International Journal of Law Analytics (IJLA)
Vol. 4, No. 4 2026: 579-594

Reconstructing Legal Standards for Artificial Intelligence in Refugee Status Determination in Indonesia: Lessons from Canada, the United Kingdom, and UNHCR

Untung Setyardi
Universitas Atma Jaya Yogyakarta
Corresponding Author: Author name untung.setyardi@uajy.ac.id

A R T I C L E I N F O

Keywords: Artificial Intelligence, Algorithmic Accountability, Indonesia, Refugee Status Determination, Procedural Fairness

Received : 09, July
Revised : 11, September
Accepted: 13, November

©2026 Setyardi: This is an open-access article distributed under the termsof the Creative Commons Atribusi 4.0 Internasional.

A B S T R A C T

Artificial intelligence (AI) may improve refugee-status administration but can jeopardize individual assessment, confidentiality, and procedural fairness. This normative legal study examines gaps in Indonesia's framework and reconstructs safeguards through comparison with Canada, the United Kingdom, and UNHCR. Statutory, conceptual, comparative, and human-rights approaches are applied to legislation, institutional policies, and scholarly literature. The findings identify fragmented authority, absent sector-specific risk classification, inadequate safeguards against algorithmic influence, and uncertain remedies. Canada offers impact assessment and limits on adjudicative AI; the United Kingdom illustrates supervised administrative assistance; and UNHCR supplies protection-centered procedural guarantees. The article proposes a rights-based, risk-tiered model that confines AI to accountable decision support, requires audits and meaningful human review, and preserves individualized protection.

DOI: https://doi.org/10.59890/ijla.v4i4.375
E-ISSN: 3032-4041
https://slamultitechpublisher.my.id/index.php/ijla

INTRODUCTION

Digital transformation has altered how public institutions collect evidence, organize administrative records, and assist officials in making consequential decisions. Artificial intelligence (AI) can summarize complex material, retrieve relevant information, identify patterns, and support decisions at a speed difficult to reproduce through manual processes. These opportunities are particularly attractive in migration administration, where agencies confront multilingual documents, dispersed country information, and resource constraints. Yet computational efficiency and legal legitimacy are different qualities. A technically plausible output can still be inaccurate, discriminatory, unexplainable, or incompatible with the rights of the person concerned. Beduschi (2021) therefore places technological innovation in migration governance alongside questions about human rights, institutional accountability, and the distribution of authority. Those questions become especially pressing when an algorithm influences the assessment of international protection needs.

Refugee status determination (RSD) differs from ordinary document processing. It evaluates whether an individual faces a well-founded fear of persecution and meets the relevant conditions for international protection, while requiring attention to the person's particular experiences and the circumstances of the country of origin. Accounts may contain gaps arising from trauma, displacement, translation difficulties, or limited access to documentary evidence. A machine-generated judgment about narrative consistency cannot safely substitute for contextual legal evaluation. Kinchin and Mougouei (2022) explore both the potential and limitations of AI in RSD, while Amoroso (2024) emphasizes the need to evaluate its compatibility with international refugee law. The central challenge is not whether AI can process refugee-related information, but under what legally defensible conditions it may influence an assessment that can affect personal safety.

AI can serve several distinct purposes along the RSD process. Relatively low-risk assistance may include organizing files, identifying duplicate records, producing preliminary translations, or locating country-of-origin information. More consequential uses include prioritizing applications, recommending credibility findings, ranking protection risks, or suggesting outcomes. These activities should not be grouped together merely because they involve the same technological label. A summarization error can omit a crucial event; a retrieval tool can privilege outdated country material; and a risk model can turn proxies for nationality, route of travel, or language into unjustified generalizations. Research on digital evidence and algorithmic border control demonstrates why reliability, transparency, and contextual verification matter (Byrne & Gammeltoft-Hansen, 2024; Forti, 2024). The appropriate legal response therefore depends upon function and likely effect, not only the sophistication of software.

Indonesia provides a distinctive institutional setting. It is not a party to the 1951 Refugee Convention or the 1967 Protocol and has no comprehensive national RSD adjudication system. UNHCR conducts mandate-based RSD, whereas Indonesian authorities carry out functions associated with the handling, accommodation, security, and immigration supervision of refugees and asylum seekers (UNHCR Indonesia, n.d.). This allocation matters because measures taken by domestic agencies may affect the information environment and practical circumstances of a protection claim without conferring upon those agencies the mandate to decide it. An AI tool introduced for registration, information exchange, or administrative support can accordingly have consequences beyond the formal competence of the institution that deploys it. The legal question extends from ownership of the software to responsibility for each actor's contribution to a potentially harmful decision.

