Closing the open public data feedback loop: the ENGAGE platform

Charalampos Alexopoulos[1], Anneke Zuiderwijk[2], Yannis Charalabidis[1], Euripides Loukis[1],
[1] University of AEGEAN, Greece
[2] Delft University of Technology, Delft, The Netherlands
{alexop,yannisx,eloukis}@aegean.gr
a.m.g.zuiderwijk-vaneijk@tudelft.nl

Abstract: One essential element of open data ecosystems concerns their development through feedback loops, discussions and dynamic supplier and user interactions. However, these elements appear barely to be part of existing open data practices. We conducted a survey which showed that most professional open data users did not know at least one open data infrastructure that enabled five specific types of discussion and feedback mechanisms. The survey showed that much can still be done to improve feedback and discussion on open data infrastructures. In this paper we discuss an open data infrastructure which has started to contribute to filling this gap. The discussed ENGAGE open data infrastructure combines functionalities to close the feedback loop and to return information to public authorities for better open data use and publication as well as establishing communication channels between stakeholders. This may effectively lead to the stimulation and facilitation of value generation from open data.

Keywords: open government data, public sector information, e-infrastructures Web 2.0, social media

This position paper is partly based on the following articles:

C. Alexopoulos, A. Zuiderwijk, E. Loukis and M. Janssen, "Designing a second generation of open data platforms: integrating open data and social media", Conference on E-Government, Forthcoming.

A. Zuiderwijk, M. Janssen, and C. Davis, "Innovation with open data: Essential elements of open data ecosystems", Information Polity, Forthcoming.

1) Introduction

Government create and collect large amounts of data in various domains, such as statistical, business, tourist, health, pollution, traffic, unemployment, crime and poverty data. If these data are released to the public they can be used for many other purposes, which may be different from the ones of their initial creation, including scientific, commercial and political purposes. Open data have been hailed for their potential to generate public value, particularly through innovation, economic growth, and transparency (for instance, Blakemore & Craglia, 2006; Charalabidis, Ntanos, & Lampathaki, 2011; European_Commission, 2003, 2011; Zhang, Dawes, & Sarkis, 2005). Moreover, another major trend in government agencies has been the exploitation of Web 2.0 social media for increasing citizens' participation in the governments decision and policy making processes, supporting networking, interaction and collaboration, collecting opinions, knowledge and ideas from citizens, and promoting government transparency and accountability (Bertot, 2012; Bonsón, 2012; Chun, 2012; Margo, 2012; Criado, 2013).

To be able to benefit from open data, several researchers have argued that open data should be seen as an on-going process or as an ecosystem rather than a product (Pollock, 2011; Zuiderwijk & Janssen, 2013; Zuiderwijk, Janssen, Choenni, Meijer, & Sheikh_Alibaks, 2012). An open data ecosystem can be defined as "a multi-level and multidimensional entity where raw material, as far as distribution and developing are concerned, is the target of cooperation"(Poikola, Kola, & Hintikka, 2011, p. 13). Open data ecosystems are characterized by the interaction of data producers,

infomediaries as intermediate consumers of data or service providers and open data users (Ding, Peristeras, & Hausenblas, 2012; Ubaldi, 2013). They consist of multiple interdependent socio-technical levels, dimensions, elements and components (Zuiderwijk, Janssen, & Davis, Forthcoming).

Charalabidis et al., 2014 has found that the data processing capabilities, a key novel feature of this new generation of OGD Infrastructures, has the strongest impact on the generation of higher level value, associated with the achievement of fundamental objectives of users, and their future behaviour. "*Another novel feature, the user-level feedback capabilities (concerning rating and commenting datasets that users download and use, and also reading other users' ratings and comments on datasets they are interested in), was found to have considerable impact on higher level value generation. Therefore, these novel Web 2.0 oriented capabilities (active data pro-sumers support) seem to be valuable and promising.*"

One essential element of open data ecosystems concerns their development "through user adaptation, feedback loops and dynamic supplier and user interactions and other interacting factors" (Zuiderwijk et al., Forthcoming). Open data ecosystems contain data cycles with feedback loops, sharing of data back to publishers and sharing between so-called infomediaries (Pollock, 2011). However, discussion and feedback loops appear barely to be part of existing open data practices and policies. Zuiderwijk and Janssen (2013) found that after open data have been used, the provision of feedback to data providers or a discussion with them is often not facilitated, while these mechanisms can be used to improve open data quality, data release processes and policies. Dawes and Helbig (2010) found that such mechanisms can help users to obtain insight in how they can use and interpret data and data value. Additionally, we believe that social media may also play an important role in the exploitation of feedback and discussion mechanisms for open data.

