Forensic Justice And Socio-Legal Research Journal
LEXFORENSICA
Vol. 2, No. 2, pp. 74-84 | December 2025
CFORSJ Journal

MODELING AI-GENERATED MISINFORMATION: SHARIAH, ETHICAL AND FORENSIC PERSPECTIVES FROM MALAYSIA

[i,*]Hadi Akbar Dahlan, [ii]Hadi Affendy Dahlan & [iii]Nur Khalidah Dahlan

[i]Leave a Nest Malaysia, Block 2330, Century Square, Jalan Usahawan, Off Persiaran Multimedia, 63000 Cyberjaya, Selangor, Malaysia
[ii]Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), 43600, Bangi, Selangor, Malaysia
[iii] Faculty of Law, Universiti Kebangsaan Malaysia (UKM), 43600, Bangi, Selangor, Malaysia

*(Corresponding author) e-mail: hadi@lne.st

Article history:
Submission date: 1 October 2025
Received in revised form: 5 November 2025
Acceptance date: 1 December 2025
Available online: 17 December 2025

Keywords:
System dynamics, AI-generated content, misinformation, Malaysia, society

Funding:
This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

Competing interest:
The author(s) have declared that no competing interests exist.

Cite as:
Dahlan, H. A., Dahlan, H. A., & Dahlan, N. K. (2025). Modeling AI-generated misinformation: Shariah, ethical and forensic perspectives from Malaysia. *LexForensica: Forensic Justice And Socio-Legal Research Journal*, *2*(2), 74-84. https://doi.org/10.33102/yy54vr91

SDG Elements:
Peace, Justice, and Strong Institutions

ABSTRACT

This study investigates the growing threat of AI-generated misinformation in Malaysia through an integrative framework combining system dynamics modeling with Shariah, ethical, and forensic analyses. A system dynamics model was developed to simulate the spread of AI-driven deceptive content among Malaysian internet users using key variables such as population size, monthly website visits, and deception rates. Simulation results show that increasing AI-generated content from 30% to 50% significantly accelerates the time required for misinformation to reach the entire internet-using population, from approximately four years to just two and a half years. This rapid propagation creates a critical mismatch between the pace of technological harm and the slower processes of legislative, forensic, and Shariah-based institutional responses. From a Shariah perspective, pervasive AI misinformation threatens the credibility of *fatwas*, *Halal* certification systems, and the authenticity of religious knowledge. Ethically, repeated exposure erodes trust, undermines civic engagement, and weakens adherence to *tabayyun*. From a forensic standpoint, the ubiquity of synthetic digital content complicates attribution, evidence validation, and prosecution, enabling both sophisticated cybercrimes and the "deepfake defense". The findings highlight the urgency for anticipatory governance, harmonized religious and legal frameworks, and strengthened digital forensic readiness to safeguard societal cohesion and the integrity of Islamic knowledge in the age of generative AI.

Introduction

The rapid proliferation of Artificial Intelligence (AI) generated content necessitates an in-depth examination of its societal impact, especially within culturally and religiously sensitive contexts like Malaysia (Brenner et al., 2024). The advent of artificial intelligence has significantly transformed digital media ecosystems, influencing how disinformation and hate speech are produced, disseminated, and consumed (Sonni, 2025). This rise of AI-driven content introduces unique challenges, making it increasingly difficult to discern truth from fiction in media consumption (Brenner et al., 2024). Within Southeast Asia, countries like Malaysia and Indonesia are actively grappling with the governance of AI, with Malaysia developing a national AI Code of Ethics and Indonesia focusing on principles such as human supervision, data governance, and fairness (Nilgiriwala et al., 2024). There is also a recognized regional need for consolidated AI governance frameworks to address these evolving concerns (Chua et al., 2023).

The ethical considerations surrounding AI extend deeply into Shariah principles, especially in Muslim-majority nations like Malaysia and Indonesia (Gorian & Osman, 2024). Research highlights the importance of developing comprehensive Shariah governance frameworks to ensure AI applications align with Islamic legal and ethical principles, thereby mitigating risks associated with impermissible practices (Zafar & Ali, 2025). Ethical deviations in AI tools, particularly in marketing practices, are critically reviewed from a *Halal* perspective, emphasizing concerns around misinformation, manipulation, and exploitation that contradict Islamic values of justice, integrity, and harm avoidance (Koswara & Herlina, 2025). Furthermore, discussions on the interaction between humans and advanced AI, such as Artificial General Intelligence, from an Islamic perspective are imperative. Issues like distinguishing between *fitna* (strife/temptation) and facts (*tabayyun*), and *fiqh* considerations related to AI and robots, particularly in areas like *Halal* food preparation and trust in AI systems, underscore the urgent need for Muslim scholars to initiate and lead these conversations. Despite a growing body of literature from Western and Eastern viewpoints, Muslim communities have been comparatively "laggard" in discussing the integration of AI technology into Muslim life, highlighting a critical gap that needs to be addressed (Dahlan, 2018). The urgent need for an Islamic frame of reference to ethically guide AI-related research and development is also underscored (Nawi et al., 2023). The region, including Indonesia, is particularly vulnerable to AI-driven disinformation due to factors such as varying levels of media literacy and growing dependence on online platforms (Putra, 2024; Sonni, 2025).

Addressing AI-generated misinformation also demands robust forensic capabilities. The ease with which generative models can produce high-quality, yet potentially malicious, images and videos necessitate advanced detection technologies to unmask AI-created visual content (Zhang et al., 2025). The field of multimedia information security is rapidly advancing in detecting AI-created visuals and deepfakes (Zhang et al., 2025). Forensic methodologies are being developed to identify image manipulations and detect fake AI-generated content, crucial for safeguarding information authenticity and social trust (Lee et al., 2023; Yu et al., 2024). This study leverages a system dynamics model to quantify the dissemination rate of AI-driven deception among Malaysian internet users, thereby providing a robust framework for understanding the mechanisms of misinformation propagation (Hossain et al., 2024). By examining these mechanisms through Shariah, ethical, and forensic lenses, this research aims to offer critical insights for understanding, mitigating, and countering the adverse effects of AI-generated misinformation within the Malaysian context, ultimately paving the way for more resilient information ecosystems that respect regional moral and legal implications (Alfiani & Santiago, 2024).

