Linguistic Fingerprints for Authorship Attribution in Human–AI Collaborative Texts

Devi Ambarwati Puspitasari[1], Dewi Nastiti Lestariningsih[2*], Bayu Permana Sukma[3], Yenny Karlina[4], Salimulloh Tegar Sanubarianto[5], Mu'awal Panji Handoko[6], Intan Pradita[7]

[1]The National Research and Innovation Agency, Jakarta Selatan 12710, Indonesia
[2]The National Research and Innovation Agency, Jakarta Selatan 12710, Indonesia
[3]PhD in Linguistics, Faculty of Humanities, Arts and Social Sciences, Lancaster University, Lancaster, LA1 4YW, United Kingdom
[4]The National Research and Innovation Agency, Jakarta Selatan 12710, Indonesia
[5]The National Research and Innovation Agency, Jakarta Selatan 12710, Indonesia
[6]The National Research and Innovation Agency, Jakarta Selatan 12710, Indonesia
[7]English Language Education, Faculty of Social and Cultural Sciences, Universitas Islam Indonesia, Yogyakarta 55584, Indonesia

ARTICLE INFO

Keywords:
authorship analysis
authorship attribution
Indonesian text
lexical choice
N-gram tracing

Article History:
Received : 22/10/2025
Revised : 24/04/2026
Accepted : 25/05/2026
Available Online: 30/05/2026

ABSTRACT

Authors leave distinctive linguistic traces that reflect their identity through consistent writing styles, particularly in morphosyntactic patterns and lexical choices. However, the increasing use of AI writing tools challenges authorship attribution because machine-generated text can imitate or obscure individual writing characteristics. This study investigates linguistic features that effectively identify authors and differentiate human-written from AI-generated texts. A corpus comprising 2,074,125 tokens and 63,414 word types was compiled from collaborative digital platforms, including instant messaging and social media. Lexical and stylistic features were extracted to develop hybrid authorship-classification models, while N-gram tracing was used to identify salient patterns. The findings demonstrate that lexical choice is the most reliable indicator for distinguishing human and AI-generated texts. Character-level N-gram analysis further demonstrates that authorship can be identified through delicate patterns involving letters, capitalization, punctuation, and other non-alphabetic characters. Diction appeared as the strongest factor in differentiating individual authors. These results enhance the reliability of authorship attribution methods and provide valuable insights for forensic investigations of digitally mediated communication involving human–AI interaction.

How to cite (in APA style): Puspitasari, D. A., Lestariningsih, D. N., Sukma, B. P., Karlina, Y., Sanubarianto, S. T., Handoko, M. P., & Pradita, I. (2026). Linguistic Fingerprints for Authorship Attribution in Human–AI Collaborative Texts. *OKARA: Jurnal Bahasa dan Sastra, 20*(1), 45–69. https://doi.org/10.19105/ojbs.v20i1.22271

1. INTRODUCTION

Machine-generated text is developing rapidly due to the rapid expansion of artificial intelligence (AI) and natural language processing (NLP). That improvement is centered on virtual assistants, automated customer service chatbots, and powerful language models

***Corresponding Author**: **Dewi Nastiti Lestariningsih** dewi041@brin.go.id

such as OpenAI's GPT and Google's BERT. Dependence on technology to perform daily activities is the driving factor behind all this improvement. On the other hand, a rare side effect of this enhancement is that there may no longer be a clear line between human- and machine-generated work (Alamleh et al., 2023; Habibzadeh, 2023). The fuzziness occurs mainly when human authors interact with or rephrase/refine the writing in a machine-generated (machine-counterpart) draft of a paper. This 2018 merger blurs the line of what makes us human. This distinction has kept humanity skeptical since its earliest days and returns to questions of creativity and authorship regarding whether machines may ever be capable of original speech, right up against the linguistic divide separating writing by man from creation by machine. One means of approaching this issue is authorship attribution, a domain within forensic linguistics (Belvisi et al., 2020).

Authors' attribution is the process of identifying the original author of a text, and it has been a fundamental research theme in linguistics/forensic linguistics for many years (Abreu & Sousa, 2022; Belvisi et al., 2020). Existing methods for authorship attribution mostly originate from the traditional writer assumption that a text is created by a lone human author or a cohort of authors. However, with the appearance of AI-generated text in development. Texts co-authored in mixed human-machine contexts — or, where an automated service generates much of the content, as well as controlled by rules set out by humans to proofread and tweak for quality reasons–are much more complicated, requiring a creative kind of attribution. Therefore, the chief research question for this study is: What are the best methods for attributing authorship to documents that contain a mixture of human- and machine-generated content?

Distinguishing between human- and AI-generated writing is important for preserving authorship authenticity, maintaining academic integrity, preventing the spread of misinformation, and promoting the transparent and responsible use of generative AI. As AI-generated content increasingly resembles human-written text, accurate detection methods are vital to promote trust, accountability, and ethical communication across journalism, education, and digital media (Katib et al., 2023; Yang et al., 2024). In the current environment, precise authorship attribution is no longer an academic question: it is central to issues of intellectual property, ethical writing practices, and digital communication itself.

To understand how to identify algorithmically generated text in mixed human-machine texts, this study also considers a secondary research question: What can be learned about distinguishing the linguistic features of human versus machine output in collaborative writing environments? This is the most important thing to understand to create models that can distinguish human-made works from machine-generated ones. Much attention has been placed on identifying how human-authored text differs (or is indeed similar) to machine-generated content but comparatively less focus has been given when these two types of writing begin to work symbiotically together.

Machine-generated text has certain stylistic and structural qualities that distinguish it from human writing. One challenge is that AI-generated content does not inherently include the contextual understanding, cultural references, and emotional complexity that characterize human communication, as AI systems lack lived experience, cultural embeddedness, and genuine emotional understanding. On the contrary, human writing is usually more complex in terms of sentence structure; it pays attention to content, allowing humans to interact freely and naturally with one another (Dou et al., 2022; Knowles, 2024). However, over time, as AI has advanced, the differences are fading, making it difficult to disguise them.

Authorship attribution has been a major theme in forensic linguistics, computational linguistics, and stylometry for a long time. Classical studies have made significant use of linguistic measures such as lexical selections, syntactic constructs, function words, and stylistic standards to find some kind of fingerprint in the way an author writes. These include Support Vector Machines (SVMs), Random Forests, and Naïve Bayes classifiers, which achieve high levels of accuracy in class separation by using linguistic features as predictors of human authorship (Aribowo et al., 2021; Machová et al., 2023). However, the landscape of authorship changed fundamentally with the rapid advancement in generative AI (artificial intelligence). Recent research has focused on detecting AI-generated texts and distinguishing them from human-generated texts using linguistic and computational features (Chen et al., 2024; Fedotova et al., 2022). While these works have produced encouraging results so far, they mostly separate human and AI texts into two different classes. Existing datasets nearly all contain either fully human-written or fully machine-generated texts, creating a methodological blind spot regarding authorship in co-writing scenarios and rendering each writing context unconducive to natural language processing (NLP) studies.

The rise of hybrid writing, which includes human-edited AI-generated content, artificial intelligence-assisted academic writing, and collaborative human-machine social media production, poses new challenges for authorship attribution. With Mixed Texts, we cannot ask attribution models how much credit to give humans and AIs. No study, however, has examined how linguistic features such as syntactic complexity, lexical selection, stylistic markers, and contextual coherence change when the authorship includes both humans and artificial intelligence. Thus, there remains a significant gap for more advanced attribution frameworks that can identify authorship processes in behavioral mixed human-machine text.

This investigation offers several critical advancements that distinguish it from earlier studies. This study includes the creation of a new, mixed human-machine corpus comprising collaboratively produced texts, AI-edited human writing, and human-edited AI-generated content. This corpus represents writing in the real world, where human and AI texts are interspersed (unlike existing corpora that are purely human or purely AI). Second, the framework establishes a multidimensional linguistic metric that simultaneously assesses four significant dimensions of authorship: syntactic complexity, lexical choice, stylistic markers, and contextual coherence. Having previously explored isolated linguistic features, this integrated framework offers a broader perspective on how authorship signals are retained, altered, or hidden across aspects of human-AI collaboration. Third, the study employs a hybrid methodology combining linguistic analysis and explainable machine learning techniques, specifically SVM-based attribution models, to quantitatively identify the top features that characterize human, AI, and mixed authorship. Besides classification accuracy, this research also aims to identify which linguistic features are strong predictors of attribution decisions, thereby promoting model transparency and interpretability. Lastly, beyond its forensic identification applications, this work broadens authorship attribution to detect digital fraud and misinformation. It serves as an alternative/automated approach to resolving intellectual property disputes or as an AI ethics code. It is among the first studies to investigate authorship attribution as a mechanism for recognizing varying levels of human and AI contribution within collaborative text production, thereby providing a foundation for more transparent and accountable human-AI interactions. Based on the identified research gaps, the study seeks to answer the following research questions:

1. What linguistic characteristics distinguish human-written, AI-generated, and human–AI collaborative texts in terms of lexical characteristics, stylistic characteristics, and character-level linguistic patterns?
2. Which linguistic features are the most salient indicators for distinguishing authorship across human, AI, and collaborative human-AI texts?
3. How do character-level N-gram patterns reveal persistent linguistic fingerprints across human-written, AI-generated, and collaborative human–AI texts?
4. What implications do the findings have for forensic linguistics, intellectual property protection, misinformation detection, and ethical human-AI collaboration?

