Data ownership, rights and controls: Reaching a common understanding

Discussions at a British Academy, Royal Society and techUK seminar on 3 October 2018

The British Academy

THE ROYAL SOCIETY

techUK

Contents

1.0 Summary

On 3 October the British Academy, techUK and the Royal Society convened a seminar, which provided an opportunity to explore and understand the concept, value and limitations of the idea of 'data ownership'. It considered the sound bases from which to consider and probe the concept of data ownership and discussed issues relating to the ability to exert rights and control over data use.

Based on the discussion during the seminar the following key points have been identified as warranting further consideration and discussion moving forward:

- Use of the term "data ownership" raises significant challenges and may be unsuitable because data is not like property and other goods that can be owned or exchanged
- Instead discussion should explore the rights and controls individuals, groups and organisations have over data, and should encompass a societal as well as individual point of view
- Broader debate could help to better describe the data rights and controls that are often associated with the concept of 'data ownership'.

This paper summarises the rich and diverse discussion at the seminar, and is followed by a set of papers, which provide further explorations of data ownership, rights and controls.

Disclaimer

This is a note summarising the discussion and debate at the British Academy, Royal Society and techUK event on *Data ownership, rights and controls: reaching a common understanding*. It is not intended to represent the views of the British Academy, the Royal Society or techUK, nor does it represent the views of individual attendees of the event. The ideas and reflections contained within are not necessarily endorsed by the British Academy, Royal Society or techUK.

2.0 Introduction

In 2017, the British Academy and the Royal Society published Data management and use: Governance in the 21st Century. This report addressed a changing data landscape and recommended the need for a new governance framework for data use, based on the principle of human flourishing, and with a need for a new body to steward the landscape as a whole.

In the 18 months since the publication of this report, there have been significant changes in the data governance landscape. The General Data Protection Regulation (GDPR) came into force in May 2018, requiring data protection 'by design and by default'. 2018 saw the UK government establish the Centre for Data Ethics and Innovation, complete a consultation on the Centre's role and activities, reflecting many of the recommendations made in *Data management and use,* and appoint the Board members that will help to steer the work of the Centre going forward. Also in the last year, the Nuffield Foundation established the Ada Lovelace Institute, which has a mission to ensure data and AI work for people and society.

On 3 October 2018 the British Academy, the Royal Society and techUK held a seminar to reflect on these changes, and to explore a key issue within the broader conversation about data management and use: data ownership, rights and controls.

The *Data management and use* report argued that some of the core concepts that underpin governance of data use are challenged by current technologies and data management practices. These include consent and the idea of data ownership. The report states: 'Uncertainties around the concept of ownership can be a barrier to effective trade and transfer of data, and leave individuals and organisations uncertain about their rights.'

Ensuring that individuals and organisations are able to exercise the appropriate rights and controls over data is essential to the data economy and is at risk unless we build the right approach to data rights and control.

The seminar was therefore held to provide the opportunity to explore and understand what is meant when individuals and groups refer to 'owning' data or 'my' data, and to explore the concept, value and limitations of data ownership from individual and organisational perspectives, in both the private and public sectors. It considered the sound bases from which to consider and probe the concept of data ownership and discussed issues relating to the ability to exert rights and control over data use and assessing and accessing the value of data.

This report includes reflections of the discussion at the seminar, and a set of contributed papers. These include papers submitted ahead of the seminar to stimulate discussion and papers submitted after the seminar to expand and open up areas for further discussion.

3.0 Data ownership, rights and controls: discussions at a British Academy, Royal Society and techUK seminar on 3 October 2018

Where are we today? Data collection has grown massively, increasingly encroaching on private spaces. We have become very used to giving up our data, consenting to its gathering and to its use, often in ways we dimly appreciate, in exchange for 'free' access to products and services that we value. This can create a feeling of unease, and this feeling is amplified by high-profile stories about the mismanagement and misuse of personal data.[1]

This sense of unease is a difficult problem to pin down, especially as it is in tension with an awareness that data can do much good – for example in its use in health research. The data governance regime faces the challenge of recognising the enormous potential for public good, but also the potential for both very specific and general harm. It also has to deal with both individual data and the holders of large data sets, and it has to balance the interests of the individual and society.

The systems for governing data are clearly under stress along with the concepts we use to talk about data. Data ownership is one of these rather problematic concepts – what does it really mean? Can you own data? Data is replicable and is not something that you use up, potentially you can share it as much as you like. If you can own data, under what circumstances and who should be able to own it? Are there different considerations in the private and public sectors? What does it mean for individuals and their ownership of data about them when the value of data comes from a collective dataset rather than from the data about one person?

The seminar highlighted that the debate and discussion around data ownership, while nascent, is becoming more and more significant given the increasingly data-enabled society in which we all live and work.

3.1 Exploring the concepts

The uncertainties that surround concepts like 'data ownership' and 'data rights' have created significant barriers to debates on data. Which aspects of these concepts are important, and which need to be revised?

3.1.1 Data ownership

The concept of 'data ownership' seems to have quite a lot of intuitive power. 'Your data' seems to be a simple shorthand for data that is about you, and because we feel as though we understand how ownership works, this seems to be a helpful way to get purchase on ideas that are otherwise difficult to talk about. Motivations for talking

[1] , Including the Facebook-Cambridge Analytica scandal in March 2018.

about ownership include privacy protection, the desire to be able to use one's own data (both for individuals and organisations), and the idea of sharing in the benefits that others get from using data that might be about you as a person.

It seems intuitively right that you should have control over 'your data', and that if it were used for financial (or even political) gain that you should be able to benefit. This was seen clearly in many of the reactions to recent high profile incidents in which many people were upset by how 'their data' had been used for political purposes without their knowledge or consent.

However, there are very significant problems with the concept of 'data ownership' that make it unsuitable for use in developing a vision for a system of data management that combats the growing sense of unease.

The idea of owning data is challenging because data is not like other goods that we can own. It is non-rivalrous – I can both give it to you and still have it myself without it costing me any of the original good. Other goods are not like this. If my bag is stolen, I no longer have it. But, generally, if your data is stolen you still have it, but someone else has it too. If I sell my house to you, it is yours, it no longer belongs to me and I cannot sell it to someone else, but this is not always the case with data, be it personal data or data that is not about people at all.

In addition, data can be about multiple people, breaking the link between the idea of data that is 'about me' and data that I therefore 'own'. For example, genetic data about me is also about my family, and data that is produced through a business or other relationship with other people or organisations also inherently involves other people. Data about your purchases and preferences is often also about friends and acquaintances. Conversely, personal data about individuals retains a connection to an individual, it is still 'about' them, even if it has in some way been transferred to someone else. It is also not clear why the subject of the data would be the data 'owner'. The parallels to other forms of property are actually easier to see if the person understood to 'own' the data is someone who holds an aggregated data set about many people.

For these reasons, there is a lack of legal basis, in common or civil law, for the idea of data ownership. Common and civil law lack a definition of 'data' and do not confer a special status on it. Only personal data is defined, non-personal data is not defined, and even with personal data there is no clarity whether it can be held or not, and the definition of 'personal data' is extremely broad. It is also a dynamic concept: what is today not personal data could be considered in the near future to be personal data if changes mean that it can be used to identify an individual. Technology evolves continuously and even machine-generated data could be considered, in some situations, as personal data.

Anthropology considers ownership in relation to the social practice of exchange. It is primarily at the moment of exchange, when one person gives something to another person, that the very question of ownership is made visible. One of the things that is at stake in debates about data ownership might be not only data's (lack of) legal status as property, but also its social status as an artefact of exchange. Could some of the problems about what constitutes appropriate exchange in fact be what is at the heart, in some of the discussions, about data ownership?

3.1.2 Consent and control

A different approach is to place the focus on consent for data use in order to give individuals a sense of control. This shifts debate from ownership to the control that people should be able to have over the data that is about them.

This approach has the advantage of directly addressing a concern that has been at the heart of data management scandals: that data about individuals is used without their informed consent.

It was raised in the discussion at the seminar that most legal protection of individuals can be waived by means of consent, which creates a risk as there are two significant problems with both this approach and with the concept of 'data ownership' discussed above. The first is that both approaches can place an unreasonable burden on individuals. Individuals may not have the ability, knowledge or time to make informed decisions about all the uses of data about them twenty-four hours of the day. It may be that we make imperfect decisions when we give consent, and this is not just about laziness, it may also be about the complexity of understanding what the implications are of giving consent for particular data collection and use.

