

LEVERAGING BIG DATA FOR INDONESIA'S IMMIGRATION POLICY: OPPORTUNITIES AND LIMITATIONS

Rafi Dewanto, Cakra Trinata, Muhammad Fahrury Romdendine

Manajemen Teknologi Keimigrasian, Politeknik Imigrasi

Jln Raya Gandul No. 4, RT.07 / RW.08, RT.5/RW.6, Kota Depok, Jawa Barat 16514

Rafinavynanda29@gmail.com

ABSTRACT

Indonesia's fragmented immigration data systems, exacerbated by its archipelagic geography and institutional complexity, pose significant challenges to effective immigration management. These issues result in inefficiencies in border security, difficulties in tracking irregular migration, and substantial economic losses due to undocumented migrant workers. For instance, discrepancies between official data and World Bank estimates reveal a gap of 5.3 million unrecorded migrant workers, highlighting systemic failures in data integration and enforcement. This study explores the potential of big data analytics to address these challenges by integrating disparate systems and enhancing decision-making processes. Using a normative juridical approach, the research examines Indonesia's legal frameworks and proposes a comprehensive implementation framework. This framework includes centralized data integration using Hadoop and Spark technologies, predictive analytics for migration patterns, and robust privacy safeguards to protect vulnerable populations. The findings emphasize that big data can significantly improve operational efficiency, enhance national security, and support evidence-based policy development. However, the study also identifies critical barriers such as privacy concerns, technical limitations, and institutional coordination gaps. By addressing these challenges, the proposed framework offers actionable insights into leveraging big data for effective immigration policing in Indonesia while safeguarding civil liberties.

Keywords: *Big Data, Immigration Policy, Data Integration, Digital Transformation, Public Administration*

1. DISCUSSION

Indonesia faces significant challenges in managing immigration data effectively, creating opportunities for innovative solutions through big data implementation. The archipelagic nature of Indonesia, with over 17,000 islands and numerous entry points, presents unique challenges for immigration management, requiring sophisticated data-driven approaches to ensure national security while facilitating legitimate movement across borders. Currently, immigration data management in Indonesia suffers from fragmentation across multiple agencies, including Customs, Immigration, and Quarantine (CIQ) services, which operate under different ministerial authorities with separate information systems. These disparate systems lack optimal coordination and centralized data collaboration, significantly affecting the effectiveness and efficiency of single-window services at border checkpoints throughout the country[1]. This fragmentation creates substantial gaps in immigration enforcement and policy development that could potentially be addressed through comprehensive big data integration. According to [2] Indonesia's immigration control at airports faces critical challenges due to the lack of integrated databases across relevant agencies.

The absence of integrated citizen data has resulted in alarming discrepancies between official statistics and actual immigration figures in Indonesia. According to data from the Agency for the Protection of Indonesian Migrant Workers (BP2MI), approximately 3.7 million Indonesian migrant workers

are employed abroad, while World Bank estimates suggest a significantly higher figure of 9 million[3]. This substantial difference of 5.3 million undocumented migrant workers not only represents a failure in data management but also correlates directly with financial losses for the nation in terms of unrealized remittances. Moreover, such discrepancies point to broader issues in immigration policing, including difficulties in tracking illegal entry and overstays, managing visa violations, and protecting vulnerable populations subjected to human trafficking or labor exploitation. The establishment of integrated data systems has thus become an urgent priority for Indonesian immigration authorities seeking to address these challenges effectively.

Current immigration data management in Indonesia represents a patchwork of systems with limited integration capabilities. The National Single Window system for immigration operates alongside agency-specific databases that are not fully compatible with one another. For instance, while BP2MI and the Directorate General of Immigration have made efforts to integrate their SISKOP2MI and SIMKIM systems, this integration remains partial and does not extend to all relevant agencies involved in migration management[3]. At Indonesia's international airports, immigration officers implement a selective policy using limited border control technologies, but these systems lack the comprehensive data integration needed for truly effective immigration enforcement[2]. The correlation between selective policy implementation and security approaches at

border controls indicates potential for improvement through advanced data analytics, yet current infrastructure falls short of exploiting this potential fully.

