House of Commons

Science and Technology Committee

Robotics and artificial intelligence

Fifth Report of Session 2016–17

Report, together with formal minutes relating to the report

Ordered by the House of Commons to be printed 13 September 2016

HC 145
Published on 12 October 2016
by authority of the House of Commons

Science and Technology Committee

The Science and Technology Committee is appointed by the House of Commons to examine the expenditure, administration and policy of the Government Office for Science and associated public bodies.

Current membership

Dr Roberta Blackman-Woods MP (*Labour, City of Durham*)

Victoria Borwick MP (*Conservative, Kensington)*

Stella Creasy MP (*Labour, Co-op, Walthamstow*)

Jim Dowd MP (*Labour, Lewisham West and Penge*)

Chris Green MP (*Conservative, Bolton West*)

Dr Tania Mathias MP (*Conservative, Twickenham*)

Carol Monaghan MP (*Scottish National Party, Glasgow North West*)

Graham Stringer MP (*Labour, Blackley and Broughton*)

Derek Thomas MP (*Conservative, St Ives*)

Matt Warman MP (*Conservative, Boston and Skegness*)

The following were also members of the committee during the parliament:

Nicola Blackwood MP (*Conservative, Oxford West and Abingdon) (Chair of the Committee until 19 July 2016*)

Liz McInnes MP (*Labour, Heywood and Middleton*)

Valerie Vaz MP (*Labour, Walsall South*)

Daniel Zeichner MP (*Labour, Cambridge*)

Powers

The Committee is one of the departmental select committees, the powers of which are set out in House of Commons Standing Orders, principally in SO No 152. These are available on the internet via www.parliament.uk.

Publication

Committee reports are published on the Committee's website at www.parliament.uk/science and by Order of the House.

Evidence relating to this report is published on the relevant inquiry page of the Committee's website.

Committee staff

The current staff of the Committee are: Simon Fiander (Clerk); Marsha David (Second Clerk); Dr Elizabeth Rough (Committee Specialist); Martin Smith (Committee Specialist); Darren Hackett (Senior Committee Assistant); Julie Storey (Committee Assistant); and Nick Davies (Media Officer).

Contents

Summary

After decades of somewhat slow progress, a succession of advances have recently occurred across the fields of robotics and artificial intelligence (AI), fuelled by the rise in computer processing power, the profusion of data, and the development of techniques such a 'deep learning'. Though the capabilities of AI systems are currently narrow and specific, they are, nevertheless, starting to have transformational impacts on everyday life: from driverless cars and supercomputers that can assist doctors with medical diagnoses, to intelligent tutoring systems that can tailor lessons to meet a student's individual cognitive needs.

Such breakthroughs raise a host of social, ethical and legal questions. Our inquiry has highlighted several that require serious, ongoing consideration. These include taking steps to minimise bias being accidentally built into AI systems; ensuring that the decisions they make are transparent; and instigating methods that can verify that AI technology is operating as intended and that unwanted, or unpredictable, behaviours are not produced. While the UK is world-leading when it comes to considering the implications of AI, and is well-placed to provide global intellectual leadership on this matter, a coordinated approach is required to harness this expertise. A standing Commission on Artificial Intelligence should be established with a remit to identify principles to govern the development and application of AI, provide advice to the Government, and foster public dialogue.

Advances in robotics and AI also hold the potential to reshape, fundamentally, the way we live and work. Improvements in productivity and efficiency, driven by the spread of these technologies, were widely predicted, yet there is no consensus about what this will mean for the UK workforce. Some expect rising unemployment as labour is substituted for AI-enabled robots and machines. Others foresee a transformation in the type of employment available—with the creation of new jobs compensating for those that were lost—and the prospect of robotics and AI augmenting existing roles, and enabling humans to achieve more than they could on their own.

Despite these differing views, there is general agreement that a much greater focus is needed on adjusting our education and training systems to deliver the skills that will enable people to adapt, and thrive, as new technology comes on stream. Government leadership in this area, however, has been lacking. It is disappointing that the Government still has not published its Digital Strategy nor set out its plans for equipping the future workforce with the digital skills it needs to flourish. The Government must commit to addressing the digital skills crisis through a Digital Strategy, published without delay.

Leadership was also found to be lacking across 'Robotics and Autonomous Systems' (RAS) which, together, form one of the Government's 'Eight Great Technologies'. The 'Eight Greats' were identified by the Government in 2013 as technologies in which the UK was set to be a global leader, yet we found that there was no Government strategy for developing the skills, and securing the critical investment, that is needed to create future growth in robotics and AI. Furthermore, there was no sign of the Government delivering on its promise, made in March 2015, to establish a 'RAS Leadership Council' to provide much needed coordination and direction. This should be remedied immediately and a Leadership Council established without further delay. The Leadership Council should

work with the Government and the Research Councils to produce a Government-backed 'National RAS Strategy', setting out the Government's ambitions, and financial support, for this 'great technology'.

1 Introduction

1. In his seminal paper, *Computing Machinery and Intelligence*, Professor Alan Turing began by posing a deceptively simple question: "Can machines think?" The question, in one guise or another, has been a source of inspiration for modern literature, drama and art, as well as being a subject of continued scientific endeavour. Yet Turing quickly dismissed it as too ambiguous, instead reformulating the question and describing his 'Imitation Game'; a test he proposed as a means to establish whether a machine could *act* indistinguishably from a human. In concluding his paper, he hoped that "that machines [would] eventually compete with men in all purely intellectual fields", perhaps beginning with "the playing of chess".[1]

2. In the 66 years since Turing published his landmark paper, the development of what we now term 'artificial intelligence' has gone through periods of optimism and progress, only to be followed by setbacks. While machines still do not compete with humans "in all purely intellectual fields"—as Turing put it—artificially intelligent machines have made extraordinary progress in the area he initially singled out: playing, and winning, at board games.

3. Early this year, for example, Google DeepMind's AlphaGo—an artificially intelligent computer programme—won a five-match series of the ancient Chinese board game 'Go' against the reigning world champion, Lee Sedol.[2] Go was "widely viewed as an unsolved 'grand challenge' for artificial intelligence" and AlphaGo's success marked a watershed moment in its ongoing development.[3] Significant progress, however, has been made across the field in recent years, linked to the rise in processing power, the profusion of data and the development of techniques such as 'deep learning'.[4] Much of that progress—such as improved automated voice recognition software, predictive text keyboards on smart phones and autonomous vehicles—has been driven by UK-based technology start-ups, founded by graduates of UK universities, as well as universities themselves.[5]

4. There is no single, agreed definition of artificial intelligence (AI), though there is a tendency to describe AI by contrasting it with human intelligence and stressing that AI does not appear 'in nature'.[6] At present, the capacity of 'AI machines' is narrow and specific; they can complete what Margaret Boden, Professor of Cognitive Science at the University of Sussex, has described as "specialised tricks".[7] For example, Google DeepMind's AlphaGo system cannot "for the moment do anything besides play Go".[8] Thus, as it currently stands, AI can be loosely thought of as:

1 A M Turing, Computing Machinery and Intelligence, Mind, vol 49 (1950), pp 433–460.
2 AlphaGo seals 4–1 victory over Go grandmaster Lee Sedol, The Guardian, 15 March 2016. Go is an abstract strategy board game for two players, in which the aim is to surround more territory than the opponent. Despite having relatively simple rules, Go is considered more complex than chess, having both a larger board with more scope for play and longer games, and, on average more alternatives to consider per move.
3 Google DeepMind (ROB0062) para 1.4.
4 The dawn of artificial intelligence, The Economist, 9 May 2015
5 10 British AI companies to look out for in 2016, Business Insider UK, 5 January 2016; Artificial Intelligence: made in the UK, Digital Catapult Centre, 20 March 2016
6 Gary Lea, Why we need a legal definition of artificial intelligence, The Conversation, 2 September 2015
7 Professor Margaret Boden, Human-level AI: Is it Looming or Illusory?, lecture at The Centre for the Study of Existential Risk, Cambridge, 19 June 2015
8 Google DeepMind (ROB0062) para 1.5

> a set of statistical tools and algorithms that combine to form, in part, intelligent software that specializes in a single area or task. This type of software is an evolving assemblage of technologies that enable computers to simulate elements of human behaviour such as learning, reasoning and classification.[9]

5. Progress has recently been made in 'machine learning'—a "way of achieving a degree of AI".[10] Machine learning involves building algorithms that can learn specific concepts for themselves, without being explicitly programmed. This, in turn, relies on those algorithms processing vast quantities of 'training data' in order to learn to identify a statistical rule that correlates inputs with the correct outputs. This type of 'narrow' AI is already found in aspects of daily life, from using voice recognition software on a smart phone, to filtering spam out of an email inbox.

6. Machines have also become more adept at translating one language into another, though they do not 'understand' language in the same way as a human. They struggle to cope, for example, with syntax and do not comprehend the meaning or implications of the language they are translating.[11] The 'general' artificial intelligence—akin to human intelligence—that this would require has not yet been developed.[12] There is continuing debate about when such general artificial intelligence might be achieved, as well as whether it is even possible. According to Professor Stephen Hawking and others, while it might be "tempting to dismiss the notion of highly intelligent machines as mere science fiction [...] this would be a mistake, and potentially our worst mistake ever".[13]

7. Robotics—machines that are "capable of carrying out a series of actions on behalf of humans"[14]—is a different topic to AI. Robots can (and, for the most part, do) operate without possessing any artificial intelligence. It is anticipated, however, that this will gradually change over time, with robots becoming the 'hardware' that use, for example, machine learning algorithms, to perform a manual or cognitive task.[15] AI and robotics will, therefore, have an important degree of interdependency. As one commentator explained, "there is no AI without robotics […] intelligence and embodiment are tightly coupled issues".[16] For these reasons, our inquiry has considered robotics and AI together.

