

ABRIDGED TRANSCRIPT OF JOHN PULLINGER'S TALK "MOBILISING THE POWER OF DATA TO HELP BRITAIN MAKE BETTER DECISIONS"

It's great to be back at DUG again. When I was here in 2008 talking about the topic of data sharing, Richard Thomas and Mark Walport had just published their data sharing report for the Prime Minister, explaining the benefit to the economy. The Statistics and Registration Service Act had just been passed and the UK Statistics Authority (UKSA) came into being.

The Eurobarometer survey had found that UK organisations were not trusted the by the public - much more so than for other countries. There had been a spate of discs lost in the post, laptops left on trains and memory sticks gone astray. The central question was: who can we trust to collect information from us and keep it safe?

What's happened since then? This data thing has really gathered pace, new and alternative sources of data are available. And we are only just starting: the pace is going to quicken and types of data sources will expand.

Government has got the message now. A year ago there was the Open Government partnership. Opening up data was important to government because, at a political level, there was the realisation that there was value to be gained if we can.

The UK Statistics Authority has come of age and, along with other organisations like Full Fact, has carved out a reputation for saying 'stop, don't mess with statistics, they are far too important for that'. So when the Prime Minister uses statistics incorrectly in a speech, here is someone there to say: don't mess with the numbers.

The groundwork is done, now is the time to mobilise the power of data to help Britain make better decisions - British parliaments and assemblies, governments, public authorities, businesses and citizens.

The best organisations are already mobilising data in real time. Is the difference between the seriousness of the Ebola outbreaks in Liberia and Sierra Leone the speed of reacting to data and mobilising forces quickly to respond?

We can already observe the best businesses drawing on all sources of data available to them, helping you to understand the customer. The core of their success is bringing together information to get customer insight, drive performance, understand and manage risk in a special way.

I'd like to share a few examples from my world. The first two were joint winners of the Royal Statistical award for Statistical Excellence and both were about data sharing.

The Ministry of Justice set up a data lab with really quite sensitive information about individual offenders. It was created to create a safe place for information to be shared with the people who needed to use it. It helped to answer questions and understanding for the benefit of people. How does a charity know how to help if it doesn't understand the problems faced by the people it's trying to help?

The Department for Energy and Climate Change combined data from a range of energy providers and government. It's a very neat system where you type in your postcode and type of house and you can find out how to save money on your bills. The government likes it because it can understand how policy is working; individuals like it because it helps them, it's very visual and presents information well.

My third example the work ONS is doing to help find solutions to problems, which are increasingly more complex and difficult. How we can best mobilise data much more adroitly, such as calibrating supermarket data to understand what price pressures are. There is an extraordinary richness we can use to produce better answers more quickly.

We are also looking at how we could use the location data from Twitter to improve our understanding of internal migration. This is particularly useful for information about students, where you can look at patterns and spot where people are, how long they are there for. Importantly it allows you ask questions.

Another source of information is mobile phone data. We are exploring how this can add to our understanding of transport flows which could be used to inform policy about the transport system and ask the question about whether it should be adapted.

There is quite a range of different sources we have access to already. We tend to start from the information we hold rather than the problem we are trying to fix and having access to this range is going to help us answer the questions we need to answer.

The next Census is going ahead after my predecessor's recommendations were accepted. There is agreement that it would be reckless to abandon the census without a viable alternative in place. But at the same time, we will need an aggressive look at administrative data.

Data sharing will be at the centre of the most fundamental evidence source in ways that previously wouldn't have been possible. Extraordinary things are happening. Data sharing is changing the way we do business. We need to make sure what do is done for good. It could go either way, good people are doing good things but there are too many people doing bad things.

How can we make sure what we are doing is good? I believe there are a range of things we can do together.

---

## 1. TO MAKE THE CASE

We must show how what we do has value, explaining why it is worthwhile. We heard the example earlier today in John Rae's talk about data sharing between mortgage lenders and how the benefits were made clear. We need more case studies which will inspire each other.

---

## 2. ENGAGE THE PUBLIC

We need to engage with the public and say how it will make life better; there is something in it for you. Unless we try to explain why we think this is a good thing people won't feel their understandable concerns are being addressed. We seem to be shy about doing this but our efforts will be doubly easy to undermine if we don't give them confidence in what we are trying to achieve.

---

## 3. GOVERNMENT

MPs are seeking to represent the interests of their constituents and we need to convince Parliament we can do it well or we won't have the laws we need to do all that we want. There is the Data Protection Act but is it enough in the current world? There is a recognition that data sharing and the use of information is part of the engine of growth but we still need to do more.

---

#### 4. PROVIDE ETHICAL OVERSIGHT

We have invested a lot in disclosure control but where you have an environment with more sources combined it becomes easier to identify individuals. What our children might be happy with sharing older generations may not be. What is right? We need to have an ethical framework, carefully looked after. Looking ahead there will be new technologies that change the landscape and we need a strong framework to protect individuals and make sure we are doing the right thing.

---

#### 5. BE GOOD STATISTICIANS

Remember you are a statistician if you are one and hire one if you're not. Understanding precision and biases, the qualities of these novel data sets and what inferences you can safely draw is important if we are not to misdirect people.

You may remember the flu trends algorithm Google produced. For the first couple of years it was accurate but then things started to go wrong because it was based on correlation equaling causality. This type of thing puts people off this work.

Statisticians need to get busy and add something to the picture.

---

#### 6. BE TRUSTWORTHY

We must think about the people who have access and make sure they do actually care about the information we are entrusted with. We need to make sure our computer systems protect information and rigorously think through risks. These protocols which can take so long to negotiate are important because without them we put the trust in what we do at risk.

It only takes one person to do one thing which can seriously undermine public confidence.

There are a lot of things we need to do. We need to fortify ourselves to go out and make the case. Only by sharing, within and between organisations will we create the social and economic benefits that will help governments, business and citizens make better decisions.