

TOWARDS RESPONSIBLE ARTIFICIAL INTELLIGENCE (AI) GOVERNANCE IN THE MALAYSIAN CRIMINAL JUSTICE SYSTEM: A LEGAL AND ETHICAL APPRAISAL OF PREDICTIVE POLICING AND AI-DRIVEN SURVEILLANCE

^{i,*}Mohamad Aniq Aiman Alias, ^{i,iii}Wan Abdul Fattah Wan Ismail, ^{i,ii,iii,iv}Ahmad Syukran Baharuddin,
^{i,ii}Hasnizam Hashim, ⁱFareed Mohd Hassan & ^vTuan Muhammad Faris Hamzi Tuan Ibrahim

ⁱFaculty of Syariah and Law, Universiti Sains Islam Malaysia, 71800 Nilai, Negeri Sembilan, Malaysia

ⁱⁱCentre of Research for Fiqh Forensics and Judiciary (CFORSJ), Faculty of Syariah and Law, Universiti Sains Islam Malaysia, 71800 Nilai, Negeri Sembilan, Malaysia

ⁱⁱⁱMaqasid Institute, United States of America (USA)

^{iv}International Advisory Board Member, Mejallat Al-Afaq wa Al-Maarif (JSKP), Faculty of Humanities, Islamic Sciences and Civilization, University Amar Telidji of Laghouat, BP 37G, Road of Ghardaia, 03000 Laghouat, Algeria

^vIslamic Civilization Academy, Faculty of Social Science and Humanities, Universiti Teknologi Malaysia (UTM), 81310, Johor Bahru, Johor, Malaysia

*(Corresponding author) e-mail: aniqalias@usim.edu.my

Article history:

Submission date: 15 January 2026

Received in revised form: 27 March 2026

Acceptance date: 17 June 2026

Available online: 31 July 2026

Keywords:

Artificial Intelligence (AI), predictive policing, surveillance, criminal justice, responsible AI governance

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:

Alias, M. A. A., Wan Ismail, W. A. F., Baharuddin, A. S., Hashim, H., Mohd Hassan, F., & Tuan Ibrahim, T. M. F. H. (2026). Towards responsible Artificial Intelligence (AI) governance in the Malaysian criminal justice system: A legal and ethical appraisal of predictive policing and AI-driven surveillance. *LexForensica: Forensic Justice and Socio-Legal Research Journal*, 3(1), 78–87. <https://doi.org/10.33102/5yknbc04>



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact pennerbit@usim.edu.my.

SDG Elements:

Peace, Justice and Strong Institutions



ABSTRACT

Artificial Intelligence (AI) is increasingly transforming criminal justice systems through technologies such as predictive policing, predictive crime mapping, facial recognition, and AI-driven surveillance to enhance operational efficiency and support evidence-based decision-making. However, the rapid expansion of these technologies raises critical legal and ethical concerns regarding algorithmic bias, privacy infringement, transparency, accountability, and the protection of fundamental human rights. In Malaysia, although digital transformation policies have progressed significantly, scholarly discussion concerning the governance of predictive policing and AI-driven surveillance remains relatively limited from a legal and regulatory perspective. Accordingly, this study examines the concepts of predictive policing and AI-driven surveillance, evaluates their operational benefits alongside associated legal and ethical challenges, and proposes practical recommendations for strengthening responsible AI governance within the Malaysian criminal justice system. Adopting a qualitative research design, the study utilizes document analysis for data collection. The data collected were then analysed and presented according to key subthemes. The findings indicate that AI possesses significant potential to enhance crime prevention, optimize resource allocation, improve investigative efficiency, and support law enforcement decision-making. Nevertheless, these operational advantages are accompanied by substantial legal and ethical risks, particularly concerning algorithmic bias, excessive surveillance, data privacy, and compromised accountability. To mitigate these risks, this study suggests that future implementation of predictive policing technologies in Malaysia should be guided by a comprehensive Responsible AI Governance Framework grounded in human oversight, transparency, accountability, fairness, and privacy protection. Overall, this research contributes to the scholarship on AI governance by providing an integrated legal appraisal and context-specific governance recommendations to inform future regulatory development and ethical AI implementation in Malaysia and other developing jurisdictions.

