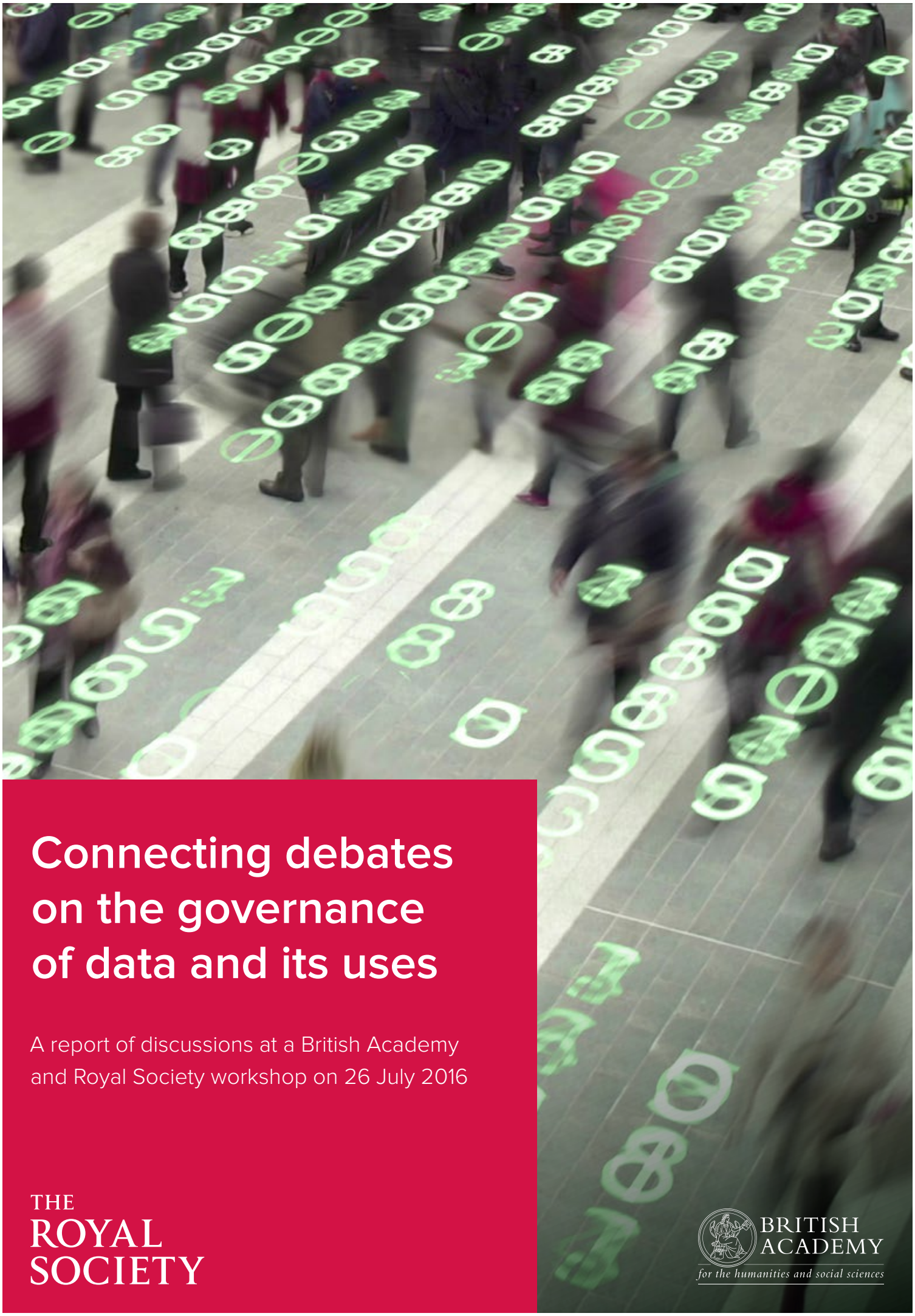

Connecting debates on the governance of data and its uses

A report of discussions at a British Academy and Royal Society workshop on 26 July 2016

THE ROYAL SOCIETY

BRITISH ACADEMY
for the humanities and social sciences

Contents

A: Introduction

In July 2016 the Royal Society and The British Academy hosted a seminar which brought existing sectoral and disciplinary debates on the governance of data and its uses. The seminar gathered leading representatives from academia, government and business; including experts in ethics, law, finance, social and data sciences, machine learning and statistics in order to build connections between existing debates, and identify key questions and gaps.

This seminar report provides a summary of discussions, together with the discussion papers written by a range of our attendees.

Building on the discussion of the seminar, the Academies have initiated a project examining new uses of data and their implications, and reviewing the data governance landscape. This project seeks to make recommendations for cross-sectoral governance arrangements that can ensure the UK remains a world leader in this area.

Terms of reference

- Identify the communities with interests in the governance of data and its uses, but which may be considering these issues in different contexts and with varied aims and assumptions, in order to facilitate dialogue between these communities. These include academia, industry and the public sector.
- Clarify where there are connections between different debates, identifying shared issues and common questions, and help to develop a common framework and shared language for debate.
- Identify which social, ethical and governance challenges arise in the context of developments in data use.
- Set out the public interests at stake in governance of data and its uses, and the relationships between them, and how the principles of responsible research and innovation (RRI) apply in the context of data use.
- Make proposals for the UK to establish a sustained and flexible platform for debating issues of data governance, developing consensus about future legal and technical frameworks, and ensuring that learning and good practice spreads as fast as possible

Further details of the project can be found on the webpages of the British Academy and the Royal Society.

B: Summary note of event

A report of discussions at a British Academy and Royal Society workshop on 26 July 2016.

Disclaimer: This is a note summarising the discussion and debate at the British Academy and Royal Society's workshop on *connecting debates on the governance of data and its uses*. It is not intended to represent the views of either the British Academy or Royal Society, nor does it represent the views of individual attendees at the event.

1. Changing debates

It is no longer just about the data – it is the use

The increasing quantities of data generated is changing the nature of debates. The scale of data collection is stimulating innovative uses of data, and the potential to link data sets and process data in new ways means that it no longer make sense to talk about data in a context disconnected from how it is used. When developing governance frameworks, consideration of data cannot take place without deliberation of how it is processed, and vice-versa. This has impact on the focus and framing of debates about data governance.

For example, in a networked world with ever more sophisticated data analytics tools, the distinction between sensitive and non-sensitive data is blurred: innocuous data can become sensitive when combined with other sets of data or when put to controversial uses.

The significance of data also depends on the wider context in which it is used. For example, the use of health data in a doctor-patient relationship may be governed by norms of confidentiality, but these norms exist in the context of professional relationship between doctor and patient. The many and varied uses of new health and fitness applications generate health data that is not interpreted and used by a medical professional – this changes the nature of our concerns about health data and how it is managed.

What is consent in a changing technology context?

In a dynamic personal data ecosystem, consent becomes problematic as a mechanism to secure legitimate use. It is context dependant and there is no one-size fits all.

Consent is a complex concept, encapsulating informed consent, specific consent, broad consent and dynamic consent. None of these, in turn, are straightforward: is it practical or even possible to have informed consent, when we cannot be sure of the uses data will be put to in the future? Do we understand the consequences of the choice we make when we click 'I accept' at the bottom of a terms and conditions page? Can we expect everyone to understand the complex science behind research using health data? Specific consent may be intended to protect against unforeseen uses, but if it is given do we limit the possibilities offered by data analytics or create burdensome bureaucracy? Conversely, if we give broad consent, have we signed a blank cheque that puts our personal information at risk in the future?

The nature of data itself can make the idea of consent challenging. In the case of genetic data, for example, an individual may be able to consent to release their own genetic information and genome sequence, but this would also amount to releasing details of their family members' genome into the public domain too. Can such data therefore be straightforwardly 'owned' by a single individual who has sole right to consent to its use?

Privacy – does it still exist?

In this new context where concerns about data are implicitly connected with the use of that data, the concept of privacy needs reconsideration for the twenty-first century. The protection of privacy may be considered essential in keeping with human rights law, data protection law and the common law requirements of confidentiality. However, this notion of privacy might be better understood in other terms – e.g. as 'intimacy' or 'secrecy' – and the relation of this concept to how data is used by evolving technologies stands in need of scrutiny.

The link between privacy and physical space is significant – the concept of privacy often refers to the idea of a space where an individual or group can be alone and unmonitored. However, physical spaces that we might have considered private are now places where data might be gathered about us: there are increasing numbers of robots in the home with the capacity to gather data about activities within the home; toys that record and relay to manufacturers what children say in the course of play; and autonomous self-driving cars that will become data gathering devices. These raise issues about the ability to protect certain spaces as private, and also the use, ownership and control of the data generated in this hitherto 'private' domain.

Privacy is often linked to the need to protect identity or identifying information. However, many services can be delivered without needing to know an individual's personal identity. If a person is not a single construct but should be thought of as different personas, can this be embedded in how we share data and deliver services, so that only relevant, historical data is used?

Not just about privacy, but about freedom and autonomy

Technological developments in the use of data mean that privacy (and protection of identity) may not be the defining ethical or governance concern in relation to data. In the context of learning algorithms that make predictions about future behaviour of people and systems, concepts of freedom and autonomy come to the fore. Issues such as 'autonomy', 'self-determination' and 'opportunity' need to be aired in discussions about governance, taking the debate beyond the idea of privacy as narrowly defined. Future concerns will likely relate to the freedom and capacity to create conditions in which we can flourish as individuals; governance will determine the social, political, legal and moral infrastructure that gives each person a sphere of protection through which they can explore who they are, with whom they want to relate and how they want to understand themselves, free from intrusion or limitation of choice.

2. Governance across the data lifecycle

Through the application and use of data, a series of challenges arise, various actors play a part at different points, and there are different stages where governance becomes important. The following issues, concerning the quality of data at different points in its lifecycle, have consequences for governance.

Data integrity

Data integrity, veracity and data provenance are of critical importance. The role of generation of data needs to be recognised as essential to avoid the risk of using inappropriate data to make decisions. This is particularly true when bringing together data from various source of differing quality.

Bias in data

Bias, or even intentional obfuscation, can be embedded into datasets. If this is not considered at the outset, it may carry through to data processing, becoming reinforced.

Accidental collection

We are not only seeing an increase of collection of online data. Other methods of collection such as CCTV footage and the use of drones in surveillance, lead to accidental or 'collateral' collection of data. Use of images, audio and video may need specific consideration in the governance of data.

Crossing sectors

Norms can be attached to particular sectors, which means that when data crosses sectors – e.g. from public to private or from health to finance – the different norms and practices in different sectors may not be respected in the new context of use.

Statistical profiling and stereotyping

With applications such as machine learning that make inferences from data, one of the challenges is when disadvantageous inferences are made about an individual. The level of harm may be low – e.g. recommendations on Amazon – but could be problematic when applied in other contexts where the stakes are higher such as determining prices people are willing to pay for services, and more significant still in other spheres.

Transparency

Decisions can be made that affect an individual, but using criteria that is unclear and opaque. Machine learning technologies have been used in the US to predict whether a convicted criminal will re-offend or not. In this type of context, an understanding of how decisions are made becomes critical. But is algorithmic transparency feasible? Can algorithms be audited when, for example, they keep learning and changing as new data come along?

Accountability

Given the increasing recognition of the potential for adverse impact, there is a need to find governance mechanisms that will secure algorithmic accountability. Some legal aspects have been untested. If someone gets hit by a driverless car, who is at fault? Who gets sued? If artificial intelligence is used in Government decision making, where accountability is usually seen as important, where does accountability rest? For example, if algorithmic systems are used to make border entry decisions, is it the Home Secretary who is ultimately accountable for those decisions, especially when they are considered wrong?

Impact

These technologies are being used right now, but there is not enough information about what is happening and what the impact may be. Therefore, there is a real need for better understanding of impact of new uses of data on individuals, groups and society.

3. Social and ethical dimensions of the debate

In order to understand the requirements of data governance, we need to understand what society values. However, society, and groups and individuals within it, can often value different things which are not always mutually attainable. Understanding the range of principles at stake in data use, their complexity, and how they relate to each other, is central to enabling data governance.

Balancing risks and benefits across individuals, groups and society

There are values and rights held on an individual level and at a public level. These are sometimes similar and sometimes in tension. For example, keeping something private might be an individual preference, but by doing so you may not be contributing to the public good – health data, which is personally private but has societal benefit when used for research, is an example. Managing the balance of risks and benefits between individuals, groups and society is a key concern of data governance.

Sensitive and non-sensitive data

The dichotomy between sensitive and nonsensitive data becomes problematic in the context of joined-up data sets. Apparently innocuous data can become revelatory or can be put to controversial uses when joined with other data sets, meaning that sensitivity can no longer be considered as an intrinsic property of data but an emergent property of data in use.

Openness and restricted access

Making data open can allow innovation in developing data use technologies and in finding beneficial uses for data. However, understanding the limits to openness requires an understanding of what data we would prefer to protect; that might be because it is potentially sensitive, or because it can be commercially valuable – both as an asset and as a commercial advantage.

Algorithmic efficiency and transparency

An algorithm may produce accurate predictions, but the ways in which algorithms evolve might inhibit the transparency and accountability of decisions made on the basis of those predictions. If there is such a trade-off, what are the criteria for valuing accuracy over transparency in different contexts?

Enabling governance

Governance and regulatory frameworks are often thought of as intending to restrict action. However, data protection regulation in Europe has its origins in the desire to allow free transport and flow of trade. The ability to access and share data can enable innovation, economic growth and unlock wider public benefits.

Human rights

Concepts of security, privacy, self-determination and freedom of speech are related to human rights. The extent to which these are mutually achievable in new technology contexts – and indeed outside of such contexts – is a matter for debate.

4. Governance

Governance does not necessarily refer to one particular set of rules, but encompasses a range of institutions and processes from the law to professional codes of conduct, regulatory frameworks and voluntary regulations, ethics and technical acceptability. It involves multiple actors with the ability to act at different times.

When should governance intervene?

At what point should governance come into play in the complex process of gathering, processing and using data? Do we simply wait until the data is used and then challenge based on the impacts – potentially impacts felt specifically by individuals – or do we try to establish governance frameworks from the earliest stages in data use?

Governance as an enabler

Governance should not only focus on managing risk, but also maximising benefit. If governance is intended as an enabler, especially of benefits to wider society, what does that tell us about when governance is needed in the development of data use?

Meaningful sanctions

Governance systems with the ability to influence behaviour are likely to be those capable of punishing undesirable behaviour. But this requires the credibility of sanctions and their ability to make a significant difference.

The role of government

What kinds of obligations relate to government vs. non-governmental bodies? Are there different duties and obligations that need to be fulfilled by government or by industry? Understanding where government needs to act, and where it is better for data users to develop their own voluntary codes, is important.

Duties, acts and omissions

What duties do organisations have to act on data they hold? For example, if robots used in the home record behaviour that shows patterns of abuse, is there a responsibility for companies to act on that information? While there is a lot of focus on protecting data, there is culpability relating to failure to act on information that is derivable from held data. How would such supposed duties be reflected in governance frameworks?

Futureproofing

A key feature of any governance framework is to be resilient to technology change and new opportunities to bring data sets together. Governance for the future is a key need.

Data protection

Data protection law is the main mechanism data has been governed by for the past 30 years, but is this regime, created in a different technological environment, appropriate for future use? For example, some of the basic principles of data protection – purpose specification and data minimisation – seem at odds with how we now use data. Big data analytics applies algorithms to massive data sets in order to generate hidden insight: big data is all about data maximisation and finding purposes for that data in ways not originally anticipated, potentially for significant public benefit.

5. Public Dialogue

How do to engage the public

Public debates can go badly awry: genetically modified crops being a paradigm example. They can also go well. For example, the Warnock Commission, based on good science and robust debate, ultimately created a situation where there was public acceptance of science, giving rise to some very difficult, ethical and moral governance issues.

The importance of trust

Building trust and trustworthiness is key to any governance approach. There are many examples of how trust can be gained or lost; a recent example being in healthcare where there was a failure of trust in handling of the care.data programme. Transparency and trust are linked – trust is more likely to exist in a context where individuals know why certain options are open to them and others not.

What people value

Discussion about governance requires an understanding of what people value and what kind of society we want to live in. To engage public debate, it is important to move past the detail of technology or law, and instead focus on what kind of world we want to live in. Through that understanding it is possible to judge what would be acceptable or unacceptable practices in data use.

Understanding public preferences

As well as explicitly asking for views on the acceptability of certain uses of data, public preference can be revealed by looking at what people are actually doing and how they are performing in relation to platforms. Looking at how people act in relation to using encryption, blocking cookies and following advertising could be a way to gain insights into how the wider public are trying to exercise and enact new forms of digital or data rights.

Platforms for debate

Genuine debate is challenging when it is mediated through social media platforms, which select what the user sees based on prior interactions. This can create a self-reinforcing algorithm where debate only happens between the like-minded.

6. Looking to the future

Can we engineer the future and create the society we want, or is technology already too far ahead? Are there core principles for governing data technologies that will stand the test of time, or will new technology merely be implemented for its immediate efficiencies rather than considering long-term implications? These questions are made more pertinent by the fact that legislation often struggles to keep pace with technology development. Data-driven technologies are developing swiftly and the landscape to be governed is in flux.

However, there is also the potential to adjust the ways systems are developed in relation to what we would like them to deliver. This involves starting by defining the core values that people would like to see in a data governance system, and working from that to identify how those values might be embedded in such a system. From a design perspective, if we decide what kind of information society we would like to live in, we can develop the ethical guidelines, the law, the framework and the governance that leads us there. Then, technology can grow in a way that yields benefit with risks managed and mitigated on the basis of clear, shared principles.

