AAAI and UKCRC Joint Response to the
UK Science and Technology Committee Inquiry in Robotics and AI

This is a joint response from two organisations:

AAAI: The Association for the Advancement of Artificial Intelligence(www.aaai.org). AAAI is an international organisation, founded in 1979, headquartered in California, the largest AI society in the world, with over 3000 members of whom around 39% are based outside North America. The response was coordinated by one of its Fellows, Prof. Anthony G Cohn, FREng FBCS FIET FAAAI (University of Leeds), who consulted with a representative selection of AAAI members, including UK based Fellows and other senior researchers. Prof. Cohn is also a member of UKCRC.

UKCRC: The UK Computing Research Committee (www.ukcrc.org.uk) is an Expert Panel of the British Computer Society (BCS), the Institution of Engineering and Technology (IET), and the Council of Professors and Heads of Computing (CPHC). It was formed in November 2000 as a policy committee for computing research in the UK. Members of UKCRC are leading computing researchers who each have an established international reputation in computing. Thus this response covers UK research in computing, which is internationally strong and vigorous, and is a major national asset, and was prepared after consultation amongst the membership of UKCRC. As such, it is an independent response on behalf of UKCRC and does not necessarily reflect the official opinion or position of the BCS or the IET. The UKCRC response was coordinated by Prof. Erol Gelenbe FACM FIEEE FIET (Imperial College), Fellow of the National Academy of Technologies of France and of the science academies of Belgium, Hungary, Poland and Turkey.

Note: in the sequel we abbreviate *Robotics and Artificial Intelligence* as *RAI*.

1. ***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.***

Scope and limitations of RAI, and the need for education of the public: Historically, both robotic systems and AI systems have performed at a high level in very specialised, narrow domains, and continue to do so, reaching ever higher levels of expert performance. Such expert levels of performance may lead some members of the public to misinterpret the general level of intelligence currently attainable by RAI, and not appreciate the challenges and fundamental difficulties that are encountered in RAI. One of these challenges is conceptual: whereas the big data revolution predominantly provides vast amounts of quantitative data, and data mining increasingly is able to find patterns and predict trends, there is a big jump to automatically generating new concepts and rules (axioms) linking them; it is instructive to consider the way progress has been made in mathematics over the centuries in which the great discoveries were made not by processing large quantities of data but by conceptual leaps to new and often very different kinds of formalisms to those found before. On the other hand, it is hard to think of areas of public life, the economy and technology which will not be impacted by RAI in the medium term. For all these reasons, education of the public and users of RAI is very likely to be necessary.

The role of humans in RAI systems: Since it is expected that RAI systems will continue to perform extremely well in narrowly defined domains for the short to medium term, there is a key role for humans in orchestrating their use, i.e., combining several systems in an appropriate way, detecting

when a system is out of scope and dealing with such "edge" cases, interfacing with customers on their requirements, etc. Used in such a way, AI systems have the potential to significantly increase productivity making goods and services cheaper and, hence, more widely available. However, as RAI develops and becomes more widely adopted over the coming decades, it is very likely that there will be many fewer unskilled jobs available, which implies that everyone aspiring to employment must be well educated and have a good understanding of the scope and limitations of the RAI systems that they work with, as already mentioned above in the context of the wider public's interaction with RAI.

Effect of RAI on the UK economy and the skill base: While RAI should enable developed economies such as the UK to become or stay competitive in a range of markets if it exploits the technology correctly, these technologies are becoming increasingly available to lower wage economies with whom the UK already competes as producers of knowledge based and other industrial products. The UK suffers in labour intensive industries because of wage costs, but automation can obviously alleviate the gap with economies that have lower wages by empowering our skilled and semi-skilled workers to be more productive. This is also the case for "information based" professions, such as legal and financial services, estate agents, financial analysts and traders who will increasingly need to rely on big data and on machine intelligence. However, the UK also needs to be aware that there is a risk that the cost gap between the UK and the lower cost economies may widen; this could happen if they are able to use smart automation and big data, including smart tutoring systems, to make up for their skill gaps and compete even more successfully with the UK.

In sectors which are critical for the UK, such as the financial industry, factors of less than a fraction of a percentage in profits on large volumes of trades, and decision times (including the time needed to retrieve data such as stock prices remotely) brought down to under a millisecond, can make a huge difference in our competitive edge. Thus in high end jobs, the UK will need the best smart technologies and competence, while, at the opposite end, our lower paid jobs will be increasingly at risk of being replaced by intelligent machines in factories, warehouses, shops, and services.

