

Integrating Analytics and Interoperability in Investigative Systems: Evaluating the ETA System Prototype in Anti Corruption Enforcement

Hasrol Hissam Osman¹, Hazwan 'Aizat Noordin², Johari Mat³, Azrul Fazwan Kharuddin⁴, Surendran Elangoven⁵

¹ Malaysian Anti-Corruption Commission (MACC), Malaysia.

² Human Capital Development Division, Public Service Department of Malaysia, Malaysia.

³ Universiti Tun Abdul Razak, Malaysia.

Email: johari.mat@unirazak.edu.my

⁴ Universiti Tun Abdul Razak, Malaysia.

Email: azrul.fazwan@ur.unirazak.edu.my

⁵ Intelli Education Group Sdn Bhd (IEG Campus), Malaysia.

Abstract

This study evaluates the Enhanced Tracking and Analysis (ETA) System Prototype, developed to modernize criminal profile data management for the Malaysian Anti Corruption Commission (MACC). The purpose was to determine how structured data integration, advanced analytics, and workflow automation improve investigative effectiveness compared to manual legacy processes. Drawing on existing literature, the study situates the ETA system within broader research on technology enabled anti corruption and digital governance that supports institutional transparency and accountability aligned with Sustainable Development Goal 16 (SDG 16) to “promote peaceful and inclusive societies... build effective, accountable and inclusive institutions” (United Nations, 2015). A mixed methods design was adopted, combining structured questionnaires and thematic interviews with investigative officers to assess dimensions of data accessibility, retrieval efficiency, analytical accuracy, and user satisfaction. Results show high performance across all dimensions, with statistically significant correlations indicating that improved data accessibility substantially reinforces analytical accuracy and user satisfaction. These outcomes demonstrate that structured ETL processes and centralized database environments enhance investigative workflows and support evidence based decision making. Findings also align with research that digital transformation and analytics reduce corruption risks by strengthening oversight capabilities. The study concludes that the ETA prototype not only meets its core objectives but also contributes to SDG 16 targets by promoting stronger, data driven institutional practices. Recommendations include full system deployment, integration with broader law enforcement data ecosystems, and future incorporation of AI analytics to further enhance pattern detection and predictive capacities....

Keywords: Criminal profile data management, Anti-corruption systems, ETA System Prototype, Digital transformation, Data analytics

INTRODUCTION

Effective management of criminal profile data has become increasingly critical for law enforcement agencies in the digital era. Within Malaysia, the Malaysian Anti-Corruption Commission (MACC) has faced mounting pressure to modernize its data systems as corruption investigations grow in scale and complexity, particularly in light of digital financial instruments, cyber-enabled schemes, and cross-border criminal behaviour that defy traditional investigative methods (Azam Baki, 2025). Recent initiatives by the MACC to integrate artificial intelligence and digital analytics into investigative processes underscore this shift toward technologically empowered enforcement, enabling faster detection of unusual transactions and anomalous behavioural patterns across large datasets (Bernama, 2025). In parallel, global research on law enforcement data systems highlights persistent challenges related to data quality, fragmentation, and interoperability, which directly impede agencies' ability to conduct comprehensive analyses and draw connections between disparate data sources (Mitar, 2004; MDPI, 2021).

Despite these technological advancements, the landscape of criminal data management remains fraught with structural and operational inefficiencies. One of the most salient challenges in contemporary policing is data fragmentation, where information relevant to investigations exists in siloed databases that do not communicate

with one another. This fragmentation hinders cross-referencing of evidence and limits an agency's capacity to develop holistic insights across cases (MDPI, 2021). Such silos can cause investigative delays, inconsistencies in case progression, and an increased likelihood of oversight particularly when diverse information types such as financial records, digital footprints, and forensic evidence reside on incompatible platforms. Moreover, challenges in data quality such as inaccuracies, incomplete entries, and lack of standardization further compound the difficulty of producing reliable analytical outputs (Mitar, 2004). In law enforcement contexts, poor data quality can directly impact strategic decision-making, resource allocation, and the ability to measure investigative progress, undermining both operational effectiveness and public trust.

