Preprints.org

Article

Can AI-Driven Plagiarism Detection Tools Uphold Academic Integrity Without Ethical Compromises? A Comprehensive Analysis of False Positives, Contextual Misunderstandings, and Dependency Issues

Edgar Eslit [*]

Article

Can AI-Driven Plagiarism Detection Tools Uphold Academic Integrity Without Ethical Compromises? A Comprehensive Analysis of False Positives, Contextual Misunderstandings, and Dependency Issues

Edgar R. Eslit

St. Michael's College of Iligan, Inc.; edgareslit@gmail.com, e.eslit@my.smciligan.edu.ph

Abstract: The rapid advancement of digital technologies has significantly impacted academic practices, particularly in the area of plagiarism detection. As universities and research institutions adopt tools to safeguard academic integrity, concerns arise about their effectiveness and potential limitations. This study investigates the role of automated plagiarism detection tools in higher education, examining how they influence academic practices and the detection of both traditional and AI-generated plagiarism. Despite the sophistication of tools like Turnitin, PlagScan, GPTZero, and QuillBot, the research finds that these systems often struggle with accurately interpreting context, resulting in false positives and overlooked instances of plagiarism. The study underscores the necessity of combining technology with human judgment, recognizing that such tools should be seen as supplementary rather than definitive measures of originality. Grounded in theoretical frameworks such as Technological Determinism, Actor-Network Theory (ANT), and Socio-Technical Systems Theory (STS), the research highlights the complex relationship between technology, academia, and societal expectations. Through a qualitative analysis of existing literature, the study identifies key challenges and suggests that hybrid approaches, blending technological tools with human oversight, may offer a more balanced and effective approach to plagiarism detection. The findings encourage further exploration into the ethical implications of reliance on automated systems in education and their broader impact on academic integrity.

Keywords: academic integrity; AI-driven plagiarism detection tools; contextual misunderstandings; dependency issues; ethical compromises; false positives

I. Introduction
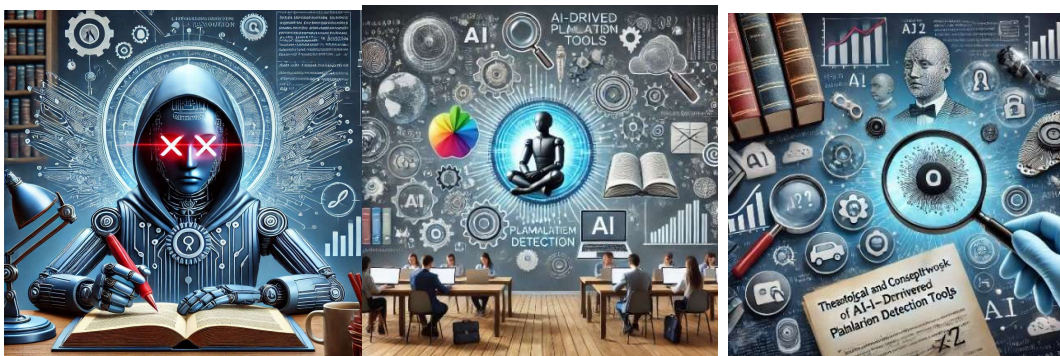
Source: Microsoft. (n.d.). Copilot [AI assistant]

Plagiarism detection tools have become indispensable in academia, where the pursuit of academic integrity is increasingly influenced by technological advancements. AI-driven tools, such

as Turnitin and PlagScan, are integral to identifying potential instances of academic dishonesty, providing educators with a means to maintain the authenticity of scholarly work. However, the widespread use of these tools has raised critical concerns related to over-reliance, ethical issues, and their limitations in detecting more subtle forms of plagiarism. These concerns necessitate a deeper exploration of how AI tools function, their implications, and their limitations in academic environments.

While tools like Turnitin, PlagScan, Copyleaks, GPTZero, and QuillBot offer significant capabilities in detecting plagiarism and AI-generated content, they cannot and should not be used as the final arbiters in determining originality in research outputs. These AI-driven tools often fall short in accurately interpreting context, leading to false positives, and they can miss nuanced cases of plagiarism. Their reliance on pattern recognition without human oversight can undermine academic integrity rather than uphold it. Therefore, it is imperative to use these tools as supplementary aids, with human judgment playing a crucial role in the evaluation process (Pudasaini, Miralles-Pechuán, Lillis, & Salvador, 2024).

The primary aim of this research article is to explore the strengths and limitations of AI-driven plagiarism detection tools in maintaining academic integrity, while addressing the ethical challenges associated with their use. Specifically, this research examines the accuracy of these tools, their contextual interpretation abilities, and the broader impact of AI on academic practices. To achieve this, the study is framed around the following questions:

How do AI-driven plagiarism detection tools contribute to upholding academic integrity?

1. What are the primary challenges and limitations, including false positives and contextual misunderstandings, associated with these tools?

2. How do these tools affect non-native speakers and their ability to produce original academic work?

This study is grounded in a theoretical framework that provides a comprehensive understanding of the interplay between technology and academic integrity. Technological Determinism helps explain how AI tools influence academic practices, suggesting that technology plays a significant role in shaping both academic behavior and the perception of plagiarism. Socio-Technical Systems Theory (STS) underscores the importance of integrating AI tools within the social and organizational contexts of educational institutions, highlighting the need for human oversight to address errors like false positives. Actor-Network Theory (ANT) offers a unique perspective on how AI tools function within complex networks, involving students, educators, and administrators, thereby influencing power dynamics and educational practices.

This exploration is essential, as AI-driven plagiarism detection tools, despite their promise of efficiency and objectivity, are not without their limitations. They may flag original content as plagiarized, misinterpret contextual nuances, and disproportionately impact non-native speakers. As such, this study argues that AI tools should not be used in isolation; human judgment must play a central role in ensuring that academic integrity is upheld. Through this analysis, the paper seeks to highlight both the benefits and challenges of AI-driven plagiarism detection systems, ultimately contributing to a more nuanced understanding of their role in modern education.

