International Journal of Education, Information Technology and Others (IJEIT), April 20242, 7 (2), 54-63
DOI: https://doi.org/10.5281/zenodo.10968289
p-ISSN: 2654-2528 e-ISSN: 2623-2324
Accredited by Directorate General of Strengthening for Research and Development
Available online at https://jurnal.peneliti.net/index.php/IJEIT

Literature Review: Threat of Artificial Intelligence (AI) Cyber Attacks on Data Security

Ferryma Arba Apriansyah[1*], Erik Iman Heri Ujianto[2], Rianto Rianto[3]

[1,2,3]Master of Information Technology, Yogyakarta Technology University

Received: 12 February 2024
Revised: 22 February 2024
Accepted: 8 March 2024

Abstract

The purpose of this study is to present an overview of the approaches and areas of concentration in data security identification. The report analyzes AI cyberattacks on Indonesian data security through a review of the literature. In order to find relevant papers, search terms like "Cyber Attacks," "Artificial Intelligence," and "Data Security" were used in this study. 80 of the 1,057 studies that were found between 2019 and 2023 were employed in this analysis. The outcomes of this evaluation of the literature can help future researchers conduct more in-depth studies on data security. A number of factors, such as extending the study's purview, incorporating a wider range of topics, and including people who may provide additional insights, might be taken into account for future data security research.

Keywords: Cyber Attacks, Artificial Intelligence, and Data Security

(*) Corresponding Author: ferryma.arba@gmail.com

How to Cite: Apriansyah, F., Ujianto, E. I. H., & Rianto, R. (2024). Literature Review: Threat of Artificial Intelligence (Ai) Cyber Attacks on Data Security. *International Journal of Education, Information Technology, and Others*, *7*(2), 54-63. https://doi.org/10.5281/zenodo.10968289

INTRODUCTION

The rapid progress of more sophisticated digital technology and various forms of artificial intelligence (AI) have been widely applied in several areas of human life. The idea that the human brain can be compared to a computer in some way provides an organic basis for artificial intelligence (AI), which is further reinforced by psychological theories regarding how humans and animals function as machines that process information through associative storage devices (Okey, 2023) .

Significant progress has been made in a number of fields, including business, health and public services, thanks to AI. However, as technology advances, the possibility of cyber attacks can also arise and cover users. Cyberattacks that use AI are considered dangerous because they endanger user privacy and data security. The development of computer technology also has an influence. Cyberspace is a term used to describe the emerging field of computer-based communications. This new area has not only positive effects but also n`egative effects; Often some people use it as a place to commit crimes. Cybercrime is usually called cyberspace. Cybercrime is any illegal act that uses a computer network and is carried out by using computers as goods or by taking advantage of other people's losses. (Ma'rufah, 2020).

Several countries and non-state institutions are making large investments in research and development (smart cities). AI as more and more countries realize the potential of AI and declare their national AI goals. The use of AI in several national fields, including health, education, food, government and

smart cities, is planned based on the National Strategy for Artificial Intelligence in Indonesia (STRANAS-KA), which was released in 2020. (Hermawan, 2023).

The aim of a systematic literature review is to summarize all the research that has been conducted on a particular subject or field of study. Carrying out a systematic literature review is very important because it can be a basis for scientific progress and can produce new concepts or understanding for further research. To collect papers regarding cyber attack intimidation through artificial intelligence related to Indonesian data security, a comprehensive literature survey was conducted for this research. Next, the collected papers will be analyzed to determine the answers to the research questions.

To increase the validity of current ideas, this research will look at a number of empirical investigations regarding the factors that contribute to intimidation of AI cyber attacks on Indonesian data security. This research also intends to determine the intimidation of AI cyber attacks with Indonesian data security as explained in previous publications.

Cyber Attack

A cyberattack refers to an attempt to gain unauthorized access to, modify, or compromise a computer system or network. Cyberattacks can be carried out by national governments, groups, or private citizens. Data theft, harassment, and extortion are just a few of the many purposes of cyberattacks (Proceedings, 2020).

Financially and reputationally, cyberattacks may have a major impact. Identity theft, fraud, and even threats of physical violence can all be committed using stolen data. Financial loss, business disruption, or even death can result from a compromised system.

Artificial Intelligence (AI)

Artificial Intelligence basically consists of software, usually using algorithms, but its functions must be represented by physical entities, such as robots, in order to play games or communicate. AI resembles the human brain in this respect. AI research has so far been largely concentrated in specific areas. (Proceedings, 2020).

