

Driving Innovation Using Big Open Linked Data (BOLD) Panel

Yogesh K. Dwivedi^{1(✉)}, Vishanth Weerakkody², Marijn Janssen³, Jeremy Millard^{2,4},
Jan Hidders⁵, Dhoya Snijders⁶, Nripendra P. Rana¹, and Emma L. Slade¹

¹ School of Management, Swansea University, Swansea, UK
{y.k.dwivedi, N.P.Rana, e.l.slade}@swansea.ac.uk

² Business School, Brunel University London, Uxbridge, UK
Vishanth.Weerakkody@brunel.ac.uk, jeremy.millard@3mg.org,
Jeremy.Millard@brunel.ac.uk

³ Section of Information & Communication Technology, Faculty of Technology,
Policy, and Management, Delft University of Technology, Delft, The Netherlands
M.F.W.H.A.Janssen@tudelft.nl

⁴ Danish Technological Institute, Aarhus, Denmark
⁵ Web Information Systems Group, The Software and Computer Technology Department,
Delft University of Technology, Delft, The Netherlands
A.J.H.Hidders@tudelft.nl

⁶ Stichting Toekomstbeeld der Techniek (STT),
The Dutch Study Center for Technology Trends, Den Haag, The Netherlands
snijders@stt.nl

Abstract. Governments have always retained public service data internally in their own systems with only limited information provided to the public and other stakeholders such as the business, charitable and NGO communities. However, the rapid advancement of ICTs coupled with electronic publishing via the Internet in the last decade in particular has enabled governments to exploit the potential of wider distribution and use of such data previously held in internal systems. The panellists will discuss how Big, Open and Linked Data (BOLD) can be utilized to drive innovation and what obstacles and challenges may be encountered. Empowering citizens, potential mis-use in identity theft, policy manipulation or market distortion, and the need to combine open data with closed sources will be discussed.

Keywords: Big data · Open data · BOLD · Panel

1 Introduction

Traditionally, governments have retained public service data internally in their own systems and only a limited amount of information was provided to the public and other stakeholders such as the business, charitable and NGO communities. However, the rapid advancement of ICTs coupled with electronic publishing via the Internet in the last

decade in particular has enabled governments to exploit the potential of wider distribution and use of such data previously held in internal systems. Recognising these developments, a European Union directive to encourage greater realisation of the economic value of public data through its reuse entered into force in 2003, thus, paving the way for governments to open up to the public previously closed data. This sought to both encourage provision and regulate licensing and charging for information by Member States. In 2004, the OECD recommended that all publicly funded research data should be made available.

The dual motivations of transparency of government and the economic potential of the reuse of data have since motivated politicians to promote the adoption of “open data” as a practice for their governments. In the US, the Presidential Open Government Directive in December 2009 required the use of open formats by federal US agencies, and in May 2010 the UK Prime Minister set out plans for opening up government data through the data.gov.uk website. Subsequently, the European Commission published a Communication on Open Data in 2011, and in the same year the USA, UK and initially six other countries were signatories to the Open Government Declaration.

Although selected public sector data has always been available to public bodies and other organisations, albeit with costs and restrictions, progressive moves by governments in countries such as the UK and US to improve availability and ease of reuse (through machine-readability and technical standards) has removed many such barriers. One of the unforeseen effects of the open data movement has been to make more data easily accessible to other actors in the policy space, including researchers, think-tanks and, most significantly, other parts of the public sector and governmental systems, including local governments. However, the availability of open data relating to various public services has led to more transparency around these services and allowed the general public to hold government departments accountable. This has encouraged stakeholders (including citizens, businesses, charities and NGOs) to take an active interest in the way services are currently delivered and stimulated thinking around how to improve services. Yet the actual use of open data is cumbersome and each stakeholder has to do it themselves as there are no proven solutions facilitating the process of using open data for services transformation or co-creation. Moreover, although the open data movement has gathered much momentum in Europe, the majority of open data repositories (e.g. data.gov.uk, etc.) are too generic and not of direct interest to individual citizens.

