Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 5 No. 2 (2026) pp: 15418-15425
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Implementation of Knowledge Graph as a Representation of Public Sentiment Analysis Toward *AI-Generated Art*

Finna Nur Nandia[1], I Kadek Dwi Nuryana[2]
[1,2]Sistem Informasi, Fakultas Teknik, Universitas Negeri Surabaya
[1]finnanur.22005@mhs.unesa.ac.id, [2]dwinuryana@unesa.ac.id

Abstract

The rapid advancement of generative AI technology has elicited diverse public responses on social media, particularly toward AI-generated art, which has significantly impacted Indonesia's creative industry. This study aims to analysed the sentiment of Indonesian society on platform X toward AI-generated art and to represent the interconnections among sentiment entities through a Knowledge Graph (KG). The approach integrates three primary methods within the Knowledge Discovery in Databases (KDD) framework: sentiment classification using the IndoBERT model, topic modelling using Latent Dirichlet Allocation (LDA), and KG construction based on Social Network Analysis (SNA). The dataset consists of 4,765 Indonesian-language tweets that underwent pre-processing. Sentiment analysis results indicate a dominance of negative sentiment (45.4%) over positive sentiment (42.5%), with the IndoBERT model achieving 69% accuracy on a three-class classification task. Topic modelling produced 18 distinctive topics (9 negative, 9 positive), validated through Two-Stage Similarity validation. Negative topics are dominated by issues of economic impact on illustrators, copyright infringement, and artistic style theft, while positive topics reflect appreciation for AI as a creative tool. The constructed KG comprises 144 nodes and 329 edges with a modularity score of 0.4556, reflecting 7 meaningful thematic communities. SNA evaluation reveals that 'ilustrasi' (illustration) is the most central entity (degree centrality = 12.621), while negative issues dominate the central positions among topic nodes. This study demonstrates that the integration of IndoBERT, LDA, and KG is capable of uncovering hidden relational patterns in public opinion that cannot be obtained through conventional sentiment analysis alone.

Keyword: *AI-Generated Art, Sentiment Analysis, IndoBERT, LDA, Knowledge Graph, SNA*

1. Introduction

The development of generative AI technology has experienced significant growth in recent years, marked by the emergence of models such as Midjourney and DALL-E, which are capable of automatically producing high-quality visual works. This phenomenon directly impacts the global creative industry, including Indonesia, which ranks third in the world in the number of AI site visits [1]. The presence of AI-generated art has not only transformed how visual content is produced, but has also triggered widespread public debate on social media regarding issues of copyright, artistic originality, and the sustainability of the illustrator profession, particularly on platform X, which serves as a frequently used discussion forum for Indonesian society.

Understanding public sentiment toward AI-generated art is critical, given that societal perceptions will directly influence the formulation of creative industry policies and the protection of artists' rights. However, conventional sentiment analysis only yields polarity classifications (positive, negative, neutral) without providing a deeper understanding of the topics underlying those sentiments. Public opinion on social media is highly dependent on the thematic context under discussion, whether it concerns copyright infringement, economic impact on creators, or appreciation for AI as a creative tool. Without adequate topic modeling, the resulting picture of public sentiment will be fragmented and insufficiently informative for decision-makers.

Several prior studies have demonstrated the superiority of the IndoBERT model over traditional classification methods in Indonesian-language sentiment analysis; however, neither has linked sentiment results to topic structure [2][3]. Nabila [4] addressed public sentiment toward AI-generated images using the Naive Bayes algorithm, yet its accuracy remains lower than deep learning-based models. Meanwhile, Sahria and Fudholi [5] explored the combination of LDA and Knowledge Graph in the health domain and application reviews. Nevertheless, these studies have not examined public opinion on social phenomena such as AI-generated art.

Accordingly, no prior study has been found that specifically integrates IndoBERT, LDA, and Knowledge Graph to analysed Indonesian-language public discourse on AI-generated art.

This study proposes an integrative approach within the Knowledge Discovery in Databases (KDD) framework that combines three core components: sentiment classification using an IndoBERT model fine-tuned on Indonesian-language tweet data, topic modelling using Latent Dirichlet Allocation (LDA) to identify major themes per sentiment group, and Knowledge Graph construction to represent semantic relationships among sentiments, topics, and keywords. The resulting graph is evaluated using Social Network Analysis (SNA) metrics to reveal the most influential entities and relational patterns in public discourse.

