Pak. j. life soc. Sci. (2024), 22(2): 21308-21323 E-ISSN: 2221-7630;P-ISSN: 1727-4915

Pakistan Journal of Life and Social Sciences
www.pjlss.edu.pk

https://doi.org/10.57239/PJLSS-2024-22.2.001505

RESEARCH ARTICLE

Civil Law Implications of Artificial Intelligence Training Using Copyrighted Data: A Case Study Of Thomson Reuters Vs. Ross Intelligence

Wulan Azalia Zanzabila[1], Maskun[2], Kamsilaniah [3], Aminuddin [4]

[1,2,4]Faculty of Law, Hasanuddin University

[3]Faculty of Law, Bosowa University

ARTICLE INFO

Received: Oct 14, 2024

Accepted: Dec 23, 2024

Keywords

Artificial intelligence
Copyright
Copyrighted data
Fair use
AI training
Thomson reuters vs. ross intelligence

***Corresponding Author:**

azalea.zanzabila11@gmail.com

ABSTRACT

The development of Artificial Intelligence (AI) technology has had a significant impact on various fields, including law. One important aspect of AI development is the use of large amounts of data, often copyrighted data. This research analyzes the legal implications of using copyrighted data in AI training, focusing on the Thomson Reuters vs. ROSS Intelligence case study. This research uses normative legal methods with statutory, case study, comparative law, and conceptual approaches. The results show that ROSS Intelligence's use of copyrighted data, such as headnotes and key number systems, to train their AI violates the principle of copyright protection. The arguments raised by ROSS regarding transformationality and the fair use doctrine cannot be justified, as the use of the data substantially affects the economic value of the original work and does not create significant new value. This study also confirms the need for clearer regulations regarding using copyrighted data in developing AI technologies, including transparent licensing mechanisms and limitations on data use. This study makes an important contribution to understanding the legal challenges in the era of artificial intelligence by highlighting the need to balance copyright protection and technological innovation. The research results will guide policymakers, technology developers, and copyright holders to create a fair and adaptive legal framework.

INTRODUCTION

Since 2011, the world is currently in the category of Industrial Revolution 4.0. The term "industrial revolution" was first used in 1799 by a French envoy in Berlin who said his country had entered the industrial revolution. Since then, the term has been used to refer to rapid progress in one industry at one particular place and time (Gordon, 2023, p. 118). The development of the concept of this revolution has entered its fourth, the Industrial Revolution 4.0, which began in the early 21st century and was introduced in 2011(Ivanova, 2022) -with a focus on the merging of digital, physical, and biological technologies, thus distinguishing it from previous revolutions in terms of its speed, scale, and impact on production, management, and governance systems. One of the key characteristics of this revolution is Artificial Intelligence (AI), which plays a crucial role in shaping the technologies and processes that drive this revolution, alongside other technologies such as robotics, the Internet of Things (IoT), driverless cars, and nanotechnology (Levente & Péter, 2023, p. 4).

AI combines various technologies that enable software and machines to sense, understand, act, and learn independently or with human collaboration (Chaurasia et al., 2024, p. 362). The working stages of AI go through several steps, namely data collection, data processing (cleaning, manipulation, and preparation), model training, data testing, and data optimization (Pragyna Karmakar et al., 2024, p. 79). AI uses inductive and deductive logical processes in data processing, where patterns from specific data are used to make general conclusions, while deduction uses existing premises to make specific conclusions (Collecchia & De Gobbi, 2024). In summary, AI collects, processes, and learns from data to perform tasks and make decisions.

Data is like "fuel" for AI, playing a crucial role in its development and operation. AI cannot be trained to make predictions or make relevant decisions without data. Pankaj Kumar Verma and Lakhbir Kaur's research shows that AI can only work effectively with high-quality and sufficient data. The more diverse and accurate the data used, the better the performance of AI systems in recognizing patterns and making more accurate predictions. Data is used not only to train models but also to test their performance. This allows AI developers to evaluate the model's accuracy before it is used in a real-world context (Verma & Kaur, 2024, pp. 111–116).

However, using data as the object of AI is crucial from a legal perspective. Data has become a highly valuable asset, but not all data can be used freely. Some of it is protected by copyright, giving rise to significant legal implications in its use. Copyright is regulated by law and granted to creators of literary works, plays, music, art, cinematographic film producers, and sound recordings (Vibha S, 2022). Copyright infringement can result in severe legal sanctions, including fines and imprisonment. Some of the international copyright infringement cases that have occurred are Robin Thicke & Pharrell Williams vs. Marvin Gaye, Queen & David Bowie vs. Vanilla Ice, and Art Rogers vs. Jeff Koons (Jindal, 2024; Manar, 2024).

There is an interesting case on copyright infringement involving AI, which is the case of Thomson Reuters vs. ROSS Intelligence. This case is a significant legal dispute regarding the use of intellectual property in AI development. Thomson Reuters, the owner of Westlaw, a widely used legal research platform, alleged that ROSS Intelligence, a startup that developed an AI-based legal research platform that allows users to ask legal questions and get answers from court opinions, used copyrighted headnotes from Westlaw (a product of Thomson Reuters) to train ROSS Intelligence's AI system. Headnotes are summaries of court decisions created and organized by Westlaw as part of its legal database (Delman, 2023). The main issue in this case is whether using copyrighted data (such as headnotes) to train an AI can be considered copyright infringement.

Some previous studies that discuss copyright infringement in technology are as follows. *First*, Ujang Badru Jaman et al. examine the importance of legal protection of digital copyrighted works in the modern era influenced by technological developments and the internet. This study aims to explain the form of legal protection of digital copyrighted works and analyze the role of the government in overcoming illegal duplication of copyrighted works in the digital era. Digital works include electronic books (e-books), music, videos, software, images, and other digital content. These works have advantages regarding ease of access, distribution, and storage. Legal protection of digital works requires a special approach that differs from physical works, especially regarding distribution and copy control. A case in point is the piracy of Andrea Hirata's "*Laskar Pelangi*" e-book by an illegal website, showing how printed works can be duplicated and disseminated without the author's permission. The government is important in providing legal protection for digital copyrighted works through strengthening regulation and supervision (Ujang Badru Jaman et al., 2021).

Second, Fenny Wulandari discusses the problem of copyright infringement in the digital era, where technological developments massively expand access and distribution of digital content. The main objective of this research is to identify the causes of copyright infringement, explore the legal system's role, and offer concrete solutions for strengthening monitoring and enforcement. Copyright infringement in the digital era is increasingly complex, mainly due to digital content's ease of distribution and reproduction. There are several types, such as music and movies, illegal software, plagiarism of literary and graphic works, and

unauthorized distribution of photography and cinematography. In addition, copyright infringement negatively impacts creators and holders, including financial losses, loss of control, and weakening of creative industries(Wulandari, 2024).