System Dynamics is a robust methodological approach and computer simulation modelling technique developed by Jay Forrester in the mid-1950s at MIT, designed to understand the complex, dynamic behavior of systems over time (Azar, 2012; Chen et al., 2022; Sterman, 2002). Its foundational principle posits that the inherent structure of a system dictates its behavior, emphasizing a continuous view of system evolution rather than discrete events (Cerabona, 2023; Fisunoğlu, 2018). Key components of SD models include "stocks" (accumulations) and "flows" (rates) that alter these stocks, interconnected by intricate feedback loops, both reinforcing and balancing, that represent causal relationships within the system (Fisunoğlu, 2018; Forliano et al., 2024). These models, often equation-based, are simulated to explore how internal structures and policies drive system performance, aiding in understanding long-term trends, identifying leverage points for intervention, and predicting unintended consequences across

diverse fields (Azar, 2012; Harrison et al., 2025). The objective of this study is twofold: first, to model the spread of AI-generated misinformation among the Malaysian population using system dynamics modelling; and second, to discuss the implications of the model's simulation results from Shariah, ethical, and forensic perspectives.

Methodology

This study employed system dynamics modelling to simulate the spread of AI-generated misinformation within the Malaysian population. Model variables were either directly derived from existing literature or synthesized from relevant scholarly sources. Subsequently, these variables underwent discussion and validation among the authors to ensure model clarity and robustness. The system dynamics model was built in Vensim PLE software version 9.4.2.

System Dynamics Model Description

Figure 1 shows the System Dynamic model constructed for this study. The model illustrates how the "Population of Malaysia" contributes to the "Monthly unique websites visit per month". This metric is then influenced by the "Rate of successful deception by AI content", leading to a flow variable called "No of successful deceived internet users monthly". This flow ultimately accumulates into the stock variable "Deceived internet users in Malaysia". In essence, the model aims to quantify how the Malaysian population, through their website visits, becomes successfully deceived by AI content over time. Model variables were either directly derived from existing literature or synthesized from relevant scholarly sources and validated among the authors.

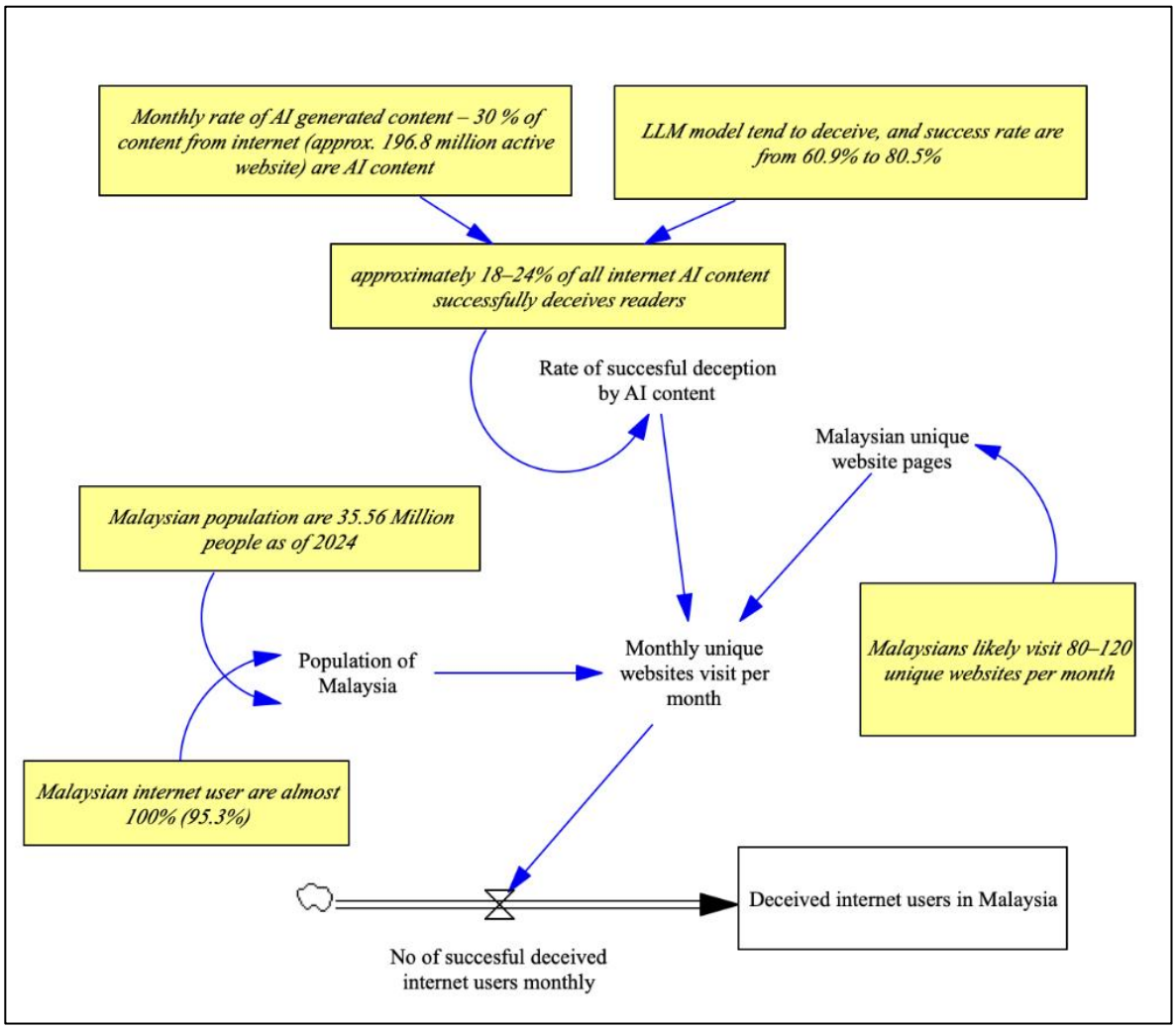

Figure 1. System dynamics model describing the simulation of Deceived internet users in Malaysia. Box in yellow denotes commentary of the variables used in the model

Variables in the Model

The key variables used in the construction of this model are the Malaysian population, the monthly unique website visits, the rate of successful deception by AI content, and the cumulative number of deceived internet users. Table 1 shows the reference for the variables.