The implementation of big data in Indonesian immigration systems has become increasingly urgent due to evolving security threats and economic considerations. In border regions like East Nusa Tenggara (NTT), which shares a direct border with Timor-Leste, government initiatives to develop Integrated Border Crossing Posts (PLBN Terpadu) have highlighted the critical need for data integration across the five ministries overseeing border operations[1]. Each service within these facilities currently maintains separate systems, creating inefficiencies in the immigration clearance process and hindering effective border management. The rise in international visitors, workers, investors, and students entering Indonesia, coupled with security concerns regarding illegal entrants, underscores the need for more sophisticated data management approaches that can analyze patterns, predict risks, and facilitate evidence-based decision making at immigration control points nationwide.

Big data implementation offers transformative potential for immigration policy and enforcement in Indonesia through several key mechanisms. Modern approaches to law enforcement increasingly prioritize crime prevention over reaction, with predictive policing models based on big data analytics being developed in various countries to anticipate migration patterns and potential security issues. Integration of citizen data at the national level would combine census information, electronic identification (e-KTP), social security data (BPJS), and immigration records to create comprehensive profiles that enable more effective immigration management[4]. Such integration would address current challenges including resource limitations, changing scopes of law enforcement, and fragmented information systems that hinder effective immigration policing. By analyzing vast quantities of structured and unstructured data, immigration authorities could accelerate processing times, develop preventive strategies, support digital transformation of border management, and significantly increase operational efficiency across the immigration system.

Despite its potential benefits, implementing big data in Indonesian immigration policing faces substantial challenges and limitations that must be carefully addressed. Privacy concerns represent a significant obstacle, as predictive policing using big data carries risks of environmental stigmatization and potential violations of individual privacy rights[5]. The surveillance capabilities enabled by comprehensive data integration may conflict with democratic principles and civil liberties if implemented without appropriate safeguards and oversight mechanisms. Technical challenges also abound, including the need for substantial

infrastructure development, interoperability between existing systems, data quality assurance, and specialized human resources capable of managing and interpreting complex data analytics. Furthermore, institutional challenges related to inter-agency coordination, legal frameworks for data sharing, and resistance to organizational change present additional barriers to successful implementation of big data solutions in immigration management.

This article aims to develop a comprehensive framework for implementing big data analytics in Indonesian immigration policing that balances security imperatives with privacy protections and practical implementation considerations. Unlike previous research that has focused primarily on either technical aspects of data integration or policy implications separately, this study will propose an integrated approach that addresses both dimensions simultaneously while considering Indonesia's unique geographical, institutional, and cultural context. By examining international best practices in immigration data management and adapting them to Indonesian conditions, this research will contribute a novel implementation roadmap that accounts for existing infrastructure limitations, legal constraints, and institutional capacities. The proposed framework will specifically address data governance issues, technical architecture requirements, privacy protection mechanisms, and inter-agency coordination protocols necessary for successful deployment of big data solutions in Indonesian immigration enforcement, ultimately enhancing both security outcomes and operational efficiency while safeguarding civil liberties.

2. LITERATURE REVIEW

Immigration remains a central issue for nation-states, presenting both economic opportunities and governance challenges [6]. Effectively managing migration flows is crucial for maintaining national security, fostering economic growth, and promoting social cohesion. Conversely, poorly managed immigration can strain public resources, exacerbate social tensions, and increase security risks. Indonesia, given its unique geographical context and strategic location, requires a robust and adaptive immigration management system. Indonesia's immigration policies are guided by Law No. 6 of 2011, which establishes the legal framework for regulating the entry, stay, and exit of foreign nationals (*Undang-Undang Republik Indonesia Nomor 6 Tahun 2011 tentang Keimigrasian*). However, the law's effectiveness is constrained by data fragmentation and limited inter-agency coordination, hindering the government's ability to efficiently track and manage migration flows. Addressing these challenges is essential for strengthening border security and maximizing the benefits of immigration while minimizing associated risks.

