Ethos and Pragmatic Law Review
Vol. 1, No. 1, 2025
e-ISSN: _____ (online)
https://sinesia.org/index.php/Ethos/index

Polmon As A Strategy To Monitor And Control Political Discourse On Ai-Based Social Media

Muh. Syah Quddus[1], Akhmad Nur Juniyanto[2], Zulfa Mutiara[3]
E-mail Korespondensi : muhsyahquddus017@gmail.com
Faculty of law, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

Info Article
|Submitted: 31 May 2025**|Revised**: 6 July 2025**|Accepted**: 10 July 2025
How it Cited : Muh. Syah Quddus, etc., "Polmon As A Strategy To Monitor And Control Political Discourse On Ai-Based Social Media", *Ethos and Pragmatic Law Review,* Vol. 1, No. 1, 2025, P. 101-113.

ABSTRACT

Data released by We Are Social in 2024 shows that around 61.7% of the global population uses social media, with the average duration of use per day reaching 2 hours 31 minutes. In Indonesia alone, the rate of social media usage reaches 78% of the total population, making it one of the countries with the largest social media users in the world. This high usage rate is directly proportional to the high potential for the spread of divisive information, especially during political moments such as elections. This study aims to analyze the strategies that Indonesia can implement to overcome the challenges of political polarization on social media. The method used in this research is juridical-normative research, by analyzing secondary data from primary, secondary, and tertiary legal materials. The results show that POLMON is predicted to help identify disinformation, hate speech, and provocative content that triggers political polarization. This system has great potential to improve the quality of political discussions on social media, but its implementation must be done carefully so as not to violate privacy rights and freedom of expression.

Keyword: *Artificial Intelligence (AI), POLMON, Political Polarization, Social media*

Introduction

In the growing digital era, social media has become one of the main platforms for people to interact and discuss, including in the political realm. However, this ease of access also brings serious challenges, one of which is political polarization. Political polarization on social media is often triggered by the spread of biased, provocative, and even untrue information, which ultimately creates divisions in society. This phenomenon is further exacerbated by social media platform algorithms that often reinforce confirmation bias, where users will only receive content that aligns with their political views and ignore other viewpoints. In this context, the use of artificial intelligence (AI) technology through systems such as POLMON becomes very relevant and important to implement in order to monitor and control the political polarization that is massively growing online.

Data released by We Are Social in 2024 shows that around 61.7% of the global population uses social media, with an average duration of use per day reaching 2 hours 31 minutes (We Are Social, 2024). In Indonesia alone, the level of social media usage reaches 78% of the total population, making Indonesia one of the countries with the largest social media users in the world (APJII, 2024). This high usage rate is directly proportional to the high potential for the spread of divisive information, especially during political moments such as elections. Research conducted by the Oxford Internet Institute found that political campaigns on social media often

involve the spread of disinformation by political actors to manipulate public opinion which ultimately triggers polarization (Oxford, 2021).

Artificial intelligence (AI) in this context can be an effective solution to moderate and monitor polarized content. POLMON, an AI-based system, can identify and analyze patterns of hate speech, false information, and discussions that lead to political divisions. Using natural language processing (NLP) technology, the AI can sift through millions of posts every day and categorize content based on its level of provocation. However, the application of AI in social media monitoring must also be balanced with a clear legal framework, especially in relation to the protection of freedom of speech and privacy.

From a legal perspective, the use of AI technology to moderate political content on social media must comply with regulations that protect human rights, especially freedom of expression guaranteed by Article 28F of the 1945 Constitution. This article states that everyone has the right to communicate and obtain information to develop their personal and social environment, and has the right to seek, obtain, own, store, process, and convey information using all available channels(Indonesia, 2024). In addition, Law No. 1 Year 2024 on Electronic Information and Transactions (ITE Law,2024), specifically in Article 27 paragraph (3), also regulates the prohibition of disseminating information that is insulting and defamatory, which is often part of political polarization. It is therefore important for AI systems such as POLMON to operate within the confines of these laws, ensuring that surveillance of content does not infringe on freedom of expression and other fundamental rights.