The second problem is that there is a risk to presuming that individuals are always able to give consent. The reality is that the data processing is integral to the delivery of many services and processes and data can be processed legally under other legal mechanisms such as legitimate rights. If individuals were genuinely expected to give permission for every use of data, they would spend their lives doing nothing else. This can in turn create unrealistic expectations of the amount of control individuals might be able to exercise over data about them, increasing the sense of unease when expectations are unavoidably unmet.

Further, when individuals say that they would like to exercise control over data about them, some might argue that they do not, in practice, take advantage of the control that they already have. For example, individuals may rarely read terms and conditions carefully before clicking 'I consent', because the terms and conditions are lengthy, and seen as hard to understand, yet they are the gateway into accessing a service or system. The fact that people may, for many understandable reasons, fail to make genuine use of opportunities to provide or withhold consent has the effect that we rarely think in detail about all the different ways in which data is being gathered, aggregated and used to make our lives better. Instead, individuals' interest in consent and control is primarily spiked only when something goes wrong.

3.1.3 Data rights

One way to tackle some of the limitations of an approach based on data ownership would be to think instead about a conceptual foundation of a bundle of data rights. This approach holds that individuals have rights over data that is about them. This does not necessarily require that they have control over data in the same ways that individuals have control over property that they own. Rather, individuals might have a right to data about them being used in only fair and reasonable ways, or a right to have personal information anonymised in any aggregation that is publicly available.

This approach may have significant power because it builds on a sophisticated and nuanced body of thought that is well-suited to considering fairness, balance and trade-offs. Furthermore, rights are generally well-understood, which might go some way towards addressing the difficulties in communication on data issues.

Additionally, a rights-based approach helps to ensure at least a minimum degree of equality, which in turn underpins the need to ensure that the system of data management is fair.

A further benefit of taking this approach is that it requires careful thought about corollary duties. This is helpful as a way of linking individuals' rights over data about themselves to the bodies that carry out stewardship functions. These bodies might hold the related duties, or at least carry out the duties on behalf of the state.

However, existing conceptions of human rights are highly individualised in their orientation and were not devised with data in mind. Hence, taking data rights as the underpinning conceptual framework requires further careful thought to answer many outstanding questions. For example:

- What is the relationship between individual data rights and collective data rights?
- Which data rights are important, and which should take priority?
- Which bodies hold the corollary duties?
- How are data rights changed by processes of anonymisation?

3.2 Moving the debate forward

An alternative approach to thinking about ownership, consent, control and rights is to place the focus on trust in the data management system. Instead of placing the burden on individuals to make decisions about data about them, we can consider what sorts of institution, accountability and regulation will give individuals trust in the system? One way to approach this is to consider the role of a fair, trustworthy steward. This approach presumes that individuals would rather feel confident that data about them is being looked after than have to make decisions about data themselves. For this approach to work, individuals need to trust the data management system and its underpinning infrastructure. This requires institutions and regulations that protect individuals, and that are accountable for their work.

This approach leaves space for considerable nuance and an understanding of the trade-offs inherent in the use of data. There are a number of ways we can move forwards in this direction.

3.2.1 Taking a societal view

A focus on data ownership tends to be individualistic. But does society have rights over data that is about us? The census, for example, is an important activity that benefits everybody and necessarily infringes on some of our rights over data about us, as does for example surveillance that enables public safety. These are things that protect us or benefit us that we have to balance against our individual rights.

Taking a societal view also highlights some of the shared risks of our data-enabled society. We are already in a world where there are the data-rich and data-poor. We can all benefit when those who are in power have access to high-quality data, because it can enable them to make better decisions. But we have to equip those who do not currently have the capability, whether that is through resources or training, in order to try to rebalance power so that everyone can use data for their own good.

3.2.2 Understanding the value of data

From an anthropological point of view ownership and exchange are key concepts for understanding societies. We need to better understand what happens when data is exchanged, and the value that is involved in an exchange of data – what we get back and what we give away. This is not only about what you might get in exchange for data immediately or in the next year, but also about how data, once processed, may have an effect on your life in the future.

A challenge is that there is a huge mismatch in size between individuals and the organisations with which they exchange data. How does the individual negotiate in those contexts? Better understanding of data value and how to exchange on fair terms is essential.

3.2.3 Focus on use

Both the value and the harm from data come from purpose for which and the way in which it is used. We have tended to focus on controlling the collection of data, but there should be a shift in focus toward the use of data and the impact of that use on individuals. Focusing on what we want at the point of data use may be more effective in reducing harms than focusing on data collection, although that too remains important. While there is no clear law on owning data, there are laws to stop people doing 'bad things' with data. Focusing how we control what people can *do* with data may be more valuable than trying to establish how we protect data as property. In addition, we need to convey that there are significant social benefits to making use of data and moving away from debates on ownership can help to open up debates on how to create a system that uses data for public good but minimises the negative consequences.

3.3 Getting to where we want to be

3.3.1 Building trust

An approach that focuses on building a trustworthy system and providing stewardship requires institutions and regulations that protect individuals, and that are accountable for their work. A significant benefit of taking this approach is that many of these institutions and regulations already exist, such as the Information Commissioner's Office (ICO) and GDPR, and place an explicit focus on principles such as transparency and fairness, which are likely to inspire trust.

3.3.2 Engaging the public and building a social vision

We do not have a sufficiently deep understanding of public attitudes with regard to their relationship with those who are using their data. But in order to find out, we must first acknowledge the complexity of the concepts in play. We can engage with the public on questions about concepts by focusing on the information that citizens need to make the choices they think they want to make, and by asking about the sorts of institutions, accountability and regulation that will build trust in the system. This goes beyond questions of data into wider questions about the kind of society that we want to create.

This requires that the individuals and institutions that make up the current data management system get out of their 'bubble'. This might include:

- Engaging specific sectors and looking at specific use cases to develop the detail of how this approach might work in practice, identifying challenges and adjusting in response.
- Ensuring a wide debate across society, across all parts of the country.
- Working closely with civil society organisations to understand what is causing the sense of unease and how best to combat it.
- Considering how to engage with groups of a wide range of sizes, from small civil society groups and charities to large private sector companies and government departments.

Taking this approach also requires serious thought about a wide range of questions for further consideration. These include the questions in section 2.1 (3) that aim to further flesh out the concept of 'data rights', which might helpfully underpin a data management system focused on stewardship. Further questions include:

- How can the system be built in a way that is ‘future proof’ and adaptable, given the speed of technological change?
- How can equality and trustworthiness be best ensured in a system that includes extremely large and powerful profit-making organisations?
- How does this framework for thinking about data management relate to broader questions about the kind of society that we wish to create?
- What does ‘stewardship’ look like in different sectors?

3.3.3 Conclusion

With a number of existing institutions exploring ethical issues in relation to data and advanced digital technologies including bodies newly created since the publication of *Data management and use*, we need a common foundation on which to build debate. This debate should move on from ownership to how we understand and manage the balance between collective and individual benefits, risks, rights, and the balance between the interests of individuals, groups and industry. If there is a concern with ownership, it might be that what people would really wish to achieve is that those balances are somehow fair, that there is a balance between who is owning or experiencing the risk, and who is owning the value from data and experiencing the benefit that come from data use. There may be a role for an overarching framework that addresses these issues, but exact resolution is likely to be different in different cases. There is no panacea through a concept of data ownership, but there is a need for a set of specific discussions on how we use data. Current institutional change and new bodies mean that we are now building the capability and capacity needed to enable these discussions to happen.

4.0 Contributor papers

The following set of contributor papers were submitted ahead of the seminar to stimulate discussion and after the seminar to open up and expand the discussion. The ideas and reflections contained within are these papers are not necessarily endorsed by the British Academy, Royal Society or techUK.

4.1 Legal notions of 'property' and 'ownership'
Professor Sarah Worthington FBA

This paper is adapted from Professor Worthington's keynote address at the seminar.

The policy and governance dilemmas associated with data rights and data control are intensely challenging. Ideas of property ownership and control – or 'data ownership' and 'data control' – are increasingly seen as providing possible solutions to these complex issues.

As a non-expert in this field, but as a lawyer whose special subject is property, I want to make some general comments about the legal notion of property. My aim is to provoke and focus discussion amongst the experts who are deeply engaged in the detail of what these concepts mean for data.

