The Royal Academy of Engineering

Autonomous Systems:
Social, Legal and Ethical Issues

The Royal Academy of Engineering

Autonomous Systems:
Social, Legal and Ethical Issues

ISBN 1-903496-48-9

August 2009

Published by
The Royal Academy of Engineering
3 Carlton House Terrace
London
SW1Y 5DG
Tel 020 7766 0600
Fax 020 7930 1549

Copies of this report are available online at:
www.raeng.org.uk/autonomoussystems

Registered Charity Number: 293074

Contents

1. Introduction

This is a report of the discussion at a roundtable meeting held at The Royal Academy of Engineering on the social, legal and ethical issues surrounding the development and use of autonomous systems. The meeting involved stakeholders from a range of areas, including medicine and healthcare, transport, defence, systems engineering, computer science, financial systems, public engagement and policy development. A full list of participants and other contributors is given in 5.1.

Autonomous systems are likely to emerge in a number of areas over the coming decades. From unmanned vehicles and robots on the battlefield, to autonomous robotic surgery devices, applications for technologies that can operate without human control, learn as they function and ostensively make decisions, are growing. These technologies can promise great benefits, replacing humans in tasks that are mundane, dangerous and dirty, or detailed and precise. They also have potential in allowing the remote performance of various functions, from defusing bombs to monitoring the ill or housebound.

However such technologies raise a number of social, legal and ethical issues, just some of which will be explored in this short document. The focus will be on two emerging areas of technology – transport, in terms of autonomous road vehicles; and personal care and support, in the form of artificial companions and smart homes. These examples are at different stages of development and will have quite different impacts on people's lives, but are both areas where autonomous systems will be quite visible to the general public.

Two main kinds of issues arise in relation to these classes of technology. First, are autonomous systems different from other complex controlled systems? Should they be regarded either as 'robotic people' – in which case they might be blamed for faults; or machines – in which case accidents would be just like accidents due to other kinds of mechanical failure. Secondly, autonomous systems allow for detailed recording of usage. Currently, road accidents, even fatal ones, attract only cursory investigation compared with air or rail accidents. But as recording quality improves and costs go down will all accidents be carefully analysed? This raises legal and privacy issues: what would happen if most road accidents, currently insurance-classified as 'accidental', were reliably blame-assigned? And how would we feel about our movements in our automated apartments being recorded and kept by a third party?

Early attention to the issues raised by the introduction of these technologies is important to ensure that their introduction has the public interest in mind and has appropriate support. Public engagement efforts on autonomous systems and debate about their impact is valuable in order to understand and address the genuine expectations and concerns of public stakeholders, and to take these into account in the development and implementation of the technologies. There are significant benefits to be gained from the development of these technologies.

1.1 What is an autonomous system?: The spectrum from user controlled, to automated to autonomous systems

Automated or automatic systems are commonplace, and cause little concern or interest. Lifts work automatically without human controllers (though they did have operators in their early days). What makes the difference between an automated system and an autonomous system?

There is a range of levels of control or involvement that human operators can have in a system. The following sets out the different grades of control:

- controlled systems: where humans have full or partial control, such as an ordinary car
- supervised systems: which do what an operator has instructed, such as a programmed lathe or other industrial machinery
- automatic systems: that carry out fixed functions without the intervention of an operator, such as an elevator
- autonomous systems that are adaptive, learn and can make 'decisions'.

These levels of control lie on a continuum from controlled to fully autonomous systems with no fixed distinction between them. The spectrum can be correlated with a proportional lessening of the degree of human intervention or interaction. The degree of autonomy employed needs to be appropriate to the task.

For many of the systems that lie on this spectrum there is a human agent with direct responsibility for how they function, whether that be the operator of the system or the designer of an automatic system. However, with fully autonomous systems it is more difficult to discern where the human responsibility lies for its functioning or malfunctioning. There will always be humans in the chain, but it is unclear in the case of injury which human in the chain bears responsibility – the designer, manufacturer, programmer, or user?

Do we expect autonomous systems to simulate humans, learning from experience and making decisions? Does the famous 'Turing Test' for a thinking machine apply here? In some circumstances, an autonomous system may be more predictable and reliable than a human. An autonomous system that 'makes decisions' will to a large extent be determined by input and past operation, and would not be 'impetuous' in the way that a human could be. This is what makes them so potentially valuable for tasks that require quick decisions in potentially dangerous circumstances. However, in highly complex situations, the breadth of human experience could give rise to better judgements than those made by a system 'programmed' by a narrow range of previous behaviour.