Table 1. List of variables and its reference used in the model

Variable	Reference
Population of Malaysia	Google Statistics
Rate of successful deception by AI Content	(Spennemann, 2025; Wu et al., 2025)
Malaysian Unique Website visit	(Rashid & Dawood, 2024) (Department of Statistics Malaysia, 2023)
Monthly Unique Website visits per month	Synthesized for this study
No of successful deceived internet users monthly	Constructed for ease of extracting and visualizing data
Deceived internet users in Malaysia	

Policy Simulation

The development of the System Dynamics Model allows the policy simulation. In this context, the model allows simulation of different parameters of the variables to assess the impact of AI-generated misinformation. For this study, the default value of the monthly rate of AI-generated content was 30%. This study simulated the impact of increasing this percentage to 50% to observe the corresponding changes in the "No of successful deceived internet users monthly". The authors set the value to 50% because such value is considered a conservative yet impactful increase, reflecting a plausible future scenario given the rapid proliferation of generative AI technologies and their widespread adoption.

The author did not increase the value of success rate of deceive by LLM since the default value is already above the 50% threshold of conservative value. Both of these parameters are reflected in the variable "Rate of successful deception by AI Content" in the model.

Results and Discussion

Simulation Model Result

Using the system dynamics model, we simulated the growth in the number of deceived internet users in Malaysia under two levels of AI-generated content on the internet: 30% and 50%. The simulated results from the model were extracted and plotted into a graph. Figure 2 shows the graph of time taken to deceived internet users in Malaysia.

The model outputs were converted to yearly cumulative counts of deceived users and compared to the 2024 population of Malaysia (≈36 million people). As shown in Figure 2, increasing the proportion of AI-generated content from 30% to 50% substantially accelerates the time required for the number of deceived users to reach the national population.

In the 50% AI-content scenario, the cumulative number of deceived users crosses the population threshold after approximately 2.5 years. In contrast, when only 30% of online content is AI-generated, the same threshold is reached only around year 4. By the end of year 4, the 50% scenario has already surpassed this level by a wide margin, while the 30% scenario is just reaching it. These results indicate that higher penetration of AI-generated content in the internet ecosystem can shorten the time needed for misinformation to affect an entire population by roughly one and a half years.

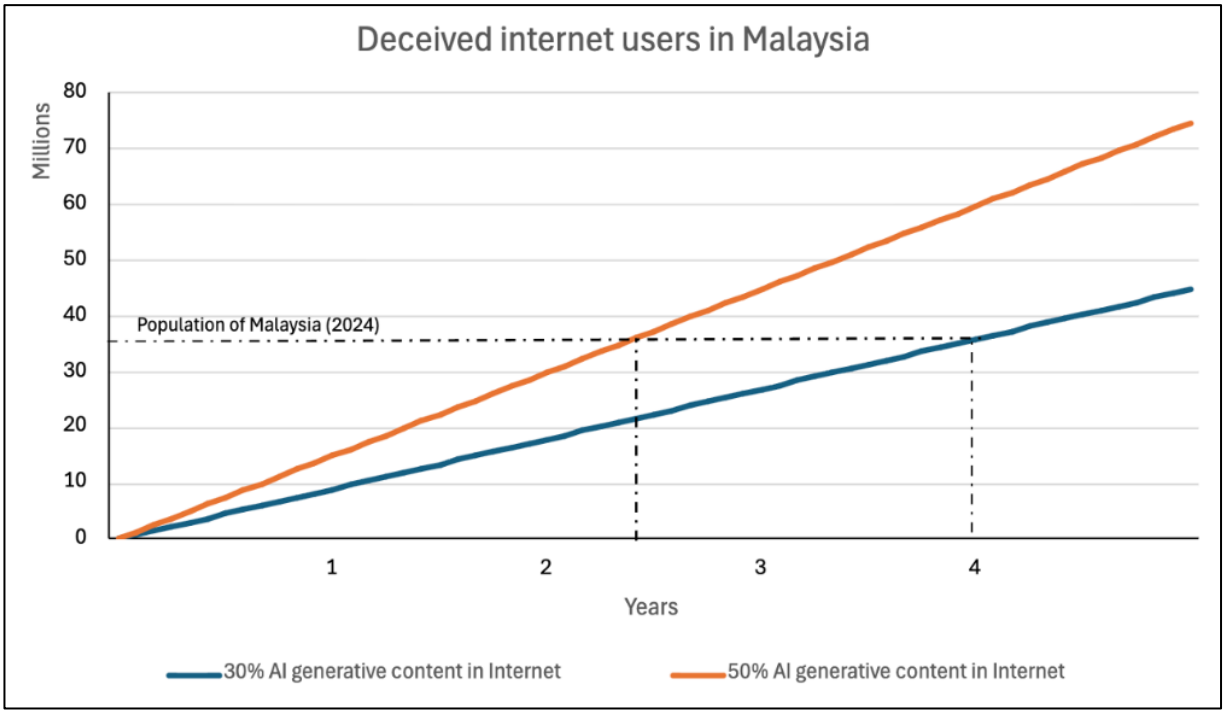

Figure 2. Graph of time taken to deceived internet users in Malaysia using extracted data from System Dynamics Simulation

