TOWARD A NEW APPROACH FOR EXAMINING ACADEMIC CITATION PRACTICES: ANALYZING CITATION PATTERN MATCHES TO DETECT STRUCTURAL AND PARAPHRASED PLAGIARISM

Salah S. Abd El-Ghani[1], Tamer Gamal Ibrahim Mansour[2*]

[1]*Department of Economics, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia.*
[2]*Department of Agriculture Economics, Agricultural and Biological Research Institution, National Research Centre, Giza, Egypt.*

Corresponding Author: Tamer Gamal Ibrahim Mansour
(tj.mansour@nrc.sci.eg, tamer_baz@yahoo.com)

ABSTRACT

The field of academic integrity is undergoing a profound transformation in the nature of scholarly plagiarism, particularly with the widespread use of paraphrasing tools and generative artificial intelligence. These technologies can now produce texts that differ linguistically from their sources while preserving the original conceptual structure. Traditional detection systems based on textual similarity, such as Turnitin and iThenticate, have shown clear limitations in identifying this emerging form of structural and paraphrased plagiarism, largely because they rely on lexical overlap rather than the ability to trace the intellectual framework of an argument. This review study aims to analyze the development trajectory of plagiarism detection tools over the past decade and to identify the methodological gaps that limit their ability to address structural plagiarism. It also evaluates the potential of the Citation Pattern Matching approach as an alternative and more reliable method for uncovering structural relationships between texts. The study relies on a systematic review of the literature from 2014 to 2024 using the PRISMA framework and applies precise criteria to select 33 studies distributed across three themes: text-matching tools, citation-based and bibliometric analysis, and semantic and AI-driven models. The results indicate that text-based tools perform well in detecting surface similarity but fail to uncover deep paraphrasing that preserves the original text's referential structure and logical sequence. By contrast, the literature shows that citation analysis, including citation sequencing, repetition of core references, and network relationships among sources, represents a cognitive fingerprint that is difficult to manipulate and proves more effective in identifying structural plagiarism even when no clear linguistic similarity exists. The review also highlights the need for hybrid models that integrate textual, semantic, and citation-based approaches to develop a new generation of detection systems capable of addressing increasingly sophisticated forms of plagiarism. The study concludes that relying exclusively on lexical similarity ratios is no longer sufficient in the era of generative artificial intelligence, and that analyzing citation structures offers a more rigorous and reliable framework for assessing the originality of scientific texts. It proposes adopting Citation Pattern Analysis as a central step toward strengthening academic integrity and developing detection systems that keep pace with rapid transformations

in knowledge production.

1. INTRODUCTION

Academic integrity has faced growing challenges in recent years due to the widespread emergence of modern forms of plagiarism, particularly paraphrased plagiarism, which is far more complex than traditional direct copying. While verbatim plagiarism can be detected with relative ease, paraphrased plagiarism reconstructs the logical structure, sequence of ideas, and core references of the original text while altering only the linguistic surface. This makes it significantly more capable of evading conventional detection tools. The literature indicates that this form of plagiarism has become increasingly common as a result of academic pressures, limited awareness of the boundaries of legitimate citation, and insufficient training in scholarly writing skills (Alhadlaq et al., 2020; Selemani et al., 2018; Welsh, 2019). Other studies emphasize that researchers' and students' misunderstanding of the concept of scientific originality often leads them to adopt flawed paraphrasing practices, frequently without recognizing their ethical implications (Newton et al., 2014).

Despite the widespread reliance on text-matching tools such as Turnitin and iThenticate, their ability to detect structural plagiarism remains limited, as they focus primarily on lexical similarity rather than the underlying conceptual organization of the text. Systematic reviews and extensive empirical studies have shown how easily such tools can be bypassed through paraphrasing or syntactic restructuring without altering the internal logic of the argument (Foltýnek et al., 2019; Higgins et al., 2016; Gallant et al., 2019). Recent research has also demonstrated that the sensitivity of these systems fluctuates over time due to continual algorithmic updates, resulting in inconsistent outputs that cannot be reliably used to detect structural plagiarism (Habibzadeh, 2023).

The emergence of generative artificial intelligence has further complicated the problem, as modern language models have become capable of producing high-quality texts that differ linguistically from the original sources while preserving their argumentative structure and internal logic. Evidence shows that traditional text-based tools fail to detect this pattern because AI systems reproduce ideas and reference structures without generating direct lexical overlap (Moravvej et al., 2023). With the growing availability of AI-neutralizing tools that obscure the stylistic fingerprint of machine-generated texts, identifying structural plagiarism has become even more challenging.

In light of this, the problem addressed in this research lies in the significant limitations of text-matching tools in detecting structural and paraphrased plagiarism, particularly in the era of generative AI. This calls for the development of alternative approaches that focus on the scientific architecture of a text rather than its linguistic surface. The literature reveals a clear research gap stemming from the absence of a unified method capable of tracing the citation structure of texts, despite substantial evidence demonstrating the effectiveness of citation-based analysis in tracking citation sequences, their density, and co-citation networks, all of which have shown strong potential for detecting plagiarism even when the linguistic expression is entirely altered (Gipp and Beel, 2010; Meuschke and Gipp, 2013). Accordingly, this review study aims to analyze the scientific literature on structural and paraphrased plagiarism, to evaluate the shortcomings of text-matching tools, and to explore the potential of Citation Pattern Matching as a more reliable alternative for detecting plagiarism in a knowledge environment that is rapidly evolving under the influence of artificial intelligence. The introduction concludes by noting that the study will examine the relevant literature, highlight the existing gaps, and present a new analytical framework capable of enhancing the accuracy and reliability of plagiarism detection in contemporary academic contexts.

2. THEORETICAL BACKGROUND

2.1. *The Concept of Scientific Citation and Its Importance in Knowledge Construction*

Scientific citation is one of the fundamental pillars of academic research, serving to document sources and situate new work within an existing body of knowledge while ensuring respect for intellectual property rights. Researchers are expected to engage in citation as an expression of scholarly integrity rather than as a mechanism for concealing plagiarism or reproducing others' ideas without proper attribution. The academic integrity literature indicates that a lack of precise understanding of citation and its boundaries is one of the main factors contributing to the spread of plagiarism, especially among students and early-career researchers who often confuse legitimate citation with unethical paraphrasing (Alhadlaq et al., 2020; Selemani et al., 2018). Behavioral studies also show that such practices are frequently linked to limited awareness or insufficient training in academic writing skills and in the proper techniques of paraphrasing (Newton et al., 2014).

2.2. Forms of Academic Plagiarism and Its Contemporary Evolution

Academic plagiarism takes multiple forms, beginning with direct verbatim copying and extending to the use of sources without attribution, eventually reaching its most complex manifestations: structural plagiarism and paraphrased plagiarism. In the latter, the plagiarist preserves the original text's logical structure, sequence of arguments, and core references while altering only the linguistic expression, which makes detection highly challenging. Recent literature confirms the growing prevalence of this pattern due to academic pressure, the availability of automatic rewriting tools, and the increasing reliance of researchers on artificial intelligence for content generation (Gasparyan et al., 2017; Zimba and Gasparyan, 2021). This shift in the nature of plagiarism indicates a transition from the level of "lexical similarity" to the level of "replicating scientific structure," a phenomenon that requires detection tools that go beyond traditional approaches.

