Open Data: Growing Up and Getting Specific

Johann Höchtl*, Tim Davies**, Marijn Janssen***, Ina Schieferdecker****

Danube University Krems, Austria; johann.hoechtl@donau-uni.ac.at

***University of Southampton, Web Science Doctoral Training Center; tim.davies@soton.ac.uk*

****Delft University of Technology, the Netherlands; m.f.w.h.a.janssen@tudelft.nll*

*****Fraunhofer FOKUS, Berlin, Germany; ina.schieferdecker@fokus.fraunhofer.de*

Abstract: Open data has been given a lot of attention in the public. In some situation 'open by default' has become established as a core principle, whereas others argue about the limited results and the lack of robust studies demonstrating the value, and point to the risk that open data might turn out to be a short lived policy fad. This special issue contains a variety of research papers addressing this topic from different views and providing recent research results on open data. The papers in this issue deepen the understanding of open data and show that the subject of open data is moving from the general to the study of specifics. The special issue also includes invited papers presented at the first public meeting of the SharePSI project. Share-PSI 2.0 is the European network for the exchange of experience and ideas around implementing open data policies in the public sector.

Keywords: Open Data, Engagement, Participation, Geo-Spatial, ICT4D.

The topic of open data has captured the attention of academics in a wide range of disciplines, sparking many papers, conference tracks, workshops, panels and journal issues. Nowadays governments have developed open government policies and have implemented open data initiatives at the global and local levels. These developments have resulted in problem-driven research in which open data has a cross-disciplinary appeal as a subject, bridging technical, political and social research domains. Whilst some of the earliest papers looked primarily at the case for changing regimes of data management, academic attention quickly turned to the practical pros and cons of open data, examining drivers and barriers, and surveying the top-down and bottom up movements which have pushed for open data. This discourse was soon joined by papers focussing on architectural considerations, addressing questions aroundow to release data, and how to structure comprehensive policies, as well as increasing work on the measurement of open data efforts. Yet, whilst some might argue that "open by default" has become established as a core principle, and that the continued rise of open data is assured, others point to the lack of robust studies demonstrating the impacts of opening data, and the risk that open data might turn out to be a short lived policy fad. We arrived at a stage where ongoing open data efforts either have to be critically revisited or disappeared, or have reached the status of de-facto established processes.

While the normative case for open by default is seldom questioned, now the issue at hand is whether there are pervasive applications and effects of open data that can justify continued investment, and the institutionalisation of open data practices. Scholarship on this has to deal with more in-depth, domain-specific issues, instead of looking at generalist open data quick wins. How can open data contribute to sustainable growth? What effects does open data have on governance structures? Does the opening of data result in transparency? What kind of data is likely to spawn the highest positive effects in developed countries vs. industrialised countries and how are those effects measured? The papers of this JeDEM special issue address these and similar research questions.

In her paper "Investigating the roots of Open Data's Social Impact" Amanda Meng focusses on the problematics of open data in a global context: examining the assets that countries may require in order to see social change result from data release. Meng first starts by describing the nature of social impact and relating various definitions of social impact to the aims of public administrations in releasing data sets. She identifies the hard challenge of assessing the impact of open data, particularly when moving beyond economic measures. One major goal of the paper is to shed light on the fact that data from the Web Foundation's Open Data Barometer (ODB), which surveyed 76 countries open data policies and practices in 2013, appears to contain a puzzling inconsistency with countries scoring relatively high in social impact not neccessarily those that high on measures of open data readiness and implementation. Meng suggests this points to unexamined preconditions that may support social impact. One interesting and relatively novel approach then taken in the paper is to discuss the distinction between social capital and political capital, with political capital expressing attitudes supportive of democratic norms and behavior that engages citizens with the state and each other in channeled ways, suggesting a re-inforcing cycle between social and political capital. To prove the hypothesis that civil society, through attributes of social capital and political capital, makes democracy more robust by increasing political participation in context of open data, the author correlates quantitative data on eight latin-american countries, applying fuzzy logic to look for relationships. The outcome of this tabulation suggests that a practice or norm of engaging with the state in channeled ways is an important factor in a country's ability to secure social impacts from open data. Technological capacity and artifacts are necessary, but do not appear to provide sufficient preconditions for change in transparency and accountability, which, again, suggests, that open data advocates should look towards civic space, public officials and political parties as important stakeholders and agents of change in the open data movement.