In this ever changing landscape, new professions will also arise: although RAI may dramatically the reduce the need for certain kinds of jobs, it is likely to open up new areas of employment and kinds of jobs as happened in the past (e.g. the ending of the profession of "hot type" printing, but there are now thousands employed as web designers). For instance, just as teachers educate and train people to acquire established and new knowledge and skills, the UK will need the people who can "design, teach and train" the assemblies of human beings and smart software that will carry out complex interdependent tasks, both in the manufacturing and service industries, and learn to be more effective as they exercise their skills and knowledge. In the human resources area we expect major changes with the use of big data and machine learning to identify talents, design or select career paths, and design individualised or company-wide training programmes for staff, raising substantial ethical and legal problems about the manner in which people are being orientated and selected.

Availability of RAI experts: There is a worldwide shortage of RAI experts at PhD level and above. Many international corporations such as IBM, Google, Apple, Amazon, Facebook, and even human resources and recruitment companies [Recruit Institute of Technology, http://recruit.ai/] are expanding. Many of them are setting up research and development laboratories in the UK, competing for RAI PhDs with University departments for staff, and offering much higher salaries. EPSRC has invested some resources in support for the research base, e.g. through Doctoral Training Centres, but these will probably not be sufficient to meet demand. Already we see many recruited to both University and industrial positions from abroad.

2. ***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.***

The need to accelerate UK innovation and exploitation in RAI: Germany has recently launched the *Industrie 4.0* programme that aims at exploiting smart technologies with *big data* and the *Internet of Things* to provide smart automation to industry sectors such as car manufacturing, and to accelerate the move to smart vehicles. This programme is accompanied by new educational and research investments in universities and technical schools. While there is increasing interest in UK industry to address such challenges, it seems that the short-term cost of investing in new software, hardware, research and education is hindering such developments in the UK. Thus it is worth examining how other countries are funding these changes, and how they make the case and find the resources for these significant investments.

3. ***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.***

Eight Great Technologies: Under the previous government *Robotics and Autonomous Systems* was identified as one of the "Eight Great Technologies" and became part of the UK's industrial strategy, welcomed by the 2015 minister for Universities, Science and Cities; this strategy should be continued and further resources should be devoted to it, not only from industry but also from government sources.

Funding models: the UK funding, research and innovation landscape in engineering relies heavily on the "catapult" model which is driven by problems that are posed by industry, including SMEs. On the other hand, many if not all of the major developments in ICT have traditionally started in research laboratories before being recognised and exploited by industry. Many major industries in ICT have grown directly from entrepreneurship that was born in the research laboratory. Thus there are arguments to support a "mixed catapult model" for RAI that is driven in part by bottom-up industry need-based research, but that also reserves an important role for basic, blue sky, innovative research and for academically driven ideas and prototypes. These would then be tested – in a second stage – as start-ups or products rolled out from laboratories to fill industry gaps or create new products.

In addition to these new catapult models, the creation of targeted calls or "transitional funding" that link up knowledge and technology providers (Universities, SMEs, robotics manufacturers etc.) with end users, to specifically build and evaluate prototype systems that solve real problems and show visible results, should be considered. While the EU's H2020 *Innovation Actions* are a step in this direction, they do not allow enough scope for new research developments and may be somewhat difficult to manage. Other forms of university-technology-business engagement, perhaps with appropriate tax incentives for business, could be considered to demonstrate and test concepts and prototypes in the field of RAI.

Need for fundamental research: It must be stressed that none of this obviates the need for fundamental research, not only in the mechanical engineering of robotic systems, but even more so in the AI aspects, including planning, perception, language understanding/generation, reasoning, multi-modal information fusion, modelling the human, and dealing with the uncertainty present in the real

world in which RAI systems must operate in general (unlike the controlled/artificial AlphaGo environment). Furthermore, very few current AI systems, even those which display impressive performance such as AlphaGo, or the growing generation of autonomous vehicles which do operate in the real world, really *understand* their own reasoning processes – let alone have the ability to explain their reasoning and be aware of the limitations. It is critical both for understanding the reproducibility of such systems and for their use in important applications such as Decision Support Systems for critical applications, including human health and safety, that research in this area is promoted. In support of such goals, the EPSRC has recently announced a priority area of *Human-Like Computing* (2016-2020). Human-Like Computing (HLC) research aims to endow machines with human-like perceptual, reasoning and learning abilities which support collaboration and communication with human beings. Such abilities should support computers in interpreting the aims and intentions of humans based on learning and accumulated background knowledge to help identify contexts and cues from human behaviour. The development of computer systems which exhibit truly human-like learning and cooperative properties will require sustained interdisciplinary collaboration between disparate and largely disconnected research communities within Psychology and Artificial Intelligence.

Formal modelling and verification of RAI: The whole field of formal modelling, verification measurement and performance evaluation of RAI systems is still very much in its infancy: 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 RAI systems before they are deployed. It should be noted that the verification of systems that adapt, plan and learn will involve the development of new modelling and verification approaches; moreover, such modelling and verification is a prerequisite for informed certification and regulation of RAI systems, which in turn is a factor in public acceptance of RAI.