Model Implications for Readiness in Islamic Legal Processes

The time frame of around 4 years can be considered challenging for a holistic intervention through law to be conducted. This because any directive from government may take some considerable time unless there is special need or cases. In Malaysia, the Department of Islamic Development Malaysia together with the National Fatwa Committee and State Mufti Department can only drafting and reviewing draft legislation for government review, but the final power to pass laws is with the elected legislative bodies. Even then, such draft legislation might take considerable times due to non standardization or acceptance of discussion or fatwas by both National Fatwa Committee Meeting and the State Mufti Department (Mohamad & Mamat, 2018). Without the fatwa by these organization, some offences can't be enforced because relevant fatwa have not been gazetted (Saifuddin et al., 2024).

Given these institutional and procedural complexities, producing a fully operational law within four years is achievable but highly unlikely. Based on the model, if the rate of AI-generated content rises to 50%, the projected timeframe decreases to approximately 2.5 years. Such an accelerated timeline would make the process even more demanding, and thus the likelihood of successful legislative completion becomes even lower. This is not to say that we need to sacrifice the quality, rigor, and public legitimacy of the resulting law.

Rapid drafting under time pressure risks inadequate consultation with stakeholders, limited expert input, and poor alignment between federal and state religious authorities. This could create enforcement gaps, judicial confusion, and pushback from state bodies that view the process as rushed. Religious laws involve deep social, historical, and theological stakes, so procedural flaws could harm the law's success and public trust in institutions.

Thus, while the model underscores the urgency of rising AI-generated content, it also highlights the need for deliberate, coordinated, and institutionally aligned legislation. Without improved inter-agency collaboration, standardized *fatwa* processes, and greater legislative capacity, meeting these tight timelines is unlikely. Ultimately, the model's projections signal the systemic reforms required to equip Malaysia's legal and religious institutions for emerging technology challenges which echoing the findings of other studies (Khairuldin et al., 2019) (Abidin & Nasohah, 2023).

Potential Impact from Ethical Perspective

The impact of repeatedly deceived by AI misinformation poses significant ethical challenges, primarily through the erosion of trust; when individuals are repeatedly exposed to deception, they may cease to believe legitimate news, authorities, and even scholars, leading to widespread apathy (Drolsbach & Pröllochs, 2025; Li et al., 2025; Storm-Mathisen, 2024). This kind of apathy is dangerous to *ummah* since it inhibit participation in communal activities. Based on previous studies, populations that is repeatedly exposed to "deception/corruption" in politics leads to reduced civic participation (Rivera et al., 2024). Although the initial cause of distrust is due to digital misinformation (AI-generated content), it will eventually affect the everyday physical life due to rampant exposure.

Another potential impact is the lack of initiative to follow the principles of *tabayyun* (searching for truth). This principle, central to Islamic epistemology, emphasizes verifying information before acceptance, a practice increasingly vital in an age saturated with AI-generated content (Dahlan, 2018) (Taha et al., 2025). When this principles is lacking in the *ummah*, this can lead to the normalization of lying and manipulation and could shift the moral boundaries where an action that one once considered haram (such as fraud, slander or fake testimony) become a tool or content for marketing (Hagendorff, 2024; Tarsney, 2025). Normalization of lying also dulls ethical sensitivity and this invite malicious actors to exploit these vulnerabilities in the *ummah*.

The consequences of this kind of normalization are already visible today. Exploitation of the elderly, low-income groups and those with low digital literacy had already happened. This is actual an impact from the current digital technology. However, as AI systems become far more powerful, the scale and severity of that exploitation could grow dramatically. For example, large-scale disinformation and fear campaigns could be aimed at specific neighbourhoods or demographic groups, with AI systems adapting in real time to people's reactions and constantly refining the most effective lies to provoke panic, hatred, or compliance. The authors afraid that this could become reality within a few years based on the model output if no significant effort is done.

Potential Impact from Forensic Perspective

From a forensic perspective and law-enforcement standpoint, the nature of AI-generated content itself already suggests a future where the very evidence itself is always suspected and subject to doubt. This situation challenged the established investigative and judicial processes. Due to pervasive amount of AI-generated misinformation, it will become difficult to make attribution and traceability of information. In cases involving harassment, online hate, terrorism propaganda, or fraud may fail in court due to "reasonable doubt" regarding the authenticity of digital evidence (Sandoval et al., 2024; Wan Ismail et al., 2024). Furthermore, the integrity of digital evidence is further compromised by deepfakes and AI-edited documents, leading to situations where genuine digital files can be claimed as "fake", and defense lawyers might argue that incriminating chats, audio, or video "could have been generated by AI" (Casu et al., 2023; Linna et al., 2024).

Another impact of proliferation of AI-generated misinformation is the dull of ethical sensitivity, which can lead to moral disengagement. This meant that a person in a community can justify doing things (including crime) that go against their morals (Bandura, 1999). In this context, universal experiencing misinformation by AI within the society will lead to the society of even condoning using AI tools to do misinformation themselves (Luo & Bussey, 2023). Together with the difficulty of attribution and traceability, this paves the way for increasingly complex and undetectable misinformation campaigns and digital crimes, making it harder to discern truth from deception (Gabriel et al., 2024), and increasing the confidence of the criminals with their crimes since prosecution will be very difficult (Klasen et al., 2024; Treleaven et al., 2023).