Third, Lidwina Dope Nyadjroh Gabsa explores the phenomenon of copyright infringement in the rapidly growing era of information and communication technology (ICT). This research aims to provide insight to copyright owners and related stakeholders regarding the concept of copyright protection, copyright infringement in the context of ICT, the basics of copyright infringement, and copyright infringement in literary works, art, and neighboring rights. According to him, copyright protects creators of works that are original and embodied in tangible form, including books, music, visual arts, computer programs, and audio-visual broadcasts. Copyright infringement in the ICT context includes music and film piracy, unauthorized duplication, illegal downloading, unauthorized distribution, and unlicensed use of images, videos, or music. In principle, copyright infringement occurs when a third party uses a work without the copyright holder's permission, making it subject to civil or criminal prosecution, depending on the applicable legal jurisdiction (Gabsa, 2024).

Fourth, Ziyan Yan and Zhao Hao examine the issue of copyright infringement arising from using artificial intelligence (AI) technology in creating paintings. Since 2022, AI has developed rapidly in the art field, enabling innovative and stylish artwork creation. However, creating AI artwork often involves using materials from other people's artworks, potentially infringing copyright. This research identifies three main types of copyright infringement by AI paintings: reproduction right infringement, adaptation right infringement, and attribution right infringement. This research shows that there are difficulties in identifying and proving copyright infringement by AI due to the difficulty in identifying the act of infringement, the difficulty in determining the liable subject, and the difficulty in the lawsuit process (Yan & Hao, 2024).

Fifth, Lijie Ai addresses the legal issues of copyright protection on the internet that arise with the development of network technology. The main focus of this research is to identify challenges, elements in need of legal protection, and measures to improve copyright legal protection in the digital environment. Some of the challenges of copyright protection on the Internet are easy and rapid dissemination, high cost of monitoring and enforcement, and non-exclusive nature. The elements of copyright legal protection on the Internet include protection of network transmission rights, technical measures, and rights management information. Improving the legal protection of copyright can be done through strengthening laws and regulations, strengthening judicial protection, and optimizing administrative protection (Ai, 2015).

Table 1. Comparison of Previous Research

No.	Researcher	Research Focus	Research Objectives
1	Ujang Badru Jaman et al. (2021)	Legal protection of digital copyrighted works in the modern era influenced by technology and the internet.	Explains the form of legal protection of digital copyrighted works as well as the role of the government in monitoring and strengthening regulations.
2	Fenny Wulandari (2024)	Problems of copyright infringement in the digital age, focusing on access and distribution of digital content.	Identify the causes of violations, explore the legal system's role, and offer supervision and enforcement solutions.
3	Lidwina Dope Nyadjroh Gabsa (2024)	Copyright infringement in the context of Information and Communication Technology (ICT) and the protection of copyrighted works.	Provide insight to copyright owners and stakeholders regarding the concept of copyright protection in the context of ICT.

4	Ziyan Yan and Zhao Hao (2024)	The issue of copyright infringement is due to AI's use in creating paintings.	Identify types of copyright infringement in creating AI paintings and examine legal challenges in the lawsuit process.
5	Lijie Ai (2015)	The legal issue of copyright protection on the internet includes challenges and elements that require legal protection.	Identify challenges, elements of legal protection, and measures to improve copyright legal protection on the internet.

Based on the comparison of previous research, it seems that no research specifically discusses the legal implications of using copyrighted data in Artificial Intelligence (AI) training. Most of the previous research focuses on copyright protection in general in the digital context, such as the protection of e-books, music, videos, software, and other digital content. While some studies address copyright infringement in the digital age, the focus is more on infringement of illegal distribution, piracy, and unauthorized copying on digital platforms or internet networks. In addition, previous studies have mostly explored the role of the government in strengthening regulations and supervision in preventing copyright infringement. No research has examined the use of copyrighted data as AI training material and its impact on the regulation of copyright law policy in the AI era. The approach used in previous research is also more normative and descriptive, with no concrete case study approach related to this legal issue.

Based on this description, this research aims to identify the legal implications of using copyrighted data in AI training. This research has the advantage of specifically analyzing the Thomson Reuters vs. ROSS Intelligence case, which deals directly with using copyrighted data (headnotes from Westlaw) in AI training. The case raises an urgent and contextual legal question: Can using copyrighted data to train AI be classified as copyright infringement? No previous research has discussed this case in depth, let alone in relation to the legal setting in Indonesia and its comparison with other jurisdictions, such as the United States and the European Union. Therefore, this research fills a gap in the legal literature (research gap) related to legal arrangements and policies governing the use of copyrighted data in the AI development process, especially by emphasizing the perspective of the legal system in Indonesia and other countries' legal systems. This research is relevant in terms of academic contribution and provides more specific policy recommendations to policymakers, AI developers, and copyright holders.

MATERIALS AND METHODS

This research is legal. The term “legal research” consists of two-word elements, namely “research” and “law." The word “research” comes from two roots, namely “re” and “search ." Re” means to return, while “search” means to find something carefully or “*examine, look carefully at, through, or into ... in order to find something.*” Meanwhile, the term “law” has various meanings depending on the perspective of each school of legal philosophy, in a neutral and simple sense, the law refers to norms that are formed, enforced, and recognized by public authorities to regulate the life of the state and society, and supported by the application of sanctions. Therefore, legal research, otherwise known as "legal research," is the process of carefully and thoroughly recovering legal materials or data in order to solve legal problems (Diantha, 2016, pp. 1–2).

The method used is normative legal research with a statutory approach, case approach, comparative approach, and conceptual approach. Normative law research is a type of legal research that aims to examine the law as norms or rules that apply in society and guide behavior for everyone (M. C. Ramadhan, 2021, p. 51). The research approach refers to strategies and methods that expand decisions from general assumptions, thus enabling thorough and optimal data collection and analysis (Armia, 2022, p. 2). The statutory approach in legal research is a method used to analyze and examine all laws and regulations related to the legal issues being studied, which are inseparable from legal research, both in the context of legal dogmatics (normative studies) and legal practice, to obtain a comprehensive picture of the applicable legal arrangements, including understanding the hierarchy of laws and regulations and the legal principles underlying them. The case approach in legal research is a method used to analyze and examine concrete

cases related to certain legal issues, especially through court decisions that have permanent legal force (inkracht van gewijsde), where researchers examine the ratio decidendi (legal considerations of judges) which are the basis for making decisions. The comparative approach in legal research is used to compare laws and regulations, court decisions, or legal practices from one country with other countries related to the same legal issues so that similarities and differences can be identified from the regulation and application of law in these countries. The conceptual approach in legal research is a method used by departing from views, doctrines, or legal theories to develop legal concepts, legal principles, and legal notions relevant to certain legal issues, especially when the legal issues under study have not been explicitly regulated in laws and regulations or court decisions, so that researchers need to develop a new concept that can be used as a foothold in their analysis (Sovia et al., 2022, pp. 25–31).

Normative legal research does not involve field research because the objects studied are legal materials. This research is library-based, with a method that prioritizes reading and analyzing primary and secondary legal materials. Therefore, normative legal research is often referred to as legal science research. The data used is entirely secondary, including primary, secondary and tertiary legal materials. Primary legal materials refer to legal materials that are legally binding. Secondary legal materials provide further explanation or understanding of primary legal materials, such as draft laws, research results, scientific works in law, opinions of legal experts, and the like. Tertiary legal materials are supporting materials for primary and secondary legal materials. In normative legal research, data presentation is inseparable from analysis. The data collected (primary, secondary, and tertiary legal materials) are analyzed to find their relevance and relation to the studied legal issues (Armia, 2022, pp. 12–14).