2. LITERATURE REVIEW

In a similar vein to the increase in 'trust' placed in artificial intelligence (AI) and machine learning, this has led to many facets of human-computer interaction changing. The problem of authorship attribution, a staple of computational linguistics, has long been used to identify the authors of human-produced texts. But AI in content creation adds layers of complexity that haven't yet been unpacked. This review discusses current research efforts in authorship attribution, particularly in the sub-discipline of distinguishing between human- and machine-generated texts, their linguistic properties, and the methods employed to perform this type of analysis.

2.1 The History and Tradition of Authorship Attribution

The first rigorous scholarship on this issue can be attributed to studies in literary research, modern forensic linguistics, and computational linguistics. This was traditionally accomplished through stylometric analysis -- identifying attributes of an author's unique writing style (e.g., diction or syntax) by scrutinizing quantifiable traits such as word frequencies, sentence structures, and syntactic patterns. This contrasts with surveys such as Holmes (1998), which describe the statistical techniques used in authorship attribution in general, or Stamatatos (2009), which also includes an overview of the development of computational methods for this task. Finally, many existing methods for authorship attribution hardly address the significant development of participants, particularly those focused on the tool's efficacy, at least when it comes to human text. Although if AI does end up generating more content, it will add another level of new and contemporary problems. Existing techniques must therefore be re-evaluated to reflect the attributes of AI-authored text and to separate human-generated from machine-generated text.

2.2 Characteristics and Problems in Automatically Generated Text

With each iteration of language models, such as GPT, BERT, and T5, the text generated very closely resembles that produced by humans like you and me. Researchers tried to assess the artificial text styles compared to human-written content. For example, Ippolito et al. (2020) and Dugan et al. (2023), looking at text generated by neural language models, found that while machine-generated text can superficially resemble human writing, it lacks the semantic depth and contextual awareness typical of human-authored text.

A further study of detecting machine-generated text, especially when models are fine-tuned to write in the style of another model. These authors suggested that in broadcasts and other works with substantial amounts of this kind of material, the standard stylometric

features (word frequencies, sentence lengths, etc.) will not, by themselves, be sufficient to distinguish between human-written and AI-generated content. This shows just how crucial it is to come up with something more in the middle that can actually identify the differences between such texts without resorting to creative dialogue tactics and whatnot.

2.3 Hybrid Texts

Hybrid texts, such as those emerging when human authors co-write with AI systems, further complicate authorship attribution. In such scenarios, determining authorship is not merely about detecting the authorship of a text but also about assessing the level of engagement between human and machine. It is especially important in areas like journalism, creative writing, and academic publishing, where AI tools are used more liberally for content augmentation, but not in totality.

They were particularly interested in how AI-assisted writing might affect academic publishing (Khalifa & Albadawy, 2023), as the use of AI tools may complicate questions about authorship and raise concerns about originality and copyright. Caliskan et al. (2017) advised that the research further supported these themes, and so too did Bender et al. (2021), who introduced the issue of automatic content generation with AI and how such systems may exhibit biases or ethical problems, also when employed to generate text or used as co-authors. Our results suggest a demand for explicit guidelines and attribution methods to detect contributions from both humans and machines.

Koppel et al. (2011) have examined authorship attribution in collaborative environments, and work on collaboration detection is becoming increasingly widespread. In one of the many texts, Baker et al. (2010) constructed models that could distinguish among several authors based on their stylistic signatures. At that time, the methods they relied on were primarily for human-authored texts. However, their work paved the way to expand such methods to hybrid text scenarios that include machine-generated narrative. These models will have to be modified not only to account for the strange recursive linguistics of AI systems, but also because the patterns they are trained on may not map cleanly or at all to the styles that readers associate with human authors.

2.4 Automatic Targeted Extraction for Authorship Attribution of Mixed Texts with Machine Learning

Machine learning (ML) methods have increasingly been applied in authorship analysis as computational technology advances. Indeed, ML techniques — and neural networks in particular – have demonstrated potential to detect subtle patterns in text that are impossible to capture through conventional analysis. For instance, Stamatatos et al. (2018) showed that deep learning models could perform very well in authorship attribution when given access to labeled written genres. However, the problem of using these models with human-machine mixed texts persists. Recent work, for example, the studies by Hovy and Yang (2021), has begun to investigate transformer models in authorship attribution due to their capacity to infer context and style within text. These studies show that, while transformer models can perform effectively in authorship identification for human-authored texts, they require additional training and feature engineering to handle mixed texts.

Another potential approach is to use ensemble methods, which combine multiple models to improve classification accuracy. Several authorship attribution studies used ensemble methods, including Kestemont (2018) and Bevendorff et al. (2020), which

combined stylometric and content-based features to increase attribution certainty. To address this, applying these methods to hybrid texts may make authorship attribution more reliable in situations where human and AI contributions are mingled.

3. METHOD

This study utilized a corpus linguistics approach to generate a comprehensive statistical representation. Although many corpora are now available worldwide, there is still a lack of a specialized corpus of personal text in Indonesian. The corpus we can obtain for Indonesia is Bahasa Indonesia, which is a very unique language for tracing authorship attribution in a linguistic profile. Multicultural and multilingual components heavily impact it. As a result, a single speech may exhibit a significant degree of linguistic variety as well as the mixing and code-switching of more than two or three languages. In this research, we construct a dataset of Indonesian personal text (Lestariningsih et al, 2025; Puspitasari et al., 2025).

3.1 Corpora

To conduct the research, we first built an extensive collection of Indonesian personal texts, primarily mixed-content data from collaborative platforms where users interact with or post-edit AI-generated text. The language data provided by CILLCO Indonesian Corpora (BRIN, 2024), a big language data platform and corpus tools analysis provided by the National Research and Innovation Agency (Pradita et al, 2026). The corpus comprises aggregated publicly available chat logs, such as online discussion forums and social media posts, where AI-assisted writing is typical. The final human-written corpus contained 2,074,125 tokens and 63,414 unique word types.

The AI text corpus was established using a stringent, multi-level assembly and validation process. The AI-generated texts were sourced from freely accessible domains: Reddit, ShareGPT, X (formerly Twitter), Facebook, LinkedIn, and Hugging Face datasets, with clear author mentions that their work was generated or heavily assisted by AI systems, as well as from medium publications on Medium and Substack. Using several search queries with keywords such as 'generated by ChatGPT,' 'AI-written,' and 'Claude output,' we kept only those texts for which it was clear they were written by AI. Data were cleaned by filtering duplicates, incomplete outputs, spam texts, non-target-language texts, and short documents. Subsequently, the data was preprocessed as follows: all texts were converted to plain text (removing HTML, deleting hyperlinks/emoji/punctuation marks), and tokenization and lemmatization were performed using NLP tools. Trained experts provided independent annotations to verify corpus quality. The final corpus consisted of 2,074,125 tokens and 63,414 unique word types, confirmed in AntConc and Python. Finally, we performed automated extraction and expert annotation of lexical richness, readability, syntactic complexity, and cohesion to derive features that facilitate comparisons between human- and AI-written texts and to train reliable authorship attribution models for Indonesian texts.

3.2 Research Procedures and Data Analysis

The research process was carefully structured to achieve the objective of accurately attributing authorship in mixed texts containing both human and machine-generated

content. The methodology was divided into three main phases (Alamleh et al., 2023; Singh et al., 2024): feature identification and extraction, model development, and N-gram tracing for text analysis, as shown in Figure 1. Each phase was crucial in building a robust framework for distinguishing between human and machine-generated text in Indonesian.

3.2.1 Phase 1: Feature Identification and Extraction

Once we had the dataset ready, we proceeded to the next step: extraction and identification, using specific linguistic features to categorize generated text as Machine-translated Text or Human-translated Text. This included an introspection of lexical choices (vocabulary frequency and diversity) displayed in Figure 1. Text written by humans has a richer lexicon and subtlety in word choice (Theophilo et al., 2023), which signals a higher ability to understand context and culture. Conversely, content produced by machines will have a repetitive structure and a less diverse vocabulary (Moneus & Sahari, 2024; Uchendu, 2020).

