WRITTEN EVIDENCE SUBMITTED BY RESPONSIBLE AI UK

(RAI0049)

Human Rights and the Regulation of AI

About this Response

1. On behalf of Responsible AI UK[1] (RAI UK) and the Citizen-Centric AI Systems[2] (CCAIS) project, we are submitting this response to the Joint Committee on Human Rights, in response to the call for evidence on Human Rights and the Regulation of AI. The submission addresses human rights issues, existing and possible changes to legal and regulatory frameworks. In summary, our response highlights:

- Cautious deployment of AI-supported age assessments for young asylum seekers to uphold Article 3 of the United Nations Convention on the Rights of the Child.
- The need to clearly state the impact of historical data on human rights in the AI Opportunities Action Plan, and how AI growth would comply with the Equality Act 2010 and Article 14 of the European Convention on Human Rights.
- For public datasets to accelerate AI research, the need to define how consent will be respected especially for datasets that were already collected.
- Embrace issues around Intellectual Property and Copyright, instead of considering them barriers to innovation, as the right to protection is covered under Article 27 of the Universal Declaration of Human Rights.
- Follow responsible innovation principles to adopt the scan-pilot-scale approach and safeguard human rights.
- The need for a robust liability framework for multi-actor scenarios, and measures for backward-looking accountability and forward-looking responsibility.
- Sector-specific legal and regulatory frameworks for safety-critical contexts and domains like healthcare and aviation.
- Initiate AI regulation discussions earlier in the AI lifecycle (at ideation and design), where interventions are likely to be most effective.

[1] RAI UK https://rai.ac.uk

[2] CCAIS https://www.ccais.ac.uk

Response Authors

- Dr Sarah Kiden, Research Fellow, University of Southampton
- Dr Vahid Yazdanpanah, Lecturer, University of Southampton
- Dr Rafael Mestre, Lecturer, University of Southampton
- Dr Ingi Iusmen, Associate Professor in Governance and Policy, University of Southampton
- Dr Joe Atkinson, Associate Professor of Employment Law, University of Southampton
- Isabela Parisio, Postdoctoral Research Associate, Kings College London
- Prof Sarvapali (Gopal) Ramchurn, Professor of Artificial Intelligence, University of Southampton; and CEO, Responsible AI UK
- Prof Sebastian Stein, Professor of Artificial Intelligence, University of Southampton

Written Evidence

Human Rights Issues

How can Artificial Intelligence (AI) affect individual human rights for good or ill, in particular in the areas of: (i) discrimination and bias and (ii) effective remedies for violations of human rights?

2. In July 2025, the UK government announced[3] its plans to deploy AI facial recognition technology for age assessments for individuals whose age is unknown or disputed, in cases where an asylum seeker claims to be under the age of 18. The Independent Chief Inspector of Borders and Immigration Report[4] concluded that the use of AI technology was the most cost-effective option. The expected AI technology is trained on millions of images to produce an age estimate with a degree of accuracy.

3. While similar technology has been used by banks and online retailers, human rights organisations like the Refugee Council[5], Human Rights Watch (HRW)[6] and other

[3] Fenwick, J., & Bentley, O. (2025, July 22). UK to use facial recognition AI to stop adult migrants posing as children. *BBC News*. https://www.bbc.com/news/articles/cglzrklp8jyo

[4] Age assessment checks https://hansard.parliament.uk/commons/2025-07-22/debates/25072227000021/IndependentChiefInspectorOfBordersAndImmigrationReportAgeAssessmentChecks

[5] Helen Bamber Foundation, Humans for Rights Network, & Refugee Council. (2024). *Forced Adulthood: The Home Office's incorrect determination of age and how this leaves child refugees at risk*. https://www-media.refugeecouncil.org.uk/media/documents/Forced-Adulthood-joint-report-on-age-disputes-January-2024.pdf

critics have cautioned that facial age estimation has not been independently evaluated in broader real-world settings, and that these technologies were not designed for this purpose.

4. Studies[7,8] have shown that children who have experienced trauma often exhibit biological signs of aging. As a result, age assessment algorithms – particularly those that rely on facial analysis to identify patterns in facial structure (e.g., distance between facial features, skin texture and bone structure) do not account for children who have aged prematurely due to trauma, violence, malnutrition and exposure to different environmental stressors. Consequently, such technologies risk misclassifying minors as adults, and may end up retraumatising young asylum seekers[9].

5. Because the proposed age assessment technology is experimental, implementation should be done with caution to uphold Article 3 of the United Nations Convention on the Rights of the Child[10] (UNCRC) to make decisions that are in the best interest of the child.

Existing legal and regulatory framework

To what extent is the Government's policy approach to deploying AI, expressed in its "AI Opportunities Action Plan", sufficiently robust in respect of safeguarding human rights?