Our inquiry

8. Both robotics and artificial intelligence are complex, and potentially transformative, emerging technologies in which the UK is playing a leading role. Yet it is often difficult to predict with any accuracy how technologies will unfold and evolve. The implications of new technologies tend, therefore, to be examined and understood by policymakers too late in the day to engage with them in any significant way.[17] As a result, technology "is sometimes presented to us as if [it] is on a relentless track in a particular direction

9 Transpolitica (ROB0044) para 1.4
10 The Royal Society, 'What is machine learning?', last accessed 31 August 2016
11 Professor Margaret Boden, Human-level AI: Is it Looming or Illusory?, lecture at The Centre for the Study of Existential Risk, Cambridge, 19 June 2015; see also Professor Tony J Prescott (ROB0020) paras 6 & 9
12 See, for example, The Royal Society (ROB0023); Professor Huw Price (ROB0031) para 7
13 Stephen Hawking, Stuart Russell, Max Tegmark, Frank Wilczek, Transcending Complacency on Superintelligent Machines, The Huffington Post, 19 June 2014
14 Innovate UK (ROB0060) para 6
15 See, for example, RACE, UK Atomic Energy Authority (ROB0041)
16 Jean-Christophe Baillie, "Why AlphaGo is not AI", March 2016
17 Jack Stilgoe, Science, ethics and shared space, The Guardian, 1 May 2013

and we have no power to move it either way".[18] We decided to examine robotics and AI after the Government was unable to produce a short statement outlining the evidence underpinning its policy on AI, which we requested as part of our 'evidence check' work.[19]

9. By undertaking our inquiry now, we hope that it will be soon enough to be productive and late enough to be relevant. Indeed, the announcement in the Queen's Speech of the *Modern Transport Bill*—with its aim to "put the UK at the forefront of autonomous and driverless vehicles ownership and use"—was a stark reminder that advances in robotics and AI are starting to make their way into the mainstream.[20] Other countries are also beginning to look at the wider issues raised by AI. During the course of our inquiry, for example, the White House Office of Science and Technology Policy ran a series of workshops on the implications of AI and launched its own review—*Preparing for the Future of Artificial Intelligence*.[21]

10. Our inquiry took a broad focus and examined robotics and AI in the round: identifying their potential value and capabilities, as well as examining prospective problems, and adverse consequences, that may require prevention, mitigation and governance. We launched our inquiry in March 2016 and sought written submissions addressing the following points:

a) The implications of robotics and artificial intelligence on the future UK workforce and job market, and the Government's preparation for the shift in the UK skills base and training that this may require.

b) The extent to which social and economic opportunities provided by emerging autonomous systems and artificial intelligence technologies are being exploited to deliver benefits to the UK.

c) The extent to which the funding, research and innovation landscape facilitates the UK maintaining a position at the forefront of these technologies, and what measures the Government should take to assist further in these areas.

d) The social, legal and ethical issues raised by developments in robotics and artificial intelligence technologies, and how they should be addressed.

11. We received 67 written submissions and took oral evidence from 12 witnesses including:

- academics working in the field of robotics and AI;
- representatives from the robotics and computing industries;

18 Q64 [Richard Moyes]
19 For further information on Evidence Check, see Science and Technology Committee, Sixth Report of Session 2016–17, Evidence Check: Smart Metering of Electricity and Gas, HC 161. See also Department for Business, Energy and Industrial Strategy (ROB0076)
20 The Queen's Speech 2016, Background Notes, last accessed 3 August 2016 at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/524040/Queen_s_Speech_2016_background_notes_.pdf
21 White House Office of Science and Technology Policy, Preparing for the Future of Artificial Intelligence, May 2016

- a non-governmental organisation concerned about the development of lethal autonomous weapons;
- representatives of Research Councils UK and Innovate UK.

We also visited Google DeepMind in King's Cross, London (see Annex). We would like to thank everyone who contributed to this inquiry. In Chapter 2 we look at the economic and social implications of robotics and AI, particularly in the context of the future of work, employment and skills. Chapter 3 focuses on the ethical and legal issues that may be raised, and what governance frameworks might be required, while Chapter 4 examines the research, funding and innovation landscape for robotics and AI.

2 Economic and social implications

12. Concerns about machines 'taking jobs' and eliminating the need for human labour have persisted for centuries.[22] It is perhaps unsurprising, then, that the potential economic and social implications of robotics artificial intelligence have been the subject of debate.[23] This chapter outlines differing views on how robotics and AI may impact upon productivity, and shape employment structures, before turning to consider how the UK might respond.

Machines versus humans?

13. The potential for robotics and AI to increase the UK's productivity, particularly in manufacturing, was repeatedly cited throughout our inquiry. The Manufacturing Technology Centre was one of many who told us that an increased uptake of robotic and artificial intelligence technologies in manufacturing would lead to "increased productivity and a stronger economy providing wealth and security to a society at large".[24]

14. Several reports have reached the same conclusion. A study undertaken by the Copenhagen Business School in 2011 modelled how much productivity in manufacturing would increase if all industries in a country had the highest found level of robot-intensity. It estimated that productivity would rise in the UK by 22%.[25] A more recent report published by Barclays Bank in 2015, based on a survey of manufacturers and its own economic modelling, estimated that "£1.24bn in automation investment could raise the overall value added by the manufacturing sector to the UK economy by £60.5bn over the next decade". [26]

15. Improvements in productivity, driven by robotics and AI, will have implications for the UK workforce. We received conflicting views, however, about precise nature of those impacts. Some predicted rising unemployment, while others foresaw a transformation in the *types* of employment available, made possible by the increasing pervasiveness of robotics and AI throughout the world of work.

Employment changes

16. Google DeepMind told us that "we should expect that new areas of economic activity and employment will be made possible" by the increased use of AI, but that certain types of work and skills will decrease in relevance.[27] Some argued that these 'new areas of economic activity and employment' would affect the structure of the workforce but would

22 In the 19th Century, for example, the Luddites—textile workers in Nottinghamshire, Yorkshire and Lancashire—led a workers' uprising throughout parts of England to protest against the introduction of new technologies, like automated looms, which were being used in the textile industry in place of their skilled labour.

23 Nesta (ROB0034)

24 Manufacturing Technology Centre (ROB0018). See also, for example, Department for Business, Innovation and Skills (BIS) (ROB0066) para 26; Professor David Lane (ROB0064) para 2.16; techUK (ROB0063) paras 2 & 15; Innovate UK (ROB0060) para 25

25 Lene Kromann, Jan Rose Skaksen, Anders Sorensen, Automation, labor productivity and employment - a cross country comparison, CEBR, Copenhagen Business School, Working Paper, 2011

26 Barclays, Future-proofing UK manufacturing. Current investment trends and future opportunities in robotic automation, November 2015, p 4

27 Google DeepMind (ROB0062) para 2.2

not diminish the overall employment rate in the UK—the creation of new jobs would more than compensate for those directly lost to robots and AI systems.[28] As the Global Priorities Project explained:

> During the industrial revolution, mechanisation did not change long-run equilibrium employment because new jobs emerged which were unimaginable at that time. Similarly, jobs lost to automation today might be replaced by jobs we cannot yet imagine.[29]

17. Deloitte was similarly optimistic about the impact of the growth of robotics and AI on the workforce, noting that:

> New jobs and, indeed, new industries have been created in the UK as technology has advanced and, looking back over the last century and a half, UK employment has more than doubled during a period of profound technological change.[30]

18. Others were not so hopeful about the future and questioned whether this 'fourth industrial revolution' would follow the same pattern as those that had gone before. Innovate UK highlighted that while "previous technologies have always resulted in a net gain in employment, there is debate about whether this generation of technologies will create the same outcome".[31] Research by Carl Benedikt Frey and Michael Osborne at the University of Oxford estimated that 35% of jobs in the UK were at high risk of automation in the next 10–20 years.[32] Conducting similar research to Frey and Osborne's, the Bank of England suggest that that up to 15 million jobs in the UK could be at risk of automation over the same time period.[33] Angus Knowles-Cutler from Deloitte clarified that Frey and Osborne's calculations were made "purely from a technology point of view", and that they did not factor in:

> social and political resistance to that change, the ease or difficulty of implementation or the cost-benefit of human labour versus investment in the technology in the first place.[34]

19. There is also debate about whether specific sectors are more susceptible than others to job destruction, or creation, by the advancement of robotics and AI. Future Advocacy pointed to the advent of autonomous cars which, it stated, could "bring redundancy to an entire industry of professional drivers".[35] Creative occupations—including musicians,

28 See, for example, Manufacturing Technology Centre (ROB0018); Robotics & Autonomous Systems Special Interest Group (ROB0027) para 15
29 Global Priorities Project (ROB0051) para 3. The Global Priorities Project is a collaboration between the Centre for Effective Altruism and the Future of Humanity Institute at the University of Oxford.
30 Deloitte (ROB0019)
31 Innovate UK (ROB0060) para 16
32 Carl Benedikt Frey, Michael A Osborne, Agiletown: the relentless march of technology and London's response, Deloitte, 2014
33 'Labour's Share' - a speech given by Andrew Haldane, Chief Economist, Bank of England, to the Trades Union Congress, London, 12 November 2015
34 Q86 [Angus Knowles-Cutler]
35 Future Advocacy (ROB0047) para 3.1

architects, and artists—were found by Frey and Osborne to be much more resistant to automation.[36] Nesta commented that the protection afforded by a job requiring creativity was unsurprising:

> When one considers that computers will most successfully be able to emulate human labour when a problem is well specified—that is, when performance can be straightforwardly quantified and therefore evaluated—and when the task environment is sufficiently simple to enable autonomous control. By contrast, they will struggle when tasks are highly interpretive (tacit), geared at products whose final form is not fully specified in advance.[37]

20. Creativity, however, may not provide long-term protection against automation. According to Dr Osborne, it was becoming "much less clear what remains as the preserve of human labour alone", particularly as we start to see the development of "algorithms that can substitute for human cognitive work", such as that done by "paralegals and junior lawyers, accountants and auditors".[38]

21. Andrew Haldane, Chief Economist at the Bank of England, reported in 2015 that those "most at risk from automation tend, on average, to have the lowest wage".[39] Angus Knowles-Cutler told us that, based on Deloitte's modelling, the "jobs that today in the UK pay £30,000 or less are five times more vulnerable to being automated than jobs that pay £100,000 or more".[40] Deloitte described this as a "potential 'hollowing out' of the labour market, in which technology impacts primarily on middle-income jobs":

> The sector with the highest number of jobs with a high risk of automation was wholesale and retail, with 2,168,000 jobs, 59% of the total current workforce, with a high chance of being automated in the next two decades. This was followed by transport & storage—1,524,000 jobs, 74% of the workforce—and human health & social work—1,351,000 jobs, 28% of the workforce.[41]

22. These prospective changes may affect income inequalities. While acknowledging that robotics and AI could generate "a host of new occupations", Dr Osborne voiced concerns that these occupations "might not be sufficiently well paid to substitute for those that are automated away […] which might lead to exacerbation of inequality".[42] Klaus Schwab, the founder of the World Economic Forum, has described such potential inequality as representing "the greatest societal concern associated with the fourth industrial revolution".[43]

23. The possible speed of such changes to the UK workforce—and whether they would be incremental or rapid—is not clear. Professor Tony Prescott from the University of Sheffield thought that "impacts can be expected to occur over several decades, allowing

36 Carl Benedikt Frey, Michael A Osborne, Agiletown: the relentless march of technology and London's response, Deloitte, 2014, p 6
37 Nesta (ROB0034)
38 Q88–89
39 'Labour's Share' - a speech given by Andrew Haldane, Chief Economist, Bank of England, to the Trades Union Congress, London, 12 November 2015
40 Q90
41 Deloitte (ROB0019)
42 Q97
43 Klaus Schwab, The Fourth Industrial Revolution: what it means, how to respond, World Economic Forum, (January 2016)

time to adapt".[44] Research by McKinsey, however, noted that AI was contributing to a transformation of society "happening ten times faster and at 300 times the scale, or roughly 3,000 times the impact" of the Industrial Revolution.[45]

Job augmentation

24. Instead of focusing on job creation or destruction, some witnesses considered the potential for robotics and AI to support, or augment, existing roles. Dave Coplin, Chief Envisioning Officer at Microsoft, thought that by framing the "conversation [as] all about humans versus machines" the discussion began "on the wrong foot". Technology, he stressed, had:

