Written evidence submitted by Professor Jo Bates, Professor. Helen Kennedy, Dr Susan Oman, Dr Itzelle Medina Perea and Lulu Pinney (UAIG0008)

Author introduction

The submission is authored by Prof. Jo Bates, Prof. Helen Kennedy, Dr Susan Oman, Dr Itzelle Medina Perea, Lulu Pinney of the University of Sheffield. The evidence is informed by our research on the Living With Data research project funded by the Nuffield Foundation (OSP/43959), on which we worked closely with the Department for Work and Pensions and the BBC, as well as Medina Perea's PhD research about the NHS.

Evidence

The Living With Data research team advocates for "socially meaningful transparency" (Bates et al, 2024) about the data and AI systems adopted within government and the public sector. Socially meaningful transparency focuses on **enhancing public understanding of AI systems for informed use and democratic engagement in datafied societies.** This is important given the widely evidenced risks of AI e.g. algorithmic bias and discrimination, that publics are increasingly aware about. Socially meaningful transparency prioritises the needs and interests of members of the public over those of AI system developers.

We identify three main barriers to socially meaningful transparency based on our empirical research with Department for Work and Pensions, BBC and NHS:

- **Information Asymmetry:** Organisations such as government departments have significantly more information about their AI systems than third parties, limiting accurate insights into what systems are being used (and planned) and how.
- **Uncertainty:** AI systems constantly evolve, making it difficult to provide stable and accurate information. Further, some AI systems are technically black-boxed by design.
- **Resourcing:** Producing socially meaningful transparent accounts of complex AI systems is resource-intensive, demanding expertise and time.

To address these challenges, we propose seven principles for socially meaningful transparency that ought to be considered in the context of AI system adoption in government:

1. Reduce information asymmetries between government and non-commercial third parties about operational or proposed AI (and related) systems, through mandated registration of AI and related systems and enhanced legal rights to access information about data and AI systems to enable informed public discourse. While the government already announced a new mandatory **Algorithmic Transparency Recording Standard** in February 2024 that would partially meet this proposal, recent reports suggest this guidance is not being followed by departments.
2. Enhance collaborative governance around what is made transparent by fostering discussion and decision making between government and non-commercial third parties, including members of the public, about **which AI systems and what about them** ought to be prioritised for making transparent for different purposes.

3. Recognise and **communicate the potential societal impacts** of these selected data and AI systems to the public, considering both potential benefits and harms, thus enhancing AI literacy and better informed public discourse.
4. Avoid obfuscation by providing **clear, concise, and accessible information**, tailored to different audiences and contexts.
5. Acknowledge and **communicate the inherent uncertainties** within data and AI systems, recognising their dynamic nature.
6. Recognise that **transparency practices can occur at various stages** of AI system development and implementation, and capitalise on opportunities for public engagement at different points.
7. **Commit necessary resources** to socially meaningful transparency, ensuring sustainability and recognising the political nature of resource allocation.

These proposals aim to shift the balance of power in transparency practices, enabling informed decision-making and democratic participation in datafied societies. We recommend policy interventions to strengthen data system disclosure requirements, establish collaborative governance mechanisms, and provide guidance for communicating uncertainty and potential harms of data systems.

Further in depth discussion can be found in our published paper:

Bates, J. , Kennedy, H., Medina Perea, I. et al. (2024) *Socially meaningful transparency in data-based systems: reflections and proposals from practice.* Journal of Documentation, 80 (1). pp. 54-72

An open access version is available here: https://eprints.whiterose.ac.uk/198835/

January 2025