AI is described by the Organization for Economic Co-operation and Development (OECD), 2016 as the capacity of computers and systems to learn, apply knowledge, and engage in intelligent behavior. It covers a broad spectrum of cognitive functions, such as observing, understanding spoken language, thinking, learning, drawing conclusions, and demonstrating appropriate movement and manipulation of objects. To run and learn, these intelligent systems include machine-to-machine (M2M) connections, cloud computing, big data analysis, and the Internet of Things (IoT).

Data Security

Across all industries, data security is a major concern for companies. Threats to data security are increasing with the development of artificial intelligence (AI). Artificial Intelligence has the potential to create cyberattacks that are more complex and challenging to detect.

In an increasingly digital world, data security is very important. It includes procedures and policies designed to protect personal data from misuse, alteration or illegal access. To reduce risks and possible vulnerabilities, sophisticated strategies combining encryption, access limits, regular security audits and proactive monitoring are used.

It is critical for communities and companies to consistently improve their security procedures, as cyber threats become more complex as technology advances. Significant financial and reputational losses are associated with data security breaches, which also compromise privacy.(Stafford, 2022). Therefore, implementing strong data security measures not only strengthens the framework of a safe and resilient digital ecosystem but also safeguards critical information by fostering trust among stakeholders.

RESEARCH METHOD

Types of Research Methods

Systematic Literature Review (SLR) or comprehensive literature observation is the research approach used in this research. The aim of methodical literature observation, according to Phua (2010), is to present a comprehensive record of all research conducted on a particular subject or field of study. A systematic literature review involves several procedures or processes. Systematic literature reviews must adhere to strict procedures consisting of three steps: planning, reviewing, and reporting (Kitchenham, 2007 in Sánchez-Aguayo, Urquiza-Aguiar, & Estrada-Jiménez, 2021).

Literature Search Strategy

Use the PICOS Framework to search for books that will be used as learning objects. Apart from being a basis for selecting literature search terms, the PICOS Framework is also used to identify literature criteria to be used. Five components make up the PICOS Framework, according to Liberati et al. (2009) and Moher et al. (2009) in Homer (2019): population/problem, intervention, comparison, results, and research design.

Databases and Keywords

Journal articles from publications indexed by Scopus are used as a source of research data. The large number of data sources were chosen because of their ability to provide full-text journals and articles, as well as the large number of articles suitable for use as research objects.

Keywords and boolean operators (AND, OR NOT, or AND NOT) are used to search for papers or journals needed for learning. With this strategy, searching for articles that are relevant to the research topic will be easier to find, thereby broadening and narrowing the search. The terms “cyber attacks” OR “artificial intelligence” AND “data security” are used.

Literature Criteria

In this study, integration criteria or inclusion and exclusion are the two sets of criteria used. Meanwhile, exclusion criteria are used to determine that the item being searched for is not included in the article to be analyzed, while inclusion criteria are characteristics or criteria that are used to involve the article so that it can be studied. The data obtained will be filtered to see whether it meets

these two requirements to be used as a research sample. The following are the inclusion criteria for this study:

1. This journal is indexed by Scopus.
2. Articles about AI-related cyber attacks regarding information protection.
3. Articles are written in English.
4. Articles published in print media in 2019–2023.
5. The full text of this publication is available for open access.
6. Articles about artificial intelligence and computer science.
7. The final publishing stage is at this stage.

Literature Selection

The technique used by Sánchez-Aguayo, Urquiza-Aguiar, and Estrada-Jiménez (2021) was followed through several stages of the literature selection process. There are four steps in the literature selection process: identification, filtration, formality, and deciding how many articles to study. Based on the findings of a literature search conducted on the Scopus database using the terms ("artificial intelligence" OR "cyber attack") AND ("data security")). The assessment was carried out at 22.37 WIB on January 20 2024. Of the 1,057 articles identified, (ii) 765 papers were obtained by adding the phrase "Computer Science"; an additional keyword, "Artificial Intelligence", is also included. After obtaining 497 papers, finally found the right keyword (iii): Found eighty documents in the field of computer science and artificial intelligence.

