

2025-26



AI FOR SOCIAL PURPOSE

ADDRESSING KEY INDUSTRY AND POLICY CHALLENGES

Authors

Nisreen Ameen
Tippayanet Sorosrungruang
Lucy Gill-Simmen

ISBN 978-1-0369-7408-4

Publisher:

Digital Organisation and Society (DOS) Research
Centre, Royal Holloway University of London.

Publication year: 2025

A report by

**Digital Organisation and Society
(DOS) Research Centre**

Executive Summary

This report synthesises key insights and recommendations emerging from a strategic roundtable convened by the Digital Organisation and Society (DOS) research centre at Royal Holloway, University of London, held at the House of Lords in April 2025.

The roundtable brought together leading academics, policymakers, industry practitioners and community voices to examine how the UK can shape the development and governance of artificial intelligence (AI) to serve public interest, uphold democratic values and drive inclusive innovation.

At a time of rapid technological acceleration and deepening social inequality, participants stressed the need for AI systems and strategies that prioritise transparency, fairness and societal well-being. While AI offers immense potential to transform public services, productivity and problem-solving, it also poses urgent risks, ranging from algorithmic bias and data misuse to job displacement and environmental costs.

This report addresses these challenges

by focusing on five interrelated domains:

1. Education and workforce development
2. Trust and governance
3. Data and inclusion
4. Responsible innovation
5. Economic resilience



The discussions revealed a broad consensus on the need to embed AI literacy across all levels of education, reform data governance practices to be more inclusive and ethical, and establish national strategies for workforce upskilling aligned with the realities of automation. Strong calls were also made for collaborative frameworks that bridge sectors and support shared accountability in AI deployment. Environmental sustainability and public sector readiness were highlighted as often overlooked aspects requiring urgent investment.

Drawing on these discussions, the report puts forward targeted, practical recommendations aimed at government bodies, regulators, educational institutions, and employers.

As a university committed to social purpose, Royal Holloway, University of London and its Digital Organisation and Society (DOS) research centre at Royal Holloway Business School are uniquely placed to drive forward this agenda. This report represents a significant step in fostering evidence-based dialogue and promoting co-created solutions that respond to the ethical, social and economic dimensions of AI. It calls for a national commitment to ensuring that the future of AI in the UK is not only technologically advanced but also socially just and democratically accountable.

These include:

1. Strengthening multi-stakeholder collaboration and co-creation of solutions between academia, employers, policymakers, and civil society, with research centres playing a pivotal role in bridging evidence, practice, and policy to drive impactful and inclusive AI adoption.
2. The integration of foundational AI education in schools/FE/HE
3. Cross-sector knowledge hubs for innovation
4. Inclusive data infrastructures
5. Stronger mechanisms to support SMEs and workers at risk of displacement.





Acknowledgement

The Digital Organisation and Society (DOS) Research Centre

at Royal Holloway, University of London, directed by Dr. Nisreen Ameen, brings together an interdisciplinary community of researchers dedicated to exploring the critical intersections between digital technologies, organisations, and society.

Grounded in the principles of inclusivity and impact, DOS promotes collaborative inquiry into how digital innovation shapes work, governance, ethics, and everyday life.

This report builds upon the shared expertise, values, and research clusters within DOS, which span areas such as AI in services, digital inequality, cybersecurity, the digital economy, and business analytics. It reflects the Centre's commitment to co-producing knowledge that is both academically rigorous and socially purposeful.

We are especially grateful to **Baroness Manzila Pola Uddin** for her support and for graciously joining the roundtable discussion. Her long-standing commitment to equity and social justice provided an invaluable foundation for the event and its purpose.



**BARONESS
MANZILA POLA
UDDIN**

We would like to thank all members of the DOS community and its extended network of national and international collaborators who have contributed to shaping this vision of responsible and inclusive AI. The insights presented here are the product of ongoing dialogue across sectors, and we remain committed to fostering partnerships that generate real-world impact in line with DOS's mission.

FOREWORD



"The pace and extent of developments in Artificial Intelligence (AI) are astonishing. With such far-reaching implications, this report is hugely important in highlighting the opportunities in tandem with our shared responsibilities here at Royal Holloway. Exploring how inclusivity and social purpose – both of which are so much at the heart of our university strategy – must shape our collective work in this space, this report offers critical insights into enabling equitable progress and driving public benefits. I encourage all those working in this dynamic field to read it and reflect on its recommendations."

Professor Julie Sanders

Vice-Chancellor and Principal
Royal Holloway, University of London



"At Royal Holloway Business School, we are proud to lead conversations that shape the future of Artificial Intelligence. This roundtable brought together experts from academia, industry, and policy to explore AI's opportunities and challenges. The insights captured in this report reflect a shared commitment to responsible innovation and collaborative progress. As a Business School, within a University of Social purpose, we remain committed to use these insights to inform our research and courses. We look forward to continued dialogue that ensures AI benefits society as a whole."

Professor Christos Tsinopoulos

Dean of the Faculty of Business and Law
Royal Holloway, University of London



"Within their given context, the purpose of our research centres is to advance knowledge and understanding, foster innovation and develop new methodologies and technologies to address complex societal problems. This report is a clear example of us leading in that regard in bringing policymakers, practitioners and the community together to address the social challenges of Artificial Intelligence. The outcomes of the work present a clear consensus re the challenges faced, that both paves the way for co-created innovations to address them and shows the value of research centres in driving that co-creation."

Professor Mark Lycett

Vice Dean, Research and Knowledge Exchange
Royal Holloway, University of London

Authors



Nisreen Ameen^{1,2}



Tippayanet Sorosrungruang^{1,3}



Lucy Gill-Simmen¹

Discussants and members of the Digital Organisation and Society (DOS) research centre



Mark Lycett



Philip Wu



Konstantinos (Kostas) Stathis



Aaradhana Sharma Jyawali



Anabel Sara Gutierrez Mendoza



G. 'Hari' Harindranath



Najmeh Hafezieh



Ruth Livesey



Rebecca Bolt



Sameer Hosany



Stavros Polykarpou



The Khoa Do



Xingyu (Cynthia) Huang



Yang Yang



Yu Zheng



Yuanyuan Lai



Vera Hoelscher

¹ Department of Marketing, Royal Holloway Business School, University of London, London, UK

² Department of Marketing, The University of Jordan, Amman, Jordan

³ School of Management and Marketing, Greenwich Business School, University of Greenwich, London, UK

Discussants and members of DOS industry and policy steering group committee



Alessia Clusini

Co-founder of Trybes Agency; Digital Ethnographer



Archana Dhankar

Senior Director, Sutherland; Marketing Leader and Influencer



Ben Grace

Account Development Manager, Higher Education, Oracle



Daniel Ukasoanya

E-commerce and Digital Supply Chains, Bloomberg



Jarmila Yu

Founder and MD of YUnique Marketing Ltd



Laura Cioffi

Founder of VertisMedia & Thriize



Raj Rutah

Certified Global Chartered Director and Ambassador for Institute of Directors; Lead Clinical Engagement and Service Transformation, NHS



Uche Onyekpe

Machine Learning Expert, Ofcom

Discussants and members of DOS industry and policy steering group committee



Vijay Reddi

Regulatory
Transformational
Leader as Senior
Group Director, Roche
Product Ltd



Waseem Rawat

General Manager for
the Manufacturing
Development and
Strategic Maintenance
Toyota South Africa

Additional industry experts who have contributed to the discussion



Adam Tweed

Shaw Trust



Erin Young

Institute of Directors



Gary Hudson

AI Surrey Centre



“

At the Digital Organisation and Society (DOS) Research Centre at Royal Holloway Business School, we are proud of the research we do and the way we engage with government, industry and communities to co-create solutions with strong, measurable impact. Building an inclusive AI ecosystem requires more than innovation, it demands shared accountability, investment in ethical infrastructure, and a commitment to public value. Through this work, we aim to provide a clear and actionable roadmap for policymakers, industry and civil society.”



