September 2013

BIG INNOVATION CENTRE

Big data: A lot to explore but little to share?

Ways towards a regulatory regime for big data

Gareth Quested and Sonia Sousa

The Big Innovation Centre is an initiative of The Work Foundation and Lancaster University. Launched in September 2011, it brings together a range of companies, trusts, universities and public bodies to research and propose practical reforms with the ambition of making the UK a global open innovation hub as part of the urgent task of rebalancing and growing the UK economy, and with the vision of building a world-class innovation and investment ecosystem by 2025. For further details, please visit www.biginnovationcentre.com.

Big data: A lot to explore but little to share?

Some ways towards a regulatory regime for big data

This policy briefing is an output of the event organised by Dr. Sonia Sousa, Big Innovation Centre, as part of the Big Data Week 2013, London.

We want to thank the team responsible for the organisation of the Big Data Week 2013, London for hosting the event.

The event was made possible by the support of the Economic & Social Research Council.

Executive summary

This policy briefing summarises the findings of a panel discussion on what policymakers can do for the UK economy to succeed in securing the economic benefits of big data. The event was organised by the Big Innovation Centre and the Economic & Social Research Council in April in London, as part of the Big Data Week 2013, a worldwide festival. This debate benefited from the insights of a highly engaged audience of about 200 people and was live-streamed across more than 20 cities in 16 countries.

Questions addressed

Chaired by Will Hutton, Chair of the Big Innovation Centre, a panel of six experts in data protection and intellectual property regimes, data scientists, entrepreneurs, and leaders of public bodies addressed the following questions:

- How can big data contribute to solving economic problems?
- Is policy fit for purpose? Or does it hold us back from reaping the economic potential of big data?
- What sort of progressive law-making do we need? What could deliver the regulatory regime that would enable using, sharing and combining private and public data, while addressing privacy and security issues?

Findings from the discussion

In answering these questions, the following major findings emerged from the discussion:

- There are many areas in which big data could boost innovation, new business models and better public services delivery but policy stood in the way. Examples include:
    - Better city management by making use of so called ‘smart cities’. City management could be more cost-effective for tax payers if city managers were able to make full use of the data generated in their cities;
    - Solving social problems of “complex families” at the council level. Better and faster services could be delivered if the various agencies and private contractors involved in the same process were able to share data with one another;
    - Improved health outcomes by combining and sharing patient data. More effective targeting of resources, resulting in both improved services and efficiency, could be delivered if health data could be combined and shared.
- The current legal framework for data sharing is not fit for purpose as it is ambiguous, confusing and difficult to comply with;

- There is little, if any, consensus on what needs to be changed to make the laws governing the use of data fit for purpose in exploiting the value of big data. Some contrasting views set forward include:
    - Adopting a 'wait and see' approach to avoid changing the law too fast versus adopting a progressive law-making approach instead;
    - The required changes need simple amendments to the law, opposed to the view that a fundamental change to the law is needed instead.

- Any potential legal changes will need to be accompanied by a clear explanation of the economic benefits to individuals and to society as a whole, and what safeguards need to put in place. This can be an effective way of overcoming cultural and institutional barriers to data sharing.

Policy recommendations

What exactly needs to be done to put in place a fit-for-purpose policy framework for big data? Although a one-hour debate was too short to reach firm agreement on policy recommendations, interesting themes emerged as possible ways forward:

- Firstly, build on the Creative Commons legal infrastructure to develop a Privacy Commons;
- Secondly, adopt an 'opt-in unless you opt-out' approach to personal data disclosure;
- Thirdly, the regulatory regime governing big data usage should be contingent upon what the data is and what it will be used for;
- Finally, the regulatory regime for big data should avoid loopholes, be easy to understand and to comply with.