Indonesia already has relevant legal foundations, but they do not constitute an integrated AI-in-RSD framework. Law No. 37 of 1999 on Foreign Relations addresses asylum and refugee-related policy; Law No. 39 of 1999 on Human Rights provides general protection of individual rights; and Presidential Regulation No. 125 of 2016 regulates the handling of refugees from abroad. Law No. 27 of 2022 on Personal Data Protection supplies rules concerning processing, significant automated decisions, and high-risk data processing. Ministerial Circular No. 9 of 2023 expresses general AI ethics principles. None of these instruments, however, directly specifies how algorithmic outputs may be used in mandate-based RSD, how the roles of national institutions and UNHCR should be differentiated, or what procedural guarantees are triggered when AI substantially influences an individual's protection assessment.

The resulting research gap is more precise than an assertion that Indonesia has no AI regulation at all. The country possesses dispersed protections governing human rights, public functions, and personal data, but lacks an operational bridge connecting those protections to refugee-specific procedural risks. Existing scholarship has illuminated separate dimensions of the issue: Beduschi (2021) examines AI in migration management; Byrne and Gammeltoft-Hansen (2024) analyze digital evidence in RSD; Forti (2024) addresses algorithmic error in border procedures; and Mutaqin et al. (2024) consider AI in the Rohingya humanitarian context. These contributions do not, individually, develop a sector-sensitive legal design addressing institutional competence, admissible AI functions, procedural safeguards, and remedies within Indonesia's particular allocation of RSD responsibilities. The present article integrates these dimensions into a proposed model of algorithmic accountability grounded in refugee protection.

Comparative material assists this reconstruction without implying that foreign institutions are interchangeable. Canada's Directive on Automated Decision-Making supplies a general public-administration framework for assessing algorithmic impact, while Canada's Immigration and Refugee Board (IRB) has expressly distinguished non-adjudicative AI assistance from adjudicative decision-making (Treasury Board of Canada Secretariat, 2025; Immigration and Refugee Board of Canada [IRB], 2026a). The United Kingdom's Home Office has evaluated AI tools that summarize asylum interviews and retrieve policy information while retaining responsibility for decisions with human caseworkers (Home Office, 2025). UNHCR's procedural standards and responsible AI approach place the fairness of individual assessment, confidentiality, and protection duties at the center of technology governance (UNHCR, 2020, 2025). Taken together, these frameworks supply complementary lessons rather than a ready-made foreign template.

This article addresses two linked questions: what normative, procedural, and institutional weaknesses would arise if AI were used to support refugee-status processes connected with Indonesia; and what legal standards could be reconstructed from the comparative experience of Canada, the United Kingdom, and UNHCR? It argues that the most defensible approach is a risk-tiered, rights-based decision-support framework. Under that framework, AI may assist designated functions subject to prior assessment and verification, but cannot independently determine refugee status or displace the competence of authorized decision-makers. The contribution lies in translating general AI principles into specific enforceable duties appropriate to the refugee-protection context, including explainability, auditability, data safeguards, meaningful human control, and accessible corrective procedures.

THEORETICAL REVIEW

Procedural Fairness and Refugee Protection

RSD combines evidence assessment with prospective evaluation of harm. The 1951 Convention and 1967 Protocol provide the principal international reference points for understanding refugee eligibility, although Indonesia has not ratified either instrument. The principle of non-refoulement and wider human-rights obligations remain central to assessing the consequences of state action affecting vulnerable persons. UNHCR's mandate-based procedures recognize that an applicant should be able to communicate relevant facts, obtain appropriate interpretation, understand the reasons for a negative determination, and use available review procedures (UNHCR, 2020). These are not ancillary conveniences: they help prevent factual and interpretive mistakes from becoming irreversible protection failures. Any AI-enabled administrative system must therefore be assessed by whether it preserves these practical opportunities, not merely by whether it produces consistent outputs.

Procedural fairness also entails a meaningful relationship between reasons and evidence. Decision-makers must be able to identify the material relied upon, distinguish verified facts from inferences, and explain why an applicant's account does or does not meet the applicable legal standard. Refugee narratives can be affected by memory disruption, cultural communication differences, or omissions in translation. A generic consistency score is not a substitute for assessing those circumstances. As Kinchin and Mougouei (2022) suggest, technological tools warrant careful scrutiny when they are invited to perform tasks that appear objective but depend upon normative assumptions. The relevant standard should be humanly intelligible justification: an applicant must have a genuine opportunity to challenge a material factual proposition even when the proposition originated in software.