1.2 The ethical, legal and social implications of autonomous systems

The very nature of autonomous systems raises social, legal and ethical issues. People tend to be more accepting of a technology if they can choose whether or not to adopt it and have some control over its use. Public perception of risk is very sensitive to the distinction between imposed risk and risks that individuals feel they can control. However, autonomous systems are inherently outside of users' or the public's control, and therefore it is possible that people's reactions to such technologies will be negative. Attitudes to autonomous systems are likely to vary according to the area of application, and it is not necessarily the case that acceptance of autonomous systems used in dangerous military settings transfers to the use of autonomous systems in every day life.

Autonomous systems require people to give up some of their own choice and agency. Are people happy to do this in certain contexts, or is this inherently objectionable? This question will be of particular relevance in relation to autonomous road vehicles, as discussed below.

Some general ethical questions can be asked about the acceptability of autonomous systems. One question is: should systems be designed that cannot be shut down by a human operator? It might be thought that there is always need for human intervention, but sometimes autonomous systems are needed where humans might make bad choices as a result of panic – especially in stressful situations – and therefore the human override would be problematic. Human operators are not always right nor do they always have the best intentions. Could autonomous systems be trusted more than human operators in some situations?

All technologies are liable to failure, and autonomous systems will be no exception (which is pertinent to the issue of whether autonomous systems should ever be created without manual override). Dealing with the outcomes of such failures will raise legal issues. If a person is killed by an autonomous system – say, an unmanned vehicle – who is responsible for that death? Does the law require that someone be held responsible? Legal and regulatory models based on systems with human operators may not transfer well to the governance of autonomous systems. In addition, the law currently distinguishes between human operators and technical systems and requires a human agent to be responsible for an automated or autonomous system. However, technologies which are used to extend human capabilities or compensate for cognitive or motor impairment may give rise to hybrid agents – humans with autonomous prostheses which support their physical or cognitive functioning. Without a legal framework for autonomous technologies, there is a risk that such essentially human agents could not be held legally responsible for their actions – so who should be responsible?

The concern with potential failures of autonomous systems could mean that such technologies are held back until they are believed perfect. But is this too strong a requirement? Autonomous systems tend to be compared with the paradigm of perfection, and not with the systems that they will replace. Reaction to failures of autonomous systems is somewhat different to reaction to failures of manned systems. For example, a failure of a surgical robot that has

genuinely learnt from past use cannot be attributed to any one person. Is it therefore the case that no one can be held responsible? This has been a barrier to introducing adaptive surgical robots since in the health sector it is usually the surgeon who has responsibility when things go wrong during a procedure. As well as the concern about the legal framework covering autonomous systems, without an understanding of responsibility for the functioning of autonomous systems and a formal process for their certification, it may prove difficult to secure insurance for autonomous systems. These issues must be addressed, as it is almost inevitable that autonomous systems will experience failures, but if those failures are less frequent than for human-operated systems it would be counterproductive if they were a barrier to their development and use.

These kinds of social, legal and ethical questions need to be debated so that autonomous systems can be developed in a way that is legally and socially acceptable. Some general questions for debate are:

What are the consequences of giving up human choice and independence?

When is an autonomous system good enough? Must there be higher standards for autonomous systems than those with human operators? How much higher?

Can the law keep up with technology? How can the law be changed to accommodate the range of autonomous systems?

What are the criteria for assigning responsibility, or degree of responsibility, for the failure of an autonomous system and the harm it may cause?

Who will be responsible for certification of autonomous systems?

2. Autonomous road vehicles

2.1 Technologies – from GPS and car-to-car communication to centrally controlled autonomous highways

There is a spectrum of technologies between driver-operated vehicles and autonomous road vehicles. The following sets out the likely timeline for the introduction of autonomous vehicles:

Driver information systems, including driver warning systems
These provide warnings and alerts to drivers and leave drivers with a choice as to how to act on those warnings

↓

Advanced driver assistance
Some automation of responses to warnings
Self-parking systems

↓

Cooperative vehicle highway systems
Allows information to be relayed between vehicles as well as from operators to vehicles, for instance when cars are too close together, to prevent collisions

↓

Automated highway
The ultimate vision, with the highways and the vehicles within them being managed and controlled as a system, instead of individual drivers making decisions about speeds, stopping distances, routes and so on

2.2 Timescales and transformation – what can this technology offer? When will it appear?

Automated and autonomous controls in vehicles can help to prevent human error as a result of tiredness and lack of concentration. Road accidents are more likely to be due to a driver not reacting appropriately or quickly enough than as a result of vehicle failure. With automated vehicles that can select their own routes, congestion can be controlled and driving made more efficient, leading to reductions in pollution.

