

Orwellian Odyssey: Smart Borders and the Imperative for Explainability

Uchenna Nnawuchi¹ and Carlisle George²

¹ PhD Researcher, ALERT Research Group, Department of Computer Science, Faculty of Science and Technology, Middlesex University, London UK
{UN055@live.mdx.ac.uk }

² Associate Professor, ALERT Research Group, Department of Computer Science, Faculty of Science and Technology, Middlesex University, London UK
{c.george@mdx.ac.uk }

Abstract Sunlight, they say is the best disinfectant. Therefore, it does not come as a surprise that transparency vis-a-vis explainability of Machine Learning (ML) algorithms used in the context of border control constitutes a key societal desideratum. Integrating ML algorithms in surveillance technologies, particularly algorithms deployed in border and asylum control and management (smart borders), has transformed security measures. This has evoked profound ethical and legal concerns as a result of ML opacity. This paper seeks to contribute to the ongoing discourse on safeguarding fundamental rights in an increasingly algorithmic world, specifically in the context of migration. The paper argues that embedding explainability into AI systems used in migration control is essential to prevent the emergence of an Orwellian surveillance state, ensuring that security measures do not erode civil liberties. It proposes several technical solutions to enhance AI system explainability and address power asymmetries between affected individuals and system users, including developers. It argues that by mandating explainability of ML algorithms, we can better balance power dynamics and protect the rights of individuals impacted by AI systems.

Keywords: Machine learning, Explainability, Automated decision-making, GDPR, AI Act.

1 Introduction

George Orwell's vision of a dystopian future [1], where "the eyes [are] always watching you and the voice enveloping you," ominously reflects contemporary realities with advanced surveillance technologies. Machine Learning (ML) systems have transformed individuals into subjects of relentless scrutiny, akin to "walking barcodes" [2]. This extends Orwell's warning, impacting even the sanctity of personal privacy as border control agencies increasingly use ML technologies. Smart borders now incorporate drones [3], robo-dogs [4], and extensive biometric databases including fingerprints, retina scans, and facial recognition [5], allowing precise tracking and monitoring of migrants' journeys. The advent of these technologies, specifically automated decision-

making systems (ADS), has revolutionised immigration and refugee processing, with these systems predicting aid delivery, human behaviour, and border crossings [5]. ML algorithms now influence decisions on visas and detentions, tasks previously managed by human officials [5]. These ADS —ranging from Emotion Recognition Systems (ERS) to advanced surveillance and data analytics [9]—aim to enhance efficiency and security by reducing waiting times and targeting interventions [10]. They improve border checks and detect illicit activities like smuggling and terrorism, thereby bolstering national security [11]. However, the opacity of some of these ADS poses significant ethical challenges. The ‘black box’ nature of these AI systems means that the reasoning behind algorithmic outputs is obscured, making it difficult to understand or question the factors influencing decisions [12]. This lack of transparency compromises procedural rights, privacy, and data protection [13], exacerbating power imbalances between state authorities and migrants.

This paper will investigate the opacity of decision-making algorithms in border control, particularly focusing on risk assessment tools and ERS. It will first examine the effects of ADS on migrants to highlight potential algorithmic harm. Next, it will review regulatory regimes governing border and asylum control, identifying gaps and shortcomings in existing regulations, such as the GDPR [4], EU AI Act [15], PNR [16], and ETIAS [17]. Finally, the paper will advocate for developing explainability guidelines as a solution to address ML opacity and protect individual rights.

2 Automated decision-making in Border Control

The opacity of ADS such as risk assessment tools and ERS raises significant ethical and legal concerns about their impact on migrants [18]. Tools like the Visa Information System (VIS) and the European Travel Information and Authorization System (ETIAS) analyse risk profiles for travellers to the Schengen area [19]. VIS processes information manually despite automation, whereas ETIAS issues travel authorisations automatically unless flagged [20]. Despite their advanced capabilities, these systems struggle with transparency; understanding why an individual is flagged by any of these systems can be elusive [9]. This opacity undermines fairness and due process, as migrants often cannot challenge or understand the criteria behind these decisions, potentially leading to significant rights violations [13].