I make four main points. First, it is essential to be very clear about the end goal before selecting the appropriate legal means of getting there. Once in place, a legal rule *will* have consequences. It is important to ensure these are the intended consequences, not unintended ones. Secondly, having 'property' is not as protective as one might think. This may have important consequences in thinking about data rights and data control. Thirdly, and on the plus side, even though 'property' may not be quite as protective as often assumed, English property law is nevertheless rather remarkable, and worth a little investigation. Fourthly, the obvious alternatives to 'property' thinking in data governance may be worth deeper investigation.

4.1.1 What is the regulatory end goal?

Defining the end goal is typically the hardest part of any project. Here is no exception. If we want the right answers, then we have to make sure we ask the right questions.

In discussions about data, focus typically centres on the personal information we hand over to third parties and the use third parties make of that information, in particular the use they make of aggregated data sets or the resulting 'data infrastructure'.

Until recently there has been relatively little public concern about individual data collection. The law does not ban the mere observation of individuals going about their daily business. This means that publicly observable shopping habits, movie watching habits, newspaper reading habits, height and weight estimations, etc, might all be gathered without restraint (although use of the gathered data might be more constrained). But now this data collection has encroached on 'private' places: machines – not people – log all the above habits and more besides, whether these activities are carried out in public or at home. Moreover, these diverse data points can all be integrated, again by machines, aggregating face recognition data, credit card usage data, mobile phone location data, and the list goes on. This integration can be done almost instantaneously, when previously the aggregation might have taken months or years to produce by private investigators focusing on one individual at a time. That possibility feels invasive in and of itself.

Modern data usage goes still further, with each of us being increasingly complicit in it. We bargain – or ‘barter’, as Gillian Tett described it in last weekend’s FT (6 October 2018) – with our own data records, giving up the data or consenting to its gathering and its use, often in ways we dimly appreciate, in exchange for free access to products we value.

In any event, our own individual data point is, by itself, of very little value (either to us or to the data collector), so we perhaps quite rightly feel as though we are obtaining the online services for free. Yet the service providers freely acknowledge that it is the data we deliver that is valued, not any funds we pay over (FT, 13 October 2018 with the focus on Monzo).

Only very recently have we begun to understand that this may be a pact with the devil. We are now increasingly likely to be offered on line for sale only what we have already indicated we wanted to buy, or the types of movies we once enjoyed, or indeed the news and political views that we once chased down. And from there it is a very short and slippery slope to the ‘infrastructure of industrialised persuasion’ that Onora O’Neill rightly flags as a public risk to our political and social institutions, not just a private risk to individual data subjects.

And yet, dramatically, on the other side of the leger, the public benefits of data collection and its use are legendary. A great deal of both ancient and modern medical research is based on the analysis of data sets. The same is true of old and new social policy interventions. Think of Charles Booth’s poverty maps of London, or even the Doomsday Book, alongside their modern equivalents.

No doubt we want all the good and none of the bad. It is always so. In the industrial revolution people wanted the jobs and cheap goods, but not the pollution and slave wage push.

Here the hard question is what we will put up with, or even welcome, and what we will bar. This is difficult in the extreme, and yet this question is one that must have a clear answer before we can have any hope of designing a data governance regime or a legal infrastructure that will deliver the desired ends.

4.1.2 ‘Property’ is not as protective as one might think

Because property is viewed as being especially protected by the law, concepts of data ownership seem to be the answer to any difficult question in this area. However, property is not as protected as most people think.

Take a simple illustration. I own my bicycle. Most people would expect that if it were stolen the law would ensure that I could get back. However, English law holds to the line that even if I can find the thief I am only entitled to money, not to the bicycle. Most non-lawyers find that response startling and completely counterintuitive: what is the point of owning things if you cannot even recover them from a thief?

Predictably, a statute creates exceptions (the Torts (Interference with Goods) Act 1977 s. 3). If it is not my bicycle that has been stolen, but a Picasso painting, then I will indeed be able to recover the painting. You might think that personal data is more like the Picasso painting than the bicycle, but there are further difficulties with data that I shall come to shortly.

It is also important to note that rights, entitlements and the power to control need not necessarily be associated with ownership. The right to light is not associated with ownership of light, and control over the export of national art treasures does not indicate that the government owns all these art treasures. If rights and entitlements are identified as valuable, they can be allocated and protected without any intermediating notion of ‘property ownership’. That simple fact ought not to be lost in discussions.

Notwithstanding this simple fact, ‘ownership’ and ‘control’ are increasingly prevalent in discussions about regulating data usage. If this is where the focus is to lie, then it is crucial to understand something about legal notions of ‘property’. The law’s approach to property is not necessarily aligned with the general public’s assumptions.

4.1.3 How do lawyers think about ‘property’?

English law, and all the common law systems derived from English law, have long had an approach to ideas of ‘property’ that is quite different from civilian jurisdictions. Nowadays these latter jurisdictions increasingly seek to mimic by statute what the English courts created independently centuries ago.

The first thing to notice is that English law, unlike its civilian counterparts, no longer holds to a sharp divide between tangible assets (such as land, bicycles and Picasso paintings) and intangible assets (such as shares, bonds and debts). Put in legal terms, English law no longer draws hard lines between ‘property’ and ‘obligation’ or ‘property’ and ‘contract’. All these different types of rights are ‘*assets’*: they are all different forms of wealth. All are valuable, all can be controlled in similar ways, all are protected by the law in similar ways, all can be used in commercial transactions in similar ways, and so on. Recall all the commercial assets in issue in the last financial crisis. They were not tangible ‘things’; they were merely contractual rights, yet they could be traded internationally, secured, made the subject matter of trusts, and so on.

If that is true, then what makes ‘property’ special in English law? The answer, it now seems to me, is simple but rather surprising. It is an idea I only settled on relatively recently, but it seems to have a good deal of explanatory force. English property law is not about classifying assets as ‘property’ or ‘not property’, with some differential protection accorded to each class. Rather, it is all about the *sharing* of assets, whatever their type.

We are very used to this idea in the context of tangible assets. The ability to split legal ownership from possession enables legal owners to lease land, hire out a car, pledge a Picasso painting as security for a loan, and so on. All these options are possible simply because we – along with every other legal regime no matter how primitive – recognise the commercial and social advantages of being able to split ownership from possession and share an asset in various ways, retaining its ownership in one person but permitting someone else to have possession on specified terms.

Notice that such an arrangement is necessarily a transactional arrangement between the two parties concerned in the sharing arrangement. The terms of the sharing have to be agreed, whether the arrangement is a contract of hire, or a gratuitous loan, or a pledge to secure a loan. And if it is possession that is being granted, then the subject matter – the asset – has to be a tangible asset that can indeed be possessed.

English law also recognises other forms of sharing that are less obvious. For example, it recognises various forms of security interests, which can be taken over tangibles, such as houses and machinery, but also over debts, or shared, or indeed entire businesses. And if that were not enough, English law has gone still further and allowed the creation of trusts. These too are possible over any type of asset at all. The trust enables a legal owner to specify that certain economic benefits that might be derived from a designated asset will be held for some nominated third party (the beneficiary of the trust) rather than for the owner. It is possible to have trusts of historic country estates and grand family art collections, or trusts of shares held for nominees, pension trusts, superannuation trusts, or client account monies or investment funds held on trust. The term 'trust' simply signals that although one individual is the legal owner of the asset, the nominated economic benefits that might be derived -from the asset are held for other parties, not for the owner – another form of sharing.

To summarise, 'property' in English law is not some special classificatory system that divides assets into 'property' and 'not-property' classes and protects each class differently from the other. Instead, 'property' in English law is a set of rules that enables legal owners to *share* the benefits of their assets with third parties by way of different types of derivative interests, whether those derivative interests are possession, security interests or trust interests. These legally enabled sharing regimes have proved invaluable, both socially and economically.

4.1.4 Data as property?

But notice one crucial limitation. In all of this it is essential that there is 'an asset' of which one might say 'I own this asset' or 'I am determined to share this asset with X'. And it is here that the crunch point arises in relation to any discussion of 'data ownership' and 'property in data'. In most legal systems, information, or 'data', is not an asset.

Why is that? For a start, with information there is not the same ability to control access, or assignment, or – crucially it seems – sharing. If you steal my bicycle, I may still have legal title but I no longer have possession. It is very clear what I have lost and what you have gained by your criminally enforced sharing with me. By contrast, if you steal information, we both have it. We then need to think very carefully about what I have 'lost' by way of involuntary sharing that the law should remedy.