Beyond technical constraints, law enforcement agencies like the MACC must contend with workflow inefficiencies that arise from manual data handling and inadequate automation. Traditional investigative workflows often involve repetitive tasks such as manual tagging, cross-checking records, and report generation that are time consuming and prone to human error (Braekt et al., 2017). These bottlenecks not only slow investigative velocity but also diminish the overall capacity for analytical depth, particularly when dealing with high volumes of evidence or complex legal frameworks. Additionally, the growing use of digital communication and social media in criminal contexts has introduced new dimensions of data, prompting agencies to process and validate vast quantities of unstructured information. This influx of data further stresses outdated systems that lack robust query and analytics functions capable of deriving meaningful patterns or predictive insights.

In the Malaysian context, these systemic issues are compounded by the evolving nature of corruption crimes themselves, which increasingly leverage digital systems and anonymity networks to conceal illicit activity. While MACC has begun adopting AI and blockchain tools to keep pace with these developments, concerns related to data governance, security, and integration persist, necessitating a more unified approach to managing investigative information across the organization (The Star, 2025). Such challenges suggest that without a structured, scalable platform designed specifically for criminal profile data management, investigative outcomes may remain suboptimal, and organizational responsiveness limited in the face of escalating demands.

The purpose of this study, therefore, is to explore the design and implementation of the ETA (Enhanced Tracking and Analysis) System Prototype as a strategic response to the aforementioned challenges, with a focus on its potential to unify criminal profile data, improve analytical accuracy, and support evidence-based investigative practices within the MACC.

Literature Review

Research on criminal data management systems and related variables such as data integration, analytics, quality, and adoption has grown substantially in recent years as law enforcement agencies worldwide embrace digital technologies for investigative work. A key focus of this literature has been on *data quality and interoperability*, which directly influence analytical credibility and operational effectiveness. Kusak (2022) highlights that poor data quality undermines the performance of big data and AI applications in law enforcement, leading to inaccurate insights and potentially eroding trust in automated decision-making systems, thus underscoring the importance of rigorous data management processes in analytics-driven criminal systems (ERA Forum, 2022). Complementary research on computational analysis of unstructured police data demonstrates that methods such as natural language processing and text mining have been increasingly applied to extract actionable insights from narrative reports, yet outcomes strongly depend on structured data practices and system integration (Lukmanjaya et al., 2026). These findings align with the broader trend toward *big data policing*, where the integration of diverse datasets and analytical tools supports more nuanced intelligence products and strategic operations (Neth. Police study, 2024).

Prior system-oriented research further illustrates how *data consolidation and analytics platforms* improve investigative outcomes. VALCRI a data analytics tool developed under the European Commission—aggregates information from multiple sources and supports visual sense-making for criminal analysts, showing that structured analytics can enhance pattern recognition in case data (VALCRI, n.d.). Likewise, India's *Crime and Criminal Tracking Network and Systems (CCTNS)* project demonstrates how integrating police records across thousands of entities can support single-search access and analytical consistency, though challenges remain in maintaining uniformity and real-time updates across decentralized inputs (CCTNS, n.d.). These system examples inform the need for unified data ecosystems that facilitate efficient querying and cross-case correlation, variables central to the ETA prototype's design objectives.

Studies on *technology adoption frameworks* provide further theoretical grounding for understanding how law enforcement agencies integrate new systems. Information systems adoption research frequently draws upon the *Technology–Organization–Environment (TOE)* framework, *Technology Acceptance Model (TAM)*, and related theories to explain organizational readiness and barriers to adoption. For instance, Ismail et al. (2025) synthesize multiple theoretical lenses including TOE and TAM to map technology adoption determinants, emphasizing factors such as technological compatibility, organizational context, and environmental pressures in shaping adoption decisions across various digital systems. Similarly, conceptual frameworks for big data analytics adoption articulate how organizational readiness, resource capacity, and governance structures influence the effective integration of analytics platforms, with data quality and management capabilities highlighted as key enablers of performance outcomes (Muhammad, 2022; Muhammad et al., 2021). These theoretical models underpin the logic for embedding structured ETL processes, role-based access controls, and scalable architecture in criminal data systems like ETA, linking system capabilities to organizational adoption and performance outcomes.

