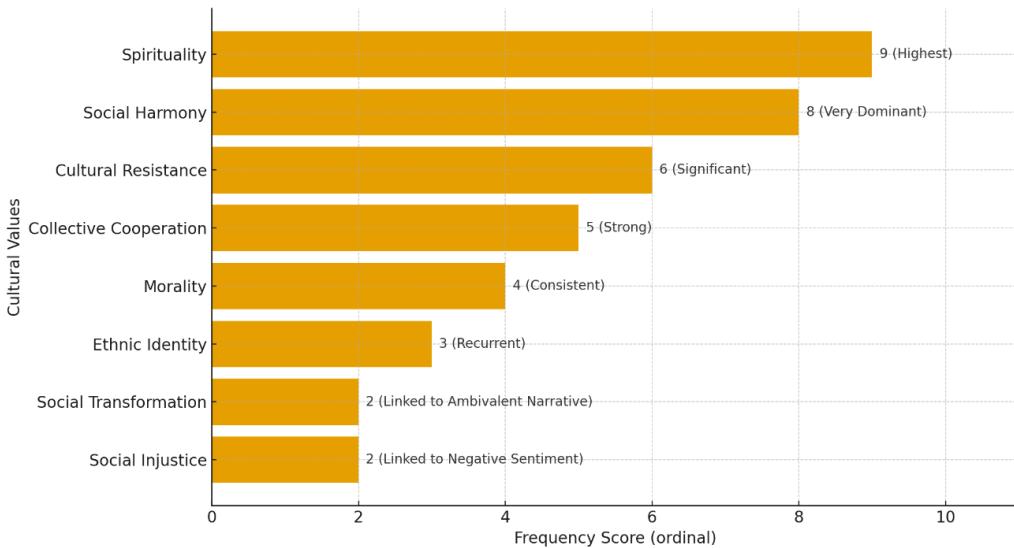


Figure 2. Frequency of cultural values in the digital literary corpus.

The frequency visualization of cultural values in the corpus indicates that Indonesian digital narratives serve purposes beyond mere personal expression. They function as collective platforms for articulating and renegotiating dynamic and living social values. Classification data reveal that spirituality and social harmony emerge as the two most dominant values, each appearing at the highest proportion among all categories. This prominence affirms that digital texts are often framed within ethical and religious orientations, emphasizing themes such as the human–divine relationship, sincerity, compassion, and efforts to maintain social equilibrium through tolerance and coexistence. The centrality of these values suggests that the cultural emotions expressed in the corpus are deeply rooted in both universal and contextually grounded moral ideals.

In addition, cultural resistance and value adaptation occur with notable frequency. Cultural resistance is most visible in narratives critiquing hegemonic power structures, interrogating colonial legacies, or amplifying marginalized voices resisting forced assimilation. In contrast, value adaptation appears in texts depicting negotiations between traditional values and modern challenges—such as representations of urban female protagonists, contemporary religious practices, or intergenerational dynamics. The prominence of these two values underscores the role of digital literature as a dialogic arena where the tensions between cultural preservation and renewal are actively explored, serving as a medium for responding to the disruptions brought about by globalization.

Although not the most dominant in terms of frequency, the value of *gotong royong* exhibits a strong and consistent association with positive sentiment. Narratives embodying this value are almost invariably framed within contexts of social solidarity, collective assistance, and inter-identity collaboration. This persistence indicates that *gotong royong* remains a widely resonant and positively received cultural value among readers and writers of digital literature. It appears not only as a normative ideal but also as a social practice symbolically reproduced through digital storytelling.

Taken together, the classification of cultural values—based on both their frequency of occurrence and emotional associations—demonstrates that the corpus reflects a broad spectrum of collective values, ranging from normative and affirmative to critical and transformative. These findings reinforce the position of Indonesian digital literature as a dynamic arena for cultural articulation, where social values are not only preserved but also negotiated and, at times, challenged within the evolving landscape of Indonesia's multicultural society.

3.1.4. Cross-Analysis: Sentiment Polarity vs. Cultural Values

The cross-tabulation of cultural values and sentiment polarity reveals a consistent pattern linking specific cultural dimensions to corresponding affective tendencies (see **Figure 3**). The data show that social harmony and *gotong royong* are predominantly associated with positive sentiment, suggesting that these values are typically embedded within affirming narratives emphasizing empathy, tolerance, and cross-group solidarity. By contrast, values such as value adaptation and cultural identity frequently appear in ambivalent contexts, indicating that narratives centered on these themes often portray tension, dilemmas, or processes of value negotiation in the face of modernity, globalization, or generational change.