The reliance on ADS risks perpetuating biases and systemic discrimination necessitating a rigorous ethical oversight in the deployment of AI at the borders. The black-box nature of these algorithms means that once a system flags a person, the specific factors influencing this decision are often opaque, making it difficult to address errors or biases [9]. The GDPR and EU AI Act mandate some human involvement, but this is ineffective if the human reviewer cannot understand the algorithm’s reasoning [21]. Thus, migrants face challenges in contesting decisions, risking detention or deportation without a fair hearing or procedural rights [20].

Furthermore, ERS tools like iBorderCtrl and AVATAR, which analyse micro-gestures and physiological responses to assess truthfulness, also face scrutiny. These systems, while designed to enhance border security, have been criticised for their questionable scientific validity and potential to infringe on human rights [19]. The use of

such tools may violate privacy and non-discrimination rights, as their opaque decision-making processes undermine the presumption of innocence and the right to a fair hearing [22][23]. Recently, iBorderCtrl's algorithm mistakenly flagged an honest journalist as lying, exposing its flaws. These systems' simplistic judgments of truthfulness, based on unreliable interpretations of micro-expressions and physiological responses, are scientifically flawed and ethically troubling, raising concerns about their validity and potential human rights violations in migration contexts [22] [23].

Calls for regulation are growing. Some propose a six-month moratorium on AI development [24], while the UN High Commissioner for Human Rights seeks a ban on untested AI tools and 'black box' applications [25]. The European Digital Rights (EDRi) network advocates prohibiting AI-based risk assessment in migration, citing inherent biases and human rights violations. EDRi also recommends classifying these tools as 'high-risk' under the EU AI Act and banning remote biometric identification [26]. While such measures aim to protect rights, they may impede AI's benefits in security, and efficiency, and stifle innovation [27].

Thus, this paper argues for integrating and mandating explainability into AI regulatory frameworks. Explainability ensures transparency and accountability in AI systems, enabling clear insights into decision-making processes and facilitating corrective actions. This approach bridges ethical concerns with security needs.

3 The Imperative for Explainability in Automated Systems.

This section addresses the importance of explainability in AI and ML systems within migration control, emphasising its role in protecting civil liberties, ensuring trust and accountability, and mitigating power asymmetries.

A Explainability as a Safeguard for Civil Liberties and Human Rights.

Integrating explainability into AI systems is a fundamental safeguard against an Orwellian surveillance state, where opaque and unaccountable decision-making processes can systematically erode civil liberties. In contexts like predictive border policing and risk profiling, the lack of transparency could foster discriminatory practices and disproportionate targeting of certain groups, thereby perpetuating biases and fostering discrimination. A right to explanation and explainability empowers stakeholders to challenge these decisions, ensuring they are based on transparent and equitable criteria. This not only diminishes the risk of discriminatory practices but also enhances trust in AI systems and protects individual rights. Explainable AI models—capable of elucidating decision-making processes without compromising accuracy—are crucial for ensuring legal and ethical compliance [28]. As emphasised by the High-Level Expert Group on AI ethics, explainability covers both the technical processes of AI systems and the associated human decisions, making it indispensable for upholding the rights of affected individuals [29]. Mandating explainability and the use of explainable models within AI systems—particularly in sensitive areas like migration management—could provide a balanced solution. Such mandates would ensure that these systems are either inherently or externally explainable, safeguarding the right to explanation and mitigating potential harm to civil liberties and human rights like non-discrimination, liberty etc. When AI models are transparent and understandable, stakeholders, including affected individuals and oversight bodies, are better equipped to assess and challenge decisions that impact

rights and freedoms. This capacity for challenge is not merely a technical requirement but a fundamental right that empowers individuals to hold AI systems accountable [28]. Therefore, embedding explainability into the design and deployment of AI is crucial to protecting civil liberties in an increasingly automated world.