Big data has emerged as a transformative technology for processing and analyzing vast volumes of structured and unstructured data, providing new insights and capabilities across various domains [7]. Recent advancements in big data analytics have enabled organizations to improve decision-making, enhance operational efficiency, and gain a competitive edge. In immigration management, big data offers the potential to revolutionize traditional approaches by enabling predictive analytics, real-time monitoring, and evidence-based policymaking. The application of big data in immigration policing raises important considerations related to privacy, algorithmic bias, and civil liberties. Implementing big data solutions requires careful attention to data governance, transparency, and accountability to ensure the technology is used ethically and responsibly. Striking a balance between leveraging big data for security and protecting individual rights is crucial for maintaining public trust and upholding democratic values.

Effective data management is essential for organizations to harness the value of their information assets and achieve their strategic objectives [8]. Data management encompasses a range of activities, including data governance, data quality management, data integration, and data security [9]. In immigration management, robust data management practices are critical for ensuring the accuracy, reliability, and accessibility of immigration-related information. Indonesia's immigration data management faces significant challenges due to fragmented systems, limited interoperability, and a lack of standardized data models. These challenges hinder effective information sharing among relevant agencies, leading to inefficiencies in border control, difficulties in tracking irregular migration, and challenges in combating transnational crime. Integrating immigration data into a centralized platform, as proposed in this study, can address these challenges and improve the overall effectiveness of Indonesia's immigration system.

3. RESEARCH METHOD

The research method used in this study follows a descriptive analytical approach with a qualitative normative juridical focus. This method is selected to provide a comprehensive analysis of immigration data management challenges in Indonesia while also evaluating the legal and policy frameworks governing immigration and national security. The normative juridical approach is employed to examine laws, regulations, and policies related to immigration data governance, inter-agency coordination, and the implementation of big data solutions in border management.

This research is qualitative, prioritizing an in-depth understanding of the subject matter rather than broad generalizations, unfolding in five sequential stages to trace the research flow from start to finish. First, we map the current state of immigration data management in Indonesia by conducting library

research on primary legal materials (immigration laws, government regulations), secondary sources (academic articles, policy reports), and tertiary references (legal dictionaries, encyclopedias). Second, we examine existing integration efforts and their limitations, using thematic synthesis and legal interpretation to identify gaps in inter-agency coordination and data governance. Third, we explore opportunities for big data implementation in immigration policing, drawing on best practices and policy analyses to pinpoint high-value use cases. Fourth, we assess challenges and limitations in big data implementation, including privacy concerns, technical constraints, and institutional complexities, through in-depth qualitative categorization. Finally, we develop a proposed framework for big data implementation in Indonesian immigration policing, integrating insights from each prior stage to recommend governance structures, phased rollouts, and privacy safeguards.

4. DISCUSSION

4.1. The Current State of Immigration Data Management in Indonesia

Indonesia's management of immigration data currently operates within a fragmented ecosystem characterized by multiple disparate systems, limited integration, and significant operational inefficiencies. The archipelagic geography of the nation, comprising over 17,000 islands with numerous international entry points, compounds the challenges of maintaining comprehensive immigration records and enforcing consistent policies across diverse locations. Immigration authorities operate under the supervision of the Directorate General of Immigration within the Ministry of Law and Human Rights, yet their data systems intersect with numerous other agencies including the Agency for the Protection of Indonesian Migrant Workers (BP2MI), Ministry of Foreign Affairs, Ministry of Manpower, National Police, and customs authorities[3]. This institutional complexity creates substantial barriers to information sharing and coordinated enforcement actions that could otherwise be facilitated through integrated data systems. [10] emphasize that Indonesia's migration and border governance remain constrained by a lack of defined policy structures and poor data integration among key institutions. Their study highlights that overlapping mandates and fragmented data systems hinder effective decision-making, leading to inefficiencies in border security and immigration control.

The limitations of current immigration data infrastructure become evident when examining specific cases of information discrepancy within the system. According to BP2MI records, approximately 3.7 million Indonesian migrant workers are officially registered as working abroad, yet World Bank estimates place this figure at approximately 9 million individuals. This striking difference of 5.3 million unaccounted workers represents not merely a

statistical anomaly but indicates systematic failures in tracking population movements across Indonesia's borders. Former BP2MI head Benny Rhamdani has directly linked this data discrepancy to the activities of human trafficking syndicates and irregular migration channels that operate outside official government oversight[3]. The economic implications are equally concerning, as unrecorded migrant workers translate to unrealized remittances that could otherwise contribute significantly to Indonesia's economy. These discrepancies demonstrate that current data management approaches are insufficient to capture the full scope of migration flows affecting the nation.