Nonetheless, there are significant challenges in implementing AI surveillance systems. One is the lack of transparency in the AI algorithms used by social media platforms. Many large tech companies are not open about how their algorithms work, making it difficult to assess whether AI is truly fair in moderating political content. This raises concerns about the potential for algorithmic bias to amplify political polarization rather than reduce it. This is where clear legal intervention is needed to ensure that the use of AI in social media moderation is done with high transparency and accountability.

The objective of this research is to explore the legal dimensions surrounding the implementation of artificial intelligence (AI) systems such as POLMON in monitoring political polarization on social media platforms. This study aims to analyze the balance between the need for technological intervention to curb the spread of disinformation and the obligation to uphold fundamental human rights, particularly freedom of expression and privacy. The findings of this research are expected to contribute to the formulation of a more comprehensive legal framework that not only ensures transparency and accountability in the use of AI for political content moderation but also protects democratic values in the digital public sphere.

The establishment of a policy framework that supports the responsible use of AI technology is crucial. Such policies should include transparency requirements for social media companies, where they are required to disclose how their AI algorithms work, especially in relation to political content. In addition, regulations should also include mechanisms for independent auditing of AI systems, to ensure that these technologies are not used to manipulate public opinion or eliminate legitimate voices in political discussions.

Research Method

This type of research is juridical-normative. Yuridis Normatif is research conducted by examining library materials or secondary materials (Marzuki, 2005). This research will examine issues related to the transformation of digital political surveillance through POLMON as a strategy for using AI in the face of social media polarization. The Data Analysis Method will be carried out by collecting secondary data through a review of library materials which include legal materials, both primary namely, the 1945 Constitution, Article 28 F and G, Law Number 1 of 2024 concerning Electronic Information and Transactions (ITE Law) Article 27 paragraph (3). Secondary legal materials are legal materials that explain primary legal materials, namely books, journals, articles, and documents related to the research theme. Tertiary Legal Materials are legal materials that complement primary and secondary legal materials, namely: Legal Dictionary, Indonesian Dictionary, and English Dictionary. The secondary data collection method is done by document study. After the data is collected, the data analysis used is qualitative data analysis. Qualitative data analysis is applied to find and describe problems in the field or structures and processes in routines and practices (Irwansyah, 2020).

This research adopts a descriptive-analytical approach in legal analysis by interpreting legal norms and principles contained within the constitutional and statutory frameworks. The qualitative data will be systematically categorized based on themes such as freedom of expression, data privacy, algorithmic accountability, and AI governance. In terms of data analysis techniques, the study applies content analysis to review relevant legal texts, policy papers, and AI system documentation (especially related to POLMON). Each data source will be examined for normative implications, legal consistency, and alignment with human rights principles. The legal arguments developed in this study will follow a logical-deductive structure to ensure clarity and legal validity. The emphasis will be placed on how legal norms respond to the technological realities of AI in moderating political discourse, including potential legal conflicts, ambiguities, and the need for interpretative flexibility.

Result And Discussion

1.1 The Role of AI Algorithms in Identifying and Moderating Polarizing Content

Artificial intelligence (AI) algorithms have become an increasingly important technology in various fields, including content monitoring and moderation on social media. One crucial aspect of concern in this context is how AI is able to detect and moderate polarizing content, such as hate speech, false information, and provocative content that can trigger political divisions. Given the increasing complexity of political discourse in cyberspace and the high use of social media as a medium of political interaction, the use of AI to monitor and control polarization is a very relevant and urgent solution.

In detecting hate speech, AI uses natural language processing (NLP) techniques that allow the system to understand the context and meaning of the text posted by social media users (Novita & Zahra, 2024). Research conducted by MIT Technology Review in 2023 showed that the ability of AI algorithms to detect hate speech has improved significantly with accuracy rates reaching 85-90% (Novita & Zahra, 2024). This includes various types of speech, including racial, religious, gender-based insults, as well as political speech that can divide society. AI systems are also capable of tracking communication patterns that lead to an increase in the intensity of negative emotions among social media users, which is often an early indicator of political polarization.