Dr Nisreen Ameen

Director, Digital Organisation and Society
(DOS) Research Centre, Royal Holloway Business School

Table of Contents

EXECUTIVE SUMMARY	1
ACKNOWLEDGEMENT	3
FOREWORD	5
1. INTRODUCTION	11
2. KEY FINDINGS	18
2.1. Challenges and opportunities in AI	18
2.2. AI in education and workforce development	21
2.3. Trust, transparency and governance	24
2.4. Data practices and digital inclusion	27
2.5. Responsible innovation and collaboration	28
2.6. Economic impact and future of work	30
3. RECOMMENDATIONS FOR ACADEMIA, POLICY AND PRACTICE	31
3.1. Strengthening collaborative frameworks	32
3.2. Embedding responsible AI in education	35
3.3. Investing in ethical AI infrastructure	39
3.4. Promoting inclusive and ethical data governance	40
3.5. Certification for AI that promotes human well-being and flourishing	42
4. SUMMARY AND NEXT STEPS	44

1. Introduction

Artificial intelligence (AI) is reshaping the foundations of modern economies, labour markets and public services. The UK government has recognised the strategic significance of AI through the publication of its AI Opportunities Action Plan¹ in 2025, which outlines 50 recommendations grouped into three missions: investing in foundational infrastructure, adopting AI across sectors, and building homegrown AI capability.

**AI contribution
to the UK
economy:**



**£400
billion**

by **2030**
Projected

The Plan emphasises key themes including infrastructure investment, data accessibility, public-private collaboration, and safe, trusted deployment of AI. These areas underscore the need for national readiness that matches the pace of AI innovation, with an emphasis on social purpose and long-term economic resilience.

AI is projected to contribute over £400 billion to the UK economy by 2030², but adoption remains uneven and often disconnected from robust policy and education frameworks. According to the Office for National Statistics (2024), only 15% of UK businesses currently make use of AI technologies³, with wide disparities across sectors, regions and business sizes.

**AI adoption in
UK business**

**Only 15% of UK
businesses use
AI.**



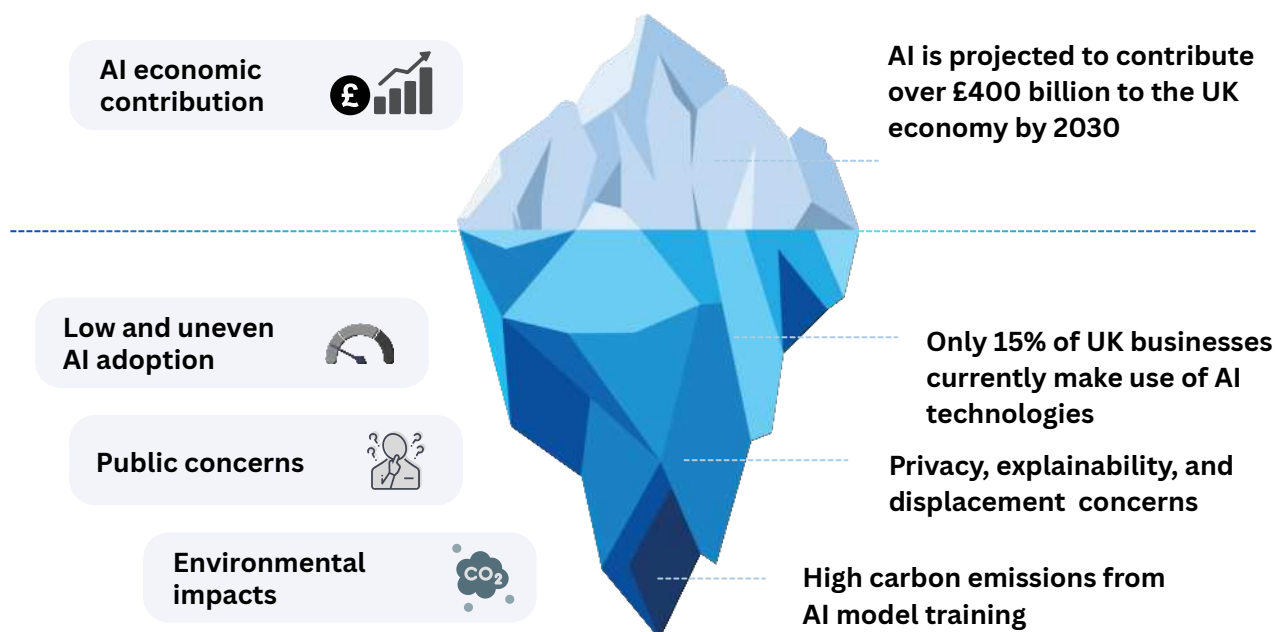
15%

^{1,2} <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan>

³ <https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/bulletins/businessinsightsandimpactontheukeconomy/previousreleases>

Meanwhile, public concerns around privacy, explainability, and displacement are growing, revealing a clear need for more inclusive and transparent AI development. Environmental impacts are also notable; training large AI models can produce substantial carbon emissions, with some estimates equating the training of a single large language model to a lifetime emission of five average cars⁴. These figures highlight the needs for policies that not only promote AI innovation but also ensure fairness, transparency and sustainability in its development and deployment.

AI economy contribution vs AI-related challenges

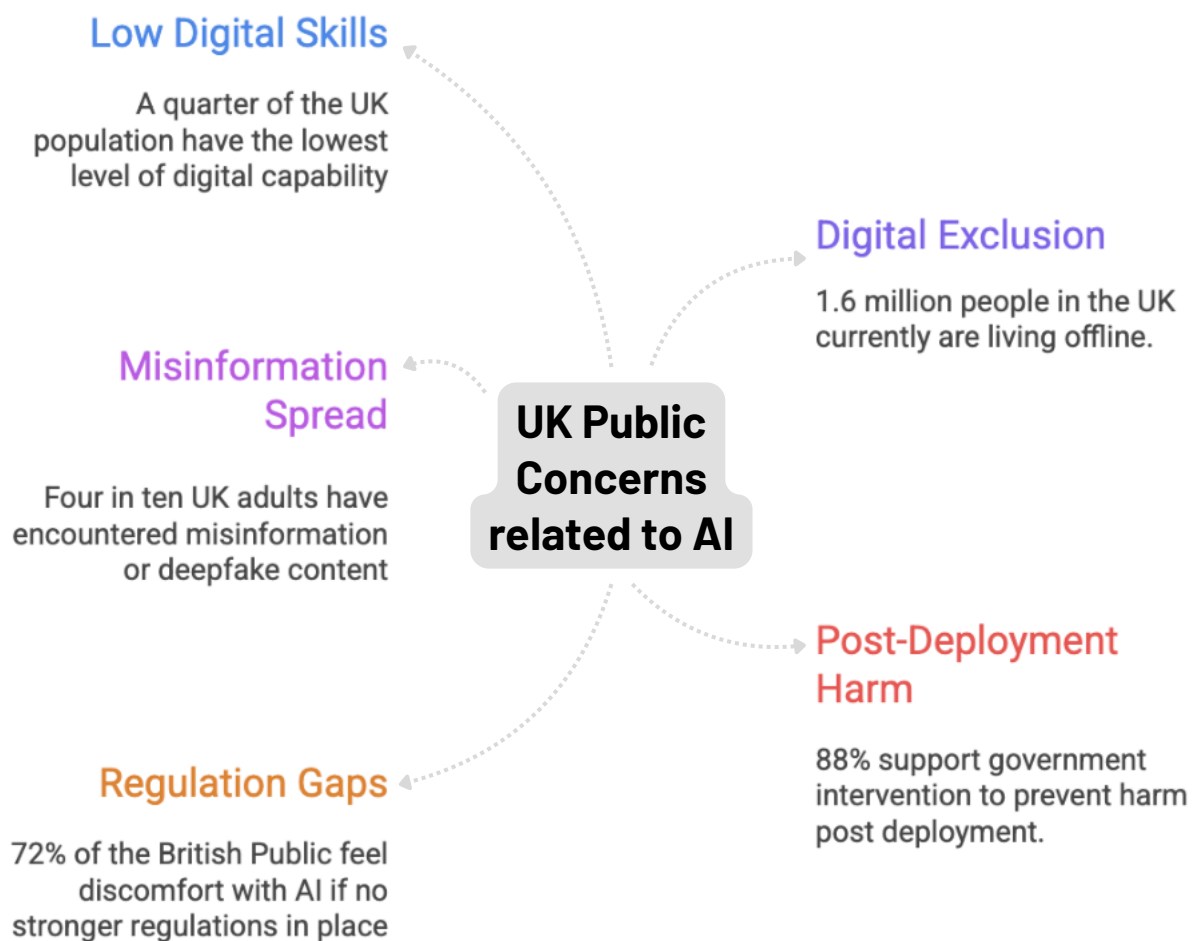


Recent data underscores the urgency of addressing several AI-related challenges. Public concern remains significant. The Ada Lovelace Institute found that 72% of the British Public would feel more comfortable with AI if stronger regulations were in place, and 88% support government intervention to prevent harm post deployment⁵.