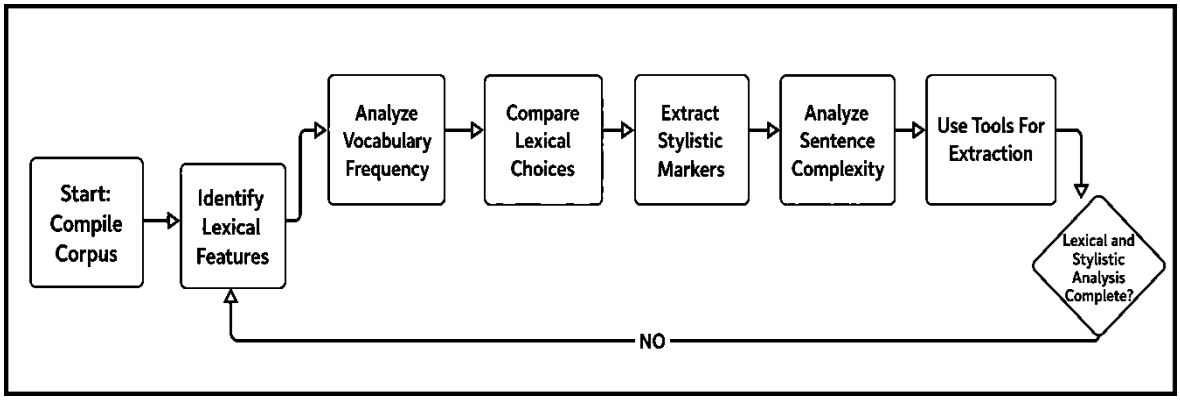

Fig.1. Phase 1 Process

The analysis explores not only lexical features but also other stylistic indicators, including sentence length, syntactic constructions, and word distribution. The human-written text demonstrates greater variation in sentence length and syntactic structure, as well as a more conversational tone. However, this characteristic should not necessarily be interpreted as an indicator of superior writing quality. Most machine-generated text from AI models would follow certain patterns or sound more authentic, making it sound formula-generated. We then systematically identified these stylistic markers, hand-annotated a portion of the data for each, and developed tools that we will describe for automatically annotating the rest of our training corpus.

3.2.2 Phase 2: Model Development and Training

Once the features were processed and extracted, we further refined them to develop and train hybrid models that could classify and attribute authorship using the input set of features. In this work, we used a set of machine learning methods (Alamleh et al., 2023; Fedotova et al., 2022), such as Support Vector Machines (SVMs) and artificial neural networks, to craft models that can differentiate between human-written and automatically generated text at the input level. The schemes are shown in Figure 2. This hybrid approach provided an avenue to employ the strengths of various algorithms, which ultimately enhanced the classification accuracy.

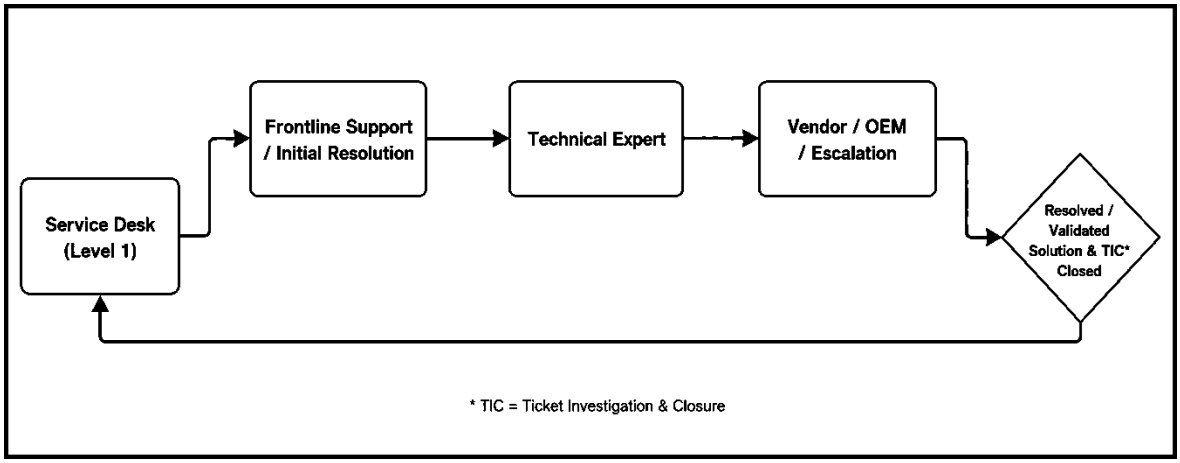

Fig. 2. Phase 2 Process

The models were trained using labeled data from the corpus (which contained examples of human-authored and machine-generated texts). Our solution was an iterative model of training, tuning parameters, and increasing the ability to categorize mixed texts correctly. In texts overlaid with human and AI inputs, hybrid models have become a valuable tool for handling these complexities. Subsequently, the interpretation of authorship attribution in this context became more refined.

3.2.3 Phase 3: N-gram Tracing for Text Analysis

N-gram tracing is an advanced text analysis method that was the final phase of the methodology used to extract features. To analyze the patterns in word usage and sentence structure, we utilized N-grams, which are adjacent sequences of n items from a given text as shown in Fig. 3 Unigrams, bigrams, trigrams, and higher-order N-grams allowed us to track a distribution of words and phrases that could differentiate between human-machine text (Grieve et al., 2019; Belvisi et al., 2020).

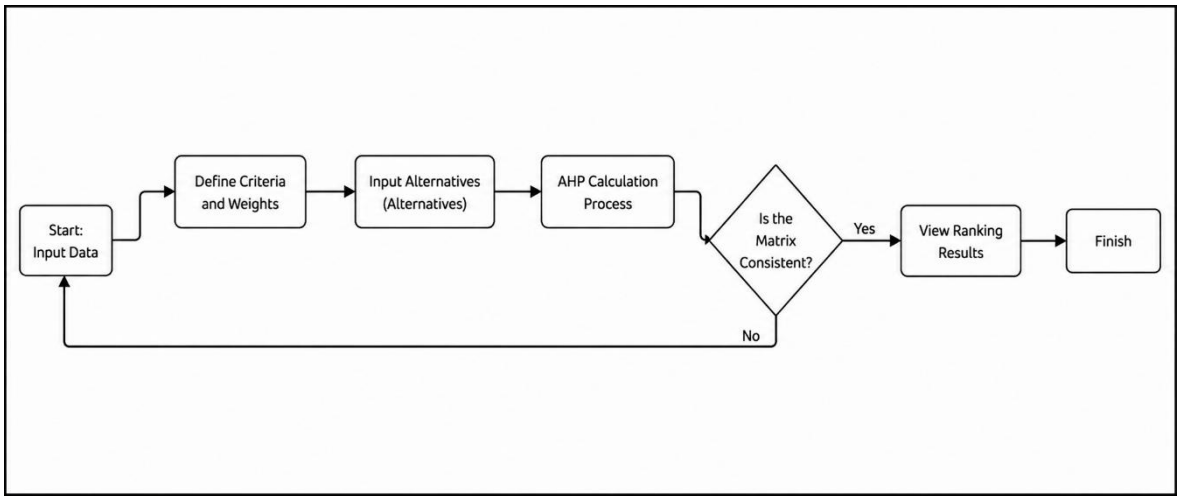

Fig. 3. Phase 3 Process

N-gram analysis was particularly successful at revealing recurring patterns unique to automatically generated text, such as revisited phrases and frequent syntactic constructions not typically found in human-generated content. Second, N-gram tracing supports the

detection of stylistic markers that are less salient in lexical analysis alone (Banga et al., 2018; Santos et al., 2021), enabling a more in-depth understanding of the differences between human and AI writing. The results of N-gram analysis were combined with the features obtained in previous steps, thereby improving the hybrid models' capacity for accurate authorship determination in mix-authored texts. The research methodology and data analysis aimed at rigorously distinguishing human from machine-generated Indonesian text, combining corpus construction, feature extraction, model development, and N-gram traces. Our broad experimental strategy led us to define successful hybrid models for authorship attribution, offering new ways to understand the distinctions between human- and machine-generated content in collaborative writing.

4. RESULTS

The corpus analysis systematically indicated a large disparity between human- and machine-generated texts in the linguistic features used and word selection. Finally, word choice emerged as the most significant feature by which authentic human-generated text could be distinguished from machine-generated text. This was especially evident when examining the smallest unit of text (characters) using different N-gram analyses. The results show the relevance of lexical characteristics and how character-level variations, even subtle ones, significantly contribute to authorship identification.

4.1 Linguistic Characteristics of Human-Written, AI-Generated, and Human–AI Collaborative Texts

This study first aimed to identify the linguistic features that differentiate human, AI, and human–AI collaborative writing. This analysis asks whether particular linguistic fingerprints can still be observed across text production modes, rather than suggesting or assuming that these text categories differ only in overall writing quality. The comparison centers on three complementary linguistic dimensions that vary meaningfully within the corpus: lexical characteristics, stylistic characteristics, and character-level linguistic patterns. Together, these dimensions provide the concrete evidence of what human writing styles do—and how they result even with AI assistance. Evidence presented in the following sections shows that many differences between human and AI writing are more than isolated deviations in usage, but systematic regularities at the levels of lexical selection, stylistic preferences, and orthographic behavior that can combine to discriminate authorship accurately.