B Explainability Increases Trust, Transparency, and Accountability.

Empirical evidence shows that increased transparency in AI systems significantly enhances user trust and system reliability [30], both of which are critical for successful deployment in sensitive areas like migration control. A "right to explanation" framework asserts individuals' right to understand and contest decisions significantly affecting them, safeguarding civil liberties and ensuring accountability [31]. Explainability, therefore, serves as a cornerstone for fostering ethical AI development, aligning with democratic values and societal norms. Explainability plays a pivotal role in translating technical solutions into legally relevant actions. By embedding explainable models into the architecture of AI systems—whether through inherently interpretable models such as decision trees and K-nearest neighbours or post-hoc explanation techniques like surrogate models, counterfactual explanations, and saliency maps—AI systems can achieve greater transparency in decision-making processes and in turn public trust [32]. Also, this level of transparency allows for enhanced control over systemic risks, the identification of biases in training data, improved safety, and heightened trust in the overall functionality of AI systems.

This transparency also empowers affected individuals and oversight bodies. It operationalises the "right to explanation," crucial for preventing the unfair and unethical application of algorithmic outputs. This facilitates due process for affected individuals and enables NGOs, collective redress organisations, and supervisory authorities to conduct legal reviews, scrutinising AI systems for unlawful behaviour such as manipulation or discrimination. In this manner, explainability ensures 'rights-enabling transparency' [30], allowing individuals to contest and comprehend decisions that significantly impact their lives, and enthrones accountability.

Furthermore, systems with higher levels of explainability foster public trust and engagement, essential for ensuring that AI technologies serve the broader public interest without compromising individual rights [30]. This promotes AI that enhances democratic values and aligns with human rights principles, fostering accountability among AI developers and deployers, and prioritising ethical considerations throughout the development and implementation process.

C Mitigating Power Asymmetries and Preventing an Orwellian State.

Machine Learning (ML) algorithms often function as "black boxes," reinforcing power asymmetries between individuals and system users, including developers and deploying organisations. This opacity concentrates power with developers and system users, who, under their access and control over the technology, can make decisions that significantly impact individuals and communities without meaningful oversight or accountability [33]. In sensitive areas like immigration in this context, these dynamic risks creating an Orwellian system of unchecked surveillance and decision-making. It potentially entrenches existing biases and inequalities and enables the covert perpetuation of structural injustices [33].

To counter these dystopian tendencies, embedding explainability into AI systems and enshrining the right to explanation within legal and ethical frameworks is imperative. Explainability serves as a mechanism to dismantle the opacity that fuels these power imbalances by making the reasoning behind algorithmic decisions accessible and comprehensible. This right to explanation empowers individuals to challenge outcomes that affect their rights and liberties [34].

By shifting power from technocratic elites towards more democratic governance of AI, explainability ensures these systems serve the public interest rather than reinforcing existing power structures or creating new forms of digital authoritarianism. In this context, explainability is not merely a technical solution but a societal necessity, safeguarding civil liberties, enhancing trust and accountability [34], and mitigating power asymmetries in smart borders [35].

While explainability is vital for promoting transparency, accountability, and trust in AI systems, especially in contexts with significant human rights implications, it is equally essential to examine the existing regulatory frameworks that govern these technologies. Understanding the extent to which current regulatory provisions in migration control offer a right to explanation will help identify gaps and opportunities for strengthening accountability and safeguarding civil liberties.

4 Regulatory Frameworks in Migration Control

Regulations governing automated decision-making in the context of migration include the GDPR, EU AI Act, PNR 2016, and ETIAS regulation; however, none seem to effectively provide for explainability as a redress to the issue of ML opacity raised in this paper.