Introduction

Artificial Intelligence (AI) has emerged as one of the most transformative technologies of the twenty-first century, significantly influencing sectors such as healthcare, finance, education, public administration, and law enforcement (Alias et al., 2025; Dwivedi et al., 2021). Within the criminal justice system, AI has evolved from a data-processing tool to an important decision-support technology capable of analysing large volumes of information, identifying crime patterns, and assisting law enforcement agencies in crime prevention, criminal investigation, and resource allocation (Oswald et al., 2017). Technologies such as predictive crime mapping, facial recognition, biometric identification, and intelligent surveillance systems have increasingly been adopted in many jurisdictions to enhance policing efficiency and improve public safety (Perry et al., 2013; Klare et al., 2012; Jain et al., 2004; Valera & Velastin, 2005). Consequently, the integration of AI into criminal justice has generated considerable interest among policymakers, legal scholars, and law enforcement agencies due to its potential to improve operational effectiveness and transform traditional approaches to policing and criminal justice administration.

Despite these potential benefits, the growing adoption of AI within criminal justice has simultaneously raised significant legal and ethical concerns. Existing studies have highlighted various issues related to algorithmic bias, lack of transparency and accountability, privacy infringement, and potential violations of fundamental human rights arising from the deployment of predictive policing and AI-powered surveillance technologies (Richardson, Schultz, & Crawford, 2019; Oswald et al., 2018). Furthermore, the reliability of AI-driven decision-making remains heavily dependent upon the quality, representativeness, and neutrality of the datasets used for algorithm development (Barocas & Selbst, 2016; Mehrabi et al., 2021). Consequently, while AI has demonstrated its capability to enhance policing efficiency, it also presents substantial risks that may undermine public trust and the principles of fairness and justice if implemented without appropriate legal safeguards and human oversight. These concerns have prompted many jurisdictions, particularly within the European Union and the United Kingdom, to strengthen regulatory and governance frameworks governing the deployment of AI in law enforcement.

In the Malaysian context, the adoption of AI within the criminal justice ecosystem is gradually evolving alongside broader national digital transformation initiatives. Current applications primarily involve surveillance technologies, facial recognition, crime analytics and other digital innovations introduced by enforcement agencies to enhance operational efficiency and public service delivery (Ministry of Digital, 2024; National AI Office (NAIO), 2025). Although Malaysia has introduced several national initiatives relating to AI governance, including the establishment of the National AI Office (NAIO) and the National Guidelines on Artificial Intelligence Governance and Ethics (AIGE), existing legal and policy discussions have largely focused on responsible AI governance, personal data protection and digital transformation (Ministry of Digital, 2024; NAIO, 2025). Comparatively, limited attention has been devoted to examining the legal and ethical implications of predictive policing and AI-driven surveillance within the Malaysian criminal justice system. This indicates an important research gap, particularly in identifying an appropriate governance approach that balances technological innovation with the protection of fundamental rights and the rule of law.

Accordingly, this article aims to examine the concepts of predictive policing and AI-driven surveillance within the context of criminal justice, analyse the current applications of AI in crime prevention and law enforcement, and evaluate the legal and ethical issues arising from their implementation, particularly in relation to algorithmic bias, privacy, accountability and the protection of fundamental rights. Building upon these discussions, the article proposes several recommendations for developing a responsible AI governance framework to guide the future implementation of AI-enabled policing within the Malaysian criminal justice system. It is anticipated that this study will contribute to the growing body of scholarship on AI and criminal justice while providing valuable insights for policymakers, legal practitioners and enforcement agencies in formulating a balanced legal and ethical framework for the responsible adoption of AI technologies.

The Concept of Predictive Policing and Surveillance in an Artificial Intelligence-Driven Criminal Justice System

The advancement of AI has transformed the criminal justice system from a model that primarily relies on the experience and intuition of law enforcement officers to one that is increasingly data-driven. Within this context, AI does not function as a substitute for police officers or other competent authorities; rather, it serves as a decision-support system that assists in analysing large volumes of data to support crime prevention, criminal investigation, and law enforcement activities. By processing data automatically, AI enables crime patterns to be identified more rapidly than conventional approaches, thereby facilitating more systematic and evidence-based decision-making. Nevertheless, AI does not possess the capacity to exercise legal or moral judgement autonomously, as final decisions continue to require human assessment, professional expertise and discretionary authority (Mandalapu et al., 2023).