4.1.1 Lexical Characteristics

According to the analysis results, human-authored texts differ from machine-generated texts in that words are a major distinguishing feature. The human authors have a wider and more nuanced vocabulary representative of a deeper comprehension and understanding of the content. This variety of wording primarily serves a stylistic function; it can depend on the author (who also belongs to an ideological community), their cultural capital, and the subject being written about. In comparison, the machine-generated texts sound coherent and grammatically correct. However, they frequently repeat certain words or phrases, as this model works based on statistical patterns rather than understanding of content.

For instance, idioms, colloquialisms, and domain-specific lexicons were observed more frequently in the human-authored texts of the corpora than in machine-generated text. This is challenging for AI models to replicate because they need to interpret cultural and contextual subtleties that go far beyond mere word prediction. In contrast, machine-generated texts often employed much more generic terms and phrases, which hinted at the model's preference for words or phrases with higher statistical probability rather than what would be more appropriate for the context.

Personal pronouns and their patterns are the main marker traits of authorship used to differentiate human from machine-generated writing. Personal pronouns in Indonesian are N-units that examine how writers express social identities, hierarchies, or certain social-action standing in their writing (MacLeod & Grant, 2012), and this identity is who each writer identifies with. Authorship attributes can also be detected based on the pronoun variant(s) a writer employs, in addition to our choice of first- and second-person pronoun "pairs". My previous algorithmic approach revealed this; e.g., there are five such instances in which *aku* [I, me, my] shows up in tokenizations. It contains five slight variations: *aku, qu, ak, Aku*, and *q*. However, every writer uses a single variant throughout. Character-level N-unit-based findings related to writer uniqueness can serve as markers for authorship attribution (McMenamin, 2002). These, as well as the type and number of characters in the text, are also choices of the author and form part of his writing style. A writer's style is composed of a unique combination of grammatical patterns that are commonly seen in most or all of the writer's writings, owing to repeated use or habits on the part of the author (Beyer, 2014). Correcting Word Choice: At the character level, you also correct word choice as students choose words carefully and precisely as they write. The corpus of this research data indicates that the word '*aku*' is written in five variants, the word '*saya*' is written in three variants, and the word '*gue*' is written in four variants, as shown below.

Table 1
Word Variant of Personal Pronoun of Human-authored Text

Word Choice	N-gram	Token Variant	Found in… Text Set/s	Frequency (in Corpora)
aku	n-1	*q*	5	176
[I/me/my]	n-2	*ak*	5	9
		aq	1	7
	n-3	*aku*	14	145
		aqu	1	6
saya	n-2	*sy*	2	5
[I/me/my]	n-3	*sya*	1	8
	n-4	*saya*	19	52
gue	n-2	*gw*	2	4
[I/me/my]	n-3	*gue*	6	22
		gua	3	16
	n-5	*guweh*	1	3

According to Table 1, the word saya is a popular choice, appearing in 19 texts and totaling 52 tokens. In turn, 19 authors always wrote *saya,* thus indicating that this is a systematic stylistic option. Another example is around the word *aku*, which is another way of saying *saya* or *aku,* and it appears in 14 texts, so 14 authors have this in common. These patterns identified specific words that multiple authors liked, indicating stylistic elements they all seem to prefer. The repeated use of identifiable vocabulary by unrelated authors suggests a common linguistic style.

The study found that human authors use specific words that reflect their individual, social, and cultural backgrounds, while AI-generated texts often employ more formal, standardized language. Lexical variation of this kind makes word choice a useful signal for authorship attribution. Moreover, analyzing word- and character-level n-grams improves the detection of unique writing patterns in human- and AI-generated texts.

By contrast, high-unique words such as *aq, aqu, sya*, and *guweh* appear in only one text set, suggesting they may belong to an author with a very personal writing style. Non-standard or colloquial forms of personal pronouns are typically found in informal speech or within certain social groups. While the rarity of these is exemplified by their presence only once across text sets, it is indicative of a distribution that separates authors who are distinguishable by their individual language choices. This specificity in terminology and its accompanying style of writing can provide a unique linguistic signature that helps detect or attribute authorship.

Stylometric analysts can then examine the frequency of lexical variants in texts produced by a single author or compare their distribution across texts written by different authors. Furthermore, authors may exhibit distinctive lexical preferences reflecting their regional linguistic background or social context. For instance, the words *ane, I, gue*, and *ulun* are found in this research corpus. The writer from Jakarta used the word *ane* (and also *gue*). The word chosen by a Chinese Indonesian author from Banjarmasin (a city on Borneo Island) is *ulun*. To be more precise, each of these words is unique in the corpus, and for *gue*, only 6 out of 273 text sets contained the word. Since most of the data sources in this study are from Jakarta and surrounding areas, it is common for them to use *gue*. For the six authors with *gue,* determining which identifier is still needed requires further investigation to confirm their uniqueness. In contrast, for the authors using *ane, I,* and *ulun*, there is straightforward proof of their uniqueness.

This is in stark contrast to the standardized language style you generally find in machine-generated text, which cannot capture the personal idiosyncrasies that characterize human authorship. The texts generated by machines typically use a more technical, consistent language, as machine learning's statistical nature prefers common constructs that are broadly accepted. On the other hand, human writing is filled with informal or playful nuances characterized by an individual author and their setting (background, social environment, and culture). This variation already makes human writing different from machine writing, which further suggests that vocabulary and stylistic characteristics are the most important tokens of context.

4.1.2 Stylistic Characteristics

In addition to lexical diversity, stylistic features appeared as another key aspect that differentiated human-written from AI-generated and human–AI collaborative texts. Though lexical choice illustrates an author's vocabulary preferences, stylistic markers demonstrate

habitual writing behaviors—often produced unconsciously—and as such make for reliable linguistic fingerprints in authorship attribution. The corpus analysis pointed out that capitalization, yawning gaps, writing punctuation, orthographic vestiges, and spirit character sequences differentiation, mediating mostly human interventions, were what accounted for the differences in its chi-squared form tracking between texts written by humans and those transformed by embedding transformation to AI. Such stylistic features are a combination of personal writing style, intention, emotion, and social interaction, and would still be difficult to consistently reproduce by AI-driven systems.

The most striking stylistic variation between them was the use of uppercase letters. The texts created by humans showed great flexibility in capitalization, sometimes completely ignoring conventional grammatical rules to express intensity, emotion, or meaning (pragmatics). In contrast, AI-generated texts usually produce grammatically well-formed sentences written in a standardized manner with comparatively less stylistic variety, e.g., by using grammar function words such as preceder, subject wrapper. Table 2 shows representative examples extracted from the Indonesian digital communication corpus.

Table 2.
Representative capitalization patterns in human-written and AI-generated texts

Text Type	Representative Example	Stylistic Function
Human-written	*JAGA DIRIMU BAIK-BAIK, ANAK* (Please take care of yourself, son)	Full capitalization for emphasis and emotional intensity
Human-written	*ya ntar kl gw ud dikirimin gw kabari* (Yes, once they have sent it to me, I will let you know)	Deliberate lowercase and informal orthography reflecting conversational style
AI-generated	*Baik pak, sudah saya kirim dengan faktur pajaknya.* (Dear sir, I have sent it, along with the tax invoice.)	Standard capitalization following formal grammatical conventions

Results indicate that human writers use capitalization and punctuation not only to fulfill grammatical roles but also to express emotion, emphasize, convey urgency, and take an interpersonal stance. The informal characteristics of repeated punctuation, ellipses, and clumps of unconventional symbols generate recognizable stylistic fingerprints. In contrast, AI-generated texts follow standardized writing conventions, yielding precise but less expressive punctuation and capitalization patterns (e.g., higher grammatical consistency), making these features valuable candidates for authorship attribution.

In addition, the corpus displayed significant orthographic variation in human-authored text, through expressive character sequences such as *wkwk, wkwkwkk* for laughter, or *haha* and *heuheu*. This set of variants encapsulates a socially and culturally embedded manner of expressing laughter, emotion, and interpersonal solidarity in Indonesian digital communication. Authors differed in the frequency, length, and structure of their references – suggesting that they are highly idiosyncratic stylistic markers rather than random orthographic slips. These patterns were less typical of the part of the text that a computer would write, which tended to favor more standardized lexical and orthographic forms. Overall, these results indicate that stylistic markers are more than simple formatting choices, but rather long-standing patterns of behavior which can adequately discriminate between human, AI, and human–AI claims in text forms; and therefore support their continued value as indicators for forensic authorship attribution.

