

White paper

Personal Big Data: Is there a missing third party in our emerging Big Data society?

Dr Duncan R Shaw

March 2014

Executive summary

New Big Data technologies are rapidly changing marketing, healthcare, government, financial services, retailers and whole supply chains.

We are rushing towards a 'Big Data society' that is using data analytics to more efficiently target resources and to deliver incredibly personalised user experiences. But the precise use of resources and the personalised delivery of services require access to deeply personal consumer data.

Personal data can be used for great harm as well as for great good. The more that data is shared between organisations then the more value it can create – and the more difficult it is to control who uses it and what they use it for.

The change in how organisations use our personal data is happening whether we like it or not and we risk destroying trust if consumers are harmed or even surprised, by how their personal data is used. We need consumers to trust how their data is used or they will be slower to engage by sharing their data. This will delay the benefits of a Big Data society and leave the UK to be potentially overtaken by other countries with a different view of the importance of consumer trust.

But current systems of legislation and regulation are based on older technologies and ways of working that did not include cheap access to mass data sharing capabilities and personalised data analysis in real-time.

Our investigation incorporates the views of experts from regulators, government, commercial data firms and consumer privacy organisations. It concludes that there are several missing roles in our emerging Big Data society – a missing 'Third Party'.

This 'Third Party' would support individual consumers to deal with networks of large and small firms; help firms to share and use data in new ways *in return for doing so appropriately*; aid regulators to bridge the gap between the market and individual consumers, staff and firms; and give privacy and consumer organisations a platform to help more consumers and to engage with more firms.

We propose a solution, a design for a 'Third Party' that engages the attention and resources of the different stakeholders to watch and help each other. Firms would have a strong interest in behaving appropriately; and in turn they would encourage their staff to behave appropriately and become more successful in the process.

Contents

Acknowledgements	3
1 Introduction.....	4
1.1 The opportunities and risks in our emerging Big Data Society.....	4
2 Problem: How to share data and preserve consumer trust?	6
2.1 Preserving consumer trust.....	6
2.1.1 The difficulties with preserving consumer trust	6
2.1.2 Preserving consumer trust using legislation and regulation.....	7
2.2 The problems with sharing data.....	8
3 A Third Party to promote preserving trust through appropriate data sharing	9
4 Conclusion.....	11
5 Next steps	13
6 About the author.....	14
Appendix 1: Themes from Roundtable Two on 5 September 2013.	15
Appendix 2: Themes from Roundtable One on 12 March 2013.....	17

Acknowledgements

The author gratefully acknowledges the time, help and advice from the people who have supported this study and who work for the regulators, government, commercial data firms and in the area of consumer interest and privacy.

This white paper makes use of the expert perspectives and experiences of the different people who attended the Personal Big Data roundtables and who were interviewed. Each of the attendees and interviewees gave their personal perspective and this is not necessarily shared by their employer.

1 Introduction

This white paper describes the outcomes of two roundtable discussions about the opportunities and dangers that are arising from Big Data technologies in our society. The roundtables were made up of experts from regulators, government and corporate data users. The roundtables were followed by one-to-one interviews with the roundtable experts and others to more deeply investigate the main themes that came up.

The findings of the roundtables and the one-to-one interviews were then written up into this white paper which was also validated by the interviewees at the draft stage.

1.1 The opportunities and risks in our emerging Big Data Society

We have all heard examples of the increasing scale and sophistication of how Big Data technologies are being used in many aspects of our society to develop completely new products and services. Also, much has been written about how Big Data is changing marketing, healthcare, government, financial services, retailers and supply chains¹.

In our roundtables we discussed how new analytical techniques and technologies are leading to a much more precise use of resources in firms and in government. This greater precision in how we do things, and who we do it for, will revolutionise the different sectors of our society by enabling much greater customisation and personalisation of services. And by generating completely new services that have never existed before.

For consumers and citizens this means that more people will get the services that they need, when and how they need them and in ways that are most convenient. For firms and government this means vast increases in efficiency from only targeting people who need a service and by only doing what they need to do and no more. It also means many new services will be invented to more fully satisfy all the different needs that people generate as they go about their lives – using the efficiency savings to fuel this. And everything that can be done to help a business-to-consumer customer can also be done for a staff member in each of the business-to-business relationships in a supply chain.