As mentioned in the previous section, AI-generated misinformation can lead to engineering targeted hatred toward a specific community. Inversely, AI-generated misinformation can also create targeted propaganda to bolster support for certain ideologies or political agendas, making the manipulation of public opinion more subtle and pervasive (Walker et al., 2024). It is highly possible with the advent of AI, creation of digital personas, and sophisticated impersonation tools will make it increasingly difficult to discern authentic human communication from synthetic fabrications, further complicating forensic

investigations (Shoaib et al., 2023). This erosion of trust, coupled with the "deepfake defense" phenomenon (a phenomenon where AI-generated evidence is submitted and used in the courtroom), poses significant challenges to the legal system, as the very concept of digital evidence veracity becomes inherently questionable (Tortora, 2024). The ease with which general-purpose AI systems can generate convincing but fabricated content, including deepfake pornography and other reputation-damaging materials, further complicates forensic analysis and introduces new avenues for criminal exploitation (Bengio et al., 2024). The proliferation of AI-generated misinformation at that stage is now profoundly confronting and undermining the sacred doctrines of the Quran and the foundational principles of Muslim life (Dahlan, 2018).

Potential Impact from Shariah Perspective

From a Shariah perspective, the proliferation of AI-generated misinformation directly threatens the integrity of religious discourse and the authenticity of Islamic knowledge, particularly concerning online *fatwas* and *Halal* certification (Dahlan, 2022). Documentation and traceability of *Halal* certification will be very important since AI-generated misinformation can easily sow doubts among the *ummah*. Likewise, online *fatwas* and religious knowledge must be readily verifiable by the *ummah* (Masruha et al., 2025). This is important because it will serve as a defense against proliferation of deepfakes involving Islamic figures. AI ability to produce convincing audio and visual content, including deep-fake voices for clerics and video lectures, means that "seeing is believing" has transformed into "seeing is deceiving," leading to widespread misinformation and distrust in authentic religious sources (Ilyas et al., 2025; Wahab, 2025).

There are some Islamic scholars and Muslim developers who are exploring the AI technology to better serve the *ummah*. However, the very tools that facilitate the dissemination of religious content also carry the risk of generating spurious *da'wah* texts or altering the voices of respected clerics, thereby undermining the credibility of religious guidance (Ilyas et al., 2025). This because we are still debating the ethics of using AI for religious purposes. This raises significant concerns about accuracy, credibility, and the potential erosion of established religious guidance, as AI-generated interpretations often lack the required depth, historical, and religious context (Al-Janabi, 2024; Maha et al., 2025; Malik, 2023; Niam, 2024).

In the future where AI-generated misinformation is ubiquitous and religious authority is apathy, there will be decline in trust not only in religious institutions but also in the authenticity of religious knowledge, potentially fostering agnosticism or even atheism within the Muslim community (Dahlan, 2018). Alternatively, the Muslim world at that stage could face an unexpected spiritual drift. A growing segment of believers might come to rely almost entirely on AI systems for religious guidance, treating algorithmic responses as unquestionable truth. Detached from traditional scholarship and unaware of subtle fabrications, such as invented Quranic or hadith texts. The creation of this "pseudo-Muslim" identity could emerge as a synthetic faith built on unverified digital authority. At the same time, broader apathy within the *ummah* may allow popular sentiment, amplified by progressive digital narratives, to overshadow the influence of classical *ulama'*. As vox populi quietly redirects religious norms, the fragmentation of the *ummah* could remain hidden beneath the rhetoric of modernization. In such an environment, even practices once considered unimaginable, like AI led long-distance congregational prayer might gain acceptance, reflecting a community reshaped not through deliberate reform, but through gradual surrender of authority to technology and social trends

AI Misinformation Outpacing Institutional Defences

Taken together, the system dynamics model and the subsequent Shariah, ethical, and forensic analyses reveal a deeply interconnected risk landscape surrounding AI-generated misinformation in Malaysia. The model simulation demonstrates that increasing the proportion of AI-generated content in websites from 30% to 50% compresses the time for deception to permeate the entire internet-using population from roughly four years to about two and a half years. This acceleration sharply contrasts with the slower pace of Islamic legal and legislative processes, where *fatwa* development, gazettement, and harmonization between federal and state level religious authorities already face structural and procedural constraints. Within such a compressed timeframe and no intervention, the *ummah* is likely to experience repeated exposure to deception that erodes trust in institutions, weakens adherence to *tabayyun*, and normalizes ethically impermissible behaviours such as lying, fraud, and defamation.

At the same time, forensic institutions are confronted with an evidence environment where attribution and traceability of digital content become increasingly difficult, enabling both sophisticated crimes and the "deepfake defence" that undermines the credibility of genuine evidence (Alias et al., 2024). From a Shariah perspective, these dynamics threaten the credibility of religious authorities, the integrity of *Halal* and *fatwa* ecosystems, and the authenticity of transmitted Islamic knowledge. This can lead to the possibility of both spiritual apathy and the emergence of AI-mediated pseudo-authorities that reshape religious life without due scholarly oversight.

The author hopes that this model will instil a sense of urgency among Islamic and Muslim scholars, motivating them to take action in service of the *ummah*. The results and discussion expose a temporal mismatch between the rapid pace of AI-driven deception and the response capacities of legal, ethical, and religious institutions, thereby underscoring the urgent need for anticipatory, cross-sectoral interventions into Shariah governance, digital ethics, and forensic readiness.

Conclusion

This research employed a system dynamics methodology to simulate the propagation of AI-generated misinformation among Malaysian internet users, while analyzing the simulation outcomes through Shariah, ethical, and forensic frameworks. The model reveals that a modest rise in the share of AI-generated content from 30% to 50% markedly hastens the penetration of misinformation across the entire online population, compressing the interval for institutional countermeasures from roughly four years to approximately two and a half years. This constricted timeframe presents formidable obstacles for Malaysia's Islamic juridical and regulatory mechanisms, which must traverse multifaceted institutions including the Department of Islamic Development Malaysia, the National *Fatwa* Committee, State Mufti Departments, and legislative authorities, prior to the full operationalization and enforcement of new regulations.