C: Discussion papers

For the seminar, a number of attendees produced papers in advance to help provide background and stimulate debate. They present summaries of current activities relating to data use and its governance, perspectives from individuals working in this space, and potential models for governing data use into the future.

The Royal Society and the British Academy would like to thank the authors of the papers for their contribution to the seminar and for allowing us to reproduce them in this report.

Medical Confidentiality in Historical Perspective

Professor Holger Maehle, University of Durham

Medical confidentiality is widely seen as the cornerstone of an effective physician-patient relationship. Patients' assumption that the personal information which they give to their doctor will be kept in confidence encourages them to be open about the details of their illness. This personal information will help the doctor in making the correct diagnosis and in prescribing an appropriate treatment. Keeping sensitive knowledge about patients secret is also regarded as an expression of respect for their privacy. This is the ideal picture. Sometimes the precept of secrecy in the Hippocratic Oath is cited here, in order to emphasise the ancient authority of this rule for the medical profession.

History, however, teaches a different lesson. Medical secrecy was, and remains, a controversial subject. Doctors have breached confidentiality in the interest of persons other than the patient; or they have been forced by the state to disclose information on patients in the name of the law or of public health. Simultaneously, secrecy has assiduously been guarded by the medical profession. Recent projects, such as the now closed 'care.data' programme, to make large databases of patient-related information accessible for research and audit purposes, have again drawn much attention to medical confidentiality, and its limits, in the UK.

Supported by the Leverhulme Trust, I have undertaken a book project, for University of Chicago Press, about the origins of the debate on medical confidentiality in the nineteenth and early twentieth centuries. This historical period is critical for the issue because it was then for the first time that the traditional medical ethos of keeping patient information secret was challenged in a big way by the demands of public health and the law. In particular, campaigns against the spread of venereal diseases (VD) and against illegal abortion put pressure on medical secrecy. Another arena where medical confidentiality was contested, were the courts of law, when doctors were called to give testimony about details of their patients. These issues were not only debated in Britain, but similarly also in the USA and in Germany, which has allowed me to organise my findings as a three-country comparison.

What were the main results and conclusions? One feature that emerged clearly for all three countries was an antagonism between the medical profession and the legal profession in the question of confidentiality in court. This antagonism was especially apparent in Britain. In the trial of the Duchess of Kingston for bigamy, in 1776, Lord Chief Justice Mansfield had compelled her surgeon to give evidence about her personal circumstances. Since then, British courts have refused to recognise a medical privilege in court that would have protected doctor-patient communications in the way that the legal privilege protected communications between lawyers and their clients. If doctors wanted to avoid punishment for contempt of court, they had to breach confidentiality. The issue became particularly acute in the 1920s, when judges forced doctors who had been subpoenaed as witnesses in divorce proceedings to testify to the alleged venereal disease of one of the spouses. The judges did this, although confidentiality had been guaranteed to patients in the government-sponsored VD treatment centres of the time. The argument of the judiciary prevailed that doctors cannot be permitted to obstruct the course to justice by refusing testimony.

Similar problems had arisen at the beginning of the twentieth century in Germany, although breaches of medical confidentiality were punishable according to the Penal Code of 1871 and doctors had been granted (in 1877) an entitlement to refuse giving evidence about their patients. In divorce proceedings, lower courts tried to compel doctors to confirm the venereal disease of one of the spouses, but the German Supreme Court (*Reichsgericht*) ruled in 1903 that medical witnesses could use their discretion in deciding whether they wanted to testify, even against their patients' wishes. At the same time, the Supreme Court envisaged a 'higher moral duty' that might entitle doctors to warn contacts of VD patients.

In the USA, the situation for medical witnesses differed, depending on whether the relevant state had recognised a medical privilege in court or not. In 1828, the state of New York had been the first to introduce a statute ruling that physicians and surgeons shall not be allowed to disclose any information which they had acquired in attending any patient in a professional capacity. Doctors in other states campaigned for a medical privilege also in their jurisdictions, sometimes against bitter resistance from parts of the legal profession. Critics of the New York statute argued for example that it was abused in personal injury or contested life insurance cases by silencing the medical witness if this lay in the interest of the plaintiff. By the turn to the twentieth century, about half of the American states had adopted statutes similar to that of New York. Still, by the start of the First World War, twenty-one American states continued to adhere to the English common law rule that doctors had no right to refuse giving evidence in court.

While the medical profession, in all three countries, was largely united in the aim to preserve confidentiality in legal proceedings, opinions were more divided on the question whether venereal diseases should be made notifiable by law in the same way as, for example, smallpox or diphtheria. Some medical anti-VD campaigners, in Germany, USA and Britain, argued that the collective interest in protecting public health must override the interest in secrecy of the individual VD patient. The choice, however, was not simply one between collective or private interest. Defenders of confidentiality expressed concerns that reporting of VD cases to the authorities might undermine the doctor-patient relationship to such an extent that VD patients would go to unlicensed healers, resort to self-treatment, or forgo any treatment – to the detriment not only of their personal health but also of public health. Therefore, medical secrecy in VD cases was in the public interest, not just in the interest of the individual patient.

Similar arguments were exchanged with regard to legal suggestions that doctors should report cases of abortion to the police or the public prosecutor as soon as they became aware of them – for example, if a woman sought their help after a botched abortion attempt. In all three countries abortion was illegal in the period concerned, except if the intervention was made to save the woman's life. Reporting cases, so the argument ran, would help bringing dangerous abortionists to justice and thus serve public health. The majority of doctors, however, seem to have resisted such calls for disclosure.

In parts of the USA and in Germany the existing medical privilege in court supported the position against the reporting of abortion cases. In Britain, the medical professional organisations declared their resistance by arguing that the doctor's first duty was to help the woman in such a situation, not to 'play the detective'. Only if the woman had died from the abortion, the case should be reported to the authorities. At the heart of doctors' reluctance to report abortions was a twofold concern about reputation: firstly about the woman's reputation, but secondly about the doctor's own reputation; his practice might suffer if he was known to be an 'informer'.

As my analysis of these historical debates has shown, the issue of confidentiality required intricate acts of balancing between individual and public interests. The notion of a 'right to privacy' was then still rather new: the Boston attorneys Warrren and Brandeis published their nowadays famous article with this title in 1890. Today, privacy concerns are paramount in discussions about health data bases, like 'care.data'. However, from a historical perspective, it seems to me that fears of reputational damage, if data sets are insufficiently anonymised to rule out identification of individuals, still are at the heart of the matter. And as in the past, it might be a twofold damage: not only to the patient concerned but also to the medical practice that has provided the data.

The Use and Reuse of Data

Hugh Whittall – Nuffield Council on Bioethics

The use and re-use of personal data, especially through 'big data' and data-linking initiatives, have the potential to deliver significant public and individual benefits. In the context of healthcare, these can include improved direct care; better management systems; and opportunities for research that can deliver future health and medical advances.

There are, of course, risks with the use of personal data, especially where they have been collected in a context where there is an expectation and duty of confidentiality. These include loss of privacy, distress, and loss of trust in the professional/patient relationship.

At the same time, there are risks associated with failure to share and re-use data, such as the possibility of suboptimal care and lost opportunities for potentially valuable research.

The privacy interests of those to whom data relates are generally protected through the instruments of the Human Rights Act, Data Protection legislation, and common law duties of confidentiality; and the usual way in which these protections are operationalised in the context of secondary data use has been through the dual approach of consent or anonymise.

However, there are problems with this approach. First, consent, normally taken at the time data is collected, cannot always project all possible future uses (and might not therefore achieve what it might claim), can be burdensome, and it does not necessarily protect all possible interests of the individual concerned. Nor is it always necessary, for example where a further data use does not affect a person's relevant interests. Second, anonymisation cannot always be assured in circumstances where different data sets can be linked, making re-identification possible.

Reliance on high-level instruments for the protection of people's privacy and other interests in data use are therefore insufficient because (i) they do not recognise the broader interests of the individuals from whom the data originated – whilst they may be legally sound, they are not necessarily ethically justifiable; (ii) they offer a one-size-fits-all framework that is not able to deliver the adapted governance that would be needed for different circumstances of data use; and (iii) the 'consent or anonymise' paradigm is not in itself adequate to the task.

Data initiatives (linking and uses of data that go beyond the purposes for which the data was originally collected) can therefore deliver significant benefits, but to be sustainable and publicly supported should, whilst complying with basic legal requirements, also establish a sound moral basis that secures public trust. Rather than simply complying with the law, they should be recognised as social practices that engage a wide set of interests, norms and expectations in ways that can differ between different initiatives, and which can be dynamic and evolving.

As well as ensuring legal compliance, data initiatives should therefore, if they are to meet dual (and inter-related) private and public interests, be co-produced in a way that involves participation, governance, accountability and, ultimately, trustworthiness.

Participation. Inclusive (of those with morally relevant interests), engaged and sustained deliberation regarding design as well as governance in data initiatives is important to identify the interest, norms and expectations of those whose interests might be affected by the use of data.

Governance systems should be adapted (and adaptive) to secure the trust of those whose interests are at stake through participation, adequate and appropriate security (which can include anonymisation) and continuing reflection on privacy norms and expectations. Consent (whether specific, broad, or opt-out) can form part of the governance system, but is not, alone, sufficient to protect individuals' interests, nor does it obviate the moral responsibility of the data user to take account of the interests of those who might be affected.

Accountability means that systems should be transparent; responsibilities should be clearly identified; criminal sanctions should be established for the deliberate misuse of data; and privacy or security breaches should be reported, whether or not actual harms are identified.

Through these approaches, trustworthy data governance can (and, if data exploitation is to be sustainable, must) become the basis on which innovative data uses both deliver on the public interest and protect the interests of the public. Maintaining trust necessarily involves professional competence, the ongoing engagement and alignment of relevant interests, and the recognition of public concerns such as the involvement of commercial interests.

With these in mind, the Nuffield Council on Bioethics has proposed a number of precepts for those approaching or developing a data initiative:

- Identify prospectively the relevant values and interests in any data initiative;
- Take special care to identify those interests that may be especially at risk or that arise from diverse values;
- Do not rely simply on compliance with the law to secure that data use is morally appropriate, particularly where it does not fully reflect moral norms;
- Establish what existing privacy norms are engaged by the contemplated uses of data;
- Involve a range of those with morally relevant interests in the design of data initiatives in order to arrive at a publicly statable set of expectations about how data will be used;
- State explicitly the set of morally reasonable expectations about the use of data in the initiative; and
- Involve a range of those with morally relevant interests in the continuing governance and review of data initiatives.

Data governance for Infrastructure and Utilities

Professor Martyn Thomas, The Royal Academy of Engineering

The Smart Grid

For the UK to meet its commitments to reducing carbon emissions, there must be substantial moves away from hydrocarbon energy sources towards lower-carbon sources, principally wind, solar, wave and nuclear generation of electricity. In turn, this requires that much more of the energy used in transport and in heating must be electricity rather than petroleum or gas and implies that there will be very many more electricity generators that will necessarily be geographically distributed and an increased adoption of electric vehicles and heat pumps.

The existing electricity transmission and distribution grids were designed to carry electricity from a small number of large generators to a large number of consumers and not to deal with intermittent and uncertain capacity from large numbers of distributed renewables sources. The changes outlined above require a more complex flow of electricity, and they bring challenging problems of power quality and of balancing supply and demand. The monitoring and control that will be needed to protect existing infrastructure and ensure safety (plus the strategic introduction of storage and other network innovation) are known as the *smart grid.* The alternative of replacing substantial parts of the existing infrastructure of underground and overhead cables is not attractive, financially or environmentally.

One of the main sources of data for managing the smart grid will be the millions of smart meters that the Government and industry have committed to install over the next 5 years. The specifications for these meters and for the management of the commands and data flows between meters and data users has been negotiated between DECC and industry. It has been a lengthy and complex process that is not yet fully complete and that will continue to evolve. I was involved in discussions with DECC about the security architecture for the smart metering system, on behalf of the IET and the RAEng.

I failed to persuade DECC to specify rigorously the required security properties of the architecture and the semantics of the commands and data flows. DECC and CESG argued that there was no point in doing this, as they did not trust the industry to implement the specifications correctly in any case. There is no overall systems architect for the smart grid, despite strong representations by the IET and RAEng (though the IET is now involved with the Smart Grid Forum[1]). Assurance that the smart metering system is fit for purpose will rely largely on testing; in my opinion this is a weak foundation on which to place the privacy of millions of householders and the security of critical energy infrastructure. The same seems likely to be true about the security of the smart grid.

Open Data

At a recent Foundation for Science and Technology Meeting, Sir Nigel Shadbolt described Open Data as "part of the UK's critical national infrastructure".

Open data is defined as data that anyone can use, process and share. The UK Government has taken the view that data collected by Government departments should be made available openly wherever possible; as a result the Government open data website, data.gov.uk contains over 22,000 datasets that are free for use under the Government Open Data Licence. Other organisation have also made their data available – Transport for London and other transport providers have created the data feeds that support the transport planning apps that provide real-time data to our smartphones, for example.

1 https://www.ofgem.gov.uk/electricity/distribution-networks/forums-seminars-and-working-groups/decc-and-ofgem-smart-grid-forum

A few years ago, a study in the RAEng reviewed the extent to which an extraordinary number and diversity of industries and services in the UK had become dependent on one source of open data: the location and timing data that is freely available from satellite navigation systems, GPS being the prime example. The report[2] described and illustrated the vulnerabilities that this had already created (the Academy agreed to remove some of the details that were considered too sensitive, even though these had been taken from publicly available sources).

GPS is an example of a common point-of-failure for many different activities (some of which are assumed to be resilient because they have backup from other services, even though these are also dependent on the same GPS signal).

Other open data sources could become single points-of-failure for multiple services. No one is in a position to know who is using each data source or what the full impact would be if a data source was falsified or unavailable. Therefore, no-one is able to assess how important it is that the security of these open data sources is assured. If Nigel Shadbolt is right that open data is critical national infrastructure, the weak governance of open data may become a problem.

2 http://www.raeng.org.uk/publications/reports/global-navigation-space-systems

Protecting Innovations and Protecting Data: Can We Have Surveillance and Innovation?

Dr Gus Hosein, Privacy International

International

Privacy can be seen as a reflex of innovation. One of the seminal pieces on the right to privacy as the 'right to be let alone' emerged in response to the camera and its use by the tabloid media. Seminal jurisprudence is in response to new surveillance innovations, though often with significant delays.

While one approach would be to say that privacy is a norm and that with modern technologies the norm must be reconsidered and if necessary, abandoned; I think there is an interesting idea around the question of protecting privacy as a protection of innovation.

Protecting data to protect innovation

At its most conservative, creating the basic rules of the land to ensure that people can have confidence in our technologies and transactions is why we have seen the spread of data protection laws – now in over 100 countries – and encryption – now considered essential to commerce. Neither safeguard was easy to establish as powerful interests have and continue to try to push back against them.

Studies repeatedly show that the lack of confidence in our infrastructure actually inhibits commerce. A recent survey from the Information Commissioner's Office shows a lack of confidence in Industry and Government to manage our information so people are 'taking matters into their own hands'[3]. The ICO interprets this as an explanation for the rise of adblockers.

As the theory goes, policy can create the optimal framework for trust. A recent survey of Americans found that they all wanted similar rules applied to all of industry, though the U.S. Congress continues to be unable to move on data protection rules after decades of inactivity with no sign of change on the horizon[4].

Protecting people's rights over innovation that they are being served

It is not law alone that people need. People do not transact just because they know there are laws in place. They are likely to want an element of control.

Another recent survey from Pew Research Center found that "when presented with a scenario in which they might save money on their energy bill by installing a 'smart thermostat' that would monitor their movements around the home, most adults consider this an unacceptable trade-off (by a 55% to 27% margin)"[5]. People do want control over their electricity and most likely want to reduce their costs. But because the data generation is uncontrollable, in this scenario, the concern rises.

Many of the innovations today do not give people that ability. Most technology development actually removes the ability for the individual to even know what data is being generated. Previously we could understand what was occurring over a telegram, telephone communication, and internet transaction. Now, with the numerous sensors, scripts and cookies on our devices and services, we cannot tell what data is being generated and what is happening with that data. Even as you imagine that, you tend to focus on the devices that we actually own – now think about the future smart city and what data is being generated by what technologies owned by which institutions? Today we have signs for all the CCTVs, but what will we have in the future for all the sensor networks?

Transparency and control are going to be radically different going forward.

3 https://ico.org.uk/about-the-ico/our-information/research-and-reports/information-rights-research/

4 http://www.progressivepolicy.org/issues/communications/ppi-poll-recent-national-survey-internet-users/

5 http://www.pewinternet.org/2016/01/14/privacy-and-information-sharing

Protecting innovation from surveillance

If we want to see a world where there is data everywhere, we need technology to be distributed everywhere. This is the idea that there will be sensors everywhere generating data, collecting data, sharing data. Such a world comes with new potentials but also a myriad of new risks.