Table 1. Methodology Screening

Database Scopus	Screening	Publication
Meta Search	*Keyword : Cyber Attack, Artificial Intelligence, Data Security*	1.057
Inclusion Criteria	*Keyword : Computer Science*	765
	Keyword : Artificial Intelligence	497
	Keyword : Article, English	410
	Keyword : English	1.023
	Keyword : Final	1.030
Screening	*Keyword : Computer Science;Artificial Intelligence*	80

Source : Scopus

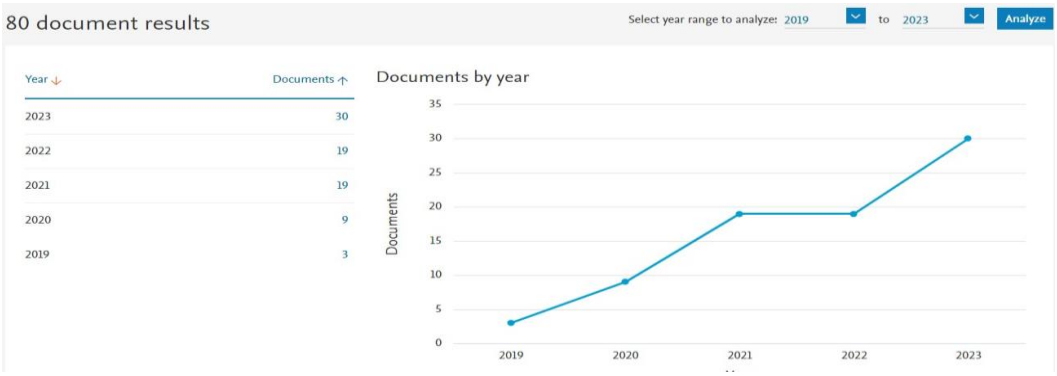
RESULTS AND DISCUSSION

General description

The purpose of this Systematic Literature Review (SLR) or Systematic Literature Observation is to determine the factors that participate in the risk of AI cyber attacks on data security. Microsoft Excel was used to analyze articles collected from various data sources for this research. Each article studied will be categorized according to the research findings, techniques used, and the researcher's recommendations for further research.

Research Findings

From the articles that will be researched, the following researchers will show research trends from 2019 to 2023:



Picture 1. Research Trends
Source : Scopus

Summary and Chronological Overview

The chronological summary of the papers published in this study is presented in Figure 1, showing the number of publications related to the threat of AI attacks on data security from 2019 to 2023. The graph demonstrates an increasing trend in this field of study. This surge has been evident in recent years, particularly since 2021, when a significant rise in the number of published papers began to occur. The majority of the reviews included in this analysis were released after 2020. In this subject, the highest number of papers (30) was published during the research period in 2023, followed by articles in 2022 and 2023. With only five papers reviewed in 2019, it's apparent that there was a lower level of publication in that year.

Synthesis Findings

The findings of the data synthesis are presented in this section to address the research questions derived from the selected publications. Therefore, the purpose of this section is to answer the research questions related to the Systematic Literature Review (SLR).

RQ1: What are the types of cyber attacks involving the use of Artificial Intelligence (AI) that can threaten data security in Indonesia?

The use of artificial intelligence (AI) in various cyber attacks poses risks to Indonesia's data security. One example is AI-enhanced phishing attacks(Akash, N.S, 2022) . Phishing efforts, involving the sending of fake emails or the creation of counterfeit websites, can be designed to leverage artificial intelligence (AI) to generate more convincing messages that are difficult to identify as phishing attempts. This can result in the loss of consumer personal data. Moreover, network intrusions utilizing AI also pose risks. Cybercriminals can hack systems, steal data, or damage digital infrastructure more quickly and effectively by using AI to exploit security vulnerabilities in networks (Wahyuni, 2023) .

AI-based malware cyber attacks also present a significant risk. AI-enhanced malware can exploit security system weaknesses in more complex and adaptable ways, making it more difficult to stop its spread, evade detection, and cause substantial damage to systems or data.

Deepfake is a technique capable of constructing human-like images based on artificial intelligence (AI). Deepfake attacks use Generative Adversarial Networks (GANs). This technology employs the Generative Adversarial Network approach to connect and integrate existing images and films into a source image or video. Ian Goodfellow discovered GANs in 2014 as an algorithmic means to generate actual data from existing data. GANs can also produce new text from existing text and actual audio from existing audio. (Kasita, 2022) . In 2018, a program called FakeApp was widely launched, allowing anyone to create deepfakes. The characteristics of FakeApp can be misused to spread hate speech and falsehoods throughout Indonesia. Similar to an incident in 2022, deepfake also claimed the life of the artist Nagita Slavina. A 61-second film featuring a naked scene with a face resembling Nagita's could be seen.