AI can also detect and flag false information or hoaxes. According to a report from The Global Disinformation Index (GDI), the spread of disinformation on social media has increased drastically over the past decade, especially in the context of political campaigns (GDI, 2021). AI algorithms equipped with machine learning techniques can be trained to recognize certain patterns in the spread of false information, such as the use of certain keywords, suspicious links, or information sources that are not credible. In this context, AI not only acts as a content detector, but also as a guardian to prevent further dissemination of potentially misleading content and muddy the political atmosphere.

A study conducted by Carnegie Mellon University in 2022 found that by utilizing historical data on the patterns of information spread on social media, AI is able to predict the spread of political disinformation with an accuracy rate of about 75-80% (CMU, 2020). This allows social media platforms to take swift countermeasures before such false information goes viral and causes further polarization. Thus, AI serves as a proactive tool to maintain information integrity in the digital public sphere.

Provocative content that is often polarizing is also targeted by AI technology. This kind of content is usually characterized by extreme rhetoric, personal attacks, and the use of sensitive issues to provoke emotional responses from social media users. With algorithms capable of recognizing this kind of communication pattern,

AI can moderate and limit its spread, especially in a political context. For example, Facebook has used AI to lower the visibility of content that is algorithmically categorized as "provocative" or "inflammatory," with the aim of reducing political polarization on its platform. As a result, research from the Center for Social Media and Politics at NYU in 2021 found that the use of AI by social media platforms was able to reduce user interactions with provocative content by 35% (NYU, 2024).

POLMON or Political Monitoring AI is an artificial intelligence-based system specifically designed to m onitor and control political polarization on social media. The system works by analyzing millions of interactions on social media in real time, flagging content deemed to lead to political divisions, and providing early warnings to authorities or social media platforms to take immediate moderation action. With the ability to detect polarizing content such as hate speech, disinformation, and political provocation, POLMON can be used to minimize the negative impact of unhealthy political discussions online.

The specific benefits of using POLMON in a political context are extensive. Firstly, it enables early detection of extremist political movements that may use social media as a tool to organize or influence the public. Second, POLMON can help the government or relevant authorities to monitor the spread of political propaganda aimed at manipulating public opinion, especially ahead of crucial moments such as elections. Third, active use of POLMON can help maintain political stability by ensuring that political discourse on social media remains healthy and does not trigger conflict in the real world.

The POLMON system also holds great potential in helping to combat the "filter bubble" phenomenon or echo chambers on social media, where users are exposed only to information that supports their own political views. By utilizing AI capable of mapping patterns of political interaction on social media platforms, POLMON can help introduce more balanced perspectives into digital conversations in a way that is non-intrusive yet effective.

On the other hand, the implementation of POLMON must comply with existing legal frameworks to ensure it consistently contributes positively. One of the main challenges in moderating political content on social media is maintaining a balance between freedom of expression and protecting the public from the effects of polarization. As stated in Article 28F of the 1945 Constitution of Indonesia, every individual has the right to freedom of expression and access to information. However, this freedom cannot be exercised absolutely if it negatively impacts public order or political stability (Indonesia, 1945). In this context, the Electronic Information and Transactions Law (UU ITE,2024), particularly Article 27, which regulates content deemed insulting or provocative, provides a legal basis for content moderation using technologies like POLMON (Indonesia, 2024). Therefore, the implementation of POLMON must be designed to align with human rights

principles and prevailing laws. AI as a content moderation tool must be accompanied by human oversight to ensure that no algorithmic bias harms any particular group. In addition, transparency in the use of POLMON must be ensured so that the public can understand how the system works and why moderation decisions are made.

The POLMON (Political Monitoring AI) system operates through a series of artificial intelligence-based processes that integrate various techniques such as big data processing, natural language processing (NLP), and machine learning (Mirza, 2023). This system is designed to monitor the dynamics of political discourse on social media, detect patterns of interaction that tend to be polarizing, and carry out moderation or preventive actions in accordance with pre-configured algorithms. Technically, POLMON functions by analyzing data collected from various social media platforms in real time, categorizing content based on the level of polarization risk, and identifying the influence of political actors who may exacerbate the situation (Nurfajri et al., 2025).