⁴ <https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/>

⁵ <https://www.adalovelaceinstitute.org/press-release/ada-turing-survey-2025/>

Moreover, Ofcom's 2024 research indicates that four in ten UK adults have encountered misinformation or deepfake content online, raising concerns about the integrity of information in the digital age⁶. Furthermore, Lloyds estimates that 1.6 million people in the UK currently are living offline and around a quarter of the UK population have the lowest level of digital capability, meaning they are likely to struggle to use online services⁷.



⁶ <https://www.ofcom.org.uk/media-use-and-attitudes/attitudes-to-news/four-in-10-uk-adults-encounter-misinformation>

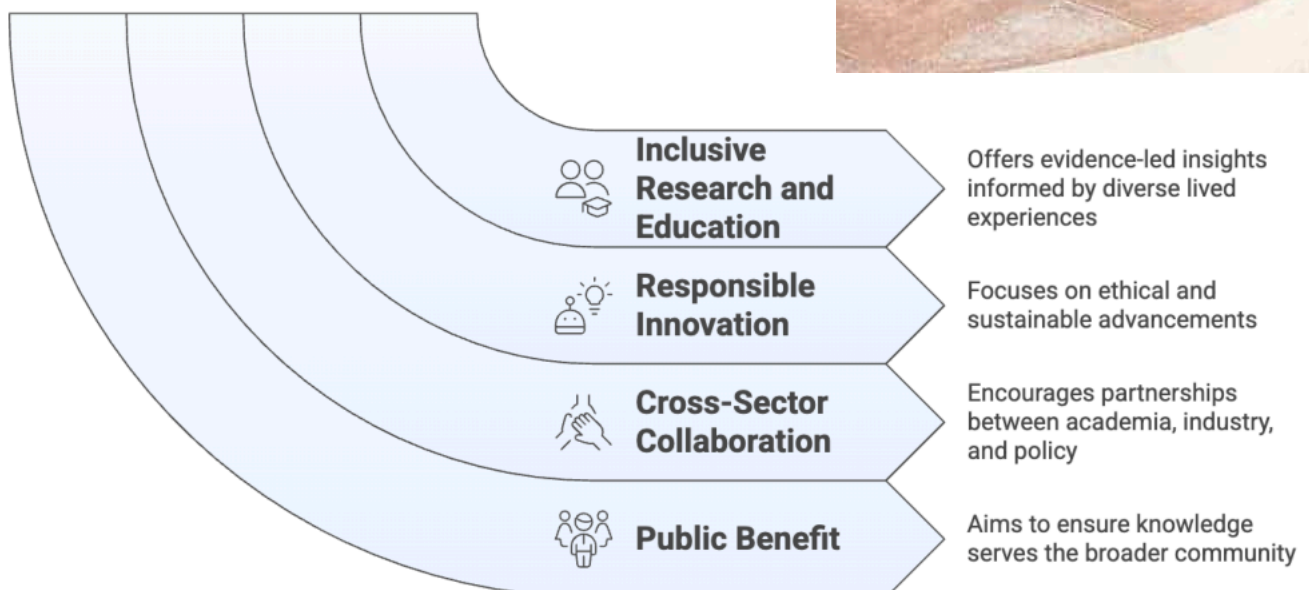
⁷ <https://www.lloydsbank.com/consumer-digital-index.html>

As a university for social purpose, **Royal Holloway, University of London**, is guided by a strategic commitment to addressing societal challenges through inclusive research, transformative education and meaningful external engagement. The university's strategy (RH2030s strategy⁸) places strong emphasis on inclusive research and education, responsible innovation, collaboration across sectors, and ensuring that knowledge benefits the wider public.

Within this context, academic–industry–policy collaboration becomes essential. Universities are uniquely positioned to offer evidence-led insights into the social, ethical and technical dimensions of AI, yet these insights must be better connected to business practices and policymaking processes. New models of engagement are needed to ensure that innovation is governed responsibly, informed by diverse lived experiences, and aligned with long-term public value.



RHUL 2030s Strategy



⁸ <https://intranet.royalholloway.ac.uk/students/news-events/our-rh2030s-strategy.aspx>

In response to the aforementioned challenges, the Digital Organisation and Society (DOS) research centre⁹ at Royal Holloway, University of London (RHUL) convened a roundtable at the House of Lords on 13 April 2025, bringing together researchers, industry leaders and policymakers.

The objective was to identify shared priorities, surface gaps in the current AI ecosystem, and discuss how collaborative strategies could strengthen ethical and inclusive AI adoption in the UK. The session focused on AI's potential for public good, while critically examining current barriers in education, data governance, regulation, and workforce transformation.



The event was held at a crucial moment in the national AI agenda, as attention shifts from strategic vision to implementation. It offered a timely opportunity to ask how academic research can be embedded in real-world solutions, how businesses can better engage with innovation frameworks, how educators can ensure AI literacy for the next generation and how government policies can reflect lived realities across the country.

⁹ <https://www.royalholloway.ac.uk/research-and-education/departments-and-schools/business-school/research/our-research-areas/digital-organisation-and-society-research-centre/>



Education emerged throughout the roundtable as **a cross-cutting enabler of responsible AI adoption**. It underpins the success of governance, workforce resilience, innovation, and ethical data practices.

Without equipping all citizens, from school children to mid-career workers to public leaders, with critical AI literacy, ethical reasoning, the broader policy and governance ambitions of the UK's AI strategy will falter. Embedding education as a foundation of the national AI ecosystem is therefore not simply desirable; it is essential for realising a socially just AI future.

Embedding education as a foundation of the national AI ecosystem is therefore not simply desirable; it is essential for realising a socially just AI future.

This report synthesises the key findings from the roundtable and **offers practical, policy-relevant recommendations across five thematic areas:**

Theme

01

Responsible innovation

Theme

02

Workforce resilience

Theme

03

Ethical data governance

Theme

04

Trust and inclusion

Theme

05

Collaborative infrastructure

It reflects a shared commitment to building a more inclusive, transparent and future-oriented AI landscape for the UK, one that prioritises societal well-being and democratic values alongside technological advancement and recognises that promoting societal benefit must be aligned with supporting sustainable innovation and competitiveness across the UK's AI ecosystem. We collectively offer key recommendations to academics, employers and policymakers.