AI technology allows hackers to scale up and automate their operations, which may complicate the efforts of conventional security systems to prevent and identify hacks. As a result, all forms of these attacks pose a significant danger to data security in Indonesia. Therefore, for the business world and individuals in Indonesia, strong data security, cyber education, and improved detection and response capabilities to attacks utilizing AI are crucial (Samtani, 2020) .

RQ2: What is the level of awareness among the Indonesian public about the risk of AI cyber attacks on data security?

Although more attention is still needed, the Indonesian public is increasingly aware of the dangers of cyber attacks using artificial intelligence (AI) on data. Through initiatives to raise awareness, seminars, and workshops, there has been an increase in efforts in Indonesia in recent years to enhance awareness of the threats posed by cyber attacks. These initiatives have been actively undertaken by various public and commercial entities to raise awareness of cyber security threats, particularly the potential of AI attacks on data security (Stafford, 2022) .

However, there are still challenges in the effort to increase the overall knowledge of the public. It is possible that a large part of the Indonesian population is not aware of the complexity of the threats posed by AI-driven cyber attacks. The public's perception of these dangers can be motivated by several factors, including a lack of digital literacy, limited access to cyber security information, and ignorance of practical data protection measures (Cherviakov, 2020) .

The public must become more aware of the dangers of AI hacking on data security, a responsibility that falls on the government, academic institutions, the commercial sector, and the media. Reducing vulnerability to cyber attacks can be achieved through more extensive education campaigns, digital learning in schools, internet user training, and the promotion of safe cyber security practices. Therefore, protecting institutional and personal data from the risks of AI-based cyber attacks requires a sustained increase in awareness among the Indonesian population.

RQ3: What are the impacts of Artificial Intelligence (AI) cyber attacks on data security in Indonesia?

AI-based cyber attacks have the potential to cause serious impacts on Indonesia's data security. One of the primary impacts is the risk of sensitive data theft. Attackers can more easily breach security systems and steal important data, including financial and personal information, by leveraging AI technology. This can endanger the security and privacy of individual identities and result in substantial financial losses for citizens, businesses, and the government.

Moreover, vital digital infrastructure may be impacted by AI hacking. AI technology can be used by attackers to infiltrate or destroy critical information systems, including public services, financial institutions, and communication networks. Public services might be disrupted, everyday life for the population could be affected, and economic instability might result from these disruptions.

Furthermore, ransomware attacks can be developed using Artificial Intelligence (AI) to become more complex and sophisticated. The victims' data is encrypted by these attacks, and a bitcoin ransom is demanded to unlock the data (Wahyuni, 2023) . In 2017, a hospital in Jakarta became a target of an AI-based ransomware attack. The hospital's system encrypted patient data, and the hackers demanded a substantial payment to decrypt it.

Additionally, misinformation can be spread and reputations damaged through social media and other internet channels using Artificial Intelligence (AI) attacks known as Information Manipulation. The political, economic, and social stability of Indonesia could be disrupted by these attacks. Specifically, in 2019, there was a cyber attack that used AI to obtain personal information from individuals. As a result of this attack, the personal information of millions of people was published and maliciously exploited.

Another impact is the diminishing public trust in technology systems. AI hacking has the potential to create distrust in general technology if it consistently breaches security systems. Declining trust could hinder the acceptance of new technologies and impede progress in various economic and social sectors in Indonesia.

Lastly, sophisticated geopolitical impacts could be caused by AI cyber attacks. Cyber attacks leveraging artificial intelligence have the potential to become tools used by states or specific organizations to conduct espionage, sabotage, or interfere with the critical infrastructure of other countries (Maksim, 2020) . This can lead to international disputes and tensions, disrupt diplomatic relations, and pave the way for greater dangers to domestic and global security. Therefore, strong data protection and cybersecurity are crucial for the long-term political and economic stability of Indonesia.

RQ4: What efforts and measures can Indonesia take to protect data from Artificial Intelligence cyber attacks?