One of the principal applications of AI in the criminal justice system is predictive policing, which uses machine learning algorithms and data analytics to estimate the likelihood of future criminal activity based on historical crime data. This approach analyses multiple variables, including crime location, time of occurrence, offence type, crime patterns and environmental factors, to identify geographical areas or situations that are potentially at a higher risk of criminal activity (see Figure 1).

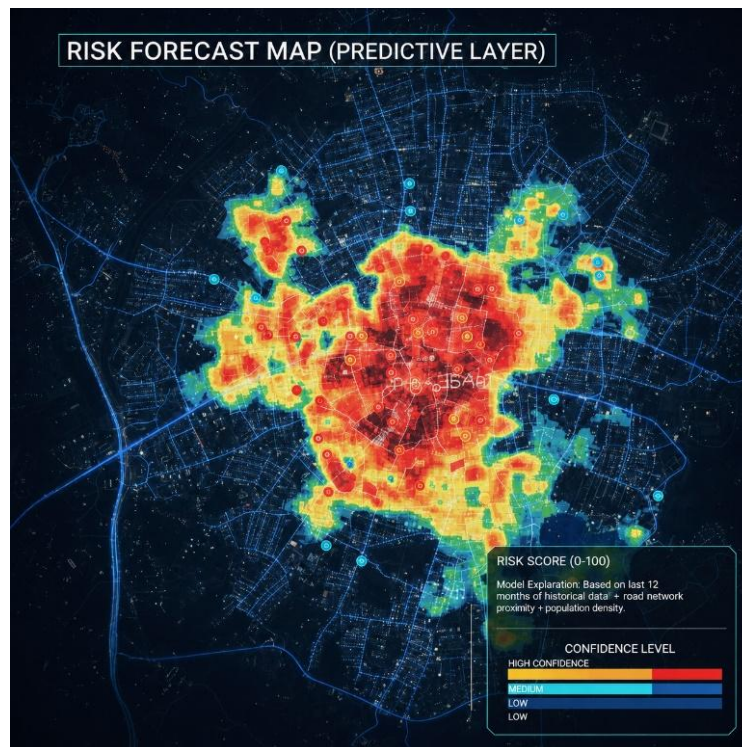


Figure 1. Illustration of an AI-Based Predictive Crime Risk Map Showing Crime Risk Forecasting Based on Historical Crime Data to Support Hotspot Identification, Patrol Planning and Resource Allocation

As illustrated in Figure 1, predictive policing systems generate predictive crime risk maps by applying AI techniques to historical crime data. Areas highlighted in red indicate locations with a higher predicted crime risk, whereas yellow and green areas represent moderate and lower-risk locations, respectively. In addition, the system provides risk scores and confidence levels to assist law enforcement agencies in prioritising patrol deployment and resource allocation more effectively. Importantly, these predictions are probabilistic assessments of crime risk, not definitive determinations that a crime will occur or that a particular individual will commit a criminal offence.

Perry et al., (2013) explain that predictive policing is not intended to predict future crimes with certainty; instead, it generates probabilistic predictions that enable law enforcement agencies to plan patrol strategies and allocate operational resources more efficiently. These findings are further supported by Mohler et al., (2015), who demonstrated that location-based predictive models improved the effectiveness of police patrol deployment across several case study areas in the United States. Accordingly, predictive

policing should be understood as a strategic decision-support mechanism, rather than a system capable of determining that a particular individual will inevitably commit a crime.

Beyond predictive policing, AI has also been extensively applied in surveillance, which refers to the continuous monitoring of activities through digital technologies to detect, identify, and monitor behaviour that may pose threats to public safety. Unlike conventional surveillance methods that rely largely on human observation, AI enables automatic surveillance through technologies such as smart CCTV, facial recognition, biometric analysis, object detection, and real-time monitoring. According to Almeida et al., (2022), advances in AI have expanded the role of surveillance from merely recording evidence to analysing visual information, identifying individuals, and automatically detecting suspicious behaviour. Consequently, AI-driven surveillance has become an important component of modern criminal justice systems, not only by supporting criminal investigations but also by facilitating the early detection of potential security threats.