4.1.3 Character-Level Linguistic Patterns

Character-level N-gram analysis was performed to discriminate human from AI-generated texts beyond lexical criteria. The results showed that human writing exhibits greater variation in alphabetic and non-alphabetic characters, capitalization, punctuation, numerals, symbols, and emojis, indicating richer stylistic expression. The study examined N-grams from n-1 to n-4 and found consistent punctuation and character sequences that were unique to each author. These micro-level habits produce consistent linguistic signatures, enhancing the accuracy of authorship attribution for human, AI-generated, and collaborative texts.

The decrease in the number of characters with respect to character usage characteristics depends on the smallest n-unit alphabet range: from n-1 to n-4 (Table 2). A writer will use *y* to mean *iya* [yes], *q* [I/my/me] to indicate anything in the 1st person, and *g* to mean *enggak* [no]. Another common thing in n-2 is that we often find an author replacing two characters that represent a word; for example, the *ya* character instead of *iya* [yes], and the character *ga* rather than *enggak* [no]. This is shown in Table 3.

Table 3
Examples of N-Units of Non-Alphabet Character-Level

Token	n-gram	Text
..	n-2	**Data 1:** Teks 1: *dia berlari dari lorong sampai di cegat pramugari karna sudah tidak bisa* ***turun.*** [Text 1: he ran from the hallway until a flight attendant intercepted him because he couldn't **get off** anymore.] Teks 2: *dan semakin* ***sepi..*** *pesawat sudah sangat* ***siap..*** *tapi ebel belum* ***datang..*** [Text 2: and it is getting **quieter..** the plane is very **ready..** but Ebel has not **arrived** yet**..**]
!!	n-2	**Data 2:** Teks 1: *Iya* ***dong!!*** [Text 1: Yes, **please!**!] Teks 2: *Gw wisuda 2013* ***anzeeengg!!*** [Text 2: I graduated in 2013, **anzeeegg!!]**

To illustrate the structure of the character-based n-gram analysis, **Table 4** presents several representative examples of n-units consisting of one- and two-character sequences extracted from the corpus. For each token, the table indicates its corresponding n-gram category, the number of text sets in which it appears, and its overall frequency across the entire corpus. These examples demonstrate how individual characters (n-1) and adjacent character pairs (n-2) are distributed within the dataset, providing the basis for subsequent analyses of stylistic and authorship-related linguistic patterns.

Table 4.
Examples of N-Units of the alphabet that consist of 1-2 characters

Token	n-gram	Found in…Text Set/s	Frequency (in Corpora)
q	n-1	5	176
aq	n-2	1	7
g	n-1	2	3
ga	n-2	17	136
y	n-1	1	4
ya	n-2	33	172

Table 4 presents examples of alphabetic *n*-units consisting of one- and two-character sequences extracted from the corpus. The table shows each token, its corresponding *n*-gram level, the number of text sets in which it appears, and its overall frequency across the corpora. Single-character (*n*-1) and two-character (*n*-2) units exhibit different distribution patterns, with some tokens, such as **"*ya*"** and **"*ga*",** occurring across many text sets and with relatively high frequencies. These findings indicate that certain short character sequences are consistently used throughout the corpus and provide useful low-level linguistic features for subsequent authorship classification.

The data also revealed that authors tend to have rich writing styles with a large number of n-units, namely n-3 and more. The n-units are punctuation characters, letters of the alphabet, or a combination of both. Based on the data, authors tend to use the letters 'w' and 'k' to express laughter, but the number of characters remains consistent. Differently, some authors tend to use a combination of letters 'h' and 'a' or 'e', sometimes followed by 'u' for Sundanese authors, as shown in Table 5.

Table 5.
Examples of Character N-units

Token	n-gram	Frequency in Text Set	Frequency in Corpora
wkk	n-3	4	12
wkwk	n-4	3	15
wkwkwkk	n-7	3	9
wkwkwkwkk	n-9	2	1
haha	n-4	4	13
haha..	n-6	3	7
heuheu	n-6	4	2
heuheu..	n-8	2	2

Automatically generated texts, on the other hand, were found to have more regular structures using a high percentage of alphabetic characters and a low percentage of digits or symbols. This leads to a pattern in which machine-generated text is narrow and less context-sensitive than human-generated text, tailored to the platform, audience, or communication purpose. The pattern of capitalization was also associated with the writers' identities. Human writers would make use of capitalization to function, for instance, exactly where a word has to be given emphasis; within an informal or uncommon context, and in

wide-ranging text.IM Options included all-caps (for shouting or emphasis) and forced lowercase in cases where standard capitalization might likely be expected, as a stylistic decision.

Table 6.
Examples of differences between human-written and machine-generated texts

Data	Author	Text
Data 3	human	*JAGA DIRI**MU** BAIK-BAIK, ANAK* *ISTERI**MU** JAGA BAIK-BAIK, PELAN TAPI PASTI, MAUT AKAN MENUNGGU**MU*** [TAKE CARE OF YOURSELF, CHILD TAKE CARE OF YOUR WIFE, SLOWLY BUT SURELY, DEATH WILL AWAIT YOU]
Data 4	machine	*Baik **pak**, sudah saya kirim dgn faktur pajaknya.* *Maaf **pak**, apa tidak apa-apa **saya** lampirkan satu dulu?* [Ok, sir, I have sent it with the tax invoice. Sorry, sir, is it ok if I attach one first?]
Data 5	human	*penawaran dari subcon tadi **aku** ada titip di yono,trus katanya ditaruh di meja**mu*** [I left the offer from Subcon at Yono, and he said that he put it on your table]
Data 6	human	*ya ntar kl **gw** ud dikirimin **gw** kabari.* ***lu** tunggu aj pe urusan tender kelar, fokus situ aj.* [Yes, when I have sent it, I will let you know Just wait for the tender to finish, just focus on that]

There was more variety and randomness in capitalization and punctuation between human and AI-generated texts. Writers also abandoned standard capitalization rules for emphasis, emotion, or stylized effect, which AI did not do well, as it always followed grammar rules. Likewise, human writers used ellipses and dashes–along with repeated punctuation and exclamations in informal settings–to indicate tone, pausing, or emotional intensity. On the other hand, texts generated by AI exhibited rather stereotyped and homogenous punctuation rules; though grammatically correct, they were much less original in their writing style. Such disparities in style provide important features for distinguishing human from machine-generated authorship.

4.2 Salience of Linguistic Features for Authorship Attribution

Lexical choice was the most robust linguistic “fingerprint” across all tested variables, uniquely reflecting stable stylistic differences in individuals' writing across the corpus. The second position was taken by character-level N-grams, which did not vary significantly with AI, as they reflect unconscious orthographic habits. Punctuation (especially the number of localization-specific characters) and capitalization metrics also yielded consistent information that, when used combinatorially in unison with word patterns, augmented the authorship attribution.

Table 7
Relative Contribution of Linguistic Features to Authorship Attribution in Human-Written, AI-Generated, and Human–AI Collaborative Texts

Linguistic Feature	Evidence from Corpus	Salience	Contribution to Authorship Attribution
Lexical choice	Stable vocabulary preferences, pronouns, and discourse particles	Very High	Primary linguistic fingerprint
Character N-grams	Character sequences, repeated letters, laughter expressions	High	Captures unconscious writing habits
Stylistic markers	Capitalization, punctuation, spelling variation	Moderate	Reinforces individual writing style
Orthographic variation	Informal spellings and abbreviations	Moderate	Reflects sociolinguistic identity

As seen in Table 7, lexical choice is the strongest linguistic fingerprint for authorship attribution, corresponding to consistent language preferences, discourse particles, pronouns, and recurrent parts of speech across texts, even in AI-assisted writing. The next most informative elements were character-level patterns (punctuation, capitalization, repeating characters, expressive spellings, because they capture unconscious orthographic habits. The discriminative strength of stylistic indicators alone is limited, but combining them with lexical information considerably improves the attribution. The results show that combining multiple linguistic variables is crucial for credible authorship attribution, providing a solid foundation for investigating human-AI collaborative writing and validating machine learning models (Alamleh et al., 2023; Abreu & Sousa, 2022; Theophilo et al., 2023).

4.3 Character-Level N-gram Patterns as Persistent Linguistic Fingerprints

Results show that character-level N-grams reveal stable micro-level writing behaviors across human, AI, and human–AI collaborative texts. Unlike lexical features, they are automatic habits that are difficult to break: spelling, punctuation, capitalization, and character sequences. AI interactions produce more regular character patterns, while human writing retains clear sub-word-level regularities, meaning that character-level features constitute an important complementary (though less reliable) basis for authorship attribution (Fedotova et al., 2022; Uchendu, 2020; Abreu & Sousa, 2022).