Table of Contents

Introduction

Big data is a new general purpose technology, with the potential of huge economic opportunities that could transform our lives. Simply keeping up with the numbers is a task in itself. The figures below show the magnitude of what data is available, but with the rate at which data is growing, these figures are already out of date:

- 30 billion pieces of content are shared on Facebook every month[1]
- Google processes 24,000 terabytes of data every day[2]
- There are more than a billion transistors *per human* in the world[3]
- Only 0.5% of data in the "digital universe" is estimated to have been analysed[4]

But for all the discussion on the potential economic benefits of big data, there has been little time spent considering the need for a policy framework fit for purpose. Without this, we will miss the opportunity to make the most out of big data. How do we balance reaping the benefits of big data with protecting the privacy of citizens? Should we even consider this in the first place? What sort of progressive law should we put in place?

These are just some of the questions that were addressed at our event at the Big Data Week 2013[5].

Big Data Week is an annual, week-long international event looking at all things relating to Big Data. In April 2013, the Big Innovation Centre and the Economic and Social Research Council (ESRC) hosted a panel discussion in London for Big Data Week. The event had an audience of about 200 people and was live-streamed across more than 20 cities in 16 countries. Interestingly, among the hundreds of events hosted around the world as part of the Big Data Week festival, our event was the only one discussing the regulatory regime for big data. Clearly, the need for a fit-for-purpose big data regulatory regime has been largely absent from the big data agenda, which has so far been dominated by discussions on technological issues.

In this policy briefing, we will look at the major conclusions and policy implications of the event. Firstly, what practical challenges were identified that can lead to developing

---

[1] Manyika et. al (2011), "Big data: The next frontier for innovation, competition, and productivity", McKinsey Global Institute, p.vi
[2] Davenport, T., Barth, P. & Bean, R. (2012), "How 'Big Data' is Different", in MIT Sloan Management Review, Vol.54, No. 1, p. 22
[3] Zikopoulos et.al (2012), "Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data", McGraw Hill, p.xxiv
[4] http://uk.emc.com/leadership/digital-universe/iview/executive-summary-a-universe-of.htm
[5] See the video of the event here: http://www.biginnovationcentre.com/Media/video/346/Big-Data-A-lot-to-explore-little-to-share

interesting case studies? Secondly, what legal hurdles are holding us back from creating value out of big data? Thirdly, what social and economic issues were raised that highlighted areas of concern, and how did the audience vote on a number of questions put to them? Finally, we will lay out tentative policy recommendations and possible ways forward.

The panel discussion was chaired by Will Hutton, Chair of the Big Innovation Centre, and the panellists were:

- Birgitte Andersen, Director, Big Innovation Centre
- Roger Bickerstaff, Partner and International Head of IT, Bird & Bird
- Tom Heath, Data Scientist, Open Data Institute
- Daniel Hulme, Chief Executive, Satalia
- Peter Reid, Chief Executive, Digital Economy Catapult
- Chris Widgery, Head of Strategy and Support, Camden Council

1. Practical challenges

Interesting practical examples of areas where big data could unleash value but was held back by policy were raised by both panellists and the audience.

One example raised was ‘smart cities’ and the prospect of city management using big data. More integrated cities could be more energy efficient, with better transport. By using data properly, citizens would be able to increase engagement with each other and government. Through a combination of both aggregated and personalised data, quality of life could be increased and city management could be more cost-effective for tax payers. Although some of these concepts are already in use in some cities around the world[6], policy will need to change so that city managers make full use of the data being generated and make fully integrated ‘smart cities’ a reality.

A second example is the work developed by councils when dealing with “complex families”. Camden Council, for example, has to deal with a number of agencies and private contractors while addressing social problems. However, while working together to solve the same social problems, these partners cannot share data among themselves. Camden is given data by central government, but has not been allowed to share this with a range of partners on their project. Although the data involved is not necessarily ‘big’, it highlights that the sharing of data, vital for our exploitation of big data, is often not possible despite clear benefits for doing so. Data within the council and other government departments is often ‘siloed’, so that different sections of the same council are not able to share data with other areas, even when working towards a common goal. From a privacy perspective, many of these cases should not have any problem sharing data. Yet, because of institutions erring on the side of caution, data privacy is seen as a default.