> never been about humans versus machines. The story, certainly from our perspective, in the personal computer revolution is about how we augment humanity and how we enable human beings to rise up and achieve more than they could on their own.[46]

25. Angus Knowles-Cutler from Deloitte similarly recognised that while "we often think of this as […] a human versus a machine or a robot" it was, in fact, far "subtler" and about providing support for "tasks within jobs". Also pointing to the advent of the personal computer, he noted that although a computer had helped to eliminate the repetitive tasks from his day:

> it did not destroy my industry [consultancy], and in fact my industry is much larger than it was back in 1985, so there is a subtlety there that is very important. These are tasks where technology is enabling us to be more effective as productive workers.[47]

26. In a similar vein, Professor Nick Jennings, representing the Royal Society's Machine Learning Working Group, described robotics and AI as "an augmenter of many of the professional white collar activities".[48] Rather than replacing humans, he emphasised that the "key future for AI" lay in its potential to "work in partnership" with people.[49] Drawing on the example of applying AI to the medical sciences, Professor Stephen Muggleton highlighted how, with "large amounts of data from genome projects and testing, […] machines [were] able to go through millions of hypotheses and select the best out of a large space and then present it to scientists". This approach, he explained, did not replace scientists but it could amplify "what they can do, much in the same way as a telescope amplifies what astronomers could do".[50]

Education and skills

27. Though we heard a wide range of views on how the nature of work may change, our witnesses generally agreed that learning new skills, and adapting our education system,

44 Professor Tony J Prescott (ROB0020) para 1
45 Richard Dobbs, James Manyika, Jonathan Woetzel, 'The four global forces breaking all the trends', McKinsey Global Institute (April 2015)
46 Q118 [Dave Coplin]
47 Q91
48 Q47
49 Q9 [Professor Jennings]
50 Q9 [Professor Muggleton]

would help to ensure that the UK realised the full range of opportunities presented by robotics and AI, while also managing its potential risks. Deloitte, for example, argued that the UK's ongoing success would depend on the ability of businesses, educators and government to anticipate future skills requirements and provide the right training and education for the coming decades.[51] TechUK stated that the key skills needed in robotics and AI were "in areas such as software development, systems design, engineering, programming and data science".[52] It added, however, that all these "have been reported areas of domestic shortage right across tech firms in the UK".[53]

28. Addressing the UK's digital skills 'crisis' (discussed in detail in our reports on the *Digital Skills Crisis* and the *Big Data Dilemma*) was repeatedly identified in written submissions as essential in order to mitigate some of the more potentially negative impacts of robotics and AI on employment.[54] As a Committee, we have been clear that digital exclusion, and systemic problems with digital education and training, need to be addressed as a matter of urgency in the Government's Digital Strategy; a document that was due to be published in January 2016 but which has been subject to a series of delays.[55]

29. According to Google DeepMind, one of the "most important steps we must take is [ensuring] that current and future workforces are sufficiently skilled and well-versed in digital skills and technologies, particularly STEM subjects".[56] Achieving this goal may require the current workforce to be re-skilled, or up-skilled. As the EPSRC's Robotics and Autonomous Systems Network explained:

> the Government needs to tangibly support the workforce in adjusting their skills and business in creating opportunities based on new technologies. Training in digital skills and re-educating the existing workforce are essential to maintain the competitiveness of the UK.[57]

30. Professor Rose Luckin from the UCL Institute of Education made a similar point, noting that "whether your entire job has been replaced" or "certain parts of your job are automated", "different skills" will subsequently be required and workers will need to be retrained.[58] She questioned, however, what progress the UK had made in this area:

> I do not feel that at the moment we are equipping either students in school or workers in the workforce with the requisite skills to know how to adapt themselves to use the automation they are being offered to best effect. We need to take that on board and make some changes to address it.[59]

31. Professor Luckin stressed that the structure, focus and delivery of the UK's education system needed to evolve, in order to prepare students for a future where robotics and AI were commonplace. She explained that, currently, the "very things on which we focus our education system are the routine cognitive skills that are the easiest to automate".[60] Future

51 Deloitte (ROB0019)
52 techUK (ROB0063) para 24
53 techUK (ROB0063) para 24
54 See, for example, Google DeepMind (ROB0062) paras 2.2–2.6; Innovate UK (ROB0060) para 17; Geoff Pegman (ROB0059) para 2.5; Autonomous Intelligent Systems Partnership (ROB0049)
55 Science and Technology Committee, Second Report of Session 2016–17, Digital skills crisis, HC 270, paras 14 &15
56 Google DeepMind (ROB0062) para 2.2
57 EPSRC UK-RAS Network (ROB0032) para 5
58 Q96 [Professor Luckin]
59 Q96 [Professor Luckin]; see also Manufacturing Technology Centre (ROB0018)
60 Q99

Advocacy suggested that the education system should be adapted to "focus on things that machines will be less good at for longer" such as "creativity, ideation, judgement, inter-personal skills".[61] Dave Coplin from Microsoft also emphasised the importance of creativity. While recognising that STEM skills were "really important" he noted that:

> Without art and creativity, innovation is dead. We could have a bunch of scientists, which would be brilliant, but their ability to be creative in the future world of work is the thing that makes them successful.[62]

32. By talking about STEM "in exclusivity of the other skills that will be required", Mr Coplin suggested that "our ability to be successful" was being curtailed. Instead he advocated showing young people, via inspirational role models, how individuals can be creative through:

> a combination of their human skills—empathy and creativity—and their ability to manipulate the technology to deliver a great outcome […] We do not need to frighten them off with a bunch of science; we need to show them how creative they can be and how it is a blended world.[63]

33. Professor Luckin highlighted that bringing AI techniques into education also held "unique potential to mitigate [changes to the jobs market] by providing lifelong skills development to the workforce".[64] She gave the example of 'Intelligent Tutoring Systems' that use AI techniques "to simulate one-to-one human tutoring, delivering learning activities best matched to a learner's cognitive needs and providing targeted and timely feedback", without a teacher having to be present.[65] Yet, according to Professor Luckin, there was "little awareness […] in government of the existence of AIEd [AI in education] or of the implications of AIEd for teaching and training the current and future UK workforce".[66]

34. The Government's less than wholehearted engagement does not appear to be limited to AI in education. In its initial written evidence, the Government simply commented that it recognised "the broader impact of RAS [Robotics and Autonomous Systems] on the UK economy, including employment" and that the:

> discussion of these issues involves experts in law and computer science, the National Academies, the Alan Turing Institute, the Information Commissioner's Office and other relevant bodies.[67]

There was no mention, however, of the Government in these discussions. When we asked the Government to clarify its role in addressing the implications of AI on society, it stated that it did have "a role to play in managing and mitigating any risks that might arise", adding that it would:

61 Future Advocacy (ROB0047) para 3.2
62 Q119
63 Q119
64 Professor Rose Luckin (ROB0043)
65 Professor Rose Luckin (ROB0043)
66 Professor Rose Luckin (ROB0043)
67 Department for Business, Innovation and Skills (BIS) (ROB0066) para 24

> continue to work with the Royal Academies, the Government Office for Science, and others such as the Aerospace Technology Institute and Alan Turing Institute to inform decisions about ethical issues and appropriate governance issues for AI.[68]

35. Though it was not referred to in the Government's written evidence, the UK Commission for Employment and Skills (UKCES)—an executive, non-departmental body sponsored by the Department of Business, Energy and Industrial Strategy—reported on *The Future of Work* in 2014.[69] The publication considered several scenarios for a more automated future, as well as the steps that policy makers, employers and individuals could take to prepare for tomorrow's world of work. UKCES reports, however, do not receive a formal Government response, and the Government provided no indication that it had engaged with its findings. Instead, the Government announced on 21 July 2016 that "all operational activities of UKCES will be concluded by the end of 2016 and it is expected the organisation will be wound up in line with the end of its financial year, 2016–17".[70]

36. **Advances in robotics and AI hold the potential to reshape fundamentally the way we live and work. While we cannot yet foresee exactly how this 'fourth industrial revolution' will play out, we know that gains in productivity and efficiency, new services and jobs, and improved support in existing roles are all on the horizon, alongside the potential loss of well-established occupations. Such transitions will be challenging. As a nation, we must respond with a readiness to re-skill, and up-skill, on a continuing basis. This requires a commitment by the Government to ensure that our education and training systems are flexible, so that they can adapt as the demands on the workforce change, and are geared up for lifelong learning. Leadership in this area, however, has been lacking. It is disappointing that the Government has still not published its Digital Strategy and set out its plans for equipping the future workforce with the digital skills it needs to thrive.**

37. ***Digital exclusion has no place in 21st century Britain. As we recommended in our Big Data Dilemma, Digital Skills Crisis, and Satellites and Space reports, the Government must commit to addressing the digital skills crisis through a Digital Strategy, published without delay.***

68 Department for Business, Energy and Industrial Strategy (ROB0076)
69 UK Commission for Employment and Skills, The future of work: jobs and skills in 2030, February 2014
70 HCWS121 [Update on the UK Commission for Employment and Skills: Written statement] 21 July 2016

3 Ethical and legal issues

38. The development, programming and use of robotics and AI raises a host of ethical and legal issues. Our witnesses were clear that these need to be identified and addressed now, so that the societal benefits of the technologies can be maximised while also mitigating the potential risks. Both steps are essential to building public trust, particularly as robotics and AI diffuse into more aspects of everyday life. In this chapter we consider safety and control, and how society can make sure that the outcomes of robotics and AI are beneficial, intentional and transparent. We then examine what roles standards, regulation and public dialogue might play.