In some areas the law has taken up with a vengeance this challenge of dealing with information and ideas. This is where intellectual property has a crucial role to play. All intellectual property rights are created by statute, not by the courts. Notably, despite the 'property' terminology, the protection delivered by these statutory means is not dependent on any idea of there being 'property' in the creative idea or endeavour. Instead, the statute itself defines rights, and then defines remedies for their infringement, and it is these statutory rights that are then 'assets' that may be assigned or shared in all the ways that other assets can be dealt with at law.

In creating these statutory rights, the relevant statute defines the scope and extent of the rights in issue (note how carefully that is done) and then provides for time-limited monopolies over those rights to the creator or inventor. The commercial privilege of a limited monopoly, a monopoly that can itself be sold or shared by licence, is given by the law in exchange for full public access to the rights and their inherent creative and knowledge benefits in the longer term. But in the interim, before public release, the statute provides the creator with remedies against those using these statutory protected benefits without consent.

As these rights are currently structured, they do not help very much with data protection. The individual subject's data points of the sort collected so commonly these days do not fall within any existing statutory definition of protected 'intellectual property'. There is some limited protection afforded to data sets by way of database rights, but this does not protect the individual subject of the data from the collection of the data nor from its aggregation, but instead protects the holder of the aggregated database from *its* use by competing third parties who might want access to the compilation either in whole or in part.

We should not really be surprised that general data – the individual subject's data points – is left unprotected by this legislation. Recall the practical and policy considerations raised earlier. If protective legislation *were* to be enacted, then it would likely be very difficult to define what controls and what permissions we would want embedded in the legislation, especially given the recognised public benefits of data gathering and analysis, and notwithstanding that these benefits need to be balanced against the potential for substantial personal and public detriments.

Perhaps more difficult still is that it would be quite hard to define what sort of initial data collection would be constrained. As noted earlier, English law has long held to the idea that there is no property, and indeed no right of any individual, in a 'public spectacle'. So passers-by are free to observe other individuals going about their daily activities, including their shopping and movie watching and internet browsing, all without interfering with the subject's legal rights. All the more so if the observation and recording is done with the data subject's consent, as it now so often is. Then it is irrelevant whether the observation is of public or private 'spectacles' or behaviours and information.

If the law is to have anything to say about this sort of general data collection and use – and there appears to be general agreement that it should – then it would seem to be essential to define some new kind of right for the data subject that is protected in specified ways. In doing that, I suggest it would be quite unnecessary to label any such statutory protection as 'property' or 'data ownership', notwithstanding the attractions of such terminology. The data subject would simply have a legally protected right, or – more likely still – the data user would be subject to specified legal constraints in its activities. The notion that this can be done is not at all difficult. What is difficult is the precise settling of the desired limits and permissions. That fraught debate goes to my first point, which must of necessity be the issue that is addressed first.

4.1.5 Alternative strategies beyond 'data ownership'

There are any number of alternative strategies that might be used in regulating data collection and usage. I cannot range as widely as one might, but I want to make two comments, one about consent and one about privacy.

Most legal protections of individuals can be waived by consent. Increasingly we give consent to the collection and use of our data, often without investigating the terms of that consent or understanding the ends to which the data might be used. Indeed, the data gatherers themselves may not comprehend the potential ends for which the data might be used. The parallels with the last major financial crash and the trading in derivative interests that no one really understood is plain. In line with a great deal of consumer protection legislation, it may be worth considering whether these 'take it or leave it' agreements with consumers in relation to their data collection and usage should be subject to constraints on what terms can be taken to be agreed by a consumer simply clicking 'I have read and accept the terms'.

Secondly, privacy and its legal protection is a much more ephemeral concept than property. Rights to privacy are enshrined in the European Convention on Human Rights (Article 8 ECHR delivering 'the right to respect for your family and private life, your home and your correspondence') and in the UK Human Rights Act. But the meaning of privacy is not defined, and nor is its protection absolute – it is a qualified right requiring the balancing of its protection against the need to protect other similarly important personal rights. The practical analogies with data usage and its governance are clear, and thus the legal means that have been used to address the basic concept of privacy and then undertake the necessary balancing act in engaging with its protection may hold some important lessons in thinking about protection of an individual's data and its use.

4.1.6 Conclusion

In summary, any data governance regime faces the difficult task of dealing with the individual subject of the data and also the holder of the aggregated datasets, recognising the enormous potential for public good, but also for specific and generalised harm. Whether the 'property model' of data ownership and control provides the best legal approach to the necessary governance regime is a question discussed in detail in other sections of this report.

4.2 Data ownership: is it an appropriate concept? Eleonora Harwich, Director of Research and Head of Digital and Tech Innovation, Reform

The digital trails that individuals leave when they go about their daily lives give clues about who they are, what they do, what they want and need.[2] This information is valuable to both the public and private sector.[3] Data on how individuals interact with public services can be used to deliver better services and outcomes for individuals.[4] Nevertheless, data is not always used with the principle of human flourishing in mind and there have been notable instances of misuse both in the public and private sector.[5] These often spark off debates around the notion of ownership as individuals feel that data about them was used for something they object to or have not been informed of. Despite the legitimacy of this reaction, is data ownership a useful and appropriate concept?

People tend to assume an implicit understanding of the concept of data ownership, which means it is often not defined. However, this may be because the notion of data ownership does not exist in legal terms. A scan through the *Data Protection Bill* or the *General Data Protection Regulation* shows no mention of the concept. As highlighted in the House of Lords report *AI in the UK: ready, willing and able?*, the notion of ownership cannot be easily applied to data and is therefore not fit for purpose.[6] Data is legally inert in itself.[7]

Instead, the law speaks about the rights and duties that arise in relation to data.[8] These need to be upheld by data subjects (i.e. individual who is the subject of personal data), data controllers (i.e. a person who, either alone or jointly or in common with other persons, determines the purposes for which and the manner in which any personal data are, or are to be, processed) and data processors (i.e. any person, other than an employee of the data controller, who processes the data on behalf of the data controller).

Data is produced through transactional relationships so it is difficult to ascribe it property rights and treat it like any other asset. Data is created through an interaction - an individual with their doctor or an individual and their Internet provider, for example. In addition, in certain contexts data about an individual can be revealing about that individual's family, which means that data about a person is not always only data about that person.

Despite this, there is still a vocal group, particularly in the distributed ledger technology community, that upholds data ownership as moral imperative. Their definition of data ownership can be boiled down to three core elements: the right to access, to control

[2] Daniel "Dazza" Greenwood et al., 'Reshaphing the Social Contract: The New Deal on Data', in *Trust:: Data, A New Framework for Identity and Data Sharing*, ed. Thomas Hardjono, David Shrier, and Alex Pentland (Visionary Future, 2016),103.

[3] Ibid.

[4] Sarah Timmis, Heselwood Luke and Harwich Eleonora, Sharing the Benefits : How to use data effectively in the public sector, (Reform, 2018).

[5] Health and Social Care Committee, Oral Evidence - Memorandum of Understanding on Data-Sharing between NHS Digital and the Home Office, 2018; Zoe Kleinman, 'Cambridge Analytica: The Story so Far', BBC News, 21 March 2018

[6] Select Committee on Artificial Intelligence, *AI in the UK: Ready, Willing and Able*, Report of Session 2017–19 (House of Lords, 2018), 28.

[7] Richard Kemp, Hinton, Paul and Garland, Paul, 'Legal rights in data', *Computer Law and Security Review*, 27, 2001, 142.

[8] Ibid.

and to distribute one's data.[9] This would be akin to the data subject becoming the data controller. Individuals could access data about them at any time; individuals would have full control over how data about them is used if they do not agree with the terms they could entirely remove their data; finally individuals would be able to decide with whom to share their data with. [10] They are the distributors data about them.

This model has some attractive and desirable features such as giving people more agency and control over what can be done to data about them. These features are crucial for building a trustworthy data infrastructure.[11] Nevertheless, building this type of system does throw up challenges of its own. Firstly, there is a certain amount of data that is necessary to share for services to be able to work. For people not be incessantly pestered with notifications about services wanting to access data about them, rules would have to be put in place in a smart contract (i.e. a computer protocol that allows for the transfer of a digital asset from one party to another automatically under agreed upon stipulations and terms) about what is the minimum amount of data that should be shared with different services. This already means giving up a bit of control. In addition, deciding what is the minimum amount of data a service needs to function might not always be a clear and objective decision. Secondly, who will build this system, approve of its accuracy and fitness for purpose and who will carry out the oversight of it? These are crucial question that do not seem to have a clear answer yet.