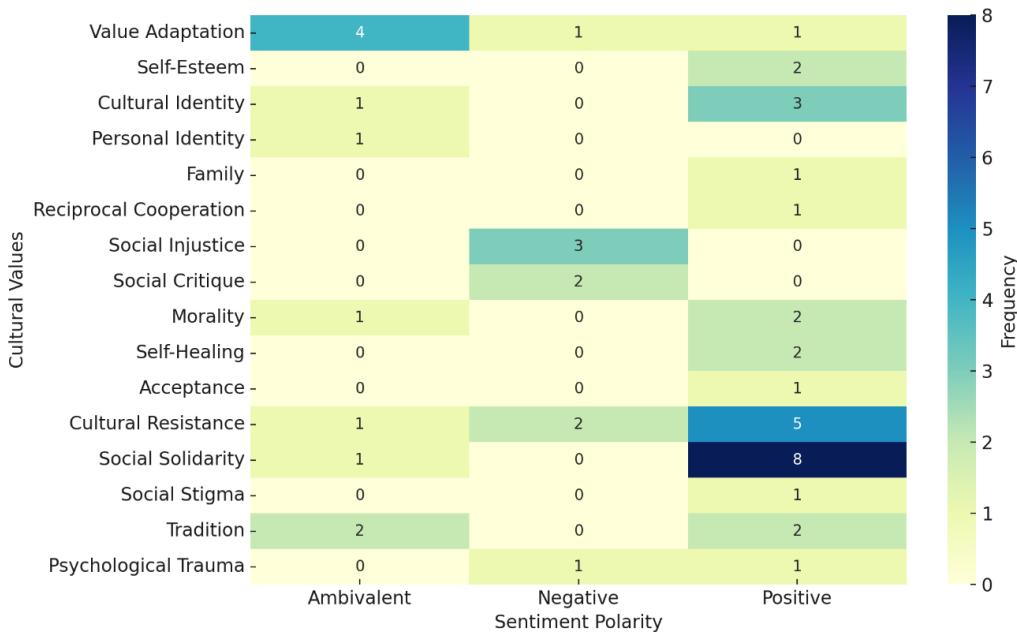


Figure 3. Heatmap of sentiment polarity toward cultural values.

Conversely, cultural resistance and self-worth exhibit stronger correlations with negative sentiment, particularly in texts addressing marginalization, subordination of local identities, or critiques of mainstream cultural dominance. These relationships are visualized in the accompanying heatmap, which highlights the intensity of association between each cultural value category and its corresponding sentiment polarity.

This visualization underscores that cultural values function not merely as thematic frameworks but also as emotional markers representing distinct social and ideological positions within the discourse of digital literature. As such, the relationship between cultural values and sentiment polarity in this corpus is not neutral but instead reflects a dynamic, context-dependent configuration of emotions and values.

The main findings from the heatmap—summarized in **Table 2**—reveal significant patterns of association between cultural value types and sentiment polarity tendencies within the multicultural digital literature corpus. Narratives embodying social harmony and *gotong royong* consistently project positive sentiment, constructing discourses of coexistence, tolerance, and collective solidarity as solutions to complex social realities. By contrast, values such as cultural resistance and self-worth are more frequently aligned with negative sentiment, reflecting narratives of struggle against structural injustice, cultural domination, and the erasure of local identities. Meanwhile, value adaptation and cultural identity are predominantly expressed in ambivalent tones, reflecting a reflective and nonpartisan narrative stance—particularly in portraying tensions between tradition and modernity or between local and global contexts. This heatmap not only illustrates the quantitative distribution of these associations but also reveals the emotional configurations embedded in the representation of cultural values. This offers a deeper understanding of how collective emotions are constructed through digital narratives as responses to contemporary sociocultural conditions.

Table 2. Key findings from the sentiment polarity heatmap towards cultural values.

No	Cultural Value	Dominant Polarity	Interpretation
1	Social Harmony	Positive	Narratives of coexistence and cross-identity empathy.
2	Value Adaptation	Ambivalent	Narratives of value negotiation in modern/global contexts.
3	<i>Gotong Royong</i>	Positive	Consistently associated with collective acts of solidarity.
4	Self-worth	Ambivalent/Positive	Narratives of struggle for social recognition and autonomy.
5	Cultural Identity	Ambivalent	Narratives of hybridity and dilemmas between locality and modernity.

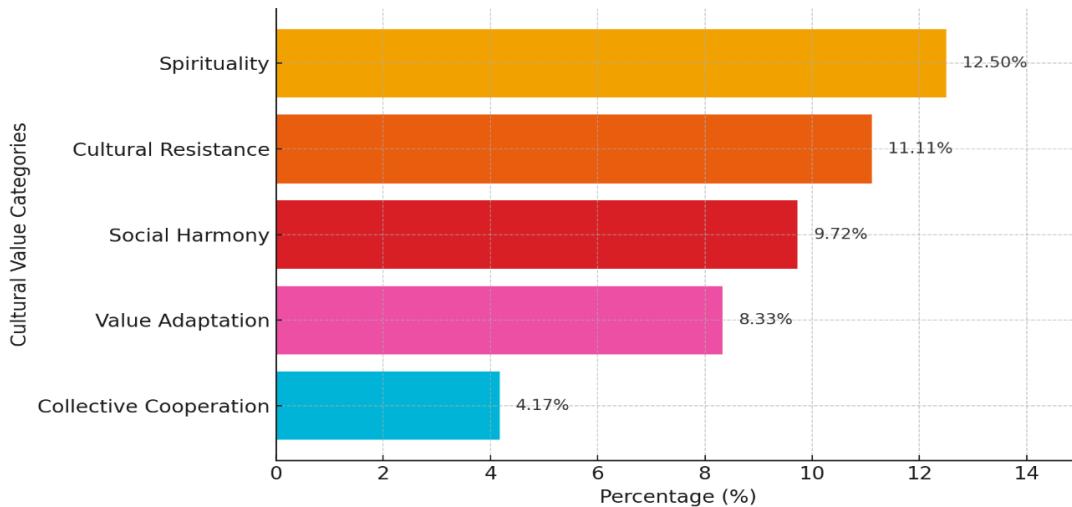
The results of this polarity distribution analysis and its relationship to cultural values show that positive sentiment consistently appears in narratives that promote universal ethical values such as honesty, sincerity, social empathy, and togetherness. These narratives typically feature characters, events, or social interactions that emphasize solidarity, tolerance, and harmonious collective living—particularly through values such as *gotong royong* and social harmony. The presence of these values not only reinforces the function of texts as educational or affirming instruments but also suggests that digital literature plays an active role in strengthening social cohesion and constructing positive, solution-oriented cultural representations amidst diversity.

On the other hand, negative and ambivalent sentiments in the corpus should not be narrowly interpreted as destructive expressions. Instead, they indicate the presence of healthy spaces for social critique, as well as reflective arenas for value negotiation and transformation. Narratives with negative polarity frequently emerge in contexts involving cultural resistance and self-worth, reflecting critical awareness toward the dominance of majority cultures, historical injustices, or the pressures of globalization. Meanwhile, values such as value adaptation and cultural identity are often associated with ambivalent sentiment, demonstrating that digital narratives also serve as transitional spaces, where cultural meaning is not fixed but continuously renegotiated.

