

Artificial Intelligence Similarity Index Detection App as an Anti-Plagiarism Campaign against Intellectual Property Crime

M. Iman Wahyudi

Universitas Islam Negeri Sultan Maulana Hasanuddin Banten, Indonesia
iman.wahyudi@uinbanten.ac.id

Sabo Jamilu Sa'idu

Ahmadu Bello University zaria, Nigeria

Abstract

This paper aims to discuss plagiarism according to Indonesian and Islamic laws and develop applicable solutions for detecting similarity index plagiarism. Plagiarism is the act of taking someone else's work and claiming it as theirs without giving recognition or credit to the original owner. This practice is against positive and Islamic laws. The method used in this paper is RAD (Rapid Application Development), one of the software development methods that emphasises the speed and flexibility of the results of the trials. We found that this application could detect the similarity index of a text composed by an AI tool like ChatGPT. This application can be implemented in education but still has limitations. This application cannot detect plagiarism from sources unlisted in the database. Therefore, further research must address this problem.

Keywords: AI Chatbot, plagiarism, Similarity Index Detection App

Introduction

The digital era is marked by ongoing changes in business transformation that affect various aspects of life. These advancements, facilitated by artificial intelligence, have made life more convenient for humans through applications like navigation, autopilot, smart homes, and smart cities. According to Romansky and Noninska (2020), the Digital Age is defined by technological solutions in virtual environments, digital services, intelligent applications, machine learning, and knowledge-based systems.

Numerous opportunities arise in this digital era, including the utilisation of the Internet for online business purposes (Ordoñez de Pablos, 2022), the ability to work remotely from anywhere, engaging in online learning (Saykili, 2019), leveraging blockchain technology (Heo et al., 2021), and experiencing various conveniences through the digitisation of business processes.

However, the advantages offered by technology come with some challenges. In the digital era, there is a competition to maintain traditional skills while increasingly relying on technology. It is crucial to possess the ability to use technology effectively and understand digital issues. Digital literacy plays a key role in actively and productively participating in an ever-connected society.

The challenges of the digital era are incredibly diverse and complex. Data has become immensely valuable, making security and privacy issues, like data protection, illegal access, and even data theft, critical challenges (Romansky & Noninska, 2020). Perhaps the most significant challenge is developing innovative and effective technologies without compromising security and privacy (Mantoro et al., 2016). Technology developers must consistently consider the ethical implications and social impact of the technologies they create. Another hurdle is the effective and accurate management of information.

In the digital realm, plagiarism can profoundly impact an individual's or organization's reputation. Many online platforms have anti-plagiarism policies and can take necessary action if a violation occurs. Despite being condemned, plagiarism remains a prevalent form of academic misconduct (Gerald, 2021). It involves taking someone else's work and presenting it as one's own without recognising or crediting the original author. In the digital age, plagiarism has become easier due to the vast amount of available online information and its easy accessibility. The concentration of plagiarism in the digital world is high because of the abundance of resources, such as articles, books, journals, websites, and blogs.

Various forms of plagiarism in the digital world include copying and pasting text from another source without providing attribution or acknowledgement, taking text from another source and altering words or phrases to make it appear original, or paying someone to write content and using it without crediting the original author. Additionally, using copyrighted content without permission or acknowledgement is another common form of plagiarism. Machines can also contribute to this, and nowadays, institutions are growing concerned about the potential negative impacts that large language models (LLMs) like ChatGPT could have on society, such as the spread of fake news, plagiarism, and issues

related to social security (Guo et al., 2023).

Plagiarism involves taking or using someone else's work without giving due recognition or obtaining permission from the original author. It is considered a form of intellectual crime because appropriating someone else's work and presenting it as our own can harm the intellectual property rights of the original author and violate copyright. Additionally, plagiarism is deemed unethical and has the potential to tarnish an individual's reputation in academic, writing, and other professional spheres. Therefore, everyone needs to respect intellectual property rights and ensure that the works they use do not infringe on the copyrights of others.

Data from the Ministry of Research and Technology-Higher Education of Indonesia indicates a concerning rise in cases of plagiarism of scientific works in Indonesia over the last decade (Sukaesih, 2018). While there have been sanctions for academic crimes in Indonesia based on the Law Act No. 20 Year 2003 about the National Education System (Sholihah, 2019). Research by Weriansyah and Ramadani (2022) highlights the unpreparedness of existing intellectual property regulations to address potential corporate criminal activities.