A growing body of research also illustrates *conceptual models* that bridge advanced analytics and law enforcement needs. For example, integrated forensic ontology and data harmonization tools have been proposed to enhance data integration and analytical flexibility, enriching the query and reporting functions critical for investigative systems (MDPI, 2021). At the same time, broader AI and big data applications in crime prediction and pattern analysis reveal how machine learning methods when fed high-quality, interoperable data can advance predictive capabilities and strategic decision-making in policing (Mandalapu et al., 2023). Collectively, these studies reinforce that effective criminal data systems require not only advanced analytical technologies but also robust data governance structures, interoperability mechanisms, and adoption-aware frameworks that together support reliable, secure, and actionable intelligence.

Research Methodology

This study employs a mixed methods research design, combining both quantitative and qualitative techniques to investigate the development and evaluation of the ETA (Enhanced Tracking and Analysis) system prototype for criminal profile data management. A mixed methods approach, which integrates numerical and contextual data, allows for a more holistic understanding of complex phenomena by drawing on the strengths of both methodological paradigms (Williams, 2023). In the context of information systems research, mixed methods enable researchers to capture not only the statistical performance and usability metrics of technological artifacts but also user perceptions, experiences, and implementation challenges, thereby enhancing analytical depth and validity. Such integrative designs mitigate the limitations inherent in purely quantitative or qualitative studies, and are particularly useful for evaluating prototype systems that require both performance assessment and user-oriented insights (Kharuddin & Ismail, 2017).

The procedure for this study comprises several sequential stages. First, a prototype development phase was executed, during which data architecture, user interface design, and ETL (Extract, Transform, Load) processes were constructed and iteratively refined. This phase informed the creation of measurement tools and data collection instruments. Second, pilot testing was carried out with a select group of MACC investigative officers to validate the usability features and ensure the instruments were reliable. Pilot testing is a recognized practice in research methodology that enhances data accuracy by identifying procedural errors and refining measurement tools before the main data collection (Palanisamy et al., 2025). Third, actual data collection took place following ethical clearance and informed consent processes to assure privacy and participation rights (Kharuddin et al., 2018).

For sampling and data collection, a purposive sampling strategy was adopted. Purposive sampling allows researchers to select participants who have direct and relevant expertise with criminal investigation systems, ensuring that insights collected are highly pertinent to the study's objectives. Investigators and data analysts from the MACC or related enforcement units were invited to participate based on their operational experience with investigative data systems. In mixed methods research, purposive sampling is frequently employed to ensure that both qualitative and quantitative elements are contextually rich and theoretically significant (Williams, 2023). Data collection involved two primary instruments: (a) structured questionnaires to collect user ratings on system performance, data accessibility, and analytical efficiency; and (b) semi-structured interviews to obtain narrative

insights into the practical challenges of data management and suggestions for system improvement. Questionnaires provide quantifiable metrics that can be subjected to statistical analysis, while interviews yield rich qualitative data that help interpret user perceptions and organizational contextual factors.

Data analysis procedures were aligned with the study's dual methodological orientation. Quantitative data from questionnaires were subjected to descriptive and inferential statistical analysis to evaluate system performance metrics such as retrieval speed, accuracy indexes, and user satisfaction scores. Standard analytical software packages (e.g., SPSS, R) were used to compute measures such as means, standard deviations, and correlation coefficients, facilitating comparisons between prototype outcomes and historical benchmarks. Qualitative interview data were analysed through thematic analysis, an established interpretive technique that identifies recurring patterns, concepts, and user sentiments across textual data. Thematic analysis enhances the explanatory power of qualitative findings by systematically categorizing responses into core themes relevant to system utility, integration challenges, and organizational implications. By triangulating the findings from both quantitative and qualitative analyses, this methodology supports robust inferences regarding the prototype's effectiveness and user acceptance, thereby strengthening the overall research validity and practical relevance.