The analysis also found that AI-generated texts exhibited much less orthographic variability than those written by humans. The Indonesian digital communication corpus often encountered informal spellings, repeated vowels or consonants, expressive punctuation, laughter words (*wkwk*, *wkwkwk*), and non-standard abbreviations (e.g., LDR). These orthographic variants are not randomly distributed; rather, they appear consistently within individual authors in a manner indicative of stable personal writing habits. Conversely, the texts generated by an AI were more likely to use spelling, punctuation, and capitalization in a probabilistic manner, according to established language patterns—without any emotionally infused style—and, as a result, they produced much more homogeneous outputs. It is also intriguing that AI-assisted collaborative human–AI texts retained so many character-level stylometric habits exhibited by the original authors, since these micro-textual rules seem to allow certain linguistic behaviors to be preserved even when lexical or

syntactic content has been altered. Character–level N-grams have been shown to model latent stylistic information that is presumably more resistant than higher-level linguistic structures (Alamleh et al., 2023; Theophilo et al., 2023), which would explain this persistence.

Table 8.
Representative Character-Level N-gram Patterns Identified in Human-Written, AI-Generated, and Human–AI Collaborative Texts

Character-Level Feature	Human-written Text	AI-generated Text	Human–AI Collaborative Text	Forensic Interpretation
Character repetition	*baguuuusss*, *okeeee*	Rarely repeated	*baguuus*	Reflects expressive emotional style and habitual emphasis
Laughter expressions	*wkwk*, *wkwkwk*, *hehe*, *heuheu*	Mostly standardized (*haha*)	*wkwk* retained after AI revision	Indicates culturally embedded and author-specific stylistic behavior
Capitalization	*JANGAN YA!!!*	Standard sentence case	Mixed capitalization	Reveals habitual emphasis strategies
Punctuation sequences	*!!!*, *??*, *...*	Conventional punctuation	Human punctuation retained	Indicates individual interaction style
Orthographic variation	*gk*, *ga*, *nggak*, *aq*, *gw*	Standardized spelling	Mixed variants	Reflects stable sociolinguistic and personal writing preferences
Non-alphabetic symbols	*:)*, 😂, ^^, ~	Limited or normalized use	Human symbols retained	Captures interpersonal and affective communication style

Table 8 summarizes the representative patterns extracted from the corpus, distinguishing how different character-level features serve as discriminative linguistic fingerprints on writing conditions. Our results suggest that authorship attribution should not rely solely on lexical evidence but also incorporate orthographic and character-level behaviors that reflect routine writing habits. Cumulatively, from a forensic linguistic perspective, this type of feature appears to offer quite strong evidence, as it is very hard to copy (but often found to be stable across subsequent samples). This means that the mixture of lexical analysis and character-level N-gram tracing adds value for detecting personal writing styles in digitally mediated communication, especially when human–AI collaborative writing scenarios are more prevalent; Belvisi et al. (2020), Abreu & Sousa (2022).

5. DISCUSSION

The results show that human, AI, and human–AI texts have different linguistic fingerprints at the lexical, stylistic, and Character levels. We found that our results corroborated previous findings highlighting the importance of patterns for authorship attribution in forensic linguistics, and we established that character-level features are

particularly robust. The findings also emphasize applications such as IP protection, misinformation detection, and the ethical use of AI, while noting limitations in the studies conducted and areas for future research.

5.1 Persistence of Linguistic Fingerprints in Human–AI Collaborative Writing

The results show that human–AI collaborative writing still retains detectable linguistic fingerprints. Although AI generates fluent, coherent text, those patterns are still present at the lexical, stylistic, and even character level for human authors. Stable features like lexical choices, punctuation, capitalization, and orthographic variation support the idea that authorship is represented through the repetition of combinations of linguistic features, rather than isolated markers (Grant 2007; Grieve 2023; Holmes 1998).

The robustness of these linguistic "fingerprints" has critical implications for authorship detection in the age of generative AI. Prior studies have shown that texts generated by AI systems are increasingly similar to human writing and that, when used in isolation, traditional stylometric indicators become less predictive (Ippolito et al., 2020; Alamleh et al., 2023). The results indicate that the use of word- and character-level attributes creates a more powerful authorship representation than style alone While we see an effect from AI in writing, individual linguistic patterns remain strong, especially with respect to lexical and orthographic choices. These findings provide further confirmation of the utility of corpus-based forensic linguistics for inferring more consistent and multidimensional authorship in human, AI-assisted, and digitally mediated communication (Belvisi et al., 2020; Abreu & Sousa, 2022; Hovy & Yang, 2021).

5.2 Lexical Choice as the Strongest Indicator of Authorship

Diction is a key feature for distinguishing between human, AI-generated, and mixed texts because it serves as a stable linguistic signature. It encompasses vocabulary, phrase selection, idiomatic expressions, and sentence flow, all of which are shaped by factors such as education, culture, profession, and psychological state (Baker, 2010; Grieve, 2023). Although AI can imitate human language, it cannot consistently reproduce the subtle lexical preferences that characterize individual human writers, making diction a valuable indicator for authorship attribution.

Diction is a particularly strong marker of authorship because it constitutes a central aspect of an individual writer's language use. Previous studies have shown that human authors often leave distinctive verbal fingerprints, enabling higher-confidence authorship attribution even in texts containing substantial AI-generated content. Furthermore, human authors tend to exhibit consistent preferences in expressing ideas through the recurrent use of particular words, phrases, and syntactic constructions (García-Díaz et al., 2021). These recurring lexical and structural choices function as reliable stylistic indicators of authorship. This study reveals that authors preserve unique lexical choices and sentence structures, especially in an empirical domain where tokenization into words is possible, such as digital communication. Moreover, there are some of these habits that can be observed in writing aided by AI, such as a certain combination of words and phrases you struggle to get rid of. These stable linguistic features can be thought of as sufficient fingerprints for authorship — that individual writing styles persist even alongside AI-generated material.

5.3 Character-Level N-grams as Persistent Linguistic Fingerprints

Our results suggest that character N-grams capture some of the most powerful linguistic fingerprints for authorship attribution, revealing unconscious writing patterns beyond vocabulary choice. Character-level patterns, including capitalization, punctuation patterns, repeated letters, abbreviations, and orthographic variants, are invariant across texts, even with the help of AI, unlike lexical traits, which are dependent on topic and genre. These micro-level automatic behaviors are difficult to control, and so they represent important evidence for forensic authorship analysis. The study also finds that AI can normalize grammar and increase fluency in human-AI collaborative texts, but many of these orthographic tendencies are still distinctive. The results indicate that AI cannot properly replicate the nuanced writing styles of particular authors. Therefore, authorship should be considered a multidimensional construct influenced by lexical, stylistic, and character-level patterns. Character-level N-grams are thus a linguistically meaningful and reliable way to boost authorship attribution in the emerging domain of human-AI collaborative writing (Grant, 2007; Abreu & Sousa, 2022; Alamleh et al., 2023).

5.4 Implications for Forensic Linguistics, Intellectual Property Protection, Misinformation Detection, and Ethical Human–AI Collaboration

These study findings have critical forensic linguistic implications by demonstrating in multiple human–AI collaborative writing environments that authorship is still detectable through stable linguistic fingerprints. Corpus analysis suggests that lexical preferences, stylistic markers, and character-level N-gram patterns collectively retain fingerprints of writing behavior even in AI-supported text output. These results reinforce the commonsense foundation of forensic linguistics, which (some) authors implicitly generate linguistic ruts to be later used in attributing authorship (Belvisi et al., 2020; Abreu & Sousa, 2022). The integration of corpus-based linguistic analysis with character-level stylistic traits provides a transparent, interpretable framework for forensic authorship identification, anonymous communication, cybercrime investigations, and digital evidence processing. The findings also have crucial implications for intellectual property protection, showing that stable human linguistic fingerprints remain in AI-assisted writing. Hence, in the future, authorship verification should rely on persistent language features, not only on the AI-detection technologies, to help, for example, in copyright disputes and academic integrity investigations (Alamleh et al., 2023; Habibzadeh, 2023).

Table 9.
Practical Implications of Persistent Linguistic Fingerprints in Human–AI Collaborative Writing

Application Domain	Key Findings from This Study	Practical Implications
Forensic Linguistics	Lexical, stylistic, and character-level features remain stable across writing contexts	Supports authorship attribution, forensic investigation, anonymous text identification, and cybercrime analysis
Intellectual Property Protection	Human linguistic fingerprints remain observable despite AI assistance	Assists with copyright disputes, authorship verification, and academic integrity assessment

(*Continue on the next page*)

Table 9 (*Continued*)

Application Domain	Key Findings from This Study	Practical Implications
Misinformation Detection	AI-generated texts exhibit comparatively standardized linguistic patterns	Supports the identification of AI-assisted misinformation and coordinated online manipulation
Ethical Human–AI Collaboration	Human authors preserve distinctive stylistic behaviors during AI-assisted writing.	Encourages transparent AI disclosure and responsible human–AI collaboration in digital communication.

