

Submission to the UK Parliament Joint Committee on Human Rights

Response to a Call for Evidence on
Artificial Intelligence and Human Rights

Submission to the Joint Committee on Human Rights: Response to the Call for Evidence on Artificial Intelligence and Human Rights

This submission by the International Panel on the Information Environment (IPIE) sets out recommendations for the regulation of Artificial Intelligence (AI) to safeguard human rights in the United Kingdom. Its purpose is to offer guidance to policymakers on how to establish robust accountability frameworks, ensure independent oversight of AI systems, and protect democratic processes while fostering innovation in AI development and deployment.

Summary of IPIE recommendations

Recommendation 1: Establish human rights impact assessments for AI deployments:

The IPIE recommends mandatory human rights impact assessments for high-risk AI systems, with particular scrutiny warranted for those integrated into social networking platforms. These assessments should evaluate risks to privacy, dignity, democratic participation, and access to accurate information, and include independent oversight, monitoring, and enforceable safeguards.

Recommendation 2: Establish accountability frameworks for Generative AI deployment in electoral contexts:

The IPIE recommends accountability frameworks that ensure electoral content integrity and address attribution challenges in democratic processes. This should include oversight and platform obligations during election periods such as content labelling and detection systems for synthetic electoral material, clear responsibility chains for developers and deployers in political communication contexts, accessible redress mechanisms, and international coordination against cross-border electoral misuse.

Recommendation 3: Establish system access requirements for independent AI auditing:

The IPIE recommends mandatory data access for independent auditors, supported by trusted intermediaries, to evaluate risks across the AI lifecycle. Auditing processes must provide sufficient access to documentation and datasets while safeguarding trust, privacy, and security.

Background on the IPIE

The IPIE is an independent and global science organization providing scientific knowledge about the health of the world's information environment. Based in Switzerland, the IPIE offers policymakers, industry, and civil society actionable scientific assessments about threats to the information environment, including AI bias, algorithmic manipulation, and disinformation. The IPIE is the only scientific body systematically organizing, evaluating, and elevating research with the broad aim of improving the global information environment. Hundreds of researchers worldwide contribute to the IPIE's activities.

This submission draws on multiple IPIE research outputs, including *The Role of Generative AI Use in 2024 Elections Worldwide*, which documented AI-related incidents in electoral contexts across 80% of countries holding elections.¹ The submission also references the IPIE technical paper *Towards A Global AI Auditing Framework: Assessment and Recommendations*,² which identified significant structural challenges in contemporary AI auditing practices that limit the effectiveness of independent oversight mechanisms. The submission further draws on the IPIE's *Trends in the Global Information Environment: 2024 Expert Survey Results*,³ which found that three-quarters of surveyed researchers identified barriers to data access as being among the greatest obstacles to studying the information environment.

Our evidence demonstrates the urgent need for comprehensive regulatory frameworks that can address the dual challenges of ensuring AI systems respect human rights while maintaining the transparency and accountability necessary for democratic oversight and effective remedies for individuals whose rights have been violated.

IPIE Recommendations

Recommendation 1: Establish human rights impact assessments for AI deployments

Given the documented scale and severity of AI-related harms to human rights, the IPIE recommends that policymakers consider mandatory human rights impact assessments for AI systems before deployment in high-risk contexts. Expert research from the IPIE's 2024 global survey reveals that 63% of researchers are very or extremely concerned about Generative AI's potential to perpetuate biases, amplify harassment, and increase misinformation.⁴ Experts predict these negative impacts will intensify, with 60% expecting AI-generated videos to greatly or somewhat worsen the information environment in the next five years.⁵

Establishing a preventive framework would enable identification and mitigation of human rights risks before deployment, particularly for AI systems integrated with social networking platforms, which experts identify as representing a significant threat to the information environment.⁶ Human rights impact assessments should apply to any AI system capable of generating synthetic content, ensuring risks are addressed in advance of public use. Enshrining this process in future regulation would create consistency across sectors and provide policymakers with clearer benchmarks for action.

To ensure that human rights impact assessments are meaningful and enforceable, the IPIE recommends that policymakers adopt the following measures:

- Assessment of potential impacts on privacy, dignity, democratic participation, and access to accurate information, with particular attention to the risks faced by vulnerable communities.
- Explicit evaluation of the risks arising from the integration of AI systems into major social networking platforms, given expert findings that platforms are among the greatest threats to the information environment.⁷
- Requirements for technical safeguards where potential human rights harms are identified.
- Independent oversight of assessment processes to ensure accountability.
- Ongoing monitoring systems to identify emerging harms as AI systems evolve.
- Appropriate enforcement mechanisms where adequate safeguards are not implemented.

Recommendation 2: Establish accountability frameworks for Generative AI deployment in electoral contexts

The rapid deployment of Generative AI in electoral contexts raises urgent concerns for democratic integrity and human rights by enabling actors to create and spread persuasive content at an unprecedented scale. In electoral settings, AI-generated content poses severe risks: It can be used to incite violence against vulnerable communities through fabricated campaign materials, create deepfakes targeting political candidates or voters with

manipulated imagery, and spread disinformation designed to suppress voter participation or delegitimize electoral processes.