2. Methods

This study constitutes quantitative, computationally-based research of an exploratory and analytical nature, designed to extract knowledge from unstructured data in the form of social media text. The methodological framework employed is Knowledge Discovery in Databases (KDD), encompassing five systematic stages, Data Selection, Data Cleaning and Preprocessing, Data Transformation, Data Mining, and Pattern Evaluation and Interpretation. Data were collected through a crawling technique using the Python library Tweet Harvest [6] with relevant keywords such as 'gambar AI' (AI image), 'ilustrasi AI' (AI illustration), and 'karya AI' (AI artwork) from platform X over the period January 2024 to October 2025, yielding 4,776 raw data records.

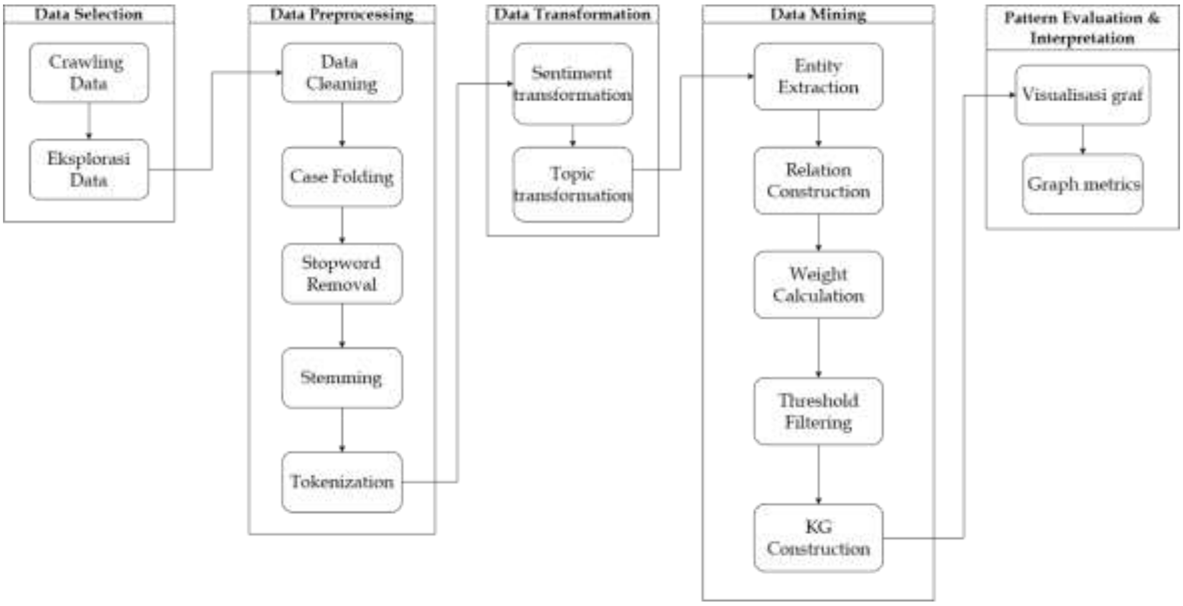

Figure 1. Knowledge Discovery in Databases Framework

The pre-processing stage differentiates workflows based on model requirements. IndoBERT employs cleaned text without stemming in order to preserve linguistic context, whereas for LDA, full pre-processing is applied, including case folding, stopword removal, and stemming using the Sastrawi and NLTK libraries, in accordance with recommendations in [7]. At the Data Transformation stage, sentiment classification was performed using the IndoBERT model (indobenchmark/indobert-base-p2), fine-tuned on manually labeled data with a three-class scheme (positive, negative, neutral) and an 80:20 train-test split ratio using stratified sampling, implemented via the PyTorch framework and the HuggingFace Transformers library [8]; model performance was subsequently evaluated using accuracy, precision, recall, and F1-score metrics.

Topic modelling was conducted using Latent Dirichlet Allocation (LDA) based on the Gensim library with a Bag-of-Words representation, wherein the optimal number of topics (k) was determined by testing the coherence score using the C_v metric with a maximum of 200 training iterations, in accordance with the characteristics of short Twitter text [9]. Topic quality was validated using the Two-Stage Similarity method, combining Cosine Similarity ($\tau_1 \geq 0.70$) and Jensen-Shannon Divergence (JSD ≤ 0.50), following Blair et al. [14].

At the Data Mining stage, the Knowledge Graph was constructed using NetworkX, with nodes representing dominant keywords per LDA topic and edges representing co-occurrence relationships and topic associations. Edge weights were combined from LDA distribution probabilities, co-occurrence frequencies, and aggregated sentiment scores normalized via Min-Max Scaling. Low-weight edges were filtered using a threshold based on

DOI: https://doi.org/10.31004/riggs.v5i2.11349

the mean and standard deviation of the weight distribution ($\tau = \mu + \sigma$), following the backbone extraction principle [10].