Surveillance may actually be one of those risks. I would like to suggest that surveillance laws may chill the spread of these technologies. If we look at the current 'communications surveillance' discourse, it is focused on issues around the internet where the user possesses a mobile phone or a computer. But how do communications surveillance powers apply to the sensor-rich environment, or the Internet of Things? While it is in the realm of our imaginations to allow governments access to our locations and telephone call logs for the purpose of an important investigation, how about these scenarios:

- Would we as a society be willing to permit governments to know our movements at all times, here and abroad, monitor our heartbeats, and keep records of all our interactions and what we watch and what we read and to be able to go back in time and see everything we've done? This is already the law but rarely discussed publicly in such a way.
- Should we allow law enforcement to operate fake websites of our banks, pretend to be our utility providers, our friends, provide fake internet or mobile service so that they can monitor groups of people? They are doing much of this already on a widespread basis with no clear legal framework.
- Should we allow any institution to be able to search our lives and conversations and find out on any given moment if we like a given minister, or have a preference that some other interested individual or group may have? This has been done for years, under social media monitoring and GCHQ's operation Squeaky Dolphin.
- Should we ensure that all our personal databases, home CCTV security networks, baby monitors, wearable devices and fitbits, smart TVs and fridges, smart sensors, and all future services have secret back doors that allow for any government anywhere to get access to the information inside? This is all permissible already, according to the government.
- Should we hack into someone's home even if they're not a suspect? Hack a transport infrastructure provider? A telecommunications company? An airline? A car? A car manufacturer? Some of this is what governments are already doing and are claiming it is within the law.
- Finally, should we ensure that all these capabilities are built into all technologies sold everywhere and even the capabilities of surveillance be made available to any government who seeks it? That's the state of affairs now with surveillance technology where standards ensure that surveillance capabilities are in there by design.

These are simple examples because all they do is apply existing extraordinary powers to new infrastructure. This is the very same infrastructure that most consider essential to our future economic and social growth. And yet the internet has already been turned into a significant tool for collection and analysis of data by so many third parties. Why should we expect this to not happen again?

In this context, every demand and excitement around data-driven innovation is one that invites an infrastructure whose security and integrity will be undermined for the purpose of surveillance. The logic therefore says that if we want a secure infrastructure where individuals are in control, in order to have the confidence in a society where data is well-governed, then we cannot allow for it to be undermined.

Radical innovation or radical protections?

Innovation today seems to infer anything around the realm of data generation. It is interesting if we think about some of the concepts in our data world that we would not have if they were to be invented today. So deep is our excitement and discourse around data innovation that we would not ever consider creating limitations on collection and the purpose of processing, data protection regulators, limitations on data-sharing, and even warrant regimes. But these all pre-date the data and innovation discourse – how hard are we going to fight to change these 'risks' to innovation, instead fighting the ones posed by surveillance? Is it possible that the ability to collect and share data only arose because the legal context provided safeguards that the innovators are now removing?

Therefore, it is noteworthy that I feel like the 'radical' defending privacy in the context of data governance, when the ideas that should be branded as radical are the ones that are trying to undermine all of these practices, protections, norms and rights we have built up for good reason. Those are the innovations that should run deep, that give rise to all the others.

The future has to be bright: if we want all the things we want, we need the frameworks to provide them and prevent the things that will undermine them. At best, we will be able to develop a new discourse and new safeguards. At worst, we will continue the cycle we have long been stuck in: we build it, we take it to market, we promote it, and we act aghast when abuse arises.

Is Data Governance Needed for Robotics?

Professor Noel Sharkey, University of Sheffield

According to the International Federation for Robotics, there has been a dramatic upsurge in service robots for everything from healthcare to the care of children and the elderly, from cooking and preparing food to making and serving cocktails, from domestic cleaning to agriculture and farming, from policing, security and killing in armed conflict. 4.7 million robots were sold for personal and domestic use in 2014 including a 542% increase for assistive robots for the elderly and disabled. The forecast is a rise to 38 million by 2018 at a conservative estimate. And the predictions do not include rapid developments of driverless technology.

Governments and corporations are viewing robotics as a powerful economic that needs considerable funding. Many more companies and startups are getting in on the act to create a multitude of new robot applications in what is becoming a highly competitive market that will drive innovation. There is reason to be optimistic that many robotics applications will greatly benefit society. However, It is difficult to predict societal risks as robotics joins the Internet of Things and new developments in big data and machine learning are incorporated.

Yet, there is little or no government or international joined up thinking about how these developments might impinge on our privacy and fundamental civil rights. What can we do about it? What sorts of data governance protections can and should be put in place? Five potentially problematic areas are briefly described below to initiate discussion.

1. Accidental surveillance:

In 2012, the US Airforce created a furore when they warned that, when flying over US soil "Collected imagery may incidentally include US persons or private property without consent". Google ran into similar problem with its street view cameras. They were forced to pixilate the faces of the people they accidentally photographed. As robots become more common in our society the accidental video capture will be a major feature in our lives. Images are recorded by most outdoor robots from self-driving transport to home delivery robots and road work robots and even cleaning robots. When you receive your robot pizza delivery or talk to a robot shop assistant you may be recorded.

Q: Would it be possible to regulate accidental surveillance or should we just live with it?

2. Police surveillance

The UK police have been using small quadcopter drones for surveillance since 2007. In 2010 the Guardian used a freedom of information request to transcripts of meetings about a UK national drone surveillance strategy. Kent Police and five police other police forces were working with the home office and military contractor BAE systems. The PR was that drone surveillance would be used for maritime surveillance of our borders for illegal border crossing. But transcripts of the meetings revealed plans for the routine monitoring of antisocial motorists, protesters, agricultural thieves and fly-tippers, fly poster and a range of minor offences.

Q: Are there problems lurking here? Should the police use of drone surveillance be limited or constrained?

3. Home Robots

This is one of the largest development areas in robotics. Robot toys are now being developed with internet connectivity that could see the erosion of children's right to privacy. The new internet-connected Hello Barbie is a prime example where parents can sign away their child's privacy. All conversations between the child and Hello Barbie are recorded and stored on ToyTalk's cloud servers and used anonymously by the company to create more pre-recorded response options. The company says that children will sometimes be asked about their preferences and this will be used for occasional marketing through the doll.

There is also a major push to develop assistive robots in the home to enable the elderly to stay independent for longer. This involves home monitoring, but who can see the footage and who can hear what the elderly person is saying? There is a drive in Japan for developing conversational companion robots and we need to ask who will be able to hear recordings of the conversations.

And what about the increase of other interconnected home robots like Jibo or home security robots that can be operated remotely on the internet? What guarantees are there about the ways that the collected data will be used?

Q: What protections do we need to put in place to guarantee that what goes on in our homes, stays in our homes? Can we protect our data against hackers and others with criminal intent?

4. Bio sensing and behaviour recognition

This is an area fast gaining traction in security applications. It is part of the UK Home Office strategy (2011) for monitoring in high security situation.

The US is already ahead of the game with their SPOT system at airports for recognition of facial micro expressions that are purported to tell if someone is acting suspiciously or trying to hide something. This has led to protests from the civil rights community about the number and type of people being taken off for further processing. And the scientific community has complained about the lack of science underlying the technique.

In addition, the FAST system, again in US airports, is being used for biosensing of heart rate, sweat and breathing in an effort to find terrorists. Data has been presented to show that this is no more effective than tossing a coin for now. Yet is still being rolled out.

The *what if* question is, could this type of equipment be used in the consumer world? For example the Aldebaran Pepper robot has been designed with sales assistants in mind. It is at present a cute novelty that can video record the customers and answer questions about products. But it could be equipped with biosensing and micro-expression monitoring to determine what customers are interested in or how much they might be prepared to pay. And this information could be recorded for future use.

Q: Do we want to stop the use of these techniques from penetrating into the civilian world and if so, can we do it?

5. Non-security drones

Personal drones are essentially flying HD video recorders. They can be bought easily and cheaply from many retailers. It is very difficult to get an estimate of UK figures but they were Mapplin's biggest sellers over Christmas 2015 and in the same period in the US more than a million were sold. In 2014 there was an 80% increase in the number of permits issued by the Civil Aviation Authority for flying larger drones. Drones are also being increasingly being used by the media. Drone journalism has become a term and several training courses have been set up.

Q: (i) How can we regulate the use of personal drones to prevent 'peeping Toms' and other criminal uses? With these numbers, does it mean an end to privacy?

(ii) Should we let the media self-regulate their use of drones rather than simply letting them go wherever and do whatever is possible?

On the Importance of Data Governance, with Special Reference to Finance

John Taysom, Web Science Trust

Legendary US banker Walter Wriston, writing in 1992 in his book *The Twilight of Sovereignty: How the Information Revolution Is Transforming Our World* described how the basis for wealth had evolved from land to labour to information. "Information about money has become almost as important as money itself," he wrote, a quarter of a century ago.

Financial institutions today, banks and insurance companies, use internal and external data, both structured and unstructured, to develop personalised products. Much of this is collected from on-line activity. More personalised products result and customers stay loyal longer. Data analytics have an increasingly important part to play in the granting of personal loans, having been used in commercial loans for some years. The now central role of data in finance is reflected by a new 'C' suite appointment – the Chief Data Officer – whose role is to ensure data provenance, veracity, consistency, legality and regulatory compliance.

One could say that data in finance, once used to record the past, is now used principally to predict the future. Where once robbers broke into banks to steal money, now they use cyber-attacks to breach security to steal data. The data they want to steal is data about you and me, held in their customer systems, not about the operations or market positions taken by the banks held in the banks' accounting systems. Walter Wriston was right. But data is easier to steal.

Data in finance is heavily regulated in most areas; self-governing professional accounting firms ensure that data about companies is accurately and fairly recorded. Sarbanes Oxley in US, and similar laws and regulations elsewhere, protect the small investor from misleading information generated by companies. Rules on disclosure in most jurisdictions control the timing of the release of data on listed companies and laws or regulations ensure that only experienced investors invest in more opaque private companies. Because information is nearly money. Although it should be noted that reports this week say 20% of news-driven major stock market moves appear to have been anticipated by some market participants, which probably implies that this is still a work in progress. An industry exists to gather non-traditional data to be sold to hedge funds and traders to enable them legally to anticipate market movements. Classical economics was wrong, not all information is priced into markets. What sort of data? Everything from sensor data in shopping malls, to satellite images of crop growth, to cell phone traffic, to web activity. Suddenly there is an appreciation of what this data is now worth. The same is true of data about you held by telcos. Most have been developing new revenue streams around this data as a corporate asset. Yet, all this data about me is absent from the balance sheets of those companies. If the bits and bytes that represent my cash holdings are stolen from a bank, its balance sheet is reduced: if the bits and bytes that represent my personal profile is stolen, there is no such impact.

Consider the information accumulated by the newest entrants. The new capital created in the 21st Century by Google, Facebook, Amazon and others, ranks alongside other seismic periods of new capital formation for example during the US Gold Rush or the formation around discovered oil assets of new companies like Saudi Aramco. Yet, Thomas Piketty in his great work *Capital in the C21st* is silent on 'information capital'. There is no reference to it even in the index. Clearly the value of Google is not in the server farms nor in the code that runs on them to collect data, which is surely intellectual capital but is not information capital. The value is in the data. This asset is not a current asset, which is defined as one consumed or sold within a year. It is some form of long term asset, a capital item, the value of which lasts even after I cease to be a customer myself – though balance sheets do not mention it, and we have no accounting tools to recognise it, nor any entry in our National Accounts for what could be characterised as the most extraordinary hidden exportation of capital ever: to the West Coast.

Which makes me wonder where the information assets are?

I was certainly born with some: my date and place of birth, my parentage, my genomic make-up, and I have certainly acquired many since. Qualifications, experience, and all the elements of my medical phenotype. And all the activity that makes up my financial phenotype. Yet, I have myself no information 'capital' to speak of derived from those assets. I have some intellectual capital, and some monetary capital, and those combined with my activity and my intentions generate value for organisations, public and private, that collect data on my activity and 'monetise' it, that is, generate capital value from it. I may have access to a service in return for the data they collected. And that may be a fair exchange. But it is certainly not a transparent exchange. Often, I have no effective choice in the exchange. I was fortunate to chair a seminar at Harvard as part of the input to the 2014 Podesta review on Big Data for The White House. One conclusion was that when we are faced with a choice to have a benefit now with a potential cost later, we are bad at making rational decisions. We over consume. When we are offered a personalised cheaper insurance rate if we allow closer monitoring of travel or health, are we over-sharing now to gain an immediate benefit, with a potential cost in future?

The unexpected impact of the Web is not in how it has empowered distributed content production: "we are all journalists now", or so the slogan went, but in how the inexorable emergent property of the Web is the unexpected concentration of power in a few walled gardens bricked-in by control of metadata. This has not been done by bad people with bad intent – quite the reverse. The well intentioned have allowed us to be happily caught in our own echo chamber – our 'Filter Bubble' as Ali Pariser wrote – meanwhile allowing data about us to be collected in order to provide us with a better service. Did we make rational choices? Is it not the role of regulators and law makers to intervene when the market is systematically imperfect?

Regulators and law makers have taken note. The US had led a review of the data brokers that collate data and a draft Bill to regulate their activity was proposed in 2015. "What was a business of data keeping has become a business of data reaping", said its proposer. The EU General Data Protection Regulation, with extensive input from the UK, has led the debate about the implications of indirectly attributable data derived from running analytics, for example AI, against personal data we thought had been de-identified. The US has launched this week a major challenge to develop better privacy preserving tools. However, this is all a long way behind the web of laws and regulations that guide the use of data created and shared by companies. This is particularly important for the finance sector as it develops non-traditional methods of determining credit worthiness based not only on your actions in the "Intent Economy", to quote Doc Searls from the Berkeley School at Harvard, but also on the actions of your friends in that on-line space. Your actions, and those of your friends, when you and they intend to transact but haven't yet. When you are just thinking. There are serious problems of bias and non-inclusion in the applications of some analytics that are just being understood. For example, the impact of ethnic names on job interviews screened on-line. These biases are hard to correct. The algorithms can be unbiased but the database corpus will almost certainly not reflect any explicit editorial selection but only the motivations and economic status of the people who put the data there in the first place by their actions. You can hear the voice of Prof. Joan Robinson here: "We can no more detect our own ideology than we can smell our own breath." This data and these algorithms are determining your credit worthiness and are teaching your children. They may also determine your access to health via assisted medical triage. If you doubt the seriousness of this try the following search. Type 'Freedom Fighter' into Google or Bing and click on 'images'. Then do the same but for 'Terrorist' and compare the images presented. I am not sure how any bias, if you perceive there to be one, can be 'corrected'.

Therefore, in a finance setting we have issues that need to be considered around provenance: How did the data get there, under what business model? And around unintentional re-identification: given enough parameters it is possible to re-identify individuals from data that had been previously stripped of name, address, and obvious identifiers. There is a secondary concentration issue: the more data I can collect the better I can predict; I can feed my AI better input and make my service better, so I can collect better data on you. And the process is self-reinforcing. This puts the Google acquisition of Deepmind into some strategic setting – it was not, I think, about helping the NHS and developing new revenue lines, but about deepening its existing dominant market position in its current products. AI techniques will just serve to add an extra layer onto the inherent concentration that unexpectedly results as an emergent property from a web of peered nodes. However, the

GDPR, which may or may not affect processing of data on UK residents over time, is likely to do so before we Brexit, given that it becomes law here in September 2017 under the EU directive and Brexit is scheduled for end-2018, we now learn. The GDPR will still apply after that date for data on EU residents, whether processed here or in US. The GDPR makes serious demands on algorithmic decision taking including requiring a right to an explanation of how a decision that materially affects me was made. This is very much still an area of research as recent work at the OII has highlighted. (Goodman and Flaxman, June 2016).

So what policy prescriptions are implied?

I am particularly focussed on the issue of information capital because I think it unlocks many other issues. Referencing the Latin American economist Hernando de Soto Polar, we know that to turn assets into capital we need property rights and relevant institutions. I would say that it is not yet clear who owns data about me, except that it is clear that in many circumstances it is not me. In commerce, significant value is created from my data and I get little opportunity to share it. The value is only created when the data is shared and I cannot act alone to capture the value. Clearer laws on the ownership of data and the metadata that is created from it are needed. Government have been no better. The recent withdrawal of care.data, an attempt by the Government to derive value from aggregated NHS-collected health data about you and me, is a reaction to the angry backlash that greeted that initiative. Policy needs to recognise that where personal data is concerned personal choice is not a good base. We make bad decisions. This is true whether the motivation to share personal data is monetary reward or altruism. Sharing my genetic data or my medical data, because it feels right to do so, can also share data about my parents and grandparents, my siblings and cousins, and my children and grandchildren, though they were not party to the act of altruism. There are some indications already that consumers are starting to take action. Advertising revenues on-line have been severely impacted by the mass adoption of on-line ad-blockers which suppress all advertisements, in some countries up to 50% of users now use these applications to remove from the content the advertising that pays for it to be promulgated. The two most frequently cited reasons for adopting an ad-blocker are a feeling of being always surveilled and most importantly also a sense that the value created by their personal data, covertly gathered, is not being fairly shared with them by the service operator, not even by the provision of the service. The more ad-blockers are adopted the worse the data for finance company models, and the weaker the broad coverage of independent news on-line becomes. Yet, this is where millennials get their news.

A cooperative institution, independent of either commerce or Government, would seem to be the solution. Acting for me as 'producer' of personal data, to level the playing field with the broker or on-line service provider (the collector or data buyer), just as cooperatives do in agriculture. New technology allows the data that identifies me to be separated from the data that is needed in many applications to derive valuable insights from my data. Even the large on-line companies accept that the data that identifies me is toxic waste to them. They do not need to know who I am to develop relevant advertisements and within large financial institutions many applications do not need access to my identity to monitor anomalies or to generate relevant personalised product offerings. That allows many applications in finance organisations to be re-architected to minimise the 'attack surface' for potential cyber-attacks. The value is in the crowd. Even in medicine we speak now of 'precision medicine' and not personalised medicine. The value from data about me comes from sharing the data with data about you. What is valuable about us is our similarities and not the differences that make us identifiably unique. The Royal Society and British Academy can be the midwife for this essential new institution.