Robotic vehicles are technologically closer than people might think. The DARPA grand challenge in the US, a competition to build a driverless car, has been held three times since 2004. A number of vehicles passed the challenges of navigating complex courses over long distances within a set period of time, with no driver control. Such vehicles are still some way from the mainstream, but there are vehicles available with intelligent sensor driven cruise control. Vehicles with a greater degree of autonomy are technologically possible, but their development is hampered by the lack of a legal framework for them – at present, they would be impossible to insure.

While completely driverless cars may not be a significant proportion of the traffic for some time, convoys of heavy trucks that are controlled like aircraft are potentially nearer the horizon. So, while the ultimate vision of the fully automated, driverless highway is still in the distance, the steps for getting there are being addressed. A great deal of the technology needed is already available, and under test. A model currently being investigated is that of long lines of regulated vehicles over which the individual drivers have limited control. The technology to achieve this is potentially close to market, however, it is of limited value unless there wide adoption. Highway operators need to be

aware of when they are likely to reach mass market, so that systems can be implemented to make the most of them. However, there is a chicken-and-egg relationship between the uptake of such vehicles and the implementation of a centralised highway-control system.

However, a top-down model of an automated highway, with ultimate control of vehicles left to a centralised control system, may not be the only option. It might be that the move to such a system is prompted by individual drivers seeing the benefit of collision avoidance and driver warning systems, so that the system emerges from the use of such devices in individuals' cars. For example, GPS systems in cars became popular very quickly and if driver warning systems and similar devices are marketed in a similar way they could be taken up very swiftly. The possibility of this is a reason why Government has chosen not to legislate but to let the market dictate the growth of the automated highway.

The latter approach would mean that the automated highway is achieved by degrees. A first step could be technologies for allowing vehicles to communicate with each other. Once such systems are in the majority of cars then Government can take control and demand those vehicles comply with set standards and can form a unified system.

But the order in which technologies will appear is not fixed. More important for Government is the timescale in which they will appear and the point at which changes need to be made in the way that the highway system is managed. This point may not be so far away with the prevalence of GPS and the growth of car-to-car communication. Therefore, the time to establish a common approach to vehicle automation may be near. This would be by a process of 'smoothing' the system – perhaps starting with control of speed since there is a good case for this.

2.3 Barriers: ethical, legal and social – administrative structures, enforced compliance and conformity

It is worth considering the steps that have already been made toward autonomous vehicles. Most vehicles are controlled to a significant extent by engine management systems, drivers are reminded to close doors and put on seat belts, and many people give over responsibility for their route to GPS. Driving is a quite different experience now to what it was 20 or 30 years ago. So, are we just separated from autonomous vehicles by a matter of degree, and is the acceptance of such vehicles therefore likely?

There are a number of ethical, legal and social issues surrounding autonomous vehicles. First, there is the potential for exclusion of those who do not want to be part of the system. With an automated highway with centralised control, some routes will simply be shut to those who do not buy in. Is that fair? Even with co-operative control, drivers in vehicles without such systems are likely to be treated quite differently, especially by insurers or in the case of an accident where culpability is under question.

This raises the issue of the impact of failure of such a system. How will accidents be dealt with, and how will insurance models have to be changed? Will such a system be able to deal with different vehicles that do not all work perfectly, or that host different versions of the technology? The technology needs to be kept up to date for the system to work well, which is a challenge when it takes up to 10 years to turn the fleet around.

Regulatory models will also struggle to fit this system, as can be seen from the analogous case of automatic trains. Automated trains are most easy to implement on metros with only one operator. They become more complex in situations as in the UK where the infrastructure is owned by a different company to the freight owners and there are a number of different operators. This suggests that the situation for road transport with individually owned cars will be even more challenging, as individual drivers are likely to be reluctant to give up their autonomy.