KG evaluation was conducted using Social Network Analysis (SNA) with metrics including degree centrality, betweenness centrality, closeness centrality, eigenvector centrality, clustering coefficient, and modularity, to identify community structures and the most influential entities in the network [11]. The entire implementation was executed in the Google Colab environment using the Python programming language, and the final KG visualization was exported in GEXF format for rendering in Gephi software.

3. Results and Discussion

This section presents the results and discussion of each stage of the applied pipeline, encompassing sentiment classification using IndoBERT, topic modeling using LDA, Knowledge Graph construction, and graph structure evaluation using SNA.

3.1. Sentiment Classification

The IndoBERT model (indobenchmark/indobert-base-p2), fine-tuned on 997 manually labelled records, yielded an accuracy of 69.00% on a test set of 200 records, with a precision value of 0.7049, recall of 0.6900, and F1-score of 0.6500. This performance is consistent with the literature for three-class sentiment classification on informal social media text, wherein inter-class ambiguity and label distribution imbalance are recognized as performance-suppressing factors [3][12]. Application of the model to the entire dataset (4,765 tweets) produced a sentiment distribution in which negative sentiment dominated with 2,163 records (45.4%), followed by positive sentiment with 2,025 records (42.5%), and neutral sentiment with 577 records (12.1%).

The dominance of negative sentiment indicates that Indonesian public discourse on platform X regarding AI-generated art more frequently expresses criticism, concern, and rejection than appreciation. This pattern is consistent with the findings of Christopher [13], who reported a protective stance among the artist community toward their rights in response to developments in generative AI technology. Nevertheless, the relatively high proportion of positive sentiment (42.5%) suggests that public appreciation for this technology is also substantial, rendering the discourse bipolar in nature and not entirely dominated by a single pole.

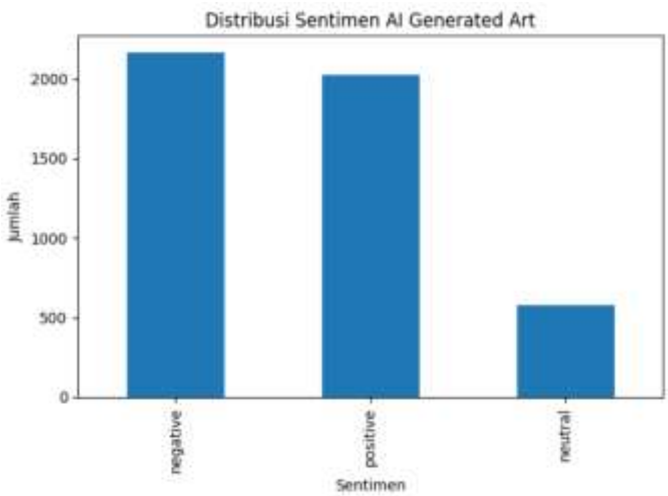

Figure 2. Sentiment Distribution

3.2. LDA Topic Modelling

Topic modeling was conducted separately on the negative sentiment group (2,163 documents) and the positive sentiment group (2,025 documents). Based on coherence score C_v testing for k = 2 through k = 9, the optimal number of topics obtained was k = 9 for both sentiment groups, with the highest C_v values of 0.6199 for negative sentiment and 0.5033 for positive sentiment. All 18 resulting topics were declared valid through Two-Stage Similarity validation [14]; none of the 36 topic pairs in either group simultaneously exceeded the Cosine Similarity threshold $\tau_1 \geq 0.70$ alongside JSD ≤ 0.50, confirming that all topics are semantically distinctive.

Table 1 presents the interpretations of the 9 negative sentiment topics along with their dominant keywords and confidence values.