Energy consumption: In addition, programmes should be launched to understand how big data and RAI can be used to reduce energy consumption by ICT (which has already reached some 1000 Terawatt-hours per year worldwide, more than the annual electricity consumption of Japan), and dynamically minimise the energy needs of ICT, maximise the benefits ICT provides, and dynamically match the available (including renewable) energy to the operation of ICT systems.

Research Council leadership: All these areas offer great opportunities for the UK Research Councils to assert a leadership role and more investments in research will be needed. The Engineering and Physical Sciences Research Council (EPSRC) have recently set up a UK Network on the "Verification and Validation of Autonomous Systems" cgi.csc.liv.ac.uk/~michael/VaVAS/ [http://gow.epsrc.ac.uk/NGBOViewGrant.aspx?GrantRef=EP/M027309/1] which is a first step towards addressing the safety, legal, ethical and trust issues surrounding RAI.

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

Robots and Autonomous Systems as a Force for Good: there is tremendous potential for RAI to improve human life and wellbeing and make our own tasks more interesting. For example autonomous cars have the potential to reduce the carnage on roads owing to driver errors, and to empower those currently unable to own and drive their own vehicle. Another example may be in health care and nursing scenarios. But there is also a danger that they will be seen as a threat and

rejected in the way GM crops are seen by many in the UK. Early UK programmes should therefore, in part, focus on applications which have positive societal benefits with openly disseminated results to act as exemplars that engage the public in a positive manner. In all applications, including for the military and security forces, RAI needs to be placed into a framework of usage which ensures monitoring and oversight by human beings, and compliance with ethical rules and regulations. Moreover, there are many arguments in favour of a world-wide ban on completely autonomous (i.e. human independent) weapons system, and a ban on other safety critical RAI not under human control. A formally verified system would mitigate some aspects of a lack of human control, but the difficulties of achieving this are immense, in particular to ensure accurate recognition of targets (and non-targets). The UK could take a critical role in this area by calling for a pre-emptive ban on lethal autonomous weapons without meaningful human control at the UN Convention on Certain Weapons (CCW). Without a robust system of ethical values, and verifiable performance embedded in RAI, giving complete autonomy to any autonomous system operating in a critical environment carries great risk.

The dangers of AI being too dumb rather than too smart: there have been a number of recent reports in the scientific and popular media of the dangers to humanity of "super intelligent AI". Any such danger if, at all, is certainly a long way off – we are many decades away from super intelligent AI – as has been shown by the last 50 years of progress in ICT, the exponential grown in hardware performance and capacities has not been matched by exponential growth in machine intelligence. Rather, the danger is in the public seeing isolated examples of narrow AI (e.g. AlphaGo) and assuming greater levels of comprehension and intelligence on the RAI systems they encounter in their everyday and professional lives. There is a need to ensure that deployed RAI systems have appropriate safety checks, humans in the loop, and are capable of knowing and advertising their strengths and limitations.

Robotics, Privacy, Exploitation and Acceptance: While ICT based industries often begin with start-ups, the more successful ones turn into behemoths that dominate their market (Google, Apple, Facebook, Uber) while yet other companies (even in Europe) reach a large enough size (SAP, ATOS) and dominate whole sectors of the economy. Thus the UK should consider how it can occupy a more prominent place in this development. However, there should also be a concern that large corporations may come to dominate the creation and deployment of robotics and rich sensor grids in homes, workplaces and the wider environment and exploit personal data generated through those platforms for their own gain rather than for social good. In the long term, robotics and intelligent sensor grids may come to support people closely at all stages of their lives from baby care, through child minding and school, to their working lives, when health care is needed and beyond into retirement and elderly care. The current trends erode privacy by "mining" personal and social data for commercial use and sale to third parties. Ownership over such data is frequently asserted by global corporations as an imposed condition of access to the most basic and widely deployed services on the Internet. Consideration needs to be given as to whether this should extend to robotics and their supporting sensor appliances. The means for wider personal data ownership, privacy, necessary and proportionate data sharing, data controls and ethics must be carefully considered and will probably require legislation to prevent abuse.

RAI Ethics and Legal Issues: The UK government stance on ethics for RAI platforms and research so far has set a good example. The UK is well placed to use its RAI aware scientific and engineering community and institution members to oversee ethics issues that will arise. There are many ethical issues which arise from the deployment of RAI, sometimes obvious ones such as in the care of the

young, elderly or mentally incapacitated, which are also present when humans perform such functions, but also new ones which will arise (e.g. if an autonomous car realises it can't avoid an accident, how should it mediate between hitting different things?). The question of legal responsibility also arises which is not straightforward: who is liable: the owner of the robot, its user, the designers, the standards bodies, the verification body, … ?