For the creator of the original work, plagiarism can diminish the credibility and recognition of their produced work, impacting their reputation and career and hindering progress in science and technology. On the other hand, perpetrators of plagiarism face legal consequences, disqualification from competitions or scientific recognition, and potential damage to their future careers and reputations. Beyond legal implications, plagiarism also reflects a lack of integrity, morality, and ethics in the academic and professional realms.

Research by Spajc et al. (2017) indicates that students may engage in plagiarism due to the convenience offered by advances in information and communication technology. The convenience includes easy access to advanced artificial intelligence-based tools like ChatGPT (Perkins, 2023). Sakirin (Sakirin & Ben Said, 2023) finds that 70% of users prefer ChatGPT-powered conversational interfaces over traditional techniques, citing convenience, efficiency, and personalization. This suggests a growing desire among users to incorporate these technologies into their daily lives.

The study also reveals that historical standards were not necessarily lower than current ones, and the potential for condemning plagiarists has always existed (Moten, 2014). Therefore, everyone must understand that plagiarism is unacceptable in any context. All produced work should be grounded in original ideas and research while respecting the copyrights of others.

In Indonesia, the issue of plagiarism has been addressed in Article 70 of Law Number 20 of 2003 concerning the National Education System (Alfazri, 2014), and it is explicitly implied in the Copyright Law (Hakim, 2018). Articles 41, 42, 43, and 44 of UUHC No. 28 2014 affirm the illegality of plagiarism, both in positive law and Islamic law (Pitrian Ramadhan, 2019).

The anti-plagiarism campaign aims to raise awareness about the issue of plagiarism and promote actions to prevent and combat it. Various methods are employed in this cam-

paign, including educating students, lecturers, and the general public on what plagiarism entails and how to avoid it. The campaign highlights the detrimental effects of plagiarism on academic and personal reputation, emphasizes the presence of plagiarism detection software, provides examples of plagiarism actions and their consequences, and boosts social media outreach through banners, posters, and brochures on campuses. Educators play a crucial role in educating students, monitoring for suspicious activity (Ventayen, 2023), whether on social media or in organizations. Strict sanctions against plagiarism perpetrators are implemented by establishing clear rules and offering knowledge about copyright and intellectual work protection.

Several solutions are proposed to curb this academic crime. Kolhar & Alameen (2021) have developed anti-plagiarism software to detect plagiarism in students' assignments. Edward Tian, a senior at Princeton University, created an app to identify text written by ChatGPT (Bowman, 2023). Additionally, research by Mostofa et al. (2021) analyses researchers' awareness of plagiarism and the impact of plagiarism detection tools on preventive actions. Some opt for existing internet-based applications (Chandere et al., 2021), whether free access (Gamba et al., 2014) or paid (Dong & Shi, 2021). Saini's research (Saini et al., 2016) involves indexing the original document and using cosine similarity to compare plagiarised documents with a set of previously saved documents, mainly focusing on the similarity index on documents compared with those stored on the Internet.

This paper proposes an application serving as a tool to check the similarity index based on artificial intelligence, building upon Suwaid Aslam's work (Salam, 2023). This application, developed as a simple Python prototype, is user-friendly, facilitating easy implementation in its intended use.

The initial analysis in this paper involves data collection using the literature study and conducting reviews from various sources such as articles, journals, scientific research, and other relevant materials. The application development follows the RAD (Rapid Application Development) method, emphasizing speed and flexibility. This method aims to produce software quickly, with stages including planning, design, construction, integration, and implementation, chosen for its time efficiency in application development (Olorunshola & Ogwueleka, 2022).

Plagiarism in Indonesian and Islamic Laws

This paper explores chatbots as potential sources of plagiarism and provides a brief overview of their functioning. Chatbots are computer programs capable of interacting with humans through chat or text messages. The concept of an AI chatbot involves leveraging artificial intelligence (AI) to create smarter chatbots that can better understand and respond to users.