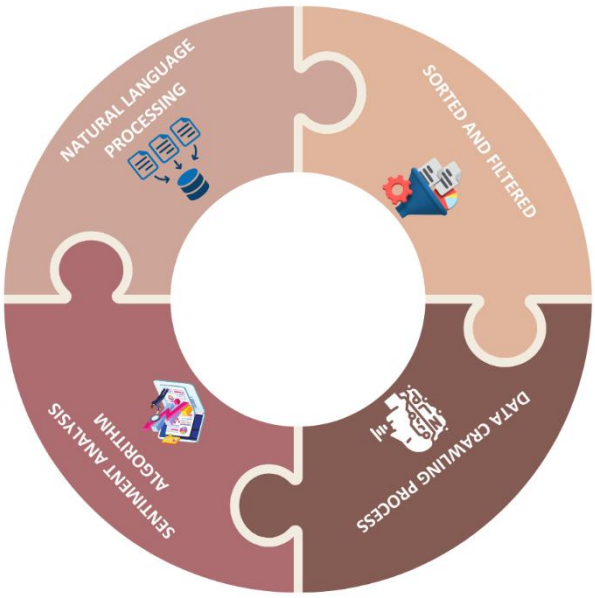


Figure 1. How the Polmon System Works

POLMON's workflow begins with a data crawling process, which involves collecting data in the form of text, images, videos, and metadata from social media platforms such as Twitter, Facebook, Instagram, or YouTube. This data is then sorted and filtered based on its relevance to specific political issues, such as election topics, public policy, or political debates. Once the data is gathered, POLMON uses natural language processing (NLP) algorithms to perform in-depth analysis of textual content, searching for keywords, phrases, or linguistic patterns that indicate provocative tendencies, disinformation, or hate speech.

The next step involves training POLMON's machine learning algorithms to recognize various types of content based on previously categorized datasets. For example, the system has been trained on datasets containing examples of hate speech, disinformation, and political propaganda from around the world. In this way, POLMON can predict whether a particular post has the potential to exacerbate political polarization within society. In addition, the system employs sentiment analysis algorithms to assess the emotional tone of the content—whether it is positive, negative, or neutral. The higher the level of negative emotion in a piece of content, the more likely it is to serve as a source of polarization.

In the next stage, POLMON applies a network analysis model to map the distribution patterns of content and identify key accounts that play a significant role in disseminating information deemed polarizing. Through this network analysis, the system can detect "main actors" or "influencers" who have a substantial impact on political polarization on social media. Once these actors and content are identified, the system can trigger alerts to social media platforms or relevant authorities to take action—such as removing content, blocking accounts, or adding warning labels to the content.

Overall, POLMON operates within a multi-tiered architecture, where the system first filters content based on relevance, then analyzes the content using AI-based algorithms, and finally issues moderation or intervention actions when necessary. This approach allows POLMON to work quickly and efficiently in moderating thousands to millions of pieces of content in a short period, without the need for direct human intervention at each step.

1.2 Policy Design to Prevent Political Polarization Amplified by Technology

The use of artificial intelligence (AI) in moderating content on social media presents both opportunities and challenges for legal policies, especially in the context of preventing political polarization, which is increasingly amplified by digital interactions. Policy design that supports the use of systems like POLMON (Political Monitoring AI) must consider various crucial legal and ethical aspects. Political polarization, often triggered by the spread of disinformation, hate speech, and provocation on social media, requires technology-based oversight to be effectively controlled (Purboyo et al., 2024). However, the implementation of AI in this regard presents significant legal implications, particularly concerning privacy, freedom of expression, and the responsibilities of social media platforms.

The use of AI in social media monitoring, particularly in moderating political discussions, requires a thorough analysis of its legal implications. First, privacy issues become a primary concern. Systems like POLMON operate by collecting and analyzing data from millions of social media users, including both private and

public conversations. This raises concerns regarding violations of Article 28G of the 1945 Constitution, which guarantees every individual's right to personal protection, family, honor, dignity, and the right to feel secure and protected from threats (Indonesia, 1945). Policies underlying the use of AI in monitoring political discourse must ensure that data collection does not arbitrarily violate users' privacy rights.