Safety and Control

Verification and validation

39. It is important to ensure that AI technology is operating as intended and that unwanted, or unpredictable, behaviours are not produced, either by accident or maliciously. Methods are therefore required to verify that the system is functioning correctly. According to the Association for the Advancement of Artificial Intelligence:

> it is critical that one should be able to prove, test, measure and validate the reliability, performance, safety and ethical compliance—both logically and statistically/probabilistically—of such robotics and artificial intelligence systems before they are deployed.[71]

Similarly, Professor Stephen Muggleton saw a pressing need:

> to ensure that we can develop a methodology by which testing can be done and the systems can be retrained, if they are machine learning systems, by identifying precisely where the element of failure was.[72]

40. The EPSRC UK-RAS Network noted that the verification and validation of autonomous systems was "extremely challenging" since they were increasingly designed to learn, adapt and self-improve during their deployment.[73] Innovate UK highlighted that "no clear paths exist for the verification and validation of autonomous systems whose behaviour changes with time"[74] while Professor David Lane from Heriot-Watt University emphasised that "traditional methods of software verification cannot extend to these situations".[75]

41. Part of the problem, according to Dr Michael Osborne, was that researchers' efforts had previously been focused on "achieving slightly better performance on well-defined problems, such as the classification of images or the translation of text" while the "interpretation of the algorithms that [were] produced to achieve those goals [had] been

71 AAAI and UKCRC (ROB0021)
72 Q15
73 EPSRC UK-RAS Network (ROB0032) para 4.4
74 Innovate UK (ROB0060) para 20
75 Professor David Lane (ROB0064) para 4.3

left as a secondary goal".[76] As a result, Dr Osborne considered that "we are not where we would want to be in ensuring that the algorithms we deliver are completely verifiable and validated". He added, however, that progress was now being made.[77]

42. Google DeepMind, for example, was reported in June 2016 to be working with academics at the University of Oxford to develop a 'kill switch'; code that would ensure an AI system could "be repeatedly and safely interrupted by human overseers without [the system] learning how to avoid or manipulate these interventions".[78] In the same month, researchers from Google, Open AI, Stanford University and UC Berkeley in the United States, together published a paper which examined potential AI safety challenges and considered how to engineer AI systems so that they operated safely and reliably.[79]

Decision-making transparency

43. It is currently rare for AI systems to be set up to provide a reason for reaching a particular decision.[80] For example, when Google DeepMind's AlphaGo played Lee Sedol in March 2016 (see paragraph 3), the machine was able to beat its human opponent in one match by playing a highly unusual move that prompted match commentators to assume that AlphaGo had malfunctioned.[81] AlphaGo cannot express why it made this move and, at present, humans cannot fully understand or unpick its rationale. As Dr Owen Cotton-Barratt from the Future of Humanity Institute reflected, we do not "really know how the machine was better than the best human Go player".[82]

44. When the stakes are low—such as in a board game like Go—this lack of transparency does not matter. Yet, as Tony Prescott, Professor of Cognitive Neuroscience at the University of Sheffield noted, "machine learning and probabilistic reasoning will lead to algorithms that replace human decision-makers in many areas", from financial decision-making to the development of more effective medical diagnostics.[83] Nesta suggested that in these types of applications, where the stakes are far higher, an absence of transparency can lead to a "level of [public] mistrust in its outputs" since the reasoning behind the decision is opaque.[84] Patients, for example, may be unwilling to simply accept the "supposed quality of [an] algorithm" where their treatment is concerned and may instead want a clear justification from a human.[85]

45. Dr Cotton-Barratt was one of a number of witnesses who supported "a push towards developing meaningful transparency of the decision-making processes".[86] Dave Coplin from Microsoft, for example, stated that:

> The building blocks […] the way in which we create the algorithms […] They must be transparent. I must be able to see the pattern or rules that have

76 Q110
77 Q110
78 Google developing kill switch for AI, BBC News Online, 8 June 2016
79 Dario Amodei et al, Concrete Problems in AI Safety, June 2016
80 AAAI and UKCRC (ROB0021)
81 Global Priorities Project (ROB0051) para 18
82 Q63; see also Q2
83 Professor Tony J. Prescott (ROB0020) para 3; see also Nutmeg Saving and Investment Ltd (ROB0035)
84 Harry Armstrong, Machines That Learn in the Wild, Nesta, July 2015, p14
85 Harry Armstrong, Machines That Learn in the Wild, Nesta, July 2015, p14
86 Q66

> been used to create the outcome. As a human I need to be able to inspect that, as much as the algorithms need to understand what the humans may choose to do with that information.[87]

Similarly, Professor Alan Winfield from the Bristol Robotics Laboratory emphasised the importance of being able to 'inspect' algorithms so that, if an AI system made a decision that "[turned] out to be disastrously wrong […] the logic by which the decision was made" could be investigated.[88]

46. As we noted in our *Big Data Dilemma* report, the European Union's new General Data Protection Regulation is due to come into effect across the EU in 2018.[89] It will create a "right to explanation," whereby a user can ask for an explanation of an automated algorithmic decision that was made about them. Whether, and how, this will be transposed into UK law is unclear following the EU Referendum.

Minimising bias

47. Instances of bias and discrimination being accidentally built into AI systems have recently come to light. Last year, for example, Google's photo app, which automatically applies labels to pictures in digital photo albums, was reported to have classified images of black people as gorillas.[90] The app learnt from training data and, according to Kate Crawford from Microsoft, the AI system built "a model of the world based on those [training] images".[91] Yet, as Drs Koene and Hatada from the University of Nottingham explained "all data-driven systems are susceptible to bias based on factors such as the choice of training data sets, which are likely to reflect subconscious cultural biases".[92] So, if a system was "trained on photos of people who are overwhelmingly white, it will have a harder time recognizing nonwhite faces".[93]

48. It is not clear how much attention the design of AI systems—and the potential for bias and discrimination to be introduced—is receiving. John Naughton, Emeritus Professor of the Public Understanding of Technology at the Open University, was reported as saying that these types of biases can go unrecognised because developers take "a technocratic attitude that assumes data-driven decision-making is good and algorithms are neutral".[94]

49. Dave Coplin from Microsoft, however, acknowledged that "in AI every time an algorithm is written, embedded within it will be all the biases that exist in the humans who created it".[95] He emphasised a need "to be mindful of the philosophies, morals and ethics of the organisations […] creating the algorithms that increasingly we rely on every day" but added that our understanding of "how we as humans imbue human bias in artificial

87 Q156
88 Professor Alan Winfield (ROB0070) para 10
89 Science and Technology Committee, Fourth Report of Session 2015–16, The big data dilemma, HC 468, paras 83–102
90 Frankenstein's paperclips; Ethics, The Economist, 25 June 2016 (US Edition)
91 Artificial Intelligence's White Guy Problem, The New York Times, 25 June 2016
92 Dr Ansgar Koene and Dr Yohko Hatada (ROB0057)
93 Artificial Intelligence's White Guy Problem, The New York Times, 25 June 2016
94 Forget killer robots: This is the future of supersmart machines, New Scientist, 22 June 2016
95 Q124

intelligence" was still "relatively new".[96] Safeguards against discriminatory, data-driven 'profiling' are included in the EU's forthcoming General Data Protection Regulation, as discussed in our *Big Data Dilemma* report.[97]

Privacy and consent

50. During the course of our inquiry, there were reports in the media about Google DeepMind working with NHS hospitals to improve patient diagnoses and care.[98] Media commentary focused not just on the work that was underway—such as building an app that helps clinicians detect cases of acute kidney injury, or using machine learning techniques to identify common eye diseases—but also on DeepMind's access to patient data: namely how much data the company could access, whether patient consent had been obtained, and the ownership of that data. Such concerns are not new. As we highlighted in our *Big Data Dilemma* report, the anonymisation and re-use of data is an issue that urgently needs to be addressed.[99] In the same report, we also drew attention to potential improvements in NHS efficiency, planning and healthcare quality that could be realised through greater use of data analytics.[100] Data, as Professor Jennings explained during our current inquiry, is the "fuel for all the algorithms to do their stuff and make smart decisions and learn". Yet he stressed that there remained "a whole load of issues associated with appropriate management of data to make sure that it is ethically sourced and used under appropriate consent regimes".[101]

51. Dr Cotton-Barratt identified the "large benefits", as well as the challenges, that arise when AI is applied in healthcare:

> If it can automate the processes and increase consistency in judgments and reduce the workload for doctors, it could improve health outcomes. To the extent that there are challenges, essentially it means there is less privacy from the same amount of shared data, in that people can get more information out of a limited amount of data.[102]

He added that ways to handle those privacy challenges needed to be found, and suggested that responses should include "making sure that the data is held in the right places and is properly handled and controlled".[103] Similar points were raised by Dave Coplin from Microsoft who told us that if AI was "going to work successfully for us as a society, we need some intelligent privacy and we need to figure out how to do that".[104] One approach—which we recommended in our *Big Data Dilemma* report—is to establish a 'Council of Data Ethics' to address the difficulties associated with balancing privacy, anonymisation,

96 Q156
97 Science and Technology Committee, Fourth Report of Session 2015–16, The big data dilemma, HC 468, para 95
98 "Revealed: Google AI has access to huge haul of NHS patient data", New Scientist, 29 April 2016; "Google handed patients' files without permission", Daily Mail, 3 May 2016; "Google given access to London patient records for research", BBC News Online, 3 May 2016; "Google's DeepMind to analyse one million NHS eye records to detect signs of blindness", Daily Telegraph, 5 July 2016
99 Science and Technology Committee, Fourth Report of Session 2015–16, The big data dilemma, HC 468, para 101
100 Science and Technology Committee, Fourth Report of Session 2015–16, The big data dilemma, HC 468, para 43
101 Q7
102 Q78
103 Q78
104 Q124

security and public benefit. We were pleased that the Government agreed with this step and is in the process of setting up the Council within the Alan Turing Institute, the UK's national institute for data science.[105]

Accountability and liability

52. For some aspects of robotics and AI, questions of accountability and liability are particularly pertinent. To date, these have predominately been discussed in the context of autonomous vehicles ('driverless cars') and autonomous weapons systems. The key question is 'if something goes wrong, who is responsible?'[106] Dave Coplin from Microsoft emphasised that "we need a level of accountability for the algorithms. The people making the algorithm and the AI need to be held accountable for the outcome".[107] He suggested that a "safety net" provided by Government was required "so that people can be held to account in how we build" AI systems.[108]

53. The debate on driverless cars has also focused on liability. The Law Society highlighted that situations may arise in which a driverless car takes action that causes one form of harm in order to avoid other harm. This raises:

> issues of civil, and potentially even criminal liability [as well as] the ownership of that liability, whether the manufacturer of the vehicle, the software developers, the owner of the vehicle and so on. The questions multiply.[109]

54. Whether such questions can be decided in the courts, and solutions developed through case law, or if new legislation will be needed, remains under discussion. The Law Society noted that "one of the disadvantages of leaving it to the courts […] is that the common law only develops by applying legal principles after the event when something untoward has already happened. This can be very expensive and stressful for all those affected".[110]

55. After we had concluded our evidence taking, the Government set out its proposal for addressing liability for automated vehicles. It stated that:

> Our proposal is to extend compulsory motor insurance to cover product liability to give motorists cover when they have handed full control over to the vehicle (ie they are out-of-the-loop). And, that motorists (or their insurers) rely on courts to apply the existing rules of product liability—under the Consumer Protection Act, and negligence—under the common law, to determine who should be responsible.[111]

105 Science and Technology committee, Fifth special report of session 2015–16, The big data dilemma: Government Response to the Committee's Fourth Report of Session 2015–16, HC 992, para 57
106 Q156
107 Q156
108 Q150
109 The Law Society (ROB0037) para 10
110 The Law Society (ROB0037) para 11
111 Centre for Connected & Autonomous Vehicles, Pathway to Driverless Cars: Proposals to support advanced driver assistance systems and automated vehicle technologies, July 2016, para 1.3

Consultation on this and other proposals for automated vehicles runs until 9 September 2016.[112]

56. Accountability is also critically important for autonomous weapons and, more specifically, 'lethal autonomous weapons systems' (LAWS). These are systems that, when given a set objective, can assess the situational context and environment, then make decisions on what intervention is required, independent of human control or intervention. According to Future Advocacy, LAWS could "have the power to kill without any human intervention in the identification and prosecution of a target".[113] Google DeepMind also highlighted the "possible future role of AI in lethal autonomous weapons systems, and the implications for global stability and conflict reduction".[114] While there are "still no completely autonomous weapons systems", Innovate UK thought that "the trend towards more and more autonomy in military systems [was] clearly visible".[115]