The heatmap further reinforces these findings by highlighting that adaptive and resistant value domains are among the most narratively productive areas for conveying complex emotions, building narrative tension, and simultaneously opening up possibilities for social transformation through digital literature.

3.2. Categories of Dominant Cultural Sentiment

An analysis of the annotated corpus identified five cultural values most frequently articulated in Indonesian digital literature: spirituality, cultural resistance, social harmony, value adaptation, and cooperation (*gotong royong*) (see **Figure 4**). These categories were derived through thematic coding and are visualized in the accompanying bar chart to provide quantitative support for interpretive analysis.

**Figure 4.** Distribution of five dominant cultural sentiment categories.

Among these values, spirituality emerged as the most prominent (12.5%), indicating that Indonesian digital narratives frequently embed religious and spiritual elements as key emotional and moral framing devices. Notably,

this value is not limited to formal religious settings but also encompasses symbolic expressions, local traditions, and ancestral heritage interwoven into narrative structures. Cultural resistance (11.11%) follows closely, functioning as a narrative mechanism through which dominant hegemonic values are critiqued. Such narratives often operate as symbolic sites of resistance against marginalization and cultural homogenization.

Social harmony (9.72%) and value adaptation (8.33%) represent two complementary narrative orientations. Narratives emphasizing harmony tend to reinforce themes of coexistence and solidarity, whereas those highlighting value adaptation reflect dynamic processes of cultural negotiation in response to modernity and social change. Although cooperation (*gotong royong*) appears least frequently (4.17%), it consistently carries positive associative meaning, suggesting that collective values continue to hold a significant position in the cultural memory embedded in digital storytelling.

Thematic quotations within the corpus further illustrate the diversity of narrative types and enable classification into distinct thematic genres, as exemplified in **Table 3**.

Table 3. Examples of thematic quotations and their corresponding genres.

Thematic Quotation	Indicated Thematic Genre
"The character's inner dialogue about love and freedom..."	Psychological/Romantic
"A woman heals trauma through spirituality..."	Spirituality/Psychological
"The protagonist learns <i>karawitan</i> as a path to self-discovery..."	Local Culture/Educational
"Bullying trauma unaddressed by consequences..."	Social/Critical
"Group collaboration to find a missing person..."	Communal/Action
"Lintang treated unfairly despite being innocent..."	Resistance/Injustice
"Teacher and student find strength through life philosophy..."	Educational/Spirituality

Figure 5 depicts the results of the analysis of the distribution of thematic genres in the digital literary corpus based on keyword-based classification of thematic narratives.

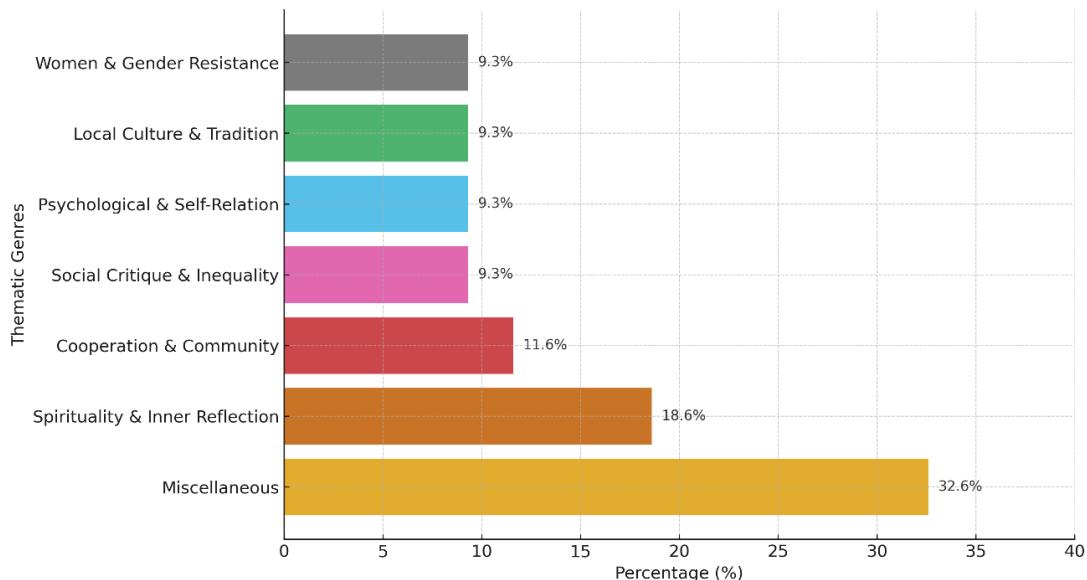


Figure 5. Distribution of thematic genres in digital literary narratives.

Through keyword-based classification, six primary thematic genres were identified as dominant in the digital literature corpus (see **Figure 6**). The Spirituality & Inner Reflection genre constitutes the largest share, representing approximately 18.6% of all narratives. This dominance underscores a strong tendency to articulate cultural values through contemplative and religious lenses. The second most prominent category, Cooperation & Community (11.6%), is characterized by narratives centered on solidarity, collective ethics, and mutual effort in addressing societal challenges.