Algorithmic Accountability and Human Oversight

Algorithmic accountability requires identifiable responsibility across the life cycle of an AI system: design, procurement, configuration, operation, verification, and correction. Systems are not legal decision-makers merely because they generate recommendations. The institution legally empowered to act remains responsible for the manner in which such recommendations are used, while developers and vendors may bear separate obligations concerning security, testing, documentation, and contractual compliance. Forti (2024) explains how algorithmic opacity can impede meaningful challenge in border governance. An accountability model must therefore demand traceable inputs, documented purposes, evaluation of foreseeable errors, and review procedures that can identify where a harmful inference entered the administrative chain.

The existence of a human signatory does not by itself establish adequate oversight. Officials may defer to systems perceived as neutral or technically superior, especially when workloads are high. Meaningful human oversight instead requires sufficient training, access to underlying evidence, power to reject outputs, and a documented independent assessment. The distinction between decision support and automated decision-making is accordingly substantive rather than formal. If a nominally human decision simply reproduces an unexamined AI recommendation, the procedural danger remains. This study uses a protection-centered model in which accountability is evaluated through the actual ability to detect, explain, contest, and remedy mistakes, while respecting the separate institutional mandates involved in Indonesian refugee governance.

Risk-Based Regulation and Contextual Proportionality

A risk-based approach connects regulatory duties to likely harm. Administrative sorting, although not risk-free, should not automatically attract the same controls as algorithmic credibility assessment or outcome prediction. Conversely, a system described as an administrative assistant may become high-risk when its output determines which evidence is viewed or which applications receive priority. Canada's Algorithmic Impact Assessment illustrates the value of examining context, likely effects, mitigation, and oversight before deployment (Treasury Board of Canada Secretariat, n.d.). For refugee protection, that assessment must also address vulnerabilities specific to displacement, including sensitive identities, language diversity, transnational data exposure, and the possibility that mistakes affect safety rather than convenience. The theoretical synthesis is therefore risk proportionality bounded by non-negotiable procedural guarantees.

METHODOLOGY

This study employs normative legal research with descriptive-analytical and prescriptive objectives. It examines the adequacy of existing rules and derives proposed standards rather than testing software or measuring institutional performance. The statute approach considers Indonesia's Constitution, Laws No. 37 and No. 39 of 1999, Presidential Regulation No. 125 of 2016, Law No. 27 of 2022, and the general AI ethics circular. International refugee instruments and UNHCR procedural materials are analyzed as reference standards while recognizing differences between binding Indonesian law, international obligations, organizational guidance, and nonbinding policy. This distinction prevents the proposed framework from being presented as an already enacted national RSD regime.

The conceptual approach examines procedural fairness, non-refoulement, nondiscrimination, privacy, algorithmic transparency, and meaningful human oversight. Comparative legal analysis uses functionally relevant materials from Canada, the United Kingdom, and UNHCR rather than assuming institutional equivalence. Canada contributes general automated-decision governance and adjudicative limits; the United Kingdom contributes documented asylum-related administrative tools and verification practices; and UNHCR contributes standards for mandate RSD and humanitarian data protection. Comparison is organized around legal purpose, institutional responsibility, permissible technological function, prevention of error, and availability of corrective procedures (Siems, 2022).

Primary materials consist of statutes, regulations, international instruments, and authoritative institutional documents. Secondary materials include peer-reviewed articles and legal-method texts, collected through documentary and library research. These materials are organized into themes addressing authority, risk, data, review, and accountability. Qualitative analysis uses systematic interpretation and deductive legal reasoning to identify gaps and formulate a coherent prescriptive model. The study does not assume that Indonesia currently uses AI to make automated refugee-status determinations, does not audit a particular product, and makes no empirical claim about the

accuracy of AI systems in Indonesian RSD operations. Its conclusions concern the standards that should govern possible future or supporting deployments.

RESEARCH RESULTS

Fragmented Legal Authority and Procedural Gaps in Indonesia

The first finding is institutional fragmentation. Indonesia's non-accession to the Refugee Convention does not erase the need to protect persons who may face persecution, but it structures the domestic allocation of responsibilities differently from countries operating their own refugee adjudication systems. Under Presidential Regulation No. 125 of 2016, national functions concentrate on coordinated handling and immigration-related measures; UNHCR undertakes substantive mandate RSD. The two activities may interact through registration, interviews, documentary verification, accommodation, and information exchange. Yet no cross-institutional AI protocol in the examined Indonesian instruments identifies which institution may use which categories of AI output, who must verify shared information, or how errors are corrected when data move between institutions. Responsibility can thus become diffuse precisely when accountability should become more stringent (Rahayu et al., 2020; Setiyono, 2017).