Table 1. LDA Topics — Negative Sentiment

Topik	Interpretasi Topik	Keyword Dominan	Confidence
Topic 0	Dampak AI terhadap Konten Kreator (Impact of AI on Content Creators)	konten (0,140), generate (0,077), tantang (0,056), karakter (0.053), narasi (0,041), content (0,035), creator (0.034), training (0.034), dampak (0.033), ngikutin (0.030)	0,9428
Topic 1	Kekhawatiran Penyalahgunaan Data (Data Misuse Concerns)	dipake (0,114), kucing (0,072), lindung (0,055), nyomot (0,054), sekolah (0,049), keluarga (0.047), nonton (0.046), informasi (0,040), anggota (0.040), orisinal_plagiat (0,026)	0,5587
Topic 2	Kritik Etika AI (AI Ethics Critique)	manusia (0,158), teknologi (0,087), standar (0,083), prompt (0,053), tanggung (0,043), global (0,040), digital (0,040), bedain (0,037), nampak (0.032), kreativitas (0,028)	0,5493
Topic 3	Pencurian Desain (Design Theft)	sengaja (0,119), cantik (0,060), tiktok (0,049), kembang (0,049), design (0,046), selamat (0.043), visual (0,043), malaysia (0.041), pandang (0.041, nyolong (0,035)	0,3727
Topic 4	Kerugian Ilustrator (Impact to Ilustrator)	ilustrasi (0,333), terima (0.047), artist (0,044), manual (0,038), ngomongin (0.022), ngasih (0.021), jadiin (0.020), berani (0.020), ilustrator (0,019), gabisa (0,019)	0,8366
Topic 5	Kritik Kualitas Gambar AI (Critique of AI Image Quality)	tangan (0,142), bentuk (0,075), posting (0,053), untung (0,051), mustahil (0,047), amerika (0.047), perintah (0,046), tukang (0.039), upload (0,035), kemaren (0.034)	0,9189
Topic 6	Style Theft	artist (0,072), sumber (0,067), percaya (0,065), cerita (0.057), serius (0.053), ngomong (0.053), realistis (0,050), negatif (0,047), muncul (0.042), banyak (0.038)	0,7779
Topic 7	Keresahan Seniman (Artist Apprehension)	seniman (0,097), ngedit (0.082), generated (0,067), otomatis (0,067), bangun (0.066), postingan (0.063), desain (0,057), commis (0.042), tracing (0,041), editor (0.033)	0,7914
Topic 8	Penyalahgunaan Sistem AI (AI System Misuse)	generate (0,141), sistem (0,084), dijadiin (0.042), orisinal (0,040), artist (0.037), cerdas (0.034), ngirim (0.033), disalahgunakan (0,033), temenku (0.033), menang (0.033)	0,9368

The negative topics with the highest confidence scores are Topic 0 (Impact of AI on Content Creators, 0.9428) and Topic 8 (AI System Misuse, 0.9368), indicating that tweets within both topics form highly cohesive and thematically focused clusters. Conversely, Topic 3 (Design Theft, 0.3727) has the lowest confidence, indicating a mixture of contexts within this topic, which constitutes a noted limitation. Overall, the 9 negative topics represent multidimensional public rejection, spanning economic concerns (harm to illustrators, impact on creators), copyright violations (style theft, data misuse), and normative critiques of AI ethics and technical quality. Table 2 presents the interpretations of the 9 positive sentiment topics.

Table 2. LDA Topics — Positive Sentiment

Topik	Interpretasi Topik	Keyword Dominan	Confidence
Topic 0	Kemampuan AI (AI Capabilities)	manusia (0,098), upload (0,079), detail (0,053), interlink (0,053), ringkas_simpan (0.036), upgrade_benefit (0.034), global (0,029), deskripsi (0,025), prompt (0,024), nampak (0.020)	0,7570
Topic 1	Industri Kreatif Indonesia (Indonesian Creative Industry)	digital (0,041), indonesia (0,033), detail (0.032), ilustrasi (0,030), inspired (0.029), visual (0,027), konten (0,027), karakter (0.026), tangan (0.024), postingan (0.019)	0,8548
Topic 2	Perkembangan Teknologi AI (AI Technology Development	manusia (0,048), teknologi (0,046), aplikasi (0,037), visual (0,035), proses (0,034), kreatif (0,027), cantik (0,025), kembang (0.024), kualitas (0,023), ilustrasi (0.022)	0,5084
Topic 3	AI sebagai Alat Bantu Ilustrator (AI as a Tool for Illustrators)	ilustrasi (0,195), prompt (0,075), training (0,033), inspirasi (0,031), manual (0.027), otomatis (0.025), lengkap (0.025), terima (0.020), bingung (0.017), ilustrator (0,017)	0,7983
Topic 4	Analisis Konten Media (Media Content Analysis)	generated (0,070), analisis (0,038), sumber (0,031), konfirmasi (0,029), billboard (0.026), indonesia (0,024), bendera (0.019), verifikasi (0,019), berita (0,018), sumber_berita (0.017)	0,4351
Topic 5	Kemudahan Generate Konten (Ease of Content Generation)	generate (0,134), ilustrasi (0,050), konten (0,049), prompt (0,033), perintah (0.028), digital (0,024), khusus (0.021), produk (0,019), animasi (0,018), tonton (0.016)	0,6594