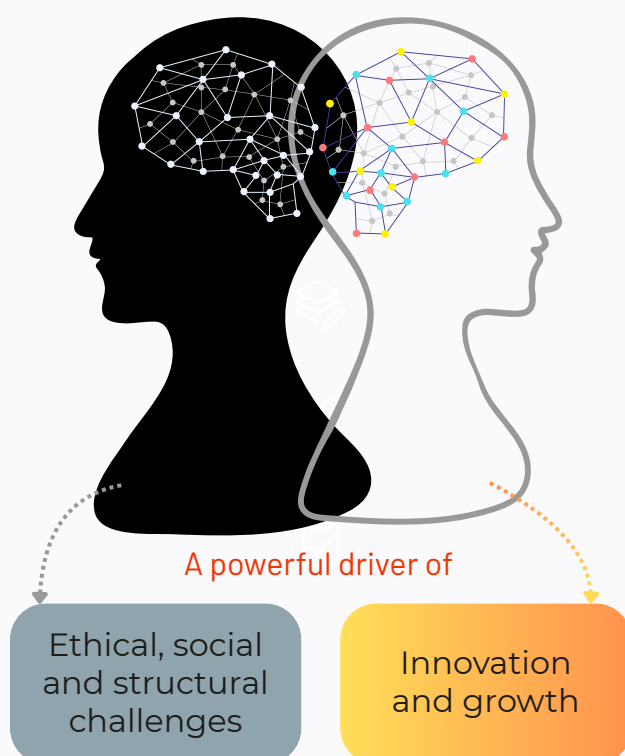
2. KEY FINDINGS

2.1. Challenges and opportunities in AI



AI presents a dual reality: a powerful driver of innovation and growth, and a source of significant ethical, social and structural challenges. As AI systems become increasingly embedded in decision-making processes, public services and daily life, the need to balance progress with responsibility has become more urgent than ever.¹⁰ The Alan Turing Institute defines AI as machines that can perform tasks that would previously have required human (or other biological) brainpower to accomplish.¹¹ AI is a broad field that incorporates many different aspects of intelligence, such as reasoning, making decisions, learning from mistakes, communicating, solving problems, and moving around the physical world.

AI presents a dual reality



¹⁰ <https://www.gov.uk/government/publications/frontier-ai-capabilities-and-risks-discussion-paper/ai-2030-scenarios-report-html-annex-c>

¹¹ <https://www.turing.ac.uk/news/data-science-and-ai-glossary>

The scale and pace of AI development can be viewed as both an inflexion point and a moment of risk. While there is shared optimism about the UK's potential to lead in responsible and innovative AI, there are strong warnings about the consequences of insufficient oversight and fragmented governance. Existing frameworks often focus on short-term innovation and commercialisation, with limited attention to the societal foundations required to support equitable and sustainable outcomes. In many cases, regulatory efforts remain reactive rather than anticipatory, struggling to keep up with the evolving capabilities of advanced AI systems.

The participants raised concerns about algorithmic bias and opaque decision-making processes. Without transparent design and rigorous testing, AI systems risk entrenching discrimination and excluding underrepresented communities from access to vital services.

There is a need for explainability and accountability, particularly in contexts such as healthcare, education, recruitment and law enforcement, where the stakes are high and human consequences are significant.

Misinformation and content manipulation also featured prominently. Generative AI tools have made it easier to fabricate realistic images, audio and text, blurring the line between fact and fiction and creating challenges for information integrity, democratic processes and public discourse. Without robust safeguards, these technologies could accelerate the spread of disinformation and deepen societal polarisation.

Without transparent design and rigorous testing, AI systems risk entrenching discrimination and excluding underrepresented communities from access to vital services.



There was agreement that green AI approaches that minimise energy consumption and prioritise resource-efficient design should become a core principle of responsible development.

Environmental impact emerged as a critical but often overlooked issue. Participants noted that training large AI models requires vast amounts of energy, often reliant on non-renewable sources. As AI adoption increases, so too will the demand for computational power and data storage.

This raises questions about how to balance innovation with environmental sustainability, particularly in light of the UK's broader net zero commitments. There was agreement that green AI approaches that minimise energy consumption and prioritise resource-efficient design should become a core principle of responsible development.

Despite these challenges, there are strong opportunities for the UK to take a leading role in shaping global AI norms.

The country benefits from a world-class research ecosystem, well-established regulatory institutions and an active community of start-ups and SMEs working at the frontier of digital technologies. However, to fully realise this potential, there must be long-term investment in inclusive digital infrastructure, cross-sector collaboration, and a stronger focus on public engagement.

It was noted during the roundtable that small and medium-sized enterprises (SMEs), in particular, face barriers in adopting AI due to cost, complexity and lack of guidance, an issue that must be addressed if the benefits of AI are to be broadly distributed.

Achieving a fair and inclusive AI future will depend on aligning innovation for profit with innovation for public good, ensuring that commercial progress and societal well-being advance together.

Achieving a fair and inclusive AI future will depend on shifting the focus from innovation for profit to innovation for public good.

This includes developing governance structures that are adaptive, participatory and able to incorporate ethical, legal and societal dimensions from the outset. Universities, regulators, civil society and employers all have roles to play in co-creating a framework that promotes trust, protects rights and ensures that technological progress enhances rather than undermines collective well-being.

AI represents a profound opportunity for national transformation, but also a responsibility to anticipate harm and prevent inequality. Harnessing the potential of AI in a way that reflects democratic values will require intentional leadership, inclusive policies and a clear commitment to shared societal goals.

2.2. AI in education and workforce development

The UK education and training ecosystem must be urgently reformed **to ensure that both current and future generations are not left behind** in an increasingly AI-driven economy and society.

The rapid expansion of AI across sectors has created a growing divide between technological capability and educational preparedness. Participants raised serious concerns about the inadequacy of current systems in equipping learners of all ages to navigate, understand and shape the future of AI. There was broad agreement that the UK education and training ecosystem must be urgently reformed to ensure that both current and future generations are not left behind in an increasingly AI-driven economy and society.

One of the central issues raised was the lack of foundational AI literacy in schools and colleges. While AI tools are already being used in classrooms and educational technologies, participants pointed out that most learners are not being taught what AI is, how it works, or how it shapes their digital environments. This absence of structured, age-appropriate learning on AI across the school curriculum was seen as a major risk, particularly in relation to critical thinking, online safety and civic engagement. Several participants stressed the importance of embedding AI understanding as a basic component of digital literacy, alongside topics such as privacy, algorithmic bias and the role of data in everyday decision-making.

At the higher education level, the discussion highlighted that while many universities are beginning to integrate AI content into degree programmes, this is often inconsistent and siloed within technical departments. Participants called for a more interdisciplinary approach, where students from business, law, healthcare, social sciences and the arts are equally exposed to AI's capabilities and implications. There was a shared sense that universities must not only provide technical knowledge, but also foster the ethical, legal and societal understanding needed to responsibly develop and use AI systems.

This absence of structured, age-appropriate learning on AI across the school curriculum was seen as a major risk, particularly in relation to critical thinking, online safety and civic engagement.

Participants emphasised the unique position of universities to lead on this, given their independence and research expertise, but noted that many institutions still lack the incentives or infrastructure to deliver this at scale.

Mid-career and older workers are at growing risk of being displaced or deskilled due to automation and algorithmic systems

Lifelong learning emerged as a particularly urgent area. Participants from both academia and industry noted that mid-career and older workers are at growing risk of being displaced or deskilled due to automation and algorithmic systems, yet few are being offered relevant, accessible retraining options. National modelling reinforces this concern, projecting that AI could ultimately displace 1–3 million UK jobs, but only gradually, with annual losses peaking at about 60,000–275,000, a figure that still falls within the labour market's typical year-to-year turnover—underscoring the urgency of scalable reskilling programmes¹².

Several participants pointed out that adult learning provision in the UK has declined significantly over the past decade, both in terms of funding and participation. Without bold action, they warned, the country could see rising digital exclusion among key parts of the labour force, including in sectors such as retail, logistics, and care, where AI tools are already being introduced.

To address this, participants called for more flexible, modular learning opportunities that are directly aligned with evolving industry needs. Online and blended delivery models were seen as key enablers, but only if accompanied by efforts to overcome digital poverty and lack of access to devices or connectivity.

AI could ultimately
displace

1–3 UK
million JOBS

Annual job losses

60K–275k

There were also strong calls for accreditation systems that recognise informal and non-traditional learning, enabling workers to build confidence and credibility in new digital skills without having to return to full-time study.

Without bold action, the country could see rising digital exclusion among key parts of the labour force, including in sectors such as retail, logistics, and care

¹² <https://institute.global/insights/economic-prosperity/the-impact-of-ai-on-the-labour-market>

Businesses must play a more active role in co-designing curricula, offering work-based learning opportunities, and supporting the upskilling of their existing workforce.

