

## Written evidence submitted by Sense about Science

Sense about Science is an independent charity that promotes the public interest in sound science and evidence. We work with communities all over the UK to make sense of evidence in key areas of decision making, and with researchers and policy makers to improve use of evidence in public life. We have over 20 years' experience in equipping people with good questions to navigate the most complex and controversial areas of evidence and policy making, - from parents of children undergoing heart surgery, to politicians faced with output from simulation models – and we have led successful international campaigns such as AllTrials, to regularise transparency standards in evidence for decisions.

We work closely with the House of Commons Library and the Parliamentary Office of Science and Technology (POST) to deliver Evidence Week in Parliament each year, bringing together the public, MPs and researchers with briefings on cutting edge research and discussions of how parliament uses, and scrutinises, evidence.

## How can Government and financial regulators strike the right balance between seizing the opportunities of AI but at the same time protecting consumers and mitigating against any threats to financial stability?

There is a recognised governance gap in the adoption of AI tools and tech. The speed of innovation has produced a situation where government, businesses and other users do not have processes to properly assess the suitability of a tool, risks of implementation and the reliability of its outputs. This is resulting in lost opportunities to innovate as well as exposure to new types of risk. The Final Report of the UN High-level Advisory Body on Artificial Intelligence<sup>1</sup> noted: *“There is, today, a global governance deficit, with respect to AI”*. The report highlights *“Stability of financial systems, critical infrastructure and institutions”* as an AI-related risk (Box 4, p31) and concludes that *“A comprehensive approach to AI safety involves understanding the capabilities of advanced AI models, adopting standards for safe design and deployment, and evaluating both the systems and their broader impacts.”* (Box 10, p57). This echoed concerns raised by the UK government in May 2024: *“We will only unlock the benefits of AI if we can manage these risks. At present, our ability to develop powerful systems outpaces our ability to make them safe. The first step is to better understand the capabilities and risks of these advanced AI systems.”*<sup>2</sup>

In their 2024 assessment in *Science*, Yoshua Bengio and 15 other leading voices on AI state that, *“Present governance initiatives lack the mechanisms and institutions to prevent misuse and recklessness...”* and call for *“adaptive governance mechanisms for a more commensurate preparation.”*<sup>3</sup>

**We ask the committee to consider advocating application in the finance sector of the Responsible Handover framework, which has recently been developed in the biomedical research sector. It is a simple governance tools that transfers sound, well-tested engineering principles to digital infrastructure<sup>4</sup>. By focusing on governing the points when tools are passed from hand to hand (handover), and guiding the exchange, it is a practical answer to the question of how AI and advanced data tools can be safely and effectively implemented, especially in the context of low and no-code settings.**

AIFS0055

## Closing the governance gap with responsible handover

In all sectors there is growing and justified concern about threats to safety resulting from AI and advanced data tools. There is already evidence of harm – to individuals, systems and infrastructure - caused where organisations lacked key information about an AI tool: AI-directed autonomous surgery used on the wrong body type, extreme weather prediction misunderstood, security systems opened to hacking and people being supplied with the wrong electrical regulations for example. Organisations are worried that they are unable to properly assess the suitability of a tool, risks of implementation and the reliability of its outputs.

Demand for practical help in commissioning and using AI is consequently high, but not being met. As AI tools pass from hand to hand (research to developer, developer to user, user to user), key information is lost that would enable them to be used safely and responsibly. Governance tools that close this gap would enable confident, responsible adoption (or rejection) of AI applications.

Responsible handover is based on sound engineering principles of commissioning. The concept of structured handover for physical infrastructure is well established, with clear handover points when building a bridge, for example, from constructor to commissioner. Working with the Wellcome Trust, Institution of Engineering and Technology (IET), Turing Institute and others, Sense about Science adapted this to digital infrastructure, tested it in different handover settings internationally and refined it.

The Responsible Handover Framework<sup>4</sup> acts as an overview, a guide to what a handover must cover. It is unique in going as far as possible with a universal framework, prompting people to identify issues and standards specific to their activities. Since its launch, the Responsible Handover Framework has been embraced by both grassroot users and major bodies and government agencies around the world as a practical answer to the pressing governance gap, including: AAPT, International Science Council, Australian national research funders, UK Quality Management System for devices and the American climate funder network. When applied to clinical settings, for example, initial work has shown that the framework maps well to ISO9001 and other regulatory requirements: clinical software developers at Guy's and St Thomas' hospital have estimated that, by following the framework, developers arrive at the gateway of ISO13485 and ISO14971 with around 70% of the work already done, reducing R&D waste.