At international airports, which serve as primary entry points for foreign nationals, immigration clearance processes rely on selective policy implementation that remains hampered by information limitations. Research examining the correlation between immigration clearance processes, selective policy implementation, and security approaches at Indonesian airports indicates strong interrelationships between these factors, suggesting that improved data integration could significantly enhance border security outcomes [2]. The current system emphasizes manual verification processes supplemented by limited digital documentation, creating bottlenecks that impact both security effectiveness and visitor experiences. Furthermore, the growing volume of international travelers has strained existing systems, highlighting the urgent need for more sophisticated data processing capabilities that can maintain security standards while facilitating efficient movement through immigration checkpoints.

The challenges in immigration data management extend to Indonesia's land borders, where integration issues are particularly evident. At Integrated Border Crossing Posts (PLBN Terpadu) in provinces such as East Nusa Tenggara that share borders with neighboring countries like Timor-Leste, the coexistence of multiple systems under different ministerial authorities creates substantial coordination challenges [1]. These border facilities house Customs, Immigration, and Quarantine (CIQ) services operating under the jurisdiction of five different ministries, each maintaining separate data systems that lack effective central coordination. This arrangement undermines the "single window service" concept that these integrated facilities were designed to implement, resulting in duplicated data collection, information silos, and reduced operational effectiveness at critical border checkpoints.

4.2. Existing Integration Efforts and Their Limitations

Th Despite these challenges, Indonesia has initiated several data integration efforts that provide foundations for more comprehensive big data implementation in immigration management. The integration between BP2MI's SISKOP2MI system and the Directorate General of Immigration's SIMKIM

platform represents one such attempt to create more cohesive information flows between agencies responsible for managing migrant workers. This limited integration has proven particularly valuable during the COVID-19 pandemic, enabling authorities to track returning migrant workers more effectively and coordinate their safe repatriation. Similarly, the implementation of the Border Control Management (BCM) system at airports has introduced elements of digital verification and biometric matching that enhance security while potentially generating valuable data for pattern analysis[2].

However, these integration efforts remain partial and insufficient to address the fundamental data management challenges facing Indonesian immigration authorities. [11] emphasize that while efforts to integrate immigration intelligence functions and early warning systems have improved monitoring, they still lack full interoperability. Their study highlights that many immigration databases remain siloed, limiting the potential for cross-agency analytics that could enhance the detection of irregular migration and security threats. Additionally, existing systems operate primarily as transaction processing platforms rather than analytical engines capable of identifying patterns, predicting trends, or supporting evidence-based policy development.

4.3. Opportunities for Big Data Implementation in Immigration Policing

The implementation of big data analytics presents transformative opportunities for enhancing immigration policing in Indonesia through multiple pathways that could address current systemic limitations. Big data technologies enable the processing and analysis of vastly larger and more diverse datasets than conventional database systems, creating possibilities for more comprehensive situational awareness across Indonesia's complex border environment. Through the integration of structured data from official government sources with unstructured information from social media, financial systems, telecommunications, and other external datasets, immigration authorities could develop significantly more nuanced understandings of migration patterns, potential threats, and policy impacts [12]. This expanded analytical scope would transcend the current transactional focus of immigration systems to enable both strategic planning and tactical responses based on evidence rather than limited sampling or anecdotal information.

Big data analytics also plays a crucial role in analyzing mass movements and predicting future migration trends, offering valuable insights for policy-making [13], [14]. Predictive analytics represents one of the most promising applications of big data for immigration policing in Indonesia. Modern law enforcement increasingly emphasizes prevention over reaction, with predictive policing models developed to anticipate potential issues before they materialize. In

the immigration context, predictive analytics could forecast migration trends based on factors including political instability in neighboring regions, economic changes, environmental disasters, or seasonal patterns that historically affect population movements. Such predictions would enable more efficient resource allocation, with personnel and technical assets deployed proactively to areas likely to experience increased migration activity. For instance, analysis of historical data combined with current indicators could predict potential surges in unauthorized border crossings at specific locations, allowing immigration enforcement to strengthen presence in those areas before significant problems develop [15].