The authors have argued elsewhere that a right to explanation and explainability mechanism has not been expressly provided for in the GDPR [32]. This argument has been echoed by other researchers [34, 36]. Art 86 of the updated version of the AI Act explicitly provides for a right to explanation of high-risk AI systems [15]. The Act refers to risk assessment tools and ERS in migration control as high-risk systems [7]. The promising appearance of the right to explanation, as currently framed in the Act, is a right in name only as it lacks practical efficacy without mandating that providers or deployers ensure these systems are explainable. The provision of a right to explanation in the AI Act does not *ipso facto* automatically balance the power asymmetries and protect the civil liberties of affected persons if explainability is not mandated. Both explanations and explainability go 'hand in glove' when high-risk AI systems are applied. One without the other is like a barrel without a bullet. A right to explanation is the legal underpinning, whereas explainability is the execution that brings this right to life. Without explainability, the right to explanation is rendered meaningless, as the affected individuals cannot comprehend the rationale for the decisions that impact their lives. Hence, a robust framework that integrates both legal and technical dimensions is essential to ensure transparency, accountability, and fairness in systems deployed for border control. It is necessary to have explainability to ensure that the right to explanation of high-risk systems in migration is achieved. This, in turn, protects the civil liberties and rights of those affected by complex decisions of border automated systems.

The ETIAS Regulation is applicable when the system is in use. Article 38(2)(c) stipulates that if an application is refused, the applicant must be provided with a "statement of the grounds for refusal of the travel authorisation indicating the applicable grounds from those listed in Article 37(1) and (2), enabling the applicant to lodge an appeal." This provision allows individuals to receive an explanation for the refusal of their applications. However, a key challenge arises: whether the human involved in the decision-making process can adequately explain the algorithm's output that contributed to the final decision. Additionally, there is no clear guideline on what constitutes an adequate explanation or how it should be obtained. Moreover, the regulation does not impose any specific compliance obligations on developers or system users, making it difficult to ensure accountability. This gap highlights the necessity of establishing a right to explanation and explainability in automated border control systems, along with ensuring regulatory compliance and the use of technical solutions to enhance explainability in migration AI systems.

5 Recommendations and Solutions for Enhancing Explainability in AI Border Control Systems

To address the opacity of AI systems in migration management, this paper recommends a multi-pronged approach that incorporates technical solutions, regulatory measures, and stakeholder engagement to ensure explainability, transparency, fairness, and accountability:

1. Establish a Regulatory Framework.

- Amend regulatory frameworks like the EU AI Act to explicitly define and protect the right to explanation and explainability
- Amend legislation mandating explainability in AI-driven migration decisions.
- Establish AI explainability regulatory board to oversee compliance and issue guidelines

2. Mandate Interpretable Models and Integrate Transparency Tools.

- Require the use of inherently interpretable models (e.g., decision trees, K-nearest neighbours) for high-stakes migration decisions. These models provide clear, understandable reasons for each decision, reducing ambiguity and enhancing transparency.
- Equip AI systems with post-hoc explanation tools like LIME and SHAP to help deconstruct complex model decisions, making them more accessible to officials and affected individuals.

3. Implement System Compliance Certificates.

- Create a certification process where AI systems must pass an explainability review before deployment.
- Establish an independent regulatory board to issue these certificates, ensuring adherence to transparency standards.

4. Conduct Regular Audits and Assessments.

- Perform periodic audits to ensure ongoing compliance with explainability requirements.

- Use these audits to identify potential biases or flaws in the systems and enable timely corrective action.
- 5. Provide Training for Border Officials.**
- Train border officials in understanding AI models and transparency tools.
 - Enable them to effectively interpret and communicate decisions to migrants, acting as an informed "human-in-the-loop" safeguard.
- 6. Develop Clear Guidelines for AI Producers and Deployers.**
- Create detailed guidelines mandating transparency and interpretability in AI systems.
 - Require incorporation of explainability into the design phase and continuous testing for bias and fairness.
- 7. Foster International Collaboration.**
- Encourage cross-border cooperation on AI explainability standards in migration contexts.

By implementing these solutions, policymakers, developers, and border authorities can ensure that AI systems at borders are both effective and aligned with fundamental human rights, fostering greater trust and accountability.