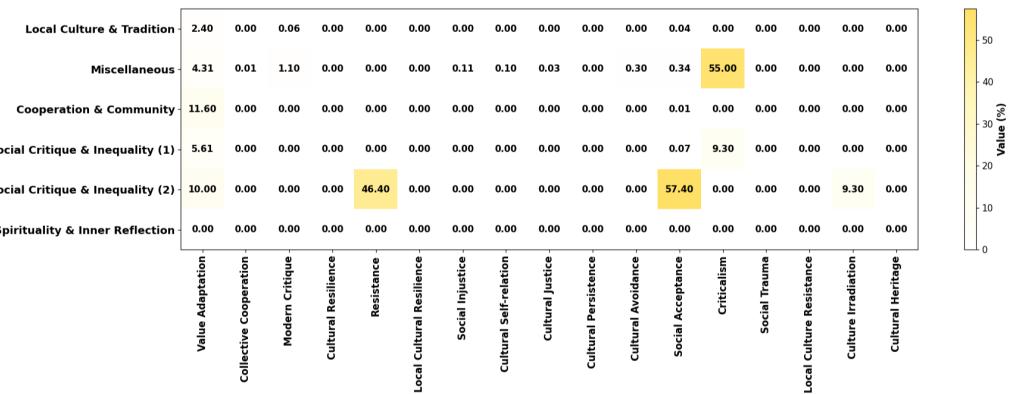


Figure 6. Heatmap of the relationship between thematic genres and cultural values.

The genres of Social Critique & Inequality and Psychological & Self-Relation each account for 9.3% of the corpus, reflecting the role of digital literature as a platform for expressing dissatisfaction with social conditions and exploring individual identity. Although appearing in smaller proportions, Local Culture & Tradition and Women & Gender Resistance remain significant, as they encapsulate strong narratives of cultural preservation and critique of patriarchal or global hegemonies.

Approximately 32.6% of narrative excerpts could not be explicitly assigned to these dominant genres and were categorized under Miscellaneous. This substantial proportion indicates the narrative diversity within the corpus and the presence of hybrid or unconventional forms that transcend predefined genre boundaries.

These findings affirm the role of narrative genre not only as a thematic container but also as a vehicle for articulating values, emotions, and cultural orientations in a structured and coherent manner.

A cross-tabulation of thematic genres and cultural values—visualized through a heatmap and supported by tabular data—reveals distinctive patterns of association, reinforcing the dual function of narrative as both ideological and emotional discourse.

The Women & Gender Resistance genre exhibits the strongest correlation with the value of Cultural Resistance (60%), reflecting the genre's engagement with themes of structural oppression and social critique. This finding aligns with feminist and postcolonial theoretical perspectives that regard the female body as a political site in struggles over identity and representation.

Similarly, the Local Culture & Tradition genre demonstrates substantial alignment with Value Adaptation and Cultural Resistance (each 33.33%), suggesting that digital representations of cultural heritage involve not only the reproduction of traditional values but also their active negotiation in the context of modernity and globalization.

The Cooperation & Community genre is most strongly associated with *gotong royong* (33.33%), and also shows meaningful relationships with Tolerance (22.22%) and Social Harmony (11.11%). This pattern underscores the centrality of communal ethics and collective resilience in narratives addressing shared challenges.

The Psychological & Self-Relation genre presents a relatively even distribution across three key values—Self-Esteem, Spirituality, and Cultural Resistance (each 28.57%)—suggesting that personal narratives serve as reflective spaces through which identity dynamics, social pressures, and existential meaning-making are explored.

Although the Miscellaneous genre does not correspond with the five major categories, it continues to reflect cultural value constructions, particularly those associated with Spirituality (17.39%) and Social Harmony (21.74%). This indicates that the expression of cultural values may also emerge through flexible, hybrid narrative forms beyond conventional thematic structures.

Collectively, these patterns demonstrate that narrative genres in Indonesian digital literature not only shape storytelling structures but also serve as critical frameworks for conveying cultural values, emotional dispositions, and ideological positions in an increasingly digitalized cultural landscape.

3.3. The Dynamics of Sentiment and Narrative Structure

The integrated application of Cultural Sentiment Analysis (CSA) and Critical Discourse Analysis (CDA) reveals that sentiment dynamics within the digital literature corpus are not purely affective, but are embedded within in-

tricate and ideologically charged discursive patterns. The findings indicate that affective expressions are closely entangled with narrative structures and the underlying cultural ideologies they reflect. Each sentiment type—positive, negative, or neutral—emerges not arbitrarily, but through specific discourse configurations that illustrate how authors construct representations of social experience, identity conflict, and intercultural dynamics. In this framework, digital literary narratives function as discursive arenas for the simultaneous articulation and negotiation of cultural meaning, enacted through both emotional resonance and linguistic structure.

Negative sentiment is most prominently observed in narrative contexts that foreground cultural dissonance—such as intergenerational conflict, ethnic tensions, and ideological value clashes. These narratives commonly draw upon rhetorical strategies including irony, symbolic opposition, and conflict-oriented dialogue to articulate affective expressions of frustration, alienation, and dissent toward dominant cultural paradigms. This pattern suggests that negative emotion functions as a medium of social critique, wherein conflict is not merely a narrative device but serves as a reflective mechanism to challenge hegemonic norms and catalyze value transformation. Structurally, such conflicts are often left unresolved, signaling the persistence of systemic tensions within the sociocultural landscape.

In contrast, positive sentiment is predominantly associated with narratives that depict intercultural collaboration, processes of reconciliation, and the construction of shared hope. These texts frequently employ storytelling techniques that highlight dialogue, empathy, and resolute plotlines, consistently foregrounding values such as cooperation (*gotong royong*), social harmony, and tolerance. Affective representations in these narratives do not merely affirm inclusive local values; they also emphasize the role of digital literature in fostering symbolic mediation and strengthening social cohesion. In numerous cases, positive sentiment is framed as a narrative response to collective trauma or prior conflict, illustrating a transformation from fragmentation to solidarity.