57. Richard Moyes from Article 36, an NGO working to prevent the "unintended, unnecessary or unacceptable harm caused by certain weapons"[116], explained that his organisation's concern was the "identification and application of force to the target being in the hands of the weapons system" rather than a human.[117] In his view, if a weapon was deployed, there should always be a human 'in the loop'. He added that "a human should be specifying the target against which force is to be applied".[118] According to Mr Moyes, military personnel may "not feel comfortable being held accountable for a system when they cannot quite understand its functioning and cannot be completely sure what it is going to do".[119] He believed that there was an opportunity for the UK in a "diplomatic landscape to have an influential position on how we orientate the role of computers in life and death decisions".[120]

58. Giving evidence to the Defence Committee in 2014, the Ministry of Defence stated that the UK complied fully with all of its obligations under international humanitarian law irrespective of the weapons systems used.[121] More recently, at the UN Convention on Conventional Weapons meeting in November 2015, the Government stated that:

> Given the uncertainties in the current debate, the United Kingdom is not convinced of the value of creating additional guidelines or legislation. Instead, the United Kingdom continues to believe that international humanitarian law remains the appropriate legal basis and framework for the assessment of the use of all weapons systems in armed conflict.[122]

112 Centre for Connected & Autonomous Vehicles, Pathway to Driverless Cars: Proposals to support advanced driver assistance systems and automated vehicle technologies, July 2016
113 Future Advocacy (ROB0047) para 4.1
114 Google DeepMind (ROB0062) para 5.3
115 Innovate UK (ROB0060) para 38
116 Article 36 (ROB0029)
117 Q69
118 Q70
119 Q64 [Richard Moyes]
120 Q82 [Richard Moyes]
121 Defence Committee, Tenth Report of Session 2013–14, Remote Control Remotely Piloted Air Systems – current and future UK use, HC 772, para 144
122 United Kingdom of Great Britain and Northern Ireland, Statement on Lethal Autonomous Weapons Systems to the CCW Meeting of the High Contracting Parties, 12–13th November 2015

Elsewhere, the Government has asserted that "the operation of weapons systems by the UK armed forces will always be under human control".[123] Article 36 reflected that "whilst such assertions seem on the surface to be reassuring" there needed to be further explanation from officials about the "form and extent of that human 'control' or 'involvement'".[124]

Governance: standards and regulations

59. Though some of the more transformational impacts of AI might still be decades away, others—like driverless cars and supercomputers that assist with cancer prediction and prognosis—have already arrived.[125] The ethical and legal issues discussed in this chapter, however, are cross-cutting and will arise in other areas as AI is applied in more and more fields. For these reasons, witnesses were clear that the ethical and legal matters raised by AI deserved attention now and that suitable governance frameworks were needed.[126]

60. TechUK believed that such frameworks were "vital" to ensure "that we have a way to ask, discuss and consider the key legal and ethical questions" such as "'What are the ethics that should underpin our use of artificial intelligence?"[127] Innovate UK expressed a similar view, stating that:

> Appropriate legal and regulatory frameworks will have to be developed to support the more widespread deployment of robots and, in particular, autonomous systems. Frameworks need to be created to establish where responsibilities lie, to ensure the safe and effective functioning of autonomous systems, and how to handle disputes in areas where no legal precedence has been set.[128]

61. Innovate UK added that there was a "genuine request from researchers and industries for a legal and ethical governance to which they can fine-tune their strategies and plans about innovative robotic applications".[129] Mike Wilson from ABB Robotics highlighted that while "the pace of development continues [in robotics] the standards and the legal frameworks around them are not keeping up with the development of the technology".[130] He stressed that this was something "that certainly needs to be addressed to ensure that people have a clear picture of where the standards are going to be".[131]

62. Having a secure regulatory environment may also help to build public trust. Drawing on the example of commercial aircraft, Professor Alan Winfield from the Bristol Robotics Laboratory thought that one of the reasons why people trust airlines was because "we know they are part of a highly regulated industry with an excellent safety record". Furthermore, when things go wrong, there are "robust processes of air accident investigation".[132]

123 UK opposes international ban on developing 'killer robots', The Guardian, 13 April 2015
124 Article 36 (ROB0029)
125 IBM, 'IBM Watson for Oncology', last accessed 31 August 2016
126 Q56 [Dr Cotton-Barratt]; EPSRC UK-RAS Network (ROB0032) para 4.4
127 techUK (ROB0063) para 42
128 Innovate UK (ROB0060) para 37
129 Innovate UK (ROB0060)
130 Q122
131 Q122
132 Professor Alan Winfield (ROB0070) para 8

Professor Nelson thought that "as technology in this area develops, a need will probably arise" for something similar to the Civil Aviation Authority to "ensure that [AI systems] are properly regulated and to build trust in the community".[133]

63. Others emphasised that a balance needed to be struck on the grounds that efforts to introduce a governance regime could curtail innovation and hold back desirable progress. Speaking in the context of developing driverless cars, Dr Buckingham told us that:

> One thing we must not do is put too much red tape around this at the wrong time and stop things developing. One of the key points is to make sure that we are doing that testing in the UK transparently and bringing the industry here so that we understand what is going on, and that we start to apply the regulation appropriately when we have more information about what the issues are. One of the risks is that, if we over-regulate, it is bad for making use of the technology.[134]

TechUK also warned that:

> over-regulation or legislation of robotics and artificial intelligence at this stage of its development, risks stalling or even stifling innovation. This could in turn risk the UK's leadership in the development of these technologies.[135]

64. Nesta noted that there were moves "in both the public and private sectors to set up ethical frameworks for best practice".[136] Such initiatives are being developed at the company level (e.g. Google DeepMind's ethics board)[137]; at an industry-wide level (e.g. the Institute of Electrical and Electronics Engineers global initiative on 'Ethical Considerations in the Design of Autonomous Systems')[138] and at the European level (e.g. the European Parliament's Committee on Legal Affairs' examination of the legal and ethical aspects of robotics and AI).[139] It is not clear, however, if any cross-fertilisation of ideas, or learning, is taking place across these layers of governance or between the public and private sectors. As the Chief Executive of Nesta has argued, "it's currently no-one's job to work out what needs to be done".[140]

65. Establishing good robotics and AI governance practices matters, both for the economy and for society as a whole. According to Dr Cotton-Barratt, the UK is well-positioned to respond to this challenge. He described a "small but growing research community looking into these questions", adding that the "UK is world-leading in this at the moment" and has the "intellectual leadership", as exemplified by the establishment of the Future of Humanity Institute at the University of Oxford and the Centre for the Study of Existential Risk at the University of Cambridge.[141] Evidence submitted jointly by these bodies suggested that the UK's expertise could be applied to best effect through

133 Q5; see also EPSRC UK-RAS Network (ROB0032) para 4.3
134 Q19 [Dr Buckingham]
135 techUK (ROB0063) para 37
136 Harry Armstrong, Machines That Learn in the Wild, Nesta, July 2015, p15
137 Google DeepMind (ROB0062) para 5.1
138 IEEE Standards Association Introduces Global Initiative for Ethical Considerations in the Design of Autonomous Systems, IEEE, 5 April 2016
139 Committee on Legal Affairs, European Parliament, Draft Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), May 2016
140 Geoff Mulgan, A machine intelligence commission for the UK: how to grow informed public trust and maximise the positive impact of smart machines, Nesta, February 2016
141 Q56 [Dr Cotton-Barratt]; see also AAAI and UKCRC (ROB0021)

a “Warnock-Style” Commission, in reference to Baroness Warnock’s examination of the ethics of IVF in the early 1980s.[142] Elsewhere, Nesta has made the case for a “Machine Intelligence Commission”, possessing powers similar to those of the now disbanded Royal Commission on Environmental Pollution.[143]

66. There has been some discussion about who should be involved in identifying, and establishing, suitable governance frameworks for robotics and AI. Kate Crawford from Microsoft has argued that:

> Like all technologies before it, artificial intelligence will reflect the values of its creators. So inclusivity matters—from who designs it to who sits on the company boards and which ethical perspectives are included.[144]

Dave Coplin from Microsoft told us that it was a task “for the tech industry, the Government, NGOs and the people who will ultimately consume the services” and emphasised that it was important “to find a way of convening those four parties together to drive forward that conversation.”[145] Dr Cotton-Barratt similarly recommended a broad “community of interest [that] would include AI researchers, social scientists and ethicists, representatives of industry and ministries”.[146]

Public dialogue

67. Professor Nick Jennings was clear that engagement with the public on robotics and AI needed “to start now so that people are aware of the facts when they are drawing up their opinions and they are given sensible views about what the future might be”.[147] He contrasted this with the approach previously taken towards GM plants which, he reflected, did “not really engage the public early and quickly enough”.[148]

68. A range of views were expressed about the role of public dialogue on robotics and AI. For some, it was a way to help build public trust and acceptance[149], and to tackle public “misconceptions”.[150] For others—including the Government—it was also about acknowledging, and improving, our understanding of the publics’ concerns.[151] A small number of witnesses suggested that the public have a role to play in *directing* the development of AI. Professor Luckin, for example, emphasised that developments in AI to date had focused predominately “on the technology and not on the problems it could solve”, adding that it would “be good if it could be more challenge-focused”.[152]

142 Future of Humanity Institute, Centre for the Study of Existential Risk, Global Priorities Project, and Future of Life Institute (ROB0052). This point was also made by Future Advocacy (ROB0047) para 2.5.
143 Geoff Mulgan, A machine intelligence commission for the UK: how to grow informed public trust and maximise the positive impact of smart machines, Nesta, February 2016
144 Artificial Intelligence’s White Guy Problem, The New York Times, 25 June 2016
145 Q124
146 Dr Owen Cotton-Barrett (ROB0074)
147 Q7
148 Q7
149 Q5; Q32 [Dr Buckingham]; Lloyd’s Register Foundation (ROB0065) para 16; Robotics & Autonomous Systems Special Interest Group (ROB0027) para 36
150 Royal Academy of Engineering (ROB0042) para 31
151 Department for Business, Innovation and Skills (BIS) (ROB0066) para 28; The Law Society (ROB0037) para 12; The Involve Foundation (ROB0025) para 2.12
152 Q108 [Professor Luckin]

69. Similarly, Paul Doyle from Hereward College—which supports young people with physical, sensory and cognitive disabilities—told us that, where assistive robotics were concerned, there remained a "massive disconnect" between "what is being produced in the University laboratory/workshop and what is needed in the thousands of homes across the UK".[153] Hereward College, he noted, had tried to bring "end users' perspectives" to the attention of research communities.[154] Robotics, as Population Matters noted, could improve the mobility of people with disabilities and "offer them a voice".[155] Pupils 2 Parliament—a group of 61 primary school children aged 9 and 10—also identified helping disabled people "move and walk" as their top priority for the "future development of robots".[156] A strong public role could thus facilitate greater scrutiny of the underlying motives behind advancements in robotics and AI, and provide a societal, rather than purely technological, perspective on how they could be developed.