2.3. Text-Matching Detection Tools and Their Methodological Limitations

For decades, academic institutions have relied on text-matching tools such as Turnitin, iThenticate, and PlagScan, which compare submitted texts against extensive databases to identify lexical similarity. Although these tools are widely used and considered central to publication and peer-review policies, their effectiveness decreases substantially when confronted with structural or paraphrased plagiarism. Systematic reviews and large-scale evaluations have shown that most commercial systems can be easily bypassed by rephrasing sentences, using synonyms, or altering stylistic patterns without affecting the scientific logic of the text (Foltýnek et al., 2019; Foltýnek et al., 2020; Gallant et al., 2019). Higgins et al. (2016) also demonstrated considerable variation in the performance of these tools across different disciplines, while Habibzadeh (2023) pointed out the absence of an agreed-upon global similarity threshold, which complicates the assessment of originality and undermines the reliability of text-based detection results.

The limitations of these systems have become even more pronounced with the rapid development of generative language models, such as GPT, which produce high-quality texts that show no lexical overlap with original sources yet replicate their analytical structure and intellectual sequencing.

Evidence indicates that text-matching tools fail to detect this form of plagiarism because AI systems do not "copy sentences" but instead "replicate reasoning" (Moravvej et al., 2023). The challenge is further exacerbated by the rise of AI-neutralizing tools that remove stylistic traces from machine-generated content, making lexical similarity a less effective indicator than ever before.

2.4. Citation-Based Analysis as an Alternative Approach for Detecting Structural Plagiarism

Given the evident limitations of text-based detection tools, researchers have turned to more advanced approaches grounded in bibliometric analysis, which examines the referential relationships embedded within texts rather than focusing on their linguistic surface. Foundational contributions such as those of Gipp and Beel (2010) and Meuschke and Gipp (2013) have demonstrated the ability of citation-based analysis to detect structural plagiarism by comparing citation sequences, their density, the identity of core references, and patterns of clustered citations. This approach is based on the idea that citation structure functions as a cognitive fingerprint that is difficult to fabricate, since individuals who paraphrase stolen material typically preserve the original referential architecture even when they alter the language.

For this reason, Citation Pattern Matching is considered one of the most promising pathways for developing the next generation of plagiarism detection systems. Its strength lies in its capacity to trace the deep knowledge relationships that remain intact despite extensive rewriting, making it particularly well suited to the era of artificial intelligence and the increasing reliance on paraphrasing tools.

3. RESEARCH GAPS

Despite notable advancements in plagiarism detection tools over the past two decades, the scientific literature shows that current systems remain unable to effectively address structural and paraphrased forms of plagiarism. This limitation is rooted in several deep research gaps that constrain the ability of traditional lexical approaches to detect the cognitive replication of texts, as outlined in the following themes:

3.1. Limitations of Text-Matching Tools in Detecting Structural Plagiarism

Most detection systems rely on text-matching techniques that analyze lexical similarity, which makes them vulnerable to failure when texts are

paraphrased in ways that alter their linguistic surface without affecting the underlying argumentative structure. Several systematic studies have confirmed that commercial tools, even the more advanced ones, fail to identify plagiarism after paraphrasing because their algorithms cannot analyze the internal structure of a text or the logic of its argument (Foltýnek et al., 2019; Foltýnek et al., 2020). Editorial analyses from academic journals have likewise indicated that a considerable proportion of manuscripts falling within "acceptable similarity thresholds" may still contain undetected forms of plagiarism, since lexical resemblance does not always reflect the true originality of the content (Higgins et al., 2016; Habibzadeh, 2023).

3.2. Lack of Capacity to Analyze Citation Structure

Citation architecture, which includes citation sequencing, citation density, and the relationships among referenced sources, represents one of the most important structural indicators of a text's originality. However, commercial detection tools do not support the analysis of such referential patterns. Foundational bibliometric studies demonstrate that citation patterns reveal similarities that are difficult to fabricate, and that individuals who paraphrase stolen material often retain the core references or the order in which they appear within the scientific argument (Gipp and Beel, 2010; Meuschke and Gipp, 2013). Despite the strength of this approach, it has yet to be integrated into widely used detection systems.

3.3. Absence of a Theoretical Framework Linking Citation Behavior to Cognitive Structure

Bibliometric research indicates that each scholar has a distinctive "citation fingerprint" that reflects their intellectual methodology and the way they construct scientific arguments. However, current detection tools do not exploit this dimension in identifying plagiarism, and no standardized models exist that can distinguish "unnatural" similarities in referential networks across texts, particularly in fields where researchers frequently draw on shared foundational literature.

3.4. The Impact of Generative Artificial Intelligence on the Evolution of Structural Plagiarism

Generative artificial intelligence has led to the emergence of texts that are "formally different yet cognitively identical," in which ideas, logic, and analytical sequencing are reconstructed without any evident lexical overlap. Studies indicate that traditional text-based detection tools fail easily to identify such texts, especially with the growing use of AI-neutralizing tools that erase stylistic traces from machine-generated content (Moravvej et al., 2023). This shift presents a fundamental challenge that weakens the effectiveness of models based on linguistic similarity.

3.5. Absence of a Standard Framework to Distinguish Legitimate Similarity from Structural Plagiarism

In many scientific disciplines, core references, theoretical frameworks, and citation positions naturally overlap, making it difficult to distinguish between legitimate similarity that arises from the nature of the field and structural similarity resulting from plagiarism. Several studies have highlighted the need for a standardized framework capable of differentiating between the two, particularly in fields with limited or highly overlapping bodies of literature (Rohwer et al., 2018; Beketov and Lebedeva, 2022).

The five gaps outlined above show that the problem no longer lies in lexical similarity but in the argumentative and cognitive structure of scientific writing. This makes reliance on traditional text-based tools of limited value when addressing structural plagiarism. Consequently, there is a clear need to adopt a new analytical approach centered on Citation Pattern Matching, which offers the strongest potential for tracing the stable referential architecture of texts in the era of generative artificial intelligence.

4. THE PROPOSED CONCEPT: CITATION PATTERN MATCHING

Citation Pattern Matching represents a promising approach for overcoming the structural limitations of traditional text-based detection tools by shifting from measuring lexical similarity to analyzing the cognitive architecture of a text's scientific argument. This approach is built on a central assumption: a researcher who paraphrases a plagiarized text may succeed in altering its language and stylistic features, yet will often fail to reconstruct its citation structure, which includes the sequence of references, the density of citations, their locations within the text, and the networked relationships among sources. Foundational studies in this field have demonstrated that citation patterns function as a "cognitive fingerprint" that is difficult to falsify even after deep paraphrasing (Gipp and Beel, 2010; Meuschke and Gipp, 2013). This concept is grounded in comparing the sequence of citations within articles, as citation sequencing is the element most capable of detecting structural plagiarism. Unlike the linguistic surface,

which can be easily rewritten, the order of citations reflects the logic of the argument: how ideas progress, and how each source is invoked to support a particular segment of analysis. Scientific evidence shows that copied or paraphrased articles often display substantial similarity in citation order even when the vocabulary and sentence structures differ (Foltýnek et al., 2019). Moreover, the density of citations in specific sections such as the introduction or theoretical framework tends to be a stable pattern that reappears in both the original and the plagiarized text, making structural plagiarism identifiable through tracking this referential configuration.

The main elements monitored within this approach include:

1. The order of citations, where the relative position of each reference in both texts is compared. This element has proven effective in detecting non-verbal forms of plagiarism (Gipp and Beel, 2010).
2. The density of citations across paragraphs and sections, as unjustified similarities may indicate structural copying.
3. The repetition of core references, since plagiarized texts often rely repeatedly on the same foundational sources.
4. The placement of citations within the argument structure, which serves as a precise indicator because the location of a citation corresponds to the cognitive role of the source within the analysis.
5. The networked relationships among references through co-citation analysis and bibliographic coupling, both of which are powerful tools for identifying unnatural similarities between texts (Meuschke and Gipp, 2013).