The principles of responsible handover are universal, and applicable across sectors and along the supply chain from discovery to implementation. Adoption of responsible handover in financial services would promote a streamlined, universal approach so that researchers, developers and adopters are working to a similar script – or set of expectations - for adoption across sectors and along the development chain feeding into financial products.

## Securing transparency and accountability in the face of rapid adoption

According to a recent survey by the Bank of England and FCA<sup>5</sup>, 75% of firms already use AI with further 10% planning to introduce it in the next three years. 46% of respondent firms say they have only partial understanding of the AI technologies they use (largely due to use of third-party models). The lack of transparency and accountability of AI systems have been identified

AIFS0055

as increasing vulnerability of financial services<sup>6</sup> and markets<sup>7</sup>, as well as posing a challenge to privacy and regulatory compliance.<sup>8</sup> Lack of transparency also has the potential to harm trust in systems and institutions,<sup>9</sup> for actual or perceived unfairness or bias,<sup>10,11</sup> or unwanted consequences. Public acceptance of financial services using AI would also be slowed by system failures resulting from ill-advised implementation.<sup>12</sup>

**Handover governance ensures that assessment and mitigation of these issues are integrated into the development and implementation of AI, enabling the fastest safe roll out of appropriate technology; we would welcome the opportunity to speak further with the committee about the application of the responsible handover framework in other sectors, and its relevance to financial services.**

***April 2025***

---

<sup>1</sup> United Nations (2024) Governing AI for Humanity (A/74/821)  
<https://doi.org/10.18356/9789211067873>

<sup>2</sup> DSIT (2024) International scientific report on the safety of advanced AI (DSIT 2024/009)  
<https://www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai>

<sup>3</sup> Bengio, Y. *et al* (2024) Managing extreme AI risks amid rapid progress: Preparation requires technical research and development, as well as adaptive, proactive governance *Science* **384**: 6698

<sup>4</sup> Sense about Science (2024) Responsible Handover: How to share and implement data analytics and AI tools. <https://senseaboutscience.org/wp-content/uploads/2024/11/Responsible-Handover-implementation-guide.pdf>

<sup>5</sup> Bank of England (2024) Artificial intelligence in UK financial services  
<https://www.bankofengland.co.uk/report/2024/artificial-intelligence-in-uk-financial-services-2024>

<sup>6</sup> Owolabi, O. *et al* (2024) Ethical Implication of Artificial Intelligence (AI) Adoption in Financial Decision Making <http://dx.doi.org/10.5539/cis.v17n1p49>

<sup>7</sup> BIS (2024) Intelligent financial system: how AI is transforming finance (Working Paper No 1194) <https://www.bis.org/publ/work1194.pdf>

<sup>8</sup> Uzougbo, NS *et al* (2024) Legal accountability and ethical considerations of AI in financial services *GSC Advanced Research and Reviews* **19**(02), 130–142.  
<https://doi.org/10.30574/gscarr.2024.19.2.0171>

<sup>9</sup> Rajesh Soundararajan, R & Shenbagaraman, VM (2024) Enhancing Financial Decision-Making Through Explainable AI And Blockchain Integration: Improving Transparency and Trust in

---

Predictive Models *Educational Administration: Theory and Practice* **30**(4), 9341-935  
<https://www.kuey.net/index.php/kuey/article/view/3672/2398>

<sup>10</sup> Chen, Z (2023) Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanit Soc Sci Commun* **10**, 567 <https://doi.org/10.1057/s41599-023-02079-x>

<sup>11</sup> Glickman, M., Sharot, T. (2025) How human–AI feedback loops alter human perceptual, emotional and social judgements. *Nat Hum Behav* **9**, 345–359 <https://doi.org/10.1038/s41562-024-02077-2>

<sup>12</sup> Habbal, A et al (2024) Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges and future research directions. *Expert Systems with Applications* **240**: 122442 <https://doi.org/10.1016/j.eswa.2023.122442>