

Precision of Academic Plagiarism Detection: A Descriptive Analysis of Artificial Intelligence Verifiers

Precisión de la detección de plagio académico: Un análisis descriptivo en verificadores de Inteligencia Artificial

Luis Ebano Amor Oliva

Universidad Juárez Autónoma de Tabasco

<https://orcid.org/0000-0002-0193-7415>

siulromaavilo@outlook.com

México

Erika Guadalupe May Guillermo

Tecnológico Nacional De México/ ITS de la Región Sierra

<https://orcid.org/0000-0001-5403-9849>

erika.mg@regionsierra.tecnm.mx

México

Abstract:

Some authors argue that the use of Artificial Intelligence (AI) should be considered plagiarism and have proposed the use of verifier tools to combat this type of plagiarism. Among these Artificial Intelligences is ChatGPT, which has revolutionized the world with its ability to produce human-like text. The purpose of this work is to identify the level of accuracy in detecting plagiarism using Artificial Intelligence detection tools in literature writing. Fifty samples of books written and published before November 2022 were collected, and ChatGPT was asked to generate another 50 literature samples from different genres. The original human and AI-generated content were analyzed using four plagiarism detection tools, which were moderately successful in identifying human content but had varying degrees of effectiveness in detecting AI-generated content. Copy Leaks scored 99% on the F-score, Content at Scale 79%, Scribber 25%, and ZeroGPT 69%. This paper has an explanatory approach with a cross-sectional design, with quantitative analysis of the data collected. ChatGPT as the potential to displace human writers and the use of these AI verifiers can aid schools and editorial houses to distinguish original human content and that generated by IA. We exhort AI verifiers to improve their algorithms used to identify plagiarism, and for schools to incorporate these types of tools in the design of strategic pedagogies for future papers.

Keywords: Artificial intelligence, ChatGPT, Large language models, Plagiarism.

Resumen:

Algunos autores afirman que el uso de Inteligencia Artificial (IA) debería considerarse como plagio y han propuesto el uso de herramientas verificadoras para combatir este tipo de plagio. Entre estas Inteligencias Artificiales se encuentra ChatGPT, que ha revolucionado el mundo con su capacidad de producir texto similar al humano. La finalidad de este trabajo es identificar el nivel de precisión al detectar el plagio mediante verificadores de IA en la escritura literaria. Se recopilaron 50 muestras de libros escritos y publicados antes de noviembre de 2022, y se solicitó a ChatGPT que generara otras 50 muestras de literatura de diferentes géneros. El contenido original humano y el generado por la IA fueron analizados empleando cuatro verificadores detectores de plagio, los cuales tuvieron éxito moderado al identificar contenido humano, pero tuvieron grados variables de efectividad en la detección de contenido generado por IA. Copy Leaks obtuvo un puntaje del 99% en el F-score, Content at Scale un 79%, Scribber un 25% y ZeroGPT un 69%. La investigación es de enfoque explicativo de diseño transversal, con análisis cuantitativo de la información recopilada. ChatGPT tiene el potencial de desplazar a los escritores humanos y el uso de verificadores de IA puede ayudar a las escuelas y editoriales a distinguir entre contenido original humano y el generado por artificialmente. Se exhorta a los



verificadores de plagio a mejorar sus algoritmos de identificación de plagio literario, así como la incorporación de estas herramientas en diseños de estrategias pedagógicas de futuras investigaciones.

Palabras clave: Inteligencia artificial, ChatGPT, Modelos grandes de lenguaje, Plagio.

Received: 04/02/2025 | **Accepted:** 12/09/2025 | **Published:** 01/06/2026 | pp. 1 - 18

DOI: 10.19136/etie.v8n16.6338

Precision of Academic Plagiarism Detection: A Descriptive Analysis of Artificial Intelligence Verifiers

Introduction

ChatGPT, as an Artificial Intelligence (AI) tool, has revolutionized the world and has been used by millions of people across multiple industries and economic sectors (Gao et al., 2022; Rathore, 2023). This tool was created by OpenAI, a non-profit organization, and was released to the public in late 2022 (Wu et al., 2023).

The nature of this AI is that of LLMs (Large Language Models) (Rathore, 2023; Thirunavukarasu et al., 2023). Generally speaking, LLMs are advanced language models based on transformer architecture, trained on large text datasets, capable of learning language patterns and context, with attention mechanisms, capturing dependencies, and capable of performing a variety of multitasking tasks, adapting to specific use cases as required (Chang et al., 2023).

With its growing popularity, teachers at schools and universities have reported that some of their students use ChatGPT to write their assignments (Adeshola & Adepoju, 2023; Khalil & Er, 2023). Authors such as (Grassini, 2023; Jarrah et al., 2023) have pointed out the problem that the use of this AI can lead to academic plagiarism.

Among research literature, Latin America has seen the retraction of a large percentage of papers due to plagiarism allegations (Almeida, et al., 2016) and with the rise of Artificial Intelligence, and the risk of copy-paste plagiarism, this issue is more important than ever, particularly in México, where plagiarism is not a phenomena widely studied, but has a history of academic papers being retracted after publishing due to plagiarism allegations (Rivera, 2023). Some authors have stated the cons and benefits of the use of ChatGPT as an academic tool in México (Landero et al., 2024), and among the cons is the ease of falling into plagiarism by students.

To combat this problem, AI-powered plagiarism detection tools have been developed (Grassini, 2023; Jarrah et al., 2023), which have been used by some teachers to detect whether student work has been written using ChatGPT (Gao et al., 2022; Cotton et al., 2023). Some authors have proposed the use of AI detection software to detect plagiarism in various fields, such as research, education, and journalism, among others (Biörck & Eriksson, 2023; Gao et al., 2022).

However, most articles have focused on plagiarism in scientific research (Currie, 2023; Hua et al., 2023; Kacena et al., 2024), or focus on the pros and cons of the use of these tools in education by students or teachers, but very few research has been conducted in evaluating the *precision* of these AI verifiers, and if we are to incorporate them at all among the educator's tools, we need to ask: How good do these verifiers even work at detecting plagiarism by AI?

And since literature is an important resource for education and culture, the objective of this article is to analyze the level of accuracy in detecting plagiarism using Artificial Intelligence checkers in academic writing, such as, Copyleaks (2024), Content at Scale (2024), Scribber (2024), and zerogpt (2024). Using

excerpts from books published before ChatGPT was created, as well as paragraphs created entirely by AI, to compare the accuracy level of detection of original and artificially created texts.

So, our research question is: Is there a statistically significant difference between the precision of the different AI verifiers analyzed in this study?

The null hypothesis (H_0) states that there is no statistically significant difference in the level of precision among the verifiers CopyLeaks, Content at Scale, Scribber and ZeroGPT.

Conversely, the alternative hypothesis (H_1) posits that at least one of the verifiers differs significantly in its level of precision.