6. We welcome the Government's approach to shape the application of AI within a modern social market through the AI Opportunities Action Plan. However, the Action Plan currently does not provide a robust policy approach to safeguarding human rights, as it focusses mostly on enabling innovation. We recognise that it may be

---

[6] Han, H. J., & Bacciarelli, A. (2025, July 31). UK Plans AI Experiment on Children Seeking Asylum. *Human Rights Watch*. https://www.hrw.org/news/2025/07/31/uk-plans-ai-experiment-on-children-seeking-asylum

[7] Canady, V. A. (2020). Experiencing childhood trauma ages body, brain faster. *Mental Health Weekly*, *30*(32), 7–8. https://doi.org/10.1002/mhw.32478

[8] Colich, N. L., Rosen, M. L., Williams, E. S., & McLaughlin, K. A. (2020). Biological aging in childhood and adolescence following experiences of threat and deprivation: A systematic review and meta-analysis. *Psychological Bulletin*, *146*(9), 721–764. https://doi.org/10.1037/bul0000270

[9] Iusmen, I., Kreppner, J., & Cook, I. (2024). *Trauma-Informing the Asylum Process .Guidelines and Recommendations Co-developed with Young People Seeking Asylum* [Monograph]. University of Southampton. https://eprints.soton.ac.uk/487622/

[10] UNCRC https://www.unicef.org.uk/what-we-do/un-convention-child-rights/

challenging to balance the promotion of innovation while providing protections for human rights for UK citizens. However, we recommend that the Action Plan specifically should clearly outline human rights issues such as how historical data may impact hiring practices, financial lending, health data management[11] and other algorithmic decisions.

7. The Action Plan acknowledges changes in workforce and proposes measures to upskill the population for new AI-enabled jobs. However, it does not address the fact many of the jobs taken over by AI may not have a direct replacement with an AI job. Even with upskilling, some of the people affected by job cuts may not have the qualifications to switch to highly qualified AI jobs. If not addressed, this could exacerbate work inequalities. Moencks et al.[12] discuss a manufacturing case study to guide human-technology integration in cases where total automation is not the preferred option, but technology is empowering the workforce.

8. The Action Plan emphasises creating public datasets to accelerate public and private AI research, and it acknowledges *"national security, privacy, ethics, and data protection considerations"*. However, it does not define how consent will be respected (especially for datasets that were already collected and where use for AI research was not a consideration). Additionally, it lacks concise steps to address data protection and exploitation of sensitive data in a responsible way. Therefore, the plan is not robust with respect to human rights as it very lightly considers issues around consent for people's data to be used for these purposes.

9. The plan addresses the important issue of safe and trusted AI through, for instance, suggesting continued support for the AI Safety Institute (AISI) - now AI Security Institute. While this is a step in the right direction, it does not safeguard against responsible application of AI tools by governments and businesses. AISI's main focuses are on misuse, safeguards when attempting to circumvent security measures, etc. It does not address how these systems could be resilient enough to issues outside of these examples, such as bias and fairness issues, and how marginalised or vulnerable groups could be protected. It does not mention how AI growth would be

[11] Lau, P. L., van Kolfschooten, H., & van Oirschot, J. (2023). Joint Statement on The Impact of Artificial Intelligence on Health Outcomes for Key Populations: Navigating Health Inequalities in the EU. EU Health Policy Platform Thematic Network. https://haiweb.org/ai-and-health-inequalities-in-the-eu/

[12] Moencks, M., Roth, E., Beitinger, G., Freigang, A., & Bohné, T. (2022). Augmented Workforce: A Case Study on integrating Operator Assistance Systems for Repair Jobs into Human-centric Production. *2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 0339–0343. https://doi.org/10.1109/IEEM55944.2022.9989927

ensured to comply with the Equality Act 2010[13] or Article 14 of the European Convention on Human Rights[14] (ECHR). Misuse of perfectly safe AI tools (which would be under AISI's remit to evaluate) by government or businesses violating (at times inadvertently), could put human rights at significant risk.

10. There is a lack of rights-based language, such as accountability, redress, dignity, equality, and so on, while emphasising economic growth, competitiveness and trust, which are not proxies for human rights protections. The Action Plan does not propose any accountability measures when something goes wrong in this fast innovation process. It misses the chance to propose an external advisory or oversight body to safeguard human rights. There is also no mechanism, or pretentions to develop one, for individuals to challenge or appeal AI-driven decisions in public or private services.

11. The scan-pilot-scale approach is focused on *"rapidly developing prototypes or fast light-touch procurement"* and achieving *"meaningful impact on productivity, effectiveness and citizen experience, as well as maximising government spending power*". This focus could go against responsible innovation[15] practices, as set up by UKRI, which operates under the remit of the Department for Science, Innovation and Technology (DSIT).