The results also have wider implications for detecting misinformation and integrating AI ethically into digital communication. While generative AI can generate fluent, grammatically correct sentences, the analysis shows that AI-generated writing is more standardized lexically and characteristically than rigorous human writing, in which interpersonal communication maintains greater stylistic coherence between a human individual or group. And more broadly, being aware of these lasting linguistic fingerprints might help flag coordinated misinformation campaigns, exposed and manipulated digital content, and bolster the efficacy of digital forensic investigations. Even more importantly, this study pins the notion that human–AI collaborative texts should be treated as complementary writing practice, not a bolt-on for human authorship. Unadorned disclosure of the use of AI, alongside linguistic evidence of explainability for authorship attribution, may help foster ethical and responsible research- and public-facing development governance of generative AI technologies (Theophilo et al., 2023; Yang et al., 2024).

The study has been limited to Indonesian digital communication, namely instant messaging and social media. Hence, its conclusions may not be generalized to other languages, writing systems, or cultural contexts. The analysis was restricted to lexical, stylistic, and character-level aspects and did not consider other evidence such as discourse structure, semantic representation, multimodal communication, temporal behavior, and interactional metadata. Future work should validate these results across multilingual corpora, combine corpus-based language analysis with explainable computational methods, and examine the effects of long-term human-AI collaboration on writing style. Such effort will enhance transparent, dependable, and ethically sound authorship attribution in increasingly AI-assisted digital communication.

6. CONCLUSION

This research shows that semantics are the key lexical feature that differentiates human/creative and machine/procedural writing, which is critical for authorship attribution. However, you can see most by doing a character-level N-gram analysis. The analytic approach that identified both cases breaks down the alphabet into alphabetic and non-alphabetic characters, taking into account case (i.e., capitalization) and punctuation use, which are important for distinguishing authorship in mixed text. This widens the scope of work for developing authorship verification models sophisticated enough to identify not only humans but also humans and AI working together in modern text production. On the one hand, this study enriches the frontier of authorship attribution, which carries important theoretical significance; on the other hand, it provides practical support and downstream tasks include, but are not limited to, digital forensics and the detection of machine-generated content across different domains, such as media. The alternative is to resist the growing

presence of AI-generated text, but then be prepared for increased disrespect for linguistic differences from those who take a harder line. The capacity to correctly identify authorship in intermingled textual content is pressingly needed, not least from a computational standpoint, but also politically, to hold people accountable for what humans produce in writing and for circulating information we want to share.

While this work offers strong corpus-based evidence for the persistence of human-like linguistic fingerprints in text produced via AI–human collaboration, its results are restricted to Indonesian digital communication and a subset of lexical, stylistic, and character-level features. Future studies based on this framework should investigate the evolution of linguistic fingerprints across increasingly complex digital communication environments over extended periods of human–AI collaboration, using multilingual and multimodal datasets.

Acknowledgment
The authors gratefully acknowledge the CILLCO Indonesian Corpora (Corpus of Indonesian Language, Literature, and Community) for providing access to the linguistic data used in this study. The corpus was developed and maintained through the Research House Program of the Research Organization for Archaeology, Language, and Literature, National Research and Innovation Agency (BRIN), Grant No. 2/III.8/HK/2025.

Availability of Data and Materials
The datasets generated and analysed in the study are available in Lestariningsih, Dewi Nastiti; Puspitasari, Devi Ambarwati; Karlina, Yenny; Handoko, Mu'awal Panji; Sukma, Bayu Permana; Sanubarianto, Salimulloh Tegar, 2025, "Eksplorasi Atribusi Kepenulisan Komunitas Multietnis di Kalimantan Barat sebagai Upaya Pemajuan Kebudayaan", https://hdl.handle.net/20.500.12690/RIN/JUBPBH, RIN Dataverse, V1 And in Puspitasari, Devi Ambarwati; Karlina, Yenny; Hernina, Hernina; Kurniawan, Kurniawan; Mulyo, Budi Mukhammad, 2025, "Rancang Bangun Text Curation Engine Forensik Kebahasaan", https://hdl.handle.net/20.500.12690/RIN/B9LJYY, RIN Dataverse, V1.

Competing Interests
The authors declare that they have no competing interests.

Funding
This work was supported by the Research House Program of the Research Organization for Archaeology, Language, and Literature, National Research and Innovation Agency (BRIN), Grant No. 2/III.8/HK/2025.

Authors' Contribution
Devi Ambarwati Puspitasari conceived the main conceptual ideas; *Dewi Nastiti Lestariningsih* validated the data and revised the manuscript for grammatical accuracy; Yenny Karlina performed the data analysis*; Bayu Permana Sukma* assisted in collecting the research data and contributed to the analysis*; Salimulloh Tegar Sanubarianto* provided conceptual contributions to the analysis; *Intan Pradita* designed the analysis framework, and Mu'awal Panji Handoko substantively edited the manuscript.

Authors' Information
DEWI NASTITI LESTARININGSIH is a junior researcher at the National Research and Innovation Agency. Her research focuses on literacy and corpus linguistics.
Email: dewi041@brin.go.id; ORCID https://orcid.org/0000-0003-1717-5539
DEVI AMBARWATI PUSPITASARI is a junior researcher *at the National Research and Innovation Agency. Her research focuses on forensic linguistics, corpus linguistics, and multilingualism.*
Email: devi018@brin.go.id; ORCID https://orcid.org/0000-0002-9093-1516
BAYU PERMANA SUKMA is a junior researcher *at the National Research and Innovation Agency. His research focuses on corpus-assisted discourse analysis, political discourse, and media discourse.*
Email: bayu025@brin.go.id; ORCID https://orcid.org/0000-0002-7873-8158
YENNY KARLINA is a junior researcher at the National Research and Innovation Agency. His research focuses on applied linguistics.
Email: yenn010@brin.go.id; ORCID https://orcid.org/0009-0006-4097-1778
SALIMULLOH TEGAR SANUBARIANTO is a junior researcher at the National Research and Innovation Agency. His research focuses on forensic linguistics.
Email: sali004@brin.go.id; ORCID https://orcid.org/0000-0002-4367-8773

MU'AWAL PANJI HANDOKO is a junior researcher at the National Research and Innovation Agency. His research focuses on political communication.
Email: muaw001@brin.go.id; ORCID https://orcid.org/0000-0001-8799-0614
INTAN PRADITA is a lecturer at English Language Education, Faculty of Social Cultural Sciences, Universitas Islam Indonesia. Her research focuses on corpus linguistics, multilingualism, and systemic functional linguistics.
Email: intan.pradita@uii.ac.id; ORCID https://orcid.org/0000-0001-9938-9842

REFERENCES

Abreu Rodrigues, S., & Sousa Silva, R. (2022). A Forensic Authorship Analysis of Threats. *RevSALUS - Revista Científica Da Rede Académica Das Ciências Da Saúde Da* Lusofonia, 4(Sup),p. 98–99. https://doi.org/10.51126/revsalus.v4isup.324

Alamleh, H., Alqahtani, A. A. S., & Elsaid, A. (2023). Distinguishing Human-Written and ChatGPT-Generated Text Using Machine Learning. *2023 Systems and Information Engineering Design Symposium, SIEDS 2023*. https://doi.org/10.1109/SIEDS58326.2023.10137767

Aribowo, A. S., Basiron, H., Yusof, N. F. A., & Khomsah, S. (2021). Cross-Domain Sentiment Analysis Model on Indonesian YouTube Comment. *International Journal of Advances in Intelligent Informatics*, *7*(1), 12–25. https://doi.org/10.26555/ijain.v7i1.554

Baker, P. (2010). Title of chapter. In J. Smith (Ed.), *Sociolinguistics and Corpus Linguistics* (pp. 25–42). Edinburgh University Press.