But these emerging benefits are balanced with risks. There are many unresolved issues to be dealt with that may reduce the benefits or slow down the implementation of the set of new technologies and services that we group under the heading of 'Big Data'.

Quite often there is a lack of precision and a lack of agreement in what is meant by 'data' – what data? Which data? Data means different things to different people and there are a lot of metaphors being used right now to help us start to understand what Big Data means to us. For example, a common metaphor is 'Data is the new oil'. Another one is 'Big Data is the minute-to-minute personal diary of everyone and everything' – because it is potentially a record of everything that you or your machines do. But metaphors only take us so far – eventually the real meaning of 'Big Data' is about uncovering the unmet needs of customers and citizens.

The most powerful and shocking part of these Big Data services is not the service – it is how using the service has to tell the service provider details about the customer's life². Where they are; when they

¹ For example: www.mckinsey.com/insights/marketing_sales/putting_big_data_and_advanced_analytics_to_work; www.strategy-business.com/blog/Big-Data-Lessons-from-Earlier-Revolutions or the references in http://www.mckinsey.com/insights/strategy/the_benefits_and_limits_of_decision_models.

² For example: <http://econsultancy.com/uk/blog/63046-google-now-it-s-about-the-data-not-the-service>.

Personal Big Data - Is there a missing third party in our emerging Big Data society?

want something; what they want; why they want it; what version do they want; and the detailed historical record of all this.

It is the detailed comparison of all the historical records of all consumers' relationships with firms government and each other that provides the ability to produce such valuable services – comparing everybody's minute-to-minute Big Data personal diary – because value comes from precision and personalisation. But there is a danger that the personal knowledge that is the prerequisite for personalised services can also be used to harm consumers.

2 Problem: How to share data and preserve consumer trust?

Our roundtable discussions generated many interesting themes but the central problem that we came up with was ‘How to share data and preserve consumer trust?’ The logic went something like this.

Our Big Data society is fuelled by quickly emerging new technologies that will develop it and change it whether we like it or not. There are vast societal benefits to be gained in terms of new services, economic development and international competitiveness. These benefits depend on sharing data between consumers and firms, and firms and other firms, at a much greater scale and with much greater complexity – the more data that is used then the more value it can then generate.

The hope is that the right data can be used *for* the right consumers, by the right organisations and in the right way. The dangers are that the wrong data could be used *against* any consumers, by the wrong organisations and in inappropriate ways.

The harm that data can cause is in *how* it is used, just as the value that it creates comes from *how* it is used. The benefits and risks are in the specifics of which particular data is used and how. In this sense data-use is the opposite of a commodity. The many different data and their uses make it best to talk in terms of a specific *outcome* rather than some general ‘data’.

The most dangerous potential outcome is that *consumer trust* would be eroded and consumers would be slower to engage by sharing their data. Talking about trust is useful because it covers all the possible positive or negative uses of consumer data that could support trust building or detract from it. Preserving consumer trust is crucial because to guarantee it would mean that government would not feel the need to constrain data sharing activities whilst organisations learn how to do so in a trustworthy manner³. Limiting data sharing would delay the benefits of a Big Data society and leave the UK to be overtaken by other countries with a different view of the importance of consumer trust.

2.1 Preserving consumer trust

2.1.1 The difficulties with preserving consumer trust

Data can be used and reused in many different ways by many different organisations and in the roundtables we talked about how consumers should be well aware of how their data is used because surprise erodes trust – especially unwelcome surprises.

But most people have very little awareness of how their data is used or of the implications of this use. Most people do not read the fine print when they download apps or when they give permission to record their data in other ways. To be fair, most people do not have the time and few people have the expertise – and this includes the experts in our roundtable. But it seems that there is an urgent need to educate consumers to help them understand the issues and implications of using these technologies. There are more and more examples of data misuse in the media, such as stories about the US National Security Agency and its huge database of hundreds of millions of mobile phone users’ data⁴. It seems that the personal data equivalent of a ‘horsemeat’ scandal is just around the corner and increasingly people are starting protest about this sort use of data by governments⁵.