The strength of this approach lies in its focus on the cognitive structure of the text rather than its linguistic surface, which allows it to detect disguised forms of plagiarism that escape text-matching tools. It also performs effectively with AI-generated texts, which often display consistency in core references or referential relationships despite differences in wording (Kulkarni et al., 2021; Moravvej et al., 2023). While generative language models can produce new sentences that bear no lexical resemblance to the original text, they frequently reproduce the same conceptual architecture, particularly when the researcher uses academic material as input for generation. Thus, Citation Pattern Matching constitutes a methodological step forward that transcends the limits of conventional tools and provides a more rigorous and reliable framework for addressing structural plagiarism. It not only offers a more accurate mechanism for detection but also paves the way for developing hybrid systems that combine textual, semantic, and citation-based analysis, enabling academic journals and universities to assess the originality of scientific texts with greater precision in an era marked by increasing reliance on paraphrasing technologies and generative artificial intelligence.

5. LITERATURE REVIEW

Scientific literature over the past two decades reveals remarkable progress in plagiarism detection methodologies, yet it also highlights a widening gap between the nature of modern plagiarism, especially paraphrased and structural forms, and the capabilities of the tools traditionally used in universities and academic journals. Most classical studies concentrated on text matching, a method limited to comparing words and phrases while disregarding the logical structure of scientific arguments. The review by Foltýnek et al. (2019) stands out as one of the most comprehensive works in this area. It showed that most commercial detection tools measure only surface-level similarity and that their ability to detect paraphrased plagiarism remains weak, even when supplemented with additional semantic algorithms. A subsequent evaluation conducted across multiple environments confirmed that advanced paraphrasing techniques can significantly reduce the accuracy of these tools (Foltýnek et al., 2020), indicating the need for approaches that move beyond textual features toward citation structure. In parallel, an advanced research direction has emerged that relies on bibliometrics and citation analysis to identify structural plagiarism. The work of Gipp and Beel (2010) marked a turning point in this field by demonstrating, for the first time, the feasibility of using citation sequence analysis to detect plagiarism even when the wording is entirely altered. Meuschke and Gipp (2013) expanded this approach by developing models based on analyzing referential networks, such as co-citation, bibliographic coupling, and citation density. These studies showed that structural similarity in citation patterns can serve as a strong indicator of an abnormal relationship between texts, particularly in cases where no direct linguistic similarity is present. These findings collectively confirm that citation-based analysis provides a detection lens that surpasses the structural limitations of text-matching tools.

Recent literature also points to the rise of hybrid approaches that integrate semantic, network, and

citation-based analyses. Moravvej et al. (2023) proposed a model that relies on deep BERT language representations combined with attention mechanisms to detect paraphrased plagiarism, while other studies, such as Kamat et al. (2024), suggest using advanced semantic models that trace the "structural meaning" of text rather than relying solely on lexical resemblance. Despite the noticeable improvement in the performance of these models, many still struggle to detect cognitive equivalence between texts, especially when the referential structure is preserved.

In another strand of research, several studies have examined researchers' and students' citation and plagiarism behaviors, offering an important perspective on the social and educational factors driving the prevalence of plagiarism. Evidence shows that low awareness of proper academic writing practices is a major factor leading students into plagiarism, particularly when they depend on paraphrasing without a sound understanding of publication ethics (Alhadlaq et al., 2020; Selemani et al., 2018; Welsh, 2019). Gallant et al. (2019) further demonstrated that similarity reported by detection tools may sometimes reflect poor writing skills rather than intentional plagiarism, underscoring the need for more precise analytical tools capable of distinguishing legitimate from illegitimate similarity.

Consistent with these observations, several editorial studies have examined the challenges faced by academic journals in interpreting detection results. Higgins et al. (2016) documented considerable variation among editors in how similarity scores are interpreted, while Habibzadeh (2023) emphasized the absence of a global standard for acceptable similarity thresholds. This lack of consensus complicates the peer-review process and reinforces the need for tools that rely on the scientific structure of a text rather than its verbal form. Additional research indicates that the technical limitations of text-matching tools conflict with recent developments in generative artificial intelligence. Several studies show that AI-generated texts, or those "refined" through fingerprint-neutralizing tools, often pass detection systems effortlessly despite lacking genuine intellectual originality (Kulkarni et al., 2021). This qualitative shift has placed scientific research in a new dilemma that requires fundamentally rethinking detection mechanisms.

Taken together, the literature reveals that the most promising and robust direction lies in analyzing citation pattern similarity, which examines referential relationships rather than textual features

and enables the development of more accurate algorithms for identifying intellectual originality. These works form the conceptual foundation upon which the model proposed in this study is built, serving as a direct response to real and well-documented gaps in the field.

6. METHODOLOGY

This review study was built on a methodological framework that combines a systematic literature review with an analytical narrative synthesis, in alignment with PRISMA 2020 principles to ensure transparency and reproducibility. The study was guided by a central research question that sought to determine the extent to which existing literature falls short in addressing structural and paraphrased plagiarism, and to explore whether citation pattern analysis offers a methodological alternative capable of filling these gaps. This question serves as the conceptual anchor linking traditional approaches with the analytical model proposed in this study.

A structured search was conducted across major databases, including Scopus, Web of Science, IEEE Xplore, SpringerLink, Wiley, Taylor & Francis, PubMed, SSRN, and arXiv. The search employed a set of compound keywords covering textual plagiarism, structural plagiarism, citation analysis, and the influence of artificial intelligence, using logical operators to ensure both breadth and precision. Clear inclusion criteria were applied, focusing on peer-reviewed studies published between 2014 and 2024 that directly addressed text-based plagiarism detection tools, structural and paraphrased plagiarism, bibliometric analysis, or the impact of artificial intelligence on academic integrity. Studies were excluded if they were non-peer-reviewed, relied exclusively on lexical matching, or lacked a coherent methodological framework.

The screening process followed PRISMA procedures rigorously. After removing duplicates, 121 studies were identified, of which 72 were excluded during the initial screening due to lack of relevance. Full-text assessment was conducted for 49 studies, and 16 were excluded due to methodological shortcomings or limited relevance to the proposed model. This process resulted in a final selection of 33 studies that form the core analytical foundation of the review.

The systematic analysis showed that these 33 studies can be organized into three major thematic clusters reflecting the evolution of plagiarism detection efforts. The first cluster focuses on evaluating traditional text-matching tools and includes 12 studies examining the performance of

systems such as Turnitin and iThenticate, revealing the limitations of lexical approaches when confronting paraphrased plagiarism. The second cluster contains 8 studies centered on citation and bibliometric analysis, emphasizing methodologies such as citation sequencing, co-citation, and bibliographic coupling, all of which demonstrated strong potential in capturing the referential structure of texts with greater accuracy. The third cluster consists of 13 studies addressing structural plagiarism, semantic models, and artificial intelligence, including BERT-based models and modern generative systems, which collectively illustrate the growing complexity of challenges faced by detection frameworks.

Strict quality assurance measures were applied, including double-review checks and verification of the authenticity of all DOIs. Employing the PRISMA framework helped ensure a high level of methodological transparency and rigor, meeting the standards required by indexed journals by clarifying the steps of selection and screening, reducing bias, and enhancing reproducibility.