Indonesia may undertake several actions and efforts to safeguard data from AI-based cyber attacks. To create stronger cybersecurity regulations, government organizations, the commercial sector, and academic institutions need to collaborate more. Strict laws are needed to control data protection, manage security incidents, and set security requirements that must be met by public and private entities. It's also important to establish an impartial organization to track and assess compliance with policies.

Secondly, investing in advanced cybersecurity technology is crucial. Enhancing early detection systems and rapidly responding to cyber attacks can be achieved with the help of artificial intelligence. Developing AI systems capable of identifying anomalous risks, monitoring network activity in real-time, and reacting quickly to attacks is vital to protect data from more sophisticated assaults.

Thirdly, another crucial element is raising public awareness about the importance of cybersecurity. Preventive measures such as public education and training on safe cyber security practices, risks associated with cyber attacks, and methods to protect personal information are essential. This can be achieved through public awareness campaigns, teaching in schools, and broad community education initiatives.

Additionally, Big Data Encryption and Access Control techniques are being used by several large companies in Indonesia to protect privacy and stop illegal access to Big Data. Data access rights are regulated and restricted through access control, and data confidentiality is well ensured through encryption techniques. Only authorized entities can view data thanks to encryption. Meanwhile, the goal of access control is to limit who has access to the information (AL-Dosari, K., 2022) .

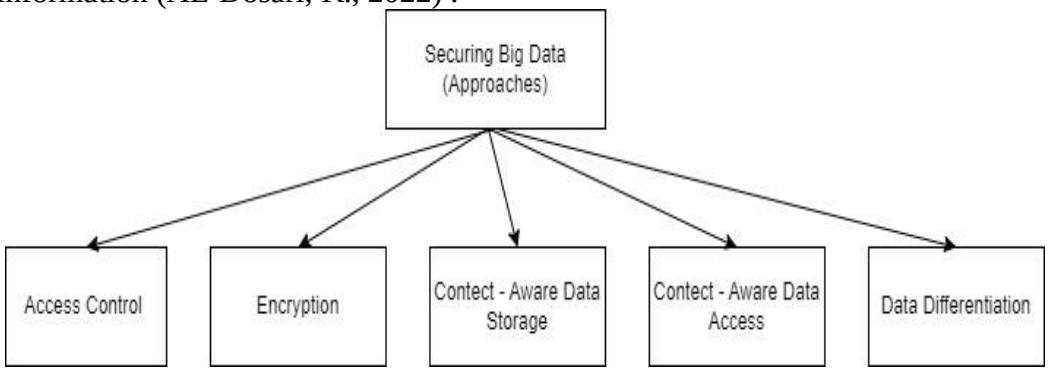


Figure 2. Big Data Security Techniques in English.

One example of an alternative use for protecting Big Data is a blockchain-based strategy that utilizes smart contracts and blockchain technology as a security mechanism.

Lastly, one of the most important aspects in preventing AI hacking is international collaboration. In the field of cybersecurity, cooperation with other countries, international organizations, and the global community can help Indonesia gain more knowledge and resources and strengthen its national cyber defense systems on a global scale. By implementing these measures, Indonesia can enhance data protection against increasingly complex and sophisticated cyber attacks.

CONCLUSION

This systematic study highlights the increasing trend of research on AI cyber attacks on data security in Indonesia from 2019 to 2023. Based on a review of papers from various data sources, it is evident that AI-based cyber attacks have significant impacts. Data security in Indonesia is now threatened by attacks including ransomware, phishing, deepfake attacks, AI-enhanced malware,

and information manipulation. Some of the impacts include data theft, damage to critical infrastructure, declining public trust in technology, and potential complex geopolitical effects.

Indonesia must adopt a comprehensive approach to safeguard data from AI attacks. It is crucial for educational institutions, businesses, and the government to collaborate in creating strict cybersecurity regulations. Investing in advanced security technology, enhancing public awareness through education, employing Big Data security strategies like encryption and access control, and encouraging international collaboration in cybersecurity are essential. Future research is suggested to extend over a longer research period, involve a larger subject matter, evaluate independent variables using alternative proxies, and consider additional data security measures.

Limitations

Several forms of data security and AI hacking have been detected in this SLR. While responding to RQs, protocols were developed with the aim of maximizing external and internal validity. However, there are a number of limitations and challenges to validity that need to be addressed, including the following:

Only conference publications and journals discussing cyber attacks related to data security qualified for this Systematic Literature Review (SLR). In the early stage of research, a search strategy was used to find and exclude a number of research publications considered irrelevant. This ensured that the selected research articles met the study's criteria. It is estimated that adding new sources, such as additional reference books, could enhance the evaluation further.