RESULTS AND DISCUSSION

The origin of artificial intelligence (AI) can be said to date back to the Greeks. AI evolved from the accumulation of human knowledge over thousands of years. Aristotle (335 BC) made an important early contribution by introducing deductive logic, a method of reasoning from general principles to specific conclusions. In the 17th century, Leibniz attempted to develop a “thinking machine” with symbolic logic. Then, in the late 19th to early 20th century, William James introduced the structure and function of the human brain, which later became the basis for the development of artificial neural networks. In the mid-20th century, Claude Shannon introduced binary numbers (1 and 0) as information representation, which became the foundation for developing digital circuits and binary information theory. Later, AI was officially born at the Dartmouth Conference in 1956. The conference was attended by prominent scientists, such as John McCarthy, Marvin Minsky, and Claude Shannon, who introduced the term "artificial intelligence" and made it an independent field of research. This conference became the starting point for AI as a scientific discipline. Shortly after, John McCarthy and Marvin Minsky founded the MIT AI LAB, the world's first laboratory dedicated to artificial intelligence research. The lab played an important role in developing AI theory and applications (Jiang, 2024, pp. 1–2).

AI is defined differently by various disciplines. For computer scientists, AI refers to developing programs that exhibit intelligent behavior. These programs can perform intelligent planning (e.g., traffic light timing), translate natural language (e.g., translate a website from Chinese to English), act as an expert (e.g., choose the best wine for dinner), or perform many other tasks. In addition, engineers' definition of AI includes the development of machines capable of performing actions often performed by humans. These machines can be as simple as computer vision systems embedded in ATMs, more complex as robotic explorers on Mars, or highly complex as automated factories that build sports machines with little human intervention.

Meanwhile, for cognitive scientists, AI means building models of human intelligence to better understand human behavior. Originally, most models of human intelligence were symbolic and closely related to cognitive psychology and philosophy. These models aimed to reflect the workings of the human brain processed through symbols (Franceschetti, 2018, pp. 12–16). In summary, AI is an artificial entity designed to mimic, replicate, or surpass human intelligence in decision making, problem solving, and adaptive behavior.

The rapid development of AI has allowed AI to perform human-like actions and deeds, which raises legal issues if AI performs actions that harm other parties. This is because one of the goals of AI is to create machines that can perform tasks that are usually done by humans, from simple tasks such as image recognition to more complex tasks such as automated factories (Franceschetti, 2018, p. 12). As a result, AI also has the ability and acts like a human which does not escape causing an action that leads to a legal act and a criminal offense that causes harm to another party (Astiti, 2023, p. 965).

In this case, it is closely related to data. The concept of "data" was initially limited to its literal meaning, which can be described as "numbers obtained through unprocessed investigation or experimentation" or "signals in the field of communication" in the Agricultural era and the Industrial era (Tong, 2021, p. 37). However, this concept has certainly evolved. In the fourth edition of the "*Encyclopedia of Information Science and Technology*" by Mehdi Khosrow-Pour, "data" in the context of information and technology is a term that refers to something that initially has no meaning in itself, but can be processed to produce useful information. Data can be anything that can be recorded, stored, and processed by a computer, including symbols, numbers, objects, and events. As symbolic representations, data describes entities, their properties and states that have been codified to enable them to be communicated. Data is often understood as raw, unevaluated facts, such as numbers or symbols separated from a specific context. In addition, data can refer to a record of transactions or events that occurred in the recent past (Khosrow-Pour, 2018, pp. 1, 220, 338, 943).

According to Syrine Ferjaoui in her work entitled "*Data: The New Form of Wealth and Power*", data can be understood as the raw form of information that is the main raw material in forming knowledge and wisdom. Knowledge is the cognitive rearrangement of information to gain a deeper understanding, which allows humans to identify patterns and understand relationships between information. At the same time, wisdom is the ability to use knowledge ethically, contextualistically, and effectively in decision-making. As a basic element of information, data plays a fundamental role in technological advancement and digital transformation in the 21st century. This concept is reinforced by the statement that "*the natural state of information is data, a raw form that catalyzed emerging markets,*" emphasizing that data is the basic form of information that triggers the emergence of new markets. In modern technology, data is seen as the raw material widely used in operating devices such as smart phones, smart watches, drones, and Internet of Things (IoT) devices. A quote from the book states, "*all invented devices share a common raw material: data*". This emphasizes that modern technology's advancement depends on using data as the basic material for its operation. The view of data as a strategic resource that can be utilized is also affirmed by Clive Humby, a British mathematician, who said that "*Data is the new oil. It's valuable, but if unrefined, it cannot be used*". This statement illustrates that data has a high value, but in its raw form, data is useless until it is processed into something more valuable. In this case, data is processed and transformed into various usable forms, such as information or insights, similar to how crude oil must be processed into fuel or other products to be used. This approach suggests that the value of data lies not only in its existence, but in the potential value that can be extracted through processing and analysis (Ferjaoui, 2020) . The following diagram can be drawn to understand the data referred to by Ferjaoui.

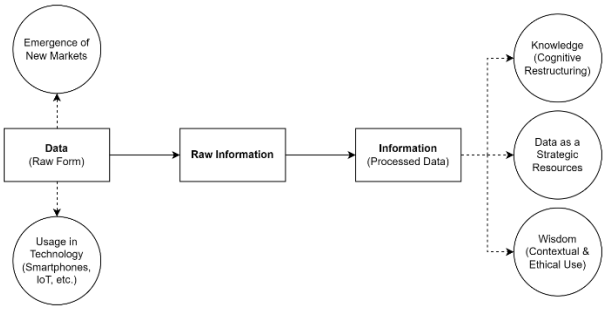

Image 1. What is Data?
(Source: processed by the author, 2024)

Data is an essential element in AI learning that acts as the main foundation in the development process of the AI system itself. In artificial neural networks, deep learning requires feeding large amounts of data into a multilevel network to train the machine for classification and pattern recognition. The algorithms used in this process instruct the neurons how to respond to improve the accuracy of the results. The existence of big data further strengthens machine learning's ability to recognize patterns and make decisions. The abundance of data allows AI to hone its ability to recognize hidden patterns from complex data. On the technology side, hardware such as graphics processing chips (GPUs) enable the simultaneous execution of thousands of deep learning processes, accelerating big data processing. In addition, facts and data serve as the basis for AI-based decision-making. Fact-based knowledge representation allows AI systems to organize data as information processed into knowledge. In artificial neural networks, data is represented as weights from a weight matrix. Updating those weights adds new knowledge to the system, thus enabling advanced learning and better decision-making. Reinforcement learning approaches also make extensive use of data. One example is the application of reinforcement algorithms by Google DeepMind, which is used to solve energy efficiency problems and decision-making in video games. Through reinforcement learning, AI can adapt from feedback obtained during learning. The utilization of big data is also found in image recognition and facial recognition. An example is facial recognition applications on social media platforms such as Facebook, which use data from photos uploaded by users to identify and tag the person in the image. This process involves algorithms that are trained using thousands to millions of image data to ensure the accuracy and speed of facial recognition (Franceschetti, 2018, pp. 14–99). In summary, data is the primary fuel in the development and operation of artificial intelligence systems, allowing algorithms to learn, adapt, and make decisions based on patterns extracted from the data. The more and more diverse the data available, the greater the opportunity for AI to achieve higher accuracy and reliability, both in practical applications and technological innovations in various fields. A depiction of the process of data in AI can be described as follows.