Individuals can have greater control over what happens to their data without the need for ownership. The debate about data ownership is one, which reduces data to an asset. This might be the way that many companies currently view data, but it might not be the most practical or desirable definition. Some have argued that data should instead be viewed as a form of "digital labour".[12] This might better reflect the complex nature of data.

4.3 Data: vital asset or toxic liability? Professor Jim Norton FREng

4.3.1 Context

The Royal Society, in its previous two reports and broader programme of work on Data Analytics and AI, has amply documented the vital role of access to data sets of appropriate scope and quality as a lubricant of the 21st century economy. There remain however unanswered questions on many aspects of data collection, ownership and exploitation. This short provocation seeks to highlight areas where good progress is being made and those worthy of much broader public debate.

[9] Greenwood et al., 'Reshaping the Social Contract: The New Deal on Data', in *Trust:: Data, A New Framework for Identity and Data Sharing*, ed. Thomas Hardjono, David Shrier, and Alex Pentland (Visionary Future, 2016),106.

[10] Ibid; Guy Zyskind, Nathan Oz, and Alex Pentland, 'Decentralizing Privacy: Using Blockchain to Protect Personal Data', in *2015 IEEE Security and Privacy Workshops*, 2015, 180–84.

[11] British Academy and Royal Society, *Data Management and Use: Governance in the 21st Century*, 2017; Timmis, Heselwood and Harwich, Sharing the Benefits : How to use data effectively in the public sector, (Reform, 2018).

[12] Eric A. Posner and E. Glen Weyl, *Radical Markets, Uprooting Capitalism and Democracy for a Just Society,* **(**Princeton University Press, 2018**)**; Imanol Arrieta Ibarra et al., 'Should We Treat Data as Labor?', *American Economic Association Papers & Proceedings, Vol. 1, No. 1,*(December, 2017),

4.3.2 Identifying data assets

One benefit of the General Data Protection Regulations (GDPR) has been to give every incentive to organisations to identify and register all their data assets properly. Once located, data assets can be graded in terms of: key material that needs to be carefully managed and fully secured (the corporate crown jewels); information that is useful, but non-personal, requiring basic security; and data that there is no need to keep, and might indeed become a liability, which needs secure disposal.

4.3.3 Personal or non-personal data?

It would be helpful to establish a clear distinction between personal data, and that which has no link to specific individuals. The exceptional value, for example of data related to in-service performance of engineering systems, has been highlighted in numerous case studies.[13] There is a strong case for the additional value that can be extracted by controlled and secure sharing or trading of such data. Other countries may already have a lead on the UK.[14] Such trading must retain secure ownership through watermarking, audit trails and demonstrably enforceable sanctions against unauthorised proliferation.

Personal data is far more challenging. Many would argue that they should own the data related to their own commercial transactions, yet in many cases they have little choice but to relinquish that ownership in order to use commercial platforms. In social media, the controls to limit personal data access and exploitation can be challenging to use or even to find... Similarly, citizens are often asked to release more personal data than is strictly required for example to establish their entitlement to access particular services. A market-led response could be through the broad use of trusted intermediaries[15] to manage and accumulate personal data on customers' behalf and perhaps to return some economic value to them?

Personal health data is perhaps the key test case. Many would be content to release their personal health data for genuine medical research purposes,[16] but would demand that it not find its way into the hands of, say, life insurance companies... Transparency and clear authorisation (e.g. opt in rather than opt out) should establish that essential trust. Innovative solutions, such as that proposed by Sensyne, where NHS Trusts retain the data sets, with strictly controlled access to data for research purposes, represent a possible way forward. A charge for access to this highly valuable asset could be used to contribute revenue to the hard-pressed Trusts and this should be explored further...

4.3.4 Data curation

At this early stage of our data-enabled economy, it is difficult to predict how non-personal (and subject to strict controls, personal) datasets collected today might be used and cross linked in the future to unlock new sources of value. This implies the creation of a new profession in data curation – 21st century librarianship? That profession would ensure that the necessary metadata is stored alongside the data set itself and that key parameters such as provenance, quality (including measurement uncertainty, data calibration and error bars), timeliness, repeatability, and so on, are

[13] See for example Royal Academy of Engineering Report Connecting Data: Driving Productivity & Innovation, Nov 2015.

[14] See the German Industrial Data Space, Fraunhofer, https://www.fraunhofer.de/content/dam/zv/en/fields-of-research/industrial-data-space/whitepaper-industrial-data-space-eng.pdf

[15] See companies such as Mydex.

[16] For example, the One Hundred Thousand Genomes project.

properly recorded. Without such, the stored datasets are essentially valueless. The establishment of standards in all these areas is an essential prerequisite...

4.3.5 Making the intangible tangible?

More than two hundred years of experience from the start of the Industrial Revolution have given us extensive accounting and business tools to identify, manage and control tangible assets. Yet increasingly corporate investment has switched to intangibles, including software and datasets. We do not yet have the equivalent tools to value and manage and fully exploit these intangibles and their absence distorts company reporting and even national assessments of the balance of payments. For example, it has been suggested that the USA balance of payments deficit would be halved if statisticians were better able to capture the value of the software developed in Silicon Valley.[17]

4.3.6 Distorting the development of Deep Learning?

The remarkable advances made in the last ten years in Deep Learning to facilitate highly targeted Artificial Intelligence (AI) applications have been driven by access to increasingly large datasets as well as commoditised computing. It is broadly recognised that such AI applications replicate any in-built biases in the training datasets used. How can more representative datasets be assembled? How can new start-ups gain fair access to such data? What impact is restricted access to such data having on fair competition? What is the role of Government in continuing to give open access to public data sets, and linked, anonymised, non-personal datasets, for example through extending the role of the Office for National Statistics?

Much has been achieved, but so much still to do...

4.4 The HAT data ownership model: first party IPR for individuals Professor Irene C L Ng

Funded through more than £3m RCUK/EPSRC grants, the HAT (Hub-of-All-Things) and its related projects set out to design and engineer a legal, economic and technological artefact (the HAT Microserver) capable of storing, processing, transforming and exchanging personal data and that also assign a set of rights to the data to individuals themselves. Its objective is that the personal data sitting within the HAT Microserver can define, sui generis, a new asset class of PPD i.e. person-controlled personal data, the personal data where intellectual property rights and excludability of the data (control) is with individuals. To create the PPD asset class, and the artefact that contains it, the HAT was designed, engineered and built around 11 design principles derived from the economic properties of data as a digital good.

4.4.1 Principle of Co-production Access Rights without lien

Personal data has the axiomatic property of co-production. It is generated through human activity, but collected through technology owned by a firm. The individual must therefore own a technology/device (the HAT Microserver) that is able to collect data in such a way that both the firm and the individual, as co-producers, would have access rights to it in real time and on demand. The data accessed must be free from lien and encumbrances and, subject to prevailing data protection laws, allow both parties to reuse and re-share.

[17] Hidden value in phones could 'cut US trade deficit in half', Financial Times 17th May 2018.

4.4.2 Principle of alienable rights

Privacy, according to many advocates, should be an inalienable right. Yet the challenge here is not privacy, but that personal data controlled by organisations often cannot even be sufficiently isolated to assign rights. While data may not be assigned rights, *databases* are protected by US copyright law and the EU database directive. Database rights are specifically coded laws on the copying and dissemination of information in computer databases. Individuals must have their own database and database rights within the HAT Microserver thereby granting alienable rights for the personal data within and individuals can grant these rights to others for a period according to their own wishes and for every data point in the database. This must be achieved by the individual executing a set of software code within the HAT Microserver to grant time and context dependent exchange of data with low effort. For rights to be assigned without ambiguity, there must also be suitable isolation of each HAT Microserver database from one another. A system of HAT Microservers must therefore be a distributed system of individual HAT Microservers owned by individuals themselves and yet fully interoperable with one another and able to be aggregated for firms to render services in a scalable manner.

4.4.3 Principle of Non-rivalrous Consumption

Personal data has an economic property of non-rivalry i.e. consumption of data by an entity does not prevent another entity from consuming it (Shapiro & Varian 1998). This implies that each co-producer may consume the data in a way that benefits itself as well as contract with other parties, without denying the other of consuming and contracting the same. That means API access from data sources into the HAT Microserver database on demand through HAT “data plugs” must create a copy of the data generated but changing the data rights once the data enters the HAT database, so as to ensure each co-producer have a set of independent rights for the data that sits within their domain.