Despite considering a significant database when reviewing research publications, there's a possibility that additional digital libraries containing related research were overlooked. This limitation was addressed by comparing search phrases and keywords against a list of leading research articles. In the process of keyword searching, certain synonyms might have been missed. To ensure no important phrases were overlooked, the SLR protocol was updated to handle this issue.

The search was limited to items written in English. This introduced a language bias, as there might have been relevant publications in this field of study in other languages. Fortunately, every publication included in our analysis was published in English. Consequently, there was no linguistic bias.

BIBLIOGRAPHY

Akash, N.S. et al. (2022) ‘Botnet Detection in IoT Devices Using Random Forest Classifier with Independent Component Analysis’, Journal of Information and Communication Technology, 21(2), pp. 201–232. Available at: https://doi.org/10.32890/jict2022.21.2.3.

AL-Dosari, K., Fetais, N. and Kucukvar, M. (2022) ‘Artificial Intelligence and Cyber Defense System for Banking Industry: A Qualitative Study of AI Applications and Challenges’, Cybernetics and Systems, 0(0), pp. 1–29. Available at:https://doi.org/10.1080/01969722.2022.2112539.

Anastasya Zalsabilla Hermawan et al. (2023) 'Studi Literatur: Ancaman Serangan Siber Artificial Intelligence (Ai) Terhadap Keamanan Data Di Indonesia', Prosiding Seminar Nasional Teknologi dan Sistem Informasi, 3(1), pp. 581–591. Available at: https://doi.org/10.33005/sitasi.v3i1.363.

Cherviakov, L.M. et al. (2020) 'Digitalization of quality management of the strategic decision-making process', Proceedings of the 2020 IEEE International Conference 'Quality Management, Transport and Information Security, Information Technologies', IT and QM and IS 2020, pp. 193–196. Available at:https://doi.org/10.1109/ITQMIS51053.2020.9322987.

Kasita, I.D. (2022) 'Deepfake Pornografi: Tren Kekerasan Gender Berbasis Online (KGBO) Di Era Pandemi Covid-19', Jurnal Wanita dan Keluarga, 3(1), pp. 16–26. Available at: https://doi.org/10.22146/jwk.5202.

Ma'rufah, N., Rahmat, H.K. and Widana, I.D.K.K. (2020) 'Degradasi Moral Sebagai Dampak Kejahatan Siber Pada Generasi Millenial di Indonesia', Nusantara: Jurnal Ilmu Pengetahuan Sosial, 7(1), pp. 191–201.

Maksim, B. et al. (2020) 'Development of a software library for game artificial intelligence', Proceedings of the 2020 IEEE International Conference 'Quality Management, Transport and Information Security, Information Technologies', IT and QM and IS 2020, pp. 188–192. Available at: https://doi.org/10.1109/ITQMIS51053.2020.9322928.

Okey, O.D. et al. (2023) 'Investigating ChatGPT and cybersecurity: A perspective on topic modeling and sentiment analysis', Computers and Security, 135(September), p. 103476. Available at: https://doi.org/10.1016/j.cose.2023.103476.

Proceedings - 2020 2nd International Conference on Machine Learning, (2020). Big Data and Business Intelligence, MLBDBI 2020' (2020) Proceedings - 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence, MLBDBI 2020 [Preprint], (October).

Samtani, S., Kantarcioglu, M. and Chen, H. (2020) 'Trailblazing the Artificial Intelligence for Cybersecurity Discipline: A Multi-Disciplinary Research Roadmap', ACM Transactions on Management Information Systems, 11(4). Available at: https://doi.org/10.1145/3430360.

Stafford, T.F. (2022) 'Platform-Dependent Computer Security Complacency: The Unrecognized Insider Threat', IEEE Transactions on Engineering Management, 69(6), pp. 3814–3825. Available at: https://doi.org/10.1109/TEM.2021.3058344.

Wahyuni, N.K.A.T. et al. (2023) 'Analisis Kerentanan Kejahatan Online Phising Menggunakan Tools Zphisher, Shellphish Dan Whphisher', Jurnal Teknik Mesin, Elektro dan Ilmu Komputer, 3(1), pp. 23–31. Available at: https://doi.org/10.55606/teknik.v3i1.915.