Innovative Ways of Communicating on Technology

Sheldon Paquin, Science Museum

At the Science Museum, barriers are everywhere; they are physical to keep your grubby hands off of our beautiful objects, they are emotional when we fail to connect visitors to human stories, and they are often intellectual, when we struggle to find a way to communicate technology appropriately. This last barrier, the intellectual block, is the most difficult to overcome. In the mind of the modern visitor, we consistently encounter the same hurtles over and over, ones which I would share below. These are the challenges that we face every day in science communication, the ongoing battle we face on a daily basis. If you visit the museum, please do two things: *stop touching that*, and please have these things in mind.

Technology is only technology

Museums and the outside world have one thing in common: everything gets a label. Unfortunately, these labels are also short, sometimes painfully so. How do you describe the Apollo 10 command module in 30 words? Do you discuss the missions' political impact, its social imagination, its ingenuity? Do you tell the human stories of the astronauts within? Do you make allusions to current events by discussing reusable rockets or the European Space Agency?

These days, the public loves to have things summarised simply – one needs only to look to the EU referendum to see how eager people were to simplify a very complex hypothetical scenario. More often than not, too, this simplification also calls for a categorical labelling of what is morally 'good' and what is 'bad'. Technology, of course, being technology, cannot be either ethically 'good' or 'bad'; where does a hammer fall in these categories? Yet, in our constant quest for brevity, we lump human developments into these groups. To take one example from the Science Museum, we had installed a display featuring a euthanasia machine used in Australia that helped four people to humanely end their lives. No judgement was made in the display of the object, its presence in the museum was simply to encourage thought and debate about man's relationship with machines. Instead, the humble showcase was met with outrage – a flurry of complaints forced curators to move the machine to a less conspicuous space. The fury that appeared from the public wasn't focused on the idea of having difficult conversations with children about death or identifying when a person has the right to die; the outrage came about because this technology was clearly 'bad' and therefore had no place in a public museum. Of course, this object was the physical expression of so much more at work, in terms of legislation, health care, and of individuals' stories, but this one object, a suitcase with some tubes and an IV drip, was branded as evil.

To be working at the Science Museum, we are constantly in an uphill battle to deliver a clear message: technology is only technology. Our inventions do not have their own morality; it is us that give them their uses. A euthanasia machine is no less ethical than a scalpel; but in the hands of a doctor or a psychopath, these can both be tools for something sinister. Our need to label the world around us is ultimately detrimental to our own understanding – we embrace the brief and overlook the nuance.

An expert is an expert

It is a curious thing when we hear scepticism towards an expert. As a museum, we work intimately with experts in their fields; physicists, data analysts, mathematicians, biologists, and chemists fill my work address book, and all text that is made publicly available is always checked first by an expert. The expert is the science communicator's bread and butter, a source of clarity when you can not figure out if it's accurate to use strong indicator words like 'every' or 'always'. There has been a movement in recent years towards the democratisation of science. The advent of the internet and better communication allows for all people to share ideas, which is fantastic in many ways, and nightmarish in others. On the one hand, we are seeing a boom in citizen science projects, people exploring science in new ways, and engaging with science through everything from actively editing Wikipedia to simply watching the occasional YouTube video.

On the other hand, this mass of voices is painfully non-discriminating, in that my thoughts hold the same weight as yours, because we are both free thinking individuals. While this may hold up in issues of ethics and philosophy, it is another matter when it comes to the scientific community. When every voice is equal, bad science is given undue attention. It is this kind of thinking that has allowed for the 'anti-vaccination' movement to take off, along with other scientific fads like climate change denial. There is nothing more beautiful in the pursuit of science to witness experts disagree; in theory, this is immediately followed by experimentation and reaching for truth. However, when a layperson opposes expert, it is given credence and that terribly toxic and lazy phrase "agree to disagree", legitimising conflicting facts and negating that genuine pursuit for truth. For science communication to be legitimate, there needs to be the inclusion of a real scientist, and that individual needs to be credited with the stronger word.

Current events are current

With the above mentioned idea that everything must be lumped together into neat piles of 'good things' and 'bad things', we miss an awful lot of nuance. As I am putting together an exhibition describing the impact of big data at the moment, we are doing plenty of research into how people perceive mass data collection and aggregation. What we have learned is that big data is understood to be either completely irrelevant, or will bring about a police state so powerful that simply mentioning George Orwell would result in your swift disappearance. While big data does have elements of surveillance and does bring up issues of privacy and control, it is also opening up the world to smarter cities, personalised medicine, catered services, and incredible efficiency in countless practical applications. With the popularised form of big data being the headlines surrounding data leaks and characters like Edward Snowden, people react very harshly to the whole process without a firm grasp of the relationship between a technology and its use. That is to say, how a technology is used now is often assumed that this is how it will be used in perpetuity.

The reaction to developments like stem cell research in the United States is similar. Indeed, the practice was for a very long time outlawed due to theoretically inhumane methods of collecting those cells. Of course, we have since made great strides in new methods for collecting stem cells, but now that the work has been branded as unethical from the beginning, it continues to face a crisis of public perception. To dismiss technologies in one way or another without the acknowledgement of either future development or new circumstance inevitably dooms scientific understanding to become static.

Perhaps in all of these notions I am suggesting, the acknowledgement of tools as tools, listening to experts, and recognising that times change, all have a common thread. For science's sake, I would ask our visitors to simply have an open mind; to think, to engage, and to discuss. Just please stop trying to touch everything.

Four Steps Towards a More Mature Infosphere

Professor Luciano Floridi, Oxford Internet Institute, University of Oxford

The ruling by the European Court of Justice in favour of the "right to be forgotten" is part of the coming of age of our information society. The tension between privacy and the value of controlling and shaping information about oneself, on the one hand, and freedom of speech and the value of having access to relevant information, on the other, has been with us for years. It is good news that we can no longer sweep it under technological or juridical carpets. Denial was the first obstacle to finding a solution.

Now that the first step has been taken and the problem has been acknowledged, the debate about how to reconcile our cherished rights and values is too important and intricate to be left to oversimplifying and sensationalist slogans. In the heat of debate, nuanced and more carefully balanced ideas may easily get lost. Therefore, the second step towards a more mature infosphere consists in recognising the delicate novelty of the problem we are facing, and hence the seriousness and complexity of the efforts needed to solve it. Data protection was developed when there was a clear divide between online and offline. Today, that divide is being bridged in favour of the 'onlife', a mixture of analogue and digital, physical and virtual experiences, like driving a car following the instructions of a navigator. There is a subtlety and scale to the challenge of our 'onlife' existence that calls for enlightened and innovative solutions. More and more, our lives are spent and shaped in the infosphere. Rather than just trying to adopt small, incremental changes in old conceptual frameworks, merely adapting previous legislation, or tinkering with current technologies, we need new and bold ideas. Sometimes, what an old technology may break, only a new technology may actually fix. Of course, this is much easier said than done, but then this is why we must welcome an open discussion.

This open discussion leads to a third step, which is to realise that – with the exception of a few fundamentalists – all parties involved have sensible points to make. The current debate over 'digital forgetting' may be exploited to fight proxy wars across the Atlantic, between different schools of political and economic thought in Europe, or between pro- and anti-European parties in the UK. But in reality, nobody is trying to destroy the internet, whitewash history, undermine an industry, or override one fundamental human right in favour of another. There are different rights, values, and interests – indeed different philosophies – at stake. We do not know yet how to harmonise them. Yet, our effort should go towards finding a collaborative solution. The temptation is to portray the debate in terms of a zero-sum game: Team Privacy vs. Team Free Speech, only one of them can win. Of course, yes/no questions are easier to explain, and zero-sum games are much more exciting to play, but this World Cup mentality is exactly the wrong way of conceptualising the issue. It is also a dangerous distraction, because it blinds us towards alternatives for reconciliation. What we must do is work towards a context in which all legitimate interests, rights, and values are represented and can find their optimal convergence. This is a very hard and tiresome task, but who ever thought that growing up was going to be entirely fun and easy?

The last preparatory step is to understand the depth of the consequences of our decisions in designing the infosphere. We would do well to pause and reflect on this moment – our moment – in the information revolution, beyond the particular scenario of the Google ruling. We increasingly live in a proxy digital world. We cannot unplug our society anymore. The result is a strange predicament: the map is becoming the territory. Anything may be 'just a click away', but in a sci-fi scenario in which one could remove all possible digital references to a closet, then that closet may practically cease to exist, regardless of whether it contains ugly skeletons or beautiful clothes. In computer science this is known as 'negation as failure': not found becomes synonymous with non-existent. This dystopia would be as unacceptable as its nemesis: keeping the wounds of history bleeding forever. Closure is not forgetfulness, but the capacity to remember without recalling. Such a capacity may require some compartmentalising and opacity. A common over-reaction to this suggestion is to speak of balkanisation: the risk of fragmenting the internet, for example, through a stricter protection for privacy rights in Europe than in other parts of the world. I would disagree with such claims. 'Fragmented' may just be another word for 'distributed'. The 'onlife' experience has been tailored for a long while. Even the same search engine already returns different results depending on the user's profile. When branches of the same bank or supermarket offer different services and products, we speak of variety, pluralism, and competition. The problem is not fragmentation in itself. The problem could be a 'divide et impera' (divide and rule): a balkanisation, yes, but one in which agents–commercial, political, or juridical–exploit walls and barriers to impose their informational monopoly locally, and have the last say on the region of the infosphere they control. Some non-democratic countries come close to this scenario. This is why gatekeeping is a headache. Who controls the controllers? The defence of privacy and the right to be forgotten should not lead to multinational corporations acting as unaccountable gatekeepers. Liberalism could become illiberal if it is so radical.

Our culture used to cope with the past through sedimentation. Today it is quietly – and, if properly overseen, desirably – acquiring a limited taste for reversibility. Letting bygones be bygones physically is not the only, and sometimes not even the main, strategy in digital contexts: there are also the edit and undo commands. It is an unprecedented opportunity in human history, which we should not abuse, and that should not promote recklessness. Instead, we must be creative, sensitive, and nuanced; not falling for convenience, nor giving up because the challenges seem too difficult or intractable. The information revolution has brought a remarkable capacity to tailor digital services and products for commercial and scientific ends. We must pay equal, if not more, attention to ethical ends, reflecting on how we may respect fundamental values and rights, beyond merely enshrining them in aspirational statements. It is trivial to remark that today we save by default and erase by choice. Yet, our memory is also very forgetful: inaccessible like your floppy disks, rewritable like your web page, fragile like your malware-prone laptop, limited like the Gigabytes in your smart phone, editable like your social media profile. Curating it is a complex problem and a difficult job with real human implications and sometimes dramatic consequences. Oversimplifying merely delays the moment when we deal with it seriously.

To think twice: this is the luxury we must afford.

Governance and Data Governance – Not Only Good, but also 'Seen to be Good'

Professor Roger Brownsword, King's College London

Introduction

The term 'governance' is used in a somewhat imprecise and varied way. However, I take it to refer to the setting (and enforcement) of rules, standards, codes, guidance, and the like, often (but not exclusively) by professions, business, industry, and sporting associations. In this sense, governance is to be differentiated from the more formal positive 'law' (or 'regulation') in the shadow of which it operates.

In the context of newly emerging technologies (whether 3D printing or gene-editing, blockchain or machine learning), it is thought that governance has an important role to play in ensuring that the overall regulatory environment (including not only formal and informal laws but also technological instruments that have a regulatory effect on human conduct) is 'fit for purpose'. Distinctively, governance is more flexible, more responsive, and more agile than formal legislation; it builds on technological and expert knowledge; and, because of its bottom-up approach, there is a real likelihood that the agreed codes will command respect.

The main points of this short paper are to suggest: (i) that good governance involves dimensions of process, product, and compliance; (ii) that 'fitness for purpose' implies (a) an acceptable three-way accommodation of interests in beneficial innovation, in human health and safety and environmental protection, and in respect for fundamental values and (b) 'effectiveness' (or compliance); and (iii) that there are some contexts – data processing in health care currently being one – where, because of a lack of public trust and confidence, governance must not only be good but also be seen to be good.

Good governance

Where a group has taken the lead in researching and developing a new technology, there will often be calls for 'responsible' application of that technology which, in turn, leads to calls for the articulation (by the group) of codes, guidelines, statements of best practice, and so on.

Of course, the first requirement is that the standards that are set are compatible with whatever background legal requirements apply. However, assuming such compatibility, governance will not be 'good' unless its processes for standard-setting, its products (the standards set), and the level of compliance that are achieved satisfy relevant expectations.

With regard to process, the expectation is that governance should be transparent and inclusive, with standard-setters being accountable. Where, as I am assuming, governance (in the context of emerging technologies with many beneficial applications but also with uncertain risks) performs a public function, these requirements become more demanding, especially in relation to the call for inclusivity.

Next, the context that I am assuming is one in which communities appreciate the benefits of new technologies and, typically, will want three things. First, they want beneficial research and development to be incentivised and supported, not obstructed. Invention and innovation should not be stifled; regulation should be light touch, proportionate, targeted and so on. Secondly, while they understand that there can be no guarantee of zero risk to human health and safety and the environment, they want such risks to be expertly assessed and managed at an 'acceptable' level. This invites various 'precautionary' measures which might delay the development of, and access to, the technology. Thirdly, they want the technology and its application to be compatible with fundamental values. This might involve some red lines being drawn (which again might constrain researchers).

If all members of the community shared the same idea of what constitutes an 'acceptable risk' as well as the same fundamental values, then accommodating invention, safety and value might not be too difficult. However, modern societies are pluralistic with many competing viewpoints in play. In these pluralistic settings, the challenge of good governance is necessarily all the greater – a challenge that modern societies increasingly try to respond to by putting the burden of justification on the integrity of the *processes* that have led to the setting of the standards (appeals to deliberative democracy, as per the US Presidential Commission on Bioethics, coming into their own here).

Failure to put in place an acceptable accommodation of the interests that are engaged or to follow acceptable processes will invite criticism as will a failure to achieve an acceptable level of compliance. Just as risk cannot be reduced to zero, no one reasonably expects perfect compliance. However, if non-compliance is compounded by corruption and the like, the charge that the regulatory environment is not fit for purpose will soon be made.

Data governance

In Europe, the general principles of data governance are largely dictated by the background law. Article 7 of the EU Charter of Fundamental Rights (in line with Article 8 of the ECHR) provides that everyone has a right to respect for his or her private and family life; and Article 8 of the Charter (broadly as in what is now the General Data Protection Regulation) provides that everyone has the right to protection of personal data. In the United Kingdom, the common law also recognises a right to confidentiality which presents a significant constraint against the onward circulation of information or data protected by the right.

The right to privacy is susceptible to a variety of interpretations. In North America, privacy is treated as a 'third rail', meaning that anyone involved in data governance would do well to take no chances or short-cuts where there is any risk of agitation relative to this value. Privacy, whatever it might mean, is to be fully respected.

The right to data protection, which has an unclear and contested relationship with privacy, is generally understood as requiring: fair and lawful collection of personal data; transparency and specificity about the purposes for which data is processed; use that is adequate, relevant, and not excessive; accuracy; retention for no longer than necessary; safe and secure holding, and so on. Confusion about data protection requirements is widespread; and, even for those who are not confused, there is plenty of scope for interpretation in relation to the meaning and application of the governing principles.

In the case of both privacy and data protection, because these are fundamental rights, many will take this as signalling the importance of data collection and processing being authorised by the individual's consent. However, there is also the view that, provided that the data is being used for 'legitimate' purposes, then it may be acceptable to proceed without consent. For all schemes of data governance, there is a difficulty: those schemes that prioritise consent will be accused of obstructing innovation that is data-dependent; and those that prioritise legitimate purposes will be accused of failing to take individual rights seriously.

Notwithstanding that the British public seems to be extremely supportive of technological advances in health care, the lesson of care.data and, possibly, the more recent agreement between the Royal Free London Hospital Trust and Google Deep Mind is that, where medical data is concerned, it is risky to rely on implied consent or opt-outs. This was also the lesson of the DeCode Genetics biobanking project in Iceland. Process matters. To restore trust and confidence, to assuage the discontents, short cuts should not be taken. Governance needs to be good *and it needs to be seen to be good* – which, no doubt, is much easier said than done.

Governing Algorithms that Learn in the Wild

Harry Armstrong and Lydia Nicholas, NESTA

While machine learning algorithms and other AI tools offer exciting new opportunities, they also present novel challenges for data governance. Transparency, historical bias, human-machine interactions, the adaptive nature of these tools and their ability to infer personal details from seemingly benign data are just some of the issues we will have to manage.

These challenges are not just technical, they span social, cultural and behavioural issues in an interconnected but subtle way. Relying solely on the individual right to privacy is not a strong enough form of public protection against the new risks. These tools are able to infer sensitive, personal information from data considered 'safe'. For example, the Cambridge Psychometric Centre, used Facebook Likes to accurately predict a range of highly sensitive personal attributes such as sexual orientation, ethnicity and political views.[6]

Transparency may be an important solution but is a challenge at multiple levels. Some machine learning algorithms are very opaque systems, a neural network can be too obscure and complex for even an expert to explain. Although the internal logic of many algorithms can be easily understood, trade secret protections can still keep them hidden from view. Even if the algorithm is openly available, it is unlikely the data that shapes the way it works will be, either because it is too valuable or sensitive to release. On top of this, the way the outputs of these tools are used will be hidden within internal organisational processes.

But perhaps we shouldn't expect too much from transparency, which can be hard to make the most of without what are still scarce technical skills. The experience of Freedom of Information laws and open data was rather different from what many expected, primarily because of very uneven capability to use the information that was freed. Dealing with this growing information asymmetry will need to be an important part of the public debate.