In this position paper we discuss to which extent existing open data infrastructures address feedback mechanisms and we discuss a case which aims to contribute to the identified gaps concerning open data feedback mechanisms.

2) Survey results

In April 2014 we conducted a survey which aimed to evaluate open data infrastructures. The survey was completed by 36 professional open data users, including researchers, civil servants/policy-makers, developers and entrepreneurs. The majority of the professional open data users were male, between 25 and 49 years old, Dutch and had some or much experience with open data use.

In the survey the respondents were asked to indicate whether they knew any open data infrastructure that enabled them to conduct various tasks related to open data discussion and giving feedback. With regard to the statement "at least one of the open data infrastructures that I know enables me to discuss what can be learned from data use by leaving a discussion post" most professional open data users indicated that they disagreed (25,0%) or strongly disagreed (19,4%) with this statement. Only 22,2% of the professional open data users agreed to a certain extent with this statement (varying from slightly agree to strongly agree). As far as sharing and discussing on social media what can be learned from data use is concerned, the survey showed that most respondents disagree (27,8%) or neither disagreed nor agreed (19,4%) with the statement that they knew at least one open data infrastructure which enabled this.

Another feedback and discussion mechanism that was investigated was the discussion of what can be learned from data use by looking at previous uses of the data (e.g. visualizations, publications and applications). To the statement that the respondents knew at least one open data infrastructure which enabled this mechanism, the majority of the professional open data users answered that they disagreed (30,6%). Opinions of other professional open data users were divided, and 25% agreed to a certain extent (ranging from slightly agree to strongly agree). With regard to the statements "I know at least one open data infrastructure that enables me to discuss what can be learned from data use by publishing experiences and articles about this on the infrastructure" and "I know at least one open data infrastructure that enables me to discuss what can be learned from the data use on a Wiki or forum"

the majority of the professional open data users were negative (52,8% and 55,5% respectively) or had a neutral attitude (16,7% and 19,4% respectively).

The survey showed that with regard to these discussion and feedback mechanisms most respondents who were professional open data users did not know at least one open data infrastructure that enabled these discussion and feedback mechanisms. The foregoing shows that much can still be done to improve feedback and discussion on open data infrastructures. In the following section we discuss an open data infrastructure which has started to contribute to filling this gap.

3) The ENGAGE open data infrastructure

The foregoing leads to the discussion about what are good ways to implement feedback and discussion loops and how we can contribute to filling this gap. In this section we describe how the ENGAGE open data infrastructure has made a first step to contribute to filling this gap. It has done so by combining the so-called classical functionalities of open data infrastructures (basic functionalities supported by most first generation open data platforms) and more novel Web 2.0 oriented functionalities.

The basic functionalities include data publication and uploading for data providers, data modeling for data providers with flat metadata descriptions based on metadata standards. In addition, these functionalities include data search for open data users via simple search via keywords, resource format, data publishers, topic categories and countries. Moreover, these basic functionalities include data visualization techniques on specific datasets (e.g. maps and charts), advanced visualization techniques on specific datasets and/or datasets mash-ups (maps, charts, plots and other) for open data users and, finally, data and metadata downloading capabilities for data users and Application Programming Interfaces for this purpose.

The novel Web 2.0 capabilities aim to support open data 'prosumers' (who are at the same time users of provided datasets, and also producers of new versions of them (through various types of processing), which are improved, enriched-extended or adapted for specific purposes, or even of new datasets), and also extensive interaction and collaboration among them. This novel functionality includes five main functionalities related to 1) grouping and interaction, 2) data processing, 3) data enhanced modeling, 4) feedback and collaboration and 5) data quality rating.

With regard to the first category, open data users can search for other open data users and data providers and exchange knowledge and experiences with them. Moreover, they can form groups and maintain datasets within these groups. As far as data processing is concerned, the ENGAGE platform provides, among other, functionalities for (meta)data enrichment, data cleansing and data format conversions. The enhanced data modeling refers to the integration of flat, contextual and detailed metadata for different metadata standards and vocabularies. The feedback and collaboration functionalities enable users to receive notifications when other open data users have processed or improved certain datasets, to discuss and comment on what can be learned from certain open datasets and to express which open data they would like to use that are not yet available or that they could not find. Finally, data quality rating refers to functionalities to assess various quality dimensions of open datasets and to be informed about the level of quality of the datasets perceived by other users through their data quality ratings per dataset.