Why do councils find it so difficult to share data among partners working on the same project? Two reasons stand out; one is that councils are not expected to share data with profit-seeking companies. The other is that the combination of different sets of data could result in revealing unintended information. This particular concern cuts across sectors and it was clear from the discussion that it is a major concern of both panellists and audience alike.

A third example is how to use health data to provide better care services. One challenge would be to find a way to safely combine the data of many individuals. At a personal level, it would enable people to have greater ownership of their own health data – and, for instance, grant them access to the health data of close relatives, with consent. At a local level, this could be used to locate and personalise social services more effectively as to meet the

[6] See some examples here: http://www.bbc.co.uk/news/technology-22538561

needs of the local population; and helping schools manage the specific health needs of their students. These approaches would not only cut costs for the NHS, but also provide individuals with better services.

All in all, these practical examples show that there are many benefits from sharing, combining and exploring data together with partners addressing similar challenges. This means that society needs to find ways to share data so that it can benefit from its full potential.

2. Legal hurdles

What are the major legal hurdles that are preventing organisations from sharing, combining and exploring data for public good? The discussion raised a wide range of views on the role of the current legal framework for data sharing.

A common concern was that the current regulatory regime governing the use of data is confusing to understand and, as a result, is damaging the potential for developing new ways to take advantage of big data. This concern was common to both the panellists and the audience. As several interventions from the audience pointed out, the law is often almost impossible to get one's head around. Due to the lack of clarity over whether or not their core business model would be legal, many potential entrepreneurs with great ideas of how to exploit new opportunities of big and open data simply do not go ahead with them.

A second common concern was that guidance from government departments is quite often different from what the law on a given topic says – either directly contradictory, or for instance requiring greater regulation than the legal minimum. This leads to a lack of clarity about what is allowed and what is not.

A third common concern revolved around the ambiguity within the law. As the British legal system is based on case law rather than a formally written constitution, the concern is that there will have to be a first 'sacrificial lamb' who would take the risk of losing in court in order to prove whether their business model was legal or not.

The discussion also revealed that there is little consensus on what changes need to be made to the law in order to make it fit for purpose in exploiting the value of big data. One view was that there is a risk of changing laws too fast, in danger of trying to fix a problem before we are fully aware of what new regulations we need. Instead, it would be better to 'wait and see' for a clearer picture of what needs changing.

Another area of disagreement relates to whether changes to the law simply require the amendment of existing legislation, or instead require a more radical overhaul. Intellectual property (IP) law, for instance, currently has a number of exemptions. For some, increasing the number of these exemptions would be one way to address current legal hurdles to data sharing. For others, IP law as it stands is simply anti-competitive and increasing the number of exemptions is not enough. According to this view, we need a fundamental change in our approach to IP.

All things considered, confusion over the interpretation of current law, and perceived gaps in addressing issues around privacy and data protection, means that we are failing to take full advantage of the potential of big data.

3. Social and economic challenges

Another major conclusion which emerged from the event is that any potential legal changes will need to factor in the response of society more widely. If there is a change in the law to allow more open access to data, it will need to be clear what the economic benefits to society are, and what safeguards are in place to reassure people of the safety of their personal information.

Without public support for opening up data, it will be quite possible that a counter reaction will lead to a backlash, possibly causing the law to be put back decades.

As many participants both on the panel and in the audience rightly pointed out, cultural and institutional barriers will play an important part in the way the legal framework will be shaped. For instance, even when data is legally allowed to be shared, it sometimes is not because organisations decide not to do so. This may be due to concerns over privacy even when there is no legal restraint. Although most of the participants seemed to accept the premise that more open data is generally a good thing, they also raised issues, particularly around privacy, that needed to be addressed.