4.4.4 Principle of Expansibility

Personal data has the economic property of infinitely expansibility (Rayna, 2008) . That means a firm’s data of a person can be copied to another space with very low marginal cost of re-production. The co-producers could hold the same copy of that data in the same instant that it is generated in their respective technological domains/devices and have the ability to contract with third parties to continue expanding its use. The HAT schema (data structure) allow infinite combinations of data values across datasets to be exchanged as a data product e.g. Tweets only in Boston, locations between 7-9am. Each of these data values and bundles can be named and then exchanged/contracted through standard APIs using standard Internet protocols and encryption in real time. In a similar way, the firm can do the same with their data (subject to prevailing laws on personal data sharing)

4.4.5 Principle of Excludability

Personal data have an economic non-excludability property, implying that it is near impossible to exclude others from consuming the data unless there is a legal (e.g. contract) or technological (e.g. encryption) framework. Excludability of personal data controlled by individuals must be based on a data contract and/or technological instrument whereby individuals are in a position to grant and/or deny rights over personal data usage. HAT Microservers create data debit contracts when granting rights of HAT data to others and data in transit is SSL encrypted from end to end, in a similar manner to emails.

4.4.6 Principle of Data Derivatives

Personal data have an economic property of recombinant and divisibility (Quah, 2003). Personal data e.g. location, combined with time e.g. 7-9am, can create a secondary, derived data product e.g. commuting journey. A new economic good can be construed as being created when different types of data are combined in such a way that can be exchanged, which means that combining personal data for new exchanges increases the underlying asset value of the database. An individual must control the permission and process of data being combined and transformed (even if it takes seconds) so that the database value increases. The individual must also control the usage of private AI tools on the HAT Microserver that creates new data.

4.4.7 Principle of Data as Store of value

Personal data use contracts cannot specify all states of nature nor all future actions and use of the data, in advance. When there are states or actions that cannot be verified ex post by third parties, they are therefore not possible to be contractible ex ante. The literature on incomplete contracts (see Grossman and Hart 1986; Hart and Moore 1990; Aghion and Bolton 1992; Dewatripont and Tirole 1994) have shown that the allocation of power matters when it is not possible to specify in advance precisely how that power should be exercised. Since the value, worth and use of the data is not known, the power to decide on future uncertain contracts must be in the hands of the individual. Therefore, the HAT Microserver has to be the store of value for the individual before a context emerges for an exchange to occur for personalisation or recommendation of products and a data contract emerges. If personal data is available in real time and on demand, every data contract will then be complete for a specific use with no ambiguity and firms have less need to hoard data.

4.4.8 Principle of Data as Medium of Exchange

The value of some personal data can expire (perish) if not used e.g. the need for Hotel recommendations. It is therefore context and time dependent. Personal data must therefore be available on demand and in real time to be a superior asset class and to be an effective medium of exchange for data contracts for personalisation and recommendation. By way of the HAT Microserver being both store of value and HAT APIs being the the vehicle for exchange, HAT data, in its standardised form, should be treated as currency (like GBP, USD). The only missing element is its ability to be unitised but that can be derived empirically through increase usage, and scale.

4.4.9 Principle of Transparency

The way personal data is stored, exchanged and processed and the way it stays at rest, in transit and used must be clear and transparently available for scrutiny. The HAT Microserver must be an open sourced technology, even if services built on it can be commercial. The processing of data within the HAT must be based on code that is open sourced and/or standard Internet technologies. The granting of data rights (usage, exclusion and alienability) must be transparent.

4.4.10 Principle of Trust Anchoring

Trust is a prerequisite of contracts (Göran and Hägg, 1994). While the HAT Microserver technology has been legally, economically and technically engineered to endow IPR of personal data to individuals, it still needs to be issued like a private data account, much like banks issuing savings or current accounts. For the market to form, HAT Microservers must still be provisioned on license by a trust anchor, which could be the data brokers or data trusts, as long as there are guarantees either by the state or through market incentives, to stay trustworthy. The difference is that, with IPR resting on individuals, the transferability of data rights can be achieved through a direct and complete contract, much like currency payments, even if it is enabled by data brokers as trust anchors. This would therefore ensure the market viability of data brokers as a service for individuals. Trust anchors could also create additional middleware services or governance mechanisms e.g. hierarchical or nested relationships between HAT owners e.g. parent and child; a power of attorney situation; or create better heuristics of data sharing practices across apps within the trust anchor's ecosystem.

4.4.11 Principle of Market Design

With HAT data having a set of transferable rights, it is now a formal economic good that is possible to create a thin crossing point (Baldwin, 2007) i.e. a transaction boundary for the transfer of rights. Matching of HAT data to apps should be dictated by market design rules of thickness, reduced congestion and safety (Niederle et al., 2008). Best practices of data exchange should be made transparent and allow different types of apps (and different levels of exposures) to play out that will optimise choice and privacy/security concerns.

4.4.12 Implementation

The HAT proof of concept was implemented in November 2016 on AWS (Amazon cloud service) as the first installation and the ability to generate a HAT Microserver (complete with a database) within 3 seconds of signing up was achieved in July 2017. The implementation of the HAT Microserver was optimised to test its cost structure and a cost of £2 to £4 per month was achieved in January 2018. The HAT is now in live use both in the innovation environment and in live commercial environment . HATs are open sourced under AGPL, portable and can be issued from most devices e.g. HATs in the cloud by different cloud operators; HATs on a Raspberry Pi or even on a PC at home, or in other devices. However, the security architecture and threat models would differ for each installation, as would be the business models. While one person per HAT would dis-incentivise hacking (a hack of one yields one HAT's data), more work could be done from the security perspective for different type of HAT installations.

4.4.13 Conclusion

The HAT Project's ultimate objective is that an explicit, primary market for personal data, similar to the emergence of a primary market for digital music in the early 2000s, would reduce illegal and inefficient personal data markets as well as reduce externalities relating to privacy, as future applications switch to using HATs as user accounts. The HAT model sets up a parallel asset class to challenge the OPD asset class through easier access, higher quality and lower friction, much like the way music licensees challenged music piracy. The HAT full technical system architecture can be seen at https://developers.hubofallthings.com and the ecosystem at https://www.hubofallthings.com/the-hat-ecosystem/. A simplified explanation of the HAT is available at https://www.hubofallthings.com/main/what-is-the-hat/. To date, there are 1500 HAT owners and the platform on AWS is live. Individuals can obtain a HAT at the applications live on HATStore https://HATDeX.org/hatstore. To date, there are 1500 HAT owners and the platform on Amazon Web Service is live, with 12 pilots that are in various stages of integration with HATs.

This paper is an excerpt submitted to the seminar on the 3 October on Data Governance with the British Academy, Royal Society and TechUK.

The full paper can be viewed at:

Ng, Irene C.L. (2018), "Can you own your personal data? The HAT Data Ownership Model", University of Warwick Service Systems Research Group Working Paper series, ISSN 2049-4297 no. 03/18, at [*http://wrap.warwick.ac.uk/108357/*](http://wrap.warwick.ac.uk/108357/)

4.5 Data ownership and data rights Roger Taylor, Chair of the UK Centre for Data Ethics and Innovation

Artificial intelligence and the algorithms that determine our experiences both offline and online are having a profound, and sometimes unexpected, impact on our lives and society. Questions about what rights and controls individuals, communities and organisations should have over data sit at the heart of how to unlock the benefits of these data-driven technologies.

How we answer these questions matters for what kind of society we want to be: who has the right to what data and for what purpose can support or undermine power dynamics. It matters for our economy: how we assign rights and values to data can help stimulate data trade and transfer and encourage innovation. It can also give rise to new business models and innovation that offer the individuals more control over data about them.

How we talk about data rights and control also matters. Relying on overly simple concepts of ownership risks shaping the narrative in unhelpful ways. Perhaps the only point of agreement at the recent *Seminar on Data Ownership, rights and controls: reaching a common understanding* was that traditional notions of ownership are going to be inadequate. That in itself shows the scale of the challenge we face.

There are two major drivers highlighting why this is a challenge we need urgently to address.