Second, existing legal sources do not adequately distinguish AI functions by effect on legal rights. Presidential Regulation No. 125 of 2016 addresses refugee handling rather than algorithmic assessment. The 2023 ministerial ethics circular supplies general values such as transparency and accountability but does not define binding RSD-specific restrictions or a structured risk-assessment process. A file-categorization application, a machine translation of sensitive testimony, and a model recommending the rejection of a claim may therefore fall under the same broad policy language despite radically different consequences. The absence of differentiated standards creates both underregulation of risky uses and uncertainty for legitimate low-risk assistance. More precise rules should classify systems by what they do, the evidence they affect, and the foreseeable seriousness of error.

Third, protection against automated decision-making is only partial. Article 10 of Law No. 27 of 2022 recognizes the data subject's right to object to decisions based solely on automated processing, including profiling, where they produce legal consequences or significant effects. Article 34 requires a personal-data protection impact assessment for processing with a high potential risk to data subjects. These provisions establish important points of entry for AI governance. Their limits arise when a caseworker formally issues the final decision but depends heavily upon an algorithmic summary, ranking, or credibility recommendation. A process may no longer be solely automated in form while still being algorithmically determined in substance. The Indonesian framework lacks RSD-specific standards for evaluating that influence, proving independent judgment, or providing understandable reasons about its use.

Fourth, AI creates layered confidentiality and data-protection risks. Refugee files may contain religious or political affiliations, family relationships, medical information, experiences of violence, biometric identifiers, and details about persons in a country of origin. Article 16 of Law No. 27 of 2022 establishes general principles of lawful, limited, transparent, purpose-oriented processing. Nevertheless, applying these requirements to AI-mediated exchange between Indonesian agencies, UNHCR, interpreters, and technology providers requires clarity about controller responsibilities, permitted secondary uses, retention, model training, and cross-border exposure. In a refugee case, inadvertent disclosure may expose the applicant or relatives to reprisals. UNHCR's specialized data-protection policy highlights why generic operational convenience cannot justify unbounded reuse of such information (UNHCR, 2015).

Fifth, interpretive reliability cannot be inferred from fluent output. Automated transcription or translation can misrepresent idioms, identities, dates, and expressions of fear. A summary may suppress qualifications that materially affect credibility. Country-of-origin searches can retrieve stale documents or omit countervailing evidence. Digital photographs and communications may also require authentication and contextual interpretation before being used to support findings. Byrne and Gammeltoft-Hansen (2024) show why digital material must be appraised for evidentiary significance rather than accepted as self-validating. These problems become more serious when the system's error cannot be detected from the final answer alone. Indonesian instruments do not provide a refugee-specific requirement for source-linked AI outputs, human comparison with originals, or documentation of corrections.

Sixth, algorithmic bias can transform contingent characteristics into apparent evidence of untrustworthiness. Nationality, dialect, route of entry, disability, traumatic experience, or access to technology may correlate with unequal data quality without demonstrating any difference in entitlement to protection. Behavioral inference based on facial expression or tone is especially unsuitable as an independent marker of deception in culturally diverse and trauma-sensitive proceedings. General human-rights and nondiscrimination protections remain relevant, but RSD-specific technical testing, disaggregated impact review where appropriate, and restrictions on unreliable credibility proxies are not articulated in the examined framework. Forti (2024) cautions that error and opacity can combine in border processes, while Amoroso (2024) places refugee eligibility criteria above technological simplification.

Finally, procedural correction lacks a distinct AI dimension. Applicants may need to challenge an inaccurate translation, omitted fact, mistaken identity, fabricated citation, or improper risk inference before it influences a determination. National decisions and UNHCR mandate decisions are subject to different review routes; a single invented appeal mechanism would not respect this division. UNHCR's existing procedural arrangements include the communication of reasons and review of negative decisions, but Indonesian administrative rules do not yet specify how an applicant affected by AI-assisted data handling can discover the role of the system and obtain prompt correction.

The gap concerns not only access to a formal remedy but the practical possibility of identifying an algorithmic mistake in time to prevent harm (UNHCR, 2020; UNHCR Indonesia, n.d.).

Comparative Findings from Canada, the United Kingdom, and UNHCR

Canada provides two distinguishable lessons. First, the federal Directive on Automated Decision-Making and its Algorithmic Impact Assessment encourage ex ante consideration of system impact, transparency, quality assurance, and oversight in relevant administrative uses. These are general governance instruments, not a declaration that AI may decide refugee claims. Second, the IRB's 2026–2027 Departmental Plan explicitly rejects adopting AI for adjudicative decision-making while considering carefully governed uses in non-adjudicative operations. Its September 2026 practice notice also requires parties to verify AI-generated material, disclose certain uses, and avoid generating or materially altering personal evidence, such as narratives or affidavits (IRB, 2026a, 2026b). The central comparative lesson is therefore a combination of prior risk assessment and clear substantive boundaries around the integrity of refugee adjudication.