From an ethical perspective, the findings highlight the risk that repeated exposure to AI-driven deception will erode public trust, discourage civic and communal participation, and dull moral sensitivity, especially where *tabayyun* is neglected and deceptive practices are normalized (Sitiris & Busari, 2024). Forensic analysis further reveals that pervasive AI-generated content threatens the reliability of digital evidence, complicates attribution, and enables both sophisticated cybercrimes and defensive strategies that cast doubt on authentic records. From a Shariah standpoint, the proliferation of AI-generated misinformation jeopardizes the credibility of religious authorities, undermines the traceability of *Halal* and *fatwa*-related information, and opens the door to AI mediated distortions of religious texts and guidance, potentially contributing to spiritual drift or the rise of pseudo-religious identities anchored in synthetic content.

Although this research provides a preliminary and conservative model within a specific national context, its methodology lends itself to broader application in other Muslim-majority societies and to refinement using more detailed behavioral and technical data. Fundamentally, the study highlights that countermeasures against AI-generated misinformation extend beyond technical measures, demanding an integrated approach encompassing legal, ethical, and religious dimensions to preserve societal cohesion and the spiritual welfare of the *ummah*.

Supplementary Data

The System Dynamics model file used in this study is accessible via the Open Science Framework (OSF) project. It can be retrieved from the following link: https://osf.io/hm9ud/overview.

References

Abidin, N. H. Z., & Nasohah, Z. (2023). The complexities of implementing gazetted fatwas in Malaysia's Federal Territories. *Mazahib, 22*(2), 485–510. https://doi.org/10.21093/mj.v22i2.5862

Alfiani, F. R. N., & Santiago, F. (2024). A comparative analysis of artificial intelligence regulatory law in Asia, Europe, and America. In *Proceeding of 1st International Graduate Conference on Digital Policy and Governance Sustainability (DiGeS-Grace 2024)*, (1-10). https://doi.org/10.1051/shsconf/202420407006

Alias, M. A. A., Wan Ismail, W. A. F., Baharuddin, A. S., & Syah Mallow, M. (2024). Wasa'il ithbat dalam undang-undang keterangan Islam: Analisis perundangan terhadap keabsahan dokumen

elektronik di Mahkamah Syariah Malaysia: Means of proof in Islamic law of evidence: A legal analysis of the admissibility of electronic documents in Malaysian Syariah Courts. *Malaysian Journal of Syariah and Law*, *12*(3), 689-700. https://doi.org/10.33102/mjsl.vol12no3.792

Al-Janabi, M. M. (2024). Artificial intelligence in Qur'anic exegesis: A critical analytical study of ChatGPT technology. *International Journal of Quranic Research*, *16*(2), 112–130. https://doi.org/10.22452/quranica.vol16no2.5

Azar, A. T. (2012). System dynamics as a useful technique for complex systems. *International Journal of Industrial and Systems Engineering*, *10*(4), 377–401. https://doi.org/10.1504/IJISE.2012.046298

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, *3*(3), 193–209. https://doi.org/10.1207/s15327957pspr0303_3

Bengio, Y., Mindermann, S., Privitera, D., Besiroglu, T., Bommasani, R., Casper, S., Choi, Y., Goldfarb, D. A., Heidari, H., Khalatbari, L., Longpre, S., Mavroudis, V., Mazeika, M., Ng, T., Okolo, C. T., Raji, D., Skeadas, T., Tramèr, F., Adekanmbi, B., … Zhang, Y.-Q. (2024). *International scientific report on the safety of advanced AI (Interim report)*. arXiv. https://doi.org/10.48550/arXiv.2412.05282

Brenner, D. A., Chaudhary, Y., Geraci, R. M., Graves, M., Griese, H., Kruger, E., Ryan, Y., & Schwarting, M. (2024). Technical and religious perspectives on AI misinformation and disinformation. *SSRN Electronic Journal*, 1-26. https://doi.org/10.2139/ssrn.4360413

Casu, M., Guarnera, L., Caponnetto, P., & Battiato, S. (2023). *GenAI mirage: The impostor bias and the deepfake detection challenge in the era of artificial illusions*. arXiv. https://doi.org/10.48550/arXiv.2312.16220

Cerabona, T. (2023). *Study of the application of a decision support approach based on classical physics principles for managing supply chain risks and opportunities within its immersive performance management cockpit* (Doctoral dissertation). HAL. https://theses.hal.science/tel-04368478

Chen, J., Chou, S., Yu, T. H., Rizqi, Z. U., & Hang, D. T. (2022). System dynamics analysis on the effectiveness of vaccination and social mobilization policies for COVID-19 in the United States. *PLOS ONE*, *17*(8), e0268443. https://doi.org/10.1371/journal.pone.0268443

Chua, L., Castillo, B. D., Alkhzaimi, H., Ramaputra, M., Dita, A., & Thong, J. L. K. (2023). A whitepaper on demystifying the perils and promises of AI in Southeast Asia: Cluster 1 of 3—Indonesia, Philippines, and Singapore. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4645031

Dahlan, H. A. (2018). Future interaction between man and robots from Islamic perspective. *International Journal of Islamic Thought*, *13*(1), 1–10. https://doi.org/10.24035/ijit.06.2018.005

Dahlan, H. A. (2022). A plausible future for high scientific literacy in Muslim societies. *UMRAN – International Journal of Islamic and Civilizational Studies*, *9*(1), 67–79.