Although big or large volumes of raw open data when published in an electronic format is machine readable and can be shared online and re-used, on its own open data offers limited potential for decision making. However, when open data is linked (Big, Open and Linked Data – BOLD) with extra information that provides more context, this offers greater opportunities for stakeholders to exploit the data for innovative purposes, for example through collaboration and co-creation. BOLD could also increase the reach of statistical and operational information, and deeper analysis of outcomes and impacts. Indeed, BOLD offers an opportunity to discover new ways of assessing policy and service outcomes, healthcare and wellbeing, measuring development and making appropriate interventions through innovative solutions where needed.

Although the use of BOLD to date remains limited and at an early stage there are many examples that show its potential. When, in 2011, the UK’s mapping authority

released all its geodata rather than selling it, the organization completely transformed its business model from being simply a purveyor of raw data for others to use to becoming its customers' partner in collecting, analysing, tailoring, visualizing and applying linked data in meeting their needs. This new value adding and problem solving role boosted and diversified the authority's income. In another example, Washington DC was one of the first cities to make BOLD available, running hackathons and competitions as part of its 'Apps for Democracy' initiative, which eventually saved the City authorities \$2.3 million at a cost of \$50,000.

In responding to the theme of The 14th IFIP Conference on e-Business, e-Services and e-Society (I3E 2015), the panellists will discuss how Big, Open and Linked Data (BOLD) can be utilised to drive innovation and what obstacles and challenges may be encountered regarding this. The discussion will be based on empirical evidence, using insights from cases; there will be a focus on analytical models, building on conceptual and theoretical arguments, and our aim is to offer directions for future research.

2 Programme

Professor Yogesh Dwivedi will open the panel by presenting an overview of the topic and subsequently will moderate the panel discussion.

Professor Vishanth Weerakkody will evaluate the role of stakeholders in the BOLD arena and discuss some of the challenges facing public agencies and potential users in exploiting public open data. The discussion will particularly centre on the Hype Vs Reality in the BOLD debate and use several European Commission funded R&D projects to discuss the present technical and user problems faced by local governments who are engaged in promoting BOLD for improved decision making. The discussion will also examine how universities, professional institutes and organisations are preparing their graduates and professionals to deal with the BOLD challenge. In the public service space, who are the main users of big and open data? Are citizens interested or prepared to exploit open data? Are conventional statistical tools equipped to deal with BOLD? Which tools are emerging as the front runners for exploiting BOLD? These are some of the other questions that will be debated.

Professor Marijn Janssen will argue that BOLD can be used to empower citizens and to create an ecosystem in which governments, businesses and citizens can strengthen each other. Yet driving innovations from data is a complex process in which both the available data and the users' demands need to be taken into account. The name "data-driven" already implies that data instead of customers' needs are leading. Furthermore all stakeholders have different interests and concerns. Policy-makers want to use the results for their policy-making, businesses want to create new services to enrich their existing products and services, whereas citizens want to know what is going on and have free access to services without having to up their privacy. For this purpose platforms creating a community, providing an overview of the data and collaborative development environment to develop new ideas, are needed. However, the opening and use of data

might make apparent the low data quality of sources and the focus might be on developing tools and platforms without having the user in mind.

Jeremy Millard's discussion will focus on two main but related issues. First, one of the wicked challenges of BOLD is its potential mis-use, either through negligence, corruption or criminality resulting, for example, in identity theft, policy manipulation or market distortion. Although this is a battle that can never be won, what are the technical, legal, regulatory and ethical challenges we face and how might they be addressed, also taking account of both the Snowden and Assange cases? Second, BOLD becomes even more interesting and disruptive when public data is mixed with data from other sources, including, for example, the private data of businesses or citizens. The smart disclosure approach starts from the premise that people, when given access to data and useful decision tools built for example by governments, can use both their own personal data disclosed by them together with other appropriate data. This can be used to make decisions about their own lives, such as healthcare choices, as well as to self-regulate and be able to hold governments and other actors to account, as well as to cooperate and engage with them.