One of the limitations of this paper is the scope of the literature analyzed, as the data we used was extracted from literature books. The reason of this is that we have the intention of starting the dialogue of how we can measure the effectiveness of the educator's tools -such as AI verifiers- and we encourage future papers analyzing the precision of AI verifiers for other types of literature, such as academic or scientific writing.

We hope the results of this study shed light on an important factor of the use of plagiarism verifiers - that is the validity of their results.

Artificial intelligence

Artificial Intelligence (AI) is a field within computer science dedicated to developing machines with human-like thinking and behavior. These systems learn from their surroundings, utilizing data to make informed decisions. AI applications extend to solving intricate challenges like medical diagnoses, autonomous vehicles, and natural language processing (Deng & Lin, 2022).

Machine Learning

Machine learning is a branch of artificial intelligence where computer systems autonomously learn and improve from experience without explicit programming. It involves algorithms that enable computers to recognize patterns, make predictions, and improve performance over time as they process and analyze data (Zhou, 2021).

Natural Language Processing

Neural networks belong to a category of machine learning systems designed to emulate the human brain's structure. They consist of interconnected units known as nodes, organized into layers. The input layer receives data, which undergoes processing in the hidden layers before being output from the output layer. Each connection between nodes is assigned a weight value, determining the connection's strength. The inputs are multiplied by these weights, summed at each node, and the resultant value undergoes transformation through an activation function (Deng & Lin, 2022).

Large Language Models

Language Models are Artificial Intelligence systems capable of understanding and generating human-like language. These systems have the ability of predicting the likelihood of word sequences and to generate their own text based on their input; however, they may exhibit biases and limitations in nuanced comprehension (Chang, et al., 2023).

There are different sizes of language models, and the term Large Language Models refers to models that have a massive number of parameters. Parameters are the internal variables that the model learns during training, and more parameters generally allow the model to capture more complex patterns in the data. This type of model is better suited than smaller Language Models for human to machine interactions, due to their more sophisticated training (Chang, et al., 2023).

ChatGPT

ChatGPT is a Natural Language Processing system created by OpenAI with the aim of producing realistic conversations. It accomplishes this by comprehending the context of a conversation and crafting suitable responses (Deng & Lin, 2022).

This LLM has demonstrated the ability to comprehend conversation context and produce fitting responses. Furthermore, it can generate responses in various languages, encompassing English, Spanish, French, and German. Additionally, ChatGPT is proficient in generating responses in diverse styles.

Academic Plagiarism

Plagiarism is defined as using someone else's ideas without their proper acknowledgment. Which can range from simple dishonesty to serious theft of complete manuscripts (Masic, 2014; Kumar, et al., 2014; Naik, et al., 2015; Sharma & Verma, 2020). Given the current widespread digitalization, copy-paste plagiarism has become widespread in academic writing (Masic, 2014; Khaled & Al-Tamimi, 2021). Some authors mention the rise of digitalization as one of the main causes of plagiarism (Dahl, 2007; Naik, et al., 2015).

Types of Academic Plagiarism

Weber-Wulff (2014) in her book "A perspective on Academic Plagiarism" classified plagiarism as copy and paste, translations, disguised plagiarism, shake and paste collections, Clause Quilts, Pawn Sacrifice, Cut and slide and self-plagiarism. Naik, et al. (2015) has a similar classification published in an article reviewing plagiarism detection tools, where they mentioned copy and paste, disguised plagiarism, by translation, shake and paste, structural plagiarism, mosaic plagiarism, metaphor plagiarism and idea plagiarism.

Plagiarism Detection Software

Lancaster (2003) proposed a methodology for identifying academic plagiarism, which is based on the effectiveness of identifying the similarities between documents, the efficiency in the use of computational resources, and the workload of academic tutors.

Other authors mention two main methods of plagiarism detection, namely external plagiarism detection, techniques which includes grammar, semantic, cluster, cross lingual, citation and character-based detection; and intrinsic plagiarism detection techniques, which incorporates grammar-semantics hybrid, structure based and syntax methods (Naik, 2015).

To combat academic plagiarism, some companies have developed Plagiarism detection software, such as Ferret, Plagiarism Catcher and Turnitin.com. This type of software compares the input text against large datasets of academic sources, web pages or books (Youmans, 2011).

Plagiarism by Artificial Intelligence Detector Software

Due to the widespread access of ChatGPT and other LLMs in academic writing, some authors have suggested the incorporation of Artificial Intelligence detector software in the editorial process of scientific papers by evaluation comities (Gao, et al., 2022). Some of these tools can be easily access by students, teachers, researchers and evaluators for free, such as: Copyleaks (2024), Content at Scale (2024), Scribber (2024), zeroGPT (2024).

| *Methodology*

The methodology of this study was based on the model proposed by Gao (2022). This study is quantitative and has an explanatory, cross-sectional design. Initially, ChatGPT 3.5 was asked to generate 50 samples of paragraphs from fiction books using the prompt: *“Write a fragment from a book, on a topic of your choice. Write the fragment as if you had chosen it from a random page in the book”*. Subsequently, a sample of 50 literary books was collected, all of them written and published before November 2022, which marks the period in which ChatGPT was introduced to the public. From this collection of books, a paragraph was randomly selected. The pages of each book were chosen by a random number generator. The sample size of 50 fragments of real books and another 50 fragments of AI generated. Sample size was determined by time limitations and the purpose of testing the methodological approach developed by the aforementioned author.

To evaluate the performance of the plagiarism by AI detectors, we employed two widely used composite measures derived from the confusion matrix: The F-score and the Area Under the Receiver Operating Characteristic Curve (AUROC). The F-score provides a balanced assessment of a detector's precision and sensitivity, which is particularly relevant when both true negatives and true positives are already known (as we already know which ones are the true positives: paragraphs made by AI; and true negatives: original work) and when false positives and false negatives affect the overall reliability of detection. The sensitivity (also known as true positive rate) and the specificity (true negative rate) are corollary measurements of the confusion matrix, which are used for the AUROC, measure that summarizes the detector's ability to discriminate between original and plagiarized text (Goutte & Gaussier, 2005; Sokolova et al., 2006; Gao, 2022). These metrics were chosen because they enable a comparative analysis of detection, precision and discriminatory power when true and false classifications are already known.

Based on these results, four AI plagiarism detection checkers were employed to evaluate their accuracy in detecting AI-generated text: (1) CopyLeaks, (2) Content at Scale, (3) Scribber's, and (4) ZeroGPT. A comparative analysis was conducted to differentiate between the results obtained with plagiarism detection tools when applied to books written by humans and those generated by ChatGPT. Statistics and graphs were generated using STATA version 14. Group statistics were reported using means and the F score. For the purpose of this study, the dependent variables are the plagiarism by AI verifiers, and the depended variables are the performance metrics F-score and AUROC. These results are interpreted as follows: the closer to 1 the score is, the more accurate the plagiarism checker is (Goutte & Gaussier, 2005). As for the AUROC, the closer the value is to 1 the more is considered to have good predictive capacity (Fan et al., 2006).