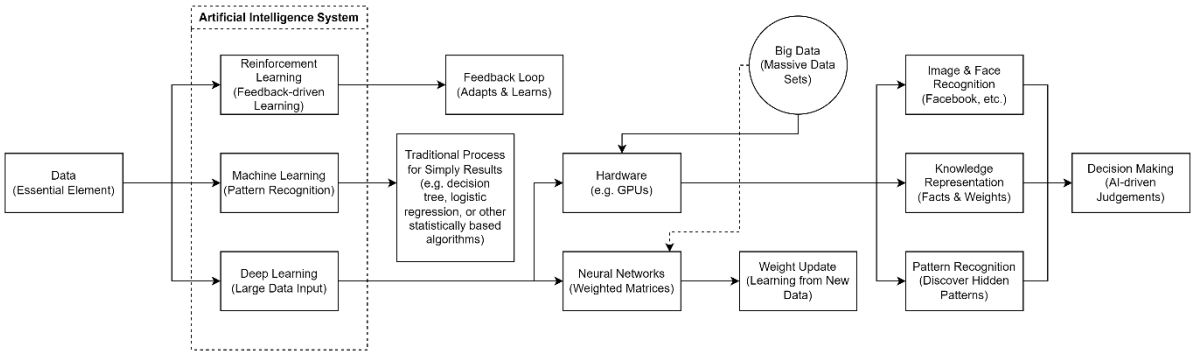

Image 2. Data Processing in Artificial Intelligence
(Source: processed by the author, 2024)

So, what is the concept of "data" from a legal perspective? The concept of "data" in law can relate to "personal data" and "information systems" in technology and information law/cyber law, and "intellectual property rights" in civil law. However, in the context of this research, the concepts of data in "information systems" and "intellectual property rights" are more relevant than the concept of "personal data". This is because "personal data" in the Indonesian legal system through Law of the Republic of Indonesia Number 27 of 2022 on Personal Data Protection, defines it as "data about an identified or identifiable natural person individually or in combination with other information either directly or indirectly through electronic or non-electronic systems." In addition, in the European Union legal system through Directive 95/46/EC, Article 2(a) refers to it as "any information relating to an identified or identifiable natural person", and in the UK legal system through the Data Protection Act 1998 (DPA 98), Article 1(1) refers to it as "data relating to a living individual who can be identified from that data or from the combination of that data with other information in the hands of the data controller (Rowland et al., 2017, pp. 344–345).

In a book entitled "*Pengantar Hukum Siber Indonesia*" by Nudirman Munir, the State of Indonesia in its regulation, namely Law of the Republic of Indonesia Number 11 of 2008 concerning Electronic Information and Transactions as amended several times last by Law of the Republic of Indonesia Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 concerning Electronic Information and Transactions, defines electronic information as one or a set of electronic data that includes, but is not limited to, writings, sounds, images, maps, designs, photographs, Electronic Data Interchange (EDI), electronic mail (email), telegram, telex, telecopy, or the like, as well as letters, signs, numbers, access codes, symbols, or perforations that have been processed and can be interpreted or understood by people who have the (Munir, 2017, p. 44). The concept of electronic information can describe "data" in the context of technology and information in the cyber world that we often encounter daily in the form of digital document files, instant messages (chat), social media posts, multimedia content such as video and audio, to elements in web-based applications and software, where this data becomes electronic information. This data is not only limited to human-readable formats, but also includes machine-readable formats, such as metadata, user activity logs, and data generated by artificial intelligence (AI) algorithms. While in the legal research by Xiuyu Tong, in the legal system in China, data is considered an economic resource and eventually considered a valuable asset so that data becomes a legal object (which was introduced in the "Decision of the Standing Committee of the People's National Congress on Strengthening Information Protection on Networks [2012]") which has an independent status in the civil law system in China, specifically regulated through the Cybersecurity Law of the People's Republic of China and the Personal Information Protection Law (Tong, 2021).

According to OK. Saidin in his book entitled "*Aspek Hukum Hak Kekayaan Intelektual (Intellectual Property Rights)*", intellectual property rights are immaterial property rights or rights to intangible objects born of human intellectual abilities, which come from the work of intellectual intelligence and human emotional intelligence that produce works in the fields of science, art, and literature. In addition, intellectual property rights are part of objects, namely intangible objects (immaterial objects), which are part of civil law, especially in Indonesia through the Civil Code as a Dutch colonial legacy, namely Article 499 which reads that objects are every item and every right that can be controlled by property rights. The concept of intellectual property rights emphasizes the protection of the copyrightability of a work, not against its manifestation (form), so that the copyrightability can take shape in the fields of art, industry and science or a blend of all three (Saidin, 2015, pp. 30–32). As for the international level, the World Intellectual Property Organization (WIPO) defines intellectual property rights include the creation of the human mind, including inventions, literary and artistic works, symbols, names, images, and designs used in commerce, and the World Trade Organization (WTO) emphasizes that intellectual property rights are exclusive rights granted to the creator of the results of his creation during a certain period, as well as in the international legal instrument in the form of Trade Related Aspects of Intellectual Property Rights (TRIPS) indicates that the term "intellectual property" refers to all categories of intellectual property covered in certain parts of the agreement, namely from section 1 to 7 (Kur & Levin, 2011, pp. 12–13 & 458), where the details of each section can be seen as follows:

Table 2. Comparison of Each Section in TRIPS
(Source: TRIPS, 1994, table processed by the author, 2024)

Section	Scope of Protection	Key Rights Conferred	Term of Protection	Condition/Limitation
Section 1: Copyright and Related Rights	Protection of literary works, computer programs, and compilations of data.	Right to authorize/prohibit reproduction, rental, distribution, and public communication.	50 years from publication or creation (for works not tied to a	Protection extends to expression, not ideas, procedures, or methods.

			person's lifetime).	
Section 2: Trademarks	Protection of signs (words, names, symbols) used to distinguish goods/services.	Exclusive right to prevent unauthorized use of the trademark in commerce.	Initial 7 years, renewable indefinitely.	Must be capable of distinguishing goods/services; registration is sometimes required.
Section 3: Geographical Indications	Protection of designations of origin for goods with qualities linked to geography.	Right to prevent misuse of geographical indications on products from other regions.	No fixed term; protection is linked to the existence and use of the geographical indication.	Only applies to geographical indications that are not generic or customary terms.
Section 4: Industrial Designs	Protection of new, original, and independent industrial designs.	Right to prevent unauthorized copying, production, and distribution of industrial designs.	At least 10 years.	Designs must be new/original and not dictated by technical function.
Section 5: Patents	Protection of novel inventions, involve inventive steps, and are industrially applicable.	Right to prevent unauthorized patented invention making, using, selling, and importing.	20 years from filing date.	Exclusions for morality, ordre public, and certain biological processes.
Section 6: Layout-Designs (Topographies) of Integrated Circuits	Protection of layout-designs of integrated circuits.	Right to prevent unauthorized reproduction, sale, and import of layout-designs.	10 years from application or first commercial exploitation.	Only unlawful if a person knows or has reason to know of unlawful reproduction.
Section 7: Protection of Undisclosed Information	Protection of undisclosed business information and test data from unfair commercial use.	Right to prevent disclosure, use, or acquisition of undisclosed information without consent.	No fixed term; protection exists as long as the information remains undisclosed.	Protection requires the information to be secret, valuable, and efforts taken to keep it secret.

