

Institution: University of Southampton
Unit of Assessment: 11 Computer Science and Informatics
Title of case study: 11-08 Leading the open data revolution
<p>1. Summary of the impact</p> <p>Open Data has lowered barriers to data access, increased government transparency and delivered significant economic, social and environmental benefits. Southampton research and leadership has led to the UK Public Data Principles, which were enshrined in the UK Government Open Data White Paper, and has led to <i>data.gov.uk</i>, which provides access to 10,000 government datasets. The open datasets are proving means for strong citizen engagement and are delivering economic benefit through the £10 million Open Data Institute. These in turn have placed the UK at the forefront of the global data revolution: the UK experience has informed open data initiatives in the USA, EU and G8.</p>
<p>2. Underpinning research</p> <p>Data is generated by many crucial social processes, yet the potential of vast swathes of information remains untapped. Successive UK governments have recognised that greater openness about spending can cut waste and increase value for the taxpayer, particularly in times of austerity. Research at the University of Southampton has driven the development of the open data movement, showing how transparency of data can revolutionise the delivery of public services, how business is conducted and how communities work together.</p> <p>From 2000, Nigel Shadbolt, Professor of Artificial Intelligence (2000-present) directed the EPSRC's six-year, £7.5m Advanced Knowledge Technologies (AKT project) which sought to harness the power of data by cutting through the jumble of different formatting in which information exists across the Internet and developing new ways to standardise how data is presented online. It influenced the development of the 'Semantic Web' (SW) a concept propounded in 2001 by World Wide Web inventor Sir Tim Berners-Lee, who took a part-time Chair at Southampton in 2007. It aims to transform the unwieldy, unstructured information online, into a machine-readable web of data.</p> <p>The AKT project which included Dr Kieron O'Hara, Senior Research Fellow (2000-present) pioneered techniques to show how academic researchers' Communities of Practice could be revealed, through analysing their social networks. Methods to harvest and integrate diverse information were pioneered, producing the first comprehensive application using linked data approaches [1]. This led to an influential reformulation of the basic principles of a Web of Linked Data [2] which served to strip out much of the complexity surrounding SW in order to focus on essentials, for example how to reference and link data across the Web.</p> <p>Critically, in 2004, AKTive PSI, a collaboration between the Office of Public Sector Information and the AKT project [3] led to the UK government identifying linked data as a way to publish public sector information. This pilot study showed how non-personal public data held by government could be released from poor data structures, legacy data formats and fragmented databases. It demonstrated how public sector bodies and government could adopt SW technology for the dissemination, sharing and use of its data, facilitated the large-scale integration and sharing of distributed data sources [4]. The project was recognised as a major success in a report to UK Parliament in 2007.</p> <p>The move to use Linked Data methods to publish government data, has inspired several national policies such as the free release of large amounts of Ordnance Survey Data and the adoption across government of Web-based, open and linked data principles [5] which were enshrined in the Open Data White Paper published in June 2012.</p> <p>Following on from AKT, Southampton was awarded the £0.8m EPSRC/DTI Market Blended Insight (MBI) project and the £1.94m EPSRC EnAKTing project, to carry out further research into the Web of Linked Data. Researchers developed services that supported linked data use by facilitating discovery, reuse, alignment and link enrichment. MBI applied linked data to business information to provide insights into markets. EnAKTing transformed data sets published on www.data.gov.uk into linked data and created simple, useful visualisations that everyone could explore [6].</p>

Impact case study (REF3b)

Most recently Shadbolt and colleagues have secured a £6.15m EPSRC Programme Grant for SOCIAM - the theory and practice of Social Machines to take work on linked data further.

3. References to the research (best three references starred)

1. Shadbolt, N., Gibbins, N., Glaser, H., Harris, S. and schraefel, m. c. (2004) CS AKTive Space or how we stopped worrying and learned to love the Semantic Web. *IEEE Intelligent Systems*, 19 (3), pp. 41-47.
2. * Shadbolt, N., Berners-Lee, T. and Hall, W. (2006) The Semantic Web Revisited. *IEEE Intelligent Systems* 21(3), pp. 96-101.
3. Alani, H., Dupplaw, D., Sheridan, J., O'Hara, K., Darlington, J., Shadbolt, N. and Tullo, C. (2007) Unlocking the Potential of Public Sector Information with Semantic Web Technology. In, *The 6th International Semantic Web Conference (ISWC), Busan, Korea*,
4. * Alani, H., Hall, W., O'Hara, K., Shadbolt, N., Chandler, P. and Szomszor, M. (2008) Building a pragmatic Semantic Web. *IEEE Intelligent Systems*, 23, (3), 61-68.
5. Shadbolt, N., O'Hara, K., Salvadores, M. and Alani, H. (2011) eGovernment. In: *John Domingue, Dieter Fensel & James Hendler (eds.), Handbook of Semantic Web Technologies*, pp. 840-900, Springer-Verlag.
6. * Shadbolt, N., O'Hara, K., Berners-Lee, T., Gibbins, N., Glaser, H., Hall, W. and schraefel, mc (2012) Open Government Data and the Linked Data Web: Lessons from data.gov.uk. *IEEE Intelligent Systems*.