The United Kingdom illustrates controlled assistance within the asylum workflow. In pilots conducted during 2024 and evaluated in 2025, the Home Office examined Asylum Case Summarisation and Asylum Policy Search. The former summarized interview transcripts; the latter helped retrieve relevant country-policy material. The evaluation reported potential time savings but also recognized limitations and recommended continuing monitoring rather than treating pilot results as proof of safe universal deployment (Home Office, 2025). A parliamentary answer in April 2026 confirmed that the summarization tool was intended to assist casework, not decide claims, and that decision-makers must read the complete relevant evidence (Home Office, 2026). The lesson is that administrative support remains contingent on verification and preservation of caseworker responsibility.

UNHCR contributes a different, protection-specific normative dimension. Its Procedural Standards for RSD under its Mandate address registration, individual interviews, assessment, confidentiality, notification, and review of negative determinations. Its data-protection policy and later AI approach foreground purpose limitation, human rights, humanitarian principles, and avoidance of harm (UNHCR, 2015, 2020, 2025). These sources should not be mistaken for an Indonesian statute or a comprehensive algorithmic adjudication code. Rather, they identify the core procedural interests that any technology used within, or materially affecting, mandate RSD must safeguard. In Indonesia, this perspective is particularly important because substantive determination rests with UNHCR, whereas public agencies may hold or process supporting information.

DISCUSSION

A Refugee-Protection-Centered Model of Algorithmic Accountability

The reconstruction proposed here begins with a competence rule. No AI system should be treated as possessing independent authority to determine refugee status, decide the credibility of an applicant, or authorize an action that exposes the individual to serious protection risks. Where UNHCR conducts mandate RSD, the substance of its determination remains governed by the relevant mandate and procedures. Indonesian bodies should not acquire indirect adjudicative power merely by attaching a predictive label to data supplied to UNHCR. Equally, national authorities remain responsible for the legality of their own administrative handling, collection, storage, and disclosure of data. The model therefore separates responsibility by function while creating explicit duties of cooperation where one actor's algorithmic error may contaminate another actor's assessment.

The first operational standard is an inventory and classification of proposed AI uses. Institutions should identify the system, operator, data categories, intended tasks, persons affected, and foreseeable effects. Administrative indexing or duplicate detection may fall into a lower tier, although even these tasks require basic security and error correction. Translation, interview summarization, and retrieval of country information warrant stronger controls because omission or distortion may influence evidence. Credibility scoring, behavioral deception detection, automatic recommendations to refuse protection, or autonomous status decisions should be prohibited or excluded from authorized functions unless a future legal framework supplies a compelling and independently defensible basis; this article recommends maintaining a prohibition on independent RSD adjudication. Classification should respond to actual use, not vendor marketing labels.

The second standard is a refugee-sensitive algorithmic impact assessment before procurement, pilot operation, or material expansion. Building upon the logic of Canada's framework, assessment should record a legitimate purpose, available non-AI alternatives, expected benefits, foreseeable errors, risks to fundamental rights, mitigation, and residual uncertainty (Treasury Board of Canada Secretariat, 2025). It must also examine whether the relevant institution has lawful competence, whether sensitive data are essential, and whether the proposed output can be meaningfully checked. A cost-saving estimate is insufficient where the system may change an applicant's opportunity to be heard. The more materially the tool affects individual rights or evidence, the more rigorous independent evaluation, documentation, and authorization should become. High-risk deployments should not proceed merely because an assessment form has been completed.

The third standard is evidentiary traceability. Each substantive AI-assisted output should, where technically feasible, identify the source document or passage from which its claim derives, the version of the source, and relevant uncertainty. Human reviewers must compare material summaries against original interviews, verify translations where ambiguity matters, and examine both supportive and adverse country information. Retrieved material should be checked for currency, provenance, and relevance to the individual's circumstances. AI-generated references or legal propositions cannot be assumed authentic; the Canadian IRB's 2026 practice notice underscores the obligation to verify such material (IRB, 2026b). The purpose of traceability is not to impose impossible mathematical certainty but to make consequential inferences contestable through ordinary legal reasoning.

The fourth standard is meaningful human control. A qualified and accountable human official or mandate decision-maker must retain the ability to inspect source material, seek clarification, correct outputs, reject recommendations, and give independent reasons. Training must include common failure modes such as hallucination, inconsistent translation, overconfident classifications, and automation bias. Review should be reflected in records rather than assumed from a signature. In especially sensitive determinations, an independent second-level check may be justified when AI has materially influenced the evidentiary picture. The British approach to case summarization provides a useful distinction: a tool can accelerate reading and orientation, but it must not replace reading the evidence needed to decide a claim (Home Office, 2025, 2026).