70. The Royal Academy of Engineering suggested that the Government "could do more to open dialogue with the public on these issues so that concerns about social, legal and ethical issues are addressed in a timely way".[157] The Academy and others pointed to the support that 'Sciencewise'—the UK's national centre for public dialogue in policy making involving science and technology issues—could provide. In March 2016, for example, Sciencewise had hosted a "RAS Policy and the Public" workshop that "identified a number of specific ethical, legal and social" issues.[158] An overview of the session on the Sciencewise website indicates that invitations to the event were sent to "Government policy makers, academics and industry leaders".[159] The Involve Foundation, however, stressed that effective public dialogue on robotics and AI required consulting as *broadly* as possible:

> Policy development around these topics should not be restricted to involving a narrow range of expert stakeholders, but should also be informed by, and responsive to, broader public opinion.[160]

71. **While it is too soon to set down sector-wide regulations for this nascent field, it is vital that careful scrutiny of the ethical, legal and societal dimensions of artificially intelligent systems begins now. Not only would this help to ensure that the UK remains focused on developing 'socially beneficial' AI systems, it would also represent an important step towards fostering public dialogue about, and trust in, such systems over time.**

72. **Our inquiry has illuminated many of the key ethical issues requiring serious consideration—verification and validation, decision-making transparency, minimising bias, increasing accountability, privacy and safety. As the field continues to advance at a rapid pace, these factors require ongoing monitoring, so that the need for effective governance is continually assessed and acted upon. The UK is world-leading when it comes to considering the implications of AI and is well-placed to provide global intellectual leadership on this matter.**

153 Hereward College (ROB0028) para 3
154 Hereward College (ROB0028) para 3
155 Population Matters (ROB0007)
156 Pupils 2 Parliament (ROB0030) para 28
157 Royal Academy of Engineering (ROB0042) para 38
158 Research Councils UK (ROB0033) para 44
159 Sciencewise event on Robotics draws wide interest in potential future dialogue, Sciencewise press notice, not dated.
160 The Involve Foundation (ROB0025) para 2.5

73. ***We recommend that a standing Commission on Artificial Intelligence be established, based at the Alan Turing Institute, to examine the social, ethical and legal implications of recent and potential developments in AI. It should focus on establishing principles to govern the development and application of AI techniques, as well as advising the Government of any regulation required on limits to its progression. It will need to be closely coordinated with the work of the Council of Data Ethics which the Government is currently setting up following the recommendation made in our Big Data Dilemma report.***

74. ***Membership of the Commission should be broad and include those with expertise in law, social science and philosophy, as well as computer scientists, natural scientists, mathematicians and engineers. Members drawn from industry, NGOs and the public, should also be included and a programme of wide ranging public dialogue instituted.***

4 Research, funding and innovation

75. We conclude this report by examining the research, funding and innovation landscape for robotics and AI and identifying barriers to progress.

Robotics and autonomous systems

76. In 2013, the Government identified 'Robotics and autonomous systems' (RAS) as one the its 'Eight Great Technologies', rather than using the 'robotics and AI' label, which is more commonly found in the digital sector. The 'Eight Greats' are technologies in which the Government anticipated that "the UK [was] set to be a global leader".[161] RAS was defined as:

> interconnected, interactive, cognitive and physical tools, able to variously perceive their environments, reason about events, make or revise plans and control their actions. They perform useful tasks for us in the real world, extending our capabilities, increasing our productivity and reducing our risks.[162]

77. Though the different labels appear essentially to cover the same technology, witnesses did highlight that there was scope for confusion between 'automated' and 'autonomous' systems.[163] Innovate UK noted that the vast majority of current robotics systems are automated, rather than autonomous, but that the concepts are sometimes used interchangeably, despite having different meanings.[164] Both refer to processes that may be executed independently, from start to finish, without any human intervention. Automated processes, however, usually involve well-defined tasks that have known, consistent outcomes and are conducted in structured, predictable environments, such as factories. More often than not, an automated process replaces a routine, manual task with software/hardware that repeatedly follows a step-by-step sequence.

78. Industrial robots performing a limited range of "physically difficult, dangerous, or repetitive tasks" currently dominate this market.[165] For example, 'Latro', a robotic spider can step around, or climb over, obstacles. Though Latro remains under development, it is anticipated that the robot will eventually be able to play an important role in UK nuclear decommissioning by grabbing waste, chopping it up and dropping it in a skip.[166]

79. An autonomous system, in contrast, will have the capability to learn, respond and adapt (within set boundaries) to situations that were not pre-programmed or anticipated in the design. This type of system may also be able to make decisions "based on external events and internal goals that lead to different courses of action, even when faced with unexpected events and unknown environments".[167] An autonomous car, for example, is designed to react to weather and traffic conditions. In this respect, there are similarities

161 HM Government, Eight Great Technologies infographic, October 2013

162 Special Interest Group, Robotics and Autonomous Systems, *RAS 2020 Robotics and Autonomous Systems*, July 2014, p 4

163 techUK (ROB0063) para 9; Innovate UK (ROB0060) para 6

164 Innovate UK (ROB0060) para 6

165 Innovate UK (ROB0060) para 6

166 Nuclear Decommissioning Authority, The role of robotics in nuclear decommissioning, June 2016

167 Special Interest Group, Robotics and Autonomous Systems, *RAS 2020 Robotics and Autonomous Systems*, July 2014, p 7

between autonomous systems and artificial intelligence. As robots become smaller, and more nimble, we are also beginning to see them working in conjunction with humans. The results to date have been encouraging, particularly in healthcare. Early in September 2016, a British surgeon restored a patient's sight by operating a robot inside their eye, via a joystick, to "remove a membrane one hundredth of a millimetre thick". The robot was described as "acting like a mechanical hand" that was "able to filter out hand tremors from the surgeon".[168]

RAS 2020 Strategy

80. To stimulate collaboration and innovation in RAS capabilities, a RAS 'Special Interest Group' (SIG), comprising academics and industrialists, was established in 2013 with support from Innovate UK. The SIG subsequently produced *RAS 2020*, a national strategy for RAS, in July 2014. Its objective was "to capture value in a cross-sector UK RAS innovation pipeline through co-ordinated development of assets, challenges, clusters and skills" and included the following eight recommendations aimed at realising that goal:

> Invest further in the five RAS strategy strands: coordination, assets, challenges, clusters and skills to build the UK's RAS capability.
>
> Establish the means for funding agencies to formally work together in execution, so that ideas, people and activity flow readily from basic investigation through early stage demonstration to fully trialled commercial product.
>
> Establish a RAS Leadership Council to engage with senior leaders across a range of sectors in industry, academia and Government, providing independent advisory oversight of planning and execution of the strategy.
>
> Further develop engagement with the EU, investors and corporate resources in the UK and overseas to fuel the development of the 5 strands.
>
> Continue to consult widely on potential Assets and cross sector Grand Challenges.
>
> Continue to develop dialogue with those involved in standards and regulation, such as BSI and CAA, to develop more detailed thinking.
>
> Extend outreach and public engagement activities to continue changing public perceptions and improve understanding of public concerns.
>
> Articulate to businesses and investors internationally (e.g. through UKTI) that the UK aims to be the best place to invest in taking RAS technologies to market.[169]

81. Drawing on analysis by McKinsey, the strategy estimated that by 2025 RAS technologies would "have an impact on global markets of between $1.9 and $6.4 trillion per annum". Witnesses thought that the UK had the potential to take a lead in developing and adopting the next generation of advanced robotics, with Innovate UK singling out

168 "Robot operates inside eye in world first", BBC News Online, 9 September 2016

169 Special Interest Group, Robotics and Autonomous Systems, *RAS 2020 Robotics and Autonomous Systems*, July 2014, p 7

"the service segment, the global market for which is expected to grow from $7 billion in 2014 to $18 billion by 2020".[170] It concluded that a "sensible aspiration would be for the UK to capture perhaps 10% of this market".[171]

82. At present, however, the take up of robotics in the UK is low when compared to other countries. Paul Mason, Director of Emerging and Enabling Technologies at Innovate UK, told us that "the numbers for last year [2015] show that installed shipments of robots in China were 75,000 compared with 2,400 in the UK".[172] Professor Philip Nelson of Research Councils UK thought that the UK was probably "off the pace […] in the deployment of robotics" and, consequently, was probably losing market share.[173] He noted that South Korea had been investing "$100 million per year [in robotics] for the past 10 years or so", while Japan had "just put $350 million into a big programme of assistive robotics".[174] The UK, in contrast, had been identified in a study by the Copenhagen Business School—which modelled how much productivity in manufacturing would increase if all industries in a country had the highest-found level of robot-intensity—as "having the greatest potential for improvement".[175]

83. Our inquiry highlighted three barriers inhibiting the type of progress envisaged in the *RAS 2020* strategy—funding, leadership and skills shortages.

RAS Funding

84. Several of our witnesses questioned whether the current level of funding for, and investment in, RAS was sufficient to improve the UK's position in the field.[176] Some, for example, indicated that European Union, rather than UK, funding was underpinning the country's strong track record in robotics and AI research. Geoff Pegman of R U Robots Ltd thought that the UK would not have its current capabilities in both robotics and artificial intelligence "without the European Framework programmes".[177] Similarly, Professor Rose Luckin from the UCL Institute of Education told us that "the funding that the research community has taken advantage of to hold its position internationally [in artificial intelligence and education research] has all come from the European Union".[178] In 2015, the Prime Minister's Council for Science and Technology calculated that 80% of funding for RAS research came from the EU and stressed that "if this funding stream was disrupted then RAS industry and research in the UK may suffer as a consequence".[179]

170 Q3; Q6; Innovate UK (ROB0060)
171 Innovate UK (ROB0060) para 11
172 Q86 [Paul Mason]
173 Q23; Q12
174 Q12
175 Q23 - the study estimated that productivity would rise in the UK by 22%. See Lene Kromann, Jan Rose Skaksen, Anders Sorensen, Automation, labor productivity and employment - a cross country comparison, CEBR, Copenhagen Business School, Working Paper, 2011
176 Manufacturing Technology Centre (ROB0018), para 14; Professor Tony J. Prescott (ROB0020); Royal Academy of Engineering (ROB0042) paras 22–23
177 Geoff Pegman (ROB0059) para 4.1
178 Q87 [Professor Luckin]
179 Prime Minister's Council for Science and Technology, Science Landscape Seminar Reports: Robotics and Autonomous Systems (RAS), June 2015

85. Others noted that the UK *had* invested in RAS, with Research Councils UK reporting that "EPSRC funding in the robotics research area [had] increased threefold since 2010 and currently [stood] at £33.8M".[180] The EPSRC's UK-RAS Network stated that this funding had:

> resulted in the creation of eight dedicated centres/facilities across the country covering key areas of transport, healthcare, manufacturing, and unmanned systems to ensure that the UK will maintain its leading engineering and research capacity in RAS.[181]