Jan Hidders will discuss the technological data management challenges of BOLD. In order for BOLD to become effective and usable there must be tools that allow ordinary users, people interested in using the data, to select, transform, combine and visualise the data in an ad-hoc fashion. In some sense this has always been the goal of data management tools, but with the advent of BOLD new challenges have been added such as the scale of the data (in number of data sources, sizes of schemas as well as sizes of actual data content), the possible lack or presence of structure and semantics, and the large heterogeneity of the data sources. This puts all kinds of new demands on the existing data management tools that are currently not being met. These demands are not only in terms of scalability and ability to deal with the semantics of the data, but also in terms of interfaces and languages that allow users to deal with such data in a more intuitive way than now is possible. A key question is what underlying data model these tools should be based upon? Should it be XML, JSON or perhaps RDF which comes with semantic abilities? Or should we perhaps go back to classical models such as the relational model? These and other issues will be discussed in this section.

Dhoya Snijders will focus on the ways in which BOLD is altering how citizens and government relate to each other. Already, BOLD pleads for new types of civil servant with different skills, different educations and, some argue, a different epistemological standpoint. The speed with which data can be created and with which correlations within the data can be made is impacting government's deep-rooted quest for causality. Similarly, BOLD is creating new types of citizen, who are data-focused and data-driven. Citizens are increasingly connecting, measuring, analysing and testing public data themselves. The relations between government and citizens are hereby mediated by BOLD. And, as BOLD is itself becoming intelligent, data is not only opened up, linked, and analysed by human actors - machine-learning is picking up speed and quality. Both citizens and governments will increasingly have to deal with non-human actors in the form of intelligent data-driven systems. To do so we need to develop what sociologists

have dubbed double contingency in which we as humans need to understand how intelligent machines will understand us and vice versa.

Nripendra Rana and Emma Slade will assist in organisation and coordination of the panel and note taking of the panel content.

3 Panellist Bios

Yogesh K Dwivedi is a Professor of Digital and Social Media and Head of Management and Systems Section (MaSS) in the School of Management at Swansea University, Wales, UK. His research interests are in the area of Information Systems (IS) including the adoption and diffusion of emerging ICTs (e.g. broadband, RFID, e-commerce, e-government, m-commerce, m-payments, m-government) and digital and social media marketing. His work on these topics has been published (more than 100 articles) in a range of leading academic journals including: CACM, EJIS, IJPR, GIQ, JORS, ISF, ISJ and IJICBM. He has co-edited more than ten books on technology adoption, e-government and IS theory and had them published by international publishers such as Springer, Routledge, and Emerald. He acted as co-editor of fourteen special issues; has organised tracks, mini-tracks and panels in leading conferences; and served as programme co-chair of IFIP WG 8.6 Conference at the prestigious IIM Bangalore, India in 2013. He is Associate Editor of GIQ, EJM and EJIS, Assistant Editor of JEIM and TGPPP, Senior Editor of Journal of Electronic Commerce Research and member of the editorial board/review board of several journals. He is a life member of the IFIP WG8.6 and 8.5.

Vishanth Weerakkody is Professor of Digital Governance in the Business School of Brunel University in London, United Kingdom. His research experience is focused in the area of public sector service transformation through technology. Professor Weerakkody has published more than 100 peer reviewed articles and has guest edited several special editions of journals and books on this theme. He is co-founder of the e-Government track at the Americas Conference on Information Systems (AMCIS) and Transforming Government Workshop at Brunel University, London. Vishanth has many years of experience in R & D projects in the area of digital governance and is currently an investigator in several European Commission and Internationally funded projects on the use of ICTs in the public sector (e.g. LiveCity, OASIS, DAREED, PolicyCompass, UBIPol, CEES, EGovPoliNet, I-MEET, SI-DRIVE). He is the Editor-in-Chief of the International Journal of Electronic Government Research and one of the two international board members of the Digital Government Society of North America. He is a fellow of the UK Higher Education Academy and combines more than 25 years of experience in industry, teaching and research leadership.

Prof. Dr. Marijn Janssen is full Professor in ICT & Governance and head of the Information and Communication Technology section of the Technology, Policy and Management Faculty of Delft University of Technology. His research interests are in the field of orchestration, (shared) services, intermediaries, open and big data and infrastructures within constellations of public and private organizations. He was involved in EU funded

projects in the past (a.o. EGovRTD2020, eGovPoliNet and Engage), is Co-Editor-in-Chief of Government Information Quarterly, Associate Editor of the International Journal of Electronic Business Research (IJEER), Electronic Journal of eGovernment (EJEG), International Journal of E-Government Research (IJEGR), is conference chair of IFIP EGOV2015 and IFIP I3E2015 conference (about big and open data innovation) and is chairing mini-tracks at the DG.o, ICEGOV, HICCS and AMCIS conferences. He was ranked as one of the leading e-government researchers in a survey in 2009 and 2014 and has published over 320 refereed publications. More information: www.tbm.tudelft.nl/marijnj.