Governments will have to play a bigger part in controlling the highways system if autonomous vehicles are to be introduced. Currently car manufacturers play an important role and Government does not yet feel ready to dictate the market – the market is currently the dominant force. Government cannot take ultimate control whilst it remains a relatively open system – how and when will the change to a regulated system occur? Another obstacle is the fact that there would have to be retrospective legislation for older cars, or the result will be exclusion of their drivers from the roads.

In summary, the main issues relating to autonomous road vehicles are:

Will the development and uptake of autonomous vehicles marginalise road users in older vehicles?

How can the autonomous highway be regulated? Who manages the road system, and how can autonomous vehicles from a variety of manufacturers and with individual owners be regulated?

How will the insurance industry deal with responsibility for failures and accidents involving autonomous vehicles?

2.4 Recommended actions – for engagement and policy development

The area of autonomous road vehicles is particularly interesting because, assuming a top-down model of highway control, it involves so many people at once ceding control to an external operator. However, is this so different to the case of passengers relinquishing control when they travel on trains or planes? Will car drivers be reluctant to give away control in this way or will they see it as a convenience? Would drivers have a different attitude to a centralised system or a cooperative system that emerges out of the technologies that drivers choose to have in their cars, so that each individual simply gives over control to their own vehicle?

Public engagement on this issue seems to be needed in order to inform government policy on this issue, especially before infrastructure is invested in. The use of autonomous vehicles links up to and raises a lot of other issues about road user charging, since such systems are likely to be put in place alongside charging systems. This is an area where public feeling runs high and hence there is value in engaging with the public to build a model that will be considered acceptable.

Autonomous road vehicles, whether controlled centrally or cooperatively, are also likely to generate a lot of data about where a car has been, how it was driven and so on. Therefore, there are potential ethical issues surrounding the

privacy and potential use of such data. Drivers' attitudes to this data gathering are difficult to judge – GPS is popular despite the fact that it involves tracking a car and could be used to record journeys. But are drivers aware of the potential privacy implications of such technology? Debate and engagement on this issue would be of great value.

In order to avoid stalling a technology that could be of significant benefit, Government should engage in early consideration of regulatory policy so that such systems can be introduced. There is also need to consider changes needed to legal frameworks that establish culpability for accidents on the road. When a driver cedes control of their car to its own navigating or driving systems, or to a centralised system controlling speed and distance between cars, they also give away some responsibility. So where would that responsibility rest in the case of an accident? Who will regulate the technology employed and how will it be certified? Will insurance companies accept the risk?

3. Artificial companions and smart homes

3.1 Technologies – from blood pressure monitoring to Second Life

The kinds of technologies under consideration here fall into the broad range of technologies for monitoring or providing companionship to human users. In terms of systems for monitoring users, these are likely to include smart homes used for older or ill people, which would monitor their movements and which could monitor items of food or medicine taken from the fridge and consumed. These could be used in order to send out alerts if the occupant had not got out of bed, had fallen and was not moving, or to automatically order food and other items that were depleted. Such systems would be autonomous in that they learned from behaviour patterns and identified unusual movements and actions. Other monitoring systems are those within the area of telemedicine, where patients' health could be monitored remotely, with checks on blood pressure or other physical symptoms recorded and relayed to healthcare workers where necessary.

In terms of companionship, autonomous systems are likely to have application in the forms of robot babysitters, artificial therapists and social or even sexual companions. There is a great deal of work being carried out on designing systems that can convey emotional expressions – or behaviours that look like the expression of emotions – and with the ability to pick up on emotions. It has been shown that automated teaching systems are improved by the ability to monitor the emotional reactions of students. Such systems could be used to provide companionship for those who are isolated, and could perform useful functions, such as reminding owners to take medicines or to perform tasks. In Japan, robot babysitters have been used to keep young children amused in shopping centres while their parents shop. There are also dispersed systems like the multi-user Second Life, where people have 'avatars' (digital alter egos) and interact with others. These systems have emergent autonomy through the behaviour of users and give companionship to users. Though there is a human user behind each avatar, the system is autonomous in so far as any society or community is.

3.2 Timescales and transformation – what can this technology offer? When will it appear?

There has been a significant amount of research in this area and these technologies are close to the horizon. The technologies for smart homes and patient monitoring are already here; indeed, these technologies have crept up on us. They offer a number of benefits to older or incapacitated people:

- enable independence, people feel less of a 'burden' and have their own autonomy restored
- health can be assessed remotely, without need to travel to appointments unless necessary
- quick action can be taken in case of a fall, an older person will not be left waiting for hours or even days
- fridge contents monitoring can ensure that inhabitants do not go hungry if they are unable to travel to shops
- appliances and lights can be switched on and off for safety, or voice prompts can remind inhabitants to switch devices off
- 24 hour automated support and advice could be made available to inhabitants.