The use of AI in moderating political discussions can also have implications for freedom of expression, as outlined in Article 28F of the 1945 Constitution. AI systems that automatically censor or moderate content deemed provocative or likely to trigger political polarization could potentially restrict the public's right to express opinions and ideas. The sophistication of AI algorithms in detecting hate speech or disinformation must be balanced with transparent and accountable mechanisms to ensure that excessive or arbitrary censorship does not occur.

Finally the responsibility of social media platforms in using AI to moderate political content must also be strictly regulated. Under the Electronic Information and Transactions Law (UU ITE, 2024), platforms are obligated to ensure that the content circulating on their platforms does not violate the law. However, the question arises as to the extent to which platforms are responsible for moderation decisions made by AI. Regulations regarding platform responsibility must be clear, including provisions on users' right to appeal if they feel harmed by AI decisions in content moderation (Quddus, 2024).

To ensure that the use of AI in moderating political content is conducted transparently and accountably, several legal and policy recommendations must be considered. First, algorithmic transparency should be a primary requirement. Social media platforms that use AI to moderate political content should be obligated to publicly disclose how the algorithms operate, including the parameters used to identify content deemed risky. Regulations should encourage technology companies to provide independent third parties with access to audit these algorithms.

Strict public oversight of AI usage on social media is also necessary. The government could establish a regulatory body tasked with overseeing the use of AI by social media platforms, ensuring that the moderation process is conducted fairly and without discrimination (Quddus et al., 2025). This body should also have mechanisms in place to receive complaints from users who feel harmed by the automated moderation actions taken by AI.

Regulations should also include an appeal mechanism for social media users. Any user whose content is removed or moderated by AI should have the right to appeal and receive a clear explanation of the reasons behind the decision. This process should be conducted swiftly and transparently to prevent violations of the right to freedom of expression.

The implementation of POLMON in Indonesia requires a well-planned and systematic approach. The following are the steps that can be taken to integrate the POLMON system into the monitoring of political content on social media:

Table 1. POLMON Implementation Steps

No	Steps	Description
1.	Technology Development	Developing AI algorithms tailored to Indonesia's local needs, including involving experts in technology and law.
2.	Regulatory Framework.	Establishing regulations that govern the use of AI in social media moderation, including provisions on privacy, freedom of expression, and platform accountability.
3.	Limited Trial	Conducting a limited trial of the POLMON system by monitoring political discussions on selected social media platforms.
4.	Stakeholder Engagement.	Involving the government, social media platforms, NGOs, and the public in open discussions about the transparency and accountability of the system.
5.	Full-Scale Implementation	Integrating POLMON into full-scale social media monitoring with a refined regulatory framework.
6.	Evaluation and Audit.	Conducting regular evaluations and audits of POLMON's performance to ensure the system operates effectively and fairly.

Conclusion

POLMON is an AI-based system designed to monitor and control political polarization on social media. With the ability to analyze millions of pieces of content in real-time, POLMON can help identify disinformation, hate speech, and provocative content that triggers political polarization. This system holds great potential to improve the quality of political discussions on social media, but its implementation must be carried out with caution to avoid violating privacy rights and freedom of expression.

The use of AI in moderating political content requires clear and transparent regulations, as well as strict public oversight. Algorithmic transparency and an appeal mechanism for users affected by moderation should be integrated into policies to ensure that the basic rights of the public are protected. With proper implementation, POLMON can become an effective tool in preventing political polarization and improving the political climate in Indonesia.

To strengthen the implementation of POLMON, it is recommended that the Indonesian government establish an independent supervisory body consisting of

multidisciplinary experts in law, technology, human rights, and political communication. This body should be tasked with regularly auditing the AI algorithms used, assessing their impact on civil liberties, and ensuring compliance with national and international legal standards. Public participation should also be encouraged through stakeholder forums and transparent reporting mechanisms to maintain accountability and trust in the system.

Suggestion

Based on the results of this study, it is recommended that future research focus on developing and evaluating policies that support the ethical and transparent implementation of the POLMON system in Indonesia. Such research needs to examine in depth the public oversight mechanism, algorithm transparency, as well as the effectiveness of appeal channels for users affected by content moderation. In addition, it is important to explore how multi-stakeholder engagement-including government, academia, civil society, and digital platforms-can create a political surveillance ecosystem that is fair, accountable, and respectful of human rights in an increasingly complex digital era.