7. RESULTS

The systematic review of the thirty-three studies that met the inclusion criteria shows a clear shift in plagiarism-detection research away from the limitations of traditional text-matching tools and toward deeper models that engage with the cognitive structure of a text rather than its verbal surface alone. Evidence drawn from 12 studies evaluating text-matching systems such as Turnitin and iThenticate indicates that these tools can detect surface-level similarity but fail to identify paraphrased or linguistically disguised plagiarism, even when sophisticated semantic techniques are employed. This limitation appeared consistently across editorial, medical, and academic contexts. Findings reported in ACM Computing Surveys, Research Integrity and Peer Review, and BMJ Open revealed that similarity scores produced by these tools are often misleading, and that human judgment, despite its potential biases and imperfections, remains indispensable in the final verification stage. By contrast, the eight studies addressing citation-based and bibliometric analysis highlighted a qualitative shift in how the problem is conceptualized. The work of Gipp and his colleagues demonstrated that citation sequencing, co-citation patterns, and bibliographic coupling form a stable cognitive fingerprint that structural plagiarism struggles to conceal even when the wording is entirely different. The success of reference-network analysis techniques suggests that the referential architecture of a text functions as a more manipulation-resistant intellectual trace than its linguistic surface, positioning it as a strong methodological candidate for next-generation detection models.

The third group of studies, consisting of thirteen papers focused on structural plagiarism and artificial intelligence, confirmed that large language models such as BERT, Transformers, and GPT have redefined the boundaries of the problem. These models generate coherent texts that bear little detectable lexical resemblance to the original materials, yet they often preserve the underlying conceptual structure and maintain similar referential relationships. Analyses published in PNAS, TechTrends, and arXiv reports demonstrated that even AI-text detection tools can be circumvented easily through "style smoothing," while the cognitive core of the text remains stable enough to be identified through citation-pattern analysis.

A holistic reading of these three lines of inquiry leads to a central conclusion. Text-matching tools, regardless of their degree of sophistication, no longer provide an adequate framework for addressing plagiarism in an era shaped by generative writing technologies and advanced paraphrasing. The findings underscore that citation patterns and referential relationships offer a more stable and reliable structural alternative because they move beyond linguistic form toward the argumentative architecture and the knowledge trajectory on which the text is built. Accordingly, the evidence synthesized from the included studies supports the development of a hybrid model that integrates textual, semantic, and citation-based analysis, leveraging network analysis of referential relationships to detect structural plagiarism that remains invisible to lexical-similarity indicators. This convergence in the literature shows that citation pattern analysis is not merely a supplementary technique.

Instead, it represents a crucial theoretical component in building a new generation of detection systems capable of distinguishing between genuine scholarly contributions and manipulated or AI-generated texts that conceal plagiarism at the linguistic surface while reproducing it at the structural level.

This table provides a comprehensive analytical overview that illustrates how the literature converges across three complementary trajectories: lexical similarity detection, structural citation-based analysis, and semantic approaches grounded in artificial intelligence. The distribution indicates that

relying solely on text-matching tools is no longer sufficient, and that referential analysis and citation-pattern structures represent the key turning point for detecting structural plagiarism in the era of generative models.

Table 1: Summary Matrix of the 33 Studies across the Three Review Strands.

Strand	Study (Author & Year)	Indexing Database	Method / Tool	Key Findings
1. Evaluation of Text-Matching Detection Tools (12 studies)				
1	Higgins et al. (2016)	Springer / Scopus	Editorial analysis and system testing	Similarity scores do not reflect structural plagiarism and require human oversight.
2	Foltýnek et al. (2020)	Scopus	Testing 15 detection tools	Tools capture similarity but not plagiarism; weak performance with paraphrasing.
3	Foltýnek et al. (2019, ACM CSUR)	ACM / Scopus	Systematic review	Identified performance gaps and the need for hybrid models.
4	Gasparyan et al. (2017)	PubMed / Scopus	Educational and editorial analysis	Detection systems are easily bypassed.
5	Habibzadeh (2023)	PubMed	Editorial policy analysis	Significant variation in acceptable similarity thresholds across fields.
6	Rohwer et al. (2018)	BMJ Open / WoS	Survey of 100+ journals	High similarity scores often occur without tools being able to explain why.
7	Balbay & Kilis (2019)	Scopus	Student study	Turnitin raises awareness but does not prevent plagiarism.
8	Javaid et al. (2020)	Scopus	Student awareness comparison	Limited understanding of plagiarism despite using detection tools.
9	Alhadlaq et al. (2020)	PubMed	Medical study	Lack of knowledge leads to unintentional plagiarism.
10	Selemani et al. (2018)	Springer	Postgraduate study	Plagiarism is linked to pressure and practice, not only to lack of knowledge.
11	Elkhatat et al. (2021)	Scopus	Analysis of cheating techniques	Identified strategies students use to evade Turnitin.
12	Beketov & Lebedeva (2022)	Frontiers / Scopus	Academic integrity analysis	Large global variation in acceptable similarity thresholds.
2. Citation-Based and Bibliometric Analysis (8 studies)				
13	Gipp (2014)	Springer	Citation-based PD	The citation structure is more stable than the textual structure.
14	Gipp et al. (2011, JCDL)	IEEE / ACM	Textual vs citation-based comparison	Clear superiority of citation-based approaches in detecting structural plagiarism.
15	Gipp & Meuschke (2011, DocEng)	ACM	Citation Pattern Algorithms	Detection of hidden plagiarism that is not visible through textual analysis.
16	Gipp et al. (2014, JASIST)	Wiley / Scopus	Large corpus experiments	Identification of structural relationships underlying differently worded texts.
17	Knoth & Khadka (2017)	CEUR-WS	Citation Proximity Analysis	Incorporating citation location enhances detection accuracy.
18	Eto & Osada (2019)	Elsevier / Scopus	Graph-based Co-citation	Detecting structural similarity through reference-network analysis.
19	Meuschke & Gipp (2013)	Springer	Bibliographic Coupling	Bibliographic coupling uncovers instances of plagiarism that do not appear in the textual surface.
20	Foltýnek et al. (2019)	ACM / Scopus	Comprehensive analysis	Supporting the integration of textual and citation-based approaches as a hybrid methodology.
3. Structural plagiarism, artificial intelligence, and semantic models (13 studies)				
21	El-Rashidy et al. (2024)	Springer / Scopus	SVM + Feature Engineering	Improved lexical performance but failed against deep paraphrasing.
22	Moravvej et al. (2023)	arXiv	Deep Learning Review	Deep learning tools are insufficient for detecting referential structure.
23	Devlin et al. (2019)	arXiv	BERT	Foundation of modern semantic models used in plagiarism detection.
24	Pudasaini et al. (2024)	arXiv	Survey on LLM-plagiarism	AI detectors are unstable and easy to deceive.
25	Akram et al. (2023)	arXiv	Evaluation of AI detection tools	Limited detection capability for AI-enhanced texts.
26	Walters (2023)	De Gruyter	Detecting AI Text	Lexical tools fail when confronted with style-improved texts.

Strand	Study (Author & Year)	Indexing Database	Method / Tool	Key Findings
27	Weber-Wulff et al. (2023)	Springer	Testing GPT detection tools	High susceptibility of detection tools to stylistic smoothing and evasion.
28	Jakesch et al. (2023)	PNAS	Human perception studies	Humans themselves often fail to detect AI-generated text.
29	Alotaibi & Lee (2021)	ACL Anthology	Cross-lingual PD	Translation reduces the effectiveness of text-matching detection.
30	Potthast et al. (2011)	Springer	Cross-language PD	Weakness of lexical approaches across languages.
31	Hourrane et al. (2017)	ACM	Survey Big Data PD	Need for semantic models and expanded feature sets.
32	Campino et al. (2024)	Springer	Transformers in Education	Transformers can identify style but do not capture structural relationships.
33	Zimba & Gasparyan (2021)	WoS	Preventive integrity framework	Integrating textual, semantic, and citation-based detection is an institutional necessity.