Jeremy Millard is a Senior Research Fellow at Brunel University (London) and Senior Policy Advisor, Danish Technological Institute (Denmark). He has forty years' global experience working with governments, development agencies, and private and civil sectors in all parts of the world. In the last twenty years he has focused on how new technical and organisational innovations transform government and the public sector. Work with the European Commission includes research and studies on eGovernment, administrative burden reduction, and on developing business models for ICT and ageing. He also recently led an impact assessment of the European eGovernment Action Plan, led the large-scale Europe-wide survey and analysis of eGovernment eParticipation, and developed the 2020 Vision Study on Future Directions of Public Service Delivery. Jeremy has also worked since 2008 as an expert for the UN on their successive global eGovernment development surveys, and provided inputs to both the World Bank and the OECD on eGovernment developments, for example through a survey on back-office developments in support of user-centred eGovernment strategies, as well as ICT-enabled public sector innovation. He works extensively outside Europe in these and related areas, including in the Gulf, the Western Balkans, Georgia, Japan, and India.

Dr Jan Hidders is Assistant Professor in the Web Information Systems group of the Software and Computer Technology department of Delft University of Technology. His research interests are in data integration, data indexing, data linking and large scale data processing, particularly in the domain of graph processing. He was involved in several EU projects such as GRAPPLE and ImREAL where he worked on data integration in the e-learning domain. He has published over 60 refereed publications in conferences such as ICDE, CCGRID, ICDT, CIKM, IUI, FoIKS, ICWE and ISWC, and in journals such as Information Systems, Fundamenta Informaticae, Theory of Computing Systems, BPM Journal and Journal of Computer and System Sciences. He has edited special issues in Fundamenta Informaticae and the Journal of Computer and System Sciences. He has co-organised several workshops such as SWEET (Scalable Workflow Enactment Engines and Technology) at the SIGMOD conference and BeyondMR (Beyond MapReduce) at the EDBT conference. He is currently one of the co-organizers of the 2015 EDBT Summer School on Graph Data Management.

Dr Dhoya Snijders works for STT, the Dutch Study Center for Technology Trends and is currently carrying out a technology foresight study on big data. He holds an MA in Philosophy from the University of Amsterdam and an MA and PhD in Organizational Sciences and Public Administration from the VU University. His research mainly

focuses on classifications and their consequences. He worked as a consultant for some years in the field of e-Government within the Dutch Ministries of Healthcare and Justice. There he carried out national and international projects on ICT governance, eHealth, Open Data, and Big Data. During this time he published on the implementation of ICT in a context of multilevel governance.

Dr Nripendra P Rana is a Lecturer at the School of Management at Swansea University in the UK. He holds a BSc in Mathematics (Hons), an MCA, an MTech, and an MPhil degree from Indian universities. He also holds an MBA with distinction and a PhD from Swansea University. His current research interest is in the area of technology and e-Government adoption and diffusion. He has published his work in some refereed journals including ISF, ISM, ESJ, IJBIS, IJICBM, IJEGR, and TGPPPP. He has varied work experience of teaching in the area of computer science at undergraduate and postgraduate levels. He also possesses a good experience of software development and leading successful software projects.

Emma Slade is a Research Officer in the School of Management at Swansea University. She holds a BSc (Hons), MSc with distinction, and PhD in Business Management from Swansea University. Emma's research interests include consumer and merchant mobile payment adoption, e-government adoption, and consumer forgiveness. Her research has been published in *Psychology & Marketing*, *Journal of Strategic Marketing*, and *Journal of Computer Information Systems*. She also presented at the 2014 Academy of Marketing Science Conference and the 2013 UK Academy for Information Systems Conference. Emma has been invited to review papers for *Information Systems Frontiers*, *Internet Research*, and a number of conferences.