Acknowledgments

Our sincere thanks go to the reviewers and anonymous editors who have provided constructive feedback for improving the quality of this paper, making it worth reading and referencing.

Biblioghraphy

APJII. (2024). APJII Jumlah Pengguna Internet Indonesia Tembus 221 Juta Orang. *Asosiasi Penyelenggara Jasa Internet Indonesia*. https://apjii.or.id/berita/d/apjii-jumlah-pengguna-internet-indonesia-tembus-221-juta-orang

CMU. (2020). Center for Informed Democracy & Social—Cybersecurity (IDeaS) CMU's center for disinformation, hate speech and extremism online. *Carnegie Mellon University*.https://www.cmu.edu/ideas-social-cybersecurity/research/disinformation.html

GDI. (2021). We exist to disrupt online disinformation. *Global Disinformation Index*. https://www.disinformationindex.org/about

Indonesia. (1945). *Undang-Undang Dasar Tahun 1945*.

Indonesia. (2024). *Undang-Undang Nomor 1 tahun 2024 Tentang Informasi dan Transaksi Elektronik*.

Irwansyah. (2020). *Penelitian Hukum : Pilihan Metode & Praktik Penulisan Artikel*. Mirra Buana Media. https://onesearch.id/Author/Home?author=Prof.+Dr.+ Irwansyah% 2C+SH.%2C+MH

Marzuki, P. M. (2005). *Penelitian hukum*. Prenadamedia Group. https://scholar.google.co.id /citations?view_op=view_citation&hl=en&user=NZea4P4AAAAJ&citation_for_vie w=NZea4P4AAAAJ:isC4tDSrTZIC

Mirza, I. M. M. (2023). Implementasi Artificial Intelligence Dalam Iklan Politik Menuju Masyarakat Indonesia 5.0. *VISUALIDEAS*, *3*(2), 65–82. https://doi.org/10.33197 /visualideas.vol3.iss2.2023.1442

Novita, Y., & Zahra, R. (2024). *Penerapan ArtificialIntelligence(AI) untuk Meningkatkan Efisiensi Operasional di Perusahaan Manufaktur: Studi Kasus PT. XYZ*. *1*(1).

Nurfajri, I., Pratama, E. T. H., Tupamahu, G. S., Saputra, R., & Erwina, Y. (2025). Dampak Algoritma AI terhadap Komunikasi Publik: Memahami Manipulasi Informasi dan Realitas. *Converse Journal Communication Science*, *1*(3), 13. https://doi.org/ 10.47134/converse.v1i3.3543

NYU. (2024). NYU's Center for Social Media and Politics works to strengthen democracy by conducting rigorous research, advancing evidence-based public policy. *Universitas New York*. https://csmapnyu.org/

Oxford. (2021, January 13). Social media manipulation by political actors an industrial scale problem—Oxford report. *Oxford Internet Institute*. https://www.ox.ac.uk/news/ 2021-01-13-social-media-manipulation-political-actors-industrial-scale-problem-ox ford-report

Purboyo, M. G., Junaidi, J., & Amriwan, A. (2024). *Eskalasi Partisanship Politik Dalam Pembentukan Polarisasi Pada Masyarakat Digital (Studi pada Platform Media Sosial TikTok)*.

Quddus, M. S., Nufitasari, M., Suraya, I., & Putranti, D. (2025). Legal Challenges in the Development of Green AI to Support Indonesia's National and International Commitments to Sustainable Technology. *Journal of Law and Policy*, *1*(2).

