

Deferring Accountability, Not Obligation: Generative Artificial Intelligence and Non-Refoulement in Asylum

 rli.blogs.sas.ac.uk/2026/06/22/deferring-accountability-not-obligation-generative-artificial-intelligence-and-non-refoulement-in-asylum

refugeelawinitiative

June 22, 2026

By Bhavya Johari

A Record Corrupted Before It Is Read

On 7 May 2026, the Council of the European Union (EU) and the European Parliament reached a [provisional political agreement](#) on the [Digital Omnibus on Artificial Intelligence](#) (AI), a proposal to amend [Regulation \(EU\) 2024/1689](#) (the EU AI Act) with a view to simplifying the implementation of its harmonised rules on AI. Among its most consequential provisions is a sixteen-month deferral (from 2 August 2026 to 2 December 2027) of the binding obligations the EU AI Act imposes on high-risk AI systems used in asylum determination.

The timing is not incidental: a few months earlier, the European Commission's [European Asylum and Migration Management Strategy](#) (29 January 2026) had announced a Forum on AI in Migration to accelerate the deployment of AI tools across asylum procedures. The EU is expanding the role of generative AI in Refugee Status Determination (RSD) while withdrawing the legal infrastructure that governs it. The [EU Pact on Asylum and Migration](#), which entered into application on 12 June 2026, represents the same legislative orientation: it had already compressed the procedural space in which protection claims are examined, through accelerated border procedures and crisis derogations, before the Digital Omnibus removed the AI governance framework from what remained of it. The deferral postpones a set of regulatory obligations, but it does not suspend the independent procedural requirements that the non-refoulement guarantee already imposes on every asylum authority using AI to construct the factual record on which protection decisions rest.

Generative AI as Epistemic Intermediary

At the core of this problem is what generative AI does to testimonial evidence. Unlike a retrieval tool that surfaces existing content, a generative model reconstructs it, producing new text shaped by its training data, the framing of its prompt, and the compression of complex, culturally embedded testimony into schematic prose. The UK Home Office's Asylum Case Summarisation (ACS) tool illustrates the legal consequences of this distinction. The [Legal Opinion](#) commissioned by the Open Rights Group (ORG) on 19 March 2026, identified its legally decisive characteristic: ACS does not simply index or organise existing information but *creates new text for the Decision-Maker to consider*. The

system does not retrieve the applicant's account; it reconstructs it, presents it to the decision-maker without attribution, and without any indication of where the reconstruction diverges from the source.

That divergence is documented. The [Home Office's own evaluation](#), published on 29 April 2025, found that 9% of summaries *were deemed to be inaccurate or had missing information*, while 23% of users reported *they were not fully confident in the summary information*. In RSD, where testimony is frequently fragmented by trauma, mediated through interpreters, and shaped by cultural registers that training data drawn predominantly from Northern jurisdictions does not adequately represent, this figure is not a software defect awaiting correction. It is a structural feature of applying a general-purpose generative model to a testimonial domain where its representational capacity is intrinsically limited: generative models produce coherent prose regardless of whether the underlying source is coherent or accurately translatable, and in asylum, coherence and accuracy are not equally consequential.

The ORG [Legal Opinion](#) establishes that applicants are not informed that the ACS has been used in their case, receive no access to the summary it generates, and have no opportunity to identify or challenge errors before the decision-maker forms their view. The system contains no cross-referencing back to the source material, meaning errors in the AI-generated summary are not flagged to the decision-maker who relies upon it.