12. Section 2.2 on adopting the scan-pilot-scale approach does not propose any measures to ensure this rapid prototyping is safeguarding human rights, or even ensuring the tools are developed with diversity, equality and inclusion in mind, and minimising potential risks, not only against misuse or national security, but also against biases, fairness and predicting unforeseen implications. It does not follow established responsible innovation principles, such as engaging with relevant stakeholders and involving them in decision-making.

13. Issues around Intellectual Property and Copyright are seen as a barrier to *"innovation and undermining our broader ambitions for AI, as well as the growth of our creative industries"*, rather than as a human rights concern, covered under Article 27[16] of the Universal Declaration of Human Rights: *"Everyone has the right to the protection of*

[13] https://www.legislation.gov.uk/ukpga/2010/15/contents

[14] https://fra.europa.eu/en/law-reference/european-convention-human-rights-article-14

[15] Responsible innovation https://www.ukri.org/manage-your-award/good-research-resource-hub/responsible-innovation/

[16] Universal Declaration of Human Rights https://www.un.org/en/about-us/universal-declaration-of-human-rights

the moral and material interests resulting from any scientific, literary or artistic production of which he is the author". This approach of rapid innovation with ethics, accountability and responsibility when seen as a barrier rather than necessities, risks developing tools that seriously put at risk human rights safeguarding.

14. The plan references that *"Just as with previous technological revolutions, the people and countries who make decisions about how these systems operate and what values they reflect - including their approach to safety - will have huge influence over our lives".* However, there is no commitment to push for value-driven AI or simply technological advancement, but rather growth- and innovation-driven AI.

Possible changes to legal and regulatory framework

Who should be held accountable for breaches of human rights resulting from uses of AI, and on what basis?

Q1. Where in the process of developing, deploying and using AI technologies should liability arise?

15. Liability should be considered across the entire AI lifecycle: from requirements gathering and design through development, deployment, and ongoing maintenance (noting that updates are common in AI software). Because AI systems are dynamic and influenced both by developer updates and user configurations, harmful outcomes can rarely be traced to a single actor.

16. Responsibility must therefore be established through contextual analysis, distinguishing potential causes (a common practice in preparing for litigation) from actual causes necessary to establish the specific actor responsible for a particular harm. For example, if harm occurs following a UK-wide software update, the updating developer may be a prima facie candidate for liability and likely the liable actor in most cases. However, closer inspection may reveal that a particular user's preference settings triggered the failure. Such user- and case-specific litigation processes are convoluted, but ongoing research on automated reasoning techniques provides a robust foundation.

17. A robust liability framework must account for such multi-actor scenarios, recognising that accountability may shift between developers, system owners, and users. In our view, embedding clear AI responsibility[17] rules into regulation will both incentivise safer design and ensure users understand where accountability lies.

Q2. What additional measures, if any, are needed to ensure that individuals have sufficient redress where they have suffered harm because of the use of AI?

18. To ensure meaningful redress, regulation should establish clear mechanisms for attributing responsibility in complex AI-driven contexts. Traditional liability approaches assume a traceable causal chain back to a human decision-makers. However, autonomous AI can break this chain as decisions often emerge from system optimisation (where the developer may set only high-level goals) rather than direct procedural instructions.

19. Measures are therefore needed to support both backward-looking accountability (which action of AI at which time resulted in a trajectory that ended in harm) and forward-looking responsibility (how future harms can be avoided through risk mitigation and management).

20. Tools from the AI research community, such as formal causation and verification methods[18], and responsibility-aware AI design[19], should underpin this effort.

21. Additionally, sector-specific frameworks are required. In some less-critical domains, minor level of harm caused by AI due to uncertainties might be seen negligible and could be covered by insurance or compensation. However, in safety-critical areas like healthcare and aviation, neglect and acting without knowledge of consequences should count as a breach of duty. This requires a context-specific "*responsibility toolbox*" to give end users transparent routes to redress.

---

[17] Dastani, M., & Yazdanpanah, V. (2023). Responsibility of AI Systems. *AI & SOCIETY*, *38*(2), 843–852. https://doi.org/10.1007/s00146-022-01481-4

[18] Yazdanpanah, V., Gerding, E. H., Stein, S., Cirstea, C., Schraefel, M. C., Norman, T. J., & Jennings, N. R. (2021). Different Forms of Responsibility in Multiagent Systems: Sociotechnical Characteristics and Requirements. *IEEE Internet Computing*, *25*(6), 15–22. https://doi.org/10.1109/MIC.2021.3107334

[19] Yazdanpanah, V., Gerding, E., Stein, S., Dastani, M., Jonker, C. M., Norman, T., & Ramchurn, S. (2022). Reasoning About Responsibility in Autonomous Systems: Challenges and Opportunities. *AI & Society*.