Using these tools changes the way people work and how they make decisions. A misguided sense of data and machine learning tools as objective can create serious problems. The most important decisions will be made by humans and machines together. However good the data and algorithms are, the way people use them will decide of their impact. Too much or too little trust in these tools will prevent the system as a whole from working.

Perhaps the biggest regulatory challenge is the speed with which these systems change and adapt to new inputs of data. The inner workings of the algorithm can change in a time scale of hours and days. Periodic auditing or constant evaluation would require unrealistic resources. We need new approaches and new tools.

Governance of algorithms at the EU level

In Germany and France, the requirement for companies or governments to provide an 'intelligible' explanation of what an algorithm is doing with people's data is gaining support. Restrictions on the use of algorithms (and machine learning) that 'significantly affect users' and the 'right to explanation' are already part of the EU's GDPR regulations set to come into force in 2018, but it is unclear how this form of transparency will be implemented.

Whether the UK leaves the EU or not, it is in a strong position to lead on data use and governance. The UK has pioneered many initiatives that we would now describe as 'data first', gathering data to discover patterns rather than waiting for theories. The recent creation of a data ethics framework by government and similar corporate initiatives are an important first step in developing the right governance structures, so what comes next?

6 http://www.pnas.org/content/110/15/5802.full.pdf

Next steps?

We need to start designing new institutions to explore and develop these new governance structures. Government has to be an important part of the conversation. Private organisations can and should attempt to embed ethics and engage the public, but ultimately the rules will have to involve accountable public power. It is not fair nor realistic to expect private firms to solve problems which are by their nature public. If government ignores the need for data governance, periodic public backlashes will make it difficult to reap the benefits of these tools.

Before we can create any kind of new regulatory system for machine learning and AI tools, we first need to know how these tools are being used and their impact. Important projects like the use of predictive risk modelling in child services are already happening, but government and regulators are not connected enough to learn the key lessons from these projects. Any new institution needs to be involved in what is actually happening on the ground.

Goeff Mulgan has proposed the creation of a Machine Intelligence Commission (MIC) to guide behaviours, understanding, norms and rules. It would not have formal regulatory powers in its initial conception. Instead, it should have strong powers of investigation, and of recommendation – much like the now disbanded Royal Commission on Environmental Pollution. It will help to establish some general principles – around accountability, visibility and control but look at specific examples. To make these powers meaningful, it would have strong technical, legal, social science and design capabilities. All of these skills will be essential if it is to analyse the whole process of the machine learning system.

Such a MIC would primarily investigate behaviours and processes. Initially, the MIC should use powers of investigation to define the spaces in which more formal and visible regulation will contribute to the public interest. While much of the work may involve dealing with risks, part of the role of a MIC could also be to drive up quality exposing bad design and encouraging higher standards.

In time it could evolve to have its own formal powers of regulation, standard-setting, fining etc. Or suggest which agencies should have these powers. It may be valuable for both the Government and industry to develop standard or certification schemes to assure the quality and fairness of systems without revealing valuable IP. As this area develops, having a clear division of labour with other bodies such as the CMA, ICO, GDS or Open Data Institute, will become ever more important.

Making Statistics Matter in a Rapidly Changing Data Landscape

John Pullinger – UK Statistics Authority

The benefits and importance of statistics

Official statistics are a vital public good for the information age, providing critical insights into the demographic, social and economic characteristics of the UK and the way these characteristics are changing over time. Earlier this month the Department for Culture, Media and Sport introduced a Bill to Parliament containing a series of provisions to enable the Office for National Statistics (ONS), to access holdings of administrative data held by public authorities and large, data-rich private organisations. This new legislation would give ONS better tools for harnessing the exponential growth in the scale and diversity of data of recent years, and for producing faster, more granular and relevant statistics. The legislation would facilitate innovation in official statistics to:

- Help policy-makers and service providers in central and local Government ensure economic and public policy remains relevant and effective;
- Help public administrators make ever greater efficiencies and ensure public money is put to the best possible use;
- Support our economy by informing the commercial decisions of UK businesses and driving competition by improving the 'information baseline' available to market entrants;
- Strengthen the UK's democracy by informing citizens and therefore helping hold elected officials to account;
- Help the UK's research community and media to produce analysis that sheds light on a broad range of economic and societal challenges; and
- Support the methodological innovations needed to overcome problems with declining survey response rates and reduce the burdens on survey respondents and those who collect and process statistical data.

Understanding an evolving data landscape

One central driver of the legislation has been an increasing recognition of the fundamental changes to the nature of the data landscape that have taken place in recent years, and a corresponding need to ensure these changes are reflected in the way statistical research is conducted.

Existing UK statistical legislation is misaligned with the reality, scale and speed of these changes. It binds ONS, for instance, to cumbersome processes for accessing administrative datasets held elsewhere in Government that severely restrict its use of these data sources. It has no provision for access to administrative data held by private sector organisations.

Sir Charles Bean's Independent Review of UK Economic Statistics emphasised that greater use of public and private administrative data has the potential to transform the provision of economic statistics. Indeed, without it we will struggle to capture the changes that impacts the world around us, from financial flows, to labour markets, and new forms of activity such as the sharing and digital economy.

Importance of an ethical framework

There is a corollary to this expansion of opportunity. The increase in the value of data increases its capacity to effect change in our societies – for good or for ill. The democratisation of both access to and the power to create new data mean that the consequences of the use and misuse of data are bigger than ever before.

We should remember too that, historically, data were created for very specific purposes – but in the data-rich world we live in now, data have become a sort of incidental by-product of the ways we interact and the technologies we deploy during those interactions. We have become data subjects just by dint of interacting and participating in society; our digital communications, our interaction with government and our consumption of

goods and services are all tracked, logged and recorded. As such, there are no inbuilt procedures and practices that delimit and govern the way those data are used. Moreover, the proliferation and diversification of data in recent years has had profound implications for the way we understand concepts such as privacy, with the world of social media, to take but one example, shifting the boundaries between our private and public realms. It also opens up as yet unanswered questions about who owns these data – a question that becomes more pressing in the context of rising awareness of the financial value of data and the increasing use of linked datasets.

In other words, the new data world confronts us not just with technical challenges (how to filter masses of data into something useable) but *ethical* challenges. Data must be collected and processed in ethical ways because of its potential impact, first and foremost. Personal data that is misused, or statistics produced on the basis of unsound methodologies, can have profoundly negative impacts on those who the data describes.

But ethics are also important, from the perspective of a National Statistical Institute (NSI), to maintain public trust in the statistics produced on the back of those data. Research carried out by NatCen in 2015 made clear the strong link that exists between statistical work that is conducted in an impartial way and with integrity and sound methodologies and public trust in statistical organisations. If we cannot demonstrate that we maintain the highest ethical standards when producing official statistics, then we risk eroding public confidence in a way that will inevitably lead to similar erosion of support for evidence-based policy decisions. In short: believing and advancing the role of evidence in our public life requires us to advocate for, and maintain, the very highest ethical standards.

How do we determine what is ethical in the production of official statistics? The NatCen research made clear that the value of statistics in public discourse rests heavily on the extent to which the public believes those statistics to have been presented in support of political agendas or not. At the heart of ONS's work is a commitment to producing statistics that serve the public good, and in acting in a transparent and accountable fashion.

The National Statistician established a data ethics committee in 2015 to consider a number of these questions in the context of statistical research conducted within ONS and across the GSS. The Committee considers proposed statistical and research projects in the context of a number of ethical principles, seeking assurance that:

- The use of data has clear benefits for users and serves the public good;
- The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately;
- The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity and quality;
- Data used and methods employed are consistent with legal requirements such as the Data Protection Act, the Human Rights Act, the Statistics and Registration Service Act and the common law duty of confidence;
- The views of the public are considered in light of the data used and the perceived benefits of the research; and
- The access, use and sharing of data is transparent, and communicated clearly and accessibly to the public.

These ethical questions will be critical to the way ONS implements the new legislation and exercises new powers under that legislation. The proposals contain provisions for the drafting of a code of practice that will outline some of the key ethical and operational principles to which ONS will have regard when exercising its new powers. Structures for scrutinising ONS's adherence to these principles will be at the heart of the legislation's implementation framework. This is crucial not simply as a means of satisfying Parliament that new powers will be exercised in a responsible and accountable way, but because ethical principles represent critical waypoints in a data landscape whose horizons are changing more rapidly than ever before.

Watching the Watchmen

Jon Akwue, Lost Boys

The conversation surrounding data governance normally revolves around the need for governments, policy makers and businesses to ensure that citizens are protected and there is a fair exchange of value.

For example, the recent Facebook commissioned report by Ctrl-Shift, *A New Paradigm for Personal Data* (June 2016) outlines five shifts that are required to maximise the contribution personal data makes to the economy, to society, and to individuals.

This involves shifting the onus of responsibility from consumers to technology providers. They argue that people do need education, but this does not mean that we should force long, detailed disclosures on people whenever they try to use a service. Instead, we should work to educate people about the issues that really matter, in ways that help them learn what is most important.

They also argue that value exchanges in existing data relationships should be reviewed and if necessary recalibrated, and we need to explore new ways of using data to add personal and social value, such as innovative service models that work on behalf of the individual.

They call for a more 'joined-up' policy making approach and a shift from compliance to sustainable customer relationships. They conclude by stating that companies need to move from good intentions to good outcomes by dropping assumptions focused around the fictional idea of a 'reasonable' decision-maker to design processes, mechanisms and services that work with the grain of actual human behaviour.[7]

What the Consumer Really Thinks

The Direct Marketing Association has been tracking changing consumer attitudes towards data privacy. Their report published in June 2015 shows a considerable change in attitudes since 2012, with significant increases overall in those willing to share data and a significant decrease in fundamentalists opposed to sharing data. This was most marked with young people. Almost two-thirds of younger consumers, claim they feel more comfortable with the idea of exchanging some personal data with companies than they did previously.[8]

The DMA identifies the rise of the Consumer Capitalist, with 23% of 18 – 24s believing that they benefit the most from the sharing of data with brands. This is in stark contrast to the majority of consumers, 80% of whom claim that businesses generally benefit the most from data sharing.

Social Everything

This difference in attitude between young people and older generations is indicative of the differing attitudes towards digital technology and social media across age groups. Research undertaken by The Center for Generational Kinetics in the US, identifies that 'Generation Z' acceptance and usage of technology is likely to be more similar to that of peers in distant countries than grandparents in their own country.[9]

They find that 42% of Generation Z respondents (who they refer to as 'iGen'), say social media affects how people see you. This 'outside looking-in' assumption affects what iGen posts, how they think about social media as a tool for making a statement and the fact that the world looks at you through the prism of social media.

Twenty-nine percent of iGen, more than any other generation, says that social media affects your popularity. 38% of iGen, slightly more than the 34% of Millennials, believe that social media affects your influence.

7 https://www.facebook.com/anewdataparadigm

8 http://dma.org.uk/uploads/ckeditor/Data-privacy-2015-what-consumers-really-thinks_final.pdf

9 http://genhq.com/wp-content/uploads/2016/01/iGen-Gen-Z-Tech-Disruption-Research-White-Paper-c-2016-Center-for-Generational-Kinetics.pdf

Inverting the Paradigm – Watching the Watchmen

This growing sense of awareness of how social media can be used to affect their lives and those of others is driving a number of behaviours. One of which is the rush towards private messaging apps such as WhatsApp and Snapchat, especially as the latter provides them with the assurance that the messages they send can be restricted and time limited. This allows for greater self-expression than on the 'open' social networks, such as Facebook, Instagram and Twitter.

The other phenomenon is a growing awareness of how mobile phones provide the opportunity to turn the tables on surveillance culture, creating a new cadre of citizen journalists who use their camera phones to record abuses of power by those in authority. This has been recently been brought to global attention by African-Americans in the US filming police shootings of black men.

On Wednesday July 6, 2016, the 10-minute Facebook Live video of the aftermath of a police officer shooting Philando Castile in Minnesota provided a graphic example of the power of video streaming.

The video taken by the victim's girlfriend and broadcast via Facebook Live was initially taken offline for about an hour. It was later restored with a warning that it was 'disturbing', following an outcry on social media. Within 24 hours, the footage had more than four million views, and had become an international news story.

This raises some complex ethical and policy issues for technology companies and society in general. According to a report in The Telegraph, Facebook are reportedly expanding the team dedicated to reviewing live content and staffing it 24 hours a day.[10] The company said it would also test the monitoring of broadcasts that go viral or are trending even before they are reported, giving Facebook a way to stop offending broadcasts quickly, in a similar way to a traditional TV network.

As an increasing cohort of technologically savvy young people enter into adulthood, we can expect to see more disruption caused by the use of technology, raising important questions regarding the governance of data in our increasing connected world.

10 http://www.telegraph.co.uk/technology/2016/07/07/facebook-live-streaming-of-us-police-shooting-of-black-man-leads/

Data Governance and Modern Computer Systems

Dr Alistair R Beresford and Martin Kleppmann – Computer Laboratory, University of Cambridge

Computer systems in the 21st Century are not neatly packaged in a box and locked in a secure room to protect the data stored on them. Today, computer systems are composed of a variety of physical pieces of hardware that are geographically distributed and connected via the Internet. When sensitive data is entered into computer systems, it legitimately flows through many digital networks, across institutional boundaries and over national borders. It is stored on, and communicated by, a variety of computing devices. These devices are owned, managed and accessible by many different people working for a variety of different organisations.

Take email as an example. If I send an email to you, who has access to its contents? It likely includes anyone with physical access to my, or your, smartphone, tablet or laptop; engineers at the Internet service providers over whose networks the data is transmitted; engineers maintaining the sending and receiving email servers; companies that provide the datacenter infrastructure; any employees who can trigger software updates on computer systems processing the email traffic; any outsourced spam and virus filtering services; and multiple government agencies with an ability to collect and analyse Internet traffic. When we talk about data governance and trust, we need to remember that governance is not only between the original sender and the intended recipient. Many other people and organisations are also involved in providing the computing infrastructure.

The story for many other online services we use today is just as complex. What happens when we upload a document to Dropbox, send data from a fitness tracker device over the Internet, or use online accounting software? Understanding the complete picture of all data flows and access rights in these systems is all but impossible. We are left with a stark choice: blindly trust the service providers or not use such services at all. This is problematic because such *'cloud'* services have a lot to offer, not just in terms of convenience, functionality and cost, but also in terms of security. Companies like Google are likely able to attract, train and retain a better team of security engineers than many smaller institutions, allowing them to respond quickly to new threats.

The evolution of computers systems continues apace and the Internet of Things is the latest development: computing technology built into physical objects, autonomously and transparently assisting us with our lives. We will soon have hundreds or thousands of computing devices embedded into most public and private spaces.

Embedded, Internet-enabled devices will collect, process, store and distribute data captured from sensors. Such systems will record much detail about our personal and professional lives, including the clinics we visit or the quantity of beer we drink in the bar. Not all such data will be recorded explicitly, but sensitive information can often be inferred from seemingly harmless data, and data intended for one system will be reused for other purposes.

Unintended interactions may happen: for example, a microphone connected to a computer, such as Amazon's Echo, may record and infer the keypresses of a password entered into a nearby laptop. The contents of a WhatsApp message written in Spanish may be translated into English using a web service, thereby revealing its contents.

The security of data also depends on the correctness of the software used to control access to it. Unfortunately writing vulnerability-free code is currently impossible. The best we can hope for is '*whackamole security*' – that is, fixing bugs and updating software on Internet-enabled devices before a malicious person can take advantage of the flaw. In the context of data governance, the move to cloud services and the Internet of Things, the consequences of poor security are profound. Can we be sure that the people with legitimate access actually fix security problems before they are exploited by people who should not? Is it possible to take remedial action after an inevitable breach of sensitive information, and what would it look like? Furthermore, what are the economic incentives for a company to provide updates in the first place? In many sectors of computing, security updates are currently rarely, if ever, provided.

Another area of computer science which is making huge strides forward at the moment is machine learning. This approach allows computers to translate documents from one language to another, automatically filter out spam email, and beat the very best human players at chess and go. Machine learning techniques typically require large quantities of data to learn from, and thus encourage the construction of computer systems that collect large quantities of data.

In other areas of computer science, network effects and economies of scale mean that companies with the largest repositories of data are the most profitable. It's no accident that there are a small number of social networking sites: individuals join social networks to connect with others, and so there is natural consolidation around a single service. Similarly, computer systems often benefit from economies of scale when there are a large number of customers, leading to a small number of large firms in any given sector.

In summary, computing infrastructure today is a complex system of many interacting components and organisations. The construction of large distributed computing systems is driven both by technical innovations (e.g. machine learning) and economic pressures.

Distributed systems and cloud computing offer significant advantages in terms of convenience, functionality, and cost. However, when thinking about data governance, we need to remember to consider not only the policies for access control at the application level, but also the many layers of infrastructure that lie beneath the surface. People may have access by design, by accident or by malice, and access to data can occur at all stages in its capture, storage, processing and distribution. Access to sensitive information could be indirect, including by inference, and also via one of many potential security flaws. With the Internet of Things, the situation is going to become significantly harder, due to the increased complexity of systems, the quantity and types of data collected, and also with the limited financial incentives to provide software updates to fix security problems.

It is tempting to suggest a return to the 20th Century – to a time when we would place data on a single machine and safely lock it away. For certain datasets today this remains feasible, but it shuns a number of positive and useful innovations, and it is going to be increasingly impractical in the future.