Firstly, there is a sense that there is a growing public concern about how data about individuals is used. This is fuelled by scandals such as the much reported-on Cambridge Analytica, by an increasing awareness of unfair practices such as price discrimination, and by a worry that a new digital society might lead to the exclusion of some groups. At the same time, there is no clear alternative emerging from calls for greater individual control nor any agreement on what this might actually mean. Key elements have been identified – greater public involvement in determining what is acceptable within society, greater individual control over how data is used, better auditing and monitoring of how systems are behaving. But there is much work to do to define scalable and enforceable solutions.

Secondly, there is a looming threat of missed opportunities unless we have the right mechanisms to enable the trade and transfer of data. For the UK to continue to be a world-leading digital economy, we need to have the right legal and governance frameworks to enable ethical innovation. This means questioning whether some of the frameworks we have relied on in the past are the right ones to create the best possible future.

The Centre for Data Ethics and Innovation will play an important part in how we go about tackling some of the challenges that lie ahead. Our high calibre Board was appointed in November and we have been tasked by the UK Government to maximise the benefits of data and AI for our society and the economy.

The Centre will work collaboratively to strengthen the data ecosystem by addressing public trust and ensuring governance is effective, enabling ethical innovation to flourish. It will do this by working across sectors providing evidence based advice and recommendations to policy makers. Our work will include analysing the landscape to map emerging harms and opportunities. We will also undertake specific projects - our first, algorithmic bias and online targeting, were announced in the recent Budget. As an independent body, our advice to Government will be robust, evidence based and reflect engagement with a wide-range of stakeholders including academics, business and the public.

4.6 Reflections on ‘data ownership’ Guy Cohen, Strategy and Policy Lead, Privitar

The arguments against data ownership are numerous and convincing. For one, data often doesn’t just relate to one person, so to whom does it belong in the first place? If person a buys something made by person b from person c in a shop owned by person d, who owns the data of that sale? Data is often about our interactions, and as such can’t easily be attributed to just one person. Even something as intensely personal as a date of birth is also information about someone else; the date their mother gave birth.

Second, it is important that individuals have rights over data about them, and that these rights cannot be sold. So, assuming that our rights would persist, what would it mean to own someone else’s data? If it is simply giving access to someone who would otherwise not have access, then what does ownership add over contract law?

So if ownership is not a helpful concept, why is it getting so much attention and support? I believe data ownership is responding to two related but distinct concerns. The first is about control. People feel that they are not in control of how their data is being used, and are concerned that they may be at risk or disadvantaged in some way, they experience some new uses as ‘creepy’. Second, people feel that companies are deriving enormous economic benefit from using their data, but they don’t receive any share of this benefit, and they think this is unfair. Seeing it as two separate issues; one on data protection rights, and one on economics, allows us to look more specifically at what we hope to achieve, and thereby identify more appropriate responses.

The first issue, around control, already has a response from data protection law. When outlining the purpose of the GDPR, Recitals 6 and 7 state that:

“Rapid technological developments and globalisation have brought new challenges for the protection of personal data. The scale of the collection and sharing of personal data has increased significantly... Those developments require a strong and more coherent data protection framework in the Union...Natural persons should have control of their own personal data. Legal and practical certainty for natural persons, economic operators and public authorities should be enhanced.”

The question of rebalancing control is one the GDPR aims to answer directly by strengthening individual rights, enforcement powers, and controller obligations. Whether the new law will be successful in this goal remains to be seen. But, given it is still so young, we must wait and see if the GDPR will sufficiently rebalance control in favour of the individual. The second question, around sharing in the economic benefit of the big data age, is not really addressed by the GDPR, and I think does require greater attention.

That the big data age is increasing inequality by reducing the role of labour and increasing the share of wealth going to capital is a well discussed topic. For data ownership to lead to the redistribution of wealth a mechanism would be needed for individuals to be paid for use of their data. This system was advanced by Jaron Lanier in his 2013 book ‘Who owns the future’, where he suggested a system of micro payments for data use. Whilst Lanier provides an insightful critique of how the concentration of wealth occurs, solutions of this kind face many challenges. For instance, the relationship between the value created by a digital product and the data contributed by a given individual is not clear. Similarly, a payments system for the amounts likely to be generated could be extremely expensive, possibly costing more to be managed than the amounts being distributed. But, more fundamentally, these approaches may have the undesired effect of inhibiting data use and reducing innovation.

The popularity of welfare capitalism in the twentieth century over other models, such as socialism or libertarianism, is arguably because it has led to better standards of living for more people. History's example is that layering protections on top of productive systems (taxing income to provide welfare), as opposed to taking steps which may inhibit their productivity (such as socialism), arguably leads to better overall results. For the data age, valuable productivity comes from the wide use of data, and so steps which make it harder to use data should be challenged, their impact appraised, and their necessity demonstrated.

Aside from being problematic as a solution, focusing just on the issue of redistribution of wealth runs the risk of being blinkered and not looking at how one response might affect the wider system. Data is often described as the fuel for the current industrial revolution. Policies which change the way in which this fuel can be used and consumed risk massively affecting other elements of that system, perhaps positively, but perhaps not. As such they shouldn't be considered in isolation, but as part of a systems thinking approach. Inequality is certainly not the only concern exacerbated by the big data age; cascading risks and antitrust issues are two other examples that share a common cause.

Data ownership may be an appealing response to pressing concerns, but it is the wrong answer to the underlying issues. To better understand what the right answer is, we need to look more closely at what these concerns are, and look at them in the context of the data driven industrial revolution and the wider set of associated issues. Policy makers may need to act, but in doing so they should think in terms of how their policies will affect the whole system, rather than looking for point solutions for individual issues, and, crucially, seek solutions which do not put the huge promise of the data age at risk. Current proposals for data ownership fall short of these objectives.

4.7 Reflections on the data ownership, rights and controls seminar from the ICO Romin Partovia, Senior Technology Officer, Information Commissioner's Office

On 3 October 2018, the Information Commissioner's Office (ICO) was invited to the British Academy, Royal Society and techUK seminar to discuss data ownership, rights and controls. It was a great opportunity for the ICO to get some feedback from businesses within the technology sector on their thoughts around data ownership, building relationships with customers and how the General Data Protection Regulation (GDPR) fits into this.

Within the GDPR data ownership isn't explicitly defined as its own concept. We have data controllers, the organisations that determine the purpose and means of processing personal data, and data processors, the organisations that can be used to process personal data on behalf of the controller. The question discussed at the seminar was who owns this data.

This isn't necessarily the first question we ask. Primarily we want to know who is in *control* of the data – that is who has overall responsibility for managing the data. Ultimately, this is the data controller, and the GDPR sets out certain responsibilities that the controller must uphold. This all sounds relatively simple, but often there can be complex layers of controllers and processors involved and identifying who has what responsibility can be no mean feat. Our Data Protection Act 2018 guidance on processors and controllers is a good starting point in establishing if you are a controller or processor. Moving on, our GDPR guidance on contract and liabilities between controllers and processors will help you manage the relationship and identify your responsibilities. These can be found on our website:

- Data controllers and data processors: what the difference is and what the governance implications are
- ICO GDPR guidance: Contracts and liabilities between controllers and processors

I haven't mentioned the most important person yet, the person who the data identifies. We refer to this person as the data subject. Does the data subject own the data that they give to controllers? They certainly have rights over the data, and these are set out in the GDPR. Right of access, rectification, erasure, portability and the right to objection are all articles within the GDPR that empower the data subject. These rights aren't always absolute, so controllers should be aware of what rights the data subjects have and when they can be exercised.

Rather than trying to define data ownership as a legal concept, perhaps it should be viewed more as a philosophy for processing personal data. If an organisation instils a culture within the organisation where data belongs to a person and that person owns the data, it's a good starting point for building better relationships with customers. This won't fit all organisations, so you should be careful in giving data subject's false expectations.

The GDPR can be used as a great tool for building trust, retaining and attracting customers and gaining competitive advantage. In the digital age that we live in, organisations should be striving for privacy and innovation; the GDPR allows us to achieve this.

Data protection by design and default provide the opportunity to include data protection practices into your processing activities and business practices. Data Protection Impact Assessments (DPIAs) allow you to identify and minimise data protection risks of a project, and Article 32 of the GDPR sets out the security requirements of controllers, so customers trust their data will be safe in your hands. These are just a few examples of how the GDPR can help build better relationships with customers. The organisations that embrace these will be the organisations that have better relationships with customers.

We had great questions from the audience about anonymisation and the increase in data protection breaches.