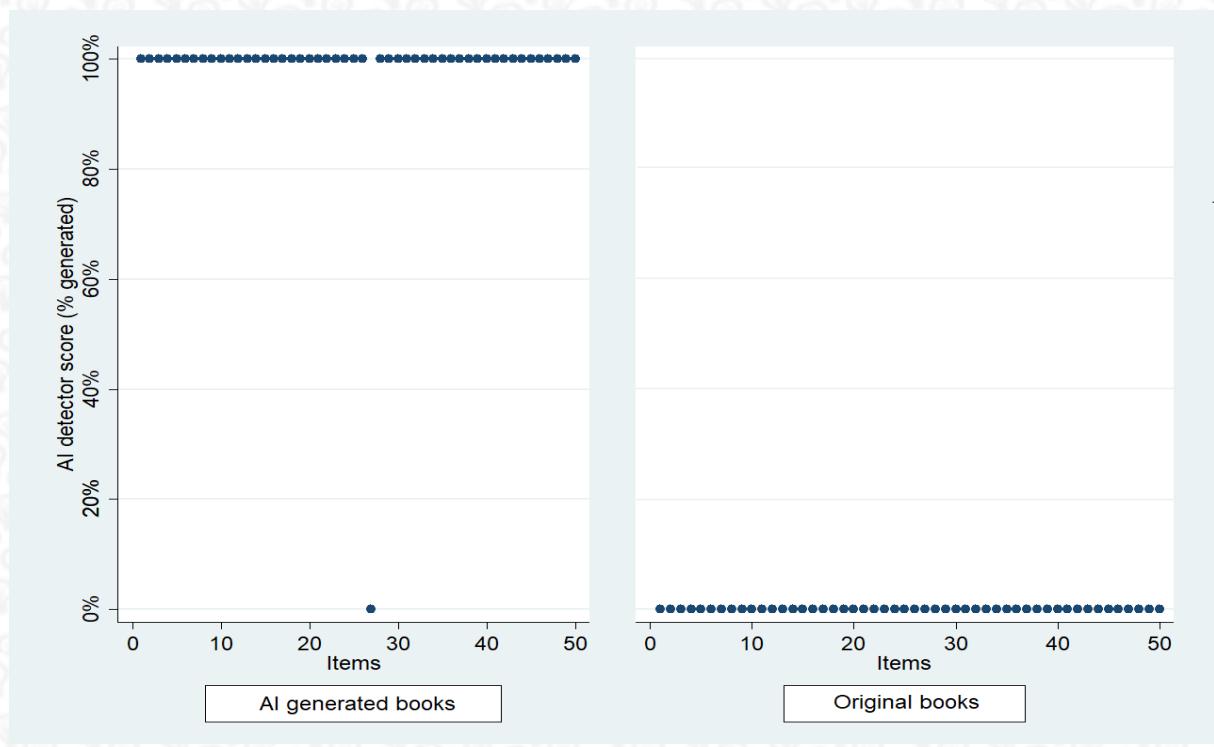
| **Results**

The following section presents the research findings, sorted by the degree of accuracy reported for each fact-checker. It begins with the analysis of CopyLeaks, followed by Content at Scale, then ZeroGPT, and finally Scribber.

The first fact-checker analyzed is CopyLeaks, which exhibited the highest accuracy in identifying original and AI-generated content, showing less data dispersion. This data is shown in Figure 1. Figure 2 illustrates an AUROC of 0.990, while the F-score obtained a percentage of 99% and a mean probability of being AI-generated for original content of 0% (0.00) and 98% (0.98) for AI-generated content.

Figure 1

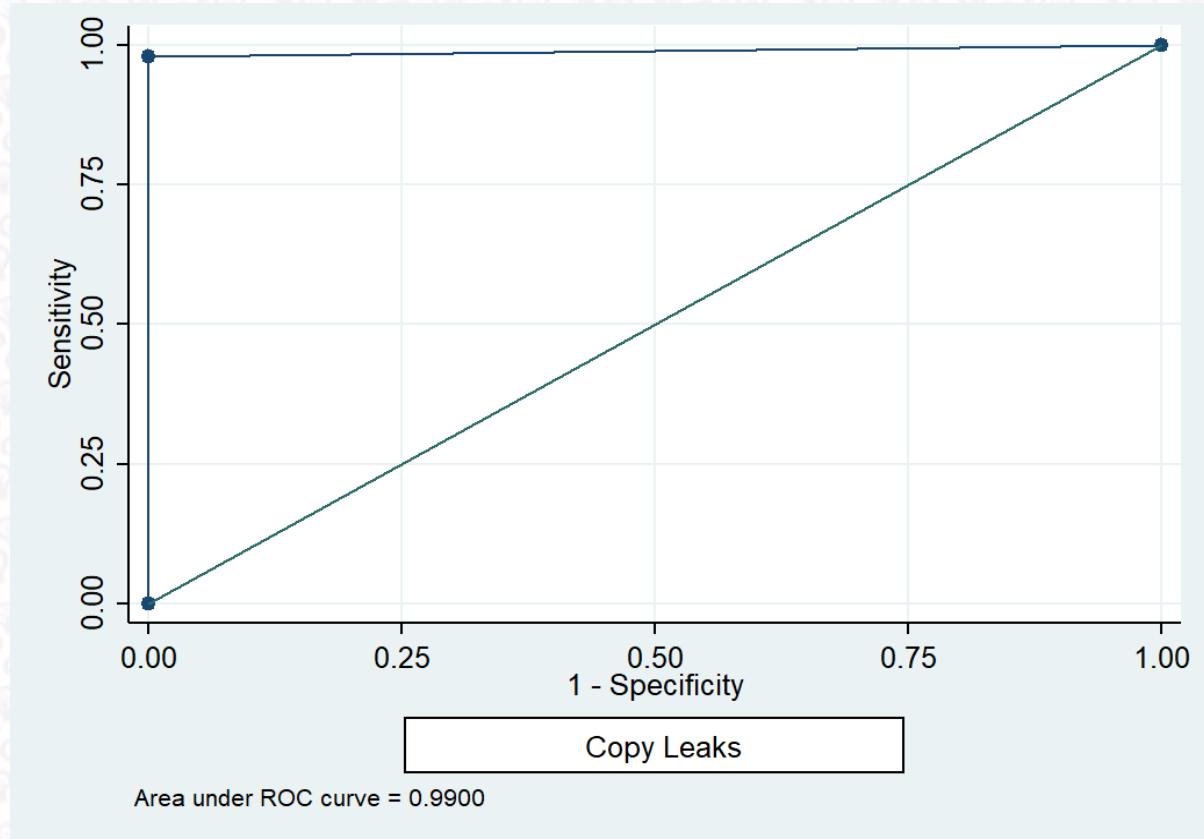
Scatterplot of CopyLeaks' output



Note. Data obtained from STATA version 14. Of the 50 items analyzed, only one was erroneously assigned as original, despite having been generated by ChatGPT. The CopyLeaks result was binary, either “This is a human text” or “AI content detected.”.