Grants supporting underpinning research

1. PI Shadbolt EPSRC Funded Advanced Knowledge Technologies (AKT) IRC £7.5m (2000-07 GR/N15764/01) AKT – rated outstanding scoring 35 out 36 at final review
2. PI Shadbolt EPSRC Funding large Grant EnAKTinG the unbounded Web of Data £1.94m (2009-2012 EP/G008493/1)
3. PI Shadbolt EPSRC/DTI Market Blended Insight £0.8m (2006-2010 DT/E007104/1)
4. PI Shadbolt EPSRC Programme Grant for SOCIAM - the theory and practice of Social Machines £6.15m (2012-2017 EP/J017728/1)

4. Details of the impact

Research at Southampton has revolutionised access to data across all sectors of UK society, in turn influencing open data movements in the United States, Europe and globally through the G8. 'Open Data' is a flagship policy of the current government, with the UK ranked #1 in the Open Knowledge Foundation's 2013 Open Data Index (index.okfn.org/country).

Policy impact

Following the success of AKTive PSI, Shadbolt and Berners-Lee were appointed by then-Prime Minister Gordon Brown in June 2009, to apply their research to making government information publicly accessible. Their brief to full Cabinet in September 2009 led to significant policy change: 40% of Ordnance Survey's geographical data were released for use by government, businesses and individuals [5.1].

In collaboration with web developers, politicians and civil servants, Shadbolt and Berners-Lee oversaw the design and implementation of *data.gov.uk* [5.2] which launched in January 2010. It became the main online access point to thousands of government datasets relating, among others, to crime, health and education. The same month, Shadbolt was appointed by the Secretary of State for Communities and Local Government to chair the Local Public Data Panel to provide independent advice on the release and use of local government public data. It oversaw some of the most significant data releases, for example, full details of local government spending.

In May 2010, Shadbolt and Berners-Lee were appointed to the coalition government's Public Sector Transparency Board, which is responsible for setting open data standards across the public sector [5.3]. They drew on their research to directly inform open data policy, drawing-up the UK Public Data Principles to guide the release of government data [5.4]. These were enshrined in the Open Data White Paper (June 2012). Prime Minister Cameron referred in a number of speeches to

Impact case study (REF3b)

the transformative opportunity of Linked Data [5.5]. In June 2013, Shadbolt was knighted for services to science and engineering – in part the citation read “He has contributed extensively to the development and practical implementation of government policy.”

Further afield, the EU’s eGovernment Action Plan (2011-2015) revolves around open data. Vice President Neelie Kroes and her Cabinet consulted Shadbolt on multiple occasions about this agenda. Shadbolt was then commissioned to write the template and technical architecture for the EU’s own open data portal, launched in December 2012 [5.6]. He also briefed the US Chief Technology Office’s Vivek Kundra on the UK Open Data Programme. Most recently, Shadbolt and one of his researchers, Tim Davies, provided input to the UK’s G8 presidency work, which resulted in the G8 signing an historic Open Data Charter in June 2013, explicitly recognising “a new era in which people can use open data to generate insights, ideas, and services to create a better world for all.”

In an increasingly open data environment, privacy is a key concern. On the back of the underpinning research, Southampton’s Kieron O’Hara was asked by the Cabinet Office to produce the 2011 report *Transparent government, not transparent citizens* [5.7], in which he stressed privacy protection should be embedded in any transparency programme in order to retain public confidence. Many respondents to the government’s Open Data Consultation endorsed the report; most of his recommendations were taken up in the 2012 White Paper and his recommendation on collating best practice on anonymising data, led directly to the Information Commissioner’s Office creating and funding the UK Anonymisation Network. O’Hara chairs the Transparency Sector Panel for Crime and Criminal Justice, an expert panel that advises the Home Office and Ministry of Justice (MoJ) about open data releases. In direct response to requests made through the sector panel, government opened up court data, particularly advance listings data. The experience has been so positive that various other departments have also set up their own panels, including International Development, DEFRA and Department for Transport. The interaction between stakeholders and officials has led to greater awareness of risks and demand for data. MoJ officials consulted academics from Southampton, Royal Holloway and LSE about an anonymised release of data which demonstrated that the initial dataset was not anonymised sufficiently. O’Hara also advised on privacy issues relating to the Home Office’s police.uk crime mapping site, a well-used flagship government transparency project, which has so far received no serious privacy complaints.