Banga, R., Bhardwaj, A., Peng, S. L., & Shrivastava, G. (2018). Authorship Attribution for Online Social Media. In *Social Network Analytics for Contemporary Business Organizations* (pp. 141–165). https://doi.org/10.4018/978-1-5225-5097-6.ch008

Belvisi, N. M. S., Muhammad, N., & Alonso-Fernandez, F. (2020). *Forensic Authorship Analysis of Microblogging Texts Using N-Grams and Stylometric Features*. In *2020, the 8th International Workshop on Biometrics and Forensics (IWBF)* (pp. 1–6). IEEE. https://doi.org/10.1109/IWBF49977.2020.9107956

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Bevendorff, J., Ghanem, B., Giachanou, A., Kestemont, M., Manjavacas, E., Markov, I., Mayerl, M., Potthast, M., Rangel, F., Rosso, P., Specht, G., Stamatatos, E., Stein, B., Wiegmann, M., & Zangerle, E. (2020). Overview of Pan 2020: Authorship Verification, Celebrity Profiling, Profiling Fake News Spreaders on Twitter, and Style Change Detection. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12260 Lncs*. https://doi.org/10.1007/978-3-030-58219-7_25

Beyer, K. (2014). Urban Language Research in South Africa: Achievements and Challenges. *Southern African Linguistics and Applied Language Studies, 32*(2), 247–254. https://doi.org/10.2989/16073614.2014.992643

BRIN. Cillco Indonesia Corpora: Corpus of Indonesia Language, Literature, and Community 2024. https://cillco-prototype.brin.go.id

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics Derived Automatically From Language Corpora Contain Human-Like Biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Chen, B., Ding, X., Zhao, Y., Fu, B., Lin, T., Qin, B., & Liu, T. (2024). Text Difficulty Study: Do Machines Behave the Same as Humans Regarding Text Difficulty? *Machine*

Intelligence Research, 21(2), 283–293. https://doi.org/10.1007/s11633-023-1424-x

Dou, Y., Forbes, M., Koncel-Kedziorski, R., Smith, N. A., & Choi, Y. (2022). Is GPT-3 Text Indistinguishable from Human Text? Scarecrow: A Framework for Scrutinizing Machine Text. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1, 7250–7274. https://doi.org/10.18653/v1/2022.acl-long.501

Dugan, L., Ippolito, D., Kirubarajan, A., Shi, S., & Callison-Burch, C. (2023). Real or Fake Text?: Investigating Human Ability to Detect Boundaries Between Human-Written and Machine-Generated Text. *Proceedings of the AAAI Conference on Artificial Intelligence, 37*(11), 12763–12771. https://doi.org/10.1609/aaai.v37i11.26501

Fedotova, A., Romanov, A., Kurtukova, A., & Shelupanov, A. (2022). Authorship Attribution of Social Media and Literary Russian-Language Texts Using Machine Learning Methods and Feature Selection. *Future Internet, 14*(1), 4. https://doi.org/10.3390/fi14010004

García-Díaz, J. A., Colomo-Palacios, R., & Valencia-García, R. (2022). Psychographic Traits Identification Based on Political Ideology: An Author Analysis Study on Spanish Politicians' Tweets Posted in 2020. *Future Generation Computer Systems, 130*, 59–74. https://doi.org/10.1016/j.future.2021.12.011

Grant, T. (2007). Quantifying Evidence in Forensic Authorship Analysis. *International Journal of Speech, Language and the Law*, *14*(1), 1–25. https://doi.org/10.1558/ijsll.v14i1.1

Grieve, J. (2023). Register Variation Explains Stylometric Authorship Analysis. *Corpus Linguistics and Linguistic Theory, 19*(1), 47–77. https://doi.org/10.1515/cllt-2022-0040

Grieve, J., Clarke, I., Chiang, E., Gideon, H., Heini, A., Nini, A., & Waibel, E. (2019). Attributing the Bixby Letter Using N-Gram Tracing. *Digital Scholarship in the Humanities*, *34*(3), 493–512. https://doi.org/10.1093/llc/fqy042

Habibzadeh, F. (2023). GPTZero Performance in Identifying Artificial Intelligence-Generated Medical Texts: A Preliminary Study. *Journal of Korean Medical Science, 38*(38), 1–12. https://doi.org/10.3346/jkms.2023.38.e319

Holmes, D. I. (1998). The Evolution of Stylometry in Humanities Scholarship. *Literary and Linguistic Computing*, *13*(3), 111–117. https://doi.org/10.1093/llc/13.3.111

Hovy, D., & Yang, D. (2021). The Importance of Modeling Social Factors of Language: Theory and Practice. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 588–602. https://doi.org/10.18653/v1/2021.naacl-main.49

Ippolito, D., Duckworth, D., Callison-Burch, C., & Eck, D. (2020). Automatic Detection of Generated Text Is Easiest When Humans Are Fooled. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 1808–1822. https://doi.org/10.18653/v1/2020.acl-main.164

Katib, I., Assiri, F. Y., Abdushkour, H. A., Hamed, D., & Ragab, M. (2023). Differentiating Chat Generative Pretrained Transformer From Humans: Detecting ChatGPT-Generated Text and Human Text Using Machine Learning. *Mathematics, 11*(15), Article 3400. https://doi.org/10.3390/math11153400

Kestemont, M. (2018). Overview of the Author Identification Task at Pan-2018: Cross-Domain Authorship Attribution and Style Change Detection. *CEUR Workshop Proceedings*, *2125*(Query date: 2024-03-21 10:08:46). https://api.elsevier.com/content/abstract/scopus_id/85051054751

Khalifa, M., & Albadawy, M. (2024). Using Artificial Intelligence in Academic Writing and Research: An Essential Productivity Tool. *Computer Methods and Programs in Biomedicine Update*, *5*, 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

Knowles, A. M. (2024). Machine-In-The-Loop Writing: Optimizing the Rhetorical Load. *Computers and Composition, 71*, Article 102826. https://doi.org/10.1016/j.compcom.2024.102826

Koppel, M., Schler, J., & Argamon, S. (2011). Authorship Attribution in the Wild. *Language Resources and Evaluation, 45*(1), 83–94. https://doi.org/10.1007/s10579-009-9111-2

Lestariningsih, Dewi Nastiti; Puspitasari, Devi Ambarwati; Karlina, Yenny; Handoko, Mu'awal Panji; Sukma, Bayu Permana; Sanubarianto, Salimulloh Tegar, 2025, "Eksplorasi Atribusi Kepenulisan Komunitas Multietnis di Kalimantan Barat Sebagai Upaya Pemajuan Kebudayaan", https://hdl.handle.net/20.500.12690/RIN/JUBPBH, RIN Dataverse, V1

Machová, K., Szabóova, M., Paralič, J., & Mičko, J. (2023). Detection of Emotion by Text Analysis Using Machine Learning. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1190326

MacLeod, N., & Grant, T. (2012). *Whose Tweet? Authorship Analysis of Micro-Blogs and Other Short-Form Messages*. In S. Tomblin, N. MacLeod, R. Sousa-Silva, & M. Coulthard (Eds.), *Proceedings of the International Association of Forensic Linguists' Tenth Biennial Conference* (pp. 210–224). Centre for Forensic Linguistics, Aston University.

McMenamin, G. R. (2002). *Forensic Linguistics: Advances in Forensic Stylistics*. CRC Press LLC.

Moneus, A. M., & Sahari, Y. (2024). Artificial Intelligence and Human Translation: A Contrastive Study Based on Legal Texts. *Heliyon, 10*(6), Article e28106. https://doi.org/10.1016/j.heliyon.2024.e28106

Pradita, I., Puspitasari, D. A., Karlina, Y., & Sukma, B. P. (2026). Introducing CILLCO: A Corpus Model of Vernacular Indonesian as a Cultural Capital. *Jurnal Komunikasi*, 20(1), 161–174. https://doi.org/10.20885/komunikasi.vol20.iss1.art10

Puspitasari, Devi Ambarwati; Karlina, Yenny; Hernina, Hernina; Kurniawan, Kurniawan; Mulyo, Budi Mukhammad, 2025, "Rancang Bangun Text Curation Engine Forensik Kebahasaan", https://hdl.handle.net/20.500.12690/RIN/B9LJYY, RIN Dataverse, V1

Santos, F. A. O., Macedo, H. T., Bispo, T. D., & Zanchettin, C. (2021). Morphological Skip-Gram: Replacing Fast Text Characters N-Gram with Morphological Knowledge. *Inteligencia Artificial, 24*(67), 1–17. https://doi.org/10.4114/intartif.vol24iss67pp1-17

Singh, A., Sharma, D., Nandy, A., & Singh, V. K. (2024). Towards a Large-Sized Curated and Annotated Corpus for Discriminating Between Human Written and AI-Generated Texts: A Case Study of Text Sourced From Wikipedia and ChatGPT. *Natural Language Processing Journal, 6*, 100050. https://doi.org/10.1016/j.nlp.2023.100050

Stamatatos, E. (2009). A Survey of Modern Authorship Attribution Methods. *Journal of the American Society for Information Science and Technology*, *60*(3), 538–556. https://doi.org/10.1002/asi.21001

Stamatatos, E. (2018). Masking Topic-Related Information to Enhance Authorship Attribution. *Journal of the Association for Information Science and Technology*, *69*(3), 461–473. https://doi.org/10.1002/asi.23968

Theophilo, A., Giot, R., & Rocha, A. (2023). Authorship Attribution of Social Media Messages. *IEEE Transactions on Computational Social Systems*, *10*(1), 10–23.

https://doi.org/10.1109/TCSS.2021.3123895

Uchendu, A. (2020). Authorship Attribution for Neural Text Generation. *EMNLP 2020 - 2020 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, *Query date: 2024-03-21 10:08:46*, 8384–8395. https://api.elsevier.com/content/abstract/scopus_id/85098761250

Yang, L., Jiang, F., & Li, H. (2024). Is ChatGPT Involved in Texts? Measure the Polish Ratio to Detect ChatGPT-Generated Text. *APSIPA Transactions on Signal and Information Processing, 13*(2), 1–19. https://doi.org/10.1561/116.00000250