³ The lack of a guarantee that consumer trust will be preserved is one reason why planned EU legislation will potentially put severe limits on data sharing <http://www.dma.org.uk/content/eu-data-protection-reform-could-be-delayed-until-2015>.

⁴ See www.bbc.co.uk/news/technology-25231757

⁵ For example the recent protest by leading authors against misuse of personal data by spy agencies: www.theguardian.com/world/2013/dec/10/surveillance-theft-worlds-leading-authors

The main way that trust can be seriously eroded is when consumers' interests are actually harmed by organisations that use their data against them. One example is personalised pricing, where firms charge higher prices⁶ or refuse to service certain market segments based on an analysis of their data. Potentially this could also happen in the insurance sector with high risk people being forced into paying very high premiums, or being left unserviceable, and in the healthcare sector with genetic information being used to bar some people from particular services.

Essentially the problem is that some consumers have different characteristics that make them less easy or less profitable to service. Any organisation that is left free to optimise its choice of target consumers will target the easiest or more profitable ones. But this is possibly unfair or even immoral if some customers are barred from a critical service based on the data that they themselves supplied. And news of it will erode trust across other types of customers.

2.1.2 Preserving consumer trust using legislation and regulation

We discussed how our society seeks to preserve consumer trust using legislation and a system of regulators to implement it. But current legislation needs to be updated to be able to preserve consumer trust in the face of new technologies and new uses of consumer data. Right now, legislation that was written for a different era is being rewritten but it is like aiming at a moving target.

The problem is that the types of consumer data that are recorded, the data's current and future uses, the methods of using it and the current and future users of it are constantly evolving. The interdependencies between data types, uses, methods and users make it very difficult to see how it will unfold. So we need to use perspectives that avoid these interdependencies and we need to seek architectural rules that help us break up the system into less complex components.

One perspective that avoids the interdependencies of guessing which different ways events can unfold is to choose desired outcomes and then work backwards – like looking how to preserve consumer trust. One set of architectural rules that helps us to break up the system along logical 'loosely-coupled' lines is to think in terms of the different stakeholders in this system and the organisational levels that they operate on, i.e. the consumers and staff on the individual level; the firms and other corporate entities on the organisational level; and the stakeholders that have market-level roles like regulators and privacy and consumer interest groups. All these stakeholders seek to influence the other levels but working across more than one level is an extreme stretch of resources unless the stakeholder in question is designed to incorporate some characteristics of the different levels that it seeks to work on.

In addition to a problem with current legislation there is another difficult challenge for regulators. Regulators support consumers and firms as part of a system with trade bodies and other organisations partnering with them. But our regulatory system is missing one or more entities, e.g. the ICO is currently consulting on a framework for endorsing a provider for a Privacy Seal scheme⁷. A third party provider for a Privacy Seal scheme would award trust marks to participating firms and check their compliance on the behalf of the ICO. This co-regulator approach splits activities into two levels. The ICO works on the strategic governance level and a privacy seal scheme operator on the firm-level to actually interact with firms and deal with consumer complaints.

⁶ See http://www.ofc.gov.uk/shared_ofc/markets-work/personalised-pricing/ofc1489.pdf

⁷ http://ico.org.uk/news/current_topics/2012.

2.2 The problems with sharing data

In our roundtables we talked a lot about how sharing data can erode consumer trust but actually firms find it quite hard to share data – *the right data* that is. Sharing data in an appropriate way that protects and helps consumers, and obtaining the appropriate data in the first place, are two sides of the same coin.

The actual ability of firms to get hold of the data that they need is limited. Firms have lots of data on their customers. But there is always a missing piece of the data puzzle and by definition they have very little data on their non-customers because they have little interaction with them. And the most interesting customers to learn about are always their non-customers.

The public sector has fewer non-customers, as a proportion of the population, and potentially much easier access to other agencies and services because there is less fundamental competition. But just making public sector data available does not solve the problem – agencies and firms need to find the specific data that *they* need and there are an increasing number of Open Data sources set alongside a limited awareness of which open data repository has the specific data that *they* are looking for.

Making data available for sharing is laudable and a prerequisite for our evolving Big Data society but an awareness of where the right data is and what to do with it is also required.