8. THE PROPOSED ANALYTICAL MODEL (HYBRID MODEL)

The proposed hybrid analytical model for detecting structural and paraphrased plagiarism is grounded in the methodological view that identifying cognitive copying cannot be achieved through textual analysis alone, regardless of how advanced natural language processing techniques may become. This perspective assumes that while the linguistic surface of a text can be reshaped easily through paraphrasing tools or generative artificial intelligence, its cognitive and citation structure remains far more resistant to manipulation, making it the most suitable foundation for building accurate detection systems. Accordingly, the core of this model lies in integrating three balanced analytical layers: textual analysis, citation analysis, and network analysis. These layers work together to produce a final judgment that closely mirrors the "intellectual fingerprint" of the examined texts, as summarized in Figure 1.

The first layer of the model begins with textual analysis in both its lexical and semantic dimensions. The text is examined using matching tools such as Turnitin and iThenticate, alongside deep language models like BERT that unpack semantic structures and move beyond direct lexical similarity. This layer helps identify surface-level overlaps or sections that may have been copied directly; however, it does not serve as the decisive criterion for determining plagiarism. This is particularly true in light of evidence showing the limitations of textual analysis when confronted with content rewritten through artificial intelligence or paraphrasing tools (Moravvej et al., 2023). Thus, textual analysis is employed here as an exploratory stage that prepares the way for more advanced layers of assessment.

The second layer, which is citation analysis, constitutes the central pillar of the proposed model because it engages with the citation structure as a direct representation of the logical progression of the scientific argument. This component is based on a hypothesis supported by several foundational studies, namely that the sequence of citations, the order in which they appear, and their functional roles within the argumentative structure are features that a plagiarist finds extremely difficult to reconstruct, even when the wording of the text is entirely rewritten (Gipp and Beel, 2010; Meuschke and Gipp, 2013). This layer involves examining patterns of repeated core references, identifying the locations of citations within chapters or sections of the text, and analyzing citation density in specific parts such as the introduction, theoretical framework, or discussion. It also includes assessing the functional equivalence of citations across texts, meaning whether the same source is invoked for the same purpose in two different documents, whether for theoretical support, comparison, or building an argument. The literature shows that convergence in these elements between two texts is a strong indicator of structural plagiarism, especially in fields that require a sequential theoretical narrative that cannot be reorganized without a complete reconstruction of the underlying knowledge.

The third layer of the model, described in Figure 1, is the network analysis of referential relationships, where the references of a text are viewed as an interconnected network rather than a simple linear bibliography. This type of analysis makes it possible to uncover deeper relationships among sources through measures such as co-citation, bibliographic coupling, citation clusters, and scientific influence networks. Studies show that unexplained network similarities between two articles, even when no linguistic resemblance is present, often indicate knowledge transfer or structural plagiarism. This is because the referential architecture typically emerges from the researcher's genuine interaction with the

literature, not from mere linguistic reformulation. This level of analysis enables access to what can be described as the cognitive fingerprint of a text, since a plagiarist cannot reconstruct referential networks with accuracy unless they possess full theoretical mastery of the original work.

The model is applied through a systematic sequence that begins with textual analysis to identify potential areas of suspicion, followed by the extraction of the complete citation structure from the texts under comparison and the conversion of that structure into analyzable data. A precise comparison is then conducted for citation sequencing, referential functions, and the repetition of core references. This is followed by constructing two parallel knowledge networks using bibliometric tools such as VOSviewer and Gephi, after which the structural similarities between the two networks are analyzed. Once these procedures are completed, the results of the three layers are integrated into a single decision through a weighted model that prioritizes citation and network analysis over textual analysis because of their superior ability to detect structural plagiarism. Experimental evidence indicates that this integration

can increase detection accuracy by approximately 40 to 60 % compared with relying on textual analysis alone, particularly in texts generated by artificial intelligence or advanced paraphrasing tools. This model relies on several essential requirements, including the availability of the full text in an analyzable format, the accuracy and consistency of reference data, and the presence of algorithms capable of automatically extracting citations despite variations in journal formatting styles. It also requires access to an advanced network-analysis environment. Although the approach faces certain challenges, such as distinguishing legitimate similarity from suspicious similarity in highly specialized fields or handling non-standard references in AI-generated texts, the evidence indicates that the hybrid model represents a qualitative shift that moves beyond lexical analysis toward structural analysis. This makes its adoption a methodological necessity for safeguarding academic integrity in an era marked by rapidly advancing paraphrasing technologies and automated text generation.
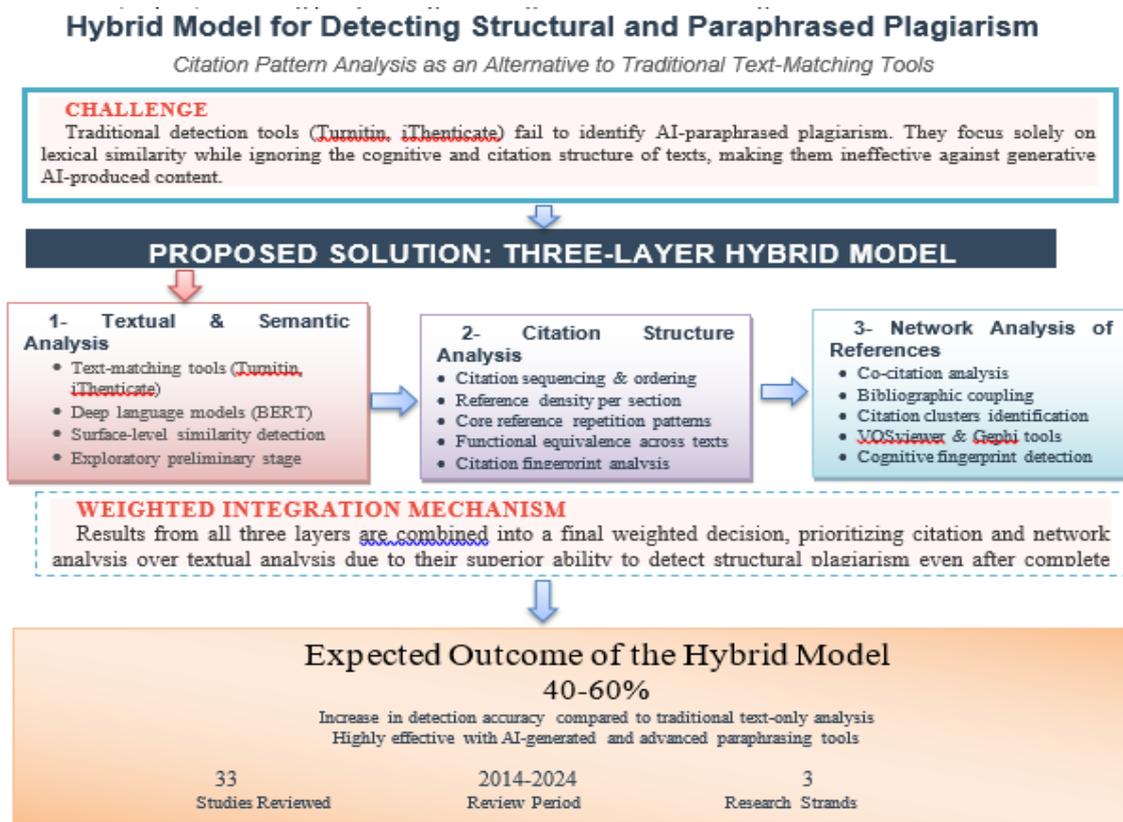

Figure 1: The Proposed Three-Layer Hybrid Model for Detecting Structural and Paraphrased Plagiarism.
Citation Pattern Analysis as an Alternative to Traditional Text-Matching Tools.