6. Conclusion

Integrating ML algorithms in border control improves security but poses ethical issues due to their opaque nature, preventing migrants from understanding or contesting decisions [37]. This paper stresses the need for explainability to ensure transparency, protect rights, maintain trust, and prevent Orwellian surveillance states. Solutions include using interpretable models, transparency tools, regular audits, and compliance certificates to address power imbalances. Current regulations, like the GDPR and EU AI Act, are insufficient in mandating explainability. Future frameworks must prioritise explainability to uphold human rights, ensuring AI in border control balances security with civil liberties.

References

1. Orwell, G.: *Nineteen Eighty-Four* (Penguin UK 2013).
2. Walker, C.: *Live Facial Recognition Technology Has No Place in a Free Society* (Politics Home 18 October 2023) <<https://www.politicshome.com/thehouse/article/live-facial-recognition-technology-no-place-free-society>> last accessed 12 August 2023
3. Cockerell, I.: *Greece Aims Long-Range Sound Cannons at Migrants across Its Border* (Coda Story 28 July 2021) <<https://www.codastory.com/authoritarian-tech/sound-cannons-greece/>> last accessed 27 July 2024.
4. Molnar, P.: *Robo Dogs and Refugees: The Future of the Global Border Industrial Complex* (www.theborderchronicle.com) <https://www.theborderchronicle.com/p/robo-dogs-and-refugees-the-future?utm_source=url> last accessed 27 July 2024.
5. Molnar, P.: *The Walls Have Eyes* (The New Press 2024).
6. Daniels, J.: *Lie-Detecting Computer Kiosks Equipped with Artificial Intelligence Look like the Future of Border Security* (CNBC 15 May 2018)

- <<https://www.cnn.com/2018/05/15/lie-detectors-with-artificial-intelligence-are-future-of-border-security.html>> last accessed 3 January 2024
7. Molnar, P.: EU's AI Act Falls Short on Protecting Rights at Borders (Just Security 20 December 2023) <<https://www.justsecurity.org/90763/eus-ai-act-falls-short-on-protecting-rights-at-borders/>> last accessed 6 April 2024
 8. Kenk, V.S.; Smart Surveillance Technologies in Border Control (2013) 4 European Journal of Law and Technology <<https://ejlt.org/index.php/ejlt/article/view/230/378>> last accessed 30 January 2023.
 9. Fink, M.: Robo Swarms and Polygraphs: The Future of European Border Management and Its Costs <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4883130> accessed 27 July 2024
 10. Nalbandian, L.: An Eye for an "I." a Critical Assessment of Artificial Intelligence Tools in Migration and Asylum Management' (2022) 10 Comparative Migration Studies <<https://comparativemigrationstudies.springeropen.com/articles/10.1186/s40878-022-00305-0>> last accessed 17 July 2024
 11. Everuss, L.: The Routledge Social Science Handbook of AI (Anthony Elliott ed, Routledge 2021).
 12. Pasquale, F.: The Black Box Society: The Secret Algorithms That Control Money and Information (Harvard University Press 2015)
 13. Digital Border Governance: A Human Rights Based Approach Digital Border Governance: A Human Rights Based Approach 2 Acknowledgments' (2023) <<https://www.ohchr.org/sites/default/files/2023-09/Digital-Border-Governance-A-Human-Rights-Based-Approach.pdf>>. last accessed 12 July 2024
 14. The European General Data Protection Regulation 2016.
 15. European Commission, 'Proposal for a Regulation Of The European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts' (Europa.eu2021) <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>>.
 16. Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, <https://eur-lex.europa.eu/eli/dir/2016/681/oj>, last accessed 3rd June 2024;
 17. 'EUR-Lex - 32016L0680 - EN - EUR-Lex' (Europa.eu2016) Regulation on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L0680>> last accessed 29 July 2024.
 18. 'Regulation - 2018/1240 - EN - EUR-Lex' (Europa.eu2018), Regulation Establishing Travel European Travel Information and Authorisation System <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32018R1240>> last accessed 29 July 2024.
 19. Kilpatrick, J., Jones, C.: A Clear and Present Danger Missing Safeguards on Migration and Asylum in the EU's AI Act 2 Authors' (2022) <<https://www.statewatch.org/media/3285/sw-a-clear-and-present-danger-ai-act-migration-11-5-22.pdf>> last accessed 20 July 2024
 20. European Digital Rights (EDRi), Migration and Technology Monitor, thePlatform for International Cooperation on Undocumented Migrants (PICUM) and Statewatch, Use of AI Migration and Border Control: A fundamental Rights Approach to the Artificial intelligence Act.