As mentioned, robot babysitters are already used in Japan, and dispersed artificial systems like Second Life have millions of users. In the UK NICE have approved an online 'Beat the Blues' counselling programme, which is in effect an artificial therapist or counsellor, which responds to a user's answers as a therapist would. Artificial companions that look very like humans may be some way in the future, but the early versions of such technologies are here.

Autonomous systems are not restricted to systems that behave like a human, but can also include systems that are embedded inside humans. Cognitive prosthetics for improving memory or other brain functions are in development, for those who have suffered brain injury or less of cognitive function through degeneration.

3.3 Barriers: ethical, legal and social

These technologies are far closer to the horizon than those for autonomous vehicles yet there has been very little debate concerning their acceptability or the ethical consequences of their use. One of the most obvious potential drawbacks to such systems is that they could lead to the social isolation of users. An older person in a smart home might be visited less often by family, neighbours and health workers and this could lead to loneliness. Conversely, creating a smart apartment can restore a person's autonomy and allow them to remain in their own home and neighbourhood, which could prevent loneliness.

There has to be engagement on this issue, to establish the value of such technologies for their intended users. This must involve finding out older people's attitudes to being on their own. Would they feel that they had more independence and freedom in such a setting, and would it free them from routine chores and tiring trips away from the home? Or do older people sometimes want to give up their independence if it leaves them isolated? Obviously, individuals will vary, but there should be engagement on this issue to establish the preferences and views of older people, not just making assumptions about their reasons for living alone, or not.

Assuming that there is a call for such ways of living, issues arise about the way that such homes and monitoring systems will work. The ethical issues surrounding the use of these technologies have not been recognised or debated prior to their implementation. Some of the potential ethical issues surrounding automated apartments and artificial companions are:

- Large amounts of personal data are generated by patient monitoring systems yet there has been no discussion about proper use of those data.
- One corporation may retain control of the technology and may not take ethical considerations seriously.
- 'Tricks' can make systems seem like artificial companions that have emotions or intelligence and these can manipulate vulnerable people, who may even believe they are communicating with another person (especially when familiar voices are used to record prompts).
- Vulnerable users may not understand what the technologies they use can and cannot do - most ethical codes for computer development assume that users are knowledgeable and sophisticated in their use of computer technologies

When it comes to technologies intended to assist people who may have cognitive impairments, consent to use of or research into technologies becomes complex. Do the research subjects understand what the technologies are being used for? And do they feel able to withhold consent, especially when asked by a doctor or someone perceived to be in authority?

It may not be only vulnerable users who are unclear about what they consent to when using these technologies. Public perceptions in this area are somewhat warped by the influence of science fiction. The fact that it is easier to get a person to act like a robot than it is to make a robot that can act like a person means that people have false ideas of the timescales and potential for autonomous companions and robotics in general.

In summary, the main issues relating to smart homes are:

Is the cost of the autonomy afforded by smart homes social isolation? How can the balance between independence and companionship be struck?

Will vulnerable people understand the nature of the technologies in smart homes, and can we be sure that they are able to consent to their use?

Who should control the data generated by systems that monitor people's movements and track their physical wellbeing?

3.4 Recommended actions – for engagement and ethical debate

There should be engagement with older people on the use of technologies for allowing people to be monitored in their homes. Are they broadly welcomed, or is it likely they will lead to people feeling more abandoned, isolated and vulnerable than they currently do? What can be done to make these technologies work for users who might not routinely use computers? Are attitudes likely to change with the generations, so that people coming to retirement now and used to working in technology-driven environments may be more comfortable, so that by the time such technologies become mature they are also more accepted? Could an artificial companion ever be seen as offering real companionship – would it be any better or worse than a pet, for instance?

The legal framework considering the use and ownership of data generated by such technologies must be considered to ensure that the data cannot be misused. Who decides what happens to the data:

- The subject of the data?
- The doctor in the case of systems that monitor health?
- The companies that run and maintain the technologies in the case of smart apartments that monitor food consumption and other behaviour?