To be direct, this article asserts that AI-driven plagiarism detection tools, while useful, are not infallible. The mere identification of plagiarism by these systems does not inherently confirm academic misconduct or imply that the work is entirely AI-generated. While these tools promise efficiency and objectivity, their inability to accurately interpret context and their tendency to generate false positives necessitate the integration of human oversight in their use, ensuring a fair and balanced approach to maintaining academic integrity.

II. Literature review

As AI-driven plagiarism detection tools become more common in academic settings, researchers are starting to explore both their advantages and their shortcomings. Tools like Turnitin and PlagScan offer valuable support in spotting potential plagiarism, but they are not without their flaws. They often struggle with interpreting context and can mistakenly flag original work or miss

subtle cases of dishonesty. This literature review brings together key studies that examine how these tools work, where they fall short, and what that means for academic integrity. By diving into these discussions, it is clear that while AI can assist in maintaining honesty in academia, it is human judgment that remains essential in ensuring fairness and accuracy.

Dependence on AI Use. In recent years, journal editors have increasingly turned to AI tools to help maintain the integrity of academic publishing. These tools, which are primarily used for plagiarism detection and identifying AI-generated content, have become essential in ensuring that manuscripts meet the standards of originality expected in scholarly work. Editors rely on these tools to flag instances of academic misconduct, offering a quick and efficient way to spot issues that might otherwise go unnoticed. However, this growing dependence on AI raises important questions about the limitations of these tools. While AI can certainly enhance the efficiency and accuracy of plagiarism detection, it cannot fully replace the critical judgment of experienced editors (Quidwai et al., 2023). The subtle nuances of academic writing, such as rewording or contextual shifts, may be overlooked by algorithms, which is why editors must strike a balance between leveraging AI tools and applying their own expertise in evaluating a manuscript's originality.

Editors vs. Writers Use of AI. When comparing the way editors and writers use AI tools, the difference in their roles becomes clear. Editors generally rely on these tools after a manuscript has been submitted, using them to verify the originality of the content. In contrast, writers often use plagiarism detection tools before submitting their work. Writers engage with these tools as a precautionary measure, ensuring that their work meets the necessary standards of academic integrity before it reaches the editorial stage. This distinction in timing and intent highlights the different ways AI tools fit into the academic workflow. While writers see these tools as a way to protect themselves from potential issues, editors view them as a way to confirm the quality of submissions (Hutson, 2024). The differences in usage reflect the varying stakes each group has in the process: for writers, it is about avoiding rejection or revision, and for editors, it is about maintaining the credibility and integrity of the journal.

Risks and Benefits of AI Reliance. The growing reliance on AI tools in the editorial process has its benefits and drawbacks. On the positive side, these tools help editors quickly and effectively identify potential issues with plagiarism, saving time and effort in reviewing large volumes of submissions. This efficiency is particularly valuable in high-pressure environments where editors are tasked with evaluating numerous manuscripts. However, there are also risks associated with over-relying on AI (Chaka, 2024).. Editors may begin to defer too much to these tools, trusting their results without questioning the nuances that may not be captured by algorithms. For example, AI tools may flag content as plagiarized based on superficial similarities, without understanding the context or intent behind the writing. As a result, editors must continue to play a central role in the review process, ensuring that AI serves as a helpful supplement rather than a replacement for their judgment.

Adaptability and Effectiveness of AI. AI tools have shown varying levels of effectiveness across academic disciplines, and their ability to detect plagiarism is often tied to the specific writing styles and practices of each field. In the natural sciences, where research tends to follow a more standardized language and format, AI tools are generally quite effective at identifying direct copying or improper citations. However, in disciplines like the humanities, where interpretation and rephrasing play a larger role, AI tools can struggle to accurately assess originality. In these fields, scholars often build on existing ideas by paraphrasing or reinterpreting sources, which can make it difficult for AI to distinguish between legitimate academic work and plagiarism (Santra & Majhi, 2023). This disparity highlights the need for more sophisticated tools that can adapt to the unique demands of different disciplines, helping to bridge the gap between the capabilities of AI and the diverse nature of academic writing.

New Forms of Plagiarism and its Adaptability. As academic dishonesty evolves, AI tools must also adapt to detect new forms of plagiarism. One of the biggest challenges is recognizing paraphrased content, which is often harder to spot than direct copying. While AI technology has made significant strides in recent years, it still struggles with detecting subtle paraphrasing, which is

4

a common form of academic misconduct. The rise of AI-generated content presents another challenge, as these tools are not always equipped to recognize text produced by artificial intelligence models (Francke & Bennett, 2019. Unlike traditional plagiarism, AI-generated text may be syntactically correct and coherent, making it difficult for detection tools to flag it. To stay relevant, AI tools will need to evolve continually, improving their ability to detect more sophisticated forms of academic dishonesty, including paraphrasing and content generated by AI.

User Feedback, Ethical Considerations, and Implications.   With the growing use of AI tools in academic settings, concerns about privacy and data security have become increasingly important. Many plagiarism detection tools require users to upload their work to external servers for analysis, which raises questions about how personal data and intellectual property are handled. These tools must ensure that they are compliant with privacy regulations and that user data is stored securely (Perkins, et al., 2024). Users, be it students, writers, or researchers, need assurances that their work will not be misused or exposed to unauthorized access. If AI tools do not take adequate precautions to protect sensitive data, they could deter individuals from using them, potentially undermining the value these tools bring to the academic process. Ensuring transparency about data handling practices is critical for building trust and maintaining the integrity of these tools.

Implications for Educational Practices. The integration of AI tools into academic settings is not just changing the way plagiarism is detected—it's also influencing teaching and learning practices. Educators now have a powerful tool at their disposal to help students understand the importance of academic integrity, but this reliance on technology could also lead to unintended consequences (Amigud & Pell, 2021). If students become overly dependent on AI tools to ensure their work is plagiarism-free, it could discourage genuine engagement with the material and critical thinking. Educators will need to find a balance between teaching students how to use AI tools responsibly and encouraging them to develop their own analytical skills and academic voice. The goal should be to use AI tools to complement the learning process, helping students understand the ethical implications of plagiarism while fostering deeper, more meaningful engagement with the content.