Big data implementation could also substantially improve visitor screening and risk assessment processes at Indonesia's international entry points. Current selective policy implementation at airports relies heavily on individual officer judgment supplemented by limited database checks. A comprehensive big data approach would enable real-time risk scoring based on numerous variables, including travel history, booking patterns, financial transactions, known associates, and behavioral indicators. This multifactorial analysis would significantly enhance detection of potential security threats while simultaneously facilitating faster processing for low-risk travelers. Evidence-based policymaking, supported by comprehensive big data analysis, ensures that decision-making is more data-driven, reducing reliance on outdated practices and improving governance structures [16].

For the management of Indonesian migrant workers, big data solutions offer particular advantages in addressing current data discrepancies and protection challenges. Integration of migration records, employment contracts, remittance data, and communications patterns could create comprehensive profiles of migrant worker populations, enabling better protection services and more accurate economic planning. Such integration would help close the substantial gap between official records and actual migrant numbers by creating multiple verification points and cross-referenced indicators of migration activity. Additionally, predictive analytics can assist in identifying high-risk employment sectors or regions where exploitation is prevalent, allowing authorities to take preventive action. This section discusses the role of big data in development policy, stating that machine learning models analyzing financial and communication patterns can detect human trafficking networks and irregular migration channels. Furthermore, pattern analysis could identify potential trafficking networks or exploitative employment situations based on unusual movement patterns, remittance behaviors, or communication anomalies, enabling more proactive protection of vulnerable citizens working abroad.

4.4. Challenges and Limitations in Big Data Implementation

Despite the significant opportunities big data presents for immigration policing in Indonesia, numerous challenges and limitations must be addressed for successful implementation. Privacy concerns stand among the most significant issues, as comprehensive data collection and analysis for immigration enforcement inevitably involves processing substantial amounts of personal information. The establishment of surveillance capabilities through big data analytics raises legitimate concerns about potential civil liberties infringements, particularly in democratic societies where individual rights must be balanced against security imperatives [17]. Recent events have highlighted these vulnerabilities, as evidenced by the massive data breach in 2023 that exposed over 34 million Indonesian passports from the Immigration Directorate General, demonstrating the current systems' security deficiencies [18]. Any implementation of big data solutions must therefore incorporate robust privacy protection frameworks that clearly delineate permitted data uses, establish consent mechanisms where appropriate, and create transparent processes for data correction and deletion when necessary [19].

Environmental stigmatization represents another significant risk associated with predictive policing models based on big data analytics. When algorithms identify geographic areas as high-risk for immigration violations based on historical data, these predictions can potentially reinforce existing biases or create self-fulfilling prophecies through intensified enforcement activities [5]. Implementing safeguards against such stigmatization requires careful algorithm design, regular bias testing, and ongoing human oversight of automated recommendations to ensure fair and equitable enforcement practices across all communities [20].

Institutional challenges further complicate the implementation of big data for immigration policing in Indonesia. Effective data sharing requires not only technical integration but also organizational agreements regarding data ownership, access rights, usage limitations, and governance frameworks. Developing such agreements across the multiple ministries involved in immigration management—including Law and Human Rights, Foreign Affairs, Labor, and Security agencies—requires navigating complex bureaucratic structures and potentially conflicting institutional priorities. Research identifies that "limitations make it challenging to share actionable data and research across agencies and between stakeholders, thus impeding evidence-based 'whole-of-government' and 'whole-of-society' responses to migration" [22]. Resistance to change within established organizational cultures presents additional challenges, as transitioning from current process-focused approaches to data-driven decision

making requires substantial shifts in operational mindsets.

4.5. Proposed Framework for Big Data Implementation in Indonesian Immigration Policing

Addressing the opportunities and challenges identified in previous sections requires a comprehensive implementation framework specifically designed for the Indonesian context. This proposed framework integrates technical architecture, governance structures, privacy safeguards, and institutional coordination mechanisms to enable effective big data utilization in immigration policing while mitigating associated risks. The foundation of this framework rests on a centralized data integration platform that would serve as a unified repository for immigration-related information from all relevant agencies, including the Directorate General of Immigration, BP2MI, National Police, Customs, and intelligence services. This approach aligns with the Indonesian government's broader strategy of using Big Data to improve e-government programs, particularly in areas of illegal activity prevention and workforce effectiveness enhancement[23]. Unlike current limited integration efforts between specific systems, this platform would implement standardized data models, consistent identification mechanisms, and comprehensive API structures to enable seamless information flow while maintaining appropriate access controls based on agency authorities and legitimate purposes.