Meanwhile, narratives characterized by neutral sentiment occupy a distinct discursive position, serving as emotionally impartial spaces of cultural introspection. These narratives tend to focus on themes of identity formation, historical memory, and cultural value negotiation, often eschewing explicit emotional intensification. Their narrative style is largely descriptive and contemplative, relying on metaphor and cultural symbolism as key expressive instruments. The presence of neutral sentiment suggests that many texts seek not to provoke extreme affective responses but to invite readers into critical reflection on the complexities of cultural experience. Within this analytical framework, neutral sentiment contributes to the creation of narrative spaces that facilitate nuanced engagement with identity dynamics, beyond the binary of emotional affirmation or resistance.

4. Discussion

The findings of this study underscore the intricate ways in which cultural sentiment is encoded in Indonesian digital literature. Emotional expressions within the corpus are rarely articulated in overtly affective terms; instead, they are embedded in layered linguistic strategies that operate both semantically and discursively. Writers strategically mobilize metaphoric expressions, culturally situated lexicons, and narrative syntactic choices to convey affective meaning. Words such as *togetherness*, *heritage*, and *struggle* emerge not simply as thematic elements but as affective markers, signaling underlying cultural value orientations that guide the reader's interpretive engagement.

Beyond quantitative distribution, the CSA outputs were interpreted through a CDA lens to reveal underlying ideological configurations [36]. Negative sentiment clusters were not merely affective reactions but discursive manifestations of power asymmetry, cultural marginalization, and resistance to dominant norms—patterns that are increasingly observable through computational methods when paired with critical interpretation [37]. Conversely, positive sentiment patterns functioned as hegemonic counterbalances, constructing narratives of cohesion and normative social order, while ambivalent sentiment reflects sites of ideological negotiation in which competing value systems coexist without resolution. This integration demonstrates that computational “distant reading” can operate as an entry point for critical “close reading.” This approach enables systematic identification of discursive formations that structure cultural meaning in digital narratives [38, 39]. Methodologically, this mixed approach aligns with recent mixed-methods and “critical computation” arguments that emphasize interpretability, contextual grounding, and the need to connect pattern detection with ideological explanation rather than treating computational outputs as neutral evidence [37].

This observation resonates with recent scholarship that highlights how cultural values and sentiment are represented through narrative structures, metaphoric expression, and discursive strategies. Bimantara et al. [26]

show how Balinese narratives integrate cultural lexicons that fulfill both aesthetic and affective functions. Iskandar et al. [40] demonstrate that metaphors in Indonesian and Acehnese proverbs serve as vehicles for transmitting shared cultural sentiments and communal identities, while Hambali et al. [41] identify war and threat metaphors as affective resources shaping collective interpretations of complex sociopolitical issues. Complementing these perspectives, Martutik et al. [42] and Rahmi et al. [43] emphasize the role of discursive framing and counter-narratives in aligning affective positions in digital texts. These findings correspond closely with Milani and Richardson's [25] conceptualization of discourse and affect as intertwined semiotic systems that mediate emotion through culturally grounded language. Taken together, these elements reveal how narrative framing, character positioning, and thematic progression operate as affective infrastructures that shape meaning construction and reception [24,26,40,44].

Within this corpus, the distribution of sentiment polarity provides further evidence of the interplay between narrative structure and emotional encoding. Narratives characterized by negative sentiment tend to emerge from sites of cultural dissonance such as intergenerational tension, ethnic marginalization, and ideological conflict. These texts often deploy irony, symbolic oppositions, and conflict-driven dialogues as rhetorical devices to articulate frustration and critique, thereby functioning as textual spaces for cultural resistance. Such patterns mirror affective arcs observed in global narrative studies, including those described by Reagan et al. [45], Hipson and Mohammad [46], and Samothrakis et al. [47], who underscore the role of conflict frames and dialogic interaction in shaping emotional valence.

From a postcolonial perspective, these resistant affective structures can be understood as counter-discursive interventions. Texts annotated with high levels of "cultural resistance" sentiment explicitly challenge dominant ideologies—including those inherited from colonial legacies, global cultural imperialism, and state-centric narratives—amplifying subaltern voices and asserting cultural agency against erasure and homogenization. This dynamic is consistent with Spivak's [48] notion of the subaltern and Bhabha's [32] theorization of hybridity, in which marginal narratives intervene critically in hegemonic cultural formations. Empirical work in the Indonesian context supports this reading. Ahlstrand et al. [49] illustrate how narratives surrounding Kartini function as postcolonial interventions, while Dwifatma and Beta [50] document how Muslim women mobilize mediated subjectivities to critique gendered and political power structures. Winarnita et al. [51] further reveal how Indonesian female journalists utilize digital platforms to enact counter-hegemonic citizenship. Wilcock's [52] concept of "networked relationality" extends the idea of hybridity into digital spaces, where resistance manifests through hybridized cultural expressions—a tendency echoed in feminist counterpublics across the Global South [53].

Contrasting with these resistant affective modes is a parallel tendency toward reconciliation and solidarity, evident in narratives characterized by positive sentiment. These narratives emphasize tolerance, shared responsibility, and communal well-being, reflecting Indonesia's pluralistic sociocultural ethos. They draw deeply from indigenous philosophies such as *gotong royong* and *Bhinneka Tunggal Ika*, positioning cultural sentiment as a cohesive rather than confrontational force [54–56].

Such narratives align closely with Berry's [5,57] multicultural paradigm and dynamic acculturation models [58], acting as affective bridges across cultural differences and fostering empathy and a shared sense of identity. Instead of erasing differences, they affirm diversity through inclusive storytelling, underscoring the dual role of sentiment in unifying and diversifying cultural discourse. Between these two poles—resistance and reconciliation—lies a third affective orientation: neutral sentiment. This emerges primarily in reflective narratives that engage with questions of identity, memory, and cultural transformation without adopting explicit ideological positions. Such narratives embody the mnemonic dimension of literary discourse as cultural repositories, reinforcing Assmann's [59] framework of cultural memory. As Arps [60] observes, these forms of storytelling act as collective spaces for negotiation and reflection, enabling cultural communities to revisit and reinterpret their shared histories in digital environments.