Table 1. Representative AI Technologies Employed in Surveillance for Criminal Justice, Illustrating Their Primary Functions and Applications in Law Enforcement

AI Surveillance Technology	Primary Function	Criminal Justice Application
Smart CCTV & Video Analytics	Detects suspicious activities and monitors environments in real time	Crime prevention and public safety monitoring
Facial Recognition	Identifies and verifies individuals	Suspect identification and identity verification
Biometric Identification	Recognises fingerprints, iris patterns, facial features or voice	Forensic investigation and identity authentication
Automatic Number Plate Recognition (ANPR)	Detects and recognises vehicle registration numbers	Vehicle tracking and law enforcement
Object Detection	Detects weapons, abandoned objects and prohibited items	Threat detection and incident response
Behaviour Analysis	Analyses abnormal or suspicious human behaviour	Early threat detection and crime prevention

Although predictive policing and surveillance are often discussed as separate concepts, in practice, they are complementary components of an AI-driven criminal justice ecosystem. Predictive policing produces analytical forecasts that identify locations or situations where crime is more likely to occur. In contrast, surveillance provides continuous monitoring and information gathering to verify incidents, detect suspicious activities and support law enforcement responses. In other words, predictive policing addresses the questions of where and when crime is likely to occur, while surveillance helps determine what is occurring and who is involved. The integration of these technologies enables law enforcement agencies to adopt a more proactive approach to crime prevention while enhancing investigative effectiveness through the utilisation of more comprehensive, data-driven information (Mandalapu et al., 2023; Almeida et al., 2022).

The foregoing discussion demonstrates that AI serves not only as a sophisticated data analytics tool but also as an enabling technology that strengthens crime prevention, surveillance and criminal investigation capabilities. Nevertheless, understanding the conceptual foundations of predictive policing and surveillance alone is insufficient without examining how these technologies are translated into practical law enforcement applications. Accordingly, the following section examines the practical applications of AI in predictive crime mapping, AI-driven policing models, facial recognition, and biometric technologies within contemporary criminal justice systems.

Applications of AI in Predictive Policing and Surveillance

The advancement of AI has significantly expanded the use of technology within criminal justice systems, particularly through the development of predictive crime mapping, which involves identifying geographic areas with a higher likelihood of criminal activity based on historical crime data and patterns. Unlike traditional policing approaches that rely heavily on the experience and judgement of police officers, predictive crime mapping employs machine learning algorithms to analyse multiple variables, including crime location, time of occurrence, offence type, demographic characteristics, and environmental factors, to generate crime hotspot maps that identify high-risk areas. According to Mandalapu et al., (2023),

predictive crime mapping represents one of the fastest-growing areas of AI research in modern policing, as it enables law enforcement agencies to identify high-risk locations and optimise patrol deployment based on objective data analysis. Nevertheless, the outputs generated by these systems represent probabilistic predictions rather than definitive certainty that a crime will occur.

Beyond predictive crime mapping, AI has also been incorporated into AI-driven policing models that support decision-making within policing operations. These models integrate multiple sources of information, including crime records, police reports, geospatial data, social information and environmental factors, to assist law enforcement agencies in determining operational priorities, allocating personnel and developing crime prevention strategies. Perry et al., (2013) explain that the primary objective of AI-based policing models is not to replace police officers but to enhance the quality and effectiveness of tactical decision-making through predictive analytics. This approach was subsequently validated by the field study conducted by Mohler et al., (2015), which demonstrated that location-based predictive models significantly improved the effectiveness of police patrol deployment in Los Angeles and Kent without requiring a substantial increase in personnel. These findings suggest that AI serves as a resource optimisation mechanism, particularly when law enforcement agencies face constraints on manpower and operational resources.

In the context of surveillance, AI has transformed monitoring practices from processes that rely primarily on human observation to intelligent surveillance systems that operate automatically. Technologies such as facial recognition, biometric analysis, object detection, and real-time video analytics enable systems to identify individuals, detect suspicious activities, and support criminal investigations more efficiently. According to Almeida, Shmarko, and Lomas (2022), advances in AI have transformed closed-circuit television (CCTV) systems from passive recording devices into intelligent analytical tools capable of analysing visual information, identifying individuals, and automatically detecting suspicious behaviour to support law enforcement operations. Within the Malaysian context, AI technologies such as facial recognition, crime data analytics, and smart surveillance have gradually been introduced through various public safety digitalisation and law enforcement initiatives. At present, however, their implementation remains primarily focused on surveillance, detection, and investigative support, while the broader adoption of predictive policing is still in its early stages.