21. Barrett et al.: Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements, 2019 <https://doi.org/10.1177/1529100619832930> last accessed 12 June 2024
22. Sánchez-Monedero, J., Dencik L.; The politics of deceptive borders: ‘biomarkers of deceit’ and the case of iBorderCtrl’ (2022) 25 Information, Communication & Society Ibid 419–423.
23. Gallagher, R., Jona, L.: We tasted Europe’s new lie detector for travelers — and immediately triggered false positives, The Intercept (July 26, 2019) <https://theintercept.com/2019/07/26/europe-border-control-ai-lie-detector/> last accessed 14 November 2022.
24. Pause Giant AI Experiments: An Open Letter’ (Future of Life Institute, 8 June 2023) <<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>> accessed 31 July 2023.
25. McGregor, Murray and Ng (2019), ‘International human rights law as a framework for algorithmic accountability’, pp. 309–43.
26. EDRi, ‘Uses of AI In Migration and Border Control: A fundamental Rights Approach to the Artificial Intelligence Act’, https://edri.org/wp-content/uploads/2022/05/Migration_2-pager-02052022-for-online.pdf.
27. Sorelle Friedler SV and others, ‘The Problems with a Moratorium on Training Large AI Systems’ (Brookings, 11 April 2023) <<https://www.brookings.edu/articles/the-problems-with-a-moratorium-on-training-large-ai-systems/>> accessed 31 July 2023.
28. Panigutti C and others, ‘The Role of Explainable AI in the Context of the AI Act’ (2023) FAccT ’23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency Pages 1139 - 1150.
29. Ethics Guidelines for Trustworthy AI | Shaping Europe’s Digital Future’ (European Commission 8 March 2021) <<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>>222.
30. Inam, R., Terra, A., Mujumdar, A., Fersman, E., Feljan, av.: Explainable AI: How humans can trust AI’, Ericsson White Paper GFTL-21:000529Uen.
31. Vredenburg, K.: the right to explanation,’ Wiley: Journal of Political Philosophy 30 (2):209-229 (2022).
32. Nnawuchi, U.: Proposing a Protocol to the CoE Convention on Artificial Intelligence, Establishing and Implementing a Legal Right to Explanation of Machine Learning Social Science Research Network [2024] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4747094> accessed 13 July 2024.
33. Eubanks, V. (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.
34. Wachter, S., Mittelstadt, B., Floridi, L.: Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation, 7 INT ’ L D ATA PRIVACY L. 76 (2017)
35. Molner, P.: The EU’s AI Act and Its Human Rights Impacts on People Crossing Borders in Brief DIALOGUE on TECH and MIGRATION a Project of the Migration Strategy Group on International Cooperation and Development’ (2022).
36. Edwards, L., Veale, M.: Enslaving the Algorithm: From a “Right to an Explanation” to a “Right to Better Decisions”?, 16 IEEE SECURITY & PRIVACY 46 (2018)
37. Fort, K.: Migration Management in the Era of AI: How Emerging Technologies Shape the Border Space (LSE International Development 12 January 2024) <<https://blogs.lse.ac.uk/internationaldevelopment/2024/01/12/migration-management-in-the-era-of-ai-how-emerging-technologies-shape-the-border-space/>> last accessed 27 July 2024.