In addition to its theoretical contributions, this study makes a methodological intervention in the field of digital literary analysis. The use of a three-layer annotation framework—integrating sentiment polarity, cultural value classification, and linguistic indicators—offers a scalable and culturally grounded model for examining affective discourse in large narrative corpora. By doing so, this approach reveals latent structures and discursive patterns that may be overlooked by traditional literary methods [28,61]. It is also aligned with contemporary affective analytics research, including emotional arc modeling [45], multidimensional linguistic interpretation [62], and the modeling of emotion dynamics through dialogue [46]. This methodological integration demonstrates how computational tools can be harmonized with culturally embedded literary interpretation to produce richer analyses of

digital texts.

Taken together, these findings contribute both theoretically and methodologically to ongoing debates in post-colonial literary studies, affective discourse theory, and digital humanities. They reveal the ambivalent role of Indonesian digital literature as both a site of ideological contestation and a medium of cultural cohesion. Theoretically, they extend Bhabha's concept of hybridity and Spivak's subaltern framework into digital narrative environments while enriching Berry's multicultural paradigm with indigenous affective frameworks such as *gotong royong* and *Bhinneka Tunggal Ika*. Methodologically, they offer a replicable annotation model applicable to multilingual and multicultural literary corpora, particularly within the Global South. This dual contribution positions Indonesian digital literature as both a culturally specific archive and a globally relevant site for examining how sentiment mediates discourse, identity, and power in the digital age.

5. Conclusion

This study has demonstrated that cultural sentiment in Indonesian digital literature is not merely an emotional expression but a narrative construction shaped through linguistic, discursive, and sociocultural strategies. By employing a multi-layered annotation scheme—encompassing sentiment polarity, cultural value classification, and linguistic indicators—this research presents a comprehensive, validated dataset that captures the richness and complexity of cultural emotions as expressed in contemporary digital narratives. The analysis reveals that metaphoric language, culturally embedded lexicons, and narrative structuring, such as character positioning and thematic progression, are integral to encoding cultural sentiment, offering insights into how emotional discourse is culturally grounded and narratively framed.

Furthermore, the findings underscore the dual function of cultural sentiment in digital literature: as a medium of resistance and as a vehicle of reconciliation. Narratives classified under “cultural resistance” challenge dominant ideologies—ranging from colonial legacies to cultural homogenization—by amplifying marginalized voices and reclaiming cultural agency. At the same time, stories emphasizing values such as social harmony and cooperation affirm the role of literature in building intercultural understanding and promoting national unity amid diversity. These seemingly divergent orientations reflect the dynamic tension within multicultural discourse, where sentiment operates as both critique and cohesion.

The integration of cultural sentiment analysis (CSA) with critical discourse analysis (CDA) in this study provides a methodological innovation for digital humanities. It bridges qualitative interpretation and data-driven computation, enabling scholars to explore affective patterns at scale while maintaining sensitivity to narrative nuance. The annotated corpus thus serves not only as a tool for empirical investigation but also as a framework for educational and policy interventions. As digital spaces become increasingly central to cultural negotiation, this research offers a foundation for developing literacy programs and cultural policies that recognize sentiment as both a mirror and a catalyst of social transformation.

In conclusion, this study contributes to expanding the methodological and theoretical landscape of sentiment analysis, particularly within culturally plural digital environments. It validates that digital literature is not a neutral textual form but an expressive site of affective discourse, identity formation, and cultural memory. While the present corpus focuses on Indonesian texts, its findings open avenues for comparative research across Southeast Asian contexts and beyond. Future studies may enrich this framework by incorporating multimodal narratives, experimenting with machine-assisted annotation models, and investigating the interplay between narrative sentiment and socio-political dynamics. Such extensions will further deepen our understanding of how culture, language, and emotion converge in the evolving landscape of digital storytelling.

Author Contributions

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

Funding

This research was funded by the Directorate of Research, Technology, and Community Service (DRTPM), Ministry of Higher Education, Science, and Technology of the Republic of Indonesia (Kemendikitsaintek), through its national research grant program.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Acknowledgments

The authors appreciate the Directorate of Research, Technology, and Community Service (DRTPM), Ministry of Higher Education, Science, and Technology of the Republic of Indonesia. The authors gratefully acknowledge the institutional support and resources that contributed significantly to the successful completion of this study.

Conflicts of Interest

The authors declare no conflict of interest.

AI Use Statement

ChatGPT (OpenAI, GPT-4.x) was used solely for linguistic purposes, specifically for grammar checking, sentence structure refinement, and improving the readability of the English text in this manuscript. The authors take full responsibility for all academic content, including all ideas, data, analyses, and conclusions presented herein. The use of AI was thoroughly reviewed and supervised by the authors.

