

PSI - The Next Ten Years

Submitted on 28 Jan 2010 by

We assembled an excellent panel to discuss Labour and Conservative PSI policies and Locus's perspective on issues surrounding PSI licensing, pricing and regulation. We also left plenty of time for attendees to ask questions, and raise their own issues for discussion.

Our panelists were:

Francis Maude MP, Conservative Shadow Cabinet Office Minister and Shadow Chancellor of the Duchy of Lancaster. The Minister leads the Party's Implementation Unit, which is tasked with determining how policies will be implemented should the Conservatives win the General Election;

Professor Nigel Shadbolt, Professor of Artificial Intelligence (AI) and Deputy head (Research) of the School of Electronics and Computer Sciences at the University of Southampton, and the Prime Minister's Information Advisor; and

Sir Bryan Carsberg, Locus President, and former Director General of the Ofcom and the Office of Fair Trading.

I chaired the event.

The presentations were thought-provoking and there was plenty of stimulating discussion. Here are some of the highlights:

Nigel Shadbolt spoke first, going through some of the key elements of the manifesto for government data that he and Tim Berners-Lee have published, (and which I referenced in my last blog; you can find it here: <http://www.guardian.co.uk/news/datablog/2010/jan/21/timbernerslee-government-data>), and talking about the new data.gov.uk website. Some interesting points:

- “We’re now moving from a web of documents to a web of data”.
- Most of the government PSI that has gone onto the website is in spreadsheet format.
- On release day, the data.gov.uk website had 700,000 hits; the web servers had to be replaced three times to keep up with demand. Within the first 40 minutes of the site going online, users had created three new web applications using government data. The site is currently in Beta: it’s not “complete” and it’s hoped it never will be.
- The site is all open source, done quickly and at low cost. It uses the same platform as Wikipedia.
- In creating the data.gov.uk site Professor Shadbolt and Tim Berners-Lee had to grapple with the issue of what data should be free. Their view is that if creation of the data has been paid for by the taxpayer, then the public should have it. Further, if information is available through a Freedom of Information (FoI) request, why should it not be published as an RSS feed?
- When it comes to determining what data should be made public, the question should be not why, but why not.

Next up was Francis Maude, who opened by stating that “there is a decent chance of continuity” — that the Conservatives would very likely continue the great work that’s been done to release government data already. Other points made:

- The Conservatives recognize that having open data is uncomfortable for government ministers. “Transparency is not always agreeable.”
- We need public money to be spent better. Government spends £4 for every £3 it generates in

revenue. That has to change. “We see transparency as a friend of that process.”

- We need to enlist the public to crowd-source consultancy, to be “armchair auditors” to scrutinize government spending.
- It’s understandable how we got to the current situation. OS and the Met Office are required to cover their costs; they then felt empowered to create added value and have squeezed out the private sector. This is not for the benefit of the economy as a whole. There’s in-built conflict in this arrangement.
- If we have the data, we should just put it out there. The Conservatives would release information on the salaries of the 35,000 most senior government employees and all contracts over £25,000 in value; provide an XML feed of council decisions; and publish the COINS database (this is the Combined Online Information System. It contains the Treasury’s detailed analysis of departmental spending under thousands of category headings). Some of this will be uncomfortable.
- We need to balance getting the data out there with making everything just right; we would be inclined to trade off some quality in the interests of making data available more quickly.
- We propose to carry forward the work Nigel and Tim have done.

Our final speaker was Sir Bryan Carsberg, Locus’s President, and formerly the Director General of Oftel (Ofcom’s predecessor) and the Office of Fair Trading.