Experience and Feedback of Users. Finally, the feedback from users, probably students, writers, and editors alike, plays a crucial role in improving the effectiveness and fairness of AI tools. Many users have reported that plagiarism detection tools are helpful for ensuring their work is original and meets academic standards, but there are also concerns about false positives or the lack of transparency in how these tools operate (Giray, 2024). Some students, for instance, may feel frustrated if their work is flagged for plagiarism due to minor issues with citation formatting or phrasing. Similarly, editors may appreciate the speed and convenience of AI tools but may also express concerns about the reliability and accuracy of the results. Understanding these user experiences is essential for refining AI tools, ensuring they are both effective and fair in their assessments. Transparent communication about how these tools work and how their results are interpreted will be key to gaining broader acceptance among users.

III. Theoretical and Conceptual Framework

The current study is grounded in several theoretical and conceptual frameworks that collectively shed light on the role of AI-driven plagiarism detection tools in shaping academic integrity and educational practices. These frameworks help us understand how technology influences the detection of plagiarism, affects student behavior, and transforms the academic environment. By integrating the theories of Technological Determinism, Socio-Technical Systems Theory (STS), and Actor-Network Theory (ANT), this study offers a multifaceted analysis of how these tools impact education, addressing both their potential and limitations.
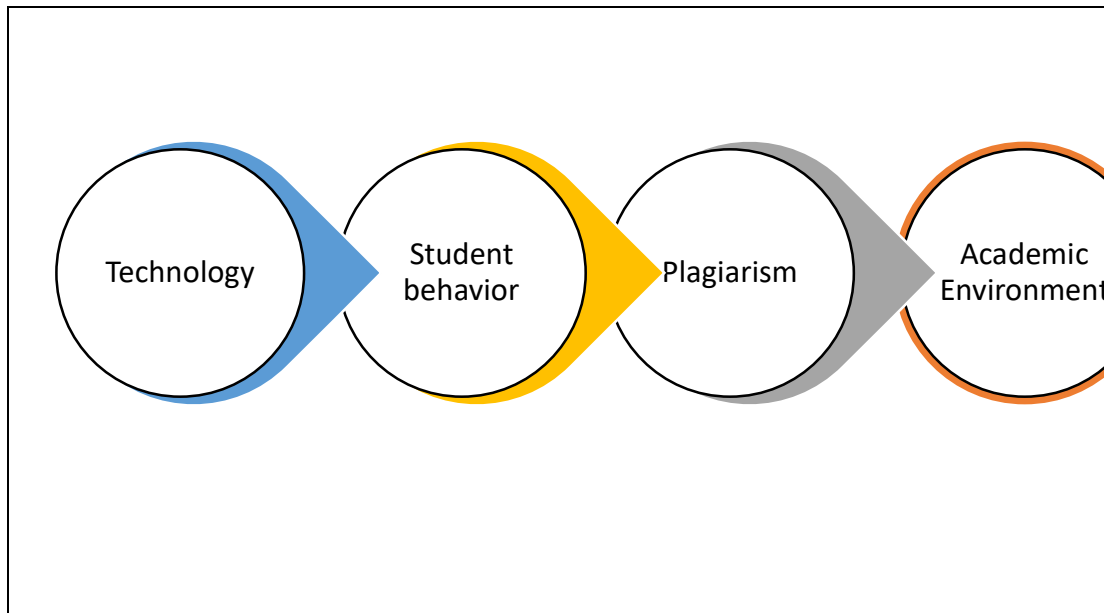
Figure 1. Conceptual framework.

Technological Determinism, as proposed by Marshall McLuhan (1964) and Andrew Feenberg (1991), provides a foundational lens for understanding how AI tools shape society and human behavior. According to this theory, technology is not merely a tool, but an active force that drives societal changes. In the context of AI-driven plagiarism detection tools, Technological Determinism can explain how these tools not only automate the detection of plagiarism but also influence how academic integrity is perceived and maintained. The growing reliance on AI tools to uphold academic honesty raises important questions about how these technologies redefine academic practices and affect both students and educators. This framework helps us examine the extent to which AI technologies are reshaping the academic landscape and influencing ethical standards in education.

Socio-Technical Systems Theory (STS), developed by Eric Trist and Fred Emery in the mid-20th century, emphasizes the interconnectedness of social and technical systems within organizational settings. This theory is particularly relevant for understanding how AI-driven plagiarism detection tools function within educational institutions. STS suggests that the success of these tools depends not only on their technical capabilities but also on how they are integrated into the social and organizational context. This framework helps us explore issues such as false positives and contextual misunderstandings, which arise from the interaction between AI tools and human users. By considering both technological and social aspects, STS highlights the need for human oversight to complement AI tools, mitigating errors and enhancing overall reliability in plagiarism detection.

Actor-Network Theory (ANT), proposed by Bruno Latour, Michel Callon, and John Law in the late 1980s, explores how human and non-human actors (including technology) form networks that shape social phenomena. In the context of AI-driven plagiarism detection tools, ANT helps analyze how these tools function as part of a broader network involving students, educators, administrators, and the technologies themselves. This theory highlights the dependencies and power dynamics that emerge when AI tools are integrated into academic settings. By examining these relationships, ANT sheds light on how the tools themselves can influence human behavior and institutional practices. It underscores the role of non-human actors (like AI tools) in shaping the educational process, offering valuable insights into the impact of dependency on technology in maintaining academic integrity.