The fifth standard involves nondiscrimination and contextual validity. Evaluation should consider whether performance varies across relevant languages, documentation conditions, and groups whose claims are affected by displacement or vulnerability. Such testing should be compatible with privacy and should not require indiscriminate collection of sensitive characteristics. When sufficient evidence of reliability is unavailable, institutions should narrow the system's functions or discontinue the risky use. A person's demeanor, facial expression, or communication style must not be converted into a generalized metric of truthfulness without an adequate scientific and legal foundation. Refugee eligibility requires the assessment of individualized experiences and prospective risk, not proxy-based inference about group identity (Amoroso, 2024; Kinchin & Mougouei, 2022).

The sixth standard is stringent data governance. In the Indonesian context, Law No. 27 of 2022 offers a baseline for evaluating lawful purpose, proportionality, security, and high-risk processing. A specialized protocol should specify the data necessary for each AI function, the identity of responsible institutional actors, retention and deletion rules, access controls, use by contractors, and conditions for any transfer between bodies. Refugee files must not be uploaded into uncontrolled services or used for unrelated model training simply because a service is convenient. The applicable responsibilities of domestic actors should be distinguished from UNHCR's own mandate-based policies, with coordination arrangements respecting their respective legal frameworks. A technology

procurement contract should never silently replace a transparent legal allocation of responsibility (UNHCR, 2015, 2024).

The seventh standard is effective notice and explanation. Applicants should be informed, in an accessible manner and without unnecessary disclosure of sensitive security details, when AI has materially assisted the processing of their information or the preparation of an assessment. An explanation must identify the function involved, the kind of data considered, any consequential factual proposition generated by the tool, and the available process for correction. This is different from requiring the public release of every line of proprietary source code. The relevant objective is an explanation adequate for legal challenge and institutional oversight. Where a system is too opaque to permit the review needed for procedural fairness, its use in a consequential function should be refused rather than normalized through vague assurances of technological accuracy.

The eighth standard is a differentiated system of correction and remedies. If an Indonesian agency records or transfers incorrect AI-generated data, it should provide an accessible route for correction and notify downstream recipients where necessary. If an AI output affects UNHCR mandate RSD, appropriate clarification and review should operate through UNHCR's procedural arrangements, without pretending that domestic administrative procedures automatically govern UNHCR determinations. Applicants should be able to identify errors and provide contextual explanations before adverse inferences become decisive. Where a possible mistake could contribute to return to persecution or another serious protection risk, verification and available review safeguards must be completed before potentially irreversible action. This connects algorithmic accountability with the protective logic of non-refoulement rather than treating it as a narrow consumer-data complaint.

The ninth standard is ongoing audit and incident reporting. Pre-deployment assessment becomes obsolete if models, sources, interfaces, or user practices change. Institutions should periodically evaluate accuracy, discriminatory effects, security incidents, and the real degree of human reliance on outputs. Material model changes or repeated errors should trigger renewed authorization. Audit records should include the system version, access logs, affected workflow, identified error, corrective action, and responsible officer, subject to appropriate confidentiality. Independent review should be proportionate to risk and institutionally credible. This approach recognizes that accountability is a continuing duty: even a tool that performs acceptably in a controlled pilot may function differently when deployed across diverse applicants and changing country conditions.

Institutional Implementation in Indonesia

Implementation should proceed in stages and avoid presenting a proposed policy as existing law. In the short term, Indonesian agencies involved in refugee handling could develop written standards for narrow, lower-risk administrative AI functions within their established competence. Such guidance should prohibit independent refugee-status adjudication by national administrative systems, limit data transfer, and require basic source verification and incident reporting. The existing personal-data framework and AI ethics circular provide useful normative points of departure, but operational safeguards should specify responsible actors, records, and procedures. Any deployment that materially influences evidence used in mandate RSD would require coordination with UNHCR and respect for its own procedures before being treated as acceptable.

In the medium term, a cross-institutional protocol could formalize assessment and responsibility at interfaces between domestic handling and UNHCR protection functions. Its provisions should identify permitted data exchanges, verify the quality of AI-generated information, establish points of contact for correction, and define what happens when a system's reliability is contested. Such a protocol would not transfer UNHCR's mandate or manufacture a new domestic appellate hierarchy. It would instead make cooperation accountable by specifying who creates data, who can use it, who must verify material inaccuracies, and who informs the applicant when an error may affect a protection process. Training should encompass operational personnel, interpreters, supervisors, and legal-support actors, with special attention to the risks of sensitive information.