3 A Third Party to promote preserving trust through appropriate data sharing

Our roundtable discussions frequently highlighted missing roles in the emerging Big Data ecosystem. We kept coming up with activities that were needed but that did not quite fit the capabilities and roles of the parties that currently exist. So we started to refer to these missing roles as if there was some other organisation that fulfilled it. We called it the 'Third Party' in the sense that there should be some third party to do things on behalf of consumers and firms that use consumers' data. There are other important stakeholders involved, like regulators, consumer interest and privacy groups, so in a sense the label should be 'Fourth Party' or 'Fifth Party'. But the label stuck even if it refers to more than one 'Third Party'.

We came up with several key roles that are either unfilled or could do with more capacity:

1. Helping consumers

- **Education:** an educator that teaches people what they need to know about the implications of sharing and using data in our unfolding Big Data society, e.g. by telling people what they need to be aware of depending on how deeply they wish to understand how their data is used. Helping people to be aware of what happens when they agree to the fine print whilst downloading apps or when they sign up to online services. Helping people to judge how much freedom they want to give to organisations that ask to use their data. Helping people with different levels of detail depending on how deeply they want to understand the issues surrounding the uses of their personal data.
- **Manage:** a manager of the ways that personal data is used on consumers' behalf, e.g. by supporting and encouraging the developing market in personal data lockers, personal data stores and facilitators of personal data management⁸. Helping people to use their own data for their own benefit and the benefit of others that they choose. Helping people to be aware of what they could allow their data to be used for and how this might be beneficial to themselves and others. Helping people who want different levels of effort, involvement and sophistication in how they decide to manage their personal data.
- **Share value:** a manager of how value is generated and shared from the use of people's personal data, e.g. by explaining how people can benefit from sharing their personal data with organisations that can be trusted and how to obtain better access to such benefits. Helping people to be aware of how the uses of their personal data benefit themselves and others. Helping people to decide how data is used to produce value and how this value is then shared. Helping people who want different levels of effort, involvement and sophistication in how they decide to manage the value that is generated from their personal data.
- **Defend and fix harm:** a defender of people against harm and a fixer of such harm, e.g. by advising people about which organisations are trustworthy, by regularly checking the trustworthiness of such organisations and by reversing the effects of harm caused by the misuse of people's data. Helping people to be aware of how they could be harmed and of how to fix or mitigate such harm. Helping people to decide how to avoid and deal with harmful situations. Helping people who want different levels of effort, involvement and sophistication in how they decide to avoid and deal with harmful situations associated with their personal data.

⁸ For example: mydex.org, personal.com, giy.nl and midatalab.org.uk.

Helping people to access insurance and data correction or recovery methods, to reverse harm.

2. Helping firms

- Staff guidance: an educator that teaches staff in firms what they need to know about the implications of sharing and using consumers' data, e.g. by communicating and explaining a code of conduct and a set of industry-wide principles. Helping firms and staff to understand the 'grey areas'. Helping staff to be aware of what happens when they use their customer's data in different ways.

Sharing: a catalyst that helps firms to find ways to work together and monetise their data or to get hold of the specific data that they need to reach their most important insights, e.g. a trusted third party for sharing data in an appropriate manner that protects the consumers that generated the data. Helping firms to find the missing piece of the data puzzle by introducing them to the organisation that is able to give them access to that data legally and ethically. Helping firms to monetise their data assets in line with the latest regulatory frameworks.

Data sharing is the activity that is of common interest to all stakeholders and how it is done differentiates value creation from harm. This is why a 'Third Party' role that spans all stakeholders and helps data sharing can also support the other aims that are listed in this white paper.

