LICS PROCEEDINGS OF LEGAL INTERNATIONAL AND STUDIES
No. ISSN : 2986-4089
Published by Universitas Islam Sultan Agung, Semarang, INDONESIA

LICS 2026
Topic: *Strengthening Legal Frameworks for Combating Human Trafficking and Enhancing Protection of Women and Children*

The Role of Artificial Intelligence in Identifying and Preventing Human Trafficking: Legal and Ethical Challenges

Joman Rabah Alkatib

Faculty of Law, Universitas Islam Sultan Agung (UNISSULA) Semarang, Indonesia, E-mail: advjoman@gmail.com

Abstract. *Human trafficking represents one of the most serious global human rights violations in the contemporary world, affecting millions of individuals annually. The emergence of artificial intelligence technologies has introduced novel opportunities for identifying, preventing, and prosecuting human trafficking cases. However, the implementation of AI-driven anti-trafficking solutions simultaneously presents profound legal and ethical challenges. This paper provides a comprehensive legal analysis of the role of artificial intelligence in combating human trafficking, examining both the beneficial applications and the significant challenges related to data privacy, algorithmic bias, legal frameworks, and fundamental human rights. Through comparative analysis of existing regulatory frameworks including the General Data Protection Regulation (GDPR), the EU AI Act, and international legal instruments, this research demonstrates that while AI technology offers unprecedented investigative capabilities, its deployment must be carefully regulated to ensure compliance with international human rights law and data protection standards. The paper concludes that effective anti-trafficking strategies require a balanced approach that harnesses AI's investigative potential while maintaining robust legal safeguards and ethical governance mechanisms.*

Keywords: *Artificial Intelligence; Algorithmic Bias; Data Privacy; Human Trafficking; Human Rights Law; Legal Frameworks.*

1. Introduction

The advancement of artificial intelligence (AI) technology has fundamentally transformed multiple sectors of modern society, ranging from healthcare to criminal justice. In recent years, law enforcement agencies, non-governmental organizations (NGOs), and international bodies have increasingly turned to

artificial intelligence tools as potential solutions to combat human trafficking a crime that the United Nations estimates affects over 27 million people globally at any given time. The introduction of machine learning algorithms, natural language processing, facial recognition systems, and data analytics platforms has created unprecedented opportunities for identifying victims, detecting trafficking networks, and prosecuting perpetrators. However, the deployment of AI in anti-trafficking contexts raises significant questions about the legal permissibility and ethical implications of such technologies. The rapid adoption of AI-driven surveillance systems, algorithmic victim identification mechanisms, and automated reporting systems has outpaced the development of comprehensive legal frameworks to govern their use. This creates a critical gap between technological capability and legal accountability. The primary objective of this paper is to provide a detailed legal and ethical analysis of artificial intelligence applications in identifying and preventing human trafficking. Specifically, this research addresses: (1) the current applications of AI technologies in anti-trafficking operations; (2) the existing legal frameworks governing AI implementation; (3) the ethical challenges and risks associated with algorithmic decision-making in this context; (4) privacy and data protection concerns under international law; and (5) recommendations for balanced and accountable AI implementation. Human trafficking constitutes a severe violation of fundamental human rights and is recognized internationally as a transnational organized crime and a form of modern slavery. The United Nations Protocol to Prevent, Suppress and Punish Trafficking in Persons defines human trafficking as 'the recruitment, transportation, transfer, harboring or receipt of persons, by means of threat or use of force or other forms of coercion for the purpose of exploitation'. This definition encompasses various forms of trafficking, including sexual exploitation, forced labor, organ harvesting, and debt bondage. The scale of human trafficking remains staggering despite international efforts to combat it. According to the most recent United Nations Office on Drugs and Crime (UNODC) Global Report on Trafficking in Persons, approximately 27.3 million people are estimated to be in situations of trafficking in persons at any given time. Women and girls comprise approximately 72% of detected trafficking victims, while children represent an increasing proportion of victims. The economic value of human trafficking is estimated at $150 billion annually, making it one of the most profitable forms of organized crime. Traditional law enforcement approaches to combating human trafficking have proven insufficient in addressing the scale and complexity of the problem. Investigators face challenges in identifying victims, especially when trafficking operations leverage digital platforms for recruitment and exploitation. The digitalization of trafficking operations has created new investigative demands that conventional forensic and investigative methodologies cannot adequately address. This

context has prompted the exploration of artificial intelligence as a potentially transformative tool for victim identification, perpetrator detection, and network analysis.