In the context of technology and information, in a book entitled "*Intellectual Property Rights In Cyberspace*" by Akash Kamal Mishra, it can be concluded that data available on the internet, especially those used as access tools and information resources, can be considered as part of intellectual property rights. This is reflected in the statement that, "*Some of the positive impacts of the internet on the intellectual property community are: ... (e) there has been a geometric increase in the amount of accessible data and collections relative to intellectual property ...*" (Mishra, 2019, p. 18). Even in the current development of trade, there are some intangible products under the protection of intellectual property rights, especially in the copyright regime, as in the statement that:

> "*Today the largest segment of business-to-consumer e-commerce involves intangible products that can be delivered directly over the network to the consumer's computer. While these intangible products, by their very nature, are difficult to measure, an increasing amount of the content that is being offered is subject to intellectual property rights.*" (Mishra, 2019, p. 24)

Some examples of intangible products protected by intellectual property rights are Spotify, Netflix, Amazon Kindle, and so on. In addition, data can also be subject to the trade secret regime if the data is transformed into a computer or in its network, as in the book entitled "*Cyber Law & IPR in the Indonesian Legal System*" by Ahmad M. Ramli, which reads: "*Pengertian rahasia dagang mencakup pula informasi rahasia yang berada dalam komputer maupun dalam jaringannya, yang tidak dapat diakses oleh sembarang orang.*" (Ramli, 2010, p. 70). This is based on Ramli's reference to the definition of trade secrets in the provisions of Article 1 point 1 of Law of the Republic of Indonesia Number 30 of 2000 concerning Trade Secrets, namely trade secrets are information that is not known by the public in the field of technology and/or business, has economic value because it is useful in business activities, and is kept confidential by the owner of trade secrets. Therefore, data in the context of technology and information receives intellectual property rights protection, which can be in the form of copyright or trade secrets.

ROSS Intelligence is a technology company that focuses on developing AI-based products to help lawyers improve their cognitive abilities in the legal process, one is by developing AI-based legal research software designed to enable lawyers and legal professionals to conduct natural language-based legal searches. Since its establishment in 2014, ROSS Intelligence has experienced rapid growth. The company was founded by two computer scientists from the University of Toronto, one of the world's leading AI research centers, and a lawyer. Their main goal is to create technology to make legal services more affordable and accessible, especially to solve complex legal problems using cutting-edge technology (ROSS Intelligence, 2020).

On the other hand, Thomson Reuters has a history that stems from two companies that later merged. It all started with Reuters, founded in London in 1851 by Paul Julius Reuter as a stock price news provider. Reuters grew rapidly and became one of the largest news agencies in the world, even being the first to report the assassination of Abraham Lincoln in Europe in 1865 (Britannica, 2024). The forerunner of the Thomson Corporation began in 1934 when Roy Thomson founded a publishing company in Ontario, Canada (Potter, 2013). Thomson later expanded its business by acquiring various media outlets in the UK, including The Scotsman in 1953 and the Sunday Times in 1959. In 1996, Thomson Corporation strategically moved by acquiring West Publishing, a legal research company that owned Westlaw (Britannica, 2024). The turning point came on April 17, 2008 when The Thomson Corporation formally acquired Reuters Group PLC to form Thomson Reuters (DeMaria et al., 2008).

Westlaw is an online legal research service and exclusive database for lawyers and legal professionals in more than 60 countries. Westlaw has a long history began with a computer-based legal research project called QUIC/LAW at Queen's University in 1968-1973. The name QUIC/LAW stands for "Queen's University Investigation of Computers and Law" led by Hugh Lawford and Richard von Briesen. The project originally used IBM's internal code, INFORM/360, but substantial modifications were required as it was incomplete. 1973 the project was commercialized into QL Systems with a product called QL/SEARCH. Three years later, in 1975, QL Systems licensed the QL/SEARCH software to West Publishing which became the basis for forming Westlaw (Bourne & Hahn, 2003; Martin, 2009). A significant development occurred when Thomson

Corporation acquired West Publishing in 1996. Since then, Westlaw has continued to grow and expand its reach to various countries. In the early days, Westlaw operated as a dial-up service with a dedicated terminal known as WALT (West Automatic Law Terminal)(Djulvezan, 2000; Mattson, 2006) . Around 1989, Westlaw began offering a program for personal computers that emulated the terminal. When internet access became available, westlaw.com emerged as an alternative to accessing this service. Their client program known as Westmate continued to grow until it was finally discontinued on June 30, 2007. On February 8, 2010, Westlaw introduced WestlawNext with significant updates, and the classic Westlaw.com platform was finally retired in August 2015. WestlawNext was renamed "Thomson Reuters Westlaw" in February 2016 (Thomson Reuters, 2015).

In 2020, Thomson Reuters, through its subsidiary West Publishing, filed a lawsuit against ROSS Intelligence in Delaware federal court. The allegation was that ROSS had illegally used copyrighted content from Westlaw to train its AI models. Thomson Reuters claimed that ROSS colluded with LegalEase Solutions, a legal research services company, to illegally access Westlaw content. According to the lawsuit, LegalEase, which had legitimate access to the Westlaw database, allegedly bulk copied Headnotes and Key Numbers from hundreds of thousands of Westlaw documents and passed them to ROSS (Hass, 2024).

Headnotes are short summaries of the main legal principles in a court decision. In the Westlaw context, Headnotes are compiled and written by legal editors from Westlaw, rather than being part of the judgment itself. They are placed at the top of the judgment to help lawyers quickly identify the key legal points of the case. These headnotes are considered original content protected by copyright as they are manually compiled by Westlaw editors (Chanana & Ing, 2023). Meanwhile, Key Numbers is a classification system Westlaw uses to group legal issues into specific categories. This system allows for easier searching and organization of legal documents by associating specific legal topics with specific key numbers. It is a “master classification system” that organizes court decisions based on relevant legal topics and issues. Lawyers can use these key numbers to find similar cases that address the same legal issues (Keefe, 2018).

The allegations claim that ROSS used the data provided by LegalEase to train their AI-based search system. ROSS then used this data to develop a natural language-based search system that allowed its users to conduct legal searches without requiring knowledge of Westlaw's Key Numbers structure. According to Thomson Reuters, this infringed on their copyright over Headnotes and the Key Number System, which they claimed was a creative and original work protected by copyright law. In response, ROSS denied the allegations and claimed that although they used material from Westlaw, their actions were protected by the fair use doctrine in US copyright law. ROSS argued that their use of the material was transformational, as they used it to train a machine learning-based AI model that produced something new and different. ROSS also emphasizes that their AI system allows lawyers to conduct legal research more effectively and efficientlymore effectively and efficiently than conventional methods (Hass, 2024).