86. The Government also drew our attention to £19 million provided to fund four 'EPSRC Centres for Doctoral Training' in Edinburgh, Bristol, Oxford and Loughborough.[182] While welcoming these investments, Research Councils UK worried that there was still a lack of "UK-wide critical mass required to move from isolated pockets of excellence to the formation of a national research, training and innovation infrastructure required to enable cross-sector exploitation".[183] It concluded that there remained "a need for greater coordination and visibility in RAS across both academia and industry".[184]

87. A similar point was made by ABB Robotics, a supplier of industrial robots. It told us that the RAS Special Interest Group (SIG) had "done a great deal of work with universities in identifying the technologies of the future and [had] invested heavily in research and development" but that the SIG's focus on universities had come at the expense of considering the needs of industry. [185] According to Mike Wilson of ABB Robotics, the *RAS 2020* strategy had been "driven largely by academic interests looking at the more advanced robotic technologies".[186] The company complained that by:

> concentrating on advanced manufacturing, the [Special Interest] Group [had] received little input from industrial robotics and UK manufacturing businesses. As such, the Group has had very little impact on productivity.[187]

RAS Leadership

88. Witnesses suggested that coordination between academia and industry could be improved through establishing cross-cutting leadership in RAS. Professor Nick Jennings, representing the Royal Society's Machine Learning Working Group, told us that "lots of excellent work" was "going on in bits and pieces around the country" but that "something that tries to bring those together and show leadership" was required.[188]

89. Establishing a 'RAS Leadership Group' had been recommended in *RAS 2020*, with the aim of engaging "with senior leaders across a range of sectors in industry, academia

180 Research Councils UK (ROB0033) para 19. The EPSRC is the Engineering and Physical Sciences Research Council.
181 EPSRC UK-RAS Network (ROB0032) para 3.1
182 Department for Business, Innovation and Skills (BIS) (ROB0066) para 10; Department for Business, Energy and Industrial Strategy (ROB0076)
183 Research Councils UK (ROB0033) para 14
184 Research Councils UK (ROB0033) para 38
185 ABB (ROB0026)
186 Q145
187 ABB (ROB0026)
188 Q28 [Professor Jennings]

and Government, providing independent advisory oversight of planning and execution of the strategy".[189] In response to *RAS 2020*, the then Universities, Science and Cities Minister, Greg Clark MP, stated in March 2015 that the:

> Government agrees with this recommendation. We will establish a Leadership Council in Robotics and Autonomous Systems. I have asked officials to implement this recommendation.[190]

90. However, Dr Rob Buckingham, one of the co-authors of *RAS 2020*, told us that setting up a Leadership Council "did not happen; there was a change of Government and a slight change of direction".[191] Professor Nelson explained that the Research Councils were "still trying to co-ordinate matters in the absence of a leadership council" adding that "something like [a leadership council] to interact with, or be a governing body for, a national initiative in this area […] is going to be very important".[192]

91. Dr Buckingham suggested that there was scope for the Government to take a much more active role in advancing the RAS sector. The RAS Strategy, he explained, "was not paid for or requested by Government; it was a slightly external thing done under EPSRC, KTNs and Innovate UK and we need it to be formally adopted".[193] Paul Mason from Innovate UK confirmed that *RAS 2020* was "not an official Government strategy".[194] On that basis, he thought it "would be worth looking at a more overarching strategy document that might bring about strong Government action in areas like regulation, standards or procurement".[195]

92. Research Councils UK advocated a complementary approach, namely establishing a RAS national flagship institute "with the dual purpose of accelerating emerging technologies to commercialisation and inspiring new disruptive academic research with the potential to open up new markets".[196] In a similar vein, Innovate UK told us that establishing an institute (or a Catapult)[197], would "provide a coherent national focal point for market-led RAS activity, presenting a visible and open front door to engage end users and international inward investment".[198]

189 Special Interest Group, Robotics and Autonomous Systems, *RAS 2020 Robotics and Autonomous Systems*, July 2014, p 3

190 Minister for Universities, Science and Cities, Response to the Robotics and Autonomous Systems Strategy, March 2015

191 Q26

192 Q28 [Professor Nelson]

193 Q27

194 Q103

195 Q103

196 Research Councils UK (ROB0033) para 6

197 Catapult Centres are physical centres, established by Innovate UK, to "transform the UK's capability for innovation in specific areas" through enabling businesses, scientists and engineers to work side by side. See http://www.catapult.org.uk/about-us/

198 Innovate UK (ROB0060) para 34; see also Research Councils UK (ROB0033) paras 38–40; Robotics & Autonomous Systems Special Interest Group (ROB0027) para 33; Northern Robotics Network (ROB0040) para 1.7

93. When we asked the Government why the RAS Leadership Council had not been established, it stated that while the:

> Rt. Hon Greg Clark MP said that the coalition Government would establish a RAS Leadership Council […] things have moved on since then and a number of activities in this area have been introduced to support the development of RAS.[199]

These activities involved strengthening "coordination and oversight between departmental interests in Robotics and Autonomous Systems, including through a cross-government senior officials meeting initiated by Cabinet Office", as well as appointing a "RAS lead" in the Department for Business, Energy and Industrial Strategy, to "coordinate the efforts within the Department and Arm's Length Bodies". The Government also added that it intended "to utilise the upcoming Industrial Strategy to drive the progression of RAS".[200]

RAS Skills shortages

94. In our reports on Big Data, Digital Skills, and Satellites and Space, we highlighted that the UK is facing a digital skills crisis.[201] Witnesses to our current inquiry emphasised that while the demand for expertise in RAS and AI was booming, the UK lacked sufficient numbers of qualified people in these fields.[202] According to Professor Nick Jennings, the UK still did "not [have] enough people […] doing the basic computer science required"[203], while Dr Buckingham stressed that there remained a "vital" need for "more people who are good at STEM".[204] The RAS Special Interest Group were concerned that the growth of the UK robotics industry was being:

> restricted by the output of engineering degree programmes and by competing industries. Of critical concern is the significant shortage of systems engineers; those engineers able to understand the complexities of developing systems and systems of systems that robotics and artificial intelligence are built on.[205]

95. There appears, however, to be a shift underway in the subject and career choices made by university graduates. Professor Nelson emphasised that the EPSRC's centres for doctoral training in RAS were "over-subscribed", with Research Councils UK reporting that some centres were receiving "an average of 100 applicants per year for around 10 places".[206] Similarly, Professor Stephen Muggleton, representing the Association for the Advancement of Artificial Intelligence (AAAI), reflected that while the majority of Imperial College graduates had previously gone into "the financial industry after they got

199 Department for Business, Energy and Industrial Strategy (ROB0076)
200 Department for Business, Energy and Industrial Strategy (ROB0076)
201 Science and Technology Committee, Second Report of Session 2016–17, Digital skills crisis, HC 270; Science and Technology Committee, Third Report of Session 2016–17, Satellites and space, HC 160, para 48; Science and Technology Committee, Fourth Report of Session 2015–16, The big data dilemma, HC 468, para 27
202 techUK (ROB0063) para 24
203 Q42
204 Q45 [Dr Buckingham]
205 Robotics & Autonomous Systems Special Interest Group (ROB0027) para 19
206 Q45 [Professor Nelson]; Research Councils UK (ROB0033) para 16

their undergraduate degree", they were now "increasingly staying on to do postgraduate work because the big take-up is in the tech industry". He added that "companies like DeepMind" were "hoovering up quality people".[207]

96. Professor Muggleton's example of London-based DeepMind is illustrative of how the UK has been at the forefront of establishing new, AI-driven companies, a number of which have subsequently been acquired by the world's biggest technology firms. DeepMind, for example, was acquired by Google for a reported £400 million in 2014.[208] Less than two years later, Microsoft bought 'Swiftkey', the makers of a predictive keyboard powered by AI, in a deal worth an estimated $250 million.[209] 'Magic Pony', a company which uses AI techniques to understand the features of an image, was bought by Twitter in the summer of 2016 "for an undisclosed sum rumoured to be about $150 million".[210]

97. An editorial in *Nature*, however, recently warned that this type of "industrial migration", out of computer science departments and into AI start-ups, could leave universities "temporarily devoid of top talent" thereby reducing "the number of students that can be trained, especially at PhD level". *Nature* also questioned whether this movement of people could "ultimately sway the field towards commercial endeavours at the expense of fundamental research".[211]

98. **Despite identifying Robotics and Autonomous Systems (RAS) as one of the 'Eight Great Technologies' in 2013, Government leadership has been noticeably lacking. There is no Government strategy for developing the skills, and securing the critical investment, that is needed to create future growth in robotics and AI. Nor is there any sign of the Government delivering on its promise to establish a 'RAS Leadership Council' to provide much needed coordination and direction. Without a Government strategy for the sector, the productivity gains that could be achieved through greater uptake of the technologies across the UK will remain unrealised.**

99. ***The Government should, without further delay, establish a RAS Leadership Council, with membership drawn from across academia, industry and, crucially, the Government. The Leadership Council should work with the Government and the Research Councils to produce a Government-backed 'National RAS Strategy'; one that clearly sets out the Government's ambitions, and financial support, for this 'great technology'. Founding a 'National RAS Institute', or Catapult, should be part of the strategy.***

207 Q41
208 Google buys UK artificial intelligence startup Deepmind for £400m, The Guardian, 27 January 2014
209 Microsoft buys British keyboard apps firm SwiftKey, The Guardian, 3 February 2016
210 Q105
211 Elizabeth Gibney, "AI firms lure academics", Nature, vol 532 (2016) p 422

Annex

Visit to Google DeepMind, London

On Monday 6 June 2016, Committee Members visited Google DeepMind—one of the largest machine learning labs in the world—based in King's Cross, London. The then Chair, Nicola Blackwood MP, Victoria Borwick MP and Matt Warman MP were present. Following a tour of the building, Members met with Mustafa Suleyman, one of the co-founders of the company, and were presented with an overview of Google DeepMind's work to date. Members were then joined by Demis Hassabis, another co-founder of the company, and took part in a question and answer session. Topics covered during the discussion included:

(1) **Background to DeepMind's work:** This included an overview of how artificial intelligence has developed over the years, from IBM's Deep Blue —which evaluated 200 million positions per second—beating Garry Kasparov at chess in 1997 to the 'Pocket Fritz' computer—which needed to search less than 20,000 chess moves per second—winning the Mercosur Cup in Argentina in 2009, and achieving a higher performance level than DeepBlue. A video was also shown highlighting DeepMind's success at building an artificial intelligence agent that could learn directly, via experience, to play classic Atari games better than humans.

(2) **AlphaGo:** the game of Go, and why it is considered the most difficult game devised by humans, was discussed together with how AlphaGo trained for the match against Lee Sedol in March 2016, and the reinforcement learning framework it used. It was noted that AlphaGo made unprecedented moves that commentators initially thought were mistakes and that, in doing so, AlphaGo has provided human Go players with new knowledge and new insights on the game. Lee Sedol, for example, has won all his games since he played AlphaGo and has stated that it taught him to be more creative in his play.