Employer involvement was identified as essential. Participants agreed that businesses must play a more active role in co-designing curricula, offering work-based learning opportunities, and supporting the upskilling of their existing workforce. However, it was also acknowledged that small and medium-sized enterprises often lack the capacity to engage with universities or training providers, and need more targeted support to do so. Participants suggested the development of regional or sector-based knowledge exchange hubs, where educators, employers and government bodies could collaborate to share resources, co-produce training, and identify emerging skills needs in real time.

For example, in the automotive industry, Jaguar Land Rover's Level 6 Digital & Technology Solutions Degree Apprenticeship offers a four-year BSc programme¹³. Apprentices spend five weeks in paid, shop-floor roles for every week on campus and follow pathways such as Data Analytics or Software Engineering with Data that cover Data Science and Machine Learning. The scheme shows how a large OEM can co-design AI-oriented curricula, fully fund tuition, and build an in-house pipeline of digital talent.

2.3. Trust, transparency and governance

Trust in AI technologies remains a critical concern among the public, policymakers and practitioners alike. Challenges around transparency, accountability and oversight continue to hinder wider adoption and public confidence. A central issue is the lack of clarity around how algorithmic systems arrive at decisions, especially in high-stakes sectors such as healthcare, recruitment and finance. Without mechanisms to ensure explainability and traceability, the public cannot evaluate whether AI outcomes are fair, accurate or aligned with democratic values.

¹³ <https://careers.jaguarlandrover.com/early-careers/apprentices/level-6-digital-technology-solutions-degree-apprenticeship>

Challenges around transparency, accountability and oversight continue to hinder wider adoption and public confidence.

Participants at the roundtable highlighted the urgent need for national guidance on AI transparency and clearer accountability standards. They expressed particular concern about the opacity of AI systems used in public services and called for tools that help users understand the rationale behind automated decisions.

The absence of standardised governance mechanisms across sectors has led to inconsistencies in oversight. While voluntary ethical guidelines and self-regulation exist, they are often fragmented and unenforceable. There is a growing need for national standards that mandate transparent and accountable AI development and deployment.

Inclusion is also essential to trustworthy AI governance. Communities that are most affected by automated decisions, including marginalised or underserved groups, must be meaningfully engaged in governance structures.

Their perspectives help to identify blind spots and ensure AI systems serve a wider public good. Multi-stakeholder models, such as citizen panels and community oversight boards, can enhance legitimacy and responsiveness. Roundtable participants noted the importance of community input to reduce blind spots in algorithmic design, and supported citizen panels as a model for strengthening democratic oversight.

Capacity-building among public institutions, regulators and civil society is needed to monitor AI systems effectively. This includes better tools for audit, redress mechanisms in cases of harm, and clearly defined responsibilities across agencies. Strengthening regulatory infrastructure and coordination will be vital to ensuring AI is governed with integrity and accountability.

Multi-stakeholder models, such as citizen panels and community oversight boards, can enhance legitimacy and responsiveness.

In the health sector, uncertainty around liability for medical errors and the complexity of safeguarding patient data throughout its lifecycle remain major barriers to adopting AI in clinical decision-making¹⁴. Liability concerns are particularly acute for autonomous medical AI systems, where it is unclear whether responsibility for harm lies with the physician, the hospital, or the AI developer and manufacturer. Currently, no legal framework explicitly assigns accountability for errors made by AI systems operating without human oversight.

Recent legal debates in the United States illustrate how unresolved liability issues can significantly affect developers and manufacturers of AI medical devices¹⁵. Some companies have opted not to release AI-powered diagnostic or decision-support tools due to the risks posed by unclear liability regimes. This ambiguity also deters investors, resulting in projects being abandoned either before market entry or after adverse events when developers cannot guarantee protection against lawsuits.

In the UK, NHS England recently (June 2025) paused development of “Foresight,” a generative AI model based on Meta’s Llama 2, following concerns raised by the British Medical Association (BMA) and the Royal College of General Practitioners (RCGP). These concerns centred on the governance of patient data in “black box” generative AI models, including using the data beyond its initial scope, insufficient safeguards, and liability for potential misuse of sensitive health information¹⁶.

Currently, no legal framework explicitly assigns accountability for errors made by AI systems operating without human oversight.

¹⁴ <https://pure.royalholloway.ac.uk/en/publications/ai-hesitancy-and-acceptability-perceptions-of-ai-chatbots-for-chro/>

¹⁵ Gerke, S., & Simon, D. A. (2025). New case law and liability risks for manufacturers of medical AI. *Science*, 388(6752), 1138–1140. <https://doi.org/10.1126/science.adu4932>

¹⁶ <https://www.digitalhealth.net/2025/06/ai-project-to-predict-health-outcomes-paused-over-gp-data-concern/>

2.4. Data practices and digital inclusion

Data remains foundational to the development and deployment of AI, but current practices often fail to deliver on inclusion, equity and trust. The use of incomplete or unrepresentative datasets continues to produce biased outcomes in AI applications. This is especially problematic in profiling, risk assessment and other contexts with significant consequences for individuals and communities. Participants highlighted the growing risks posed by unrepresentative datasets and warned that bias in training data is often amplified when applied in high-stakes decision environments.

Moreover, data is frequently extracted from individuals and communities without meaningful consent or benefit-sharing. Ethical data governance requires not only transparency but also agency, reciprocity and rights-based frameworks. Communities should be empowered to understand and control how their data is used. Digital exclusion continues to restrict access to AI benefits. Many individuals in the UK still face barriers such as inadequate connectivity, limited access to devices and insufficient digital literacy.

Bias in training data is often amplified when applied in high-stakes decision environments.

These forms of exclusion not only affect individuals' opportunities, but also skew the datasets that underpin AI tools. Ensuring high-quality, representative data demands targeted policies to address digital poverty and improve infrastructure. Participants called for the adoption of inclusive data policies and suggested models such as data cooperatives and public data trusts as ways to enhance public ownership and control over data.

Ethical data governance requires not only transparency but also agency, reciprocity and rights-based frameworks.

Alternative models, including data cooperatives and public data trusts, have the potential to democratise data access and ensure fairer distribution of AI benefits. These models provide mechanisms for collective decision-making over data use and enable greater public oversight. Inclusive data policies must prioritise access, transparency and community empowerment. Strengthening these areas will be essential for building more responsible and representative AI ecosystems in the UK.

2.5. Responsible innovation and collaboration

Responsible AI innovation requires a more expansive definition of success, in which societal value is advanced alongside technical and commercial outcomes. Embedding ethical foresight and transparency across the research, development and deployment cycle serves to strengthen the robustness, legitimacy and competitiveness of innovation, rather than constrain it.

Participants highlighted that current funding models and procurement processes often incentivise short-term gains over long-term ethical outcomes. They called for criteria that reward social impact and inclusive design. Proactive governance models are needed to anticipate risks and align AI innovation with public values. This includes ethics-informed procurement, interdisciplinary research, and the inclusion of diverse perspectives in project design. Current regulatory approaches tend to be reactive, leaving gaps in protection and oversight.

Collaboration is essential. Greater permeability between academia, industry and government can facilitate knowledge exchange and encourage more socially attuned innovation. To operationalise this, joint research initiatives should be structured with co-created agendas and shared governance models to ensure that ethical and social dimensions are embedded from the outset.

**Current funding models
and procurement
processes often
incentivise short-term
gains over long-term
ethical outcomes**

'Secondment schemes' should be expanded, for example, enabling academic researchers to spend time within industry or policy bodies, and vice versa so as to foster mutual understanding and cross-sector expertise. Shared infrastructures, such as regional AI innovation hubs or public data trusts, can provide physical and digital spaces where diverse stakeholders work together on inclusive AI solutions.

Mechanisms such as joint PhD programmes, collaborative curriculum design, and community-based AI labs can further embed socially responsive values into AI research and application. Importantly, collaboration must include civil society and underrepresented communities, whose perspectives are vital for shaping AI in a way that serves the public good.