However, in a court document entitled “*Memorandum Opinion by Judge Bibas in the case of Thomson Reuters Enterprise Center GmbH and West Publishing Corp. v. ROSS Intelligence Inc.*” in the Delaware District Court (Case Number 1:20-cv-613-SB), dated September 27, 2024, there is no explicit mention of LegalEase Solutions' involvement in the case. Instead, the document focuses on copyright infringement claims and related allegations of competition law violations by ROSS against Thomson Reuters. In these court documents, the case centers on a dispute between Thomson Reuters, a global company that provides information services, and ROSS Intelligence, a technology company that develops artificial intelligence-based legal search tools. The dispute covers two main aspects: copyright infringement allegations by Thomson Reuters against ROSS and antitrust claims by ROSS against Thomson Reuters.

Thomson Reuters owns Westlaw, a platform with an extensive legal database covering court opinions, regulations and other legal sources. The platform also comes with advanced search tools to make it easier for users to navigate through the legal dataset. According to Thomson Reuters, ROSS Intelligence used copyrighted content from the Westlaw database without permission to train their artificial intelligence

system. Using this data, ROSS developed an AI-based legal search tool that aims to be more effective and efficient than Westlaw.

In response to Thomson Reuters' copyright infringement lawsuit, ROSS filed a counterclaim alleging that Thomson Reuters engaged in anticompetitive practices violating antitrust laws. Specifically, ROSS accused Thomson Reuters of "tying," which is a practice that requires customers to purchase two products as a package, even though the products can be sold separately. ROSS argued that Thomson Reuters used Westlaw's market dominance to force consumers to buy their search tool with their legal database. ROSS claimed that this practice harmed them as a competitor and stifled innovation in the legal search tool market. This is because in legal technology and services development, ROSS is trying to introduce more innovative solutions through artificial intelligence. However, Thomson Reuters' claim that ROSS illegally used their data to train this AI system put ROSS on the defensive. Meanwhile, the antitrust claim filed by ROSS is an attempt to show that Thomson Reuters is utilizing its market power to monopolize the legal search tool market and block the entry of new players like ROSS.

The case then developed into a complex legal battle. In its lawsuit, Thomson Reuters focused on copyright infringement, while ROSS in its counterclaim tried to highlight antitrust elements. The judge in the case, Judge Stephanos Bibas, had to rule on various applications, including the claim that the legal database product and the search tool were two products that should be sold separately. In addition, ROSS had to clearly define the market to support its argument that Thomson Reuters had monopoly power.

ROSS' claims against Thomson Reuters ultimately failed in court due to a lack of supporting evidence. Judge Bibas ruled that ROSS failed to show that the products Thomson Reuters sold (legal databases and search tools) were discrete products with significant consumer demand for independent purchase. In addition, ROSS was also unable to define the relevant market sufficiently to prove Thomson Reuters' monopoly power. As a result, ROSS' antitrust claims were rejected, and the court ruled in favor of Thomson Reuters in this aspect (*Thomson Reuters Enterprise Centre GmbH and West Publishing Corp. v. ROSS Intelligence Inc.*, 2024). However, the case has yet to receive a final overall verdict as Thomson Reuters' main allegations against ROSS Intelligence regarding copyright infringement are ongoing.

However, based on the preceding analysis, it is the author's opinion that the use of copyrighted data by AIs such as headnotes and key number systems, which for Westlaw are protected by their copyright, constitutes unjustified copyright infringement, unless it meets the fair use doctrine or other legal exceptions explicitly provided for in the applicable copyright law. In the context of copyright law, works such as headnotes and key number systems meet the criteria of protected works because they have significant elements of originality and creativity. Headnotes, for example, are not mere summaries of legal facts, but are the result of interpretive analysis done by the editors at Westlaw to formulate the key legal points in court decisions. Similarly, the key number system is a complex classification system that required legal expertise in its development.

In the Thomson Reuters vs. ROSS Intelligence case, ROSS's main argument that the use of this data was transformational and fell under the fair use doctrine is questionable. The fair use doctrine in US copyright law is a principle that allows the use of copyrighted works without the copyright holder's permission as long as certain conditions are met. The doctrine aims to support "the progress of science and useful arts," as stipulated in the US Constitution (Article I, Section 8, Clause 8) (Oksana, 2016). The doctrine of fair use is defined as a concept of limitation to copyright, which allows the use of copyrighted works without permission from the creator. This use must meet certain conditions, such as for educational purposes, research, or other activities that do not harm the reasonable interests of the creator or copyright holder (G. D. Ramadhan, 2021, pp. 22–23). The doctrine of fair use in United States copyright law refers to the use of a copyrighted work that does not infringe the exclusive rights of the rights holder, if the use meets certain criteria, such as the purpose and nature of the use, the nature of the work used, the amount and substantiality of the part used, and the impact on the market value of the original work. In Indonesia, this arrangement is contained in the Law of the Republic of Indonesia Number 28 of 2014 on Copyright, namely

Article 44 paragraph (1) stipulates that the taking or duplication of a work can be done without being considered an infringement of copyright if it is for purposes such as education or research, provided that it does not harm the reasonable interests of the creator. Article 46 paragraph (1) also stipulates that duplication for personal use can be done with certain restrictions.

The author's arguments are as follows. *First*, although ROSS Intelligence claims that it uses headnotes and key number systems to train machine learning-based AI models, the purpose of its use cannot be fully considered as transformational. Transformationality refers to whether the use creates new value or a different purpose from the original work. In this case, training AI to create an alternative legal search system is not sufficiently different from the original function of headnotes as a legal research tool for Westlaw. *Secondly*, the number and substantiality of the parts used by ROSS Intelligence is huge, as claimed by Thomson Reuters, involving thousands of headnotes and their classification systems. This large-scale use directly involves the core of the protected work, making it difficult to justify under the fair use doctrine. *Third*, the economic impact on the original work cannot be ignored. By using this data without authorization, ROSS Intelligence is essentially creating a product in direct competition with Westlaw, which could harm the market value of Thomson Reuters' proprietary legal database. This violates the basic principle of copyright protection, designed to protect the right holder from economic loss resulting from unauthorized exploitation of his or her work. These three arguments by the Author can be described as follows.