Various AI technologies contribute to chatbot development, including Natural Language Processing (NLP), Machine Learning (ML), and Sentiment Analysis. NLP enhances chatbots' understanding of human language, allowing them to analyze user input and

respond more appropriately. ML enables chatbots to learn from historical data, improving their performance and understanding user preferences. Sentiment Analysis helps chatbots gauge user emotions based on entered text, allowing them to tailor responses accordingly for better support (Ashfaq, 2022).

In the development of AI chatbots, attention must be given to data security and user privacy. It is crucial to design chatbots with data security in mind and inform users about the type of data collected and its usage. Many users tend to copy and paste answers from chatbots, using AI to check for plagiarism. Online tools like Turnitin, PlagScan, Grammarly, and Quetext are popular for detecting plagiarism. They use AI algorithms to compare text with vast databases and online sources to identify suspicious similarities. For example, PlagScan and Grammarly not only serve as grammar checkers but also employ AI to detect plagiarism.

Creating an AI Chatbot detection application involves considering various factors such as the application's purpose, the type of Chatbot to be detected, and the data used to train the detection model. The accuracy of distinguishing between chatbots and humans is crucial, especially with the increasing complexity of chatbots.

This paper aims to develop an application to check the results of Chatbot-generated content, particularly in scientific papers. Plagiarism screening applications work by comparing inserted text with a database of collected documents. Typically, these applications use natural language processing technology to understand the language structure of the input document and identify similarities with other documents.

The process of plagiarism screening applications begins with collecting and inputting data into the database. It then compares the entered text with the document database, analyses the language for similarities with other documents, and generates a report detailing the percentage of similarity between the entered text and documents in the database. Plagiarism detection involves determining whether a submitted written work is original or contains elements of plagiarism, such as using ideas, words, or sentences from other sources without proper credit. This is especially critical in fields like academia, journalism, and publishing, where maintaining intellectual integrity and honesty is paramount. It is essential to note that, like any tool, this application has limitations. It cannot detect plagiarism from sources not in the database or differentiate between common or distinctive uses of words or phrases in specific contexts, potentially leading to inaccurate results.

Developed using Python, a widely-used programming language in artificial intelligence (AI), the application performs tasks like natural language processing, computer vision, and machine learning. Python's flexibility and ease of use make it a preferred choice for developing AI applications, given its powerful libraries and frameworks in machine learning, computer vision, and natural language processing.

Accessible through a web browser, the application boasts a user-friendly interface designed for ease of use. User-friendly applications have characteristics such as easy-to-understand appearance, straightforward navigation, quick response to user input, and the provision of clear information. This application aims to provide users with simple, un-

derstandable information, maintaining a consistent design for seamless navigation and offering options for customization based on user preferences and needs.

The AI Chatbot detection application addresses the growing use of chatbots in various services. As companies increasingly integrate chatbots into their offerings, these applications play a crucial role in identifying chatbots on platforms or applications. Chatbots, being computer programs facilitating user interaction through text or voice messages, offer an efficient alternative to direct customer service contact.

The AI Chatbot plagiarism detection application involves collecting data from various platforms, processing and training it using machine learning algorithms to build Chatbot detection models. These models learn patterns from real and fake chatbots to create classifiers identifying chatbots on platforms or applications. The Chatbot detection model can be integrated into other applications, such as browser extensions or mobile apps.

However, using AI Chatbot detection applications raises challenges, particularly concerning privacy and data security. Safeguarding user data confidentiality and ensuring data protection are crucial considerations. Plagiarism is viewed as an academic crime and contrary to academic integrity. Laziness often leads to plagiarism in the educational sphere, running counter to Islamic Sharia norms.

Addressing plagiarism in Indonesia, Article 70 of Law Number 20 of 2003 concerning the National Education System has been discussed by various scholars. The Indonesian Ulema Council (MUI) has issued a fatwa related to the protection of intellectual property rights, considering these rights as *huquq maliyyah* (property rights) with legal protection and economic value. MUI emphasizes that violations of intellectual property rights, including plagiarism, are tyrannical and unlawful, allowing for the inclusion of these rights in contracts and inheritance, all in accordance with Islamic law.

The MUI fatwa is grounded in the Qur'an, specifically in Surah an-Nisa verse 29 (RI, 1998), and also in Surah as-Shuara verse 183 (RI, 1998), highlighting the significance of intellectual property rights (Yusup, 2015) as a form of protected property. The Prophet Muhammad emphasised the sacredness of both one's blood and treasure, underscoring the sanctity of property.