What the Legislature Admitted by Classifying Asylum AI as High-Risk

The structural risk illustrated in the ACS case study is one that the EU legislature expressly acknowledged. Annex III(7)(c) of the [EU AI Act](#) classifies as high-risk those AI systems intended to assist competent public authorities *for the examination of applications for asylum, visa or residence permits ... including related assessments of the reliability of evidence*. That classification is affirmative evidence that the legislature recognised these systems as capable of generating precisely the kind of procedural failure that international human rights law already prohibits. In [M.S.S. v. Belgium and Greece](#) (21 January 2011), the European Court of Human Rights found violations of Article 3 (prohibition of torture and inhuman or degrading treatment) of the [European Convention on Human Rights](#) against Greece for deficiencies in the asylum procedure that resulted in the risk of expulsion without serious examination of the merits of the applicant's claim, and a violation of Article 13 (right to an effective remedy) for the procedural failures that left the applicant without recourse. The Court's reasoning, namely that procedural deficiency in an asylum system can ground a Convention violation even where no expulsion has yet occurred, supplies the structural principle that an AI tool systematically distorting the evidentiary record on which merit assessments rest, while remaining opaque and uncorrectable, raises the same order of procedural concern.

The requirements for high-risk AI systems under the [EU AI Act](#), from Articles 9 through 15, including a continuous risk management system (Article 9), representationally accurate training data (Article 10), operational transparency enabling interpretation of outputs (Article 13), and meaningful informed human oversight (Article 14), were designed precisely to prevent that failure. Yet the Act, even at full operation, contained a structural concession in the very domain it most needed to protect. Article 14(5) creates a derogation specifically for *law enforcement, migration, border control or asylum*, permitting Member States to derogate from the two-person separate-verification requirement for high-risk AI outputs where national law considers it *disproportionate*. The legislature did not strengthen human oversight when it applied it to the asylum context; it weakened it. This reflects a political calculation: that procedural rigour for AI systems is negotiable precisely where the subjects of those systems have no electoral voice and no capacity to resist. The Digital Omnibus deferral is not an aberration from this logic; it is the same political economy operating at a higher level of abstraction. The accountability infrastructure for asylum AI is treated as a cost to be deferred because those who bear the cost of its absence cannot contest the deferral. The accountability gap is therefore not only caused by the Digital Omnibus; the Act itself partially constituted it.

Why Existing Instruments Cannot Close the Gap

The predictable institutional response is that interim protection remains available under the [General Data Protection Regulation](#) (GDPR), the [Asylum Procedures Regulation](#), which entered into application on 12 June 2026, and the general human rights framework. Each addresses a distinct slice of the governance problem but fails to reach its core.

The GDPR governs the lawfulness of personal-data processing and provides rights of access and rectification; it does not impose an ex-ante obligation to audit whether a generative model's training data adequately represents the linguistic and cultural diversity of the applicant population it will assess. The Asylum Procedures Regulation contains procedural safeguards for hearings and decisions, but it contains no provision requiring disclosure of AI tool use or verification of AI-generated content. The [Committee of Ministers' recommendation](#) (2020) on the human rights impacts of algorithmic systems affirms that *rule of law standards ... such as legality, transparency, predictability, accountability and oversight, must also be maintained in the context of algorithmic systems*, but creates no enforceable obligation. None of these instruments requires what Articles 9 to 15 of the [EU AI Act](#) would have required: that asylum authorities know, before relying on an AI-generated summary, whether the model was trained on data capable of accurately representing the applicant population before them. That question remains unanswered, ungoverned, and consequential.

It is here that the doctrinal force of non-refoulement becomes operative. UNHCR Executive Committee's (ExCom) [Conclusion No. 8 \(XXVIII\) of 1977](#) on the determination of refugee

status establishes, as minimum procedural guarantees for RSD, access to information about the procedure, the services of a competent interpreter, a reasoned right of appeal, and the right to remain in the country pending a decision. These guarantees share an architectural premise: the applicant must be able to engage meaningfully with the basis on which their claim is assessed, which requires that the underlying record accurately represent what they said and experienced. The UN Human Rights Committee's [General Comment No. 36](#) (2019) reinforces this from the standpoint of the right to life under Article 6 of the [International Covenant on Civil and Political Rights](#), requiring at paragraph 31 that States allow asylum seekers claiming a real risk to their life in the country of origin access to *refugee or other individualised or group status determination procedures* capable of offering protection against refoulement. Access to a procedure, however, is only meaningful if the remedy within it is substantively reachable. The right to an effective remedy, as elaborated in [De Souza Ribeiro v. France](#) (13 December 2012), demands procedural safeguards that are substantively accessible, not merely formally available. Where an error in an AI-generated summary cannot be identified or challenged because neither the applicant nor their representative knows the summary exists, the appeal right that [ExCom's Conclusion No. 8](#) guarantees is formally present but substantively void.