Figure 2

ROC curve of CopyLeaks' output

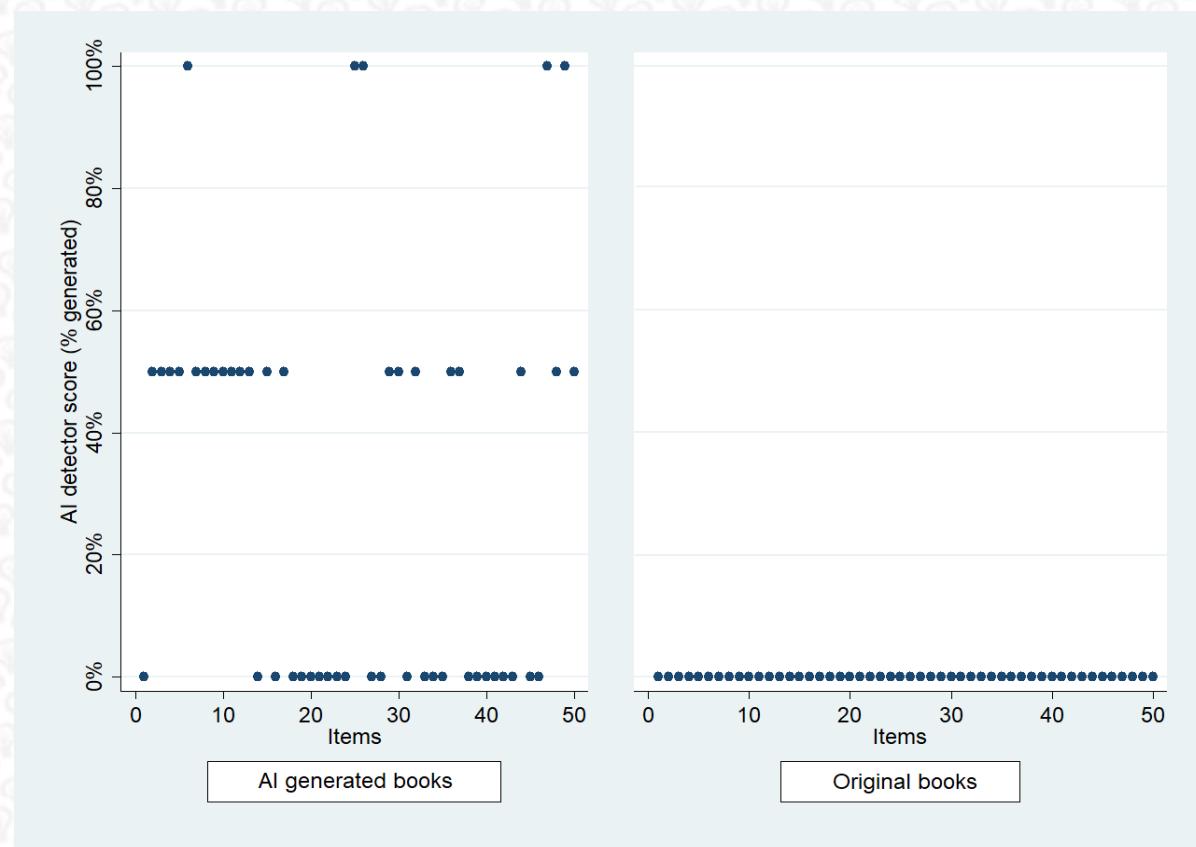


Note. Data obtained from STATA version 14.

Content at Scale was the second verifier analyzed, obtaining a lower data dispersion for texts written by humans, and a slight dispersion for texts generated by AI, as shown in Figure 3. It also obtained a score of 0.7600 in the AUROC as shown in Figure 4. It also registered a score of 79% in the F1 indicator, and successfully detected all original content as not generated by AI with a mean of 0% (0.00), but it had difficulties in detecting AI-generated content with a mean of only 31% (0.31).

Figure 3

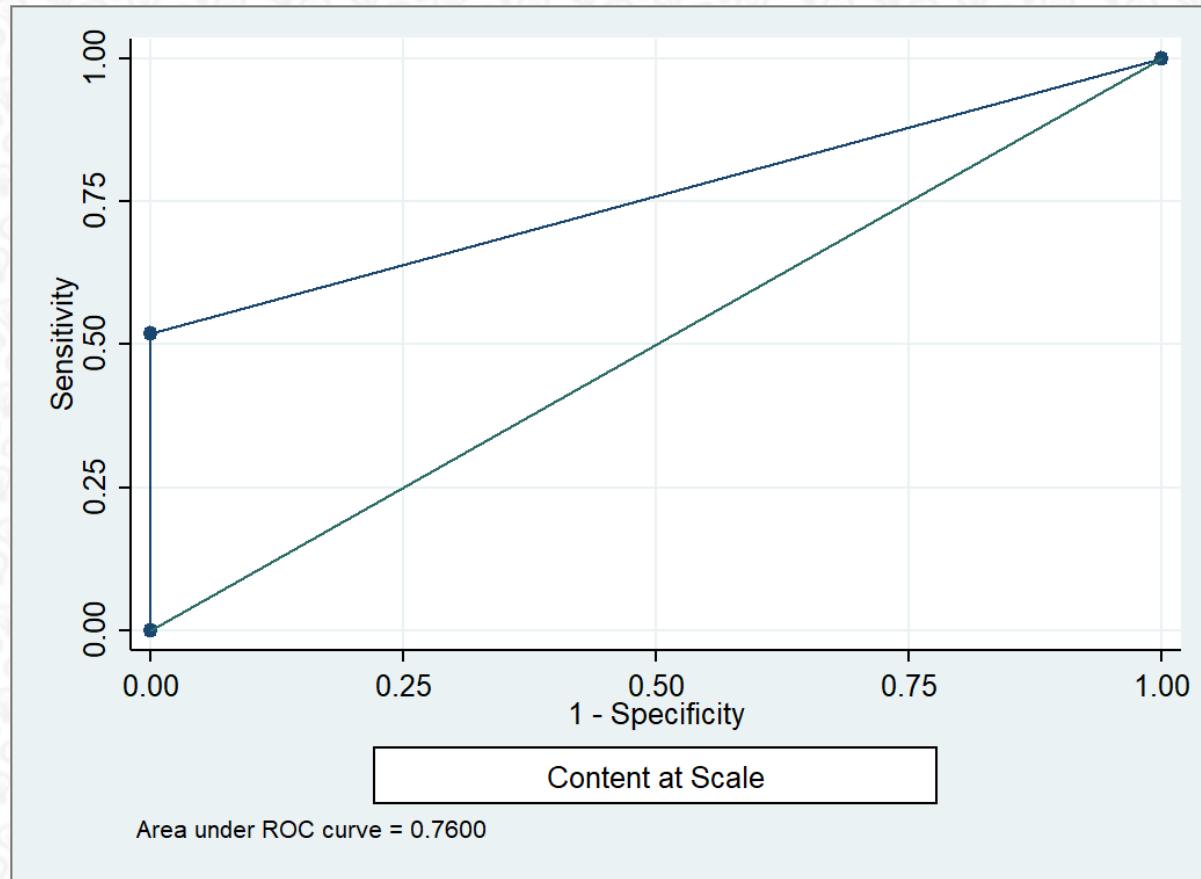
Scatterplot of Content at Scale's output