2. Research Methods

This study utilizes a legal normative research method with a comparative legal and ethical analysis approach. The research examines existing regulatory frameworks governing artificial intelligence, data protection, and human rights across multiple jurisdictions, including the General Data Protection Regulation (GDPR), the European Union Artificial Intelligence Act (EU AI Act), and key international human rights instruments. Additionally, the study incorporates qualitative document analysis of international law enforcement reports, academic literature, and case law from regional human rights courts to evaluate the operational capabilities, legal constraints, and ethical risks of AI deployment in anti-trafficking efforts.

3. Results and Discussion

3.1. Artificial Intelligence Technologies in Anti-Trafficking Operations

Natural language processing (NLP) technologies enable AI systems to analyze vast volumes of textual data to identify trafficking indicators and patterns (Li et al., 2023). Law enforcement agencies have implemented NLP-based systems to analyze online advertisements, social media communications, and messaging platforms to detect trafficking patterns and terminology commonly used by traffickers (Musto & Boyd, 2014). These systems can identify suspicious phrases, detect coded language used to advertise sexual services or labor opportunities, and analyze contextual information to flag potentially trafficking-related communications (Musto et al., 2020). The utility of NLP in anti-trafficking operations was demonstrated through the analysis of online escort advertisements, where machine learning algorithms successfully identified indicators suggestive of human trafficking with greater accuracy than manual review alone (Li et al., 2023).

In addition to text analytics, computer vision technologies, including facial recognition systems, have been proposed as mechanisms for identifying trafficking victims through analysis of online images (FRA, 2020). Law enforcement and child protection organizations have utilized facial recognition to identify victims in sexually exploitative material and to track potential trafficking victims across online platforms (NCMEC, 2024). By cross-referencing visual metadata and facial biometric signatures against missing persons databases,

computer vision accelerates the detection of victims who would otherwise remain unidentified in vast online archives (FRA, 2020).

Beyond individual victim identification, advanced data analytics and graph-based algorithms enable the identification of complex trafficking networks through analysis of financial transactions, travel patterns, communication networks, and commercial activities (Nair et al., 2024). These network mapping capabilities allow investigative bodies to construct relational graphs between perpetrators, shell corporations, and recruitment front agencies (UNODC, 2024). By uncovering hidden links within high-volume financial data, graph learning algorithms assist in disrupting the financial infrastructure of organized criminal syndicates (Nair et al., 2024).

Predictive AI systems have also been developed to assess vulnerability to trafficking and identify high-risk locations or populations by analyzing demographic data, socioeconomic factors, migration patterns, and historical trafficking incidents (Lefebvre & Benyekhlef, 2023). These machine learning tools generate spatial-temporal heatmaps that assist international relief organizations and border control units in allocating law enforcement resources efficiently (US Department of State, 2024). Through risk-scoring mechanisms, predictive platforms aim to intercept recruitment attempts prior to victim exploitation (Lefebvre & Benyekhlef, 2023).

The integration of multi-modal AI systems—combining NLP, computer vision, and network analysis—enables real-time intelligence gathering across transnational jurisdiction boundaries (Tech Against Trafficking, 2024). Digital evidence collected across border management databases and commercial web scraping platforms can be aggregated into unified investigative dashboards (Leslie, 2019). This automated cross-verification drastically reduces the operational timeframe required for multi-agency joint task forces to execute rescue interventions (Tech Against Trafficking, 2024).

However, the rapid adoption of automated tools in intelligence-led policing introduces operational dependencies on proprietary commercial software (Musto et al., 2020). Private technology firms frequently deploy closed-source algorithms in law enforcement partnerships, limiting independent auditing and public oversight regarding technical reliability (Leslie, 2019). Consequently, investigative findings derived from algorithmic processing face scrutiny when admitted as evidence in criminal court proceedings (Lefebvre & Benyekhlef, 2023).

Ultimately, while these advanced technologies enhance operational efficiency in combating human trafficking, their deployment must be continuously evaluated

against legal standards of evidence and procedural due process (Gallagher, 2010). Technological capabilities alone cannot replace comprehensive legal strategies and survivor-centered justice frameworks (Musto & Boyd, 2014). Thus, the application of AI tools in criminal investigations requires ongoing statutory refinement and judicial scrutiny to ensure operational efficacy while maintaining human rights compliance (European Commission, 2024).

3.2. Legal Framework Governing AI Implementation in Anti-Trafficking Efforts

The European Union's General Data Protection Regulation (GDPR) represents the most comprehensive legal framework governing the processing of personal data (Regulation (EU) 2016/679). The GDPR establishes fundamental principles that are directly applicable to AI systems processing personal data in anti-trafficking contexts, including data minimization, purpose limitation, transparency, and fairness (Regulation (EU) 2016/679). However, Article 9(2) restricts the processing of special categories of personal data, such as biometric or sensitive health data, and Article 22 provides data subjects with explicit rights against being subjected to decisions based solely on automated processing and profiling, presenting distinct legal challenges for AI-driven victim-identification tools (Regulation (EU) 2016/679).