Towards a Successful Data-enabled Economy: Promoting Trust in Data and Data-driven Systems

Alan Walker and Philippa Westbury, The Royal Academy of Engineering

How can the UK create a 'data-enabled economy' through the use of data analytics, supported by data science and advanced data connectivity? This was the central concern of *Connecting data: driving productivity and innovation*[11], a report by the Royal Academy of Engineering and the Institution of Engineering and Technology and two workshops that followed its publication[12,13]. The report conveyed the breadth of opportunities across and between sectors, and the social, economic and environmental benefits that data analytics can bring. It balanced these against barriers that include data privacy and security; overcoming the engineering challenges around data, analytics and data-driven systems; insufficient access to proprietary data and a reluctance of organisations to share data; poor access to broadband services; a lack of standards and a lack of data skills.

Trust in data and data-driven systems

If barriers are not adequately addressed, there is a risk that trust in the use of data and data-driven systems will be damaged. This includes public trust in how government, companies and other organisations use personal data. It also includes the trust of companies in sharing proprietary and commercially valuable data for mutual benefit, and the sharing of data between private and public organisations that has been identified as a strong potential generator of economic value[14].

Trust also relies on ensuring that individual, corporate and broader societal benefits are balanced between stakeholders. There is some evidence that the public are willing to share personal data to get a better service[15], but in many instances asymmetries still exist between organisations and consumers such that the organisation has a much better idea of how they can benefit from data than the consumer. Platforms[16] are being developed that allow individuals to control and benefit directly from their personal data, responding to the need to rebalance control of data and its benefits. These and other platforms[17] that allow organisations to share data will need frameworks that promote trust, a key element in enabling access to data and the growth of data markets.

Data as an asset

If companies consider their data as an asset, it will help them to focus on how best to develop, trade, protect and exploit it. The value of data depends on many factors including quality, integrity, provenance, timeliness and the existence of appropriate metadata. The ability to link different datasets – whether across government or company departments or from entirely different domains – also profoundly influences value. A 'data value-chain' exists whereby the value increases as data is transformed into information, knowledge and ultimately action. The nature of this valuechain varies according to the industry and type of data – these, in turn, influence the point in the value chain where there is a willingness to pay or the potential to extract value.

11 Royal Academy of Engineering and the Institution of Engineering and Technology (November 2015), *Connecting data: driving productivity and innovation*, www.raeng.org.uk/connectingdata.

12 Cyber safety and resilience: strengthening the digital systems that support critical infrastructure and the internet of things, March 2016, Royal Academy of Engineering workshop (unpublished).

13 Data as an asset: exploring how to value data better and unlock its potential for wealth creation, May 2016, Royal Academy of Engineering workshop (unpublished).

14 McKinsey and Company (2013), Open data: unlocking innovation and performance with liquid information.

15 A recent study of travellers' attitudes to intelligent mobility by the Transport Systems Catapult found that 57% of respondents would not mind sharing their personal data in order to get a better service.

16 For example, HAT - the Hub-of-All-Things, http://hubofallthings.com/.

17 For example, the MK Data Hub, a data platform for Milton Keynes, http://www.mksmart.org/data/.

It is, however, a challenge to measure the value of data as it is an intangible asset, although the question of how to do so is being addressed both for the purpose of company valuation and to measure the contribution of data for national accounting purposes[18,19]. Companies such as Google and Facebook do not have data included as a capital asset on their balance sheets even though data is central to value generation and to their stock market capitalisation. Companies could in future include information on data assets in narratives to accompany financial statements that comment on the success of a business model or key drivers of value[20]. Another approach may be to use Key Performance Indicators that represent different aspects of data assets held by a company. It is also important to understand the value of national assets such as health data and ensure the value is recognised by the appropriate stakeholders.

Data security

An important consideration is the security of data at rest or in transit, to protect its integrity and availability and to reduce the risk that it may be used for hostile purposes. Industrial espionage involving engineering design data or commercial performance data is also a risk. The security of the Internet of Things is a further concern. The appropriate security and privacy architectures must be designed and implemented from the outset.

Data breaches are becoming increasingly prevalent and awareness is growing of the risks and the subsequent costs to the companies involved[21]. Alongside this, the EU General Data Protection regulation that comes into effect on the 18 May 2018 will introduce much higher penalties to companies for data breaches[22]. Companies will need to take cyber security more seriously in future and understand how their systems could be vulnerable.

For example, TalkTalk became more vulnerable to attack following the merging of a number of companies resulting in insecure systems. Target, a US retail company, was attacked through its supply chain leading to the theft of millions of customer credit card details. A much greater awareness and control of supply chain vulnerabilities by companies is needed.

The emerging questions are:

- What will catalyse companies into becoming more strategic about data, its governance and cyber security, and to treat data as an asset?
- How can opportunities to share proprietary data between organisations best be realised and what frameworks are needed to enable this? What are the barriers?
- What more can be done by government, industry, business and others to ensure that trust is maintained and data opportunities are realised? What situations might lead to a loss of trust?
- How can broader economic and societal benefits that arise from data be reconciled with an individual's possible loss of privacy?
- Will it be possible to rebalance the information asymmetry between companies and individuals so that in future individuals are able to benefit from their own data?
- What standards and regulations are needed that promote innovation and accommodate possible new uses of data in the future?

18 CEBR (2013), *Data on the balance sheet*, a report for SAS.

19 Goodridge, P., Haskel, J. (July 2015), *How does big data affect GDP? Theory and evidence for the UK*, Discussion Paper, Imperial College Business School.

20 Financial Reporting Council (June 2014), *Guidance on the strategic report.*

21 Ponemon Institute (June 2016), 2016 Cost of Data Breach Study: United Kingdom.

22 Penalties will reach an upper limit of €20 million or 4% of annual global turnover, whichever is higher. Council of the EU, 18 December 2015, Press release – *EU data protection reform: Council confirms agreement with the European Parliament,* http://www.consilium.europa.eu/en/press/press-releases/2015/12/18-data-protection/ .

The Governance of Personal Data in an Era of Ubiquitous Computing

Professor Karen Yeung – King's College London

'Data is the pollution problem of the information age, and protecting privacy is the environmental challenge. Almost all computers produce information. It stays around, festering. How we deal with it – how we contain it and how we dispose of it – is central to the health of our information economy. Just as we look back today at the early decades of the industrial age and wonder how our ancestors could have ignored pollution in their rush to build an industrial world, our grandchildren will look back at us during these early decades of the information age and judge us on how we addressed the challenge of data collection and misuse.'

Bruce Schneier, *Data and Goliath* (2015)

Introduction: the governance of personal data

1. This brief set of reflections focuses on the governance of personal digital data, rather than personal data more generally, or other types of digital data, for it is in this realm that the need for robust governance regimes is most urgent and arguably the most difficult to design, establish and implement. My comments are concerned primarily with data collection, processing, use and transfer of personal data, rather than with data security and encryption.

2. Although the legal definition of 'personal data' has generated considerable debate, for present purposes I adopt the definition adopted in the new EU General Data Protection Regulation[23] ('GDPR') Article 4(1) which defines personal data as 'any information relating to an identified or identifiable natural person ('data subject')' i.e., 'one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.'

3. It is important to distinguish, however, between data protection law, on the one hand, and data governance regimes, on the other. Data protection law constitutes an important component of a data governance regime, the latter referring to the larger institutional configuration of ethical, professional and behavioural norms of conduct, conventions and practices that, taken together, govern the collection, storage, use and transfer of data and the institutional mechanisms by and through which those norms are established and enforced. Legitimate and effective data governance requires a holistic perspective that is based on acquiring an understanding of the overall structure, dynamics and interaction that forms the basis of the contemporary personal data eco-system. Within this broader, dynamic and highly complex ecology within which personal data now flows, the role of contemporary data protection law has come under increasing strain.

Contemporary data protection law

4. Legal scholars often distinguish between two contrasting models of data protection law: the approach taken within EU law via the enactment of general 'omnibus' laws that apply to all data collection and processing activities, on the one hand, which is typically contrasted with a sector-specific approach in which bespoke legislative provisions are established to govern the collection and handling of data in domains that are considered especially sensitive, which is the approach taken in USA federal law (notably personal medical data and financial data), on the other. While the former provides significantly stronger and more comprehensive legal protection to individuals concerning the collection and use of personal data, both the content of data protection laws in both EU and USA approach to data protection are claimed to rest on the so-called ' Fair Information Principles' (FIPs).

23 The GDPR is due to come into effect within the EU on 28 May 2018.

5. Although there are several formulations of the Fair Information Principles, the basic foundation for many data protection laws rests on the principle of 'data minimisation', which is a combination of the traditional principles of collection limitation, data quality, purpose specification and use limitation[24] as set out in Article 5 of the GDPR which requires that personal data must be

   (a) Processed fairly, lawfully and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency')

   (b) Collected for specified, explicit and legitimate purposes and not further processed for other purposes incompatible with those purposes ('purpose limitation');

   (c) Adequate, relevant and limited to what is necessary in relation to the purpose for which data is processed ('data minimisation');

   (d) Accurate and, where necessary, kept up to date ('accuracy');

   (e) Kept in a form that permits identification of data subjects for no longer than is necessary for the purposes for which the persona data is processed ('storage limitation'); and

   (f) Processed in a way that ensures appropriate security of the personal data including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality').

6. Yet, many data protection scholars doubt the adequacy of contemporary data protection laws to deliver legitimate and effective personal data governance. Although the EU data protection regime contains some of the most demanding and restrictive standards of data protection in the world (particularly in light of the willingness of the European Court of Justice to champion the data protection rights of individuals), the inability to national data protection authorities to effectively enforce these laws suggests that, in practice, these laws have been largely honoured in the breach. Moreover, as data protection scholar Bert-Jaap Koops points out, European data protection law suffers from a PR problem[25]. The business sector typically regards these laws as an obstacle and a nuisance, failing to appreciate that they are in fact intended to be empowering and facilitative, aimed at *promoting* data transfer across national borders by seeking to avoid differential data governance standards across member states that might otherwise impede technological innovation and economic development that transnational flows of digital data could enable.

24 See OECD, Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data (2013).

25 Koops, Bert-Jaap, 'The Trouble with European Data Protection Law' (2014) International Data Privacy Law 4: 250-61.

The challenges of big data analytics and ambient computing

7. At a more fundamental level, the fitness for purpose of these core principles in an era of big data and ambient computing has become a hotly contested issue. Big data's value comes from the use of algorithms to find unexpected 'hidden insight' in large data sets: hence it is antithetical to the twin core ideas of data minimisation and purpose specification. Big data is all about data *maximisation* and finding *new unexpected purposes* for which that data might be mined to generate insight[26]. Although several critics suggest that the Fair Information Principles can be revised and reformulated to fit the contemporary age,[27] others argue that we should abandon altogether and start afresh,[28] whilst still others claim that data protection law should take an entirely different tack, eschewing restrictions on data collection, and focusing instead on enhanced obligations of transparency and accountability in relation to data use.[29]

The rise of the algorithm and machine intelligence

8. The capacity of big data analytics to generate hidden insight and detect correlations between data points that were not previously identifiable can be understood as establishing a new mode of knowledge production[30] and helps explains the rapid take up of these technologies in many realms of academic inquiry. However, much of the excitement from within commerce, government and academia associated with big data can be attributed to the possibility of feeding machine learning algorithms with data sets that are automatically collected, captured and stored on cloud servers in real time by networked digital sensor technologies that are increasingly embedded into the fabric of everyday life. It is through the combined operation of these technological systems that it becomes possible to harness these algorithmic decision-making systems in order to generate predictions concerning *future* behaviour that are claimed to provide 'actionable insight', and which have been described as the 'holy grail' of big data. It is the capacity to produce actionable insight on a population-wide scale arising from the capacity of digital networked interventions to operate on a one-to-many and yet highly personalised basis that provides the technological foundations that underpin the success of giant digital platform providers such as Google, Facebook, Uber, and Amazon and which many digital entrepreneurs wish to emulate. At the same time, the potential to harness these algorithmic systems to achieve greater efficiency and precision has not been lost on government, with the increasing use of these systems to facilitate decision-making in the realm of homeland security, law enforcement, and fraud detection. Because the data is gathered automatically from myriad ubiquitous sensors embedded into the environment, algorithmic systems appear to provide objective evidence of behaviour (i.e. it is portrayed as 'game proof'), whilst avoiding the vagaries and imperfection of human judgment.[31]

26 Mantalero, Alessandro, 'The Future of Consumer Data Protection in the EU: Re-Thinking the 'Notice and Consent' Paradigm in the New Era of Predictive Analytics' (2014) *Computer Law and Security Report* 30: 643-60.

27 Ball, Kirstie, David Lyon, David Murakami Wood, Clive Norris, and Charles Raab. *A Report on the Surveillance Society*, September 2006.

28 Koops, supra n.3.

29 Mayer-Schonberger, Viktor and Kenneth Cukier, *Big Data,* (2013) London: John Murray.

30 boyd, d. , and Kate Crawford, 'Critical Questions for Big Data,' *Information, Communication and Society* (2012) 15: 662-7.

31 Yeung, Karen, 'Algorithmic Regulation and Intelligent Enforcement,' In M Lodge (ed) *Regulatory Scholarship in Crisis?* Centre for Analysis of Risk and Regulation, LSE London, September 2016, forthcoming.

Calls to secure algorithmic accountability

9. Yet, the emergence of powerful algorithmic decision-making systems has been accompanied by increasing recognition that the rise of 'algorithmic power' can have serious adverse effects on an individual's critical life opportunities. For example, Bill Davidow, writing in *The Atlantic* in 2014, claims that big data analytics are being employed by both corporations and governments to create highly granular profiles upon which many consequential decisions are being made, such that millions of people are now virtually incarcerated in 'algorithmic prisons'. He notes that many people cannot find employment due to their internet profiles, or have difficulty purchasing or accessing a range of services, ranging from the purchase of insurance, accessing loan finance, purchasing property, renting a house or boarding an airplane. Yet, few of us are aware that these automated algorithmic processes are at work, and of the tangible harm that they can produce. Moreover, as Davidow points out, even if we do – we may not know who our jailer is or even the type of behaviour that condemned us in the first place[32].

10. At the same time, critics point out that the promise of objective knowledge associated with algorithmic systems and their aura of infallibility is misplaced, given that subjective human judgment is unavoidable in the construction of these technological systems, and that errors can occur at every phase of the algorithmic cycle, from the collection of data, the construction of models, the development of algorithms and the interpretation of their results[33]. Accordingly, there is increasing recognition of the need to secure 'algorithmic accountability', and it is this challenge which data governance regimes in the 21st century must now reckon with. As the World Economic Forum put it:

> "...in many ways 'the world is now post-digital' and... the discourse is no longer about technology but how it is applied for socioeconomic change. Instead the focus is on a new nexus of control and influence: the algorithm ... As the socioeconomic impact of predictive machine learning and algorithms grows stronger, long-term concerns are emerging on the concentrated set of stakeholders (who both mediate communications and have access to powerful algorithms) and their influence over individuals. The focal point of these conversations centers on how data can be potentially abused to proactively anticipate, persuade and manipulate individuals and markets[34]...
>
> Summing up, the World Economic Forum observed that
>
> "...there is a crisis in trust. Concerns are voiced from a variety of viewpoints at a variety of scales. Industry, government and civil society are all uncertain on how to create a personal data ecosystem that is adaptive, reliable, trustworthy and fair. The shared anxieties stem from the overwhelming challenge of transitioning into a hyperconnected world. The growth of data, the sophistication of ubiquitous computing and the borderless flow of data are all outstripping the ability to effectively govern on a global basis. We need the means to effectively uphold fundamental principles in ways fit for today's world..." [35]

11. Identifying mechanisms, institutions and principles that can provide meaningful and effective algorithmic accountability are anything but simple and straightforward. Data governance now implicates complex, value laden questions and, in the words of the World Economic Forum, "give rise to some fundamental social choices. Questions about individual autonomy, the sovereignty of individuals, digital human rights, equitable distribution and free will are all part of these conversations. There are no easy answers."[36]

32 Davidow, Bill, 'Welcome to Algorithmic Prison – the Use of Big Data to Profile Citizens Is Subtly, Silently Constraining Freedom,' *The Atlantic*, 20 February 2014.

33 Krasnow Waterman, K., and Paula J Bruening, 'Big Data Analytics: Risks and Responsibilities' (2014) *International Data Privacy Law,* 4: 89-95.

34 World Economic Forum (in collaboration with A.T. Kearney) (2014) *Rethinking Personal Data: A New Lens for Strengthening Trust.*

35 Ibid.

36 Ibid.

12. The inescapably value-laden and contestable nature of the tasks associated with constructing effective data governance regimes in a world of big data and ambient computing can be illustrated by considering the scope, content and distribution of rights and obligations in relation to the following types of personal digital data[37]:

    - **Volunteered data:** refers to data that is volunteered by individuals when they explicitly share information about themselves (eg create a social network profile or enter credit card data for online purchasing) or when 'compelled' to share through governments or commercial entities – the critical feature being the data subject's awareness of the action she is taking, which is often transactional;
    - **Observed data** are behavioural data captured by the counterparty or a third party that is capable of registering machine readable behaviours that are mostly intuitive and unconscious, such as web-surf behaviour, click stream behaviour or behavioural biometrics (gait, eye movements), mobility data, the history of social networking interactions and so forth. This data is typically captured passively via machine to machine transactions and often surreptitiously without implicated individuals typically being aware of it; and
    - **Inferred data** refers to inferences (ie the patterns, clusters and correlations detected in databases or in streaming databases identified via machine learning algorithmic processing of data from multiple data sets) that form the basis of predictions. Inferred data is fabricated by organisations that mine databases in order to detect nontrivial patterns in the data that can be used to categorise persons or other entities that match those patterns.