These three AI applications should not be viewed as independent technologies but rather as complementary components within an integrated digital ecosystem that supports law enforcement operations. Predictive crime mapping identifies geographical areas that are potentially at higher risk of criminal activity; AI-driven policing assists in developing operational strategies based on analytical outputs; and AI-driven surveillance, through facial recognition and biometric technologies, provides continuous monitoring, identity verification, and evidence collection to support law enforcement actions. The integration of these technologies enables crime prevention, surveillance and criminal investigations to be conducted in a more systematic, efficient and data-driven manner. Nevertheless, the effectiveness of each application remains dependent upon algorithmic accuracy, data quality and the extent of meaningful human oversight throughout the decision-making process.

The foregoing discussion demonstrates that AI applications have introduced a more proactive and data-driven approach to criminal justice by strengthening the operational capabilities of law enforcement agencies. Nevertheless, the value of these technologies should be assessed not solely based on their technical capabilities but also on their contribution to improving crime prevention, criminal investigation, and law enforcement effectiveness. Accordingly, the following section examines the benefits of AI within the criminal justice system, as well as its potential to enhance the efficiency and effectiveness of law enforcement operations.

Balancing the Benefits and Legal-Ethical Challenges of Artificial Intelligence in Criminal Justice

The advancement of AI has transformed law enforcement from a predominantly reactive model into a more proactive, data-driven, and crime-prevention-oriented approach. Unlike traditional policing methods, which rely heavily on the experience and professional judgement of law enforcement officers, AI enables the automated processing of large volumes of data to identify crime patterns, optimise resource allocation and support decision-making. Meijer and Wessels (2019) argue that the application of AI in predictive policing has the potential to enhance operational efficiency by facilitating crime location

forecasting and more strategic patrol deployment. Nevertheless, they also emphasise that these benefits require continued empirical evaluation, as existing evidence regarding the effectiveness of predictive policing in reducing crime rates remains inconclusive. Accordingly, the assessment of AI should not be confined solely to its technological capabilities. However, it should also consider the extent to which it contributes to justice, operational effectiveness and the overall integrity of the criminal justice system.

From an operational perspective, AI offers several significant advantages to law enforcement agencies. The use of predictive crime mapping enables the identification of high-risk areas at an early stage, thereby assisting police authorities in planning patrol strategies and allocating personnel more efficiently. In addition, AI accelerates criminal investigations through the automated analysis of CCTV footage, digital images, biometric data and information obtained from multiple databases. These capabilities not only reduce investigators' workload but also improve the efficiency of suspect identification, identity verification and the linkage of digital evidence to criminal cases. In this context, AI serves as a decision-support system, assisting law enforcement officers in making informed decisions based on data analysis without replacing their professional judgement (Mandalapu et al., 2023). However, these benefits remain dependent upon the accuracy of AI algorithms and the quality of the data used for model training. Where training datasets are imbalanced or reflect historical patterns of biased law enforcement practices, AI systems may produce similarly biased outcomes (algorithmic bias), thereby undermining the principles of fairness and equality before the law.

In the context of surveillance, AI-enabled technologies such as facial recognition and biometric analysis have significantly enhanced real-time crime detection and monitoring capabilities. These technologies facilitate faster suspect identification, the detection of wanted individuals, and the systematic collection and analysis of digital evidence. Almeida, Shmarko and Lomas (2022) observe that advances in AI have transformed closed-circuit television (CCTV) systems from passive recording devices into intelligent analytical tools capable of identifying individuals, detecting suspicious behaviour and generating actionable information to support law enforcement operations. However, these technological developments also raise significant concerns relating to privacy and fundamental human rights. Smith and Miller (2022) contend that although biometric technologies possess substantial potential to enhance criminal investigations, their deployment without adequate legal safeguards may result in mass surveillance, the misuse of biometric data and the erosion of civil liberties. Consequently, the public safety benefits offered by AI cannot be separated from the state's obligation to protect individual privacy rights and to ensure that all forms of data processing are conducted lawfully, necessarily and proportionately.