Building upon data protection principles, the European Union Artificial Intelligence Act (EU AI Act) classifies AI systems used in law enforcement and the protection of vulnerable persons as 'high-risk' under Article 6(2) (Regulation (EU) 2024/1689). High-risk systems deployed in anti-trafficking operations must comply with extensive statutory requirements prior to market entry and operational deployment (European Commission, 2024). These mandatory conditions include continuous risk management protocols, rigorous technical documentation, operational transparency, robust human oversight mechanisms, and high standards of accuracy, resilience, and cybersecurity (Regulation (EU) 2024/1689).

Beyond regional European Union legislation, fundamental rights safeguards are rooted in binding international human rights instruments (Gallagher, 2010). Article 17 of the International Covenant on Civil and Political Rights (ICCPR) and Article 8 of the European Convention on Human Rights (ECHR) protect individuals against arbitrary or unlawful interference with their privacy and family life (ICCPR, 1966). Jurisprudence from international tribunals emphasizes that surveillance technologies and automated data collection mechanisms operated by state authorities must satisfy strict tests of legality, necessity, and proportionality to prevent rights infringements (Big Brother Watch v. UK, 2021).

In contrast to the EU's unified statutory approach, jurisdictions such as the United States employ a fragmented, sectoral approach to privacy and AI

governance (Blueprint for an AI Bill of Rights, 2022). While federal frameworks like the Children's Online Privacy Protection Act (COPPA) protect specific target demographics, comprehensive federal statutory rules governing law enforcement AI applications remain absent (COPPA, 1998). This sectoral division creates notable regulatory gaps and legal uncertainty regarding data retention, cross-agency data sharing, and algorithmic accountability in transnational investigations (Blueprint for an AI Bill of Rights, 2022).

Emerging AI governance frameworks in the Asia-Pacific region, such as Singapore's Model AI Governance Framework, offer alternative, soft-law approaches to technological regulation (PDPC, 2020). These regional guidelines emphasize ethical principles, risk self-assessment, and economic innovation, yet often lack enforceable statutory mandates for victim protection in criminal justice contexts (PDPC, 2020). As a result, anti-trafficking operations spanning multiple Asian jurisdictions operate under varying degrees of legal oversight and data privacy protection (UNODC, 2024).

The divergence between global legal regimes complicates international law enforcement cooperation and joint anti-trafficking task forces (UNODC, 2024). Cross-border intelligence sharing involving AI-generated data creates jurisdiction-specific legal liabilities, particularly when evidence collected in a country with lenient privacy laws is introduced in courts bound by strict data protection mandates like the GDPR (Regulation (EU) 2016/679). Harmonizing these legal standards is therefore essential to maintain the admissibility of digital evidence while preserving human rights compliance (Gallagher, 2010).

Ultimately, establishing an effective legal framework for AI in anti-trafficking efforts requires balancing law enforcement capabilities with fundamental constitutional rights (Leslie, 2019). Legislative bodies must ensure that statutory provisions evolve alongside technological advancements without diluting core legal safeguards against arbitrary surveillance (Human Rights Watch, 2024). Regulating AI tools through clear statutory oversight, mandatory impact assessments, and strict judicial review remains critical to preventing power abuses while pursuing justice for trafficking victims (European Commission, 2024).

3.3. Ethical Challenges, Algorithmic Bias, and Privacy Rights

Algorithmic bias represents one of the most serious ethical challenges in AI-driven anti-trafficking systems (Leslie, 2019). Bias frequently originates from training datasets that reflect historical disparities, selective policing, and systemic discrimination embedded in law enforcement practices (Musto et al., 2020). A critical ethical concern involves the risk that AI systems may conflate consensual sex work with human trafficking, resulting in the wrongful misidentification,

harassment, and criminalization of sex workers (Musto & Boyd, 2014). Furthermore, predictive analytics systems can create a disparate impact on marginalized populations by targeting individuals based on demographic variables, race, and socioeconomic status rather than objective trafficking indicators (Madaio et al., 2020).

From a privacy perspective, the deployment of AI-driven surveillance tools raises mass surveillance risks and severe concerns regarding the principle of proportionality under international human rights law (Human Rights Watch, 2024). Mass data collection mechanisms often process vast volumes of personal data far exceeding what is strictly necessary, directly violating data minimization principles established in privacy frameworks (Regulation (EU) 2016/679). Additionally, secondary data sharing between anti-trafficking agencies, immigration enforcement, and labor monitoring bodies undermines victim trust, discourages survivors from reporting abuse, and breaches statutory purpose limitation safeguards (Musto et al., 2020).