Note. Data obtained from STATA version 14. Content at Scale returned three outcomes: “Reads like AI,” which for the purposes of this article was considered 100% AI; “Hard to tell,” which was assigned a score of 50%; and “Passes like a human,” which scored 0%. This tool successfully identified all original books written by humans.

Figure 4

ROC curve of Content at Scale's output

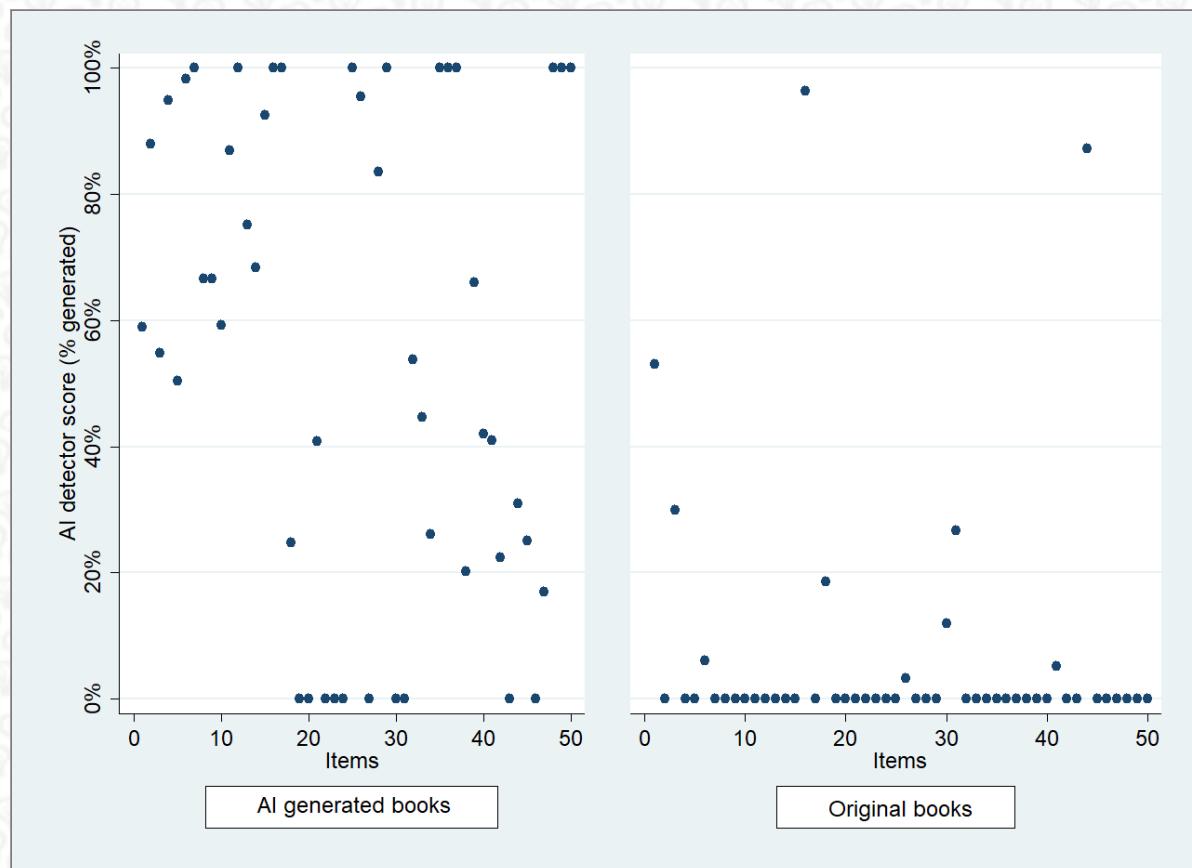


Note. Data obtained from STATA version 14.

The third verifier was ZeroGPT. It struggled to identify original content, with a mean of 6.7% (0.067), but was moderately successful in identifying AI-generated content, as shown in the scatter plot in Figure 7. It shows an AUROC of 0.8500 (Figure 8), a mean of 55.9% (0.559), and an F1 score of 69%.