Together, these frameworks provide a comprehensive understanding of how AI-driven plagiarism detection tools are transforming academic integrity and educational practices. Technological Determinism helps us explore how AI tools shape academic behavior and influence educational norms, while STS emphasizes the importance of balancing human oversight with technological capabilities to improve reliability and reduce errors. ANT provides insights into the

6

complex interactions and power dynamics that arise when AI tools are implemented, further informing the study's analysis of their role in academic settings. These frameworks collectively allow for a deeper investigation into both the benefits and limitations of AI-driven plagiarism detection tools, offering a nuanced view of their impact on academic practices and integrity. By grounding the research in these theories, the study can offer a more holistic understanding of the complexities surrounding AI's role in shaping academic integrity in the post-pandemic educational landscape.

IV. Research Methodology

Figure 3. Research Methodology.

This study took a qualitative approach to explore how AI-driven plagiarism detection tools are being used in academic settings, focusing on their influence on academic integrity and educational practices. By reviewing existing scholarly articles, the aim was to uncover key insights, recurring themes, and various perspectives that shed light on how these AI tools are reshaping academic behavior and institutional practices. A qualitative approach was selected because it allows for a deeper understanding of complex issues, capturing the nuances of how AI technologies affect academic ethics and practices in higher education (Hennink, Hutter, & Bailey, 2020).

To ensure the research was grounded in credible and relevant sources, articles were carefully chosen from well-established academic databases like Google Scholar, Mendeley, Web of Science, and Scopus. These platforms provided access to peer-reviewed articles, conference papers, and book chapters, offering a comprehensive view of AI's role in plagiarism detection. Additionally, articles from St. Michael's College, Iligan City (SMCII) were included to provide localized insights, helping to enrich the analysis with perspectives specific to the Philippine context (Savin-Baden & Major, 2023).

The selection criteria followed several key criteria to ensure high-quality literature. First, the sources had to be directly relevant to AI in plagiarism detection and academic integrity. Only works that contributed to understanding how AI tools were integrated into academic practices and their ethical implications were included. The credibility of the sources was another major factor; the study prioritized articles from peer-reviewed journals recognized for their academic rigor (Elkhatat, Elsaid, & Almeer, 2023). Timeliness was also crucial, so the research focused on publications from 2000 onwards to ensure the findings reflected the latest developments in AI technologies (Pudasaini et al., 2024).

Moreover, the methodological rigor of the studies was carefully considered. Articles using robust research methodologies—whether qualitative, quantitative, or mixed methods—were prioritized. These studies were selected because their clear data collection processes and sound

analytical frameworks added depth and credibility to the review. The research also aimed to capture a diverse range of viewpoints, drawing from various disciplines and geographic locations. This diversity ensured a comprehensive understanding of AI tools' impact globally (Gall & Maniadis, 2019).

The data for this study was gathered through a systematic search of academic databases, using specific keywords related to plagiarism detection, AI dependency, and academic integrity. This search strategy ensured that the literature review was both focused and thorough (Dalalah & Dalalah, 2023).

Once the relevant articles were identified, the data was analyzed using two primary techniques: **thematic analysis** and **comparative analysis**. Thematic analysis helped to identify key themes within the literature, such as the impact of AI tools on academic practices and the ethical challenges they present. These recurring patterns provided valuable insights into how AI is transforming academic integrity (Francke & Bennett, 2019).

In addition, a comparative analysis was conducted to explore the differences in how plagiarism detection tools are used by journal editors and writers. This analysis was important for understanding the divergent practices and perceptions within academic publishing and for highlighting the ethical concerns specific to each group (Gall & Maniadis, 2019).

By combining these two analytical techniques, the study aimed to offer a comprehensive understanding of the implications of AI-driven plagiarism detection tools on academic behavior, institutional practices, and ethical considerations. The findings provide a nuanced view of how AI is reshaping academic integrity and the educational landscape (Pudasaini et al., 2024).

Ethical consideration: Data for this study were gathered through a thorough review of scholarly articles from reputable academic databases such as Google Scholar, Mendeley, ISI, Web of Science, and Scopus, ensuring that all selected sources were relevant, credible, and of high academic quality. Ethical considerations were upheld throughout the research process by adhering to academic integrity standards, ensuring proper citation of all sources to avoid plagiarism. The analysis was conducted with transparency and objectivity, ensuring that personal biases did not influence the findings. Although the study involved secondary data sources and did not directly engage with human participants, it was mindful of the ethical implications surrounding the technologies discussed, particularly in terms of their impact on academic honesty and fairness.

V. Results

In examining the complex intersection of AI-driven plagiarism detection tools, academic integrity, and the shifting dynamics between journal editors and writers, several key themes and patterns emerged. These findings, drawn from an in-depth thematic and comparative analysis, provide insights into the evolving landscape of academic publishing in the context of rapid technological advancements. The themes not only shed light on the current practices and challenges but also offer a deeper understanding of how technology influences the roles, relationships, and ethical considerations within the academic community. Below are the answers to the three (3) research questions as cited in the introduction of the article, then followed by the ten (10) key themes identified through this study. Such, as discussed, are supported by existing scholarly literature and theoretical framework, followed by a comparative analysis of how plagiarism detection tools are utilized by journal editors versus actual writers.

1. **Contribution to Upholding Academic Integrity:** AI-driven plagiarism detection tools serve as powerful allies in upholding academic integrity by identifying potential instances of plagiarism, whether they involve directly copied text or paraphrased content. These tools compare students' submissions with an extensive database of sources, quickly highlighting any similarities. According to Chaka (2024), these tools play a pivotal role in the academic ecosystem by providing educators with quick and reliable data that helps determine whether a work is original. In this way, plagiarism detection tools maintain academic standards and ensure that research and assignments reflect authentic work. Quidwai, Li, and Dube (2023) suggest that

these tools have evolved to analyze not just individual sentences, but entire documents, offering more comprehensive checks for originality.