The Fatwa Commission, during that time, incorporated the legal principle that no action should be taken against others' property rights without permission. Scholars from the Maliki, Shafi'i, and Hambali madhhabs assert that copyright on original work is classified as valuable property. In cases where the owner no longer exists, the copyright is deemed beneficial for Muslims. The rule from Bughyatul Mustarsyidin (Abdurrahman, n.d.) states that if someone wishes to reclaim property borrowed without permission and is unaware of the owner, the property should be returned to the state treasury or utilised for the benefit of Muslims.

In Islamic law, plagiarism is associated with the concept of *riba*, portraying it as gaining profits from the work of others without permission or fair collaboration. This deems plagiarism contrary to Islamic justice and ethical principles. From the Islamic perspective, copyright is considered private property, respected as a form of acknowledgment of one's

work. Taking someone else's work without permission or modifying it and claiming it as one's own is unjustified in Islam, akin to cheating, which is also prohibited. Islamic teachings strongly emphasize honesty and justice, principles that plagiarism violates. Therefore, Islamic law views plagiarism as a severe violation of others' rights, ethics, morals, and justice, necessitating efforts for the prevention and punishment of perpetrators.

In the academic context, plagiarism is deemed contrary to the principles of honest and fair academics. As Muslims, upholding the principles of honesty and avoiding actions that harm others is essential. Plagiarism should be avoided and considered an unethical act. In the Islamic view, this discourages individuals from creating and developing works that benefit society, supporting copyright as a form of recognition for efforts and creations.

The paper proposes an intelligent application as an alternative solution to combat plagiarism. The app compares similarities between different texts or documents, involving stages such as determining the type of similarity index, collecting relevant data, calculating the similarity index, and visualizing results through graphs or tables for efficient understanding and evaluation. In tests conducted, the system successfully detected the similarity index of Chatbot ChatGPT copy-paste results.

This application boasts several advantages, including time and effort savings, high precision, detailed analysis, and user-friendliness. It finds utility across various fields, enabling quicker task completion than manual methods. The detection app incorporates an algorithm that accurately identifies similarities between documents and sentences, helping prevent errors in determining similarity.

Plagiarism detection assesses whether a written work has copied or stolen content from another. Numerous tools and technologies, such as search engines, plagiarism detection software, and online services, facilitate plagiarism detection. These technologies compare suspected plagiarism papers with original works using text benchmarking algorithms. Applications can be instructed to analyse text at a detailed level, checking for similarities in words, phrases, or even entire sentences to identify any suspicious resemblances between texts. In the case of AI Chatbot results, plagiarism detection technology can be employed to scrutinize whether the answers contain elements of plagiarism.

It is crucial to acknowledge that plagiarism detection technology is not flawless and can make errors in identifying plagiarism. Therefore, manually analyzing AI Chatbot results for plagiarism by comparing them with the information source is recommended. Overall, conducting plagiarism detection on AI Chatbot results is essential to ensure that chatbot responses are free from plagiarism and to prevent unethical actions. It's worth noting that relying solely on technology for plagiarism detection may not suffice due to factors like context and proper citation. Manual testing and evaluation are necessary to determine if plagiarism is a mere coincidence.

Conclusion

Plagiarism is more than just a violation of intellectual property rights. It is considered intellectual theft with serious legal repercussions. When an individual presents someone else's work as their own without due credit or consent, it can lead to significant legal consequences. In Islamic law, plagiarism is not only a legal issue but also a profound ethical and moral violation. Dishonestly claiming someone else's work is viewed as detrimental to the original owner, contradicting the Islamic principles that encourage honesty and fairness in all aspects of life, including intellectual endeavors.

AI Chatbot detection applications represent crucial technological advancements in distinguishing between authentic and fabricated content. However, the successful utilisation of these applications hinges on addressing challenges related to data privacy and security. In our trials, these applications have demonstrated the capability to detect similarity indexes, specifically identifying copy-paste chatbot results from ChatGPT. This indicates their potential as tools for checking artificial intelligence-based similarity indexes.