DOI: https://doi.org/10.31004/riggs.v5i2.11349

Topic 6	Aksesibilitas Gratis (Free Accessibility)	sumber (0,035), sungguh (0.027), wanita (0.026), midjourney (0,025), gratis (0,025), tiktok (0.024), realistis (0,024), generasi (0,022), detail (0,021), cerita (0.021)	0,8030
Topic 7	Kemudahan Tools AI (Ease of AI Tools)	bentuk (0,040), percaya (0,036), gampang (0,032), kucing (0.030), muncul (0.028), prompt (0,025), tunjuk (0.021), spesifik (0,019), menang (0,019), seniman (0,017),	0,7560
Topic 8	Referensi Desain dan Poster (Design and Poster References)	ilustrasi (0,103), cerita (0,052), referensi (0,036), poster (0,027), artist (0,023), watermark (0.022), desain (0,021), manfaat (0,019), mindmapping_sertifikat (0.018), leaflet_infografis (0.018)	0,7452

The positive topics with the highest confidence scores are Topic 1 (Indonesian Creative Industry, 0.8548) and Topic 6 (Free Accessibility, 0.8030). The average confidence of positive topics (≈0.70) is slightly lower than that of negative topics (≈0.74), indicating that positively-toned tweets form more distributed clusters with greater contextual diversity. This diversity of positive topics reflects that public support for AI-generated art stems from various motivations, ranging from appreciation of technological capabilities and ease of access to the utilization of AI in supporting the Indonesian creative industry. It is worth noting that Topic 4 (Media Content Analysis, confidence = 0.4351) represents a unique phenomenon characteristic of the Indonesian context, namely the use of AI for media content verification and fact-checking, which extends beyond mere aesthetics or copyright issues.

3.3. Knowledge Graph Construction

The Knowledge Graph was constructed from the integrated outputs of IndoBERT and LDA on the positive and negative sentiment groups. The final graph comprises 144 nodes and 329 edges, composed of 18 topic nodes (9 positive, 9 negative) and 126 keyword nodes. Two types of relationships were formed: MEMUAT_KATA (180 edges, connecting topics to keywords with LDA probability weights) and CO_OCCUR (149 edges after threshold filtering $\tau = \mu + \sigma = 0.106$, connecting pairs of keywords that frequently co-occur within the same document).

The keyword pairs with the highest co-occurrence (co_count) values are ilustrasi–cerita (co_count = 46), manusia–ilustrasi (co_count = 40), and ilustrasi–karakter (co_count = 34). The high co-occurrence values for pairs involving the word 'ilustrasi' (illustration) indicate that the AI-generated art narrative in Indonesia consistently associates the concept of illustration with elements of storytelling, human figures, and visual characters, both in the context of appreciation and concern. From a sentiment perspective, the resulting KG reveals that public discourse is bipolar yet asymmetric: negative topics are more cohesive and anchored to specific, identifiable concrete issues, while positive topics are more distributed across diverse AI utilization contexts.