There needs to be an awareness of the whole range of users when such technologies are developed. Programmers who are familiar with them may not understand the needs and concerns of those who are unfamiliar, or who have cognitive impairments that affect the way that they use and interact with them. Smart homes and related technologies will be of no benefit if they are not accessible to, and usable by their intended users.

4. Conclusions

4.1 Communication and engagement

There is need to think about the reality of these technologies, both positive and negative, as they are areas that are likely to be loaded with assumptions. Young people's perceptions of AI and robotics are likely to be influenced by the contexts in which they encounter them. Most young people only encounter such things in computer games where the goal is to destroy them! It is also important to find out older people's perceptions, since intelligent or autonomous systems will have application in care for older people.

People tend to have very fixed views about the kinds of robots that an engineer designs and the uses of those robots. The assumption is that robots are always designed to have some physical use. However, robots can also be designed to work as thought experiments, to give us a better understanding of how real people function and behave. Robots can be designed to give us a better understanding of the world – all of the different manifestations of AI and robotics should be explored.

When we gather the views of the wider public on autonomous systems and robotics, what are we to do with the opinions gathered? Should they be used to guide the direction of engineering? And should certain avenues of research be abandoned because there is significant objection to the idea of that research – or is technology push sometimes the right thing? After all, there were objections to the idea of heart transplants when this was a new technology but it is now an accepted and valued area of medical practice.

The value of public engagement is seeing what specific areas of research cause people concern and either these areas of research can be scrutinised more closely, or public attitudes can be addressed through communication about the research. There must also be debate and consideration about the extent to which public views should shape the path of autonomous technologies. Public opinion should not be the single criterion that directs research funding, for example. Nor should the precautionary principle be applied too liberally, with a moratorium on any area of research where they may be evidence of concern. Indeed, it is not usual for technology to be halted in the face of public concern. Many people objected to the railways and railway engines – especially because they were dangerous and killed people! But technologies like this are allowed to flourish nevertheless because of the advantages they bring. Often the public will have negative views about a technology whilst still broadly accepting them because of the advantages they bring – mobile phones, for example – particularly if it is a technology that they choose to use rather than a technology that is imposed on them.

4.2 Regulation and governance

Autonomous and even driverless vehicles are close to being a technological possibility, however in countries like the UK political structures make it difficult for them to be implemented. This involves issues concerning ownership of the system as opposed to its components and handling of risk in this quite different environment. It is important to consider that autonomous vehicles may not experience a tipping point where there is a step change in the kinds

of vehicles used, as there has in fact been an incremental reduction in driver control for cars. Therefore, regulation for these kinds of technologies has to be put in place in good time. These issues are being addressed for unmanned aircraft and they will provide some learning for the more complex ground vehicle application.

4.3 Ethical considerations

Smart homes are close to the horizon and could be of significant benefit. However, they are being developed largely without ethical research. This means that there is a danger of bad design, with assumptions about users and their behaviour embedded in programming. It is important that ethical issues are not left for programmers to decide – either implicitly or explicitly.

4.4 Looking for applications

Smart homes could be developed which control temperatures and switch appliances off and on as needed – leading to greater efficiency in the home. If these technologies could be retrofitted into existing housing, then this could be very beneficial. The wider applications of these technologies should be explored, so that they can have the maximum benefit.

4.5 The wider landscape

Autonomous systems are likely to appear in a much wider range of technology areas than discussed here – such as defence, policing, medicine, decision making systems (eg, financial services, e-commmerce etc). Technologies in all of these areas should be developed in a way that gives consideration to their potential benefits and downfalls. It is important that a lack of engagement on the wider issues does not halt the development of the technology. For example, there are significant advances on the horizon in robotic surgery. Trials will begin in Europe next year on a system in which the surgeon identifies the target area for treatment and the machine decides the best route to access the area. Issues arise regarding how to verify the system and how to assess the rules for the machine's decision making. But there is great promise that such technologies can be more accurate and less damaging than surgical procedures determined by a surgeon. These issues should be addressed so that the benefits of the technologies can be reaped.

5. Appendices

5.1 Working group and acknowledgements

This report is based on discussion at a roundtable meeting, and on presentations by Chris Elliott and William Bardo; Lesley Paterson; Ginny Clarke and Blay Whitby.