Quite often, a lack of awareness around the potential of open data means that society may not prioritise its development. As many in the audience suggested, one way to encourage people to share their own individual data, or support open data generally, is to demonstrate and explain the benefits that will accrue to them personally, and to wider society. By doing so, it was argued, the case for more openness could be made much more easily.

However, it was clear from the discussion that any move to *allowing* more data to be open must be accompanied by a cultural push to ensure data is opened up in practise when it can be.

To better understand how people feel about what data might be shared, who it might be shared with, and for what purpose, the audience were asked, firstly, whether they cared about a company like Google knowing so much about them; and secondly, whether they cared about a local council, such as Camden, knowing the same.

Perhaps unsurprisingly, most people in the audience were concerned about a corporation such as Google having so much access to their personal information, but very few were concerned about a local council doing so. In trying to understand this disparate outcome, two possible explanations arose. Firstly, people tend to be more trusting of the motives of a public sector organisation, feeling that public bodies are more likely to respect an individual's privacy. Secondly, people believe that data released to public bodies (as for example local councils) will be used for a more worthwhile purpose – the benefit will not go solely to the organisation, but for the wider public good.

One concern raised about increasing openness to data was that it may lead to people avoiding keeping a record of certain data, if this record is to be released. For instance, if a patient is given complete access to their medical records, doctors may be more reticent about what they put on patients' notes.

From an economic perspective, big data, it was argued, should be viewed as a factor of production that can only be used through Open Innovation – we need to treat data as an asset. Furthermore, by looking at practical challenges and how openness can create value, we would more easily be able to compare moral arguments over privacy with economic value.

Finally, from an economic perspective and especially in the context of tightening budgets, it was also stressed that big data is a real opportunity for a government, including local authorities, which has "to do more with less."

4. Policy recommendations: possible ways forward

What exactly do we need to put in place a fit-for-purpose policy framework that allows for the potential of big data to be unleashed?

Within a relatively short debate, there was no time for a firm agreement on policy recommendations, but clear interest emerged around four areas that can be explored further:

- Firstly, should we build on the Creative Commons legal infrastructure to develop a Privacy Commons?

  One view that emerged from the discussion was that privacy is a social construct and should be challenged. According to this view, whilst privacy is a perfectly reasonable goal, we need to be aware that society's perception of what is private can, and does, change over time. As a result, we need to consider, as a society, what we see as private and why. Are there some things we should be less private about, when the benefits of doing so far exceed the costs? When the people in the audience were asked to vote on the question, the vote was approximately 60 in favour of the statement and 50 against. As the Chair summarised, the "protect privacy at all costs" option is still rather prevalent even among big-data-oriented people.

- Secondly, should the UK adopt an 'opt-in unless you opt-out' approach to personal data disclosure?

  Another view put forward was that we should move away from a system where consent to share ones' individual data has to be explicit – that is, an individual has to actively say they are willing to share it – to a system where consent is assumed – that is, unless an individual says they are not willing to share their data, you can share it.

- Thirdly, should the regulatory regime governing data usage be contingent upon what the data is, and what it will be used for?

  As a participant put it, "What have I consented for my data to be used for?" For instance, people are more likely to be happy to share personal data on their travel patterns compared to say their health records, but are also more likely to be willing to share personal data with the NHS than a pharmaceutical company.

- Finally, how can the regulatory regime for big data avoid loopholes, be easy to understand and easy to comply with?

Acknowledgements

This report is a publication from the Big Innovation Centre, an initiative from The Work Foundation and Lancaster University. The content of this report reflects the opinions of its authors and not necessarily the views of the Big Innovation Centre or its supporters. The Big Innovation Centre is supported by the following companies, public bodies, universities and private trusts.

Contact details

Big Innovation Centre
The Work Foundation
21 Palmer Street
London SW1H 0AD

info@biginnovationcentre.com
www.biginnovationcentre.com
www.theworkfoundation.com