13. Although we may believe that volunteered data typically involves a deeper sense of unique ownership, even when that data is often more 'by me' than 'about me' (such as one's emails, biographical details and photos) to which individuals often have strong emotional ties such that the individual who generated that data should be the primary rights-holder in relation to such data. In practice, data of this kind is often stored by commercial providers in cloud storage systems who are likely to assert concomitant rights associated with such data. In relation to observational data that pertain to a specific, identifiable individuals, the party capturing this data may seek to claim entitlements to that data based on its ownership and operation of the technological systems which captured and collected that data and in which it can assert a legitimate interest. Finally, in relation to the all-important inferred data (that which forms the basis of 'actionable insight'), it is the organisations that generate this data from the development and implementation of sophisticated algorithmic systems that are inclined to assert superior claims to such data, even vis-à-vis the individual to whom those inferences relate, often asserting that they are protected by trade secrets or other IP rights and which they are free to use or transfer as any other legitimate property owner.

37 Ibid.

14. By drawing attention to the competing interests and claims in relation to these different types of personal data – it becomes readily apparent that defining and distributing rights and obligations in data will inevitably generate conflict between competing interests. These conflicts are inherent in the nature of personal digital data and the technological manner in which it is collected, to say nothing of the dilemmas associated with, for example, sector and data-specific challenges. These include determining how we should govern highly sensitive personal data (such as an individual's medical and health records) that, if aggregated on a population-wide basis, may have tremendous potential to transform medical research and treatment. At the same time, the personal data eco-system is dynamic, sprawling, complex and extremely fluid, yet the potentially permanent nature of digital data makes it extremely difficult if not practically impossible to acquire a clear view of the flow of data and the uses to which it is or may be put in future.

15. In responding to the challenges of data governance in a networked, increasingly data-driven society, a wide variety of governance tools and mechanisms have been proposed, including but not limited to:

    - risk assessment (including privacy impact assessments, data protection impact assessments and surveillance impact assessments);
    - technological protection mechanisms (privacy enhancing technologies of various kinds, personal data vaults, accountability by design, 'smart' data technologies to facilitate the creation of audit trails, profiling transparency by design, discrimination discovery and prevention by design);
    - oversight models based on the legal and ethical governance of research on human subjects (such as institutional review boards in the US and equivalents elsewhere); and
    - scrutiny, certification and auditing by expert 'algorithmists.'

    While all of these proposals warrant serious consideration, they have largely been proffered on a piecemeal basis[38] so that, at least in my view, one cannot yet meaningfully speak of comprehensive alternative 'models' of data governance.

38 One notable exception is the 'New Deal on Data' propounded by Greenwood, Daniel, Arkadiusz Stopczynski, Brian Sweat, Thomas Handjono, and Alex Pentland, 'The New Deal on Data: A Framework for Institutional Controls,' in Julia Lane, Victoria Stodder, Stefan Bender and Helen Nissenbaum (eds), *Privacy, Big Data and the Public Good,* (2014) 192-207. New York: Cambridge University Press.

16. It is worth noting, however, that various commentators (largely based in the USA) have advocated a combined legal-technological approach to data governance that entail the use of 'personal data management services' which are grounded on a model of data propertisation in which individuals would acquire transferable property rights in their personal data that would be technologically captured and stored within each individual's 'personal digital data vault', enabling individuals to sell, share, donate or otherwise dispose of their personal data at will.[39] However, the larger institutional governance framework within which such propertisation models are contemplated vary considerably[40], and they appear to me to focus primarily on 'volunteered' data, with little clarity concerning in the distribution of property rights in relation to observed and inferred data.

The commodification of personal data and the death of democracy

17. Although I am yet to form a firm view about these propertisation approaches to data governance, my instinct is one of deep scepticism. In particular, propertisation approaches to data governance are grounded in an implicit assumption that privacy is purely a matter of individual concern, failing to recognise that vital collective interests are at stake. Privacy also refers to a zone of protection around each individual's activities within a society that makes possible the capacity for individual flourishing and self-creation that allows us to play around with who we are, with whom we wish to relate and on what terms, and in which our sense of self and our individuality can emerge, mutate and stabilise.[41] Yet data governance models that treat data as a marketised commodity that can be traded at will to the highest bidder fail to recognise that the *privacy commons* is a vital element of the critical moral, political and social infrastructure that is vital to human flourishing and democratic freedom. It is this privacy commons that is at risk of continual erosion from the emergence of large-scale ambient computing systems which, ultimately, rests on an infrastructure that is intended to facilitate mass, population wide, and continuous micro-level surveillance.

39 World Economic Forum, *Personal Data: The Emergence of a New Asset Class* 2011.

40 Compare for example, World Economic Forum, *Personal Data: The Emergence of a New Asset Class* 2011; Searls, Doc, *The Intention Economy - When Customers Take Charge* (2012) Boston: Harvard Business Review Press; Lanier, Jaron, *Who Owns the Future?* (2013) London: Penguin Books; Lessig, Lawrence, *Code and Other Laws of Cyberspace* (1999) New York: Basic Books.

41 Cohen, Julie E. *Configuring the Networked Self* (2012) New Haven: Yale University Press.

18. In other words, it is not just our individual privacy, but the collective foundations and health of the democratic order that is at stake, and which forms the basis upon which our capacity for individual flourishing and democratic participation depends. As Evgeny Morozov puts it 'as Web companies and government agencies analyse ever more information about our lives, it's tempting to respond by passing new privacy laws or creating mechanisms that pay us for our data. Instead, we need a civic solution, because democracy is at risk.'[42] In other words, data governance debates are often narrowly focused on matters of individual privacy and security or, worse, technical matters best left to the engineers who design and implement our digital infrastructure and the software which powers it. As Morozov has persuasively argued, despite the conventional portrayal of networked computational techniques as value-free user-friendly tools that offer to make every aspect of lives, from the management of our personal health through to the delivery of our governmental services, more efficient and convenient, these algorithmic decision-making systems rest on a largely hidden set of ideological premises concerning the distribution of power and authority in society and the relationship between citizens, the state and the market which he refers to collectively as 'solutionism'[43]. In a 21st century, networked environment in which digital sensor technology is embedded into the very fabric of our lives and from which we cannot reasonably opt out, data governance becomes inescapably political 'all the way down'. Accordingly, it is vital that we find ways to ensure that the public understands what is at stake, can participate more fully in identifying how our digital data is governed and that the rights and responsibilities in relation to our personal data are fairly and legitimately allocated and distributed. As Alessandro Acquisti, a leading behavioural economist and privacy scholar puts it, 'one of the defining fights of our times will be the fight for the control over personal information, the fight over whether big data will become a force for freedom, rather than a force which will hiddenly manipulate us'[44]. In my view, if left unchecked and unregulated, the emerging 'digital data barons' will surely build a digital infrastructure and personal data ecosystem that will enable them to reap gargantuan profits and/or wield enormous power, and which cannot but weaken the foundations of our democracy and our capacity for individual freedom and flourishing.

42 Morozov, Evgeny, 'The Real Privacy Problem' (2013) MIT Technology Review available at https://www.technologyreview.com/s/520426/the-real-privacy-problem/

43 Morozov, Evgeny, 'The Rise of Data and the Death of Politics' *The Guardian*, 20 July 2014; Morozov, Evgeny, *To Save Everything, Click Here* (2013) London: Penguin Group.

44 Acquisti, Alessandro, 'What Will a Future without Secrets Look Like?' (October 2013) available at https://www.ted.com/talks/alessandro_acquisti_why_privacy_matters/transcript?language=en.

How to Govern: Cryptocurrencies and Police Robots

Kay Firth-Butterfield – Lucid Ethics Advisory Panel

Introduction

This submission started as a series of hypothetical concerns about current and upcoming AI issues but after two recent incidents of note became more concrete examples of the ways in which AI technology might be governed in the absence of legal regulation and what type of legal regulation might be needed. It is important to think outside the legal governance of the technology because the latter is growing at an exponential rate and lawmakers will find it difficult to keep up with the legislation required. Therefore, as well as legislation it behooves those of us working in the AI space to behave as a mature industry and look at ways of self-regulation.

The following two examples will be used as a way of examining aspects of self-regulation and legislative regulation in this space:

- The problem with DAO (decentralised autonomous organisation), cryptocurrencies and money-laundering
- The use of a robot for killing by the Dallas Police

As a start, building trust between those of us working in the AI community and the public needs to be undertaken:

- We need to establish/re-establish trust between our community and the public. The media hype does the technology a great disservice and we need to show that we understand the risks and will act appropriately whilst also ensuring the benefits. Failure to do so risks ill-informed knee jerk reactions to our technology similar to the anti-vaccine campaign
- After Brexit we can see that our highly divided politics in the western world suggests trust in institutions (which includes governments, experts, organisations like the UN and companies) is missing from large parts of the general public. Misinformation from, or misunderstanding of the media seems to be fundamental to this problem and, as most articles in the general press start with a picture of 'the terminator', it certainly seems likely to happen to AI.

DAO

The DAO is a cryptocurrency investment company operating in, what seems to be, a self-interested self-regulating community. AI, the Blockchain and Virtual currencies are a combination that not many people are thinking about but have the potential to cause great disruption in the financial space on money and money laundering. This is particularly so when combined with each other and with the ability to reason across the huge data sets of 'Big Data' using ever more powerful computers. This will upend how we think about banking, money and the regulation of money, including money laundering.

Virtual Currencies

Virtual vs real currency:

There are a number of differences between fiat and virtual currencies. These include:

- No central governing system based in a physical territory. The Bitcoin "protocol established a set of rules – in the form of distributed computations – that ensured the integrity of the data exchanged among these billions of devices without going through a trusted third-party."[45]
- So, in addition to not being controlled by countries, cryptocurrencies also do not require physical banking agencies; indeed, in many cases there are no intermediaries to virtual currency trading. Banks, or Wallets as they are called, exist in only in cyberspace.
- Mandates for anonymous distribution of the currency.
- 'Coinage' which can be divided into fractions of the original 'coin'. By way of example, a Bitcoin "... is divisible to the eight decimal. The smallest portion of bitcoin has its own name: satoshi, whereas 1 BTC = 10^8 satoshis = 100,000,000 satoshis".[46]

Are virtual currencies money?

Economists are divided on this issue but in September 2015, Bank of England economist and executive director for monetary analysis and statistics, Andrew Haldane suggested that the Bank of England might issue a State backed cryptocurrency similar to Bitcoin. In his speech Haldane commented:

"In its short life, Bitcoin has emerged as a monetary enigma. It divides opinion like nothing else (for example, Yermack (2013), Shin (2015)). Some countries have banned its use. Others have encouraged it. Some economists have denounced it as monetary snake oil. Others have proclaimed it a monetary cure-all for the sins of the state. What I think is now reasonably clear is that the distributed payment technology embodied in Bitcoin has real potential. On the face of it, it solves a deep problem in monetary economics: how to establish trust – the essence of money – in a distributed network. Bitcoin's 'blockchain' technology appears to offer an imaginative solution to that distributed trust problem (Ali, Barrdear, Clews and Southgate (2014).

Whether a variant of this technology could support central bank-issued digital currency is very much an open question. So too is whether the public would accept it as a substitute for paper currency. Central bank-issued digital currency raises big logistical and behavioural questions too. How would it work practically? What security and privacy risks would it raise? And how would public and privately-issued monies interact?

These questions do not have easy answers. That is why work on central bank-issued digital currencies forms a core part of the Bank's current research agenda (Bank of England (2015). Although the hurdles to implementation are high, so too is the potential prize if the zero lower bound (ZLB) constraint could be slackened. Perhaps central bank money is ripe for its own great technological leap forward, prompted by the pressing demands of the ZLB."[47]

In a speech Mark Carney, Governor of the Bank of England, was to have given in June 2016 at the Lord Mayor's Banquet in London he wrote that the Bank of England was exploring distributed ledger technology (DL) "In the extreme, a DL for everyone could open the possibility of creating a central bank digital currency. On some levels this is appealing. For example, it would mean people have direct access to the ultimate risk-free asset. In its extreme form, it could fundamentally and perhaps abruptly re-shape banking."[48] The Bank of England is not the only central bank considering its options when it comes to the use of blockchain but its immediate use of the technology would probably be to make safer, quicker and more transparent movements of money rather than the establishment of a 'crypto pound sterling'.

45 Don Tapscott and Alex Tapscott, *Blockchain Revolution: How the Technology behind Bitcoin Is Changing Money, Business, and the World* (New York: Portfolio / Penguin, 2016) page 5.

46 "Transactions – How Does Bitcoin Divisibility Work? – Bitcoin Stack Exchange," accessed June 30, 2016, http://bitcoin.stackexchange.com/questions/13157/how-does-bitcoin-divisibility-work.

47 "How Low Can You Go? – Speech by Andrew Haldane | Bank of England," accessed July 1, 2016, http://www.bankofengland.co.uk/publications/Pages/speeches/2015/840.aspx.

48 Mark Carney, "Enabling the FinTech Transformation: Revolution, Restoration, or Reformation?" (Mansion House, London, U.K., June 17, 2016).

Types of virtual money (Cryptocurrencies)

The most famous virtual currency is Bitcoin, perhaps because of its infamous use by 'Dread Pirate Roberts' in his 'Silk Road' business, but more likely because of it being the first of the non-gaming cryptocurrencies.[49] However, there are many similar virtual currencies including Etherium (currency is called Ether), Ripple, Litecoin and NEM, of which Bitcoin and Etherium have the largest share of the market with a combined market cap of over $11bn. In July 2016 CoinMarketCap.com listed 652 virtual currencies being traded with a combined value of $12,477,481,881.[50] As Clare Chambers-Jones points out, before Bitcoin the virtual currency was a creation of gamer and virtual lifestyle websites such as Second Life.[51] These were enabled by the development of technology like PayPal and virtual ATMs and became a safe haven for virtual money laundering.[52]

Bitcoin was invented by a person, or persons, known as Satoshi Nakamoto.[53] It enabled, as the Tapscotts eulogise, "...trusted transactions directly between two or more parties, authenticated by mass collaboration and powered by collective self-interests, rather than by large corporations motivated by profit."[54] However, it could equally be said that in this case the control of money had simply moved into the hands of those computer enthusiasts who took the time and used the necessary energy to mine Bitcoins, which is now something we see in the DAO problem. In June 2016 a hacker stole some $55m (3.6m Ether) from the cryptocurrency investment company using 'Ether' called DAO. "Like bitcoin, ether relies on a 'blockchain' – a public ledger, distributed among lots of the system's users, which records all transactions. Bitcoin's blockchain handles mainly financial transactions, but ether's can run computer code, including self-executing 'smart contracts', like those underpinning the DAO."[55]

At the time of the theft the 'miners' who control the currency were divided as to whether a way should be found to prevent the theft or whether the integrity of the coinage depended upon not being able to make changes to the Blockhain. There were those who felt it would be a more criminal action to alter the Blockchain on which Ether exists than the actual theft itself.

Therefore, it is interesting to watch how the debate developed because we see a highly disparate community choosing its morals and whether and how to self-regulate. After discussion about whether to do something changed into a decision to do something the community then had to agree on the tools necessary to 'do something'.

Initially this has been done by using a 'soft fork' which was a backwards compatible change in the code to prevent the 'stolen' money being spent. The proposed soft fork showed bugs and so the community abandoned that idea and chose the construction of a 'hard fork' which would leave anyone using the old technology cut out of using the newly coded changed technology.[56] There is now pressure on those in the community to apply the 'hard fork' change before the 21st July after which it will become more difficult.

---

49 Andy Greenberg, "Silk Road Creator Ross Ulbricht Sentenced to Life in Prison," *WIRED*, May 29, 2015, https://www.wired.com/2015/05/silk-road-creator-ross-ulbricht-sentenced-life-prison/.

50 "All Currencies | Crypto-Currency Market Capitalizations," accessed July 12, 2016, http://coinmarketcap.com/currencies/views/all/.

51 Clare Chambers-Jones, *Virtual Economies and Financial Crime: Money Laundering in Cyberspace* (Cheltenham, UK ; Northampton, MA: Edward Elgar, 2012).

52 Ibid.

53 Satoshi Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System" (Bitcoin.org, October 2008).

54 Tapscott and Tapscott, *Blockchain Revolution*. p. 5

55 "Theft Is Property," *The Economist*, June 25, 2016, http://www.economist.com/news/finance-and-economics/21701136-cyber-attacker-outsmarts-smart-contract-theft-property.

56 Russell Brandom, "Time Is Running out to Stop a $53 Million Cryptocurrency Heist," *The Verge*, June 30, 2016, http://www.theverge.com/2016/6/30/12068258/ethereum-dao-heist-theft-cryptocurrency-53-million.

Thus, it can be seen that these currencies do have some 'central control', which was also seen in the Mt. Gox bankruptcy. Of course, there is considerable debate about whether it is possible to steal a virtual currency which may or may not be property and this change does nothing to resolve that debate.[57] However, in summary, what can be observed in this vignette into control of this new technology is:

a. A hotly contested debate about whether there should be self-regulation by a community of individuals opposed in principle to centralised money systems. (Many of these users understood virtual currency to be free from any centralised mechanism of supervision or enforcement and a near failsafe for the protection of anonymity.)

b. However, once debate started about how to retrieve the money the community came together to find the right way to 'govern' the breach. They rejected the soft fork, which nearly broke the system and voted some 97% to use the hard fork and attempt to return the stolen money to the DAO.

c. Previously, efforts by the community to regulate itself had been judged insufficient by governments and international communities. This gave rise to various groupings of cryptocurreny users, for example the Bitcoin Foundation which has an Education Committee which advocates, for example on freedom to innovate in California, 'we deserve the freedom to be innovators (without state sanctions) both in the making of new decentralised systems as well as in our daily actions and exploration of new systems.'[58] Also, the virtual currency industry established a European Digital Currency & Blockchain Technology Forum, in time for consideration of the topic by the European Parliament in April 2016.[59]

d. So, is this wild west tech showing us that it can self-regulate? Kant would be proud!