The IPIE's 2024 global election analysis reveals that 90% of Generative AI incidents involved content creation, with 71% of these consisting of content creation alone, and the remaining 29% consisting of content creation as well as content proliferation and/or hypertargeting (micro-targeting individuals with personalised messages).⁸ Current approaches to countering misinformation and digital manipulation have proven inadequate against sophisticated GenAI-based tools,⁹ as AI-generated content can now be produced at speeds and scales that allow actors to shape political discourse while evading detection. The IPIE's 2024 global election analysis of 50 countries holding national elections last year found that Generative AI was used to influence elections in 80% of cases.¹⁰ Of these incidents, 69% appeared to pursue objectives potentially harmful to democratic processes, and nearly half (46%) could not be traced to specific actors,¹¹ creating structural barriers to attribution. Among cases where attribution could not be established, 79% involved suspected political manipulation, underscoring the link between anonymity and harmful practices.¹²

The IPIE recommends that policymakers establish accountability frameworks that prioritize political content integrity while addressing attribution challenges in high-risk contexts:

- Oversight and platform obligations: Require robust mechanisms for AI systems integrated with social networking platforms, including:
 - Content labeling and provenance tools to help users distinguish synthetic from authentic political content.
 - Detection systems and trust and safety protocols to identify, demote, or remove misleading synthetic content.
- Clear responsibility chains for AI developers, deployers, and end users, with mandatory disclosure requirements for political communication contexts.
- Mechanisms that provide clear pathways for challenging AI-generated content that harms reputation, privacy, or other fundamental rights.
- International coordination to address cross-border AI misuse while maintaining the ability to protect UK residents from harmful AI use by foreign actors.

Recommendation 3: Establish system access requirements for independent AI auditing

Independent AI auditing faces a fundamental information asymmetry that undermines effective oversight. The IPIE report *Towards A Global AI Auditing Framework: Assessment and Recommendations*, found that auditors conducting conformity assessments frequently depend on documentation provided by AI developers themselves, and lacked sufficient access to system documentation necessary for comprehensive evaluation across the AI lifecycle and supply chain.¹³ Such reliance on developer-curated information potentially compromises audit rigour and may enable companies to limit scrutiny of problematic system behaviours.

Access limitations create practical challenges across multiple dimensions of audit practice. Professional auditors commonly lack direct access to AI system development

documentation, forcing them to rely on potentially incomplete representations provided by the developers they are meant to assess.¹⁴ The problem is compounded by geographic and jurisdictional complexities, particularly when AI vendors operate under different regulatory environments than service providers, creating additional barriers to comprehensive assessment.¹⁵

Restrictive independent testing methodologies may also force auditors to rely on developer-provided bias assessments, potentially obscuring system behaviours that produce discriminatory outcomes.¹⁶ Furthermore, they may exacerbate dependence on developer-provided datasets and documentation by failing to require transparency about data limitations or gaps in historical records, while benchmark datasets that inadequately represent diverse demographic groups may perpetuate rather than identify embedded biases.¹⁷

To ensure effective independent auditing, the IPIE recommends the following measures:¹⁸

- Ensure access to evidence allowing auditors to evaluate whether auditees have accounted for potential risks and impacts across the AI supply chain and lifecycle.
- Enable auditors and auditees to establish secure data-sharing mechanisms that adhere to trust, identity, privacy, and security considerations.

REFERENCES

¹ International Panel on the Information Environment. 2025. *The Role of Generative AI Use in 2024 Elections Worldwide*. TP2025.2. Zurich, Switzerland: IPIE.

² International Panel on the Information Environment. 2024. *Towards A Global AI Auditing Framework: Assessment and Recommendations*. SR 2024.3. Zurich, Switzerland: IPIE

³ International Panel on the Information Environment. 2024. *Trends in the Global Information Environment: 2024 Expert Survey Results*. SR 2024.2. Zurich, Switzerland: IPIE.

⁴ Idem, p.19.

⁵ Idem, p. 17.

⁶ Idem, p.14.

⁷ Ibid.

⁸ International Panel on the Information Environment. 2025. *The Role of Generative AI Use in 2024 Elections Worldwide*. TP2025.2. Zurich, Switzerland: IPIE, p. 29.

⁹ Idem, p. 6.

¹⁰ Idem, p.25.

¹¹ Idem, p.32.

¹² International Panel on the Information Environment. 2025. *The Role of Generative AI in Electoral Campaigns: Mapping Global Patterns*. SP2025.1. Zurich, Switzerland: IPIE, p.6.

¹³ International Panel on the Information Environment. 2024. *Towards A Global AI Auditing Framework: Assessment and Recommendations*. SR 2024.3. Zurich, Switzerland: IPIE, pp. 11–12.

¹⁴ Ibid.

¹⁵ Idem, p.19.

¹⁶ Idem, p.20.

¹⁷ Idem, p.19.

¹⁸ Idem, pp.46–48.



IPIE
International Panel on the
Information Environment

International Panel on
the Information
Environment

Seefeldstrasse 123
P.O. Box 8034 Zurich
Switzerland