Support for grassroots and community-led innovation is also vital. Such initiatives often reflect local needs and perspectives more accurately, yet they frequently face systemic barriers to funding, visibility and influence. Investment in co-design, experimentation and capacity-building at the community level can enhance inclusivity and responsiveness in AI innovation.

Enabling academic researchers to spend time within industry or policy bodies, and vice versa so as to foster mutual understanding and cross-sector expertise.

There is a need for targeted support to community-based projects, particularly those led by or serving marginalised groups. Structural reforms are needed to support this vision. These include moving beyond conventional success metrics, fostering long-term innovation partnerships, and embedding responsible innovation principles in all public and private funding mechanisms.

Investment in co-design, experimentation and capacity-building at the community level can enhance inclusivity and responsiveness in AI innovation.

Ultimately, responsible innovation in AI must be guided by a systemic and inclusive vision that recognises the interconnectedness of technical advancement, ethical accountability and social equity. This requires not only institutional reform and cross-sector collaboration, but also a cultural shift towards valuing long-term societal benefit over short-term returns. By embedding inclusive design, participatory governance and ethical reflexivity into the core of innovation ecosystems, the UK can position itself as a global leader in AI that is not only powerful, but also just, sustainable and responsive to the needs of all communities.

2.6. Economic impact and future of work

The impact of AI on employment, productivity and economic structures is both profound and uneven. As AI continues to transform tasks, workflows and industries, the nature of work is being reshaped. While some roles are augmented or newly created, others are being displaced or fundamentally altered. Participants noted that job displacement is already occurring in some sectors and raised concerns about the pace of change outstripping current policy responses.

This transition risks exacerbating inequalities. Digitally advanced firms and sectors are likely to reap disproportionate benefits, while traditional industries and SMEs may struggle to adapt. Without coordinated support, regional disparities and labour market fragmentation could intensify. Employers have a central role to play in managing this transition. However, these efforts should be supported by public policy, including financial incentives, tailored training programmes and adaptive labour market strategies. Participants called for stronger incentives and support structures to help SMEs prepare for technological change.

Digitally advanced firms and sectors are likely to reap disproportionate benefits, while traditional industries and SMEs may struggle to adapt.

Social protection systems must also evolve. Existing models of employment rights, benefits and representation are not well suited to platform-based and gig-economy work structures. New frameworks that ensure security, mobility and dignity in work will be essential as AI reshapes economic participation. Anticipatory planning is needed to distribute the benefits of AI equitably. This means involving workers in decisions about technology use, linking productivity gains to improvements in well-being, and investing in shared digital infrastructure that supports innovation and inclusion.

3. RECOMMENDATIONS FOR ACADEMIA, POLICY AND PRACTICE



The following recommendations draw on key themes emerging from the roundtable discussions, offering practical pathways for embedding social purpose, ethical responsibility, and inclusion into AI development and governance. Designed to inform academic strategies, policy interventions, and industry practice, these recommendations reflect a shared understanding that addressing the risks and opportunities of AI requires coordinated, cross-sectoral efforts. The emphasis is on sustainable, values-driven innovation that serves the public good, supports workforce resilience, and ensures equitable access to digital infrastructure and skills.

Each area below highlights concrete steps to help shape a more inclusive and accountable AI ecosystem in the UK.

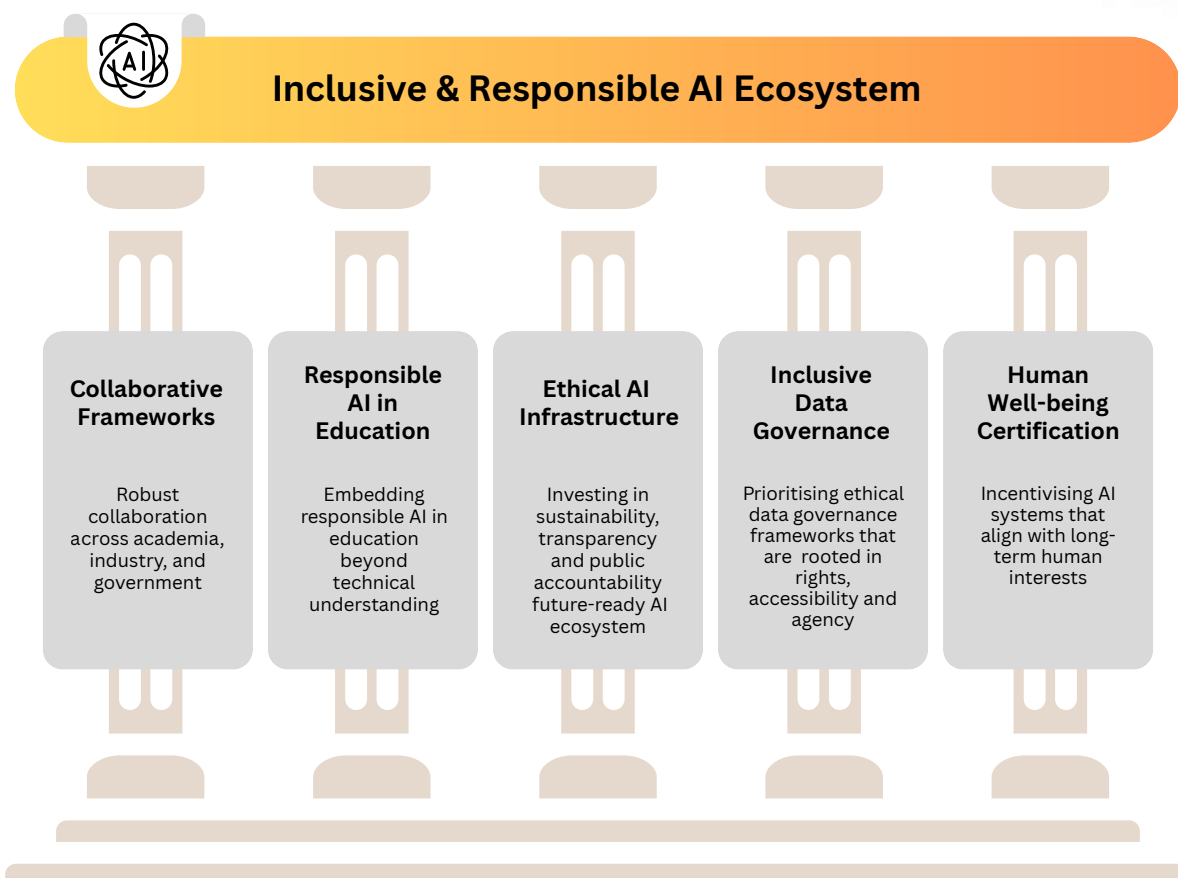


Figure 1. The five interconnected pillars of an ethical and inclusive AI ecosystem, illustrating how collaboration, education, infrastructure, data governance, and human well-being are mutually reinforcing.

3.1. Strengthening collaborative frameworks

Robust collaboration across academia, industry, and government is crucial for effective AI governance and innovation. The siloed development of AI risks producing systems that are technically proficient but misaligned with societal needs. A key insight from the roundtable was that co-creation can bridge this gap by combining practical expertise, community knowledge, and academic insight.

Beyond the workplace, growing evidence suggests that AI could intensify structural inequalities within society if inclusivity is not embedded by design.

Participants emphasised the need for long-term collaborative infrastructure, such as knowledge exchange hubs and interdisciplinary research centres. These platforms would support the co-design of policy-relevant research and the implementation of inclusive AI solutions. Joint funding calls could incentivise applied work that delivers measurable public outcomes, particularly in areas such as health, education and public service delivery.

Multi-sector forums play a vital role in facilitating dialogue, foresight, and course correction. These forums should include not only policymakers and researchers, but also community leaders, grassroots innovators, and underrepresented voices. Embedding collaboration into governance processes can improve public trust and ensure that AI development aligns with broader social goals.