Figure 5

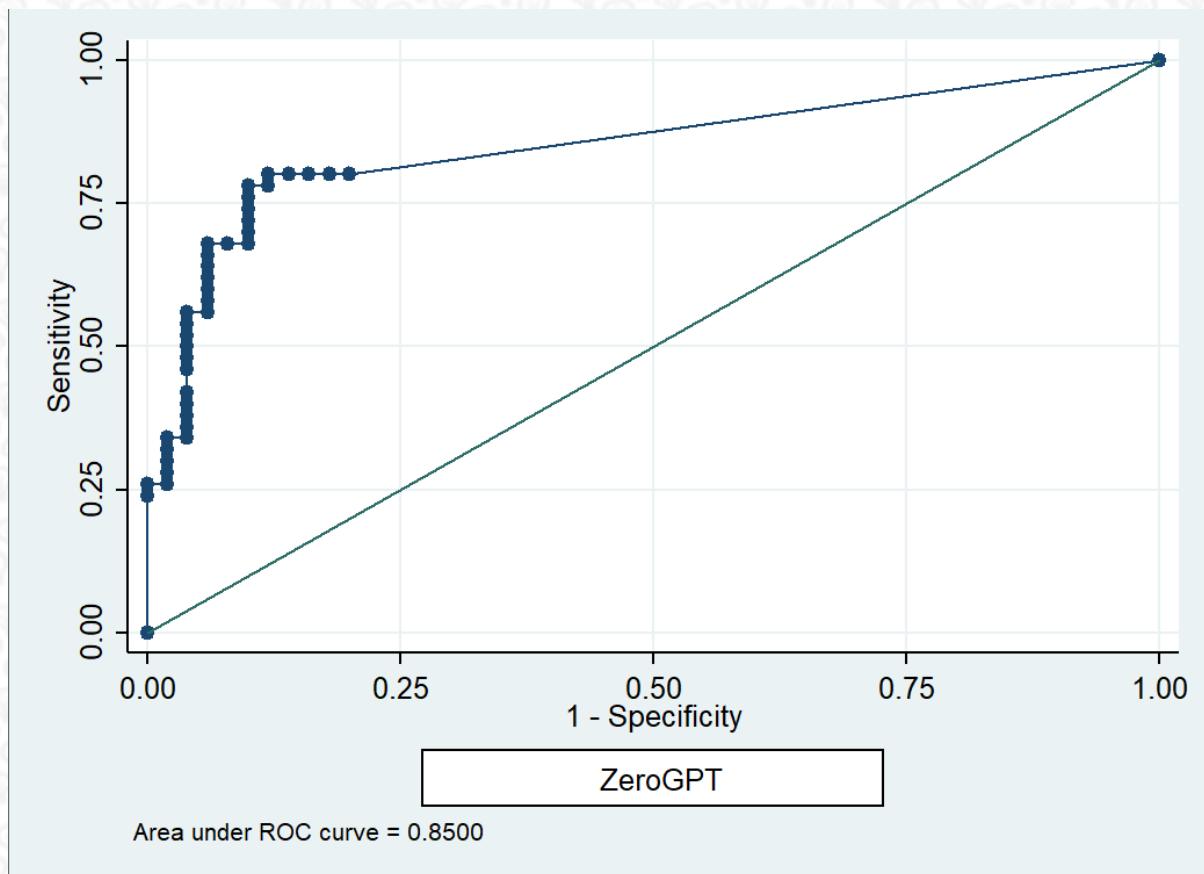
Scatter plot of ZeroGPT results



Note. Data obtained from STATA version 14. Like Scribber, this tool offered probability percentages from 0 to 100% of being generated by AI.

Figure 6

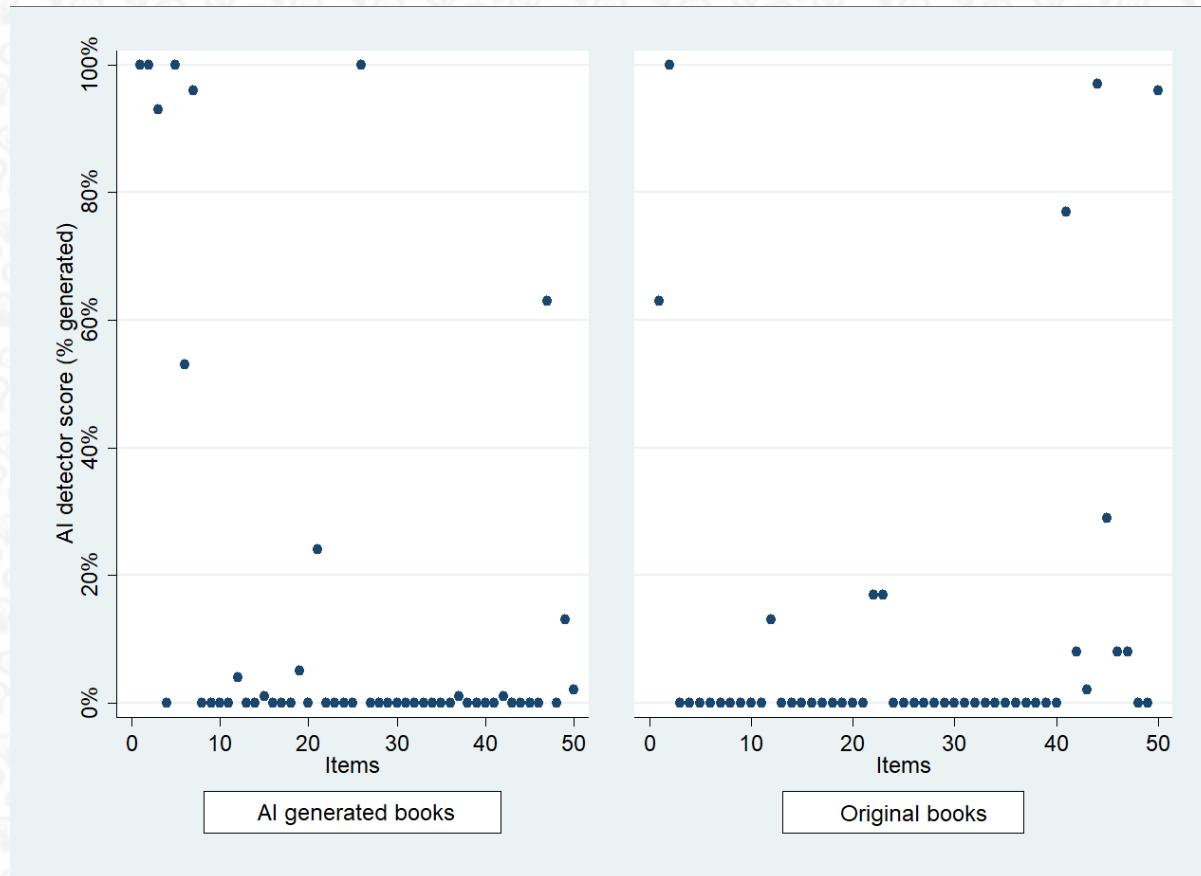
ROC curve of ZeroGPT's output



Note. Data obtained from STATA version 14. ZeroGPT ranked second in the AUROC of the four AI detection software programs analyzed. The threshold for its generation to be considered positive was +50% for the purposes of this article.

Lastly, Scribber was the fourth fact-checker analyzed. It showed inconsistent results when distinguishing between AI-generated and original content, as shown in the scatter plot in Figure 5. Furthermore, it obtained an AUROC of 0.5288 as shown in Figure 6, 25% in the F-score, and a mean of 10% (0.107) for original content and 15% (0.1512) for AI-generated content. The ROC of this tool closely resembled the diagonal of a random guess.