Beyond privacy concerns, the principal challenges associated with AI in the criminal justice system also encompass transparency, explainability and accountability. Most AI algorithms operate as highly complex systems whose underlying processes are often difficult for users and affected individuals to understand. This raises important questions regarding how algorithmic decisions can be explained, reviewed or challenged when they have legal consequences for individuals. Gstrein, Bunnik and Zwitter (2019) argue that the effectiveness of predictive policing depends not only on algorithmic accuracy but also on transparency, public trust and the existence of effective review mechanisms. Similarly, Almeida et al. (2022) recommend that the deployment of AI by law enforcement agencies should be accompanied by human rights impact assessments, periodic audits, comprehensive documentation of algorithmic decision-making processes and meaningful human oversight to ensure that AI-assisted decisions remain transparent and accountable. Such an approach is consistent with the principles of Responsible AI, which emphasise that AI should function as a decision-support tool rather than replacing human authority in decisions that carry legal implications for individuals.

Taken together, these findings demonstrate that the benefits of AI in the criminal justice system cannot be separated from the legal and ethical challenges that accompany its implementation. Although AI offers considerable potential to enhance crime prevention, criminal investigation, and resource management, its effectiveness ultimately depends on the existence of governance mechanisms capable of ensuring transparency, accountability, the protection of fundamental human rights, and adherence to the rule of law. Accordingly, the central challenge is no longer whether AI should be utilised within the criminal justice system, but rather how these technologies can be implemented in a manner that is responsible, transparent and ethically sound. This consideration forms the foundation for the proposed Responsible

Artificial Intelligence Governance Framework, which seeks to support the responsible deployment of AI within the Malaysian criminal justice system.

Recommendations for Strengthening AI Governance in the Context of the Malaysian Criminal Justice System

The preceding discussion has demonstrated that the use of AI within the criminal justice system offers significant potential to enhance crime prevention, criminal investigation and law enforcement effectiveness. Nevertheless, these benefits cannot be separated from the legal and ethical challenges associated with AI deployment, including algorithmic bias, privacy infringement, system opacity and questions concerning accountability for AI-assisted decisions. Accordingly, before technologies such as predictive policing and AI-driven surveillance are implemented more extensively in the Malaysian context, greater attention should be given to developing a clear and comprehensive AI governance framework. This approach is consistent with international developments, which emphasise that the use of AI in the public sector should be guided by the principles of fairness, transparency, accountability and respect for fundamental human rights (OECD, 2019; UNESCO, 2021).

Table 2. Summary of Proposed Recommendations for Strengthening Responsible AI Governance in the Malaysian Criminal Justice System

Strategic Recommendation	Key Focus	Expected Outcome
Strengthen the Legal and Regulatory Framework	- AI-specific legal framework - AI governance guidelines - Risk-based regulation - AI decision review mechanism	A comprehensive legal framework governing AI deployment in criminal justice
Promote Human-Centred AI Governance	- Human oversight - Transparency - Algorithmic accountability - Algorithmic Impact Assessment (AIA) - Periodic AI audits	Responsible, transparent and accountable AI-assisted decision-making
Strengthen Human Rights and Data Protection	- Privacy by design - Fairness and non-discrimination - Proportionality - Personal data protection - Independent oversight	Protection of privacy, fundamental rights and public trust

The first recommendation is to strengthen the legal and policy framework governing the use of AI within the criminal justice system. Malaysia has introduced several initiatives to enhance AI governance, including the National AI Roadmap 2021–2025, the National Guidelines on Artificial Intelligence Governance and Ethics (AIGE) and the establishment of the National AI Office (NAIO). However, a dedicated legal framework specifically regulating the use of AI in policing activities, including predictive policing, facial recognition and biometric analysis, has yet to be established. Accordingly, the relevant authorities may consider developing dedicated guidelines or a specialised legal framework that clearly defines the scope of AI deployment, the limits of law enforcement powers, standards for data processing and mechanisms for reviewing AI-generated decisions. Such an approach is consistent with the principle of risk-based regulation embodied in the European Union Artificial Intelligence Act, which classifies AI applications used in law enforcement as high-risk systems subject to more stringent regulatory requirements. Although this approach may not be adopted directly in the Malaysian legal context, its underlying risk-based regulatory principles may serve as a useful reference for developing an AI governance framework appropriate to the Malaysian legal system (European Union, 2024).