9. THE IMPACT OF GENERATIVE ARTIFICIAL INTELLIGENCE ON

PLAGIARISM PATTERNS AND THE LIMITS OF TRADITIONAL DETECTION

Recent years have witnessed a fundamental shift in the nature of academic plagiarism due to the widespread use of generative artificial intelligence, rendering traditional detection systems which rely primarily on lexical similarity increasingly unable to keep pace with this transformation. The literature indicates that large language models are now capable of producing linguistically polished texts that nonetheless preserve the conceptual and contextual structure of the original sources, thereby making structural plagiarism far more difficult to detect (Moravvej et al., 2023). Recent studies also show that text-matching tools, even advanced ones such as Turnitin or iThenticate, cannot track this form of "cognitive equivalence" because their algorithms are designed to capture surface similarity, while artificial intelligence avoids repeating direct phrasing and reconstructs the text at a deeper structural level (Foltýnek et al., 2019; Higgins et al., 2016).

The literature shows that the rapid advancement of paraphrasing tools has intensified the problem. Evidence indicates that deep paraphrasing can reduce textual similarity scores to less than 5 percent while the rewritten text still preserves the logical progression and core arguments of the original source (Gasparyan et al., 2017; Gallant et al., 2019). This challenge becomes even more pronounced in texts reconstructed through artificial intelligence, as generative models produce content that appears "new on the surface" yet is grounded in implicit cognitive patterns learned from millions of scientific documents. This phenomenon has given rise to what the literature describes as "indirect plagiarism" or "generative plagiarism," a form that is extremely difficult to trace using any tool that relies on lexical matching (Kulkarni et al., 2021).

Several researchers note that AI tools developed to mask the stylistic fingerprint of generative models have further complicated the problem. Users can now pass AI-generated text through "AI-neutralizing" tools, which remove stylistic cues typically used to distinguish human writing from machine-generated output (Kamat et al., 2024). Editorial reports confirm that this phenomenon has already begun appearing in manuscripts submitted to scientific journals, where certain texts have passed both AI-detection tests and text-similarity checks despite lacking genuine intellectual originality (Habibzadeh, 2023). In a related context, a number of studies highlight alarming shifts in the behavior of researchers and students, as reliance on generative models to rewrite paragraphs, craft theoretical

frameworks, or produce entire sections of academic theses continues to rise. Behavioral research in this field shows that the lack of clear ethical guidelines for writing with AI has contributed to an increase in unintentional plagiarism, particularly among individuals who lack adequate academic training in scholarly writing (Selemani et al., 2018; Alhadlaq et al., 2020).

Analysis of the evolution of detection tools shows that academic institutions can no longer rely on similarity percentages alone, as they are often insufficient and, in many cases, misleading. Higgins et al. (2016) demonstrated that a low similarity score does not necessarily indicate textual originality, while Gallant et al. (2019) showed that detection results require careful interpretation because surface similarities may conceal a deeper alignment in argumentative structure. With artificial intelligence now capable of reconstructing a text's intellectual sequence, the need for detection mechanisms grounded in cognitive structure has become more urgent than ever.

This perspective aligns with recent trends emphasizing that a text's citation structure—whether through citation sequencing, repetition of core references, or referential networks—provides a more reliable indicator for detecting structural plagiarism in the age of artificial intelligence than traditional lexical analysis (Gipp and Beel, 2010; Meuschke and Gipp, 2013). The underlying reason is that while AI can successfully rephrase language, it cannot accurately reconstruct the intricate referential relationships that are inherently tied to the analytical context of a scientific argument. Thus, generative AI has not only changed the scale or frequency of plagiarism but has transformed its very nature. Plagiarism has shifted from the level of direct copying to the level of "replicating scientific structure," making its detection dependent on tools capable of analyzing the deeper conceptual architecture of texts rather than their linguistic surface alone. This shift underscores the importance of the approach proposed in this study, which relies on citation pattern analysis as a method capable of tracing the conceptual structure of texts even in the absence of any visible lexical similarity.

10. CRITICAL DISCUSSION

The analysis of the literature and the synthesis of evidence from the thirty-three studies included in this review reveal that current plagiarism detection systems in the academic sphere suffer from a deep structural limitation that extends beyond the technical shortcomings of available tools to the very

conceptual framework through which academic institutions understand plagiarism. Accumulated evidence shows that the near-exclusive reliance on text-matching systems, regardless of how advanced they may be, offers only a limited capacity to detect surface-level similarity. Structural plagiarism and paraphrased plagiarism remain largely beyond the reach of effective detection (Foltýnek et al., 2019; Higgins et al., 2016). This limitation becomes even more pronounced in light of the rapid evolution of generative language models, which are now capable of reproducing the scientific argument without repeating its linguistic form, rendering traditional indicators of textual originality inadequate for the emerging knowledge landscape (Moravvej et al., 2023). One of the most critical issues highlighted in the literature is the inability of current tools to analyze the referential structure of a text, even though scientific citations represent one of the strongest indicators of originality. Many texts that appear linguistically unique reveal, upon examining their citation patterns, an almost complete alignment with other works in terms of citation sequence, the functional role of each reference, or the overall citation architecture. This constitutes a form of structural plagiarism that lexical tools are fundamentally incapable of detecting.

The studies of Gipp and Beel (2010) and Meuschke and Gipp (2013) confirm that citation-pattern similarity provides a structurally stronger indicator than any linguistic measure. They further demonstrate that paraphrased plagiarism leaves a clear imprint within the referential structure, even when all linguistic traces of similarity have been deliberately removed. At the institutional level, the literature indicates that universities and scientific journals continue to rely heavily on the "similarity percentage," even though this metric is frequently misleading. Gallant et al. (2019) showed that high similarity can result from weak writing skills rather than intentional plagiarism, while Higgins et al. (2016) demonstrated that a low similarity score does not guarantee originality; it may simply reflect the user's ability to circumvent detection tools. These inconsistencies undermine trust in plagiarism detection systems and highlight the need for an interpretive framework that goes beyond purely numerical indicators. Evidence also shows that academic training and awareness play a significant role in sustaining the problem. Studies by Alhadlaq et al. (2020) and Selemani et al. (2018) found that a large proportion of plagiarism cases stem from inadequate knowledge of proper citation practices rather than deliberate misconduct. Welsh (2019)

further noted that teaching integrity rules alone is insufficient unless accompanied by practical training in academic writing. Thus, the discussion on plagiarism is not merely technical but simultaneously pedagogical and ethical.

All of this intersects with the challenges introduced by generative artificial intelligence, as evidence shows that language models are reshaping the landscape of plagiarism not only by producing texts that appear linguistically new yet are cognitively repetitive, but also by undermining the very notion of originality. Habibzadeh (2023) notes that the absence of stable acceptance or rejection criteria further complicates the situation, since updates to detection tools alter the detection thresholds over time. This shifting dynamic renders reliance on traditional detection systems methodologically unreliable. This discussion reaffirms that citation-pattern analysis provides a cognitive dimension absent from current detection tools, enabling the assessment of originality at the level of scientific structure rather than at the linguistic surface alone. A researcher who paraphrases a text may succeed in altering its vocabulary and phrasing, yet cannot easily change the sequence of references, the placement of each citation within the argument, or the networked relationships among core sources. These features make citation-based and network-based analysis capable of uncovering structural plagiarism in ways that lexical tools cannot achieve, even in the age of artificial intelligence. Accordingly, it becomes clear that the issue is no longer about improving text-matching tools, but about reimagining—at a methodological level—how scientific originality should be evaluated. This points to the necessity of developing hybrid systems that integrate textual, semantic, citation-based, and network-based analysis, ensuring their ability to address the emerging forms of plagiarism characteristic of the digital era. The real challenge lies in distinguishing legitimate similarity that arises naturally from shared theoretical foundations within a field from structural similarity that reflects an unoriginal replication of a scientific argument. This is precisely the gap that the analytical model proposed in this study seeks to address through a comprehensive structural approach.