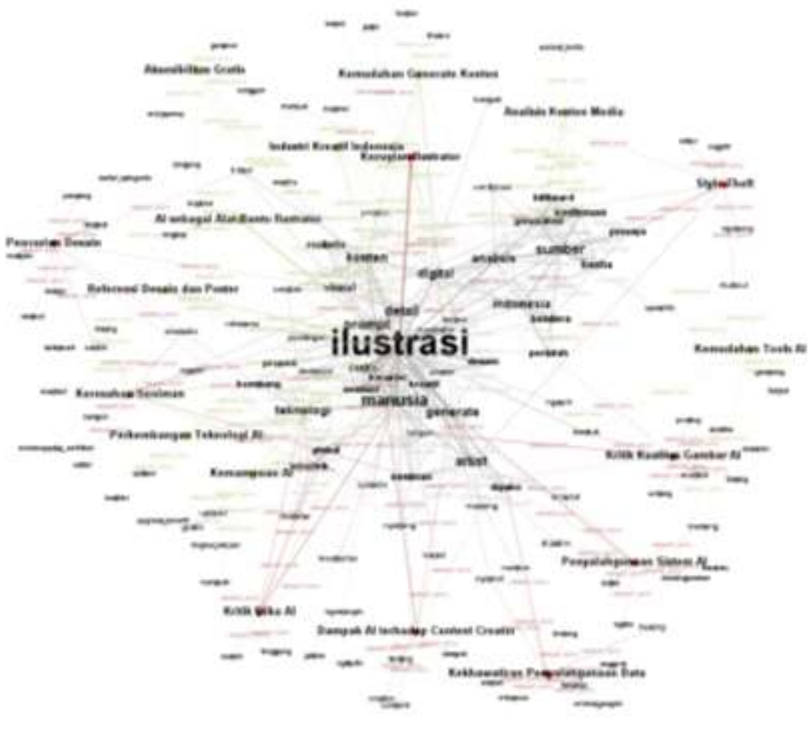

Figure 3. Knowledge Graph

3.4. Knowledge Graph Evaluation Using SNA

The KG structure evaluation using SNA metrics produced network characteristics as presented in Table 3.

Table 3. Karakteristik Knowledge Graph

Metriks	Nilai
Number of Nodes	144
Number of Edges	329
Graph Density	0,031954
Number of Communities	7
Modularity Score	0,4556

A graph density value of 0.031954 indicates that the constructed KG is sparse. This condition is commonly observed in graphs based on social media data; by way of comparison, Sadri et al. [15] reported a density value of 0.00007 for social interaction networks at the scale of tens of thousands of records, such that the value obtained in this study actually reflects a relatively favourable level of connectivity. The modularity score of 0.4556 falls within the range of 0.3 to 0.7, which is generally indicative of a significant community structure [16], demonstrating that entities within the KG have successfully clustered into meaningful thematic communities.

Community Detection. The Louvain algorithm detected 7 thematic communities in the KG. Table 4 presents the composition of each community along with its dominant topics.

Table 4. Community Detection Results

Communities	Number of Nodes	Dominant Topics
1	42	AI sebagai Alat Bantu Ilustrator, Keresahan Seniman, Kerugian Ilustrator, Referensi Desain
2	30	Kekhawatiran Penyalahgunaan Data. Kemudahan *Generate* Konten, Penyalahgunaan Sistem AI.
3	26	Aksesibilitas Gratis, Dampak AI terhadap Konten Kreator, Industri Kreatif Indonesia, Pencurian Desain.
4	18	Kemampuan AI, Kritik Etika AI, Perkembangan Teknologi AI.
5	15	Analisis Konten Media.
6	8	Kritik Kualitas Gambar AI.
7	5	*Style Theft.*

Community 1 is the largest community (42 nodes), uniting topics surrounding the role of AI in the world of illustration and its impact on the artistic profession. The coexistence of positive topics (AI as a Tool for Illustrators, Design References) alongside negative topics (Artists' Anxiety, Harm to Illustrators) within the same community demonstrates that both perspectives share the same semantic field, both centering on the domain of digital illustration as the core of discussion. Community 2 reveals an interesting relationship between the ease of using AI technology and concerns over training data misuse, indicating that ease of access and ethical risks are perceived as two inseparable sides. Community 5 represents a phenomenon characteristic of the Indonesian context, namely the issue of AI-based media content verification and fact-checking, which has developed as a distinct discourse, separate from discussions of art per se.

Centrality Analysis. Tables 5 and 6 present the centrality analysis results for topic nodes and keyword nodes with the highest values.

Table 5. Centrality Topic Node

***Node* Topik**	**Sentimen**	**WDC**	**Betweenness**	**Closeness**
Keresahan Seniman	Negatif	0,6147	0,0554	0,0254
Kritik Etika AI	Negatif	0,5999	0,0417	0,0311
Kerugian Ilustrator	Negatif	0,5831	0,0689	0,0354
Kritik Kualitas Gambar AI	Negatif	0,5686	0,0824	0,0284
Style Theft	Negatif	0,5439	0,0551	0,0252

Table 6. Centrality Keyword Node

Node Keyword	Sentiment	WDC	Betweeness	Closeness
ilustrasi	negative	12.620987	0.789323	0.039196
manusia	negative	4.659248	0.302866	0.038267

sumber	negative	3.97667	0.128041	0.035603
generate	negative	2.974021	0.147641	0.037631
artist	negative	2.242428	0.096425	0.036749
indonesia	negative	2.079187	0.024623	0.033869
generated	negative	2.003423	0.048459	0.036195
cerita	negative	1.974262	0.000000	0.037727
berita	negative	1.95165	0.000000	0.033299
analisis	negative	1.860419	0.000985	0.032771

The topic node analysis results demonstrate that all five topics with the highest Weighted Degree Centrality (WDC) belong exclusively to the negative sentiment category, with Artists' Anxiety ranking first (WDC = 0.6147). This finding indicates that negatively-nuanced issues occupy a more central position in the network structure, reflecting that expressions of concern from the Indonesian artistic community are more structured and thematically directed compared to positive topics. The topic Critique of AI Image Quality has the highest betweenness centrality among topic nodes (0.0824), signifying its role as a bridge between conceptual groups in the network.