This framework adopts a layered data-management architecture that balances the benefits of centralization with Indonesia's institutional realities by housing core immigration datasets—passport details, visa records, biometric identifiers, and entry/exit timestamps—in a secure, access-logged central repository built on Apache Hadoop Distributed File System (HDFS) for scalable, fault-tolerant storage, while connecting supplementary sources—financial records, telecommunications metadata, social media feeds, and intelligence reports—through encrypted, audited interfaces that avoid wholesale centralization of raw data. By leveraging government-controlled cloud infrastructure alongside on-premises components in Indonesia's four National Data Centers, the design preserves data sovereignty and practical deployment across agencies. A Spark Data Access Control Framework enforces fine-grained, role-based policies uniformly across big-data engines, supporting both direct SQL queries and programmatic SparkSQL DataFrame access, and a rigorous metadata layer catalogs assets, tracks lineage, and monitors quality to prevent “data swamp” degradation—ensuring that immigration authorities can flexibly, securely, and efficiently manage and analyze large volumes of structured and unstructured data.

Leveraging Apache Spark as the primary in-memory data processing engine, the proposed framework offers a comprehensive suite of analytical capabilities—descriptive, diagnostic, predictive, and prescriptive—tailored to address various dimensions of immigration management in Indonesia. Descriptive analytics would serve as the foundation for situational awareness, enabling immigration authorities to monitor real-time migration flows, entry and exit point activities, overstay statistics, and enforcement operations through interactive dashboards developed using visualization tools such as Tableau or Power BI. These dashboards, securely connected to the Spark backend via APIs, would facilitate rapid response to emerging issues and align with the Indonesian government's broader goals of using big data to enhance event management and improve public service responsiveness. Building upon this, diagnostic analytics would utilize SparkSQL for complex querying and pattern detection, allowing deeper exploration into the root causes of irregular migration trends, compliance violations, or systemic inefficiencies in current practices. Predictive analytics would further enhance decision-making by employing models trained on historical immigration data alongside external indicators—such as economic shifts, political instability, and climate-related events—to forecast potential future migration movements and areas of concern. These models could be developed using Spark MLlib or integrated with more advanced machine learning libraries, enabling nuanced, multi-variable forecasting. Lastly, prescriptive analytics would operationalize the insights generated by recommending specific policy responses or enforcement strategies, such as targeted patrol deployments or resource reallocations, ensuring that actions are evidence-based rather than driven by reactive assumptions.

The implementation would also leverage recent advancements in AI integration with big data systems, following approaches similar to those used in the development of KemenkeuGPT by the Indonesian Ministry of Finance[24]. While KemenkeuGPT focuses on financial data and regulations, similar techniques using LangChain with Retrieval-Augmented Generation (RAG), prompt engineering, and fine-tuning could be applied to immigration datasets. This would enable natural language interfaces for immigration officials to query complex data patterns and receive insights without requiring specialized technical expertise. The model could be trained on immigration regulations, historical case data, and policy documents to provide context-aware recommendations for enforcement actions while ensuring compliance with legal frameworks.

Privacy protection is a critical component of the proposed framework, featuring multiple safeguards designed to prevent misuse while enabling legitimate immigration enforcement activities. To begin with, a comprehensive data classification system would

categorize information based on sensitivity levels, imposing corresponding access restrictions and usage limitations for each category; this is particularly vital for safeguarding personal data of vulnerable populations in an era where law enforcement's reliance on AI and advanced technologies is rapidly increasing. In parallel, technical measures such as encryption, anonymization, and differential privacy techniques would create layered protection for sensitive data, while the framework would also leverage Hadoop security features—including Kerberos authentication, HDFS encryption, and Apache Ranger for centralized authorization—to establish a robust security perimeter. Complementing these technical controls, data governance policies would be codified and enforced through automated mechanisms that ensure appropriate data handling across every stage of the immigration analytics pipeline. Transparency mechanisms would inform individuals about data collection practices and usage policies, and independent oversight bodies would be established to ensure compliance with these guidelines, collectively balancing security imperatives with privacy protection to maintain public trust in immigration enforcement activities. Moreover, to address concerns related to algorithmic bias and environmental stigmatization, the framework would incorporate regular bias testing and human oversight of automated recommendations, integrating fairness metrics and bias detection algorithms within the Spark-based analytics pipeline to continuously monitor and mitigate potential discrimination, especially where AI systems in immigration enforcement may disproportionately affect minority groups.