2. **Challenges and Limitations:** Despite their value, these tools come with significant limitations. One major challenge is the occurrence of false positives, where the system incorrectly flags content as plagiarized. This often occurs when the tools fail to account for context, such as academic citations, common academic phrases, or the use of shared knowledge. As Quidwai et al. (2023) note, tools often rely on pattern recognition, which can lead to misinterpretation of the text's intent. Similarly, Gao et al. (2022) found that plagiarism detection tools could overlook nuances, leading to missed instances of nuanced plagiarism or wrongly flagging innocuous text. Moreover, these tools often lack the human judgment needed to understand the context, which is crucial in determining whether an instance of similarity constitutes plagiarism. Santra and Majhi (2023) argue that as AI-generated content becomes more sophisticated, tools are increasingly struggling to distinguish between human-written and machine-generated text, further complicating the detection process.

3. **Impact on Non-Native Speakers:** Non-native speakers face unique challenges with plagiarism detection tools. These students may use language structures or phrases that unintentionally mirror existing texts, even when they intend to produce original work. Due to their reliance on text patterns, these tools may flag these instances as potential plagiarism, leading to unnecessary consequences. As noted by Hutson (2024), non-native speakers are often penalized for language errors or lack of fluency, which are part of the learning process, not academic dishonesty. This can result in unfair academic penalties that discourage students from improving their language skills. Moreover, when the tools fail to interpret the broader context of the writing, non-native students are at a greater disadvantage, potentially misjudged for unintentional errors.

In light of these challenges, it is clear that while plagiarism detection tools provide an important service, they should be used as supplementary aids rather than definitive arbiters of originality. Human judgment remains indispensable in assessing the full context of a student's work. Further research is needed to improve the accuracy of these tools, particularly in terms of understanding context and addressing the needs of non-native speakers. Enhancing these systems could lead to more equitable academic practices and ensure that plagiarism detection supports, rather than hinders, the integrity of academic work.

Thematic Analysis

As academic integrity becomes increasingly important in today's educational world, plagiarism detection tools have become a key resource for ensuring original work. However, while these tools are widely used, their effectiveness is not without debate. In this thematic analysis, the author explores the main themes and challenges surrounding AI-driven plagiarism detection, including the difficulties of context interpretation, issues with false positives, and how these tools affect non-native speakers. This examination seeks to better understand the role these tools play in maintaining academic standards and how they influence our perceptions of academic honesty.
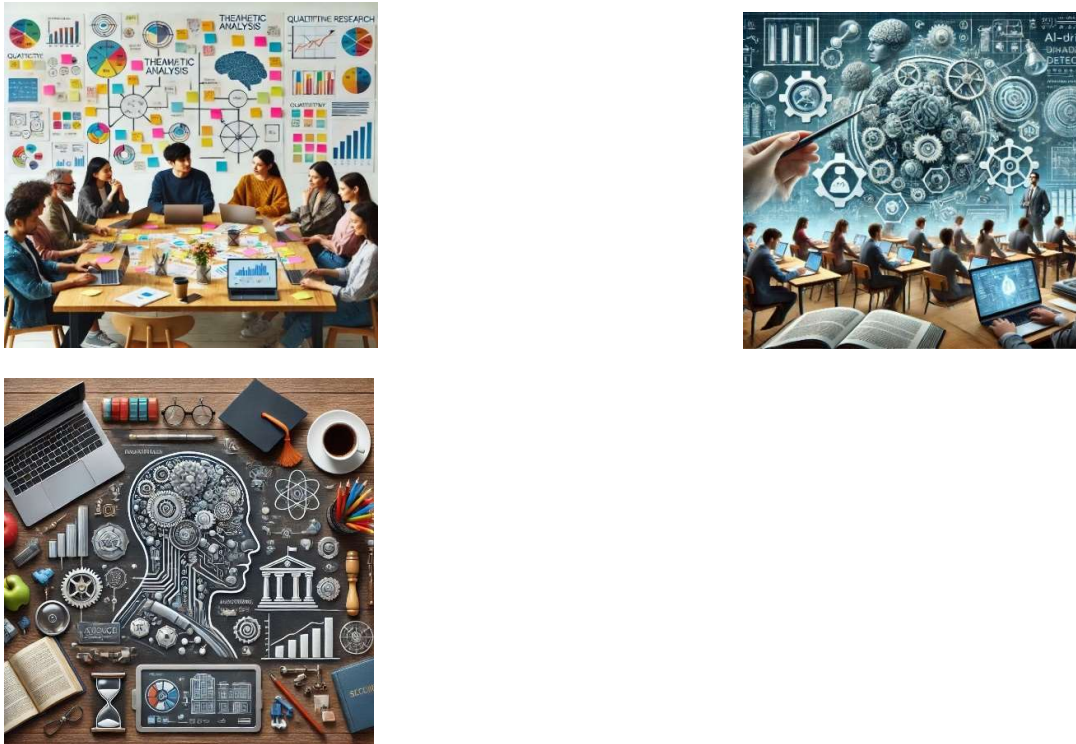
Figure 2: Thematic analysis. Microsoft. (n.d.). Copilot [AI assistant]**.**

1. **False Positives and the Challenge of Trust.** One of the most discussed issues within the realm of AI-driven plagiarism detection is the occurrence of false positives. False positives—instances where AI tools flag legitimate academic work as plagiarized—pose significant challenges to both writers and editors. Giray (2024) highlights that these tools often misinterpret complex academic writing as AI-generated, thereby undermining trust in the software. This aligns with Technological Determinism, which posits that technology, while intended to streamline processes, often shapes human behavior in unintended ways, creating a paradox where reliance on AI tools could undermine academic integrity.

2. **The Role of Academic Judgment in Detection.** The integration of AI detection tools into academic publishing has raised questions about the balance between machine-generated analysis and human judgment. Perkins et al. (2024) argue that combining AI with the expert judgment of educators and editors improves the reliability of plagiarism detection. This theme reflects Technological Determinism's influence, as AI tools are seen as shaping editorial practices, yet human oversight remains crucial to ensure accuracy in distinguishing between authentic academic work and potential misuse.