(3) **DeepMind Health:** it was explained that DeepMind Health was focused on improving patient safety and that they were adopting a user-centred approach to research so that they could learn more about what tools doctors and nurses need to improve patient safety. DeepMind Health's work on acute kidney injury (AKI) was highlighted, alongside its development of an app to help a) detect which patients are deteriorating and b) manage the subsequent intervention. Though it was anticipated that AI could become a powerful modelling tool that could be applied to solve 'wicked problems' in the future, it was noted that the AKI app does not use AI.

(4) **Ethics:** it was acknowledged that AI will have significant implications and that ethics and safety must be paramount. Google DeepMind noted that it was talking with other companies developing AI about standards and best practices.

(5) **Safety:** it was noted that safety is incredibly important to the development of AI and that responsible measures must be taken—such as by enabling 'safe exploration'—to ensure its operation within defined constraints.[212]

212 For further information about 'safe exploration' see: Dario Amodei et al, Concrete Problems in AI Safety, 21 June 2016

Conclusions and recommendations

Education and skills

1. Advances in robotics and AI hold the potential to reshape fundamentally the way we live and work. While we cannot yet foresee exactly how this 'fourth industrial revolution' will play out, we know that gains in productivity and efficiency, new services and jobs, and improved support in existing roles are all on the horizon, alongside the potential loss of well-established occupations. Such transitions will be challenging. As a nation, we must respond with a readiness to re-skill, and up-skill, on a continuing basis. This requires a commitment by the Government to ensure that our education and training systems are flexible, so that they can adapt as the demands on the workforce change, and are geared up for lifelong learning. Leadership in this area, however, has been lacking. It is disappointing that the Government has still not published its Digital Strategy and set out its plans for equipping the future workforce with the digital skills it needs to thrive. (Paragraph 36)

2. *Digital exclusion has no place in 21st century Britain. As we recommended in our Big Data Dilemma, Digital Skills Crisis, and Satellites and Space reports, the Government must commit to addressing the digital skills crisis through a Digital Strategy, published without delay.* (Paragraph 37)

Governance: standards and regulations

3. While it is too soon to set down sector-wide regulations for this nascent field, it is vital that careful scrutiny of the ethical, legal and societal dimensions of artificially intelligent systems begins now. Not only would this help to ensure that the UK remains focused on developing 'socially beneficial' AI systems, it would also represent an important step towards fostering public dialogue about, and trust in, such systems over time. (Paragraph 71)

4. Our inquiry has illuminated many of the key ethical issues requiring serious consideration—verification and validation, decision-making transparency, minimising bias, increasing accountability, privacy and safety. As the field continues to advance at a rapid pace, these factors require ongoing monitoring, so that the need for effective governance is continually assessed and acted upon. The UK is world-leading when it comes to considering the implications of AI and is well-placed to provide global intellectual leadership on this matter. (Paragraph 72)

5. *We recommend that a standing Commission on Artificial Intelligence be established, based at the Alan Turing Institute, to examine the social, ethical and legal implications of recent and potential developments in AI. It should focus on establishing principles to govern the development and application of AI techniques, as well as advising the Government of any regulation required on limits to its progression. It will need to be closely coordinated with the work of the Council of Data Ethics which the Government is currently setting up following the recommendation made in our Big Data Dilemma report.* (Paragraph 73)

6. *Membership of the Commission should be broad and include those with expertise in law, social science and philosophy, as well as computer scientists, natural scientists, mathematicians and engineers. Members drawn from industry, NGOs and the public, should also be included and a programme of wide ranging public dialogue instituted.* (Paragraph 74)

Research, funding and innovation

7. Despite identifying Robotics and Autonomous Systems (RAS) as one of the 'Eight Great Technologies' in 2013, Government leadership has been noticeably lacking. There is no Government strategy for developing the skills, and securing the critical investment, that is needed to create future growth in robotics and AI. Nor is there any sign of the Government delivering on its promise to establish a 'RAS Leadership Council' to provide much needed coordination and direction. Without a Government strategy for the sector, the productivity gains that could be achieved through greater uptake of the technologies across the UK will remain unrealised. (Paragraph 98)

8. *The Government should, without further delay, establish a RAS Leadership Council, with membership drawn from across academia, industry and, crucially, the Government. The Leadership Council should work with the Government and the Research Councils to produce a Government-backed 'National RAS Strategy'; one that clearly sets out the Government's ambitions, and financial support, for this 'great technology'. Founding a 'National RAS Institute', or Catapult, should be part of the strategy.* (Paragraph 99)

Formal Minutes

Tuesday 13 September 2016

Members present:

Victoria Borwick
Chris Green
Dr Tania Mathias
Carol Monaghan
Matt Warman

Dr Tania Mathias took the Chair, in accordance with the Resolution of the Committee of 19 July 2016.

Draft Report (*Robotics and Artificial Intelligence*), proposed by the Chair, brought up and read.

Ordered, That the draft Report be read a second time, paragraph by paragraph.

Paragraphs 1 to 99 read and agreed to.

Annex read and agreed to.

Summary agreed to.

Resolved, That the Report be the Fifth Report of the Committee to the House.

Ordered, That the Chair make the Report to the House.

Ordered, That embargoed copies of the Report be made available, in accordance with the provisions of Standing Order No. 134.

[Adjourned till Wednesday 14 September at 10.00 am

Witnesses

The following witnesses gave evidence. Transcripts can be viewed on the inquiry publications page of the Committee's website.

Tuesday 24 May 2016

Question number

Dr Rob Buckingham, Director, RACE, **Professor Stephen Muggleton**, Association for the Advancement of Artificial Intelligence, **Professor Nick Jennings**, Royal Society Working Group on Machine Learning, and **Professor Philip Nelson**, Chair, Research Councils UK Q1–53

Richard Moyes, Managing Partner, Article 36, and **Dr Owen Cotton-Barratt**, Research Fellow, Future of Humanity Institute Q54–82

Tuesday 28 June 2016

Professor Rose Luckin, Chair of Learning with Digital Technologies, UCL Institute of Education, University College London, **Angus Knowles-Cutler**, Vice Chairman and London Office Senior Partner, Deloitte, **Dr Michael Osborne**, Dyson Associate Professor in Machine Learning, University of Oxford, and **Paul Mason**, Director of Emerging and Enabling Technologies, Innovate UK Q83–117

Dave Coplin, Chief Envisioning Officer, Microsoft, and **Mike Wilson**, General Industry Sales Manager, ABB Robotics Q118–159

Published written evidence

The following written evidence was received and can be viewed on the inquiry publications page of the Committee's website.

ROB numbers are generated by the evidence processing system and so may not be complete.

1 ABB (ROB0026)

2 Andrew Brereton (ROB0004)

3 Article 36 (ROB0029)

4 Autonomous Intelligent Systems Partnership (AISP) (ROB0049)

5 AZTIX (ROB0048)

6 BAE Systems MA&I (ROB0067)

7 BioCentre (ROB0061)

8 Braintree Ltd (ROB0038)

9 Cambridge University Science Policy Exchange (ROB0001)

10 D&TforD&T (ROB0017)

11 Deloitte (ROB0019)

12 Department for Business, Innovation and Skills (ROB0066)

13 Department for Business, Energy and Industrial Strategy (ROB0076)

14 Department of Automatic Control & Systems Engineering, The University of Sheffield (ROB0016)

15 Digital Wisdom Institute (ROB0072)

16 Dr Andrew Stewart (ROB0058)

17 Dr J. Robert Michael (ROB0013)

18 Dr Owen Cotton-Barrett (ROB0074)

19 Dr Thrishantha Nanayakkara (ROB0010)

20 Dr Toby Walsh (ROB0012)

21 Dr Will Slocombe (ROB0015)

22 Edge Foundation (ROB0011)

23 EMLSRI (ROB0057)

24 EPSRC UK-RAS Network (ROB0032)

25 Future Advocacy (ROB0047)

26 Future of Humanity Institute, Centre for the Study of Existential Risk, Global Priorities Project, Future of Life Institute (ROB0052)

27 Global Priorities Project (ROB0051)

28 Google DeepMind (ROB0062)

29 Hereward College (ROB0028)

30 Informed.AI (ROB0009)

31 Innovate UK (ROB0060) and (ROB0075)

32 John Phillips (ROB0069)

33 Joint response from AAAI and UKCRC (ROB0021)
34 Lloyd's Register Foundation (ROB0065)
35 Manufacturing Technology Centre (ROB0018)
36 Microsoft (ROB0046)
37 Mr Geoff Pegman (ROB0059)
38 Nesta (ROB0034)
39 Northern Robotics Network (ROB0040)
40 Nutmeg Saving and Investment Ltd (ROB0035)
41 Optima Control Solutions Ltd. (ROB0068)
42 Ordnance Survey (ROB0039)
43 Peter Morgan (ROB0003)
44 Population Matters (ROB0007)
45 Professor Alan Winfield (ROB0070)
46 Professor David Lane (ROB0064)
47 Professor Huw Price (ROB0031)
48 Professor Ian Pyle (ROB0071)
49 Professor Rose Luckin (ROB0043) and (ROB0073)
50 Professor Sandor M Veres (ROB0055)
51 Professor Tony Prescott (ROB0020)
52 Pupils 2 Parliament (ROB0030)
53 RACE, UK Atomic Energy Authority (ROB0041)
54 Research Councils UK (ROB0033)
55 Rocket Fuel (ROB0036)
56 Royal Academy of Engineering (ROB0042)
57 Simon Beard (ROB0045)
58 techUK (ROB0063)
59 The Involve Foundation (ROB0025)
60 The Law Society (ROB0037)
61 The Robotics & Autonomous Systems Special Interest Group (ROB0027)
62 The Royal Society (ROB0023)
63 Transpolitica (ROB0044)
64 Transport Systems Catapult (ROB0014)
65 UCL Knowledge Lab (ROB0024)

List of Reports from the Committee during the current Parliament

All publications from the Committee are available on the publications page of the Committee's website.

The reference number of the Government's response to each Report is printed in brackets after the HC printing number.

Session 2016–2017

First Report	EU regulation of the life sciences	HC 158
Second Report	Digital skills crisis	HC 270
Third Report	Satellites and space	HC 160
Fourth Report	Forensic Science Strategy	HC 501
Sixth Report	Evidence Check: Smart metering of electricity and gas	HC 161

Session 2015–2016

First Report	The science budget	HC 340 (HC 729)
Second Report	Science in emergencies: UK lessons from Ebola	HC 469 (Cm 9236)
Third Report	Investigatory Powers Bill: technology issues	HC 573 (Cm 9219)
Fourth Report	The big data dilemma	HC 468 (HC 992)
First Special Report	Royal Botanic Gardens, Kew: Government Response to the Committee's Seventh Report of Session 2014–15	HC 454
Second Special Report	Current and future uses of biometric data and technologies: Government Response to the Committee's Sixth Report of Session 2014–15	HC 455
Third Special Report	Advanced genetic techniques for crop improvement: regulation, risk and precaution: Government Response to the Committee's Fifth Report of Session 2014–15	HC 519
Fourth Special Report	The science budget: Government Response to the Committee's First Report of Session 2015–16	HC 729
Fifth Special Report	The big data dilemma: Government Response to the Committee's Fourth Report of Session 2015–16	HC 992