At the keyword level, 'ilustrasi' dominates all metrics with WDC = 12.6210 and betweenness centrality = 0.7893, far surpassing all other nodes. This remarkably strong dominance confirms that 'ilustrasi' is the central semantic hub that serves as the common thread across nearly all discussions, in both positive and negative sentiment contexts. In other words, 'ilustrasi' is the primary lens through which Indonesian society interprets and discusses the phenomenon of AI-generated art. Several keywords, such as 'cerita' (story) and 'berita' (news), have betweenness centrality values of zero, indicating that these nodes function as supporting concepts relevant within specific clusters but do not serve as cross-cluster bridges.

Overall, the SNA evaluation results demonstrate that the constructed KG is capable of uncovering hidden relational patterns that cannot be obtained from conventional sentiment analysis alone. The integration of IndoBERT, LDA, and Knowledge Graph produces a representation of public discourse that not only shows what is being discussed, but also how these concepts are interrelated within a measurable semantic structure, in alignment with the role of Knowledge Graphs as a semantic knowledge layer that enriches sentiment analysis results, as argued by Vizcarra et al. [17].

4. Conclusion

This study successfully addressed both research questions posed. First, the IndoBERT model was demonstrated to classify the sentiment of Indonesian society on platform X toward AI-generated art with an accuracy of 69%, yielding a sentiment distribution dominated by negative polarity (45.4%) over positive (42.5%). Topic modelling using LDA produced 18 semantically distinctive and valid topics—9 negative and 9 positive—all verified through the Two-Stage Similarity procedure. The negative topics centred on the economic impact on the illustrator profession, copyright infringement, and artistic style theft, while the positive topics reflected appreciation for AI as a creative tool and ease of accessibility. Second, the Knowledge Graph constructed from the integrated outputs of IndoBERT and LDA successfully represented the interconnections among sentiments, topics, and keywords in a structured and comprehensive manner, with 144 nodes, 329 edges, and a modularity score of 0.4556, reflecting 7 meaningful thematic communities. From a theoretical standpoint, this study contributes an integrative methodological framework combining IndoBERT-based sentiment classification, LDA topic modelling per sentiment group, and Knowledge Graph construction within a single KDD-based analytical pipeline. This approach extends the capabilities of conventional sentiment analysis—which only yields polarity classifications—into a semantic knowledge representation capable of revealing why and within which topical context a particular sentiment is formed. The finding that the entity 'ilustrasi' dominates the entire network structure (weighted degree centrality = 12.621 and betweenness centrality = 0.789) provides empirical evidence that the domain of digital illustration constitutes the core of all Indonesian public discourse on AI-generated art, transcending mere technical debates about AI capabilities. Furthermore, the identification of a unique cluster related to media content verification as a distinctly Indonesian contextual phenomenon demonstrates that Knowledge Graphs can uncover hidden relational patterns that are inaccessible through singular sentiment analysis, consistent with the argument of Vizcarra et al. [17] that knowledge graph integration enriches the semantic layer of public opinion analysis. Looking ahead, several directions for future development may be considered. From a data perspective, expanding the volume of labelled data with the involvement of multiple annotators would improve the reliability of the classification model and reduce subjective bias. The scope of platforms could also be extended to Instagram or other digital community forums, accompanied by temporal analysis to map the evolution of sentiment in tandem with the development of generative AI technology. Regarding Knowledge Graph applications, future research may explore the use of KGs as a foundation for policy recommendation systems, semantic query mechanisms, or integration with large language models to generate more adaptive summaries of public discourse. The methodological framework produced in this study also holds potential for adaptation to other digital-social phenomena in Indonesia that require a deeper understanding of large-scale public opinion.

References

[1] N. Muhamad, “Indonesia, Penyumbang Kunjungan Aplikasi AI Terbanyak ke-3 di Dunia,” *Databoks Katadata*, Jan. 31, 2024. [Online]. Available: https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/a49ed3eb121983b/indonesia-penyumbang-kunjungan-aplikasi-ai-terbanyak-ke-3-di-dunia

[2] C. Ramadhan, V. Atina, and H. Permatasari, “Analisis Perbandingan Model CNN dan IndoBERT Dalam Sentimen Berita Politik Indonesia,” dalam *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB) 2025*, Surakarta, Indonesia, Jul. 2025, doi: 10.47701/v1r9ka69.