Data Analysis, Results, and Findings

This study aimed to evaluate the effectiveness of the ETA (Enhanced Tracking and Analysis) System Prototype in improving criminal profile data management, with specific focus on data accessibility, retrieval efficiency, analytical accuracy, and user satisfaction. Data analysis was conducted using both descriptive and inferential statistics to examine quantitative responses obtained from 50 investigative officers participating in the evaluation phase. Descriptive statistics, including means and standard deviations, were used to summarize the participants' ratings of system performance, while inferential analysis assessed relationships between variables to determine the consistency and reliability of system functionality.

System Performance Evaluation

Table 1 presents the overall system performance scores for four core dimensions: data accessibility, retrieval efficiency, analytical accuracy, and user satisfaction. Each dimension was measured on a 5-point Likert scale, with higher values indicating better performance.

Table 1: System Performance Evaluation (N=50)

Dimension	Mean	Standard Deviation	Minimum	Maximum
Data Accessibility	4.36	0.48	3	5
Retrieval Efficiency	4.12	0.52	3	5
Analytical Accuracy	4.24	0.50	3	5
User Satisfaction	4.40	0.46	3	5

The results indicate that investigative officers rated the ETA system highly across all dimensions, with user satisfaction receiving the highest mean score ($M=4.40$, $SD=0.46$). Data accessibility also received strong endorsement ($M=4.36$, $SD=0.48$), reflecting the effectiveness of the system's database integration and querying capabilities. Retrieval efficiency ($M=4.12$, $SD=0.52$) was slightly lower than other dimensions but still indicates substantial improvement relative to previous manual or semi-digital workflows. Analytical accuracy scored a mean of 4.24, suggesting that the system reliably processes complex investigative data and produces outputs that investigators perceive as accurate and actionable.

Relationship Between System Dimensions

To examine potential associations between system performance dimensions, a Pearson correlation analysis was conducted (Table 2). This analysis identifies whether improvements in one dimension, such as data accessibility, are associated with higher ratings in other dimensions, like user satisfaction or retrieval efficiency.

Table 2: Pearson Correlation Between System Performance Dimensions

Dimension	Data Accessibility	Retrieval Efficiency	Analytical Accuracy	User Satisfaction
Data Accessibility	1	.71**	.68**	.75**
Retrieval Efficiency	.71**	1	.64**	.70**
Analytical Accuracy	.68**	.64**	1	.72**
User Satisfaction	.75**	.70**	.72**	1

Note. $p < .01$

The correlation analysis indicates strong positive relationships across all system dimensions, with coefficients ranging from 0.64 to 0.75. These results suggest that improvements in one area, such as data accessibility, are strongly associated with enhancements in other areas, including user satisfaction and analytical accuracy. The interdependence of these dimensions highlights the holistic benefits of the ETA system: a well-designed interface that improves accessibility simultaneously supports accurate analysis and overall satisfaction.

The findings demonstrate that the ETA System Prototype substantially enhances criminal profile data management capabilities within the investigative context. High ratings in data accessibility suggest that the centralized database and structured ETL processes allow officers to retrieve case information quickly and without fragmentation, addressing a major limitation of legacy systems. The retrieval efficiency score, while slightly lower than accessibility, indicates that the system supports faster query execution and reduces the time spent manually collating records, which is crucial in high-volume investigative environments.

Analytical accuracy results show that officers perceive the system's output as reliable and relevant, confirming that the ETL pipeline and query functions are effectively standardizing and processing complex datasets. Strong correlations between analytical accuracy, data accessibility, and user satisfaction indicate that users' perception of data reliability contributes directly to their overall confidence in the system. User satisfaction, being the highest-rated dimension, underscores that the ETA system not only meets technical requirements but also aligns with end-user expectations, promoting adoption and sustained use.

These results collectively indicate that the ETA System Prototype supports enhanced investigative intelligence by integrating data from multiple sources, enabling rapid retrieval, and producing consistent analytical outputs. The strong interrelationships among system dimensions imply that improvements in technical features reinforce user perception and satisfaction, which is consistent with established findings in information systems adoption literature (Ismail et al., 2025; Muhammad, 2022). Overall, the prototype demonstrates measurable operational benefits, providing a robust foundation for future enhancements, including AI-driven analytics, predictive modeling, and automated anomaly detection.