Fechner and Krays in their paper on georeferenced open data" address the combination of open data with increasingly omnipresent mapping devices (smart phones and routing devices) and social media in order to address persistent challenges of effective and accessible citizen engagement.. They argue that interactive geo-visualizations that "integrate open (government) data [...], so that citizens can transport their wishes, comments or critiques on a medium that enables exploration and fosters reasoning in a visual and spatial way" have the potential to empower citizens to more effectively engage with societal and governmental processes. Yet, to realise this opportunity they note that geo-coded data, and core metadata fields are required. Fechner and Krays go beyond the simple identification of data quality barriers, to explore a quantification of the scale of the challenge. Using automated classification approaches they assess the quality of geodata provided from a number of statistical data providers, finding that "as the spatial granularity increases the quality of the metadata decreases". Their work also offers an addendum to the common call for greater 'data literacy', noting the need also to develop 'map literacy' if citizens are to be empowered by the engagement opportunities presented by the spatial and temporal dimensions of public data.

Bartenberger and Grubmüller also look at the participatory potential of open data, though this time in the context of smart cities. According to their understanding, collaborative governance is one integral part of a smart city and their paper primarily targets the question of mutual effects between open government data and collaborative governance processes. They further develop a collaborative governance model found in established literature, building on concepts of trust, knowledge and information asymmetries identified by Ansell and Gash, which the authors enrich

with the aspect of open data to overcome these asymmetries. For example, the fact that, in open data regimes, data can be released in advance can give stakeholders more time to explore relevant material, thus facilitating joint fact finding.The authors then take the case of three smart cities, namely Chicago, Barcelona and Hong Kong and try to prove the importance of open data to in the role of overcoming asymmetry, raising trust and supporting joint fact finding.

Finally, Hartog, Mulder, Spée, Visser, Gribnau focus in on two open data initiatives in the Netherlands, each of which originate within the context of geographic data, but which have experienced different trajectories, even though the two government entities studied share an overlapping geography. They note the business benefits to authorities, in terms of harmonising data policies and creating efficiency, highlighting that the agendas driving open data are often not centred on civic engagement, but may be part of ongoing modernisation and reform. The increase in comparative open data studies at a range of levels can greatly add to understanding in the field, helping to suggest the contextual variables which affect the implementation and impacts of open data initiatives, and moving analysis beyond generic lists of pros and cons. As Hartog et al. conclude there is still a need for "clarity on the actual ability of open data and, why you should act on it and what you are able to achieve with it" and this cannot be found for open data in the abstract: but must be answered with respect to particular datasets in particular settings.

We are also happy to include in this open data special issue of JeDEM papers presented at the first public meeting of the Share-PSI project. Share-PSI 2.0 is the European network for the exchange of experience and ideas around implementing open data policies in the public sector. It brings together government departments, standards bodies, academic institutions, commercial organisations, trade associations and interest groups to identify what does and doesn't work, what is and isn't practical, what can and can't be expected of different stakeholders. The first meeting took place co-located with Samos summer school and the best papers of this first meeting have been invited for publication.

In summary, the papers in this issue resonate with the deepening of open data debates: moving from the general, to the study of specifics. We hope they contribute to a continued development of interdisciplinary scholarship in this field.

About the Guest Editors

Johann Höchtl graduated from University of Vienna and Vienna University of Technology in Business Informatics. He is research fellow at Danube University Krems, Center for E-Governance, Austria His projects include EU-funded research projects and national grants in the domain of social media application in administration, and security. He is former member of OASIS SET TC standardisation group. Johann Höchtl is member of OKFN Austria and member of Cooperation Open Government Data Austria, where he is heading data quality sub working group. Apart from his professional work he was advisor to the E-Georgia strategy for the public administration.

Tim Davies is a PhD Candidate in the Web Science Doctoral Training Center at the University of Southampton, and an affiliate of the Harvard Berkman Centre for Internet and Society. He has worked with the World Wide Web Foundation as research lead for the Open Data in Developing Countries program, as well as supporting numerous other projects at the intersection of technology and governance.

Marijn Janssen is Antoni van Leeuwenhoek-Professor in "ICT and Governance" and head of the ICT section of the Technology, Policy and Management Faculty of Delft University of Technology. He serves on several editorial boards and is involved in the organization of a number of conferences including IFIP EGOV and ICEGOV. He published over 300 refereed publications. More information: www.tbm.tudelft.nl/marijnj

Ina Schieferdecker is the Open Data Coordinator at Fraunhofer FOKUS and chairs Model-Driven Engineering and Quality Assurance of Software-Based Systems at Freie Universität Berlin. She is member of the German National Academy of Science and Engineering. Besides other things, she leads the first German open data portal in Berlin and coordinated technically the German data portal GovData.de.