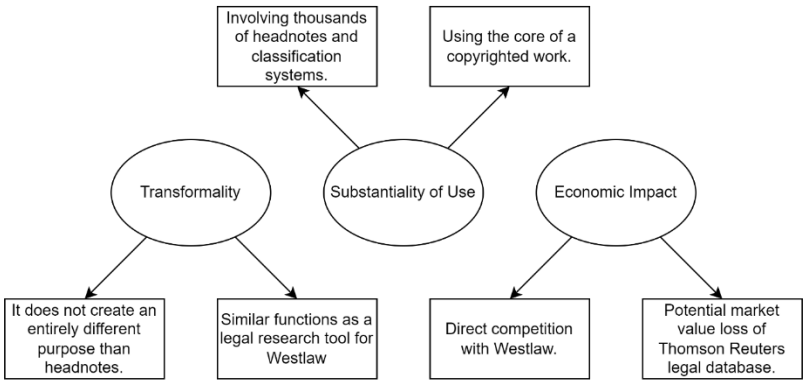

Image 3. Legal Implications of AI Training by ROSS Intelligence (Source: processed by the author, 2024)

Therefore, using copyrighted data such as headnotes and key number systems in AI training requires explicit permission from the copyright holder. Without such permission, it can be considered a violation of the law that not only harms the copyright owner, but also creates a negative precedent in using copyrighted data in the era of artificial intelligence. To prevent similar conflicts in the future, clearer regulation of the use of copyrighted data in the development of AI technologies is needed, both through specifically defined exceptions in copyright law and through transparent license contract arrangements between AI developers and copyright holders.

CONSLUSION

The development of Artificial Intelligence (AI) technology has had a huge impact in various aspects of life, including in the legal field. AI's ability to process large amounts of data and automate complex functions has driven significant changes in how humans interact with data and information. However, these advancements also present new legal challenges, particularly in relation to copyright protection of data used in the AI development process. One such case involves Thomson Reuters vs. ROSS Intelligence.

This research shows that using copyrighted data, such as headnotes and key number systems owned by Westlaw, creates a complex legal dilemma. On the one hand, these data are considered creative works protected by copyright, as they contain elements of originality and creativity in their compilation. On the

other hand, AI technology requires large amounts of data to train its models, thus encouraging technology developers to use existing data as raw material.

In the case of Thomson Reuters vs. ROSS Intelligence, using copyrighted data to train AI was judged to infringe copyright as it did not meet the fair use doctrine. The transformationality argument raised by ROSS Intelligence was difficult to accept, given that this use of the data did not create new value significantly different from the original purpose of the work. In addition, the wide scale of use and the impact on the economic value of Thomson Reuters' work strengthen the copyright infringement claim.

This case underscores the importance of clear and specific regulations governing the use of copyrighted data in developing AI technologies. Such regulations must include provisions regarding limitations on using copyrighted data, fair licensing mechanisms, and adequate protection for copyright holders. At the same time, the law should also support technological innovation by providing sufficient space for AI developers to innovate without violating the legal rights of others.

Taking these dynamics into account, this research emphasizes the need for a balance between copyright protection and the needs of technological innovation. The law must adapt quickly to meet the challenges of the artificial intelligence era, by ensuring that copyright protection does not become an obstacle to technological development. Ultimately, this research hopes to guide policymakers, technology developers, and copyright holders to create a fair, progressive legal ecosystem that supports technological advancement in the digital age. Further recommendations include:

1. Development of a national policy that balances copyright protection and technological needs;
2. Increased collaboration between AI developers and copyright holders through transparent licensing schemes; and
3. Strengthening supervision and law enforcement to prevent copyright infringement in the digital era.

Authors' Contributions

Wulan Azalia Zanzabila, as the primary author, was responsible for conducting the research, analyzing legal data, and drafting the manuscript. Maskun provided strategic direction for the research focus, motivated the primary author to undertake independent research, and contributed substantive feedback during the drafting and publication process. Aminuddin, as the academic advisor, actively guided the technical aspects, offered academic consultations, and supervised the development of the conceptual framework and research methodology. Kamsilaniah, as a collaborator from an external university, contributed through discussions and critical reviews of the research findings, enriching the comparative legal perspectives in this manuscript. All authors have read and approved the final version of this article.

Acknowledgements

This research was made possible with the invaluable support and contributions of many individuals and institutions. The authors would like to express their deepest gratitude to the Faculty of Law at Hasanuddin University and Bosowa University for providing the facilities and resources necessary for this study. We also extend our sincere thanks to our colleagues and mentors who provided insightful feedback during the preparation of this manuscript. Their guidance and constructive suggestions greatly enriched the quality of this research. Lastly, we acknowledge the inspiration and motivation from ongoing advancements in artificial intelligence and intellectual property law, which continue to challenge and broaden the scope of legal research.

REFERENCES

Ai, L. (2015). *Research on Legal Issue of Copyright Protection in the Internet*. https://doi.org/10.2991/icemet-15.2015.54

Armia, M. S. (2022). *Penentuan Metode dan Pendekatan Penelitian Hukum* (C. Fahmi (ed.)). Lembaga Kajian Konstitusi Indonesia (LKKI).

Astiti, N. M. Y. A. (2023). Strict Liability of Artificial Intelligence Pertanggungjawaban kepada Pengatur AI

ataukah AI yang Diberikan Beban Pertanggungjawaban. *Jurnal Magister Hukum Udayana (Udayana Master Law Journal)*, *12*(4), 962–980. https://doi.org/10.24843/JMHU.2023.v12.i04.p14

Bourne, C. P., & Hahn, T. B. (2003). *A History of Online Information Services, 1963-1976*. The MIT Press.

Britannica. (2024). *Thomson Reuters - Canadian company*. Britannica.Com. https://www.britannica.com/topic/Thomson-Reuters

Chanana, S., & Ing, V. K. (2023). *Fair Use Question Goes to Trial in AI Copyright Lawsuit – Thomson Reuters v. Ross Intelligence*. Farella Braun + Martel. https://www.fbm.com/publications/fair-use-question-goes-to-trial-in-ai-copyright-lawsuit-thomson-reuters-v-ross-intelligence/

Chaurasia, A., Parashar, B., & Kautish, S. (2024). Chapter 18: Artificial Intelligence and Automation for Industry 4.0. In S. Kautish, P. Chatterjee, D. Pamucar, N. Pradeep, & D. Singh (Eds.), *Computational Intelligence for Modern Business Systems*. Springer Nature Singapore. https://doi.org/10.1007/978-981-99-5354-7

Collecchia, G., & De Gobbi, R. (2024). How does Artificial Intelligence work? In *AI in Clinical Practice: A Guide to Artificial Intelligence and Digital Medicine* (pp. 7–15). Elsevier. https://doi.org/10.1016/B978-0-443-14054-9.00008-9

Delman, E. (2023). *Thomson Reuters Enterprise Centre GMBH v. Ross Intelligence Inc.* Loeb & Loeb Llp. https://www.loeb.com/en/insights/publications/2023/10/thomson-reuters-enterprise-centre-gmbh-v-ross-intelligence-inc

DeMaria, F., Hawrysh, F., Golden, F., & Brough, V. (2008). *THOMSON COMPLETES ACQUISITION OF REUTERS; THOMSON REUTERS SHARES BEGIN TRADING TODAY*. Web.Archive.Org. https://web.archive.org/web/20100819124920/http://thomsonreuters.com/content/press_room/corp/corp_news/212378

Diantha, I. M. P. (2016). *Metodologi Penelitian Hukum Normatif Dalam Justifikasi Teori Hukum*. Prenadamedia Group.

Djulvezan, M. (2000). Westlaw. *South Florida Association of Law Libraries*. https://web.archive.org/web/20120205053234/http://www.aallnet.org/chapter/sfall/newsletters/newsletter_sept2000.htm

Ferjaoui, S. (2020). Data: The New Form of Wealth and Power. *IEEE Potentials*, *39*(6), 6–10. https://doi.org/10.1109/MPOT.2020.3016359

Franceschetti, D. R. (2018). *Principles of Robotics & Artificial Intelligence*. Grey House Publishing.