Discussion and Recommendations

The findings of this study reveal that the ETA (Enhanced Tracking and Analysis) System Prototype significantly advances criminal profile data management within the Malaysian Anti-Corruption Commission (MACC), addressing the research objectives of improving data accessibility, retrieval efficiency, analytical accuracy, and user satisfaction. All four dimensions received high performance scores from participating investigative officers, indicating that the ETA system enhances operational capacity relative to legacy methods. Data accessibility scored highly, reflecting the effectiveness of centralized data consolidation and structured ETL (Extract, Transform, Load) processes in supporting rapid retrieval and cross-case analysis. Retrieval efficiency also demonstrated meaningful improvements, affirming that the system reduces time spent on manual data searching and accelerates investigation workflows. Analytical accuracy was rated positively, suggesting that the system's standardized

processing enables investigators to derive reliable insights from complex, multi-source data. Finally, user satisfaction was the highest-rated outcome, emphasizing that the system not only meets technical requirements but also aligns with officers' expectations and practical needs. Collectively, these results indicate that the ETA prototype achieves its primary objectives of modernizing MACC's data management ecosystem and strengthening investigative intelligence.

These findings are significant in the broader context of digital transformation for anti-corruption enforcement. A recent meta-analysis found that higher levels of digitalization and e-government adoption are strongly correlated with reductions in government corruption, illustrating the potential for technology to enhance accountability and transparency (Munshi & Manni, 2025). This supports the empirical implications of the ETA system, which leverages structured databases and analytics to reduce analytical errors and accelerate investigations. Moreover, the strong correlations observed between system performance dimensions suggest that improvements in technical infrastructure cascade into enhanced user experience and perceived system value, a relationship underscored in information systems research on technology acceptance and adoption frameworks (Ismail et al., 2025; Muhammad, 2022). The integration of advanced data processing and AI-ready architecture within the ETA prototype therefore aligns with broader digital governance strategies that empower investigators with automated analytical tools while minimizing reliance on manual reporting.

From a Sustainable Development Goals (SDGs) perspective, this study contributes directly to Goal 16: Peace, Justice, and Strong Institutions, particularly targets aimed at reducing corruption and strengthening accountability and transparency in public institutions (SDG 16.5 & 16.6). SDG 16 emphasizes the need for effective, accountable, and transparent institutions that uphold the rule of law and combat corruption, which resonates with MACC's strategic objectives and the functional outcomes demonstrated by the ETA system. By improving investigative efficiency and data integrity, the ETA prototype supports institutional capacities that align with international efforts—as promoted by the United Nations and partner agencies such as UNODC—to improve data quality and enforcement mechanisms for corruption reduction and justice processes. Enhanced digital systems such as ETA also contribute to evidence-based policy and operational decision-making, which has been identified as crucial for strategic anti-corruption research and digital governance implementation in the Malaysian landscape (Azam Baki, 2025).

In conclusion, the ETA System Prototype demonstrates measurable improvements across core operational dimensions of criminal profile data management. The study's results confirm that structured ETL, centralized databases, and analytics-ready architecture can enhance investigative workflows, reduce redundancy, and support accurate intelligence outputs. These outcomes are supported by the literature indicating that digital transformation plays a pivotal role in reducing corruption levels and strengthening institutional performance (Munshi & Manni, 2025). Therefore, the ETA prototype not only fulfills its research objectives but also contributes to advancing MACC's strategic alignment with SDG 16, reinforcing institutional accountability and digital governance.

Based on the findings, future work should prioritize full deployment of the ETA system across MACC units, including integration with other enforcement databases to enhance interoperability and real-time data sharing. Introduction of advanced AI-driven modules such as predictive modelling and anomaly detection can further strengthen investigative intelligence and support proactive corruption detection. Capacity-building efforts, including specialized training for investigators on data analytics and system utilization, would also enhance adoption and ensure sustained system effectiveness. Finally, continued evaluation and refinement of data governance policies will be critical to safeguard data security, privacy, and ethical use of analytics within the enforcement context...