References

1. Zheng, J.; Fan, W. Readers' reception of translated literary work: Fortress Besieged in the digital English world. *Digit. Scholarsh. Humanit.* **2023**, *38*, 408–419. [[CrossRef](#)]
2. Rebora, S.; Boot, P.; Pianzola, F.; et al. Digital humanities and digital social reading. *Digit. Scholarsh. Humanit.* **2021**, *36*, ii230–ii250. [[CrossRef](#)]
3. Chang, W.-C.; Lin, R.-T. Designing for wearable and fashionable interactions: Exploring narrative design and cultural semantics for design anthropology. *Interact. Stud.* **2020**, *21*, 200–219. [[CrossRef](#)]
4. Ricoeur, P. *Time and Narrative*; McLaughlin, K., Pellauer, D. Trans.; University of Chicago Press: Chicago, IL, USA, 2016; Volume 1.
5. Berry, J. Diversity and Equity. *Cross Cult. Strateg. Manag.* 2016, *23*, 413–430. [[CrossRef](#)]
6. Damayanti, I.L.; Moeharam, N.Y.; Asyifa, F. Unfolding layers of meanings: Visual–verbal relations in Just Ask—A children's picture book. *Indones. J. Appl. Linguist.* **2021**, *11*, 345–357. [[CrossRef](#)]
7. Cahyani, I.; Mardani, P.; Widianingsih, Y. Digital storytelling in cultural tourism. *Int. J. Manag. Entrep. Soc. Sci. Humanit.* **2023**, *6*, 45–69.
8. Nensilianti; Jahrir, A.S.; Saguni, S.S.; et al. Toxic talk and narrative power in virtual arenas: A pragmatic-narrative analysis of impoliteness in online game communication. *Forum Linguist. Stud.* **2024**, *7*, 617–631. [[CrossRef](#)]
9. APJII Internet Survey 2025; APJII: Jakarta, Indonesia, 2025. Available online: <https://survei.apjii.or.id/> (accessed on 1 December 2025). (in Indonesian)
10. Nensilianti; Hajrah; Ridwan; et al. Decoding the narrative syntax and macrostructure of Toraja folklore: A quest for meaning. *J. Lang. Teach. Res.* **2025**, *16*, 493–504. [[CrossRef](#)]

11. Adji, R.D.; Setiawan, I. Literature and identity in the digital era: A multicultural perspective. *J. Poet.* **2020**, *8*, 23–34. (in Indonesian)
12. Ismawati, S.; Suyata, S.; Marmanto, S.; et al. Multiculturalism in Indonesian literature as a teaching material for literary appreciation. *J. Pendidik. Bahas.* **2019**, *8*, 19–33. (in Indonesian)
13. Younis, E. New horizons of digital literary criticism. *Digit. Scholarsh. Humanit.* **2021**, *36*, 501–508. [CrossRef]
14. Cowen, A.S.; Brooks, J.A.; Prasad, G.; et al. How emotion is experienced and expressed in multiple cultures: A large-scale experiment across North America, Europe, and Japan. *Front. Psychol.* **2024**, *15*, 1350631. [CrossRef]
15. Kubin, E.; von Sikorski, C. The role of (social) media in political polarization: A systematic review. *Ann. Int. Commun. Assoc.* **2021**, *45*, 188–206. [CrossRef]
16. Liu, B. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*; Cambridge University Press: Cambridge, UK, 2020.
17. Mohammad, S.M.; Turney, P.D. Crowdsourcing a word–Emotion association lexicon. *Comput. Intell.* **2013**, *29*, 436–465. [CrossRef]
18. Fairclough, N. *Language and Power*; Routledge: London, UK, 2013.
19. Winata, G.I.; Aji, A.F.; Cahyawijaya, S.; et al. NusaX: Multilingual parallel sentiment dataset for 10 Indonesian local languages. arXiv preprint 2022, arXiv:2205.15960. [CrossRef]
20. Liu, C.C.; Gurevych, I.; Korhonen, A.; et al. Culturally aware and adapted natural language processing: A taxonomy and a survey of the state of the art. *Trans. Assoc. Comput. Linguist.* **2024**, *13*, 652–689. [CrossRef]
21. Santy, S.; Liang, J.T.; Le Bras, R.; et al. NLPositionality: Characterizing design biases of datasets and models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*; Rogers, A., Boyd-Graber, J., Okazaki, N., Eds.; Association for Computational Linguistics: Kerrville, TX, USA, 2023; Volume 1, pp. 9080–9102.
22. Pawar, S.; Park, J.; Jin, J.; et al. Survey of cultural awareness in language models. *Comput. Linguist.* **2025**, *51*, 907–1004. [CrossRef]
23. Kadan, A.; Padmanabhan, D.; Bhadra, S. Understanding latent affective bias in large pre-trained language models. *Nat. Lang. Process. J.* **2024**, *7*, 100062. [CrossRef]
24. Underwood, T. *Distant Horizons: Digital Evidence and Literary Change*; University of Chicago Press: Chicago, IL, USA, 2019.
25. Milani, T.M.; Richardson, J.E. Discourse and affect. *Soc. Semiotics* **2021**, *31*, 671–676. [CrossRef]
26. Bimantara, I.M.S.; Purwitasari, D.; Sanjaya ER, A.S.; et al. Balinese story texts dataset for narrative text analyses. *Data Brief* **2024**, *56*, 110781. [CrossRef]
27. Blei, D.M. Probabilistic topic models. *Commun. ACM* **2012**, *55*, 77–84.
28. Jockers, M.L. *Macroanalysis: Digital Methods and Literary History*; University of Illinois Press: Urbana, IL, USA, 2013.
29. Evidence-Based Policy: Functional Area 1 of the SDG4 High-Level Steering Committee. Available online: <https://www.unesco.org/sdg4education2030/en/evidence-and-policy> (accessed on 1 December 2025).
30. Wang, F.; Hannafin, M.J. Design-based research and technology-enhanced learning environments. *Educ. Technol. Res. Dev.* **2005**, *53*, 5–23.
31. Hall, S. *Representation: Cultural Representations and Signifying Practices*; Sage: London, UK, 1997.
32. Bhabha, H.K. *The Location of Culture*; Routledge: London, UK, 1994.
33. Anggraini, A. Lampung culture in the short story “Sebambangan” by Budi P. Hatees. *Aksara* **2017**, *29*, 49–62. (in Indonesian)
34. Amiel, T.; Reeves, T.C. Design-based research and educational technology: Rethinking technology and the research agenda. *Educ. Technol. Soc.* **2008**, *11*, 29–40.
35. Artstein, R.; Poesio, M. Inter-coder agreement for computational linguistics. *Comput. Linguist.* **2008**, *34*, 555–596. [CrossRef]
36. Catalano, T.; Waugh, L.R. *Critical Discourse Analysis, Critical Discourse Studies and Beyond*; Springer: Cham, Switzerland, 2020. [CrossRef]
37. Lukito, J. Critical computation: Mixed-methods approaches to big text data analysis. *Rev. Commun.* **2023**, *23*, 62–78.
38. Houston, L. Distant reading. In *Technology and Literature*; Hammond, A., Ed.; Cambridge University Press: Cambridge, UK, 2023.
39. Consilvio, D. Computational Close Reading: A Critique of Digital Literary Studies. PhD Thesis, University of Rhode Island, Kingston, RI, USA, 2023.