5. CONCLUSION

Our findings indicate that by integrating diverse data sources across Indonesian immigration agencies, it is possible to overcome existing fragmentation, significantly boost analytical capabilities, and transition from reactive enforcement to a more proactive, intelligence-driven approach—ultimately yielding deeper insights into migration patterns, more precisely targeted policies, optimized resource allocation, enhanced service delivery for legitimate travelers, and stronger safeguards for migrant workers. To realize these benefits, we recommend initiating a phased implementation that focuses first on a few high-impact use cases, while concurrently establishing a clear governance framework and embedding robust privacy and data-protection measures from the start. This approach should be supported by dedicated investments in scalable technical infrastructure and comprehensive staff training programs, as well as mechanisms to foster sustained interagency coordination and continuous engagement with civil society. By balancing ambition with practicality—prioritizing quick wins, ensuring accountability, and maintaining an open dialogue with stakeholders—

Indonesia's immigration authorities can harness the power of big data to improve security outcomes, enhance policy formulation, and uphold the rights and dignity of all individuals affected by immigration processes.

REFERENCE

- [1] F. M. H. Tjiptabudi, R. Bernardino, and H. Asikin, "Analisis Kelayakan Implementasi Big Data Dalam Sistem Layanan Customs, Immigrate dan Quarantine (CIQ) Pada Pos Lintas Batas Negara Terpadu," 2018, [Online]. Available: <http://www.uyelindo.ac.id>.
- [2] R. Arifin, I. Nurkumalawati, and B. Briando, *THE THEORETICAL PERSPECTIVES OF IMMIGRATION CONTROLS: IMMIGRATION CLEARANCE PROCESS, SELECTIVE POLICY AND SECURITY APPROACH AT AIRPORTS IN INDONESIA*. APIAR, 2019. [Online]. Available: www.apiar.org.au
- [3] Imigrasi, "Temui Dirjen Imigrasi, Kepala BP2MI Sampaikan Pentingnya Big Single Data."
- [4] R. Priyosantoso, ; Muhamad, E. Aminanto, and F. Fadilah, "Citizen Data Integration as the Backbone of Law Enforcement Implementation in Indonesia," 2022, doi: 10.18415/ijmmu.v9i1.3562.
- [5] C. A. Windani, "Strategi dan Tantangan Predictive Policing di Era Big Data bagi Masyarakat Modern," *Deviance Jurnal Kriminologi*, vol. 7, no. 2, p. 101, Dec. 2023, doi: 10.36080/djk.2385.
- [6] A. Triandafyllidou and B. S. A. Yeoh, "Sustainability and Resilience in Migration Governance for a Post-pandemic World," *J Immigr Refug Stud*, vol. 21, no. 1, pp. 1–14, Jan. 2023, doi: 10.1080/15562948.2022.2122649.
- [7] D. Tosi, R. Kokaj, and M. Rocchetti, "15 years of Big Data: a systematic literature review," *J Big Data*, vol. 11, no. 1, p. 73, May 2024, doi: 10.1186/s40537-024-00914-9.
- [8] H. Nugroho and S. F. S. Gumilang, "Recommendations for Improving Data Management Process in Government of Bandung Regency using COBIT 4.1 Framework," in *Proceedings of the 2020 9th International Conference on Software and Computer Applications*, New York, NY, USA: ACM, Feb. 2020, pp. 57–61. doi: 10.1145/3384544.3384588.
- [9] J. Baijens, T. Huygh, and R. Helms, "Establishing and theorising data analytics governance: a descriptive framework and a VSM-based view," *Journal of Business Analytics*, vol. 5, no. 1, pp. 101–122, Jan. 2022, doi: 10.1080/2573234X.2021.1955021.
- [10] R. Arifin, E. A. Zulfa, M. Hanita, and A. J. Simon Runturambi, "Unveiling Indonesia's migration and border governance: Challenges and imperatives post-pandemic," *Social Sciences and*