3. **Impact of Technology on Academic Integrity.** Technology has profoundly altered the way academic integrity is perceived and maintained. While AI-driven tools were developed to protect academic standards, they also raise concerns about over-reliance on technology. According to Davis (2024), the inclusion of AI in plagiarism detection is part of a broader move toward ensuring fairness in academic practices. However, this shift also risks reducing the human element of academic integrity, as

technological tools might encourage less personal engagement with the ethical standards underlying academic writing.

4. **Evolving Definitions of Plagiarism.** AI tools have forced a reevaluation of what constitutes plagiarism in academic writing. Amigud and Pell (2021) suggest that AI's growing ability to generate coherent and original text has complicated the boundaries between borrowing and plagiarism. In this context, academic institutions are adapting their policies to account for these new challenges, emphasizing the importance of transparency in using AI tools. The reliance on AI-driven systems to flag potential plagiarism has thus brought to the forefront the evolving nature of academic integrity.

5. **Challenges of Detecting Paraphrasing.** While AI tools are effective in identifying direct copying, they struggle with paraphrasing, which is often more nuanced. Walters (2023) underscores that AI lacks the ability to fully understand context, making it difficult to accurately flag paraphrased text. This gap in detection underscores the need for continuous refinement of AI systems. From a Technological Determinism perspective, this limitation reveals how technological advancements, while solving some problems, inevitably introduce new ones.

6. **The Ethics of AI-Generated Content.** As AI systems like GPT-4 become increasingly capable of generating academic content, questions arise about the ethical implications of such technology. Kendall and da Silva (2024) explore the potential misuse of AI-generated content in academic publishing, such as authorship manipulation or the use of paper mills. These concerns highlight the ethical gray areas created by technological tools, necessitating the development of clearer policies to manage AI usage in scholarly contexts.

7. **Educational Practices and Pedagogical Shifts.** The use of plagiarism detection tools has influenced educational practices, shifting the focus from teaching academic integrity to ensuring compliance with automated systems. Sefcik et al. (2020) emphasize that while these tools help maintain integrity, they might also encourage a more transactional view of learning, where students prioritize avoiding detection over engaging in genuine academic inquiry. This shift underscores the role of technology in shaping both teaching and learning dynamics.

8. **The Burden on Students and Writers.** Writers, particularly students, often bear the brunt of the pressure to avoid detection by plagiarism tools. Uzun (2023) argues that this focus on detection may lead students to engage in "game-playing," where the emphasis is on manipulating their work to avoid flags rather than understanding the ethical implications of their writing. This creates a tension between academic integrity and the tools designed to uphold it, suggesting that technology, while useful, also imposes limitations on student behavior.

9. **Technological Adaptability in the Face of New Forms of Plagiarism.** The continuous evolution of plagiarism methods, such as the rise of AI-generated content, demands that plagiarism detection tools remain adaptable. Gao et al. (2022) note that current tools are continually updated to address emerging forms of academic dishonesty. This adaptability is essential for maintaining the

effectiveness of detection systems, yet it also highlights the difficulty of keeping up with fast-evolving technology.

10. **International Perspectives on Plagiarism Detection.** The use of AI in plagiarism detection is not uniform across all institutions. A study by Quach (2023) showed that some leading universities, such as University of California, Oxford, and Harvard, have been hesitant to rely on AI tools for plagiarism detection, preferring instead to maintain traditional human oversight due to concerns about false positives and over-reliance on technology. Bretag et al. (2011) assert that these institutions prioritize academic judgment and the nuanced understanding that human reviewers bring to the process. This hesitation reflects a broader concern in the academic community about the consequences of placing too much trust in automated systems.

Comparative Analysis for Journal Editors Versus Writers

The comparative analysis of how plagiarism detection tools are used by journal editors versus writers revealed distinct differences in approach and outcomes. Editors tend to use these tools more systematically as part of a comprehensive review process, often relying on multiple software programs to ensure accuracy and prevent errors in the final publication. Writers, on the other hand, use these tools primarily as a precautionary measure to avoid unintentional plagiarism before submission. According to Perkins et al. (2024), editors possess a more in-depth understanding of academic integrity and the limitations of these tools, which allows them to apply their judgment alongside technological assistance. In contrast, writers are often less aware of the intricacies of plagiarism detection, relying on tools with the hope that they will safeguard their work. This discrepancy points to the broader influence of technological determinism, where editors' and writers' reliance on AI tools reflects the increasing dominance of technology in shaping academic practices.

VI. Findings

As the reliance on artificial intelligence (AI) for plagiarism detection in academic settings increases, several key themes and patterns have emerged from the literature. These findings underscore the complexities and potential challenges surrounding the integration of AI technologies in academic integrity practices. This section delves into the main themes identified in the literature, providing examples of false positives and contextual misunderstandings, while also discussing the impact of overreliance on AI in academic practices.

Key Themes and Patterns Identified

One of the most significant themes that emerged from the literature is the **accuracy and effectiveness** of AI-driven plagiarism detection tools. Scholars such as Mishra (2023) and Guo et al. (2023) highlighted concerns over the reliability of these tools, pointing out that they often struggle to differentiate between legitimate academic writing and plagiarized content, particularly when dealing with complex texts. For instance, AI tools sometimes flag original academic work as plagiarized due to the frequent use of technical jargon or commonly used phrases, a concern raised by Gasparyan et al., 2017). This can lead to unnecessary scrutiny, undermining the integrity of academic practices.

As pointed out by Elkhatat (2023), while tools such as Turnitin, PlagScan, Copyleaks, GPTZero, and QuillBot are widely used in academic settings to detect plagiarism and AI-generated content, their efficacy is often overestimated. These tools, despite their advanced algorithms, can still fail to accurately interpret context, leading to false positives and missed cases of nuanced plagiarism. Moreover, their overreliance on pattern recognition and lack of human judgment can undermine the very academic integrity they aim to protect. It is crucial to recognize that these tools should be used as supplementary aids rather than definitive arbiters of originality.