There was concern that organisations could sell or buy personal data that was anonymised first, especially with the increase of data-hungry Artificial Intelligence systems. Once personal data is anonymised it no longer becomes personal data and GDPR requirements no longer exist. However, with that said, truly anonymised data is difficult to achieve, and it is not to be confused with pseudonymised data. Pseudonymised data is still personal data and will require compliance with the GDPR. Where ever that personal data goes, the rights of the data subject go with it. Further guidance can be found on the ICO website.

A great point was made about the increase in personal data breaches and if this will become an upwards trajectory of ever-increasing breaches. The ICO has certainly seen an increase in calls to our helpline. We see this as a sign that people are becoming more aware of their data protection rights and organisations are coming to us for advice. The ICO has also been strengthened in number and expertise, allowing us to carry out our responsibilities and obligations to organisations and to the public.

Whereas data protection by design was once a best practice, the GDPR now requires it as a requirement of processing personal data. This will ensure the ICO can take proactive steps in ensuring data minimisation and other good practices are adopted.

We are in the process of setting up a regulatory sandbox that will support organisations to develop innovative products and services using personal data in different ways. We also have a grants programme to support innovative research and solutions focused on privacy and data protection issues.

The GDPR now requires organisations to carry out DPIAs before processing data that are likely to result in high risk to an individual's interest. We have a DPIA team that can support you and provide feedback on your DPIAs to ensure good data protection practices.

These are all tools that the ICO is championing to ensure that organisations can follow good data protection practices, whilst also improving an organisation's ability to retain and attract customers, innovate new products and services and develop new ways of thinking.

4.8 Data ownership, rights and control – an anthropological perspective Dr Hannah Knox, Associate Professor of Anthropology, University College London

This paper is the script of the presentation that Dr Knox gave at the seminar.

I have been asked to comment on some alternative ways in which we might think about data ownership. As an anthropologist I'm lucky to be part of a discipline for whom the question of what it means to own something has been very central. Given this I want to try and convey some of the ways in which anthropologists have thought about ownership and to consider what the relevance of these approaches might be for the questions that we are trying to tackle here today about the ownership and use of data.

No discussion about ownership in anthropology would be complete without locating it in relation to the social practice of exchange. For it is primarily at the moment of exchange – when one person gives something to another person, when the question of ownership is made most visible. To understand what it means to own something, means understanding the conditions under which that thing can be transferred to somebody else.

One of the things that anthropologists have repeatedly observed, is that to exchange something is an act that requires that thing become detached from the person who previously owned it – socially, materially, legally - and to pass it on to another person or body who becomes newly attached in some way to that thing (Weiner 1992). But not all things do this act of attachment and detachment in the same way. Recognising this, a key distinction that has emerged in anthropological studies of exchange, has been the difference between gifts and commodities (Gregory 1982).

Perhaps then, what is at stake in debates about data ownership and use, is not only data's legal status as property, but also its social status as an artefact of exchange. Could problems about what constitutes appropriate exchange in fact be what is at the heart of discussions about data ownership?

To consider this I want to use the example, discussed by anthropologist Marilyn Strathern, of a prior debate about public policy and ownership that has many parallels with the data debates we are having here today – the practice of organ donation (Strathern 2012). Like data ownership this has been a fraught area with many of the same ethical arguments about appropriateness of exchange being rehearsed.

Strathern points out that debates about the donation of organs have frequently hinged on whether the act of giving an organ should ideally be an altruistic act – that is, a gift - or whether people should be monetarily compensated for their body parts. For those who advocate a non-monetised form of organ donation, the act of exchange is one of giving a gift. The language of altruism implies that the gift is entirely disinterested. It also implies that after having given the gift there need be no reciprocal return. However, that is not the end of the story. Those who have received an organ often want to repay the donor, meanwhile the donor themselves often articulate their act not as simply altruistic but rather as a public act for the greater good.

At the same time the gift of giving introduces some other tensions. For example, the idea of giving away a body part as an act of altruism - and the language that is used to describe this act as one of gifting – also hides another feature of organ donation which is often commented on – that 'human donations enter the organ procurement and distribution system altruistically, and exit commercially' (Strathern 2012, 405). We might argue then, that language of the gift here does important work of concealing a commoditisation process that is at play in the world of organ donation. A recognition of this tension, has opened up discussions about whether in fact, the commodity relation – i.e. the monetisation of organs within the healthcare system – should be brought back into the moment of organ of donation. If someone is creating monetary value out of an organ, then should the act of giving not also be adequately compensated? Otherwise, the worry is that there is an unevenness in the exchange – a one-sided gift which now appears to deserve a reciprocity, but which does not receive it.

A similar dynamic can be seen at play in discussions about data. The way in which we think about data from an individual point of view often uses the language of gift exchange as well. People are asked to 'give' their consent for data to be used, or we talk about sharing our data with others. When individuals articulate that current regimes of data exchange are unproblematic for them, they frequently invoke a language of reciprocity to resolve the movement that makes their data someone else's (for example people say that they recognise that if they don't give their data they won't have access to services like Google Maps, so there is a reasonable exchange at play here). Conversely for those who see data exchange as problematic, it is precisely the mismatch between the free gift of data on the one hand and the way in which that data is monetised on the other, that is at stake. The surprise that is often expressed when data is revealed as a source of revenue generation derives from the same disjuncture between the language of gifting data at the moment of its generation and the commercialisation of data as it is monetised and exchanged that we saw in organ donation. Following from this come ideas such as giving individuals micropayments for their data to balance the seemingly uneven nature of exchange.

There is much more we could say on this. For example, when data is co-produced by corporations and so called 'prosumers' what effects does this have on the way in which people negotiate data exchange? Another question concerns personal data and the relationship that data has to the body or self. How do people let go of things that are conceived as still being a part of themselves? And what are the social implications when personal data returns as a 'digital double' (Knox et al. 2010)? Here the question become less one of who owns data, and more one of how to successfully achieve exchange, when gifts turn out to be commodities and data-objects that were given away can never be fully separated from their previous owners.

References

Gregory, C. A. 1982. *Gifts and commodities*, *Studies in political economy*. London ; New York: Academic Press.

Knox, H., D. O'Doherty, T. Vurdubakis, and C. Westrup. 2010. The Devil and Customer Relationship Management. *Journal of Cultural Economy* 3 (3):339-359.

Strathern, Marilyn. 2012. Gifts money cannot buy. *Social Anthropology* 20 (4):397-410.

Weiner, Annette B. 1992. *Inalienable possessions : the paradox of keeping-while-giving*. Berkeley ; Oxford: University of California Press.

5.0 List of speakers

Co-chairs	
Prof Genevra Richardson CBE FBA	Professor of Law, King's College London; Chair of the British Academy Public Policy Committee
Prof Dame Otteline Leyser DBE FRS	Professor of Plant Development, University of Cambridge; Chair of the Royal Society Science Policy Advisory Group
Sue Daley	Head of Programme Cloud, Data Analytics & AI, techUK

Keynote speakers	
Prof Sarah Worthington FBA	Downing Professor of Laws of England and Director of the Cambridge Private Law Centre, University of Cambridge
Roger Taylor	Chair, Centre for Data Ethics and Innovation

Panel 1: Exploring the concept of ownership, its value and limitations	
Benoit Van Asbroeck	Partner, Bird & Bird
Rachel Coldicutt	CEO, doteveryone
Dr Hannah Knox	Associate Professor of Anthropology, University College London
Jeni Tennison	CEO, Open Data Institute (ODI)

Panel 2: Building data rights and relationships between consumers, businesses and the public sector	
Simon Burall	Senior Associate, Involve
Emma Butler	Data Protection Officer, Yoti
Jacqueline Davey	Vice-President Sales in the UK and Ireland, IBM
Romin Partovnia	Senior Technology Officer, ICO

Panel 3: Reflections	
Sophia Adams-Bhatti	Director of Legal and Regulatory Policy, Law Society
Guy Cohen	Strategy and Policy Lead, Privitar
Hetan Shah	Executive Director, Royal Statistical Society
Prof Karen Yeung	Interdisciplinary Professorial Fellow in Law, Ethics and Informatics, Birmingham Law School, University of Birmingham

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.
www.thebritishacademy.ac.uk/ | @ @BritishAcademy_

The Royal Society is a Fellowship of many of the world’s most eminent scientists and is the oldest scientific academy in continuous existence.
royalsociety.org/ | @royalsociety

techUK represents the companies and technologies that are defining today the world that we will live in tomorrow. Over 900 companies are members of techUK.
techUK.org | @techUK

The British Academy

THE ROYAL SOCIETY

techUK