Beyond the workplace, growing evidence suggests that AI could intensify structural inequalities within society if inclusivity is not embedded by design. As ownership and control of AI technologies become increasingly concentrated, the economic benefits of automation may accrue disproportionately to a small group of firms and investors, widening the gap between capital and labour. Policymakers and international commentators have therefore called for new mechanisms to ensure a fairer distribution of AI-generated value.

Approaches such as public or collective investment schemes, sometimes described as universal basic capital¹⁷, have been proposed as potential ways to enable citizens to share in the productivity gains of the AI economy. While the implementation of such models remains complex, they underscore the need for proactive policy design to prevent further concentration of wealth and power in the AI era. Investing in secondments, fellowships and dual-role positions can help institutionalise collaboration and build capacity across sectors.

¹⁷ https://www.ft.com/content/9b93e02a-c693-4070-9094-a2f532dfa929?accessToken=zwAGQ_oftv04kd0bk-AqxpNAcN0QIKL1Mt-pKQ.MEYCIQCP1ZCj2QnXyAarArZB_aFEIlgJrkQAA8GesZLMAnHe4UwlhAMnjfjDLLLYlzfSB_OzLhRTGsKvwqoSMjTtJA5GQc6DI&sharetype=gift&token=4d0e7ee3-f855-4ce0-a676-d7e7ca928756

Specific priorities for strengthening collaborative frameworks include:

- Establishing co-governed AI research programmes that embed social and ethical objectives alongside technical goals, supported by joint funding calls between research councils, industry partners, and government bodies.
- Expanding structured secondment schemes to enable academic staff to gain experience within industry and public sector organisations, and vice versa, fostering mutual learning and breaking down sectoral silos.
- Developing regional AI innovation hubs that provide shared physical and digital spaces where academia, industry, government, and civil society actors can co-create inclusive AI solutions and test emerging technologies.
- Creating joint doctoral and professional education programmes that integrate technical, legal, and ethical dimensions of AI, with input from employers and civil society to ensure relevance and societal alignment.
- Embedding civil society and underrepresented community voices into AI governance structures, innovation initiatives, and ethics committees, to ensure that diverse perspectives shape AI development and deployment.
- Investing in public–academic–industry partnerships to develop accessible tools, frameworks, and public engagement activities that enhance critical AI literacy across society.

3.2. Embedding responsible AI in education

Embedding responsible AI in education is fundamental to ensuring long-term societal readiness. The absence of structured AI literacy in primary and secondary education risks excluding future generations from understanding and shaping the technologies that will define their lives. AI education must go beyond coding and technical proficiency, encompassing the ethical, societal and legal dimensions of AI use. Interdisciplinary integration is vital, students in humanities, social sciences, and vocational tracks must be given the tools to critically engage with AI, not just students in STEM fields.

To support this, AI education must go beyond technical understanding to foster what some scholars call AI fluency, a combination of critical literacy, ethical reasoning, metacognitive awareness, and creative agency. Pedagogical approaches should encourage students to interrogate AI as a constructed artefact, not a neutral tool. Drawing on principles from critical pedagogy, engaged learning, and experiential inquiry, AI education should empower learners to ask whose knowledge is embedded in AI outputs, which voices are missing, and how AI systems intersect with social justice, sustainability, and democratic values. This holistic approach will equip students not only to use AI, but to shape its future development in line with public good.

AI education must go beyond coding and technical proficiency, encompassing the ethical, societal and legal dimensions of AI use

Roundtable participants highlighted the value of a national framework that embeds age-appropriate AI literacy from an early age. This includes teaching how data is collected and used, what algorithmic bias is, and how to question the outputs of AI systems. At the university level, responsible AI should be a cross-cutting theme in curriculum design, ensuring that all disciplines are equipped to deal with AI's implications. Co-designed training with employers and civil society can help ensure relevance and responsiveness. Lifelong learning must also be prioritised, with flexible and accessible training routes for mid-career and older workers. Public investment and incentives will be critical to scaling these efforts.

The integration of AI into education offers profound opportunities for enhancing learning, creativity, and personalisation. However, it also presents serious ethical challenges that, if not addressed proactively, risk undermining core educational values such as critical thinking, independent inquiry, and social responsibility.

Policy makers should recognise that while AI can be a powerful tool for innovation, it is not a neutral one. AI systems are shaped by the data they are trained on and by the assumptions of their creators, meaning biases both historical and systemic are embedded within their outputs. Additionally, the environmental impacts of AI technologies must not be ignored. Recent cases, such as the emergence of water-intensive data centres in drought-stricken areas of Mexico, highlight how AI-driven development can exacerbate existing inequalities. In towns like Querétaro, local farmers are facing severe water shortages while nearby facilities consume vast amounts of the scarce resource to power AI operations underscoring the urgent need for critical oversight to ensure that technological advancement does not deepen social and environmental injustices.

In educational settings, these biases and challenges are often overlooked by students, who tend to view AI outputs as objective, authoritative, and complete.

Few students instinctively question whose perspectives are represented in AI responses, what information might be omitted, or what ethical assumptions underpin the results they receive. This lack of critical engagement poses a major risk to the development of independent, analytical thinkers which is an essential goal of education.

To address this, policy must not only advocate for ethical AI use but must also provide a clear framework guiding students on how and when to use AI.

Few students
instinctively question
whose perspectives are
represented in AI
responses, what
information might be
omitted, or what ethical
assumptions underpin
the results they receive.

Questions such as: *Whose voices are missing? What perspectives dominate? What ethical concerns arise?* must become routine elements of working with AI technologies.

Students should be explicitly taught that AI is appropriate for certain tasks, for example in brainstorming ideas, generating initial outlines, or suggesting new perspectives, but not a substitute for research, source validation, or original critical analysis. They must be encouraged, and equipped, to always seek corroborating information from diverse, credible sources of information and knowledge beyond AI tools.

Furthermore, a foundation in understanding AI bias must be a core part of the educational experience. Students need structured opportunities to critically examine AI outputs, learning to identify potential biases, question assumptions, and assess the trustworthiness of information. Questions such as: *Whose voices are missing? What perspectives dominate? What ethical concerns arise?* must become routine elements of working with AI technologies.

Educators also require robust support to model this critical engagement and to build students' AI literacy in tandem with creativity and innovation. AI should be framed not as a replacement for human thinking but as a partner in a broader, thoughtful learning process. When used critically and responsibly, AI can enrich students' creative capacities, enabling richer storytelling, problem-solving, and exploration.

Educators themselves have a vital role to play beyond curriculum design. As role models and mentors, they must demonstrate responsible and reflective AI use, encouraging students to approach AI outputs with curiosity and ethical caution. Universities also have a civic responsibility to foster public understanding of AI beyond their own student body—through public lectures, partnerships with civil society, and the development of open educational resources. Institutional leadership is needed to build an AI-literate academic workforce across disciplines and to ensure that AI is integrated into education in ways that enrich, rather than erode, critical inquiry and independent thought.

To meaningfully train academics in responsible AI use, institutions must go beyond basic tool training and nurture critical, creative, and pedagogically grounded engagement. This involves tiered professional development, from foundational workshops on AI literacy and ethics, to discipline-specific communities of practice that explore real-world scenarios and co-create new teaching approaches. AI champions across departments can model innovation while ensuring relevance to local pedagogic contexts. Crucially, support structures must reward reflection and experimentation, embedding AI into the heart of academic development, specifically not as a bolt-on, but as a catalyst for reimagining education.

The future of AI in education depends on maintaining a deliberate balance between innovation and responsibility. AI can help students think differently, but it must not lead them to think less. Policymakers have a critical role in shaping an educational landscape where students are empowered to use AI as a creative tool but always with critical awareness, ethical reflection, and a commitment to sourcing diverse, accurate information.

We recommend the following:

- Developing a national framework for age-appropriate AI literacy across all phases of schooling, from primary to secondary education.
- Requiring all higher education institutions to integrate AI literacy and critical AI awareness as graduate attributes across all disciplines.
- Embedding AI literacy and ethical awareness into further education curricula and apprenticeship standards, ensuring learners across vocational pathways are equipped for an AI-augmented workplace.
- Funding structured professional development programmes to build AI literacy among teachers, lecturers, and university staff.
- Investing in public education campaigns and community-based learning to promote inclusive understanding of AI among adult learners and digitally excluded groups.