Another crucial theme identified was the **ethical implications of AI in academic settings**. Studies like those by Bailey (2022) and Frąckiewicz (2023) suggest that overreliance on AI can erode human judgment in assessing academic work, replacing critical thinking with algorithmic decisions. This shift from human to machine-generated assessments is seen as a potential threat to the holistic understanding of academic integrity, which involves not only detecting plagiarism but also considering context and intent (Moriarty & Wilson, 2022). Furthermore, the **biases embedded in AI algorithms** were a frequent concern, as highlighted by Bhargava and Rattani (2021), who argued that AI tools might be skewed against certain linguistic and cultural writing styles, leading to unfairly flagged submissions.

The theme of **AI as a tool for efficiency** also emerged strongly. According to Semeijn and Jaeger (2021), AI tools are seen as highly efficient for screening vast amounts of academic content quickly, which is a clear advantage in high-volume publishing environments. However, this efficiency comes at the expense of **contextual understanding**, a limitation discussed by Smith (2023), who emphasized that AI tools fail to grasp the nuances of academic writing and often make broad assumptions based on keywords alone. This leads to concerns about their appropriateness in handling more sophisticated forms of academic dishonesty.

Examples of False Positives and Contextual Misunderstandings

False positives are a major issue with AI plagiarism detection tools. One notable example highlighted by Guo et al. (2023) involved a case where a research paper, which included standard scientific terminology and widely accepted expressions, was flagged as plagiarized by AI tools. The tool failed to differentiate between what was common academic discourse and what might have been copied from another source. Similarly, Bailey (2022), discussed how AI systems sometimes flag entire sections of papers as plagiarized simply because they match phrases found in publicly available databases, even if these phrases are commonly used in the field.

Another example comes from the work of Semeijn and Jaeger (2021), who pointed out that AI tools often fail to account for the **context** in which similar phrases appear. For example, an academic who paraphrases or synthesizes information may find their work flagged by AI systems as plagiarized, even though the synthesis was done properly. These systems struggle with recognizing subtle shifts in meaning or the transformation of ideas into original content. Such **contextual misunderstandings** are particularly problematic in disciplines that require deep critical analysis and reflection, such as the humanities, where originality is not always about phrasing but about the underlying interpretation of ideas (Frąckiewicz, 2023).

Overreliance on AI and Its Impact on Academic Practices

Overreliance on AI tools has profound implications for academic practices. While AI technologies like plagiarism detectors offer convenience and efficiency, they may inadvertently reduce the role of human judgment in assessing academic work. As Bhargava and Rattani (2021) noted, AI tools may encourage **a checklist mentality** among educators and journal editors, where the focus shifts from evaluating the integrity of the work to simply passing it through an algorithm. This could lead to a **false sense of security** and undermine the purpose of academic integrity, which involves understanding the broader context and intellectual contributions behind a piece of work (Mishra, 2023).

Furthermore, the **shift in power** from educators to machines could diminish opportunities for students to learn from their mistakes. Instead of engaging in meaningful discussions about their academic practices, students may find themselves simply complying with automated systems, without a true understanding of what constitutes ethical writing (Smith, 2023). This shift could also affect faculty roles, as they may become more reliant on AI for final decisions about academic integrity, further distancing themselves from the nuanced work of evaluating content.

The **impact on learning outcomes** is another important consideration. As noted by Çelik and Razı (2023), the use of AI tools might discourage students from developing their writing and critical thinking skills. Relying on AI for plagiarism detection may make students less engaged in

understanding the principles of academic integrity, potentially leading to a generation of students who value technological compliance over intellectual honesty.

Overall, while AI-driven plagiarism detection tools bring certain efficiencies to academic integrity practices, they also introduce challenges such as false positives, contextual misunderstandings, and an overreliance on technology. These issues have the potential to reshape academic practices, diminishing the importance of human judgment and critical engagement in evaluating academic work. As universities continue to integrate AI technologies into their practices, it is essential to balance technological advancements with a commitment to maintaining the ethical foundations of academic integrity.

VII. Conclusion and recommendation

This study openly examined the role of AI-powered plagiarism detection tools in higher education, emphasizing their potential benefits and limitations. While tools like Turnitin, PlagScan, GPTZero, and QuillBot etc. are commonly used, their effectiveness can be overstated due to issues such as false positives and misinterpretation of context. These tools should be seen as supplementary aids, not the ultimate arbiters of originality, as human judgment remains essential in upholding academic integrity. In response to the question of whether AI-driven plagiarism detection tools can uphold academic integrity without ethical compromises, the answer is nuanced. Based on the salient findings from the saturated data, these tools, while useful, cannot be used as the ultimate arbitral judge in determining or ruling out if a certain paper is AI-generated or plagiarized. While these tools offer valuable support in identifying potential instances of plagiarism, their reliance on algorithms, without human oversight, introduces the risk of ethical compromises. False positives and contextual errors can lead to unfair outcomes, making it clear that AI alone cannot ensure academic integrity without ethical concerns. Theoretical frameworks such as Technological Determinism, Actor-Network Theory (ANT), and Socio-Technical Systems Theory (STS) helped provide insight into how AI technologies reshape academic practices, emphasizing the need for a balanced approach that integrates both technological capabilities and human oversight. The qualitative methodology employed in this study revealed the challenges faced by journal editors and writers when using these tools, further underlining the risks of overreliance on technology. Furthermore, it is recommended that future research can explore hybrid models that combine AI tools with human judgment to improve plagiarism detection accuracy. Deeper investigations into the ethical implications of AI use in academia and its impact on academic culture are also necessary. Overall, while AI tools can play a significant role in supporting academic integrity, their integration must be carefully managed, with human oversight ensuring that core educational values and ethical standards are upheld.

Declarations:

References

1. Amigud, A., & Pell, D. J. (2021). When academic integrity rules should not apply: a survey of academic staff. Assessment & Evaluation in Higher Education, 46(6), 928-942.