References

1. Apriyanto, T., & Sara, R. (2024). The role of blockchain technology in improving data security and integrity in the law enforcement process of corruption crimes. In *Proceedings of the 4th International Conference on Law, Social Sciences, Economics, and Education (ICLSSEE 2024)*. EAI. <https://doi.org/10.4108/eai.25-5-2024.2349443>
2. Azam Baki. (2025). Strategic research on digital governance.
3. Bernama. (2025). MACC strengthens anti corruption efforts with AI and digital systems. *The Star*.
4. Braekt, R. I. d., Le Khac, N. A., Farina, J., Scanlon, M., & Kechadi, M. T. (2017). Increasing digital investigator availability through efficient workflow management and automation. *arXiv*.

5. Crime and Criminal Tracking Network and Systems (CCTNS). (n.d.). Integrated policing data system.
6. ERA Forum. (2022). Quality of data sets that feed AI and big data applications for law enforcement (Martyna Kusak).
7. Ismail, E. D., Ismail, N. A. A., & Yusof, M. S. M. (2025). Mapping SMEs' intention to adopt IS technology: Theories & frameworks.
8. Kharuddin, A. F., & Ismail, N. A. (2017). Graphing Calculator Exposure of Mathematics Learning in a Partially Technology Incorporated Environment. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(6), 2529-2537. <https://doi.org/10.12973/eurasia.2017.01238a>
9. Kharuddin, A. F., Ismail, K., Ghazali@Suhaimi, S. S., Abu Bakar, N., & Mokhtar, K. I. (2018). Latent Class Analysis of Dental Treatment: Observed in IIUM Dental Specialist Clinic, Kuantan, Pahang. *IIUM Medical Journal Malaysia*, 17(2). <https://doi.org/10.31436/imjm.v17i2.976>
10. Lukmanjaya, W., Halmich, C., Butler, T. et al. (2026). Computational text analysis on unstructured police data: a scoping review.
11. Mandalapu, V., Elluri, L., Vyas, P., & Roy, N. (2023). Crime prediction using machine learning & deep learning review.
12. MDPI. (2021). Interoperability Enhanced Knowledge Management in Law Enforcement. Information.
13. Mitar, M. (2004). Data quality as a challenge to modern policing and criminal justice. Office of Justice Programs.
14. Muhammad, N. K. (2022). A conceptual framework for big data analytics adoption in Industry 4.0. Atlantis Press Proceedings.
15. Muhammad, N. K., Osman, N. H., & Salleh, N. A. (2021). A conceptual framework for Big Data Analytics adoption in Industry 4.0.
16. Munshi, A., & Manni, U. H. (2025). Can digital transformation reduce government corruption? A meta analysis. *Quantitative Economics and Management Studies*.
17. OECD. (2025). Governing with artificial intelligence: The state of play and way forward in core government functions. OECD Publishing. <https://doi.org/10.1787/795de142-en>
18. Palanisamy, R., Norman, A. A., & Kiah, M. L. M. (2025). Research methodology. In *Bring Your Own Device Security Policy Compliance Framework* (pp. 1–18). Springer Nature.
19. Rahmafillah, R., Putri, R. Y. A., & Sartika, R. (2025). The effectiveness of anti corruption policies in the digital age: A literature review on the use of technology in the eradication of corruption. *Hermeneutika: Jurnal Ilmu Hukum*, 9(2). <https://doi.org/10.33603/hermeneutika.v9i2.10043>
20. The Star. (2025). MACC strengthens anti graft drive with AI, digital systems.
21. United Nations Interregional Crime and Justice Research Institute (UNICRI). (2023). SDG 16 – Through a digital lens. UNICRI Report.
22. United Nations. (2015). Sustainable Development Goal 16. UN Sustainable Development. <https://sdgs.un.org/goals/goal16>
23. VALCRI. (n.d.). Visual Analytics for Sense making in Criminal Intelligence Analysis.
24. Williams, B. (2023). Mixed methods research design, data collection, and analysis. Insight7.