Economic impact

Shadbolt and Berners-Lee’s Open Government Data programme has survived a change of government, and is part of the current UK government’s growth strategy. Minister for the Cabinet Office, Francis Maude, has said that by freeing up public sector datasets for others to reuse could contribute up to £6 billion to the UK’s economy. In his Autumn Statement in 2011, Chancellor George Osborne announced £10 million in funding to establish the Open Data Institute (ODI). Shadbolt and Berners-Lee were the ODI’s Co-Founders and are Chairman and President respectively. Launched in October 2012 it is incubating 10 start ups, one of which (Mastadon C) identified annual savings of £200 million for the NHS, by analysing drug prescription practice. Another is selling access to a whole range of transport data, which Transport for London estimates has yielded a return worth 56 times its investment. The ODI has had 3,000 visitors, received delegations from around 20 countries, provided training and raised an additional £4 million in funding. A 2012 report [5.8] by Deloitte for the Department for Business, Innovation and Skills estimated the value of public sector information to consumers, businesses and the public sector at approximately £1.8 billion, with “much larger downstream impacts affecting all areas of society beyond the direct customer.

Societal impact

Data.gov.uk now contains around 10,000 datasets that can be used to create new software applications, capable of determining anything from the nature of local authorities’ spending, to the whereabouts of the UK’s bus stops. Applications built using the data include services like *Fix My Street*, which helps those wanting to report local problems like fly-tipping or broken street lighting. The website *Who’s Lobbying* tracks who is meeting who in the higher echelons of government. *Spotlight on Spend* helps councils make information on its public spending more accessible, and thus more transparent. In effect, data.gov.uk has enabled the public – and investigative journalists

Impact case study (REF3b)

– to collectively mine data for the common good. Shadbolt has led efforts to promote the value and potential of open data to the wider public through regular media exposure (both op-eds and interviews) in national newspapers [5.9].

The research also inspired the creation of data.southampton.ac.uk, a prize-winning open data site for the University that is saving money, driving innovation and being copied by other HE institutions [5.10].

5. Sources to corroborate the impact

1. Article from No 10 Website on Berners_Lee & Shadbolt' briefing to Cabinet <http://users.ecs.soton.ac.uk/nrs/2009/09/pm-welcomes-berners-lee-to-downing-street/> and Ordnance Survey Open Data Release <http://data.ordnancesurvey.co.uk>
2. Launch of data.gov.uk <http://www.guardian.co.uk/technology/datablog/2010/jan/21/government-free-data-website-launch?intcmp=239>
3. Article from Cabinet Office on Appointment of Shadbolt and Berner's-Lee to the Coalition Government's Transparency Board http://webarchive.nationalarchives.gov.uk/+http://www.cabinetoffice.gov.uk/newsroom/news_releases/2010/100531-open.aspx
4. Published Letter to Francis Maude, MCO at start of Coalition Governments Transparency Board http://data.gov.uk/sites/default/files/Transparency%20Board%20-%20letter%20from%20Nigel%20Shadbolt%20to%20MCO%2014.06.10.pdp_0.pdf and the resultant UK public data principles <http://data.gov.uk/blog/public-data-statement-of-principles>
5. Prime Minister's speeches covering open data policy: <https://www.gov.uk/government/speeches/pm-speech-at-open-government-partnership-2013>
<https://www.gov.uk/government/news/g8-summit-prime-ministers-remarks>
<https://www.gov.uk/government/speeches/business-in-the-community-speech--2>
6. Shadbolt's report on Open Data Portals to the EU http://ec.europa.eu/information_society/policy/psi/open_data_portal/index_en.htm
7. O'Hara's report to the Cabinet Office on Open Data and Privacy <https://update.cabinetoffice.gov.uk/news/government-publishes-independent-transparency-and-privacy-review>
8. Deloitte report on economic impact following the Shakespeare Review <https://www.gov.uk/government/publications/public-sector-information-market-assessment>
9. Media coverage:

Columns by Berners-Lee and Shadbolt for The Times, London *Put in your postcode, out comes the data*, 18th November, 2009 <http://eprints.ecs.soton.ac.uk/23212/>

The Guardian, London *Our Manifesto for Government Data*, 21st January, 2010 <http://www.guardian.co.uk/news/datablog/2010/jan/21/timbernerslee-government-data>

The Times, London *There's gold to be mined from all our data*, 31st December, 2011 <http://eprints.ecs.soton.ac.uk/23090/>

Editorial in the Guardian Saturday 23 January 2010 Government Information: Creative Commons <http://www.guardian.co.uk/commentisfree/2010/jan/23/government-information-creative-commons-internet?INTCMP=SRCH>

G8 Open Data Charter by Shadbolt in the Telegraph <http://www.telegraph.co.uk/news/worldnews/g8/10128266/G8-Open-Data-Charter-why-it-matters.html>
10. THE outstanding IT Initiative of the Year <http://www.ecs.soton.ac.uk/news/4121>