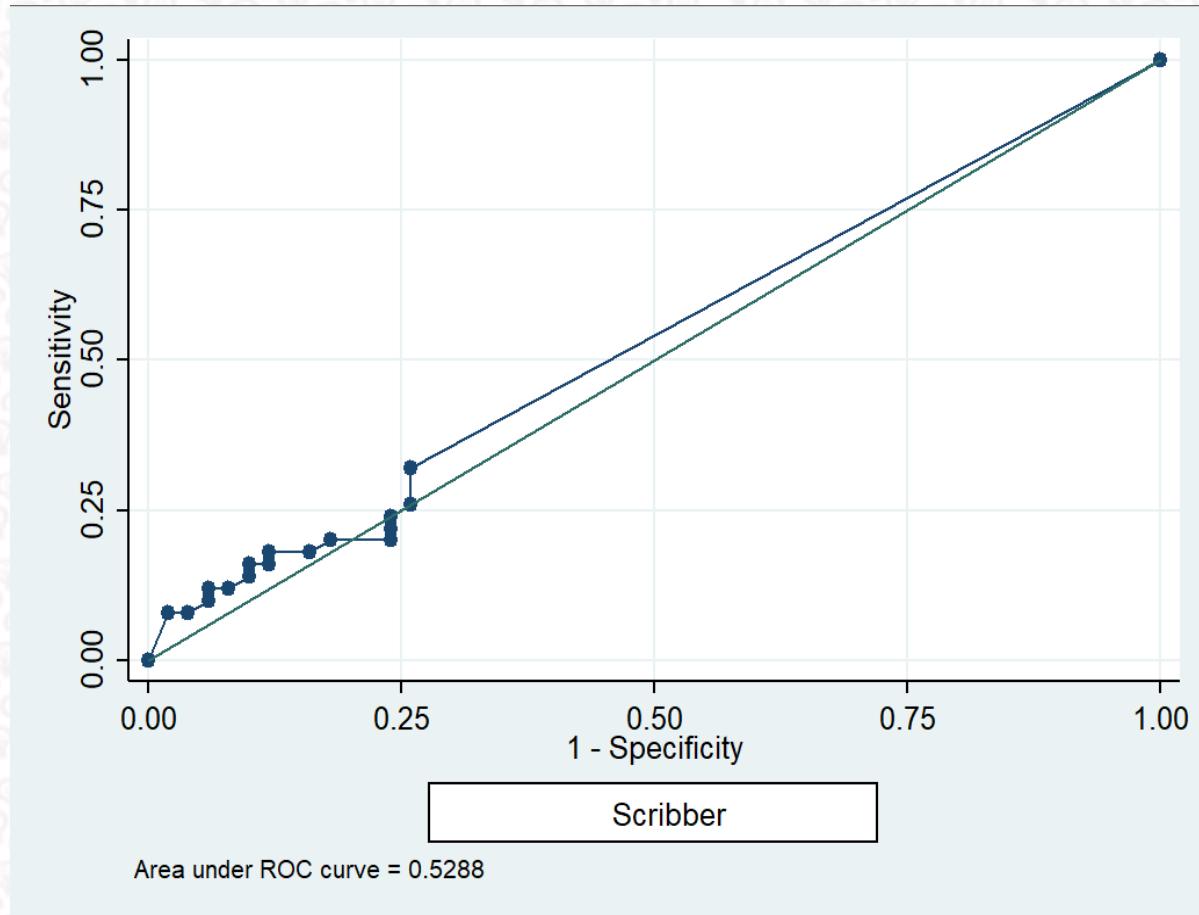
Figure 7
Scatter plot of Scribber's results



Note. Data obtained from STATA version 14. This checker detected many false negatives for AI-generated content. This tool offered probability percentages ranging from 0% to 100% that the content was AI-generated.

Figure 8

ROC curve of Scribber's output



Note. Data obtained from STATA version 14. Scribber. The AUROC closely matched the diagonal of a random guess.

Overall, these four tools demonstrated moderate success in identifying original content but exhibited varying degrees of effectiveness in distinguishing AI-generated content. CopyLeaks stood out with the highest accuracy, achieving an F-score of 99%. Content at Scale performed well in detecting original content, but struggled with AI-generated content, with an F-score of 79%. Meanwhile, Scribber showed inconsistent results, indicating challenges in accurately discerning between original and AI-generated content, with an F-score of 25%. And ZeroGPT showed mixed performance, with notable success in identifying AI-generated content, but less accuracy in detecting original content, with an F-score of 69%. These results are shown in Table 1.

Table 1

Means and F1 scores of AI plagiarism detection software.

Software	Copy Leaks	Content at Scale	Scribber	ZeroGPT
F1 Score	99%	79%	25%	69%
AUROC	0.99	0.76	0.5288	0.85

Note. A summary of the results.

| Discussion

The four AI-powered plagiarism checkers analyzed in this article were mostly able to identify original human texts, but they performed differently when identifying AI-generated content. The most accurate tool was Copy Leaks, which successfully identified all human-generated text and almost all AI-generated content. With the results of the analysis, we proved the alternative hypothesis "at least one of the verifiers differs significantly in its level of precision." Considering Copy Leaks showed a great level of precision on identifying AI generated work.

The limitations of this work are the small sample size, and the type of content analyzed, as ChatGPT was asked to generate literary samples from various genres, including romance, thriller, horror, science fiction, and others. The results of AI detection tools may vary depending on the type of content being analyzed, and the internal mechanisms of each tool may vary depending on the input type. Another limitation is the lack of explanation as to what parameters did the AI verifiers use to identify plagiarism by AI. It seems for the free tier of their services, they only show a percentage of likelihood for plagiarism, but not their actual criteria. During the time of this study, these checkers were free to use, and each offered a paid subscription that included access to more features. For this paper we used the paid version for trial purposes, and with the paid versions some of the parameters for identifying plagiarism were shown, which were mostly the percentages of matches of sentences and words of known AI work or other online material. For the sake of brevity and the scope of this paper, we decided not to use the paid versions, however analyzing the paid's version results could be used for other future papers.

ChatGPT not only has the potential to disrupt academia, but also education and culture. The first is due to the use of this tool to cheat on assignments, essays, and exams; the second is due to its ability to mass-produce literary works, displacing human writers. Among educators, AI checkers can complement plagiarism detection tools to comprehensively assess all aspects of academic plagiarism. They can also be used by publishers to verify the authorship of content. Caution should be exercised when using these checkers, as the information entered can be leaked or used to train other artificial intelligences.

Supplemental materials

Sample of the generated text

"He had been running for hours, but he knew he couldn't stop. The forest was dark and full of dangers, but it was better than being caught by the soldiers who were after him. He had stolen a precious artifact



from the king's palace, a golden orb that was said to have magical powers. He didn't know what it could do, but he knew it was worth a fortune. He had planned to sell it to the highest bidder, but his escape had gone wrong. He had been spotted by a guard, who had raised the alarm. Now, he was a fugitive, hunted by the king's men and their dogs.

He stumbled over a root and fell to the ground. He quickly got up and checked his backpack. The orb was still there, shining in the moonlight. He felt a surge of hope. Maybe he could still make it. He had heard of a secret passage that led to the border, where he could find a smuggler who would take him to another country. He just had to find it before they found him.