- A trust mark: an awardee of a badge that shows that a firm has regularly proved that it uses consumer data in line with the latest regulatory frameworks, e.g. the firm could display a logo to show its trustworthiness to consumers in return for regular checks on how it uses their data. Helping firms to prove to consumers that they are trustworthy. Helping firms to develop systems of governance and processes that lead to trustworthiness and appropriate uses of data *as well as to more effective uses of data.*

3. Supporting regulators

- A consumer portal: a large-scale source of information, advice and guidance for consumers as well as a place to complain to, e.g. a two-way channel for registering complaints, getting action and answering questions. Helping regulators by acting as an interface between systemic or market-level phenomena and consumer-level issues, problems, complaints and questions.
- A firm level portal: a large scale source of information, advice and guidance for firms as well a place to complain to, e.g. a two-way channel for influencing regulatory change and checking compliance. Helping regulators by acting as an interface between systemic or market-level phenomena and firm-level issues, problems, complaints and questions. Helping firms to use data appropriately by making it *in their best interests to do so*, i.e. inappropriate use of data would bar them from using the sharing service and from holding a trust mark. So they would be cut off from the data that their need as well as *the* major proof of their trustworthiness.

4. Supporting privacy and consumer organisations

- Transparency: giving privacy and consumer organisations a direct view of how data is shared at the scales that the Third Party would seek to operate on. Helping to ensure openness and influence, e.g. a non-executive role to act as a neutral, non-commercial and apolitical eyes and voice in the Third Party's board.

The ultimate objective is to preserve consumer trust and access to external thinking from experts in the privacy and consumer rights community is of paramount importance because this domain is complex and evolving quickly. External thinking always provides a clear way through local complexities.

4 Conclusion

The defining characteristic of 'data' is its specificity which means that different data and different uses of data create different outcomes. The value of data comes from its use, as does its harm, and whether value is created or harm is done depends on exactly what data is used and exactly how it is used. The reason for this is that ultimately, data is used to help us generate options and choose between them. It is data that enables us to uncover and choose between available resources according to our personal context.

The problem of preserving consumer trust whilst promoting data sharing

As our Big Data society develops we will find new ways to use data and these will generate new services and new ways of creating value for consumers. All this will involve sharing data with organisations that have new ideas for using it and are that are better positioned to use it. But sharing data makes it harder to preserve consumer trust because it multiplies the ways in which data might be used.

The problem is that data sharing hugely expands the value that any particular data might generate but it also makes it very difficult to predict who could use any particular data or what they might use it for. On the one hand the more we share our data then the more it can be used but on the other hand the more likely it is that it might be misused as well.

The problem of data sharing is that it involves more people and organisations than can be easily watched or controlled; by either the consumers that generated the data or even the regulators. Consumers can engage with one retailer that they bought goods from and which generated data. But it is difficult to engage with all the retailers that they go to and impossible to monitor and deal with all the supply chain partners of all these retailers.

For firms

Part of the solution to this problem is getting organisations that use consumer data to watch and control themselves and each other. We can make it in their best interests for firms and the staff in them to use data appropriately *by helping them to use it successfully*. This would motivate firms and other organisations to behave appropriately on the firm level and on the level of each individual staff member.

For the staff this means providing education and advice about the implications of using customer data as well as how to do it successfully. For the firms this means helping their staff, giving them access to external data sharing opportunities and giving them a trust mark to demonstrate their trustworthiness to their customers.

For the regulators

Market regulators are designed to work at the market-level, to guide and nurture whole markets. Of course they interact with individual consumers and staff members but they are creatures with complex responsibilities that span the whole market. They have limited resources that must be focused on systemic and large scale issues and they are anchored to the market level by their basis in a market-level legal framework. So reaching down to help individuals on a large scale is difficult.

There is a role for a Third Party that supports the regulators by helping firms to be more successful, by helping the firms' staff to be successful and by helping consumers to trust and be confident in how their data is being used. Supporting consumers and all other stakeholders according to their different needs and levels of focus is the most robust solution to the problem of preserving consumer trust whilst helping data sharing.

The reason for this is that it is a multi-level solution for a multi-level problem. It engages the attention and resources of each of the stakeholders to watch and help each other. Firms have a strong interest in behaving appropriately in order to preserve their access to data services and a trust mark. In turn they encourage their staff to behave appropriately and use help from the Third Party to do so; and be more successful in the process.

For privacy organisations and consumer organisations

The objectives of the role of a Third Party and the objectives of privacy and consumer organisations overlap most strongly on the preservation of consumer trust. Consumer trust is the prerequisite for scale and sustainability in data sharing in our emerging Big Data society. We know data will be increasingly shared but do not know how exactly what form it will take. The journey towards a Big Data society is complex and it is accelerating. So the advice and perspective of an external viewpoint would help to steer a successful course as much as it would guarantee trustworthy behaviour by the Third Party itself.