4) Conclusions

In this position paper we aimed to discuss to which extent existing open data infrastructures address feedback mechanisms and to how the ENGAGE open data infrastructure contributes to filling the identified gaps concerning open data feedback mechanisms. Our survey showed that most respondents (professional open data users) did not know at least one open data infrastructure that enabled five specific types of discussion and feedback mechanisms. The survey showed that much can still be done to improve feedback and discussion on open data infrastructures. We discussed an open data infrastructure which has started to contribute to filling this gap. The so-called ENGAGE open data

infrastructure combines functionalities to close the feedback loop and to return information to public authorities for better open data use and publication as well as establishing communication channels between stakeholders. The ENGAGE open data infrastructure provides 'classical' first generation open data functionalities as well as a comprehensive set of additional Web 2.0 social media oriented capabilities. The novel functionalities relate to five categories, namely 1) grouping and interaction, 2) data processing, 3) data enhanced modeling, 4) feedback and collaboration and 5) data quality rating. The additional functionalities may lead to the stimulation and facilitation of value generation from open data.

References

Bertot, J. C., Jaeger, P. T. and Grimes J. M. (2012). Promoting transparency and accountability through ICTs, social media, and collaborative e-government. TransformingGovernment: People, Process and Policy, 6(1), 78 – 91.

Blakemore, M., & Craglia, M. (2006). Access to Public-Sector Information in Europe: Policy, Rights and Obligations. *The Information Society, 22*(1), 13-24.

Bonsón, E., Torres, L., Royo, S. and Flores, F. (2012). Local e-government 2.0: Social media and corporate transparency in municipalities. Government Information Quarterly, 29, 123–132.

Charalabidis, Y., Ntanos, E., & Lampathaki, F. (2011). An architectural framework for open governmental data for researchers and citizens. In M. Janssen, A. Macintosh, J. Scholl, E. Tambouris, M. Wimmer, H. d. Bruijn & Y. H. Tan (Eds.), *Electronic government and electronic participation joint proceedings of ongoing research and projects of IFIP EGOV and ePart 2011* (pp. 77-85). Delft

Y Charalabidis, E Loukis, C Alexopoulos. (2014). Evaluating Second Generation Open Government Data Infrastructures Using Value Models. (HICSS). 47th Hawaii International Conference on System Sciences, 2114-2126

Chun, S. A. and Luna Reyes, L. F. (2012). Editorial - Social media in government. Government Information Quarterly, 29, 441–445.

Criado, J. I., Sandoval-Almazan, R., Gil-Garcia, J. R. (2013). Government innovation through social media. Government Information Quarterly, 30, 319-326.

Dawes, S., & Helbig, N. (2010). *Information strategies for open government: Challenges and prospects for deriving public value from government transparency*. Paper presented at the 9th International Conference on e-government (EGOV), Lausanne, Switzerland.

Ding, L., Peristeras, V., & Hausenblas, M. (2012). Linked Open Government Data [Guest editors' introduction]. *Intelligent Systems, IEEE, 27*(3), 11-15. doi: 10.1109/MIS.2012.56

European_Commission. (2003). Directive 2003/98/EC of the European Parliament and of the council of 17 November 2003 on the re-use of public sector information Retrieved December 12, 2012, from http://ec.europa.eu/information_society/policy/psi/rules/eu/index_en.htm.

European_Commission. (2011). Digital agenda: Turning government data into gold. Brussels: European_Commission.

Margo, M. J. (2012) A Review of Social Media Use in E-Government. Administrative Sciences, Administrative Sciences, 2, 2, 148-161.

Poikola, A., Kola, P., & Hintikka, K. A. (2011). Public data. An introduction to opening information sources. Helsinki, Finland, : Ministry of Transport and Communications.

Pollock, R. (2011). Building the (Open) Data Ecosystem Retrieved September 13, 2013, from http://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem/

Ubaldi, B. (2013). Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. In N. OECD Working Papers on Public Governance (Ed.): OECD.

Zhang, J., Dawes, S., & Sarkis, J. (2005). Exploring Stakeholders' Expectations of the Benefits and Barriers of E-Government Knowledge Sharing. *Journal of Enterprise Information Management, 18* (5), 548-567.

Zuiderwijk, A., & Janssen, M. (2013). *A coordination theory perspective to improve the use of open data in policy-making*. Paper presented at the 12th conference on Electronic Government (EGOV), Koblenz, Germany.

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Sheikh_Alibaks, R. (2012). Socio-technical impediments of open data. *Electronic Journal of eGovernment, 10*(2), 156 - 172.

Zuiderwijk, A., Janssen, M., & Davis, C. (Forthcoming). Innovation with open data: Essential elements of open data ecosystems *Information Polity*.