- Humanities Open*, vol. 10, Jan. 2024, doi: 10.1016/j.ssaho.2024.101202.
- [11] W. Purnamasari, A. Josias, and S. Runturambi, "Optimizing the Functions of Immigration Intelligence and Immigration Assisted Villages in the Prevention of Trafficking in Persons: An Integrative Approach to Enhancing Community Resilience in Indonesia-A Systematic Literature Review," 2025. [Online]. Available: <https://ajesh.ph/index.php/gp>
- [12] A. M. Hamdi, B. Briando, and F. Santiago, "The Role of Artificial Intelligence in Immigration Law Enforcement: Balancing Efficiency, Transparency, and Ethical Accountability," *Journal of Multidisciplinary Sustainability Asean*, vol. 1, no. 6, pp. 376–385, 2024, doi: 10.70177/ijmsa.v1i6.1835.
- [13] E. E. Korkmaz, *Smart Borders, Digital Identity and Big Data: How Surveillance Technologies Are Used Against Migrants*. Bristol, UK: Bristol University Press, 2023. doi: 10.51952/9781529233520.
- [14] A. W. Ahmad Yar and T. Bircan, "Big data for official migration statistics: Evidence from 29 national statistical institutions," *Big Data Soc*, vol. 10, no. 2, Jul. 2023, doi: 10.1177/20539517231210244.
- [15] A. de N. G. de Rubalcava *et al.*, "Modeling Population Movements under Uncertainty at the Border in Humanitarian Crises: A Situational Analysis Tool," Mar. 2023.
- [16] L. Stielike, "Producing migration knowledge: From big data to evidence-based policy?," in *Evidence in Action between Science and Society: Constructing, Validating, and Contesting Knowledge*, Taylor and Francis, 2022, pp. 185–200. doi: 10.4324/9781003188612-13.
- [17] B. Underwood and H. Saiedian, "Mass surveillance: A study of past practices and technologies to predict future directions," *SECURITY AND PRIVACY*, vol. 4, no. 2, Mar. 2021, doi: 10.1002/spy2.142.
- [18] A. Hope, "34 million Indonesian Passports Exposed in a Massive Immigration Directorate Data Breach." Accessed: Mar. 12, 2025. [Online]. Available: <https://www.cpomagazine.com/cyber-security/34-million-indonesian-passports-exposed-in-a-massive-immigration-directorate-data-breach/>
- [19] J. Park and D. S. N. Heriyanto, "In Favor Of An Immigration Data Protection Law In Indonesia And Its Utilization For Contact Tracing," *Prophetic Law Review*, vol. 4, no. 1, pp. 1–21, Jun. 2022, doi: 10.20885/PLR.vol4.iss1.art1.
- [20] N. Andalibi, C. Pyle, K. Barta, L. Xian, A. Z. Jacobs, and M. S. Ackerman, "Conceptualizing Algorithmic Stigmatization," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, Apr. 2023, pp. 1–18. doi: 10.1145/3544548.3580970.
- [21] T. J. Ma, R. J. Garcia, F. Danford, L. Patrizi, J. Galasso, and J. Loyd, "Big data actionable intelligence architecture," *J Big Data*, vol. 7, no. 1, p. 103, Dec. 2020, doi: 10.1186/s40537-020-00378-7.
- [22] International Organization For Migration, *IOM Strategy for Indonesia (2022–2025)*. 2023. doi: doi:
- [23] OpenGov, "Indonesia: Big Data and AI in e-government," [opengovasia](https://opengovasia.com). Accessed: Mar. 19, 2025. [Online]. Available: Indonesia: Big Data and AI in e-government
- [24] G. Fajar Febrian and G. Figueredo, "KEMENKEUGPT: LEVERAGING A LARGE LANGUAGE MODEL ON INDONESIA'S GOVERNMENT FINANCIAL DATA AND REGULATIONS TO ENHANCE DECISION MAKING," 2024