Written evidence submitted by the British Academy (BIG0064)

The British Academy welcomes this important inquiry into the opportunities and risks of big data. The UK's investment in a world-class data infrastructure is unrivalled by almost any other country. As the national academy for the humanities and social sciences, we have a longstanding interest in the opportunities afforded to research and innovation through investments such as this.

In considering the next stages of the inquiry, we would urge members to prioritise the main issue that underlies our ability to take advantage of these opportunities (and to mitigate associated risks): our capacity to understand and interpret data. Our submission refers to the UK's critical quantitative skills deficit and the interventions that the British Academy has initiated. Successive governments have recognised the importance of this agenda, but the urgency of this challenge demands that much more should be done. Government leadership is required for investment in, and coordination and evaluation of, interventions across education, employment, research and innovation.

We would welcome the opportunity to engage with you further on this inquiry.

INTRODUCTION

1. Big data is revolutionising both how we see the world and how we engage with it. Examples range from the near real-time tracking of global deforestation and the human genome project's mapping of our genes, to the monitoring of people's consumer preferences through their use of social media. The ubiquity of data makes the understanding and interpreting of statistics an essential feature of life in the 21st century.

2. We welcome the attention of the Committee on this important issue. In considering the next stages of the inquiry, we would urge members to prioritise the main issue that underlies our ability to take advantage of big data (and to mitigate associated risks): our capacity to understand and interpret data.

OPPORTUNITIES OF BIG DATA

3. *An economic prize.* Big data has been described by the Government as one of the 'Eight Great Technologies' in which the UK has the chance of becoming a world leader. The way in which we handle the volume and variety of information now available – together with the technology required to process it – will reap rewards for individuals, companies and the economy as a whole. Quantifying the precise value to our economy is difficult, but studies have estimated that between 2012 and 2017, 58,000 new jobs a year may be created in the UK in the big data marketplace; and big and open data are predicted to contribute an extra £147 billion to GDP across the EU economies per year by 2020.

4. *Understanding the world in which we live.* The sheer volume of data currently being generated provides significant opportunities for policymakers and researchers. Statistical analyses of large and complex data sets underpin the deciphering of social patterns and trends, and the evaluation of the impact of social interventions.

5. As a world leader in the social sciences, pride of place in the UK has gone to empirical studies of social phenomena founded on rigorous, scientific data collection and innovative analysis. Over the last six decades, the UK has invested in a world-class social science data infrastructure unrivalled by almost any other country. This includes the British birth cohort studies, which have collected vital data on education, employment, social attitudes, physical and mental health. Other big data has been amassed by national statistical agencies, for example the Census or General Household Survey, and still more by government departments. More open access to large scale databases and the increase in data generated by a digital society – combined with our increasing data-processing power – create tremendous opportunities for both research and evidence-based policy making.

6. *Data for better policy making.* Used well, big data offers the opportunity for enhancing the process of evidence-based policy making[1]. Through it, governments are better able to understand the needs of the population, and in turn exploit data to both improve policy making and evaluate the impact of policies.

7. *Administrative data.* The British Academy recently held a series of roundtable discussions in partnership with the British Society for Population Studies to examine how data from population studies can help tackle some of the most pressing policy problems of our day. Discussions highlighted the importance of administrative and transactional data to research and policymaking. Administrative data serve to show us what people actually do (how they spend their money, demonstrate social media preferences or communicate via email) in a way that allows us to develop a much richer understanding of the way people live. The Administrative Data Research Network[2] is helping to facilitate better use of data for research and policy.

CHALLENGES AND RISKS

8. *Understanding 'big data'.* The umbrella term 'big data' refers to data sets of different varieties, with different uses (e.g. consumer choices, monitoring trends in health, crime etc). While there are several challenges to the use of data more generally, a greater understanding of the different kinds of data is required in order to appreciate the challenges and risks unique to each context. It is important, too, to be realistic about what big data can capture. The uncertainties built into

---

[1] See the Royal Statistical Society's Data Manifesto: http://www.rss.org.uk/Images/PDF/influencing-change/rss-data-manifesto-2014.pdf
[2] See http://adrn.ac.uk/

modelling techniques demand careful interpretation of trends and forecasts. Interpretation of this kind requires bringing together computational and mathematical expertise and diverse qualitative and interpretive skills in the handling of big data.