The chain is doctrinal: non-refoulement requires genuine capacity to identify protection needs; genuine capacity requires an accurate record; an accurate record requires, at a minimum, that the tool generating it be trained on representative data and that its outputs be disclosed so errors can be corrected. The AI Act's deferred obligations would have operationalised these requirements precisely. Their deferral does not suspend the underlying standard; it removes the mechanism through which compliance with it would have been enforced.

The Cost of the Gap and the Law That Already Governs It

The gap is not experienced neutrally across the applicant population. According to the [European Union Agency for Asylum's 2025 report](#), the five largest nationality groups of asylum applicants in EU+ states in 2024 were Syrians, Afghans, Venezuelans, Turks and Colombians. The subjective fear standard under [Article 1A\(2\)](#) of the 1951 Refugee Convention [resists algorithmic capture](#) because it is grounded in the individual's narrated, culturally embedded testimony, a standard that presupposes the kind of personal, contextual assessment that a generative model cannot reliably replicate. The representational limits of a generative model trained predominantly on English-language data are sharpest precisely where the protection need is greatest: the applicant whose testimony is mediated through multiple layers of linguistic and cultural distance from the model's training corpus bears the highest risk of an inaccurate summary and the lowest capacity to detect or contest it. The ungoverned deployment of AI tools whose error risk falls asymmetrically on Global South applicants instantiates the structural tendency of Northern-designed refugee law frameworks to serve [containment rather than protection](#):

not proportionate regulation, but structurally discriminatory governance that the deferral makes legally sustainable for sixteen months longer than it should have been.

A legal vacuum does not absorb that cost. [Article 33](#) (prohibition of refoulement) of the 1951 Refugee Convention depends on the procedural architecture through which protection decisions are made; its guarantees are [constitutive of protection, not administrative supplements to it](#). The procedural content of non-refoulement, as established through the authorities examined above, already encompasses the requirements of data accuracy, operational transparency, meaningful oversight, and access to the evidentiary basis of the decision that the AI Act's deferred obligations would have codified. The deferral removes the regulatory scaffold; it does not remove the obligation the scaffold was designed to serve. Member States deploying AI tools in RSD during the accountability gap are not operating in an unregulated space; they are operating in one where the operative standard is non-refoulement itself, and where reliance on the deferred AI Act timeline as justification for ungoverned deployment is a political argument that the law of refugee protection does not support.

For someone seeking asylum in the EU today, this is not an abstract regulatory question: their account of persecution may be summarised by an AI tool that distorts or omits what they said, placed before a decision-maker without disclosure, with no binding legal requirement for sixteen months that the system producing it be accurate or representative. This is not the consequence of a single regulation: the EU Pact on Asylum and Migration, which entered into application on 12 June 2026, had already redirected asylum procedures toward administrative speed rather than the careful, individual review that refugee law requires. The Digital Omnibus extends that logic to AI governance, making the right to asylum harder to access and contest for those with the least means to do so.

Bhavya Johari is a Lecturer at Jindal Global Law School, O.P. Jindal Global University, India, and serves as an Assistant Editor of the Jindal Global Law Review. He is also Consulting Director at IDIA (Increasing Diversity by Increasing Access), a national non-profit organisation that empowers students from underprivileged and marginalised backgrounds to pursue legal education. He earned his undergraduate law degree from NALSAR University of Law, Hyderabad, graduating with 10 gold medals awarded by the President of India. He holds an LL.M. from Melbourne Law School, University of Melbourne, where he was an Alex Chernov Scholar and received the Danny Sandor Prize in Children's Rights.

The views expressed in this article belong to the author/s and do not necessarily reflect those of the [Refugee Law Initiative](#). We welcome comments and contributions to this blog – please comment below and see [here](#) for contribution guidelines.