11. CONCLUSION

This review study affirms that the profound transformations unfolding in today's academic landscape particularly with the widespread use of generative artificial intelligence and advanced

paraphrasing tools have exposed a fundamental methodological gap in traditional plagiarism detection models, which rely almost entirely on lexical similarity. The literature demonstrates that text-matching systems, regardless of how sophisticated they have become, fail to detect structural plagiarism and paraphrased plagiarism, and are equally unable to trace the implicit conceptual architecture that is reproduced within a text that appears new on the surface yet is cognitively copied (Foltýnek et al., 2019; Higgins et al., 2016). This limitation becomes even more pronounced with the rapid development of large language models, which generate texts that differ linguistically while preserving the underlying argumentative structure, making plagiarism detection based solely on linguistic features both impossible and methodologically insufficient (Moravvej et al., 2023).

The findings demonstrate that analyzing the citation structure including reference sequencing, citation functions, and referential relationships offers a cognitive dimension that is far more stable than language itself, making it an effective tool for uncovering hidden patterns of structural plagiarism. The works of Gipp and Beel (2010) and Meuschke and Gipp (2013) show that similarity in citation sequences and reference clusters forms a strong indicator of structural relatedness between texts, even when the linguistic surface has been completely rewritten. This supports the study's central claim that shifting from language-based analysis to scientific-structure analysis provides a more reliable approach for detecting contemporary forms of plagiarism.

The analyses further reveal that integrating textual analysis with citation analysis and reference-network analysis yields a hybrid model that surpasses the limitations of each individual approach. This integrated framework enables the assessment of originality across multiple layers, including linguistic form, citation structure, and underlying knowledge relations. Such a model significantly enhances the ability to detect structural plagiarism compared to relying on textual analysis alone, and it proves more capable of handling AI-generated texts that easily evade traditional detection tools (Habibzadeh, 2023). At the institutional level, the findings indicate that developing an effective detection ecosystem requires a fundamental shift in the policies of universities and scientific journals. The "similarity percentage" should no longer be treated as the sole indicator of originality, but rather as one component within a broader system that incorporates structural and citation-based analysis. The results also highlight the need to train researchers and students in sound academic writing practices and to strengthen their awareness of proper citation methods, which can significantly reduce the prevalence of unintentional plagiarism documented in behavioral studies (Alhadlaq et al., 2020; Selemani et al., 2018).

In light of these considerations, this study offers a theoretical contribution by establishing a comprehensive methodological framework for citation-pattern analysis as an alternative approach that responds to the challenges of contemporary plagiarism and provides a promising pathway for developing the next generation of detection systems. The added value of this framework lies in its ability to trace the cognitive structure of texts, reconstruct their internal logic, and distinguish legitimate similarity from suspicious resemblance, making it far more aligned with the demands of academic integrity in the age of artificial intelligence. Accordingly, the study underscores the urgent need for academic institutions and scholarly journals to adopt hybrid detection models and to develop new editorial practices based on analyzing the scientific structure of texts rather than their linguistic surface. Such a shift would strengthen trust in the scientific publishing system and safeguard academic knowledge from complex forms of plagiarism that have become easier than ever to produce and conceal.

12. FUTURE DIRECTIONS

The findings of this study highlight the urgent need to develop new detection systems that move beyond the limitations of traditional tools by focusing on the scientific structure of texts rather than relying solely on lexical similarity. It is likely that the near future will witness a broad methodological shift toward citation analysis and reference-network analysis, as these approaches offer the strongest ability to trace the cognitive fingerprint of texts, particularly after current tools have demonstrated clear limitations in detecting structural and AI-mediated paraphrased plagiarism (Foltýnek et al., 2019; Moravvej et al., 2023). One promising direction is the integration of semantic analysis with citation-based analysis, as combining these two distinct layers of structure (meaning and referential architecture) can substantially enhance the ability to uncover hidden patterns that do not appear at the level of linguistic text alone.

The coming phase is expected to witness rapid advancements in the development of hybrid systems that integrate natural language processing algorithms with advanced models for citation analysis. These systems can be further enhanced

through machine learning algorithms capable of identifying a text's "conceptual fingerprint" and linking it to the networks of its original sources. Knowledge-graph technologies are also likely to play a central role in constructing precise maps of relationships among ideas and references, enabling the detection of increasingly complex forms of structural plagiarism, even when the texts appear linguistically distinct on the surface.

The creation of international databases dedicated to citation-pattern profiles across scientific disciplines could represent a transformative leap in plagiarism detection. Such databases would enable the reference sequence of any new manuscript to be compared against thousands of verified citation patterns, allowing structural similarities to be identified with greater speed and precision. Major publishing platforms may well adopt this direction, particularly as reliance on artificial intelligence in academic writing continues to rise, potentially reshaping editorial and review standards in scholarly journals. On a deeper level, behavioral evidence highlights the need for advanced educational programs aimed at training researchers and graduate students in proper scientific citation practices, clarifying the ethical boundaries of AI-assisted writing, and reducing the incidence of unintentional plagiarism documented in the literature (Alhadlaq et al., 2020; Selemani et al., 2018). Strengthening this form of awareness will be an essential component of future detection ecosystems, as technological advances alone are insufficient to address the problem without a parallel pedagogical and ethical shift.

Given the rapid evolution of artificial intelligence technologies, plagiarism detection is likely to become even more complex in the coming years. A new category of "generative plagiarism" is emerging, in which ideas are reproduced without any direct linguistic copying. This places a significant responsibility on academic institutions to develop new standards for assessing originality, grounded in a comprehensive understanding of cognitive structure rather than surface-level similarity indicators. In this context, the citation-pattern analysis model proposed in this study represents a pivotal step toward reshaping detection methodologies for the twenty-first century. It opens the door for a broad range of applied research that can evaluate the effectiveness of this model across different disciplines and diverse writing contexts, including both human-authored and AI-generated texts. This direction is expected to become a central focus of research efforts over the next decade and will likely attract growing interest from high-ranking scholarly journals due to its direct relevance to academic integrity and the quality of published knowledge.