A longer-term national framework could codify sector-specific restrictions for high-impact AI in administrative decisions affecting fundamental rights. It should provide legal authority for risk classification, independent assessment, enforceable duties of explanation and review, proportional sanctions, and oversight of technology vendors. Refugee-related functions should be explicitly recognized as sensitive because administrative errors can indirectly affect the capacity to obtain international protection. Indonesian policymakers could adopt the functional logic of Canadian impact assessment without transplanting Canada's adjudicative architecture. Likewise, they could draw on British testing and monitoring without assuming that Indonesia currently operates an equivalent asylum caseworker structure. Comparative borrowing is defensible only where the receiving institution possesses lawful competence and the adopted safeguard fits the relevant rights risk (Siems, 2022).

An illustrative workflow clarifies the limits of permissible assistance. Suppose a support office uses AI to summarize a multilingual account and identify relevant country material. The responsible operator records the purpose and system version, limits access to personal data, links each summary statement to the underlying interview, and flags uncertain translations. A trained human then checks the original evidence and offers the applicant an opportunity to correct omissions where they could matter. If information is communicated to UNHCR, the receiving actor understands its provenance and retains responsibility for evaluating it under mandate procedures. The algorithm's output remains neither

a credibility judgment nor a substitute for an individualized legal decision. This example illustrates a proposed safeguard architecture, not a claim that such a workflow presently exists in Indonesia.

CONCLUSIONS AND RECOMMENDATIONS

Indonesia's legal framework provides important general safeguards for refugee handling, human rights, personal data, and AI ethics, but it has not integrated them into a sector-specific regime for the use of AI in processes connected with refugee status determination. The principal weaknesses are dispersed institutional authority, a missing function-based risk classification, insufficient controls on consequential AI recommendations, limited explanation and verification standards, uncertainty over interinstitutional data responsibility, and the absence of clearly coordinated pathways for remedying algorithmic error. These are prospective regulatory vulnerabilities, not findings that Indonesia already operates an autonomous AI system deciding refugee claims. Distinguishing those propositions is essential to the validity of the legal analysis.

Comparative analysis supports a refugee-protection-centered model of algorithmic accountability. Canada's experience contributes ex ante impact assessment and clear limits on adjudicative AI; the United Kingdom contributes verifiable, supervised administrative use and continued evaluation; and UNHCR supplies the substantive procedural guarantees of individual assessment, confidentiality, reasoned decisions, and available review. For Indonesia, AI should be confined to carefully authorized decision-support functions, subject to source traceability, sensitive-data safeguards, bias testing, meaningful human review, continuing audit, and accessible correction. It must neither acquire independent competence to determine refugee status nor indirectly displace the authority of the legally responsible human or institution.

It is recommended that responsible Indonesian authorities establish immediate operational guidance for lower-risk uses, develop coordination arrangements with UNHCR for information that may influence mandate RSD, and progressively codify enforceable rules governing high-impact AI. Future policies should specify accountability at each stage, prohibit unsupported credibility proxies and autonomous adjudication, and link remedy mechanisms to the institution competent for the affected action. The protection of applicants must remain the criterion by which technological efficiency is judged. The proposed reconstruction thereby offers an anticipatory legal standard for responsible innovation without presuming a power of national refugee determination that Indonesian law has not established.

ADVANCED RESEARCH

The study is limited to normative and documentary analysis; it does not test an AI product, evaluate operational systems, or gather interviews with applicants or caseworkers. Future research should examine how automated summaries, translation tools, and country-information retrieval perform across languages and vulnerability contexts, with secure access and appropriate ethics safeguards. It should also investigate practical responsibility-sharing between national authorities and mandate institutions and assess whether notice, explanation, and correction procedures are genuinely accessible to persons seeking protection. Such empirical work could refine the proposed risk tiers and translate legal principles into measurable implementation criteria.

REFERENCES

Amoroso, M. P. (2024). Intelligent borders: Exploring the suitability of artificial intelligence systems in refugee status determination under international law. Refugee Survey Quarterly, 43(4), 410–426. https://doi.org/10.1093/rsq/hdae021.

Beduschi, A. (2021). International migration management in the age of artificial intelligence. Migration Studies, 9(3), 576–596. https://doi.org/10.1093/migration/mnaa003.

Byrne, W. H., & Gammeltoft-Hansen, T. (2024). Digital evidence in refugee status determination. AJIL Unbound, 118, 62–66. https://doi.org/10.1017/aju.2024.6.

Disemadi, H. S. (2021). Urgensi regulasi khusus dan pemanfaatan artificial intelligence dalam mewujudkan perlindungan data pribadi di Indonesia. Jurnal Wawasan Yuridika, 5(2), 177–199. https://doi.org/10.25072/jwy.v5i2.460.

Forti, M. (2024). Addressing algorithmic errors in data-driven border control procedures. German Law Journal, 25(4), 635–645. https://doi.org/10.1017/glj.2023.102.