Department of Statistics Malaysia. (2023). *"ICT use and access by individuals and households survey report"*. https://storage.dosm.gov.my/icths/icths_2023.pdf

Drolsbach, C., & Pröllochs, N. (2025). *Characterizing AI-generated misinformation on social media.* arXiv. https://doi.org/10.48550/arXiv.2505.10266

Fisunoğlu, A. (2018). System dynamics modeling in international relations. *All Azimuth: A Journal of Foreign Policy and Peace*, *7*(2), 1–15. https://doi.org/10.20991/allazimuth.476884

Forliano, C., Bernardi, P. D., Rózsa, Z., & Bertello, A. (2024). Systems dynamics research in management and organization studies: Overview and research agenda. *Journal of Innovation & Knowledge*, *9*(3), 100512. https://doi.org/10.1016/j.jik.2024.100512

Gabriel, I., Manzini, A., Keeling, G., Hendricks, L. A., Rieser, V., Iqbal, H., Tomašev, N., Ktena, S. I., Kenton, Z., Rodríguez, M. B., El-Sayed, S., Brown, S., Akbulut, C., Trask, A., Hughes, E., Bergman, A. S., Shelby, R., Marchal, N., Griffin, C., … Manyika, J. (2024). *The ethics of advanced AI assistants.* arXiv. https://doi.org/10.48550/arXiv.2404.16244

Gorian, E., & Osman, N. D. (2024). Digital ethics of artificial intelligence (AI) in Saudi Arabia and United Arab Emirates. *Malaysian Journal of Syariah and Law*, *12*(3), 583-597. https://doi.org/10.33102/mjsl.vol12no3.798

Hagendorff, T. (2024). Deception abilities emerged in large language models. In *Proceedings of the National Academy of Sciences*, (p. e2317967121). https://doi.org/10.1073/pnas.2317967121

Harrison, G., Gühnemann, A., Jittrapirom, P., Pfaffenbichler, P., Shepherd, S., & Vilchez, J. J. G. (2025). In the loop: The application of system dynamics in transport. In *Lecture notes in mobility* (pp. 723–731). Springer. https://doi.org/10.1007/978-3-031-85578-8_98

Hossain, M. B., Miraz, M. H., & Ya'u, A. (2024). From legality to responsibility: Charting the course for AI regulation in Malaysia. *IIUM Law Journal*, *32*(1), 397–430. https://doi.org/10.31436/iiumlj.v32i1.927

Ilyas, H., Fatmal, A. B., & Ahmad, L. O. I. (2025). Digital jihad in Qur'anic perspective: An Islamic response to the challenges of cyberspace in the age of artificial intelligence (AI). *QOF*, *9*(2), 189–210. https://doi.org/10.30762/qof.v9i2.3091

Khairuldin, W. M. K. F. W., Anas, W. N. I. W. N., Embong, A. H., İsmail, D., & Hanapi, M. S. (2019). The challenges of fatwa institutions in Malaysia in facing the progress of science and technology. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3453554

Klasen, L. M., Fock, N., & Forchheimer, R. (2024). The invisible evidence: Digital forensics as key to solving crimes in the digital age. *Forensic Science International*, *362*, 112133. https://doi.org/10.1016/j.forsciint.2024.112133

Koswara, A., & Herlina, L. (2025). The ethical deviations of AI in marketing practices: A critical review from halal perspectives. *Research of Islamic Economics*, *2*(2), 105–118. https://doi.org/10.58777/rie.v2i2.393

Lee, J., Jeon, S., Park, Y., Chung, J., & Jeong, D. (2023). A forensic methodology for detecting image manipulations. *arXiv*. https://doi.org/10.48550/arXiv.2308.04723

Li, J., Qu, L., Cai, T., Zhao, Z., Haldar, N. A. H., Krishna, A., Kong, X., Macau, F. R., Chakraborty, T., Deroy, A., Lin, B., Blackmore, K., Noman, N., Cheng, J., Cui, N., & Xu, J. (2025). AI-generated content in cross-domain applications: Research trends, challenges, and propositions. *arXiv*. https://doi.org/10.48550/arXiv.2509.11151

Linna, D. W., Dalal, A., Gao, C., Grimm, P., Grossman, M. R., Pulice, C., Subrahmanian, V. S., & Tunheim, J. (2024). Deepfakes in court: How judges can proactively manage alleged AI-generated material in national security cases. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4943841

Luo, A., & Bussey, K. (2023). Moral disengagement in youth: A meta-analytic review. *Developmental Review*, *70*, 101101. https://doi.org/10.1016/j.dr.2023.101101

Maha, A. T., Habtiter, Z. A., Ahmed, H., & Mijbel, Y. K. (2025). Employing artificial intelligence applications in the service of Islamic religion and belief. In *Lecture Notes in Networks and Systems* (pp. 167–176). Springer. https://doi.org/10.1007/978-3-031-82377-0_15

Malik, S. A. (2023). Artificial intelligence and Islamic thought: Two distinctive challenges. *Journal of Islamic and Muslim Studies*, *8*(2), 108–124. https://doi.org/10.2979/jims.00020

Masruha, M., Munawar, S., Zulmi, F., Abdullah, A., & Husaini, H. (2025). Digitization of fatwas and religious authority: A study on the role of social media in the interpretation of Islamic law. *Journal of Mujaddid Nusantara*, *2*(2), 99–115. https://doi.org/10.62568/jomn.v2i2.421

Mohamad, M. N., & Mamat, Z. (2018). Penerimaan dan penyelarasan fatwa Muzakarah Jawatankuasa Fatwa Majlis Kebangsaan bagi Hal Ehwal Ugama Islam Malaysia oleh negeri-negeri. *Journal of Fatwa Management and Research*, *4*(1), 33–50. https://doi.org/10.33102/jfatwa.vol4no1.93

Nawi, A., Khamis, N. Y., Yaakob, M. F. M., Samuri, M. A. A., & Zakaria, G. A. N. (2023). Exploring opportunities and risks of artificial intelligence research for Islamic ethical guidelines. *Jurnal Akidah & Pemikiran Islam*, *25*(2), 1–25. https://doi.org/10.22452/afkar.vol25no2.1

Niam, M. F. (2024). Does artificial intelligence go beyond the limits of religious authority? An ethical review on IslamGPT. *Al-ʿAdālah*, *27*(1), 71–90. https://doi.org/10.35719/aladalah.v27i1.477