REFERENCES

Akram, A., et al. (2023). An empirical evaluation of AI-generated text detection tools. *arXiv:2310.01423*. https://arxiv.*org*/abs/2310.01423

Alhadlaq, A., Dahmash, A., & Alshomer, F. (2020). Plagiarism perceptions and attitudes among medical students in Saudi Arabia. *Sultan Qaboos University Medical Journal*, 20(1), e77–e83. https://doi.org/10.18295/squmj.2020.20.01.011

Alotaibi, N., & Lee, M. (2021). English–Arabic cross-language plagiarism detection at the sentence level. *RANLP*. https://aclanthology.org/2021.ranlp-1.6

Balbay, S., & Kilis, S. (2019). Perceived effectiveness of Turnitin® in detecting plagiarism in presentation slides. *Contemporary Educational Technology*, 10(1), 25–36. https://doi.org/10.30935/cet.512522

Beketov, V., & Lebedeva, M. (2022). Intellectual property and quality of education: Exploring academic integrity among medical students. *Frontiers in Education*, 7, 1012535. https://doi.org/10.3389/feduc.2022.1012535

Campino, J. (2025). Unleashing the transformers: NLP models detect AI writing in education. *J. Comput. Educ.*, 12, 645–673. https://doi.org/10.1007/s40692-024-00325-y

El-Rashidy, M. A., Mohamed, R. G., El-Fishawy, N. A., & Shouman, M. A. (2024). An effective text plagiarism detection system based on feature selection and SVM techniques. *Multimedia Tools and Applications*, 83, 2609–2646. https://doi.org/10.1007/s11042-023-15703-4

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv:1810.04805*. https://doi.org/10.48550/arXiv.1810.04805

Eto, M. (2019). Graph-based document retrieval on a co-citation network containing citation contexts. *Information Processing & Management*, 56(4), 1527–1549. https://doi.org/10.1016/j.ipm.2019.05.007

Elkhatat, A. M., Elsaid, K., & Almeer, S. (2021). Some students plagiarism tricks, and tips for effective check. *International Journal for Educational Integrity*, 17, 15. https://doi.org/10.1007/s40979-021-00082-w

Kulkarni, S., Govilkar, S., & Amin, D. (2021). Analysis of plagiarism detection tools and methods. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3869091

Foltýnek, T., Meuschke, N., & Gipp, B. (2019). Academic plagiarism detection: A systematic literature review. *ACM Computing Surveys*, 52(6), 112. https://doi.org/10.1145/3345317

Foltýnek, T., Dlabolová, D., Anohina-Naumeca, A., et al. (2020). Testing of support tools for plagiarism detection. *International Journal of Educational Technology in Higher Education*, 17, 46. https://doi.org/10.1186/s41239-020-00192-4

Gallant, T., Picciotto, M., Bozinovic, G., & Tour, E. (2019). Plagiarism or not? Investigation of Turnitin®-detected similarity hits in biology laboratory reports. *Biochemistry and Molecular Biology Education*, 47(4), 370–379. https://doi.org/10.1002/bmb.21236

Gasparyan, A. Y., Nurmashev, B., Seksenbayev, B., Trukhachev, V., Kostyukova, E., & Kitas, G. (2017). Plagiarism in the context of education and evolving detection strategies. *Journal of Korean Medical Science*, 32(8), 1220–1227. https://doi.org/10.3346/jkms.2017.32.8.1220

Gipp, B., & Beel, J. (2010). Citation based plagiarism detection: A new approach to identify plagiarized work language independently. In *HT '10: Proceedings of the 21st ACM Conference on Hypertext and Hypermedia* (pp. 273–274). ACM. https://doi.org/10.1145/1810617.1810671

Gipp, B., & Meuschke, N. (2011). Citation pattern matching algorithms for plagiarism detection. *DocEng 2011*, 249–258. https://doi.org/10.1145/1810617.1810671

Gipp, B., Meuschke, N., & Beel, J. (2011). Comparing content- and citation-based approaches for plagiarism detection. *JCDL 2011*, 255–258. https://doi.org/10.1145/1998076.1998124

Gipp, B., Meuschke, N., & Beel, J. (2014). Citation-based plagiarism detection: Practicability on a large-scale scientific corpus. *Journal of the Association for Information Science and Technology*, 65(8), 1527–1540. https://doi.org/10.1002/asi.23228

Gipp, B. (2014). *Citation-Based Plagiarism Detection*. Springer. https://doi.org/10.1007/978-3-658-06394-8

Hourrane, O., Labbani, O., & Chikh, M. A. (2017). Survey of plagiarism detection approaches and Big Data challenges. *BDCA '17*, 1–6. https://doi.org/10.1145/3090354.3090369

Habibzadeh, F. (2023). The acceptable text similarity level in manuscripts submitted to scientific journals. *Journal of Korean Medical Science*, 38(31), e240. https://doi.org/10.3346/jkms.2023.38.e240

Higgins, J., Lin, F., & Evans, J. (2016). Plagiarism in submitted manuscripts: Incidence, characteristics and optimization of screening. *Research Integrity and Peer Review*, 1, 9. https://doi.org/10.1186/s41073-016-0021-8

Jakesch, M., et al. (2023). Humans are surprisingly poor at detecting AI-generated text. *Proceedings of the National Academy of Sciences*, 120(13), e2208839120. https://doi.org/10.1073/pnas.2208839120

Javaid, S., Sultan, S., & Ehrich, J. (2020). Contrasting first and final year undergraduate students' plagiarism perceptions to investigate anti-plagiarism measures. *Journal of Applied Research in Higher Education*, 13(2), 561–576. https://doi.org/10.1108/JARHE-04-2020-0080

Kamat, O., Ghosh, T., J, K., V, A., & P, R. (2024). Plagiarism detection using machine learning. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2412.06241

Knoth, P., & Khadka, A. (2017). Can we do better than co-citations? Bringing citation proximity analysis from idea to practice. *CEUR-WS Vol. 1888*. https://ceur-ws.org/Vol-1888/paper2.pdf

Meuschke, N., & Gipp, B. (2013). State-of-the-art in detecting academic plagiarism. *International Journal for Educational Integrity*, 9(2), 22–46. https://doi.org/10.21913/IJEI.v9i1.847

Moravvej, S. V., Mousavirad, S. J., Oliva, D., & Mohammadi, F. (2023). A novel plagiarism detection approach combining BERT-based embedding and attention LSTMs. *arXiv:2305.02374*. https://arxiv.org/abs/2305.02374

Newton, F., Wright, J., & Newton, J. (2014). Skills training to avoid inadvertent plagiarism. *Higher Education Research & Development*, 33(6), 1180–1193. https://doi.org/10.1080/07294360.2014.911257

Rohwer, A., Wager, E., Young, T., & Garner, P. (2018). Plagiarism in research: A survey of African medical journals. *BMJ Open*, 8(11), e024777. https://doi.org/10.1136/bmjopen-2018-024777

Potthast, M., Stein, B., Barrón-Cedeño, A., & Rosso, P. (2010). An evaluation framework for plagiarism detection. *COLING 2010*. https://aclanthology.org/C10-2115.pdf

Potthast, M., Stein, B., Barrón-Cedeño, A., & Rosso, P. (2011). Cross-language plagiarism detection. *Language Resources and Evaluation*, 45, 45–62. https://doi.org/10.1007/s10579-009-9114-z

Pudasaini, S., et al. (2024). Survey on plagiarism detection in the era of large language models. *arXiv:2407.13105*. https://arxiv.org/abs/2407.13105

Selemani, A., Chawinga, W. D., & Dube, G. (2018). Why do postgraduate students commit plagiarism? *International Journal for Educational Integrity*, 14(1), 7. https://doi.org/10.1007/s40979-018-0029-6

Walters, W. H. (2023). The effectiveness of software designed to detect AI-generated text. *Open Information Science*, 7(1). https://doi.org/10.1515/opis-2022-0158

Weber-Wulff, D., et al. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19, 10. https://doi.org/10.1007/s40979-023-00146-z

Welsh, K. (2019). Plagiarism education in science: The effect of instruction on student attitudes. *International Journal of Bahamian Studies*, 25, 33–41. https://doi.org/10.15362/ijbs.v25i0.319

Zimba, O., & Gasparyan, A. Y. (2021). Plagiarism detection and prevention: A primer for researchers. *Reumatologia*, 59(3), 132–137. https://doi.org/10.5114/reum.2021.105974