Using self-interest to help each of the stakeholders

The missing roles in our emerging Big Data ecosystem are roles that span levels and link stakeholders together more strongly. They are:

- a supporter of individual consumers to deal with networks of large and small firms;
- a helper of firms and other organisations to share and use data in new ways in return for doing so appropriately;
- an aid to the regulators that helps them to bridge the inter-level gap between macro and micro, between the market-level and individual consumers, and between staff and firms;
- and a promoter of the objectives of privacy organisations and consumer organisations that gives them a platform to help more consumers and engage with more firms.

For consumers, the ability to engage with firms and to test them for trustworthiness in how they use consumer data would help to generate consumer trust. This is the main outcome that we need to support the development of our Big Data society. A Third Party organisation that sits at the intersection of consumers, firms, regulators, privacy organisations and consumer organisations is fundamentally best placed to support consumers with the resources of all the other stakeholders.

For firms, help in enabling them to share data – to get the right data and to use it more successfully – would produce some income for the Third Party that could be used to fund the other roles. Helping firms in this way would motivate them to behave appropriately since non-compliance would bar them from data sharing opportunities as well as from any other advice and assistance on how to use data. The advantages of working with a Third Party far outweigh any compliance efforts and even these are paid for by a trust mark that increases their customers' trust in their brands.

For regulators, this Third Party would help regulators to extend their reach deeper into the micro levels of the market in terms of more individual consumer connections – two-way connections – and more frequent or timely interactions. This would help regulators in their efforts to deal with group and systemic issues. A Third Party would check process quality and the compliance of its member organisations and a regulator such as the ICO would check the Third Party's own process quality and adherence⁹. The Third Party would enhance the interface between the regulator and the firms in the market by supporting regulatory objectives at well-defined touch points and times. Also, a trust mark and audit scheme that is linked to the data sharing and assistance services that are described above would be more likely to have greater participation and compliance. Making membership attractive would make it scale up more quickly and more robustly.

For privacy organisations and consumer organisations, the Third Party would be a platform to augment what they already do, a direct access channel to monitor large scale data sharing activities and a means of ensuring that the Third Party itself is trustworthy.

⁹ This might be part of the trust mark service called 'Privacy Seals' that the ICO is currently considering.

5 Next steps

This white paper discusses several missing roles in our emerging Big Data society. A missing 'Third Party', 'Fourth Party' or 'Fifth Party'.

Now I propose that we fill these roles by following this process:

1. Publish this white paper as you see fit. Pass it on to the people that you think will be interested in its ideas.
2. Contact the author to register your interest and describe the capabilities that you can contribute to this project.

6 About the author

Duncan Shaw

After working in logistics, for the Total Group, and in engineering Duncan did an MBA and went into management consultancy, where he worked for clients including Xerox, Coca-Cola, Danone and Shell on business transformation projects.

Then Duncan joined Motorola and was made Customer Satisfaction Manager for the EMEA region and reported jointly to the Marketing Vice President EMEA and the Director of Quality EMEA. He led a team of internal consultants that worked with customers and across Motorola's country teams and business functions. The team profoundly increased Customer Satisfaction for Motorola's \$2 billion handset division in Europe and revolutionised relationships with customers that accounted for 90% of annual business.

Now Duncan is a Lecturer in Information Systems at Nottingham University Business School and his research and consultancy interests are focused on uses of Big Data analytics, customer loyalty, customer relationship management, service design, and network orchestration. Duncan consults to private companies and public sector organisations and teaches workshops for senior business people and executive MBAs.

Duncan has a PhD in modelling Network Orchestration and an MBA from Manchester Business School, an MSc in Logistics gained whilst working for the Total Group and a BEng Hons degree in Engineering. He blogs at www.duncanrshaw.co.uk.

Contact: duncan.shaw@nottingham.ac.uk.

Appendix 1: Themes from Roundtable Two on 5 September 2013.

The themes that arose in our discussions and the suggestions for next steps that we all talked about.