The second recommendation is to ensure that the principles of human oversight, transparency and algorithmic accountability govern all AI applications within the criminal justice system. AI should function as a decision-support tool rather than replacing the professional judgement of law enforcement officers or judicial authorities. Accordingly, any AI-assisted decision that may affect individuals' rights and liberties should remain subject to review by competent human decision-makers. In addition, the relevant authorities may consider implementing algorithmic impact assessments (AIAs), conducting periodic audits of AI systems and maintaining comprehensive documentation concerning data sources, processing methodologies and the rationale underlying algorithmic decisions. Such measures would not only enhance transparency but also strengthen institutional accountability in the event of system errors or misuse (NIST, 2023; UNESCO, 2021).

The third recommendation is to strengthen the protection of fundamental human rights and personal data throughout every stage of AI development and implementation. Technologies such as facial recognition, biometric analysis and real-time surveillance have considerable potential to improve public safety; however, they also present significant risks to privacy rights and individual freedoms. Accordingly, AI development should be guided by the principles of privacy by design, fairness, non-discrimination, and proportionality, while ensuring that personal data are processed lawfully, necessary, and proportionate. Within the Malaysian context, the implementation of these principles should also be aligned with the existing personal data protection framework, including the Personal Data Protection Act 2010 (Act 709), as well as ongoing developments in national AI policy. Furthermore, the relevant authorities may consider establishing or involving an independent oversight body to assess compliance with AI ethical principles, thereby enhancing public confidence in the use of AI within the criminal justice system (OECD, 2019; Almeida et al., 2022).

Taken together, the analysis presented in this section suggests that AI governance within the criminal justice system should be regarded as a continuous process that evolves alongside technological advancements and developments in the legal landscape. Within the context of the Malaysian criminal justice system, this study proposes that a Responsible AI Governance Framework should be underpinned by five core principles: human oversight, transparency, accountability, fairness, and privacy protection. These principles were selected because they are consistently reflected in leading international governance instruments, including the OECD AI Principles (2019), the UNESCO Recommendation on the Ethics of Artificial Intelligence (2021), the NIST AI Risk Management Framework (2023) and Malaysia's National Guidelines on Artificial Intelligence Governance and Ethics (AIGE). Collectively, these principles provide a valuable reference for the future development of policies, regulatory guidelines and legal reforms governing AI within the criminal justice system, while supporting the responsible, transparent and lawful implementation of AI in a manner that upholds the rule of law and protects fundamental human rights. Accordingly, the framework proposed in this study is not intended to represent a definitive regulatory model; rather, it provides a conceptual foundation that may serve as a reference for developing a more responsible and contextually appropriate approach to AI governance within the Malaysian criminal justice system.

Conclusion

AI is increasingly reshaping the criminal justice system by enabling data-driven approaches to crime prevention, criminal investigation and law enforcement through applications such as predictive policing, AI-driven surveillance, facial recognition and intelligent decision support. This article has demonstrated that while these technologies offer considerable potential to enhance operational efficiency and public safety, their deployment also raises significant legal and ethical challenges, particularly regarding algorithmic bias, privacy, transparency, accountability, and the protection of fundamental rights. These findings suggest that the successful implementation of AI in criminal justice depends not only on technological capability but also on an appropriate governance framework that ensures AI systems remain lawful, transparent, accountable, and subject to meaningful human oversight. Building upon the issues identified throughout this study, the article proposes a Responsible AI Governance Framework for the Malaysian criminal justice system founded upon five core principles, namely human oversight, transparency, accountability, fairness and privacy protection. Rather than prescribing a definitive regulatory model, the proposed framework provides a conceptual foundation that may assist policymakers, legislators, enforcement agencies and other relevant stakeholders in developing legal and

institutional safeguards for the responsible use of AI within the Malaysian criminal justice system. It is anticipated that this study will contribute to the growing body of scholarship on AI governance and encourage further research into developing context-specific regulatory approaches that balance technological innovation with the protection of fundamental rights and the rule of law.