What this shows to me is that the AI community should also come together to agree whether it is necessary to do something to self-regulate and then chose, together, what that form or forms might be.

Blockchain

There is often confusion between the Blockchain and Bitcoin, this is because the concept of blockchain was invented to enable bitcoin to become an assured currency. However, blockchains are simply ways of ensuring security through distributed networks and so can be used to safeguard and date stamp all sorts of legal documents, make smart contracts (signatures are fully verified and timestamped by the distributed ledger network) and private data, for example health records. 'Smart contracts' are small pieces of computer code which require the computer to do X if Y happens.[60] For example, in a simple commercial contract – if X gives £100 then send to X Y dress". IBM "has contributed code to the [Hyperledger] project to help developers easily build secure distributed ledgers that can be used to exchange most anything of value."[61] IBM sees their technology ADEPT as a way of using the blockchain to allow devices within products connected in the Internet of Things to make contracts with one another.[62] Businesses can use blockchain to increase transparency into their business by allowing the public access or by allowing certain individuals with necessary 'keys' to access data, this latter is the more common form used currently.

57 "Theft Is Property."

58 http://www.coindesk.com/californias-bitcoin-bill-shelved-by-state-senator/

59 http://edcab.eu/press/new-voice-of-the-european-virtual-currency-and-blockchain-industry-edcab-set-up-to-lead-engagement-with-policymakers

60 Antony Lewis 2015, "A Gentle Introduction to Smart Contracts," *Bits on Blocks*, February 1, 2016, https://bitsonblocks.net/2016/02/01/a-gentle-introduction-to-smart-contracts/.

61 "IBM Blockchain," February 11, 2016, https://www.ibm.com/blockchain/.

62 "IBM - Device Democracy – United States," May 2, 2016, http://www-935.ibm.com/services/us/gbs/thoughtleadership/internetofthings/; "IBM Reveals Proof of Concept for Blockchain-Powered Internet of Things," *CoinDesk*, January 17, 2015, http://www.coindesk.com/ibm-reveals-proof-concept-blockchain-powered-internet-things/.

Using cryptocurrencies to launder money?

There is no doubt that cryptocurrencies provide an anonymity to the user which would, on the face of it appear to be very enticing to the criminal wishing to launder the proceeds of crime. In 2012, the FBI published an analysis of the risks of criminal activity using Bitcoin as follows: "The FBI assess with low confidence, based on current user and vendor acceptance that malicious actors will exploit bitcoin to launder money. This assessment is based on observed criminal activities, investigations, and prosecutions of individuals exploiting other virtual currencies, such as e-Gold and WebMoney. A lack of current reporting specific to Bitcoin restricts the confidence level."[63]

Clearly these currencies cannot be regulated like fiat countries because, whilst there are businesses (such as Coinbase) which trade the currencies, much of the trading happens purely online without the intervention of a third party. The Tapscotts suggest that the blockchain technology could discourage rather than encourage criminals because they have to 'publish all their bitcoin transactions in the blockchain' this, they argue, makes the dealings more transparent to law enforcement.[64] However, the anonymity of the bitcoin owner was guaranteed in the technology as detailed in Nakamoto's paper[65]

"The necessity to announce all transactions publicly precludes this method, but privacy can still be maintained by breaking the flow of information in another place: by keeping public keys anonymous. The public can see that someone is sending an amount to someone else, but without information linking the transaction to anyone. This is similar to the level of information released by stock exchanges, where the time and size of individual trades, the 'tape', is made public, but without telling who the parties were. As an additional firewall, a new key pair should be used for each transaction to keep them from being linked to a common owner. Some linking is still unavoidable with multi-input transactions, which necessarily reveal that their inputs were owned by the same owner. The risk is that if the owner of a key is revealed, linking could reveal other transactions that belonged to the same owner."

So the problem for authorities seeking to find money laundering amongst the many public, but anonymised, transactions of any cryptocurrency is finding the occasional linking which might destroy anonymity. Things are further complicated by a host of ways in which cryptocurrencies can be used in the real world or exchanged for fiat money. By way of example, the Bitcoin debit card which works in all countries without any ID required and can be used interchangeably with a PayPal account. [66] Additionally, the website BitLaunder is one of many which offers bitcoin tumbling because bitcoin is not sufficiently anonymous.[67] Therefore, these websites offer an anonymising process for their clients' use of bitcoin by breaking down the connection between the sender's address and the recipient's address. Likewise, use of this cryptocurrency within the DarkWeb using the Tor Browser adds levels of complexity to the laundering of money which cannot be undone.[68]

Obviously, the legal challenges associated with cryptocurrency and money laundering are manifold. As can be seen below, some jurisdictions are attempting to make sense of this growing phenomena however the following represent only some of the obvious questions arising in this area of law:

- questions of jurisdiction.
- the point at which the virtual currency actually becomes a fiat currency.
- whether even if third parties are regulated, the actual 'miners' who control the distributed ledgers can be regulated.
- whether there should be acceptance of these currencies as real currencies with the 'miner' as the 'central bank' in which case it might be possible to place regulatory obligations upon them.

63 FBI, "Bitcoin Virtual Currency: Unique Features Present Distinct Challenges for," 2012

64 Tapscott and Tapscott, *Blockchain Revolution*.

65 Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System."

66 "Bitcoin to Paypal," accessed June 30, 2016, https://bitplastic.com/bitcoin-to-paypal-exchange.

67 "Launder Bitcoin," accessed July 2, 2016, https://bitlaunder.com/.

68 The Tor Project Inc, "Tor Project: Anonymity Online," accessed July 2, 2016, https://www.torproject.org/.

Interaction between AI and cryptocurrencies

It is often said that there is no cyber technology which cannot be hacked and the theft from DAO certainly proves the theory against a technology which was believed to be extremely secure. Whether it is today's 'narrow' AI or tomorrow's strong AI, an AI system never sleeps and works faster and more efficiently than multitudes of humans. Machine learning systems can be trained to 'follow the money' and semantic AI is used to make human-like connections of evidence. As AI becomes more powerful there will be fewer places for criminals to hide money anywhere that an AI cannot pierce. Take, for example, an AI application developed to assist companies report their efforts to check supply chains for human trafficking under the Modern Slavery Act 2015; such an application will chase down contracts, read newspapers, police and other reports, work with materials supplied by NGOs and IGOs and any other material which might be relevant to its researches of a customers' supply chain. Once the crime has been revealed it is much easier for the AI to follow the money. However, whilst the exemplar is very laudable, an AI which searches for ways of looking within cryptocurrency wallets and works out where money has come from and is going to would be invasive of the privacy of law abiding individuals as well as those who are not. Society will need to decide what protections citizens, criminals and the state deserve vis a vis one another as these and similar technologies outstrip the ability of the way we currently think of our laws to regulate them and give justice.

Dallas Police – "robot used to kill"

On 8th July 2016, the Dallas police used a bomb disposal robot to kill as suspect. One can only imagine the extremis which was being felt by the officers seeing their fellows gunned down but this seems to be the first known use of re-purposing a robot to be a lethal, albeit non-autonomous, machine. In a discussion about AI it could be said to be of no importance. However, I think it raises a number of serious questions:

a. Should civilian officers be allowed to use the technology in this way?

   Elizabeth Joh (UC Davis) said she was worried that the decision by police to use robots to end lives had been arrived at far too casually. "Lethally armed police robots raise all sorts of new legal, ethical, and technical questions we haven't decided upon in any systematic way," she said. "Under federal constitutional law, excessive-force claims against the police are governed by the fourth amendment. But we typically examine deadly force by the police in terms of an immediate threat to the officer or others. It's not clear how we should apply that if the threat is to a robot – and the police may be far away." "In other words, I don't think we have a framework for deciding objectively reasonable robotic force. And we need to develop regulations and policies now, because this surely won't be the last instance we see police robots." Of course, the alternative argument is that a rifle shot by a Policeman is equally killing from a distance using a tool. Is the robot just a tool? Is a robot powered by AI a tool if it doesn't have autonomy?

b. Whilst this could be said to be a slippery slope argument, we know that bomb disposal robots came into civilian use from the military sector; if lethal autonomous weapons are permitted in military service it seems inevitable that there will be creep into civilian use. Why put police officers at risk when a robot could be used.

c. If re-purposing of this technology is deemed appropriate what is to stop the police similarly repurposing other technology, and particularly, for our discussion I will mention three, which with AI provide for a type of predictive policing:

    i. Mining Twitter and other feeds to forecast unrest

    ii. CCTV and face recognition

    iii. Internet of Things – cars, home robots *et al*

    i. On the 10th July, 2016, the Washington Post reported an application created by a hacker called the Jester which "Using IBM's Watson AI, ... not only examines large collections of tweets but – somewhat eerily – also can go through a single user's [Twitter] timeline and, with Watson's machine learning technology, offer an analysis of that user's 'trustworthiness, propensity toward violence [and] openness,' the Jester said. That information, he said, could hold clues to a criminal's intentions."[69] This raises another debate about whether the Open Source AI movement or indeed the selling of AI to anyone is adding to the control problems which we seek to consider here or decreasing them. Sadly, I do not have time to consider them further here.

    ii. Predictive policing through networked facial recognition/CCTV and AI. My example is something our AI could already do if given access to the data; it raises a number of issues from privacy to protection of the public by government. The example posits a small amalgamation of data – access to credit card transactions, access to CCTV, access to the network on which the ignition of X's car is located and facial recognition. It is an example from a tragic event at SXSW in Austin where a drunk driver killed 6 people. Consider this scenario. It is known, through credit card data, that X has bought (and through CCTV been seen to consume) many drinks. After leaving the Bar X is caught on CCTV falling over and trying to get into his car. Once in the car he is seen on CCTV to drive into another car in the car park. The question, then, is whether we use the AI, which has gathered this data, to shut down his car before he causes an accident or to decide when he should be arrested. We know, though various hacks that it is possible to stop a car engine in a networked car and government could partner with car manufacturers to do just that because it is perceived to be seen as in the public's best interests. However, surely, that perception is a huge question.

In fact, it is not necessary to have the credit card data in this example. A human being would infer from a person walking erratically, falling down and driving the car erratically that the person was drunk or drugged or at least ill. They might even ask for the keys to prevent an accident. All of that could be inferred by AI just from the CCTV footage. In 5-10years we may have self-driving cars and so this example may become irrelevant so we should think about whether police should use AI to scan for 'deviant' behaviour, such as street crimes like pickpocketing or larger issues like bank robbery, by way of example, the is one major US bank which has a branch robbed every day. Should AI be allowed to examine the behavior of each person coming towards the bank for 'telltale' signs of deviant behavior and deny that person entry as a result? The owner of the Bank could do that but the question for government is whether there would be sufficient evidence of criminal intent to be inferred for the commission of an inchoate crime?

69 "How Artificial Intelligence Could Help Warn Us of Another Dallas," *Washington Post*, accessed July 12, 2016, https://www.washingtonpost.com/news/the-switch/wp/2016/07/10/how-artificial-intelligence-could-help-warn-us-of-another-dallas/.

iii. The Internet of Things: the implications for the compact between government and citizen. In a world where Government can use AI to gather and discover huge volumes or all data on its population – what then would be the government/citizen compact? How can personal privacy be protected in a world where all citizens' activities can be monitored via pervasive CCTV and other big-data electronic surveillances and then dynamically distilled and understood by an increasingly knowledgeable AI which derives its citizen-oriented intelligence from the big data being collected by the government? Professor Cheung has suggested that we should use the Data and Society Research Institute's four dimensions of transparency to evaluate new technology. A conversation about the technology is advocated. However, it seems to me that this technology is becoming more able at a faster pace than a citizen which is largely ignorant of it can have true awareness. Take robots in the home which as Ryan Calo says allows surveillance into a hitherto private and protected space. This may be good for the victim of domestic violence but generally Calo suggests it could "operate to dampen constitutional privacy guarantees by shifting citizen expectations." Indeed, where the robot owner voluntarily permits the robot to send data to third parties they lose all or part of protection of that information under Fourth Amendment in the US; the third party doctrine.

Whilst the events in Dallas could be considered extreme we need laws and a regulatory body or bodies which address these situations before it is too late to go back.

Regulation and Self-Regulation

Case law will be able to accommodate some of the regulation needed in common law jurisdictions; it may be more difficult for civil law countries to adjust as swiftly.

There is a place for both regulation and self-regulation but as I said earlier, because the process of passing laws is so time consuming and so difficult the technologies continue to outpace any such attempts. Thus, those of us working in the AI industry need to develop a mature attitude and self-regulate.

Ethics Advisory Panel Model

At Lucid we have an Ethics Advisory Panel which allows me to talk about it and discuss ways forward in workshops such as this one. My EAP members are not paid and have access to all the company data they feel that they need. Deep Mind also have an ethics board and, for DeepMindHealth use Independent Reviewers. These people are not paid and are independent from the company but have access to all aspects of the company to review its actions in the healthcare space. In both examples, however, the public should be encouraged to look behind the statements at the actual power such boards have and, for example, ask about the Non-Disclosure Agreements Board members have signed and how they might whistle blow.

Company Law for the 21st century

Especially in the US, we are operating in a company law system with is based in the 19th century. We need governments to think about how companies of the 21st century might be created to owe duties to stakeholders as well as shareholders.

Likewise, as we had to re-think sustainability within our companies and created sustainability officers we could consider creating Chief Values Officers to help guide and direct development of the use of AI systems in the multiple fields in which they will be used.

Independent Think tanks and Associations

Organisations, such as the Royal Society, which have unparalleled integrity have a role to play in 'testing' by white paper and conversation the ways in which the use of AI is regulated by governments, industry and others. In some cases, it is only pressure from the public, informed by such independent sources, on government and companies which might bring about change. In this I am thinking about the green debate.

One such body of experts from different specialist areas who are interested in the ethical design of autonomous systems and who work for an independent organisation can be found at the IEEE. We are working with such experts who come from industry, academia, government and international organisations to create a Charter for Ethical Considerations in the Design of Autonomous Systems. Encouraging designers and creators of the systems to follow an ethical design approach is a form of internal self-regulation which currently seems to be missing from some of the AI products on the market today which range from endangering our lives to diminishing our privacy.

From the discussion above on cryptocurrencies and money laundering we can see that in certain cases the technology and the 'community' involved with it has probably gone beyond anything which can be externally regulated and so we need to find some way to work with such communities of hackers, miners and others at the cutting edge of creating technologies which invade hitherto highly regulated industries such as finance and law.

Government

Self-regulation can only take the governance of the myriad of ways in which AI will affect our lives so far.

Privacy:

Legislation is needed particularly in areas where there is mixing of personal privacy with AI. I have highlighted some above but one can also think of the amount of personal data which might be collected from a conversation and road map of a journey in an autonomous car.

Safety and AI:

Areas of safety and the use of AI is an area where the law needs to catch up fast. For example, in the US, the NHTSA which regulates the safety of cars, cannot create regulations but must wait for authorisation from Congress. In the absence of such authorisation it is reduced to inviting Google to apply for exemptions to its current standards which do not apply to autonomous cars rather than being able to actually regulate safety. A similar safety issue is the beta-testing of technology on human subjects exemplified by the recent Tesla crash.

Safety of AI:

I am uncertain of the type of legislation which might secure such an issue because of the international nature of AI companies and the fact that work on AI is increasingly dispersed. Changing company law as suggested might be some small help.

Use of AI technology by government:

As can be seen from the Dallas case and extra comments above there are major issues which require the government to self-police their use of the technology and some guidelines for such governance in the form of legislation would be helpful.

International Organisations

AI is an international technology, thus working on Treaties which may assist in working out ways of providing international governance is necessary.

All views expressed are my own and not those of my employers or other organisations with which I am affiliated.

The Royal Society is a self-governing Fellowship of many of the world's most distinguished scientists drawn from all areas of science, engineering, and medicine. The Society's fundamental purpose, as it has been since its foundation in 1660, is to recognise, promote, and support excellence in science and to encourage the development and use of science for the benefit of humanity.

The Society's strategic priorities emphasise its commitment to the highest quality science, to curiosity-driven research, and to the development and use of science for the benefit of society. These priorities are:

- Promoting science and its benefits
- Recognising excellence in science
- Supporting outstanding science
- Providing scientific advice for policy
- Fostering international and global cooperation
- Education and public engagement

For further information
The Royal Society
6 – 9 Carlton House Terrace
London SW1Y 5AG

T +44 20 7451 2500
W royalsociety.org

@royalsociety
@theroyalsociety

Registered Charity No 207043
Issued: November 2016 DES4610

The British Academy is the UK's national body for the humanities and social sciences – the study of peoples, cultures and societies, past, present and future. We have three principle roles: as an independent Fellowship of world-leading scholars and researchers; a Funding Body that supports the best ideas, nationally and internationally; and a Forum for debate and engagement – a voice that champions the humanities and social sciences.

For further information
The British Academy
10 – 11 Carlton House Terrace
London SW1Y 5AH

T +44 20 7969 5200
W britishacademy.ac.uk

@britac_news
@TheBritishAcademy
Britacfilm
BritishAcademy

Registered Charity No 233176

Cover image