1. The fundamental requirement for consumer trust and guidance for staff and organisations.

We talked about the importance of consumer trust as a pre-requisite for consumers giving organisations their data and then allowing them to use it. Trust is a basic foundation of the new data services that organisations are developing and any damage to this trust would hinder the development of our future Big Data society.

The attendees felt that larger and more prominent organisations would be less prone to misusing personal data. But there is a risk that some of staff, like any human being, might make the wrong choices of how to use data even if they do it out of loyalty to the organisation or in seeking to benefit their customers.

We discussed how Aimia uses a set of values and best practices called TACT* to help to guide all their staff into doing what customers would expect and to ensure that they put customers first. This could possibly be used as a start in developing guidance for organisations and their staff in the digital supply chains that are now forming – to promote awareness of the issues, appropriate behaviours and the implications for the people concerned.

* http://www.aimia.com/files/doc_downloads/WhitepaperUKDataValuesFINAL.pdf

2. The problems caused by the limitless mixing, use and reuse of personal data.

We noted that the potential uses of 'base' data, like names and ages, to create *inferences* are a form of using data to create new data. And the ownership, usage rights and capabilities of the new data are even more unclear than those of the base data – especially if this new data is then used to create more inferences.

This problem applies to finding beneficial ways to use the data, designing frameworks of appropriate use and even communicating guidance about appropriate use to staff and organisations that use such data, e.g. making staff aware of the many potential implications and unforeseen consequences of using the data.

We discussed how some of these complexities can be avoided by talking about the *use* of the data rather than 'the data' and by using a 'risk-based' approach. A risk-based approach seeks to deal with outcomes and the probabilities of these outcomes rather than just a simplistic focus on the tools and resources that could be used in many different ways.

We talked about how this subject has for some time been a critical issue in the world of credit reference agencies. Perhaps we might learn from their longer experience of this area, how they share data, how they deal with negative consequences for customers and how they ensure appropriate corporate behaviours.

3. More detail on the Third Party(s) role and its potential forms

We continued to develop a notion of the currently unfilled niche, or niches, in the big data ecosystem that we have labelled the 'Third Party'. There is a need to help to reduce the danger of losing consumer trust by helping to ensure the appropriate uses of personal data. On the customer side, this could be educating customers to raise awareness of the issues; helping them to control, manage and use their data; and as insurance, or a safety net, to helping people to fix things if they go wrong. On the data supply chain side, i.e. the public and private organisations that together provide services, this could be educating staff about appropriate use of data; supporting regulators and helping to ensure appropriate corporate-level behaviours – especially in newer and smaller firms.

It seemed that what is required in order to ensure consumer trust is much more like a group of people than just a passive and inflexible declaration of good practice. Certainly there is a strong need for clear

descriptions and awareness-raising communications, in plain-English. To what extent are customers actually making real and informed choices? It is very difficult to make people aware of all the parameters and issues involved in their decisions.

Also, it is not possible to forecast the different ways in which any data will be used by organisations in the future and what guidance customers may need. So an actual group of people, expert stakeholders, is required to watch, to analyse and to act on these issues.

We talked about various *forms* that the Third Party or Parties could take:

- a trade body – which we heard is already being formed with some support from government
- an exchange - stock exchanges help to ensure appropriate behaviours as a requirement for membership
- a supporter of a regulator or programmes like midata – new roles are being discussed internally but there was an interest in understanding how government and other parties could evolve their relationships

These different forms of a Third Party would act to *protect customers' trust* and also to speed the development of our Big Data society, e.g. by helping organisations to release savings via increased precision in the use of marketing budgets, operational resources and new product development expertise. They could also speed development by helping organisations to create wholly new classes of services via a much better understanding of customers' lives. They could also contain capabilities such as a centre for best practice, a code of conduct or a form of accreditation

4. A paper to develop our ideas on a Third Party(s) and how to create them

We all agreed that it would be worth producing a paper to develop our thinking on what a Third Party could and should do as well as the forms it could take and the ways of implementing these. This would take in the experiences and expertise of all the attendees, help us to develop our thinking in the areas that different people are working on right now and to potentially move the wider debate towards solving some of the problems that we have highlighted.