[3] D. Alfatah, “Application Of Transformer Model For Sentiment Detection On Indonesian Twitter Data,” *Jurnal Komputer*, vol. 2, no. 2, 2024, doi: 10.70963/jk.v2i2.99.

[4] S. S. Nabila, “Analisis Sentimen Publik terhadap Gambar Berbasis AI di Twitter Menggunakan Metode Naive Bayes,” preprint, 2025. [Online]. Available: https://www.researchgate.net/publication/391848610_Analisis_Sentimen_Publik_terhadap_Gambar_Berbasis_AI_di_Twitter_Menggunakan_Metode_Naive_Bayes/citation/download.

[5] Y. Sahria and D. H. Fudholi, “Pemodelan Pengetahuan Graph Database Untuk Jejaring Penelitian Kesehatan di Indonesia,” *Jurnal Media Informatika Budidarma*, vol. 4, no. 3, pp. 604–611, Jul. 2020, doi: 10.30865/mib.v4i3.2183.

[6] D. Prismala and I. K. D. Nuryana, “Analisis Opinion Mining pada Topik ChatGPT di Aplikasi X dengan Pendekatan Algoritma SVM Berbasis Lexicon,” *Journal of Informatics and Computer Science (JINACS)*, vol. 6, no. 2, pp. 479–492, Jul. 2024.

[7] U. Khairani, V. Mutiawani, and H. Ahmadian, “Pengaruh Tahapan Preprocessing Terhadap Model IndoBERT dan IndoBERTweet untuk Mendeteksi Emosi pada Komentar Akun Berita Instagram,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 4, pp. 887–894, Aug. 2024, doi: 10.25126/jtiik.1148315.

[8] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," in *Proc. 1st Conf. Asia-Pacific Chapter Assoc. Comput. Linguist. 10th Int. Joint Conf. Natural Language Process. (AACL-IJCNLP)*, Suzhou, China, Dec. 2020, pp. 843–857.

[9] W. X. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li, "Comparing Twitter and traditional media using topic models," in *Proc. 33rd Eur. Conf. Inf. Retrieval (ECIR 2011)*, Lecture Notes in Computer Science, vol. 6611, Dublin, Ireland, Apr. 2011, pp. 338–349.

[10] M. Á. Serrano, M. Boguñá, and A. Vespignani, "Extracting the multiscale backbone of complex weighted networks," *Proc. Natl. Acad. Sci. USA*, vol. 106, no. 16, pp. 6483–6488, Apr. 2009.

[11] N. Mamatha, S. S. Sunitha, and M. D. Shivakumar, "Graph theory and its role in social network analysis," *World J. Adv. Res. Rev.*, vol. 21, no. 2, pp. 2088–2093, 2024.

[12] V. E. Sidauruk and W. Herowati, "IndoBERT-based sentiment analysis of political discourse on platform X: The case of Prabowo-Gibran administration," *J. Appl. Inform. Comput. (JAIC)*, vol. 10, no. 1, pp. 673–683, 2024.

[13] B. Christopher, *Analisis Sentimen Penggunaan AI-Generated Artwork di Twitter Menggunakan IndoBERT*. Tangerang: Universitas Multimedia Nusantara, 2025.

[14] S. J. Blair, Y. Bi, and M. D. Mulvenna, "Aggregated topic models for increasing social media topic coherence," *Appl. Intell.*, vol. 50, no. 1, pp. 138–156, Jan. 2020.

[15] A. M. Sadri, S. Hasan, S. V. Ukkusuri, and J. E. S. Lopez, “Analyzing Social Interaction Networks from Twitter for Planned Special Events,” *arXiv preprint* arXiv:1704.02489, Apr. 2017.

[16] J. Xie and B. K. Szymanski, "Community detection using a neighborhood strength driven label propagation algorithm," in *Proc. 2011 IEEE Netw. Sci. Workshop (NSW)*, New York, NY, USA, Jun. 2011, pp. 188–195.

[17] J. Vizcarra, K. Kozaki, M. Torres-Ruiz, and R. Quintero, “Knowledge-Based Sentiment Analysis and Visualization on Social Networks,” *New Generation Computing*, vol. 38, no. 4, pp. 701–742, Oct. 2020.

DOI: https://doi.org/10.31004/riggs.v5i2.11349