References

- Alias, M. A. A., Wan Ismail, W. A. F., Baharuddin, A. S., Hasnizam, H., Tuan Ibrahim, T. M. F. H., & Tuan Muhammad Faris Hamzi, T. M. F. H., & Mohamad Sukri, M. N. (2025). The integration of artificial intelligence (AI) into the shari'ah judiciary: A maqāṣid al-shari'ah approach to ethical and legal transformation. In *CFORSJ Procedia*, (pp. 119-127). <https://alnadwah.usim.edu.my/cforsjprocedia/paper/view/204>
- Almasoud, A. S., & Idowu, J. A. (2025). Algorithmic fairness in predictive policing. *AI and Ethics*, 5, 2323–2337. <https://doi.org/10.1007/s43681-024-00541-3>
- Almeida, D., Shmarko, K., & Lomas, E. (2022). The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: A comparative analysis of US, EU, and UK regulatory frameworks. *AI and Ethics*, 2(3), 377–387. <https://doi.org/10.1007/s43681-021-00077-w>
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671–732.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chong, S. Z., & Kuek, C. Y. (2022). Facial recognition technology in Malaysia: Concerns and legal issues. In *Proceedings of the International Conference on Law and Digitalization (ICLD 2022)* (pp. 101–109). Atlantis Press. https://doi.org/10.2991/978-2-494069-59-6_10
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- European Union. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union.
- Garvie, C., Bedoya, A. M., & Frankle, J. (2016). *The perpetual line-up: Unregulated police face recognition in America*. Georgetown Law Center on Privacy & Technology.
- Gstrein, O. J., Bunnik, A., & Zwitter, A. (2019). Ethical, legal and social challenges of predictive policing. *Católica Law Review*, 3(3), 77–98.
- Hutchinson, T., & Duncan, N. (2012). Defining and describing what we do: Doctrinal legal research. *Deakin Law Review*, 17(1), 83–119. <https://doi.org/10.21153/dlr2012vol17no1art70>
- Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4–20.
- Lim, A. C. M., Ng, L. H. X., & Taeihagh, A. (2025). Biometric data landscape in Southeast Asia: Challenges and opportunities for effective regulation. *Computer Law & Security Review*, 56, 106095. <https://doi.org/10.1016/j.clsr.2024.106095>
- Mandalapu, V., Elluri, L., Vyas, P., & Roy, N. (2023). Crime prediction using machine learning and deep learning: A systematic review and future directions. *IEEE Access*, 11, 60153–60170. <https://doi.org/10.1109/ACCESS.2023.3286344>
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35.
- Meijer, A. J., & Wessels, M. (2019). Predictive policing: Review of benefits and drawbacks. *International Journal of Public Administration*, 42(12), 1031–1039. <https://doi.org/10.1080/01900692.2019.1575664>
- Ministry of Digital. (2024). *National guidelines on artificial intelligence governance and ethics (AIGE)*. Ministry of Digital.
- Mohd Nor, M. A., Mohd Tasrib, M. A., Francis, B., Hesham, N. I., & Othman, M. B. (2021). A study on the laws governing facial recognition technology and data privacy in Malaysia. *Malaysian*

- Journal of Social Sciences and Humanities*, 6(10), 480–489.
<https://doi.org/10.47405/mjssh.v6i10.1086>
- Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., & Brantingham, P. J. (2015). Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, 110(512), 1399–1411. <https://doi.org/10.1080/01621459.2015.1077710>
- National AI Office. (2025). “About the National AI Office (NAIO)”. <https://ai.gov.my/about-naio/>
- National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)*. U.S. Department of Commerce.
- Organisation for Economic Co-operation and Development. (2019). “OECD principles on artificial intelligence”. <https://oecd.ai/en/ai-principles>
- Oswald, M., Grace, J., Urwin, S., & Barnes, G. C. (2018). Algorithmic risk assessment policing models: Lessons from the Durham HART model and "experimental" proportionality. *Information & Communications Technology Law*, 27(2), 223–250.
<https://doi.org/10.1080/13600834.2018.1458455>
- Perry, W. L., McInnis, B., Price, C. C., Smith, S. C., & Hollywood, J. S. (2013). *Predictive policing: The role of crime forecasting in law enforcement operations*. RAND Corporation.
- Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *New York University Law Review Online*, 94, 15–55.
- Saunders, J., Hunt, P., & Hollywood, J. S. (2016). Predictions put into practice: A quasi-experimental evaluation of Chicago's predictive policing pilot. *Journal of Experimental Criminology*, 12(3), 347–371. <https://doi.org/10.1007/s11292-016-9272-0>
- Smith, M., & Miller, S. (2022). The ethical application of biometric facial recognition technology. *AI & Society*, 37(1), 167–175. <https://doi.org/10.1007/s00146-021-01199-9>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO.
- Valera, M., & Velastin, S. A. (2005). Intelligent distributed surveillance systems: A review. *IEE Proceedings – Vision, Image and Signal Processing*, 152(2), 192–204.