Gabsa, L. D. N. (2024). Copyright Infringement in the Information and Communication Technology (ICT) Era. *Law and Humanities Quarterly Reviews*, *3*(1). https://doi.org/10.31014/aior.1996.03.01.106

Gordon, C. (2023). *Many Possible Worlds*. Springer Nature Singapore. https://doi.org/10.1007/978-981-19-9281-0

Hass, M. (2024). *Copyright, AI, and Fair Use on Trial: Thomson Reuters Enterprise Centre GmbH et al. v. ROSS Intelligence Inc.* Holon Law Partners. https://holonlaw.com/copyright-ai-and-fair-use-on-trial-thomson-reuters-enterprise-centre-gmbh-et-al-v-ross-intelligence-inc/

Ivanova, O. (2022). Industry 4.0 - a revolution requiring technology and national industrial development strategies. *Russian Journal of Management*, *10*(3), 56–60. https://doi.org/10.29039/2409-6024-2022-10-3-56-60

Jiang, Z. (2024). Artificial intelligence: a review of development, application and prospects. *Science and Technology of Engineering, Chemistry and Environmental Protection*, *1*(7). https://doi.org/10.61173/y38yzk38

Jindal, M. (2024). *31 Notable Copyright Infringement Cases*. Bytescare.Com. https://bytescare.com/blog/copyright-infringement-cases

Keefe, M. (2018). *Free vs. Westlaw : Why you need the West Key Number System*. Thomson Reuters. https://legal.thomsonreuters.com/en/insights/articles/using-the-west-key-numbers-system

Khosrow-Pour, M. (2018). *Encyclopedia of Information Science and Technology, Fourth Edition* (M. Khosrow-Pour, D.B.A. (ed.)). IGI Global. https://doi.org/10.4018/978-1-5225-2255-3

Kur, A., & Levin, M. (2011). *Intellectual Property Rights in a Fair World Trade System*. Edward Elgar Publishing Limited.

Levente, H., & Péter, K. (2023). Industrial Revolution 4.0 - A New World Order? *Journal of Trade Science*, 3–14. https://doi.org/10.54404/JTS.2023.11.01.01

Manar, R. M. (2024). *21 Famous Copyright Infringement Cases That Will Make You Cringe*. Copyrighted.Com. https://www.copyrighted.com/blog/copyright-infringement-cases

Martin, S. (2009, September 26). Prof's system revolutionized legal practice. *The Globe and Mail*.

Mattson, K. (2006). *American Legal Research: Evolution to Revolution*. https://web.archive.org/web/20080228235751/http://www.aallnet.org/chapter/sandall/workshop_SANDALLAmericanLegalResearchSept2006.ppt

Mishra, A. K. (2019). *Intellectual Property Rights In Cyberspace*. Cyberlekh Publication's.

Munir, N. (2017). *Pengantar Hukum Siber Indonesia* (3rd ed.). Rajawali Pers.

Oksana, L. (2016). The Fair Use Doctrine in the Contemporary US Copyright Law. *The Journal of the Higher School of Economics*, 186–199. https://doi.org/10.17323/2072-8166.2016.2.186.199

Potter, J. (2013). *Thomson Reuters*. The Canadian Encyclopedia. https://www.thecanadianencyclopedia.ca/en/article/thomson-reuters

Pragyna Karmakar, Satarupa Sinha, & Debrupa Pal. (2024). Artificial Intelligence. *International Journal of Advanced Research in Science, Communication and Technology*, 79–87. https://doi.org/10.48175/IJARSCT-19613

Ramadhan, G. D. (2021). Perlindungan Hukum dan Implementasi Doktrin Fair Use Terhadap Costume Play (Cosplay) dalam Hak Cipta dan Desain Industri. *Journal of Intellectual Property*, *4*(1), 13–26.

Ramadhan, M. C. (2021). *Metode Penelitian Hukum* (I. Mujtaba (ed.)). Kaizen Sarana Edukasi.

Ramli, A. M. (2010). *Cyber Law & HAKI Dalam Sistem Hukum Indonesia*. PT Refika Aditama.

ROSS Intelligence. (2020). *About Us - ROSS Intelligence*. ROSS Intelligence Inc. https://www.rossintelligence.com/about-us

Rowland, D., Kohl, U., & Charlesworth, A. (2017). *Information Technology Law* (5th ed.). Routledge.

Saidin, O. K. (2015). *Aspek Hukum Hak Kekayaan Intelektual (Intellectual Property Rights)*. Rajawali Pers.

Sovia, S. N., Hasbullah, A. R., Mustakim, A. A., Setiawan, Rachmatulloh, M. A., Rais, P., Rizal, M. C., Zahro', F., Dermawan, R., Widodo, M. F. S., Qamaria, R. S., & Kamil, H. (2022). *Ragam Metode Penelitian Hukum*. Lembaga Studi Hukum Pidana.

Thomson Reuters. (2015). *WestlawNext has been renamed Thomson Reuters Westlaw*. Thomson Reuters. https://store.legal.thomsonreuters.com/law-products/westlaw-legal-research/westlawnext-name-change

Thomson Reuters Enterprise Centre GmbH and West Publishing Corp. v. Ross Intelligence Inc. (2024). https://www.lawnext.com/wp-content/uploads/2024/09/D.-Del.-20-cv-00613-dckt-000669_000-filed-2024-09-27.pdf

Tong, X. (2021). The Emergence, Development and Trend of Data Law in China. *Journal of Commercial Law Review*, *5*(1), 36–57. https://doi.org/10.47297/wspjclrWSP2516-249702.20210501

Ujang Badru Jaman, Galuh Ratna Putri, & Tiara Azzahra Anzani. (2021). Urgensi Perlindungan Hukum Terhadap Hak Cipta Karya Digital. *Jurnal Rechten : Riset Hukum Dan Hak Asasi Manusia*, *3*(1), 9–17. https://doi.org/10.52005/rechten.v3i1.22

Verma, P. K., & Kaur, L. (2024). Research Paper on Role of Data Features and Data Collection Tools in Artificial Intelligence. *International Journal for Research in Applied Science and Engineering Technology*, *12*(3), 111–116. https://doi.org/10.22214/ijraset.2024.58693

Vibha S, B. (2022). INTELLECTUAL PROPERTY RIGHT - COPYRIGHT : GENERAL REVIEW. *INDIAN JOURNAL OF APPLIED RESEARCH*, 65–67. https://doi.org/10.36106/ijar/4213768

Wulandari, F. (2024). Problematika Pelanggaran Hak Cipta di Era Digital. *Journal of Contemporary Law Studies*, *2*(2), 99–114. https://doi.org/https://doi.org/10.47134/lawstudies.v2i2.222

Yan, Z., & Hao, Z. (2024). Study On Copyright Infringement of Artificial Intelligence Paintings. *Journal of Education, Humanities and Social Sciences*, *28*, 816–822. https://doi.org/10.54097/vacy2b08