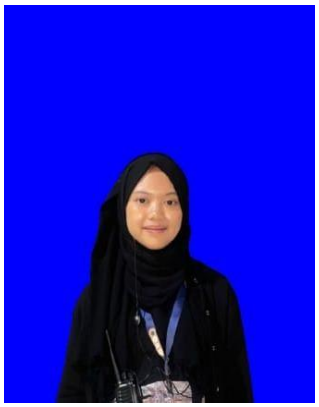
Quddus, Muh. S. (2024). Implications law in ethics sustainable business in Indonesia. *Gema Wiralodra*, *15*(1), 596–605. https://doi.org/10.31943/gw.v15i1.693

We Are Social. (2024, January 31). Digital 2024: 5 miliar pengguna media sosial. *We Are Social*. https://wearesocial.com/id/blog/2024/01/digital-2024-5-billion-social-media - users/

Biografi Singkat Penulis

Muh. Syah Quddus, S.E., adalah lulusan S1 Manajemen Bisnis dari Universitas Cendekia Mitra Indonesia tahun 2024 dan saat ini sedang melanjutkan studi S1 Ilmu Hukum di Universitas Ahmad Dahlan sejak tahun 2021. Ia memiliki minat yang mendalam dalam bidang hukum, khususnya pada Hukum Bisnis, Hukum Perdagangan Internasional atau isu-isu Hukum Internasional, dan Hukum Teknologi. Dalam perjalanan akademiknya, ia telah menerbitkan beberapa publikasi ilmiah yang terindeks baik nasional maupun internasional, di antaranya yang relevan dengan topik penelitian berjudul *Proposing Digital Copyright Infringements Enforcement through Internet Court in Indonesia:Fostering the Existing E-Court System*, Pemanfaatan *Digital Twin* dalam Perlindungan Hak Kekayaan Intelektual Sebagai Strategi dalam Mengoptimalkan Potensi Ekonomi Halal di Indonesia, *Copyright Challenges and Opportunities for the Integration of Generative Artificial Intelligence (GenAI) in Indonesian Higher Education Learning*, Serta *Sustainable Ip-Tourism* Sebagai Langkah Inovatif Dalam Mengakselerasi Pertumbuhan Ekonomi Pariwisata Indonesia. Ia juga Aktif mengikuti berbagai kegiatan konferensi nasional dan internasional serta kegiatan pengabdian masyarakat internasional. Ia saat ini bekerja sebagai *Junior Editorial* pada salah satu jurnal hukum yang terindeks Sinta 1 & Scopus Q3 di Yogyakarta.

Akhmad Nur Juniyanto lahir di Cirebon pada 1 Juni 2005 dan tinggal di Dusun Banaran, Sidoharjo, Sruweng, Kebumen, Jawa Tengah. Ia menyelesaikan pendidikan di SD IT VIP Al-Huda Adimulyo, SMP Somolangu Kebumen, dan SMAN 1 Pejagoan. Pada tahun 2023, ia diterima di Fakultas Hukum Universitas Ahmad Dahlan melalui jalur PMDK-SKL dan saat ini aktif sebagai mahasiswa semester 3. Selama menjadi mahasiswa, ia terlibat aktif dalam organisasi kampus, di antaranya sebagai pengurus Lantern *Law Community* dan *Liaison Officer* (LO) Mahasiswa Baru 2024.



Zulfa Mutiara lahir di Bondowoso, 8 Juni 2003, dan tinggal di Desa Sucolor, Kecamatan Maesan, Bondowoso. Ia merupakan anak kedua dari dua bersaudara. Pendidikan formalnya dimulai di MI Miftahul Ulum Sucolor, dilanjutkan ke MTsN Manggarai Barat, dan SMAN 1 Darussholah Singojuruh Banyuwangi. Pada 2023, ia diterima di Fakultas Hukum Universitas Ahmad Dahlan Yogyakarta melalui jalur PMKD-SKL. Selama kuliah, ia aktif dalam berbagai organisasi dan kegiatan, seperti menjadi pengurus Barisan Anti Korupsi Ahmad Dahlan dan Lantern Law Community. Ia pernah menerbitkan buku antologi puisi berjudul Masih Hujan Tetap Hujan dan Akan Selalu Hujan, serta meraih Juara 2 Debat TKD-PAN DIY. Pada 2022, bersama timnya, ia memperoleh Medali Silver di ajang Indonesia Inventors Day atas inovasi media pembelajaran The Reporter, yang mengangkat isu pencegahan perkawinan anak, kesehatan seksual, dan gender berbasis boardgame kearifan lokal.