40. Iskandar, D.; Islahuddin; Harun, M.; et al. Metaphors in Indonesian and Acehnese proverbs of similar meanings: Semantic and cultural analyses. *Indones. J. Appl. Linguist.* **2025**, *14*, 512–523. [[CrossRef](#)]

41. Hambali, M.; Istianah, A.; Susilowati, N.E.; et al. Battling the climate crisis: WAR and THREAT metaphors in Indonesian news media through a corpus-ecolinguistics lens. *Cogent Arts Humanit.* **2025**, *12*, 2526143. [[CrossRef](#)]

42. Martutik; Setiawan, A.; Rani, A.; et al. Exploring flaming in the discourse of negative judgment: Invoked strategies used by Indonesian netizens in Instagram comments. *Cogent Arts Humanit.* **2024**, *11*, 2333601. [[CrossRef](#)]

43. Rahmi, R.; Suhardijanto, T.; Pratama, H.S.; et al. Tales of health crises: Indonesia's dynamic narratives and counter-narratives about pandemics. *Cogent Arts Humanit.* **2025**, *12*, 2451507. [[CrossRef](#)]

44. Winchcombe, Z. Time in the Gutter: A Narratological Approach to the Comics Medium. Master's Thesis, University of Ottawa, Ottawa, ON, Canada, 2018.

45. Reagan, A.J.; Mitchell, L.; Kiley, D.; et al. The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Sci.* **2016**, *5*, 31. [[CrossRef](#)]

46. Hipson, W.E.; Mohammad, S.M. Emotion dynamics in movie dialogues. *Plos One* **2021**, *16*, e0256153. [[CrossRef](#)]

47. Samothrakis, S.; Tsakalidis, A.; Fasli, M. Emotional sentence annotation helps predict fiction genre. *Plos One* **2015**, *10*, e0141922. [[CrossRef](#)]

48. Spivak, G.C. Can the subaltern speak? In *Marxism and the Interpretation of Culture*; Nelson, C., Grossberg, L., Eds.; University of Illinois Press: Urbana, IL, USA, 1988; pp. 271–313.

49. Ahlstrand, J.; Maniam, V. Kartini, online media, and the politics of the Jokowi era: A critical discourse analysis. *Asian Stud. Rev.* **2025**, *49*, 78–98. [[CrossRef](#)]

50. Dwifatma, A.; Beta, A.R. The “Funny Line Veil” and the mediated political subjectivity of Muslim women in Indonesia. *Asian J. Commun.* **2024**, *34*, 284–297. [[CrossRef](#)]

51. Winarnita, M.; Bahfen, N.; Mintarsih, A.R.; et al. Gendered digital citizenship: How Indonesian female journalists participate in gender activism. *Journalism Pract.* **2022**, *16*, 621–636. [[CrossRef](#)]

52. Wilcock, C.A. From hybridity to networked relationality: Actors, ideologies and the legacies of sudan's comprehensive peace agreement. *J. Interv. Statebuild.* **2021**, *15*, 221–243. [[CrossRef](#)]

53. Osei-Appiah, S.; Mutsvairo, B.; Orgeret, K.S. Women and digital political communication in non-Western societies. *Inf. Commun. Soc.* **2023**, *26*, 2507–2517. [[CrossRef](#)]

54. Lukiyanto, K.; Wijayaningtyas, M.; Bernarto, I.; et al. Gotong royong as social capital to overcome micro and small enterprises' problems. *Heliyon* **2020**, *6*, e04879. [[CrossRef](#)]

55. Syahputra, W.; Dhowi, B.; Sianipar, S.M.; et al. The psychological perspective of Bhinneka Tunggal Ika: A national unifying tool. *Open Psychol. J.* **2023**, *16*. [[CrossRef](#)]

56. Humaedi, M.A.; Wibowo, D.P.; Hariyanto, W.; et al. Shifting collective values: The role of rural women and gotong royong in village fund policy. *Humanit. Soc. Sci. Commun.* **2025**, *12*, 411. [[CrossRef](#)]

57. Berry, J.W. Acculturation: Living successfully in two cultures. *Int. J. Intercult. Relat.* **2005**, *29*, 697–712. [[CrossRef](#)]

58. Sherman, S.R.; Koven, M.J. *Folklore/Cinema: Popular Film as Vernacular Culture*; University Press of Colorado: Boulder, CO, USA, 2007. [[CrossRef](#)]

59. Assmann, J. *Cultural Memory and Early Civilization: Writing, Remembrance, and Political Imagination*; Cambridge University Press: Cambridge, UK, 2011.

60. Arps, A. Memori melompat ('jumping memory'): The mnemonic motion of Indonesian popular culture and beyond. *Memory Stud.* **2023**, *16*, 1579–1594. [[CrossRef](#)]

61. Piper, A. *Enumerations: Data and Literary Study*; University of Chicago Press: Chicago, IL, USA, 2018. [[CrossRef](#)]

62. Liu, Z.; Wan, G.; Zuo, X.; et al. Sentiment analysis of Chinese ancient poetry based on multidimensional knowledge attention. *Digit. Scholarsh. Humanit.* **2025**, *40*, 214–226. [[CrossRef](#)]



Copyright © 2026 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.