This application offers various advantages, such as saving time and effort, providing high precision, enabling detailed analysis, and being user-friendly. Its versatility makes it applicable in diverse fields, particularly within the realm of education. Yet, like any tool, this application has its limitations. Notably, it cannot identify plagiarism from sources outside the listed database. Further research on this issue is warranted, exploring big data as a potential source for enhancing the application's capabilities.

Reference

- Abdurrahman, S. (n.d.). *Bughyatul Mustarsyidin*. Al Haramain.
- Alfarisi, U. (2018). *Kajian Plagiarisme: Studi Perbandingan Hukum Islam dan Hukum Positif di Indonesia*. JURISDICTION, 9(1). <https://doi.org/10.18860/j.v9i1.5134>
- Alfazri, R. (2014). *Analisis Yuridis Terhadap Tindak Pidana Penjiplakan Karya Ilmiah Orang Lain (Plagiat) dalam Hukum Pidana*. ETD Unsyiah, 1(1).
- Ashfaque, M. W. (2022). *Analysis of Different Trends in Chatbot Designing and Development: A Review*. ECS Transactions, 107(1). <https://doi.org/10.1149/10701.7215ecst>
- Bowman, E. (2023). *A College Student Created an App that can Tell Whether AI wrote an Essay*. NPR.
- Chandere, V., Satish, S., & Lakshminarayanan, R. (2021). "Online Plagiarism Detection Tools in the Digital Age: A Review." *Annals of the Romanian Society for Cell Biology*, 25(1).

- Dong, Y., & Shi, L. (2021). *Using Grammarly to Support Students' Source-based Writing Practices. Assessing Writing*, 50. <https://doi.org/10.1016/j.asw.2021.100564>
- Gamba, M., Ocbian, M., & Gamba, M. (2014). "Plagiarism Checker as Best Free On-line Plagiarism Detection Software." *JPAIR Multidisciplinary Research*, 18(1). <https://doi.org/10.7719/jpair.v18i1.292>
- Geraldi, J. (2021). "Plagiarism in Project Studies". *Project Management Journal* (Vol. 52, Issue 1). <https://doi.org/10.1177/8756972820982443>
- Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., Yue, J., & Wu, Y. (2023). *How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection*. November, 1–20. <http://arxiv.org/abs/2301.07597>
- Hakim, G. (2018). "Perlindungan Hukum Pencipta yang Dirugikan Haknya atas Tindakan Plagiarisme." *Halu Oleo Law Review*, 2(1). <https://doi.org/10.33561/holrev.v2i1.4199>
- Heo, G., Yang, D., Doh, I., & Chae, K. (2021). Efficient and Secure Blockchain System for Digital Content Trading. *IEEE Access*, 9. <https://doi.org/10.1109/ACCESS.2021.3082215>
- Hs, D. R. (2017). "Plagiarisme Dan Korupsi Ilmu Pengetahuan." *Kordinat: Jurnal Komunikasi Antar Perguruan Tinggi Agama Islam*, 16(1).
- Kolhar, M., & Alameen, A. (2021). "University Learning with Anti-Plagiarism Systems." *Accountability in Research*, 28(4). <https://doi.org/10.1080/08989621.2020.1822171>
- Mantoro, T., Wahyudi, M. I., Ayu, M. A., & Usino, W. (2016). "Real-Time Printed Document Authentication Using Watermarked QR Code." *Proceedings - 4th International Conference on Cyber Security, Cyber Warfare, and Digital Forensics, CyberSec 2015*. <https://doi.org/10.1109/CyberSec.2015.22>
- Mostofa, S. M., Tabassum, M., & Ahmed, S. M. Z. (2021). "Researchers' awareness about plagiarism and impact of plagiarism detection tools – does awareness effect the actions towards preventing plagiarism?" *Digital Library Perspectives*, 37(3). <https://doi.org/10.1108/DLP-10-2020-0100>
- Moten, A. R. (2014). "Academic Dishonesty and Misconduct: Curbing Plagiarism in the Muslim World." *Intellectual Discourse*, 22(2).
- Olorunshola, O. E., & Ogwueleka, F. N. (2022). "Review of System Development Life Cycle (SDLC) Models for Effective Application Delivery." *Lecture Notes in Networks and Systems* (Vol. 191). https://doi.org/10.1007/978-981-16-0739-4_28
- Ordoñez de Pablos, P. (2022). "Editorial: "Digital innovation to tackle business opportunities in the digital age." *Journal of Science and Technology Policy Management* (Vol. 13, Issue 3). <https://doi.org/10.1108/JSTPM-08-2022-186>
- Perkins, M. (2023). "Academic Integrity considerations of AI Large Language Models in