To mitigate these ethical and privacy risks, AI anti-trafficking systems must prioritize victim-centered design principles (Tech Against Trafficking, 2024). Involving survivor-led organizations, human rights advocates, and multidisciplinary legal experts in all stages of system design, development, and deployment ensures that automated tools prioritize victim protection over punitive enforcement (Madaio et al., 2020). Survivor-centered input helps developers identify subtle operational risks and prevents technological interventions from inadvertently inflicting secondary trauma on vulnerable populations (Human Rights Watch, 2024).

Addressing structural flaws within algorithmic models requires rigorous bias assessment and mitigation protocols (Madaio et al., 2020). Organizations deploying AI tools must conduct continuous audits of training datasets and monitor system performance across diverse demographic groups to detect discriminatory outputs (Leslie, 2019). Establishing standardized metrics for fairness, accuracy, and error rates ensures that predictive algorithms do not reinforce historical prejudices or disproportionately target marginalized communities during investigations (Madaio et al., 2020).

Operational integrity further depends on maintaining human oversight and accountability mechanisms across all decision-making levels (Regulation (EU) 2024/1689). All legal and operational determinations significantly affecting individuals—such as victim identification, surveillance authorization, or criminal profiling—must remain subject to meaningful human review (FRA, 2020). AI models should operate strictly as decision-support instruments rather than autonomous law enforcement agents, preserving human agency and procedural due process (Leslie, 2019).

From a technical architecture standpoint, adopting privacy-preserving technologies is essential to safeguard sensitive victim data (Leslie, 2019). Incorporating federated learning, differential privacy, and homomorphic encryption allows law enforcement and non-profit organizations to analyze patterns across decentralized databases without exposing underlying personal identifying information (Tech Against Trafficking, 2024). These cryptographic techniques minimize centralized data repositories, reducing the potential impact of data breaches and unauthorized surveillance (Regulation (EU) 2016/679).

Finally, achieving long-term ethical sustainability requires systemic transparency, public accountability, and international legal harmonization (European Commission, 2024). Government bodies and private developers must publish comprehensive algorithmic impact assessments and subject their technologies to independent oversight committees (OSTP, 2022). Concurrently, international organizations such as the United Nations must establish binding legal standards to harmonize global AI governance in criminal justice (UNODC, 2024). Aligning technological capabilities with international human rights frameworks ensures that AI serves as a lawful, accountable, and effective tool in eradicating human trafficking worldwide (Gallagher, 2010).

4. Conclusion

Artificial intelligence technology presents both extraordinary opportunities and significant risks in human trafficking prevention and victim identification. While machine learning, NLP, and data analytics offer unprecedented capabilities to disrupt trafficking networks, their deployment must occur within robust legal and ethical frameworks that protect fundamental rights, safeguard data privacy, and prevent systemic bias. Existing regulations such as the GDPR and EU AI Act establish crucial baselines, but global legal harmonization remains necessary. Moving forward, anti-trafficking initiatives must adopt a balanced, survivor-centered approach that pairs technological capabilities with human oversight, legal accountability, and strict adherence to international human rights standards.

5. References

Journals:

Lefebvre, G., & Benyekhlef, K. (2023). Predictive Policing in Canada. International Review of Penal Law, 94, 87.

Li, R., et al. (2023). Detecting Human Trafficking: Automated Classification of Online Customer Reviews of Massage Businesses. Manufacturing & Service Operations Management, 25(3), 1051.

Musto, J., & Boyd, D. (2014). The Trafficking-Technology Nexus. Social Politics, 21(3), 461.

Musto, J., Thakor, M., & Gerasimov, B. (2020). Between Hope and Hype: Critical Evaluations of Technology's Role in Anti-Trafficking. Anti-Trafficking Review, 14, 1.

Books:

Gallagher, A. T. (2010). The International Law of Human Trafficking. Cambridge University Press.

Leslie, D. (2019). Understanding Artificial Intelligence Ethics and Safety. Alan Turing Institute.

Internet:

European Commission. (2024). Regulatory Framework Proposal on Artificial Intelligence. European Commission.

Human Rights Watch. (2024, July 1). Using AI to Fight Trafficking Is Dangerous. Human Rights Watch. https://www.hrw.org/news/2024/07/01/using-ai-fight-trafficking-dangerous

Regulation:

Children's Online Privacy Protection Act, 15 U.S.C. §§ 6501-6506 (1998).

Protocol to Prevent, Suppress and Punish Trafficking in Persons, Especially Women and Children, Nov. 15, 2000, 2237 U.N.T.S. 319.

Regulation (EU) 2016/679 (General Data Protection Regulation), 2016 O.J. (L 119) 1.

Regulation (EU) 2024/1689 (Artificial Intelligence Act), 2024 O.J. (L 2024/1689).

International Covenant on Civil and Political Rights, Dec. 16, 1966, 999 U.N.T.S. 171.