9. *Data capability*. The British Academy recently published a comprehensive review[3] of the demand and supply of quantitative skills – which we define as the ability to reason using numbers. Our evidence shows that the UK is facing a crisis in levels of numeracy, whilst there is a rise in the number of jobs requiring these skills. About seven in 10 employees say that quantitative skills are essential or important in carrying out their work. While roughly three in 10 jobs require basic arithmetic skills, a further four in 10 require the ability to apply quantitative skills to a more advanced level. There is evidence that demand for more advanced skills, which might range from the ability to use descriptive statistics to highly complex mathematical procedures, has risen sharply in the past two decades, with the proportion of employees saying advanced mathematics or statistics are important in their jobs rising from 29 per cent in 1997 to 38 per cent in 2012. Correspondingly, the number of people reporting that arithmetical skills are not at all important in their jobs has declined.

10. This presents long-term challenges for our education system (more on this below), which is charged with improving the data skills of new entrants to the labour market – but also immediate challenges to the skill sets of our existing workforce.

11. *Risks of administrative data.* While administrative data is increasingly an important source for understanding many social issues, its collection and use also present a number of risks and challenges. The use of administrative data for research purposes is necessarily *secondary* analysis, in that it draws on data gathered for a different purpose – so there may be mismatches between the data and research questions. There may also be issues of consent if the data is being used for something other than its original purpose. Administrative and transactional data is also subject to external influences on its accuracy and range, and cannot be controlled in the way that research data can. Therefore it must be handled in a way that is opposite to how it has been generated and collected.

SKILLS GAPS AND PUBLIC UNDERSTANDING

12. Our ability to harness the opportunities of big data, while also understanding the challenges and risks associated with it, hinge on one of the underpinning factors of this inquiry: data capability. There is an urgent need to raise the data skills of the current population, as well ensure that there is a structure in place to ensure these skills flourish at all levels.

---

[3] See: http://www.britac.ac.uk/policy/count_us_in_report.cfm

13. There are serious, long-term deficits in the UK's data capacity. The UK's performance in mathematics generally makes it only middle-ranking among developed nations. Undergraduates embark on degree courses with varying, and often weak, fluency in statistics. In universities in the USA, Germany, the Netherlands, Belgium and Switzerland students typically develop a much better grasp of data than in even the best UK degree programmes.

14. One of the Academy's major concerns regarding data capability within education is in stressing that this is not just an issue for science, technology, engineering and mathematics (STEM). While we recognise the efforts of policymakers in emphasising the importance of STEM in reforms to curricula, big data makes statistical and numerical understanding relevant across *all* disciplines.

15. These skills gaps are closely linked to issues relating to broader public understanding of data, its collection and use. Providing citizens with the means to understand, analyse and criticise data becomes ever more integral to the functioning of a democracy. Our recent report, *Count Us In*, offers a vision of how the UK can rise to the transformational potential of big data, creating a generation of citizens, consumers, students and workers as comfortable with numbers as they are with words. The general up skilling of the population in data-handling will inevitably have an impact on the issue of building public trust around the use of people's personal data.

16. Concerned about the skills deficits that exist in the UK's data and quantitative capacity, the British Academy convenes a High Level Strategy Group comprised of major funders and users of data – including the Economic and Social Research Council, Higher Education Funding Council for England, UK Statistics Authority, Department for Business Innovation and Skills, Department for Education and Universities UK – with the aim of ensuring a joint and coordinated approach to addressing the deficit across education and employment.

17. Successive governments have recognised the importance of this agenda, putting in place interventions aimed at improving the confidence and fluency in handling and interpreting data. But the potential and urgency of the big data dilemma demands that much more should be done. There needs to be a concerted and coordinated national effort – with leadership from government.

September 2015