Roundtable participants:

Professor Igor Aleksander FREng, Emeritus Professor Imperial College London
Rob Alexander, Department of Computer Science, the University of York
Professor William Bardo FREng, Systems Engineering for Autonomous Systems Defence Technology Centre
Ginny Clarke, the Highways Agency
Lambert Dopping-Hepenstal FREng, BAE Systems
Professor Martin Earwicker FREng, Vice-Chancellor, London Southbank University
Dr Patrick Finley, Medimatron
Professor Peter Gore, Institute of Ageing and Health, University of Newcastle
Professor Roger Kemp FREng, Engineering Department, University of Lancaster
Dr Lesley Paterson, The Royal Academy of Engineering
Dr Mike Steeden, Defence Science Technology Laboratory
Professor Will Stewart FREng
Dr Alan Walker, The Royal Academy of Engineering
Dr Blay Whitby, The University of Sussex
Michael Wong, IBM

Contributors:

Dr Chris Elliott FREng, Pitchill Consulting
Professor Noel Sharkey, University of Sheffield

The discussion meeting and report were commissioned by The Royal Academy of Engineering's Engineering Ethics Working Group:

Professor Richard Bowen FREng, University of Wales, Swansea
Professor Brian Davies FREng, Emeritus Professor, Imperial College London
Lambert Dopping-Hepenstal FREng, BAE Systems
Wendy Harrison, Shell International
Bob Harvey, the British Computer Society
Philippa Foster-Back OBE, Institute for Business Ethics
Gordon Masterton FREng, Jacobs Babtie Group
Richard Maudslay CBE FREng (Chair), National Nuclear Laboratory and Hardy & Greys Ltd
Professor Christopher Megone, University of Leeds
Andrew Ramsay, Engineering Council UK
Professor Ernest Shannon FREng
Dr Alastair Soane, CROSS
Professor John Uff CBE QC FREng, King's College London
Stephen Vranch FREng, Jacobs Engineering

5.2 Statement of Ethical Principles

The Royal Academy of Engineering and the Engineering Council UK's Statement of Ethical Principles (SEP) for engineers is available online here: **http://www.raeng.org.uk/societygov/engineeringethics/pdf/Statement_of_Ethical_Principles.pdf**

The SEP sets out a number of duties for engineers. In the area of emergent technologies like autonomous systems, there is a particular duty of engineers to:

- identify and evaluate and, where possible, quantify risks.
- be aware of the issues that engineering and technology raise for society, and listen to the aspirations and concerns of others.
- actively promote public awareness and understanding of the impact and benefits of engineering achievements.

This report was prepared by Dr Natasha McCarthy

The Royal Academy of Engineering

As Britain's national academy for engineering, we bring together the country's most eminent engineers from all disciplines to promote excellence in the science, art and practice of engineering. Our strategic priorities are to enhance the UK's engineering capabilities, to celebrate excellence and inspire the next generation, and to lead debate by guiding informed thinking and influencing public policy.

The Academy's work programmes are driven by three strategic priorities, each of which provides a key contribution to a strong and vibrant engineering sector and to the health and wealth of society.

Enhancing national capabilities

As a priority, we encourage, support and facilitate links between academia and industry. Through targeted national and international programmes, we enhance – and reflect abroad – the UK's performance in the application of science, technology transfer, and the promotion and exploitation of innovation. We support high quality engineering research, encourage an interdisciplinary ethos, facilitate international exchange and provide a means of determining and disseminating best practice. In particular, our activities focus on complex and multidisciplinary areas of rapid development.

Recognising excellence and inspiring the next generation

Excellence breeds excellence. We celebrate engineering excellence and use it to inspire, support and challenge tomorrow's engineering leaders. We focus our initiatives to develop excellence and, through creative and collaborative activity, we demonstrate to the young, and those who influence them, the relevance of engineering to society.

Leading debate

Using the leadership and expertise of our Fellowship, we guide informed thinking, influence public policy making, provide a forum for the mutual exchange of ideas, and pursue effective engagement with society on matters within our competence. The Academy advocates progressive, forward-looking solutions based on impartial advice and quality foundations, and works to enhance appreciation of the positive role of engineering and its contribution to the economic strength of the nation.

The Royal Academy of Engineering promotes excellence in the science, art and practice of engineering.

Registered charity number 293074

The Royal Academy of Engineering
3 Carlton House Terrace, London SW1Y 5DG

Tel: 020 7766 0600 Fax: 020 7930 1549
www.raeng.org.uk