3.3. Investing in ethical AI infrastructure

Building an ethical AI infrastructure requires investment not just in technologies, but in the principles and capacities that guide their use. The roundtable revealed strong support for sustainability, transparency and public accountability as foundational elements of any future-ready AI ecosystem.

Participants noted that many AI systems are energy-intensive and environmentally costly, underscoring the importance of sustainable design. Funding mechanisms should prioritise low-carbon AI development and support research into green computing techniques. Infrastructure should also support traceability and accountability—whether through robust audit trails, interoperable platforms, or open-source development practices.

Public sector bodies, especially at local levels, require tailored guidance and tools to assess and procure ethical AI. Investing in open access models and shared procurement frameworks can ensure value for money and reduce duplication. AI systems that serve the public must be interoperable, secure, and aligned with human rights standards.

Funding mechanisms should prioritise low-carbon AI development and support research into green computing techniques.



Ethical AI infrastructure also includes social infrastructure: institutions and norms that support ethical decision-making, participatory governance, and transparent evaluation. Embedding ethics within the design and deployment process, rather than treating it as an afterthought, will be critical for long-term societal trust in AI systems. Independent auditing should be adopted where resources allow to address ethical issues, which are often socio-culturally sensitive.

We recommend the following:

- Prioritise sustainable AI development by investing in low-carbon technologies, green computing, and environmentally responsible infrastructure.
- Embed transparency and accountability through interoperable systems, robust audit trails, and open-source development practices.
- Strengthen public sector capacity with clear procurement guidance, shared frameworks, and practical ethical AI toolkits at local and national levels.
- Institutionalise ethics from the outset by integrating ethical standards throughout AI design, deployment, and evaluation processes.
- Enable independent oversight through third-party auditing and participatory governance models to build public trust and accountability.

3.4. Promoting inclusive and ethical data governance

AI systems are only as fair and effective as the data that powers them. Current data governance models often privilege corporate or state interests over those of individuals and communities. Participants identified significant gaps in consent processes, transparency, and representativeness in data collection and usage.

To create a more inclusive AI ecosystem, the UK must prioritise ethical data governance frameworks that are rooted in rights, accessibility and agency.¹⁸ This includes clear standards for data sharing, consent, and accountability. Investment in local data initiatives, such as community data trusts and cooperatives, can help embed lived experience into AI development. These models offer mechanisms for democratic oversight and ensure that benefits derived from data are shared more equitably.

Digital exclusion remains a significant barrier to inclusive data practices. Millions still lack the devices, connectivity or skills required to participate in the digital economy. This not only entrenches inequality but also results in skewed data that fails to represent the full spectrum of society. Addressing data poverty requires a mix of infrastructure investment, inclusive education, and targeted policy interventions.

Participants also called for broader data literacy among citizens to strengthen consent processes and empower users. Ensuring people understand how their data is collected, processed and used will be key to enabling more equitable participation in the AI economy.

We recommend the following:

- Establish rights-based data governance frameworks that prioritise transparency, consent, and accountability over corporate or state interests.
- Invest in community-led data initiatives such as data trusts and cooperatives to embed lived experience and support democratic oversight.
- Address digital exclusion through targeted investment in infrastructure, connectivity, and skills to ensure representative and equitable datasets.
- Enhance public data literacy so individuals can understand, question, and control how their data is collected and used.
- Set clear national standards for ethical data sharing to promote trust, interoperability, and equitable distribution of AI benefits.

¹⁸ <https://a.storyblok.com/f/313404/x/a709567d3f/costar-national-lab-ai-copyright-framework-report-may-2025.pdf>

3.5. Certification for AI that promotes human well-being and flourishing

One of the central issues identified during the roundtable was the misalignment between the core incentives driving AI development and the actual needs of human beings and societies. Currently, many AI systems are evaluated and monetised based on metrics such as user engagement, time-on-platform, and behavioural manipulation, rather than their capacity to foster well-being, autonomy, mental health, or social connection.

Participants proposed the development of a certification scheme for AI systems that demonstrably support human well-being and flourishing. The aim is to provide both a practical framework and a market incentive for companies to prioritise outcomes that align with long-term human interests. This would not only help distinguish ethical and socially beneficial innovations but also empower investors, consumers, and policymakers to reward companies that act in the public interest.

The certification would serve multiple purposes:

- Offer a voluntary but rigorous benchmark for AI products, platforms, and organisations committed to enhancing human flourishing
- Shift business incentives by linking ethical and human-centred design to market advantage
- Encourage public and private funders to use well-being impact as a criterion for investment or procurement
- Provide a basis for cross-sector collaboration on measurement standards, including mental health, digital dignity, civic trust, and relational well-being

Ethical data governance and infrastructure investments

- Establish rights-based data governance frameworks
- Invest in community-led data initiatives
- Address digital exclusion through targeted investment in infrastructure, connectivity, and skills
- Enhance public data literacy
- Invest in low-carbon technologies and environmentally responsible infrastructure.
- Embed transparency and accountability through interoperable systems

Human well-being, trust & accountability outcomes

- Offer a voluntary but rigorous benchmark for AI products, platforms, and organisations committed to enhancing human flourishing
- Shift business incentives by linking ethical and human-centred design to market advantage
- Encourage public and private funders to use well-being impact as a criterion for investment or procurement
- Provide a basis for cross-sector collaboration on measurement standards



Equitable and transparent AI development and deployment for diverse communities and long-term societal resilience

Responsible AI in Education & Collaborative Framework

- Establish co-governed AI research programmes
- Expand structured secondment schemes
- Create joint doctoral and professional education programmes
- Develop a national framework for age-appropriate AI literacy
- Fund structured professional development programmes
- Invest in public education campaigns and community-based learning

Figure 2. A high-level implementation pathway showing how ethical data governance and infrastructure investments lay the foundation for education and collaboration, leading to trust, accountability, and human well-being.

4. SUMMARY AND NEXT STEPS

AI is reshaping the societal, economic and ethical fabric of the UK. This report has highlighted key opportunities and risks in ensuring AI development and deployment are equitable, transparent and grounded in public interest. Across the themes explored, education, governance, data, innovation and employment, a consistent message emerged: the UK must adopt an inclusive and socially purposeful approach to AI that reflects the values of its diverse communities and promotes long-term societal resilience.

The discussions presented in this report reflect a clear imperative: we cannot leave AI to evolve in isolation from democratic oversight, public engagement and ethical scrutiny. Universities, as convenors of knowledge and trusted civic institutions, have a central role to play. Royal Holloway, University of London, through the Digital Organisation and Society (DOS) research centre, is committed to advancing this agenda by fostering collaboration between academia, industry and government. As a university guided by social purpose, RHUL is uniquely positioned to ensure that AI innovation serves the common good.

NEXT STEPS

Looking ahead, the DOS Research Centre will continue to lead the conversation on AI for societal readiness. Our next phase of work will bring together government, industry, and education partners to co-design solutions that build inclusion, ethical awareness and creativity across every stage. We see education not as a downstream beneficiary of AI, but as its moral compass and catalyst, shaping how intelligent technologies serve people and society. Guided by the values of our institution, openness, innovation and social purpose, DOS is committed to ensuring that the UK's AI future is one where every individual is empowered to thrive with curiosity, integrity, and imagination.

Presented by DOS Research Centre

AI FOR SOCIAL PURPOSE

ADDRESSING KEY INDUSTRY
AND POLICY CHALLENGES

**THE DIGITAL
ORGANISATION
AND SOCIETY**

**(DOS) RESEARCH
CENTRE**

Website :
DOSrhul.org

Address :
Royal Holloway, University of
London

Email address :
dosdirectors@rhul.ac.uk

ISBN 978-1-0369-7408-4

Publisher:

Digital Organisation and Society (DOS) Research
Centre, Royal Holloway University of London.

Publication year: 2025