He heard a howl in the distance. They were getting closer. He ran faster, ignoring the pain in his legs and the blood on his arm. He had to get away. He had to survive. He had to finish his adventure."

| **Bibliography**

Adeshola, I., & Adepoju, A. P. (2023). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments*, 32(10), 6159–6172. <https://doi.org/10.1080/10494820.2023.2253858>

Almeida, R. M. V., de Albuquerque Rocha, K., Catelani, F., Fontes-Pereira, A. J., & Vasconcelos, S. M. (2016). Plagiarism allegations account for most retractions in major Latin American/Caribbean databases. *Science and engineering ethics*, 22(5), 1447-1456. <https://doi.org/10.1007/s11948-015-9714-5>

Biörck, J., & Eriksson, S. (2023). *Detecting Plagiarism with ChatGPT Using Prompt Engineering* [Tesis de licenciatura, Escuela Electrical Engineering and Computer Science]. <https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-330996>

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., & Xie, X. (2018). A survey on evaluation of large language models. *ACM transactions on intelligent systems and technology*, 37(4), 1-45. <https://doi.org/10.48550/arXiv.2307.03109>

Content at Scale (2024). *Introducing WriteWell to Humanize Your AI Writing with Unlimited AI Detector Credits*. BrandWell. <https://brandwell.ai/ai-content-detector/>

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>

Copyleaks (2024). *AI Detector*. Copyleaks. <https://copyleaks.com/ai-content-detector>

Currie, G. M. (2023). Academic integrity and artificial intelligence: is ChatGPT hype, hero or heresy? *Seminars in Nuclear Medicine*. WB Saunders, 53 (5) <https://doi.org/10.1053/j.semnuclmed.2023.04.008>

Dahl, S. (2007). Turnitin®: The student perspective on using plagiarism detection software. *Active Learning in Higher Education*, 8(2), 173-191. <https://doi.org/10.1177/1469787407074110>

Deng, J., & Lin, Y. (2022). The benefits and challenges of ChatGPT: An overview. *Frontiers in Computing and Intelligent Systems*, 2(2), 81-83. <https://doi.org/10.54097/fcis.v2i2.4465>

Fan, J., Upadhye, S., & Worster, A. (2006). Understanding receiver operating characteristic (ROC) curves. *Canadian Journal of Emergency Medicine*, 8(1), 19-20. [doi:10.1017/S1481803500013336](https://doi.org/10.1017/S1481803500013336)

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*, 1-18. <https://doi.org/10.1101/2022.12.23.521610>

Goutte, C., & Gaussier, E. (2005). *A probabilistic interpretation of precision, recall and F-score, with implication for evaluation*. En 27th European Conference on IR Research, ECIR 2005, Santiago de Compostela, Spain. https://doi.org/10.1007/978-3-540-31865-1_25

Grassini, S. (2023). Shaping the future of education: exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), 692. <https://doi.org/10.3390/educsci13070692>

Hua, H. U., Kaakour, A. H., Rachitskaya, A., Srivastava, S., Sharma, S., & Mammo, D. A. (2023). Evaluation and comparison of ophthalmic scientific abstracts and references by current artificial intelligence chatbots. *JAMA ophthalmology*, 141(9), 819-824. DOI:10.1001/jamaophthalmol.2023.3119

Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature say. *Online Journal of Communication and Media Technologies*, 13(4), e202346. <https://doi.org/10.30935/ocjmt/13572>

Kacena, M. A., Plotkin, L. I., & Fehrenbacher, J. C. (2024). The Use of Artificial Intelligence in Writing Scientific Review Articles. *Current Osteoporosis Reports*, 22, 115-121. <https://doi.org/10.1007/s11914-023-00852-0>

Khalil, M., & Er, E. (2023). *Will chatgpt get you caught? rethinking of plagiarism detection*. En In: Zaphiris, P., Ioannou, A. (Coordinadores), 10th International Conference, LCT 2023, Held as Part of the 25th HCI International Conference, HCII 2023, Copenhagen, Denmark. https://doi.org/10.1007/978-3-031-34411-4_32

Khaled, F., & Al-Tamimi, M. S. H. (2021). Plagiarism detection methods and tools: An overview. *Iraqi Journal of Science*, 62(8), 2771-2783. <https://doi.org/10.24996/ijss.2021.62.8.30>

Kumar, P. M., Priya, N. S., Musalaiah, S. V. V. S., & Nagasree, M. (2014). Knowing and avoiding plagiarism during scientific writing. *Annals of medical and health sciences research*, 4(3), 193-198. DOI: 10.4103/2141-9248.141957

Lancaster, T. (2003). *Effective and efficient plagiarism detection* [Tesis de Doctorado, School of Computing, Information Systems and Mathematics South Bank University]. https://www.academia.edu/168972/Effective_and_Efficient_Plagiarism_Detection

Landero, E. A. R., Basurto, A. C., & Ramírez, S. G. V. (2024). ChatGPT en la elaboración de trabajos académicos. *Ingenio y Conciencia Boletín Científico de la Escuela Superior Ciudad Sahagún*, 11(22), 157-161. <https://doi.org/10.29057/escs.v11i22.12553>

Masic, I. (2012). Plagiarism in scientific publishing. *Acta Inform Med.*, 20(4), 208-213. DOI:10.5455/aim.2012.20.208-213

Naik, R. R., Landge, M. B., & Mahender, C. N. (2015). A review on plagiarism detection tools. *International Journal of Computer Applications*, 125(11). DOI: [10.5120/ijca2015906113](https://doi.org/10.5120/ijca2015906113)

Rathore, B. (2023). Future of AI & generation alpha: ChatGPT beyond boundaries. *Eduzone: International Peer Reviewed/Refereed Multidisciplinary Journal*, 12(1), 63-68. <https://doi.org/10.56614/eiprmj.v12i1y23.254>

Rivera, H. (2023). La integridad científica ante los plagios fabricados con el ChatGPT. *Revista Médica del Instituto Mexicano del Seguro Social*, 61(6), 857-862. doi: 10.5281/zenodo.10064458

Scribber (2024). *Free AI Detector*. Scribber. <https://www.scribbr.com/ai-detector/>

Sharma, H., & Verma, S. (2020). Insight into modern-day plagiarism: The science of pseudo research. *Tzu-Chi Medical Journal*, 32(3), 240-244. DOI: 10.4103/tcmj.tcmj_210_19

Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). *Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation*. En Sattar, A., Kang, B. (Coordinadores), AI 2006: Advances in Artificial Intelligence. Berlin Heidelberg. https://doi.org/10.1007/11941439_114

Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature medicine*, 29, 1930-1940. <https://doi.org/10.1038/s41591-023-02448-8>

Weber-Wulff, D. (2014). *False feathers: A perspective on academic plagiarism*. Springer Science & Business. DOI: 10.1007/978-3-642-39961-9

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, 10(5), 1122-1136. DOI: 10.1109/JAS.2023.123618

Youmans, R. J. (2011). Does the adoption of plagiarism-detection software in higher education reduce plagiarism?. *Studies in Higher Education*, 36(7), 749-761. <https://doi.org/10.1080/03075079.2010.523457>

ZeroGPT (2024). *Trusted GPT-4 & ChatGPT AI Content Detector AI Checker Tool*. ZEROGPT.cc. <https://zerogpt.cc/>

Zhou, Z. H. (2021). *Machine learning*. Springer Nature. Singapore <https://doi.org/10.1007/978-981-15-1967-3>