Duncan will arrange to interview the attendees individually, to understand their perspectives and the questions that this paper could help them think through. First he will talk to the attendees from government departments followed by the commercial attendees. Other experts will be interviewed as needed. He will present the findings at the next roundtable.

Appendix 2: Themes from Roundtable One on 12 March 2013.

1. How to make a basic framework for achieving the societal benefits of Big Data services without compromising consumers' basic rights to privacy and data protection?

We talked about how the benefits and the risks of Big Data technology are significant and how one risk is that a lack of consumer protection will lead to some 'horsemeat' moment that will harm consumer trust and actually slow down the development of these services. But a better understanding of how benefits and risks can be balanced may help the development of Big Data services, e.g. as a logic for making EU and other legislation more flexible in this area.

Several of us noted how complex the issues are and some common points were:

- There are different types of data – Facebook data may be less sensitive than health data
- Not just Big data – 'small data' use also has the same potential benefits and risks
- People exist on a spectrum of sensitivity with some not caring about data privacy and some being highly sensitive
- The same data could be used in different ways, some of which might be objectionable to the person who originated it. In a related point we noted that the future uses of any data is unknown and still developing
- People have different levels of awareness and knowledge of these issues – although most have very little understanding of Personal Big Data and its implications
- People rarely read or understand the 'small print' in user agreements and thus expose their data, e.g. in downloading phone apps
- When things do 'go wrong' what comeback do consumers have? How can things be 'fixed'? Can this 'fixing' at least be speeded up to minimise harm and give support for those that are wronged?
- Some issues can be solved; you can do some things but not all – what are 'the basics'?

We thought it more practical not to aim to develop some comprehensive model, but a basic framework of initial guidelines would be a valuable first step where little presently exists. This list of the various dimensions of complexity might be the start of a basic framework that we can develop.

2. What is the role or roles of a 'third party' that does not exist yet?

The idea of some 'third party' came up many times in our discussion – a broker that would manage a person's data and deal with firms on their behalf; a defender of a person against harm; a fixer of such harm; an educator that teaches people what they need to know about our unfolding Big Data society; a new form of regulator to uphold public interest; or even a third party 'dating agency' for firms and their data.

These third parties roles might not exist within the same organisation and it could be that we were noticing vacant niches in some sort of 'Big Data ecosystem'. Indeed some of these roles may be taken up by regulators or they may fall to multiple competing Third Parties rather than a single organisation.

We noted how Google, Facebook and others are aiming to take up some of these roles and how this might not necessarily be for the best. It may be that such a third party may be some new organisation which is made up of a grouping of different organisations which combine the capability to be trustworthy as well as the required technical and commercial capabilities.

Such a grouping might combine the aims and values of the people who were invited to our roundtable. And the presence of one trusted 'good' player, of a big enough scale, may raise the standards of the others.

3. How to motivate firms to share their data?

Data creates value when it is used and different organisations could use the same data in different ways. So firms 'opening up' their data would create greater benefits than keeping it hidden and underused. But we talked about how firms see this as giving away the 'crown jewels'.

This has a parallel to the issue of individuals giving away their data in return for some fair recompense. Not 'sweating' their data assets is an issue for individuals as well as firms and we talked about how it is common practice in the credit checking sector to swap data in clubs like those run by Experian. Maybe this will provide some best practice ideas.

4. The reform of EU data protection legislation and global Big Data

We talked about the potential form of new EU legislation, how different mainland European countries have very different attitudes and histories when it comes to data protection and how this may impact us in the UK.

We also discussed other factors that will influence how this develops, that are part of what is actually a global Big Data revolution.

5. The relationship between personalisation and fairness

There was a concern about the appropriateness of the algorithms that have the power to make decisions over what offers are made to a consumer, e.g. should speeding tickets be proportional to salary? Is personalised pricing fair? Especially since the aims of the firm may not always be consistent with those of the individual. There was also concern about the accuracy of the data that is used.

Some industries have a long history of tailoring their services, e.g. insurance, and some have recently become less personalised, e.g. women's car insurance. But increasingly we will see more and more sectors use Big Data analytics to tailor their offers and this will raise the sort of issues that personalised pricing is raising now.