Sir Bryan opened with “Carsberg’s principles of regulation”: The best form of regulation is the promotion of competition. Monopolists often have an interest in postponing uses of new technology. Liberalization has been a tremendous boon in the telecom sector; there would be enormous benefits if we can open up use of information. Other points included:

- Pricing: Locus has always been cautious on price. It’s a no brainer that data financed by tax-payers should be released free; In other cases, we believe data should be provided at marginal cost.
- Locus is mainly seeking fair competition, whether or not the information is free.
- Information should be made available in an easy way: in telecom, we had the CLASS license. That provided ready access to the base infrastructure and the constraint that providers compete fairly. All players had access to the raw materials on the same terms, at the same price, with no-cross-subsidization by the incumbent.
- You could argue that many of the Trading Funds shouldn’t provide value-added services at all.
- No matter what the outcome of the OS consultation, “This will not work without regulation.”
- We can understand the reluctance to set up a regulator: it could be costly. Right now, though, we have OPSI, but it doesn’t have the power. So we’re incurring the cost and not getting the benefit.
- It needn’t be enormously expensive to regulate this sector. If there are big penalties for breaking the rules, that can be very effective.

The panel discussion was followed by a very lively question and answer session.

I asked to panel to comment on Sir Bryan’s assertion that we must have regulation if current PSI holders are to continue to act commercially. Nigel Shadbolt’s response was that quite a lot of time had been spent on the management of data to be released, the technology, and getting it all out under a Creative Commons license; how to regulate is still to be determined.

Francis Maude said he was reluctant to set up a range of new bodies: He feels that the OFT is good at this. But, “we’re open-minded.”

One participant said it would be “terrible if the whole environment changed on the first of April.” We’ve had years of consultation, but we do need to be able to prepare for whatever changes are

brought in. We also need a set of principles that we can be sure will survive changes of government

Other topics of discussion included:

The consultation on the future of Ordnance Survey: A key point: How much Ordnance Survey data should be made available for free, and the potential negative impact on existing commercial providers if, for example, OS maps, rather than raw data, are made freely available. Nigel Shadbolt responded that he and Tim Berners-Lee couldn't avoid the issue of location and place when looking to open up government data. He felt the OS consultation didn't need to be "as complex as it is." He also finds it "odd to pay for information that you need to live your life." There was general agreement in the room that raw data, such as boundaries, should be released. There is an argument that detailed mapping should not be released for free, but that all players, including OS, should have access to the raw data to create those maps, and should compete on the same terms to provide them.

Nigel said, "Tim and I were told not to deal with OS. But we knew we had to. We will have to grasp the nettle.... We must decide what are the 'totemic datasets'".

How to find the right cost model to make data available for free: Francis Maude expressed relief that those of us who spend time of this issue also find it complex. Sir Bryan referenced the Cambridge Study, and noted that the cost of making basic data sets available is actually small. Some participants, though, disagree with the findings of the Cambridge Study, which assumes price elasticity of demand: it was noted that while there has been a 30% real price fall in MasterMap over the last few years, there has been no increase in use.

This led to a discussion of the whether or not Ramsey pricing (pricing according to the elasticity of demand of different customer groups through price discrimination) is appropriate where data is charged for. Sir Bryan noted that the OFT study concluded that the market could be greatly expanded if data was made available at marginal cost. He could see how Ramsey pricing could distort the market. Others considered it a mixed blessing for the mapping sector. Nigel Shadbolt suggested more work was required from economists.

Postcodes: There was a fair amount of discussion surrounding access to the postcode database. One participant said, "There will be trouble if we don't release postcodes soon: OS and the Royal Mail have crucial pieces of data. Have we got a plan for dealing with postcodes and addresses?" The Ordnance Survey consultation document is vague on this issue. Nigel Shadbolt suggested that we should determine what the developer community would find most useful. However, if government doesn't find a way to release postcodes, it's likely the community will crowd-source them. Other participants made the point that there is a cost of maintaining the postcode address file, and that must be covered somehow.

Our Locus event highlighted the considerable progress that is being made in PSI policy in this country, but also emphasized that there is a long way to go in the next decade to unlock the full potential of the UK PSI marketplace. In particular, government must "grasp the nettle" of freeing up access to the postcode database and Ordnance Survey raw data so all can benefit fully from the PSI that has already been released.