

Supplementary written evidence submitted by The Royal Society (ADM0021)

1. The Royal Society is the National Academy of Science for the UK and the Commonwealth. It is a self-governing Fellowship of many of the world's most distinguished scientists working across a broad range of disciplines in academia and industry. The Society draws on the expertise of its Fellows and Foreign Members to provide independent and authoritative scientific advice to UK, European and international decision makers
2. The Society responded to the original call for evidence for this inquiry in April of this year, and welcomes the opportunity to reiterate some key messages and provide further information. Our original submission drew heavily on the findings of our report *Machine learning: the power and promise of machines that learn by example*. We have since published a report jointly with the British Academy, *Data management and use: Governance for the 21st Century* that makes recommendations for a 21st century data governance system that are relevant to this inquiry.
3. Key messages from our report on *Machine learning* were summarised in our original submission to the Committee. We outlined that machine learning is a potentially transformative technology, which brings with it both opportunities and challenges, and whose risks and benefits need to be navigated as its use becomes more central to everyday activities. Our report recommended action in key areas to:
 - Create an amenable data environment;
 - Build skills at all levels;
 - Support businesses in using machine learning;
 - Encourage public dialogue; and
 - Support machine learning research.
4. The Royal Society set out a new wave of machine learning research, which – if pursued – could accelerate the development of the field, while addressing societal challenges and helping to ensure continued public confidence in these systems. These areas include algorithm interpretability, robustness, privacy, fairness, inference of causality, human-machine interactions, and security. Funding bodies should encourage and support research applications in these areas, though not to the exclusion of other areas of machine learning research.
5. The technological capabilities of machine learning enable new uses of data, which challenge existing data governance systems; its new applications raise questions about public confidence and acceptability. There are governance issues surrounding the use of data, including those concerning the sources of data, and the purposes for which they are used. For this, a new framework for data governance¹ – one that can keep pace with the challenge of data governance in the 21st century – is necessary to address the novel questions arising in the new digital environment. We make recommendation for the nature of this framework in the *Data management and use* report. It hinges on establishing a set of principles for data governance, set out in the annex below, which flow from the central objective of making human flourishing central to data governance and seek to ensure trustworthiness and trust in the management and use of data as a whole; and on establishing a stewardship body that will oversee the data governance landscape to ensure that there is adequate governance for new uses of data.

¹ We consider data governance to mean everything designed to inform the extent of confidence in data management, data use and the technologies derived from it. We cannot properly consider this by treating data management or data use individually, or separately from each other. Therefore throughout this submission we use the term 'data governance' as shorthand to refer to the governance of data management and data use. In instances where the distinction between the governance of data management and the governance of data use is relevant, we aim to make that clear.

Supplementary written evidence submitted by The Royal Society (ADM0021)

6. There are two important aspects of our recommendations that will be helpful to the Committee's inquiry. First is that it is not appropriate to set up governance structures for machine learning *per se*. The range and extent of use of machine learning algorithms is potentially extremely broad and extremely diverse. In many or most contexts machine learning is generally uncontroversial, and where questions do arise, many of them are very context specific. This means that it would be unhelpful to create a general governance framework or governance body for all machine learning applications. While there may be specific questions about the use of machine learning in specific circumstances, these should be handled in a sector-specific way, rather than via an overarching framework for all uses of machine learning; some sectors may have existing regulatory mechanisms that can manage, while in others there may not be these existing systems.
7. Secondly, we need to take a *connected* approach to the governance of data and its uses. There are already significant governance arrangements in place across the landscape of data use and the framework that we set out in the *Data management and use* report is premised on this. What is now needed is a means for ensuring that debate on the governance of data and its uses is *connected* across different sectors (while retaining the flexibility to respond and to support innovation in a wide variety of contexts), as the proposed stewardship body would do.
8. We would welcome the opportunity to clarify these points.

November 2017

Annex: Principles for data governance:

The promotion of human flourishing is the overarching principle that should guide the development of systems of data governance. The four principles that follow provide practical support for this overarching principle across the varied ways data is managed and used:

- protect individual and collective rights and interests
- ensure that trade-offs affected by data management and data use are made
- transparently, accountably and inclusively
- seek out good practices and learn from success and failure
- enhance existing democratic governance.