How might regulation match the pace of AI technology development, such as the emergence of agentic AI, to ensure that human rights are preserved as technology continues to develop?

22. Typically, technology design and policy development have occurred separately in different venues and with different audiences. About a decade ago, Jackson et al.[20] recommended that policy (and regulation by extension) should be considered as a third factor of Computer-Supported Cooperative Work (CSCW)'s alongside traditional orientations to design and practice. However, as expressed by Munger and Van Dael[21], research in the area still predominately focuses on technological innovation rather than innovation in policy specifically. More recently, there is a growing community of Human-Computer Interaction (HCI) and design researchers and practitioners interested in the technology and policy design integration. For example, Yang et al.[22] outline a research agenda on designing technology and policy simultaneously, highlighting some HCI issues (e.g., foundation models and their ecosystems) that require technology designers, developers and engineers to work at societal level, where public policy plays a big role in shaping society.

23. There is an opportunity for regulation to match the pace of AI technology development by initiating regulatory discussions in earlier stages of the AI lifecycle (i.e., at AI ideation or design), where interventions are likely to be most effective. Bringing design (from industry and academic research) and regulatory processes together opens up the space to integrate human rights and related values into both AI design and regulation. Anticipatory governance[23,24] approaches promote forecasting and scenario planning to explore possible futures.

---

[20] Jackson, S. J., Gillespie, T., & Payette, S. (2014). The policy knot: Re-integrating policy, practice and design in cscw studies of social computing. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 588–602. https://doi.org/10.1145/2531602.2531674

[21] Munger, J., & Van Dael, R. (2020). *Putting People at the Heart of Policy Design*. Asian Development Bank (ADB). http://dx.doi.org/10.22617/TCS200281-2

[22] Yang, Q., Wong, R. Y., Gilbert, T., Hagan, M. D., Jackson, S., Junginger, S., & Zimmerman, J. (2023). Designing Technology and Policy Simultaneously: Towards A Research Agenda and New Practice. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3544549.3573827

[23] Guston, D. H. (2014). Understanding 'anticipatory governance'. *Social Studies of Science*, *44*(2), 218–242. https://doi.org/10.1177/0306312713508669

[24] OECD. (2025). Steering AI's future: : Strategies for anticipatory governance. *OECD Publishing*, *OECD Artificial Intelligence Papers* (32). https://doi.org/10.1787/5480ff0a-en

24. The Policy Lab[25] has made strides in this direction by partnering with government departments, agencies and international organisations to understand policy challenges, engage publics in policy development, and collaboratively experiment ways to address policy challenges. Similar approaches (e.g., speculative design,[26] futuring,[27] co-design,[28] and ethnography[29]) may be adapted for AI regulation. In the context of AI regulations, anticipatory governance may focus on including a wide range of voices in speculating best-case and worst-case scenarios of AI futures and developing strategies to mitigate possible risks.

About CCAIS

CCAIS is a 5-year UKRI Turing AI Acceleration Fellowship led by Professor Sebastian Stein is developing the fundamental science needed to build AI systems that can be trusted by citizen users.

About RAI UK

RAI UK brings together researchers from across the four nations of the UK to understand how we should shape the development of AI to benefit people, communities and society. It is an open, multidisciplinary network, drawing on a wide range of academic disciplines. This stems from our conviction that developing responsible AI will require as much focus on the human, and human societies, as it does on AI. Funded by the Technology Missions Fund, we convene researchers, industry professionals, policy makers and civil society organisations.

(Sep 2025)

---

[25] Policy Lab https://openpolicy.blog.gov.uk

[26] Dunne, A., & Raby, F. (2024). Speculative Everything: Design, Fiction, and Social Dreaming. MIT Press.

[27] Hoffman, J., Pelzer, P., Albert, L., Béneker, T., Hajer, M., & Mangnus, A. (2021). A futuring approach to teaching wicked problems. *Journal of Geography in Higher Education*, *45*(4), 576–593. https://doi.org/10.1080/03098265.2020.1869923

[28] Steen, M., de Boer, J., Kuiper-Hoyng, L., & Visser, F. S. (2008). Co-design: Practices, challenges and lessons learned. *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services*, 561–562. https://doi.org/10.1145/1409240.1409350

[29] Atkinson, P., Coffey, A., Delamont, S., Lofland, J., & Lofland, L. (Eds). (2001). *Handbook of Ethnography*. SAGE Publications Ltd.

DOI: https://doi.org/10.5258/SOTON/PP0152