- the Post-Pandemic Era: ChatGPT and Beyond.” *Journal of University Teaching and Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>
- Pitrian Ramadhan, A. (2019). *Sanksi Plagiat Dalam Penulisan Karya Ilmiah Menurut Peraturan Menteri Pendidikan Nasional Nomor 17 Tahun 2010 Dan Tinjauan Hukum Islam*. UIN Raden Fatah Palembang.
- RI, D. A. (1998). *Al-Qur'an Al Karim dan Terjemahannya*. PT. Karya Toha Putra.
- Romansky, R. P., & Noninska, I. S. (2020). Challenges of the Digital Age for Privacy and Personal Data Protection. *Mathematical Biosciences and Engineering*, 17(5). <https://doi.org/10.3934/MBE.2020286>
- Saini, A., Bahl, A., Kumari, S., & Singh, M. (2016). “Plagiarism Checker: Text Mining.” *International Journal of Computer Applications*, 134(3). <https://doi.org/10.5120/ijca2016907833>
- Sakirin, T., & Ben Said, R. (2023). “User Preferences for ChatGPT-Powered Conversational Interfaces Versus Traditional Methods.” *Mesopotamian Journal of Computer Science*. <https://doi.org/10.58496/mjcs/2023/006>
- Salam, S. (2023). *AI Generated Text Checker App*. https://github.com/SuwaidAslam/AI_Generated_Text_Checker_App
- SAYKILI, A. (2019). *Higher Education in The Digital Age: The Impact of Digital Connective Technologies*. *Journal of Educational Technology and Online Learning*. <https://doi.org/10.31681/jetol.516971>
- Shah Abd ul Raof, & Dr. Muhammad Feroz ud Din Shah Khagga. (2021). “The Legal Status Of Electronic Theft In The Light Of The Quran And Hadith, A Research Study.” *International Research Journal on Islamic Studies (IRJIS)*, 3(02). <https://doi.org/10.54262/irjis.03.02.e05>
- Sholihah, K. (2019). *Upaya Perpustakaan Perguruan Tinggi dalam Mewujudkan Integritas Akademik*. *Pustablibia: Journal of Library and Information Science*, 3(2). <https://doi.org/10.18326/pustablibia.v3i2.189-203>
- Šprajc, P., Urh, M., Jerebic, J., Trivan, D., & Jereb, E. (2017). Reasons for Plagiarism in Higher Education. *Organization*, 50(1). <https://doi.org/10.1515/orga-2017-0002>
- Sukaesih, S. (2018). “Permasalahan Plagiarisme dalam Penelitian Kualitatif di Indonesia.” *Jurnal Politikom Indonesiana*, 3(1).
- Ventayen, R. J. M. (2023). “OpenAI ChatGPT Generated Results: Similarity Index of Artificial Intelligence-Based Contents.” *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4332664>
- Weriansyah, A., & Ramadani, A. P. (2022). “The Analysis of Corporate Crime in Indonesia’s Intellectual Property Laws.” *Global Legal Review*, 2(1). <https://doi.org/10.19166/>

- Widyastuti, Y. (2020). "Analisis Fatwa Mui No. 1/Munas Vii/Mui/5/2005 Tentang Perlindungan Hak Kekayaan Intelektual (Hak Cipta) Dalam Perspektif Maá1çLaá, ¢Ah Mursalah." *Jurnal Al-Hakim: Jurnal Ilmiah Mahasiswa, Studi Syariah, Hukum Dan Filantropi*, 2(1). <https://doi.org/10.22515/alhakim.v2i1.2496>
- Yusup, D. K. (2015). "Norma-Norma Hukum Hak Asasi Manusia Dalam Al-Quran: Studi Kritis Atas Tafsir Al-Quran Surat Al-Nisá' Ayat 135." *Ayy-Syari'ah*, 17(2). <https://doi.org/10.15575/as.v17i2.649>