Home Office. (2025). Evaluation of AI trials in the asylum decision making process. UK Government.

Home Office. (2026, April 28). Written answer to parliamentary question 129697 concerning Asylum Case Summarisation. UK Parliament. https://questions-statements.parliament.uk/written-questions/detail/2026-04-24/129697.

Immigration and Refugee Board of Canada. (2026a). Immigration and Refugee Board of Canada's 2026 to 2027 departmental plan.

Immigration and Refugee Board of Canada. (2026b). Practice notice: Use of artificial intelligence in IRB proceedings.

Kinchin, N., & Mougouei, D. (2022). What can artificial intelligence do for refugee status determination? A proposal for removing subjective fear. International Journal of Refugee Law, 34(3–4), 373–397. https://doi.org/10.1093/ijrl/eeac040.

Mutaqin, E. Z., Yahya, M. S., & Fian, K. (2024). Artificial intelligence efforts in addressing the humanitarian crisis: A study on Rohingya refugees. ISLAMICA: Jurnal Studi Keislaman, 18(2), 95–122. https://doi.org/10.15642/islamica.2024.18.2.95-122.

Peraturan Presiden Republik Indonesia Nomor 125 Tahun 2016 tentang Penanganan Pengungsi dari Luar Negeri. (2016). Lembaran Negara Republik Indonesia Tahun 2016 Nomor 368.

Rahayu, R., Roisah, K., & Susetyorini, P. (2020). Perlindungan hak asasi manusia pengungsi dan pencari suaka di Indonesia. Masalah-Masalah Hukum, 49(2), 202–212. https://doi.org/10.14710/mmh.49.2.2020.202-212

Setiyono, J. (2017). Kontribusi UNHCR dalam penanganan pengungsi internasional di Indonesia. Masalah-Masalah Hukum, 46(3), 275–281. https://doi.org/10.14710/mmh.46.3.2017.275-281.

Siems, M. (2022). Comparative law (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108892766.

Soekanto, S., & Mamudji, S. (2015). Penelitian hukum normatif: Suatu tinjauan singkat. Rajawali Pers.

Sukma, N. A. S., Nurlukman, A. D., Amiludin, A., & Ahmad, A. (2021). Bantuan pemerintah Indonesia dalam penanganan pengungsi internasional dari negara konflik. Masalah-Masalah Hukum, 50(2), 184–195. https://doi.org/10.14710/mmh.50.2.2021.184-195.

Surat Edaran Menteri Komunikasi dan Informatika Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial. (2023). Kementerian Komunikasi dan Informatika. https://jdih.komdigi.go.id/produk_hukum/view/id/883.

Treasury Board of Canada Secretariat. (2025). Directive on automated decision-making. Government of Canada. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592.

Treasury Board of Canada Secretariat. (n.d.). Algorithmic impact assessment. Government of Canada.

Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi. (2022). Lembaran Negara Republik Indonesia Tahun 2022 Nomor 196, Tambahan Lembaran Negara Republik Indonesia Nomor 6820.

Undang-Undang Republik Indonesia Nomor 37 Tahun 1999 tentang Hubungan Luar Negeri. (1999). Lembaran Negara Republik Indonesia Tahun 1999 Nomor 156, Tambahan Lembaran Negara Republik Indonesia Nomor 3882.

Undang-Undang Republik Indonesia Nomor 39 Tahun 1999 tentang Hak Asasi Manusia. (1999). Lembaran Negara Republik Indonesia Tahun 1999 Nomor 165, Tambahan Lembaran Negara Republik Indonesia Nomor 3886.

United Nations High Commissioner for Refugees Indonesia. (n.d.). Refugee status determination (RSD). https://help.unhcr.org/indonesia/refugee-status-determination.

United Nations High Commissioner for Refugees. (2015). Policy on the protection of personal data of persons of concern to UNHCR. https://www.refworld.org/docid/55643c1d4.html.

United Nations High Commissioner for Refugees. (2020). Procedural standards for refugee status determination under UNHCR's mandate. https://www.refworld.org/policy/legalguidance/unhcr/2020/en/123306.

United Nations High Commissioner for Refugees. (2024). Data protection impact assessment. UNHCR Information Integrity Toolkit.

United Nations High Commissioner for Refugees. (2025). UNHCR AI approach. https://www.unhcr.org/digitalstrategy/ai-approach.

United Nations. (1951). Convention relating to the status of refugees. https://www.unhcr.org/media/convention-and-protocol-relating-status-refugees.

United Nations. (1967). Protocol relating to the status of refugees. https://www.unhcr.org/media/convention-and-protocol-relating-status-refugees.