2. Bailey, J. (2022). The Growth of AI in Plagiarism Detection. Plagiarism Today. Moya, Beatriz A., et al. (2023). Academic Integrity and Artificial Intelligence in Higher Education Contexts: A Rapid Scoping Review Protocol. Canadian Perspectives on Academic Integrity, 5(2). doi: https://doi.org/10.11575/cpai.75990

3.   Bhargava, A., & Rattani, A. (2021). Plagiarism detection: Traditional techniques and recent advances. In Data Analytics for Enhanced Educational Outcomes, 211-227.

4.   Bretag, T., Mahmud, S., Wallace, M., Walker, R., James, C., Green, M., ... & Patridge, L. (2011). Core elements of exemplary academic integrity policy in Australian higher education. International Journal for Educational Integrity, 7(2).

5.   Çelik, Ö., Razı, S. Facilitators and barriers to creating a culture of academic integrity at secondary schools: an exploratory case study. Int J Educ Integr 19, 4 (2023). https://doi.org/10.1007/s40979-023-00125-4

6.   Chaka, C. (2024). Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review. Journal of Applied Learning and Teaching, 7(1).

7.   Dalalah, D., & Dalalah, O. M. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. The International Journal of Management Education, 21(2), 100822.

8.   Davis, M. (2024). Inclusion Within a Holistic Approach to Academic Integrity: Improving Policy, Pedagogy, and Wider Practice for All Students. In Second Handbook of Academic Integrity (pp. 1129-1147). Cham: Springer Nature Switzerland.

9.   Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. International Journal for Educational Integrity, 19(1), 17 https://doi.org/10.1007/s40979-023-00140-5

10. Frąckiewicz, M. (2023). The Influence of AI on Academic Integrity and Plagiarism Detection. Retrieved 12, May 2023, from https://ts2.space/en/the-influence-of-ai-onacademic-integrity-and-plagiarism-detection

11. Francke, E., & Bennett, A. (2019, October). The potential influence of artificial intelligence on plagiarism: A higher education perspective. In European conference on the impact of artificial intelligence and robotics (ECIAIR 2019) (Vol. 31, pp. 131-140).

12. Gall, T., & Maniadis, Z. (2019). Evaluating solutions to the problem of false positives. Research Policy, 48(2), 506-515.

13. Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. BioRxiv, 2022-12.

14. Gasparyan, A. Y., Nurmashev, B., Seksenbayev, B., Trukhachev, V. I., Kostyukova, E. I., & Kitas, G. D. (2017). Plagiarism in the context of education and evolving detection strategies. Journal of Korean medical science, 32(8), 1220-1227.

15. Giray, L. (2024). The Problem with False Positives: AI Detection Unfairly Accuses Scholars of AI Plagiarism. The Serials Librarian, 1-9.

16. Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., Yue, J., & Wu, Y. (2023). How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection. ArXiv. https://arxiv.org/abs/2301.07597

17. Hennink, M., Hutter, I., & Bailey, A. (2020). Qualitative research methods. Sage.

18. Hutson, J. (2024). Rethinking Plagiarism in the Era of Generative AI. Journal of Intelligent Communication, 4(1), 20-31.

19. iStock. (n.d). Stock illustration. https://www.istockphoto.com/vector/seo-reporting-data-monitoring-web-traffic-analytics-big-data-flat-vector-gm1041244364-278766086

20. Kendall, G., & da Silva, J. A. T. (2024). Risks of abuse of large language models, like ChatGPT, in scientific publishing: Authorship, predatory publishing, and paper mills. Learn. Publ., 37(1), 55-62.

21. Microsoft. (n.d.). Copilot [AI assistant]. Microsoft. https://www.microsoft.com/copilot

22. Mishra, S. (2023). Enhancing Plagiarism Detection: The Role of Artificial Intelligence in Upholding Academic Integrity. Library Philosophy and Practice (e-journal). 7809. https://digitalcommons.unl.edu/libphilprac/7809

23. Moriarty C, Wilson B (2022) Justice and consistency in academic integrity: Philosophical and practical considerations in policy making. Journal of College and Character 23(1):21–31. https://doi.org/10.1080/2194587X.2021.2017971

24. Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2024). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. Journal of Academic Ethics, 22(1), 89-113.

25. Pudasaini, S., Miralles-Pechuán, L., Lillis, D., & Salvador, M. L. (2024). Survey on AI-Generated Plagiarism Detection: The Impact of Large Language Models on Academic Integrity. Journal of Academic Ethics. https://doi.org/10.1007/s10805-024-09576-x

26. Quach, K. (2023). Some universities reject Turnitin's AI-writing detector over fears it'll wrongly accuse students. The Register. Retrieved from https://www.theregister.com/ 2023/09/23/turnitin

27. Quidwai, A., Li, C., & Dube, P. (2023). Beyond Black Box AI generated Plagiarism Detection: From Sentence to Document Level. Association for Computational Linguistics.

28. Santra, P. P., & Majhi, D. (2023). Scholarly communication and machine-generated text: is it finally ai vs ai in plagiarism detection?. Journal of Information and Knowledge, 175-183.

29. Savin-Baden, M., & Major, C. (2023). Qualitative research: The essential guide to theory and practice. Routledge.

30. Sefcik, L., Striepe, M., & Yorke, J. (2020). Mapping the landscape of academic integrity education programs: what approaches are effective?. Assessment & evaluation in higher education.

31. Semeijn, J. H., & Jaeger, M. (2021). Technology for promoting academic integrity: Current state and future directions. Frontiers in Education, 6, 648120

32. Smith, J. A. (2023). AI and Academic Integrity: Navigating the New Landscape of Plagiarism Detection. Academic Press.

33. Uzun, L. (2023). ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content. Language Education and Technology, 3(1).

34. Walters, W. H. (2023). The effectiveness of software designed to detect AI-generated writing: A comparison of 16 AI text detectors. Open Information Science, 7(1), 20220158.