Nilgiriwala, K., Mahajan, U., Ahmad, R. A., Castro, R. de, Lazo, L., Kong, J. D., Lee, A., Veerakumarasivam, A., Sharef, N. M., & Demidenko, S. (2024). Navigating the governance of artificial intelligence (AI) in Asian nations: A focus on India, Indonesia, Malaysia and the Philippines. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4735279

Putra, B. A. (2024). Fake news and disinformation in Southeast Asia: How should ASEAN respond? *Frontiers in Communication*, *9*, 1380944. https://doi.org/10.3389/fcomm.2024.1380944

Rashid, M. F. A., & Dawood, S. R. S. (2024). Statistical evaluation of webometric analysis of tourism websites in ASEAN countries. *International Journal of Research and Innovation in Social Science*, *8*(2), 389–402. https://doi.org/10.47772/ijriss.2024.802171

Rivera, E., Seira, E., & Jha, S. (2024). Democracy corrupted: Apex corruption and the erosion of democratic values. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4828243

Saifuddin, S., Mat Salleh, M. K., Ali Tajuddin, H. H., & Yahya, M. A. (2024). Importance of gazettement of fatwa in the enforcement of Syariah criminal law in Malaysia: An analysis. *Al-Qanatir: International Journal of Islamic Studies*, *33*(4), 20–31. https://al-qanatir.com/aq/article/view/935

Sandoval, M.-P., Vau, M. de A., Solaas, J., & Rodrigues, L. F. D. (2024). Threat of deepfakes to the criminal justice system: A systematic review. *Crime Science*, *13*(1), 39.

Shoaib, M. R., Wang, Z., Ahvanooey, M. T., & Zhao, J. (2023). Deepfakes, misinformation, and disinformation in the era of frontier AI, generative AI, and large AI models. *arXiv*. https://doi.org/10.48550/arXiv.2311.17394

Sitiris, M., & Busari, S. A. (2024). The legal capacity (al-ahliyyah) of artificial intelligence from an Islamic jurisprudential perspective. *Malaysian Journal of Syariah and Law*, *12*(1), 31-42.

Sonni, A. F. (2025). AI-based disinformation and hate speech amplification: Analysis of Indonesia's digital media ecosystem. *Frontiers in Communication*, *10*, 1603534. https://doi.org/10.3389/fcomm.2025.1603534

Spennemann, D. H. (2025). *Delving into: The quantification of AI-generated content on the internet (synthetic data)*. arXiv. https://doi.org/10.48550/arXiv.2504.08755

Sterman, J. D. (2002). *System dynamics: Systems thinking and modeling for a complex world*. MIT OpenCourseWare. http://hdl.handle.net/1721.1/102741

Storm-Mathisen, F. (2024). 'Violence is completely normal': Managing violence through narrative normalization. *The British Journal of Criminology*, *65*(1), 37–55.

Taha, J. Z. A., Sulaiman, S., & Jima'ain, M. T. A. (2025). The use of artificial intelligence (ChatGPT) in Islamic education: An analysis of challenges, opportunities, and ethical considerations based on maqāṣid al-sharīʿah. *Journal of Research Innovation and Strategies for Education (RISE)*, *2*(3), 35–50. https://doi.org/10.70148/rise.v2i3.3

Tarsney, C. (2025). Deception and manipulation in generative AI. *Philosophical Studies*. Advance online publication. https://doi.org/10.1007/s11098-024-02259-8

Tortora, L. (2024). Beyond discrimination: Generative AI applications and ethical challenges in forensic psychiatry. *Frontiers in Psychiatry*, *15*, 1346059. https://doi.org/10.3389/fpsyt.2024.1346059

Treleaven, P., Barnett, J., Brown, D. W., Bud, A., Fenoglio, E., Kerrigan, C., Koshiyama, A., Sfeir-Tait, S., & Schoernig, M. (2023). The future of cybercrime: AI and emerging technologies are creating a cybercrime tsunami. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4507244

Wahab, A. (2025). Futures of deepfake and society: Myths, metaphors, and future implications for a trustworthy digital future. *Futures*, *173*, 103672. https://doi.org/10.1016/j.futures.2025.103672

Walker, C., Schiff, D., & Schiff, K. J. (2024). Merging AI incidents research with political misinformation research: Introducing the political deepfakes incidents database. In *Proceedings of the AAAI Conference on Artificial Intelligence*, (pp. 23053–23060).

Wan Ismail, W. A. F., Abdul Mutalib, L., Mamat, Z., Hashim, H., Baharuddin, A. S., Mohammed Hasan, B. M., & Alias, M. A. A. (2024). Analisis terhadap konsep penerimaan dan pengesahan E-Kitabah sebagai kaedah pembuktian menurut perundangan Islam di Malaysia. *LexForensica: Forensic Justice And Socio-Legal Research Journal*, *1*(1), 1-12.

Wu, Y., Pan, X., Hong, G., & Yang, M. (2025). *OpenDeception: Benchmarking and investigating AI deceptive behaviors via open-ended interaction simulation*. arXiv. https://doi.org/10.48550/arXiv.2504.13707

Yu, X., Wang, Y., Chen, Y., Tao, Z., Xi, D., Song, S., Niu, S., & Li, Z. (2024). *Fake artificial intelligence generated contents (FAIGC): A survey of theories, detection methods, and opportunities*. arXiv. https://doi.org/10.48550/arXiv.2405.00711

Zafar, M. B., & Ali, H. (2025). Shariah governance standard on generative AI for Islamic financial institutions. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5143165

Zhang, Y., Pang, Z., Huang, S., Wang, C., & Zhou, X. (2025). Unmasking AI-created visual content: A review of generated images and deepfake detection technologies. *Journal of King Saud University – Computer and Information Sciences*, *37*(6), 101544. https://doi.org/10.1007/s44443-025-00154-8