JeDEM 6(1): 62-68, 2014
ISSN 2075-9517
http://www.jedem.org

A Platform for Closing the Open Data Feedback Loop based on Web2.0 functionality

Charalampos Alexopoulos*, Euripidis Loukis*, Yannis Charalabidis*

** University of the Aegean, Department of Information and Communication Systems Engineering, Greece, {alexop, eloukis, yannisx}@aegean.gr*

Abstract: *One essential characteristic of open data ecosystems is their development through feedback loops, discussions and dynamic data suppliers - user interactions. These user-centric features communicate the users' needs to the open data community, as well to the public sector organizations responsible for data publication. Addressing these needs by the corresponding public sector organizations, or even by utilising the power of the community as ENGAGE supports, can significantly promote and accelerate innovation. However, such elements appear barely to be part of existing open data practices in the public sector. A survey we conducted has shown that professional open data users find the feedback and discussion on open data infrastructures from their users to their providers as highly useful and important, but they state that they do not know at least one open data infrastructure that provides various types of discussion, and feedback mechanisms. In this paper we describe and discuss an open data platform, which contributes to filling this gap and also present a usage scenario of it, explaining the sequence of using its functionality. The discussed open data infrastructure combines functionalities that aim to close the feedback loop and to return information to public authorities that can be useful for better government data opening and publication, as well as establishing communication channels between all stakeholders. This may effectively lead to the stimulation and facilitation of value generation from open data, as such functionality positions the user at the centre of the open data publication process.*

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

Acknowledgement: The authors would like to thank their colleagues of the ENGAGE project for their input for this paper (see www.engage-project.eu and www.engagedata.eu), although the views expressed are the views of the authors and not necessarily of the project.

Introduction

Governments create and collect large amounts of data in various domains, such as statistical, business, tourism, 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 various 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).

In order be able to benefit from open government data and generate value from them, several researchers have argued that opening and using these data should be seen as an on-going process performed by an ecosystem of multiple collaborating entities (Pollock, 2011; Zuiderwijk, Janssen, Choenni, Meijer, & Sheikh_Alibaks, 2012). Pelet (2013) states that currently open data is an early experiment of a promising idea, and that it is important to understand that an effective open data program requires time and patience to grow. The open data ecosystem is "a multi-level and multi-dimensional 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, and elements.

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., 2014). Open data ecosystems perform data production and usage cycles with feedback loops, sharing of data back to publishers and also with the so-called infomediaries (Pollock, 2011). However, discussion and feedback loops appear barely to be part of existing open data practices and infrastructures. 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 might be useful for improving 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 open government data and generate value from them.

At the same time another major trend in government agencies has been the exploitation of Web 2.0 social media for increasing citizens' participation in the government decision and policy making processes, supporting networking, interaction and collaboration, and also collecting opinions, knowledge and ideas from citizens, and promoting government transparency and accountability (Bertot, 2012; Bonsón, 2012; Chun, 2012; Margo, 2012; Criado, 2013). Therefore it would be interesting to investigate the use of Web 2.0 social media oriented capabilities in open government data platforms for the collection of feedback from their users, and in general for enabling and promoting discussion both between providers and users, and among users, in order to facilitate value generation from them and accelerate innovation.

In this paper (initially in section 2) we discuss to which extent existing open data platforms provide mechanisms for bi-directional communication with users, and for collecting feedback from them, based on a survey conducted with expert users of open data. Next in section 3 we present the ENGAGE platform[1] case, which aims to contribute to filling the identified gaps concerning citizens' feedback mechanisms, and then in section 4 we provide a usage scenario explaining the sequence of using its functionality. Finally, in section 5 the conclusions are summarized and future research directions are proposed.

1. Initial Survey Results

In April 2014 we conducted a survey, which aimed to evaluate existing open government data infrastructures and identify weaknesses. The survey was completed by 36 professional open data users, including researchers, civil servants/policy-makers, application 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 this survey the respondents were asked (among others) to indicate whether they knew any open data infrastructure that enabled them to conduct various tasks related to feedback provision and discussion among open data users. 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 allowed 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,

[1] http://www.engagedata.eu

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%), while the opinions of the remaining were divided, with 25% being in agreement 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 most respondents, who were professional and highly knowledgeable open data users, did not know at least one open data infrastructure that provided such discussion and feedback mechanisms. This shows that much can still be done in order to improve feedback and discussion on open data infrastructures. In the following section we discuss an open data infrastructure which contributes in this direction towards filling this gap.

2. ENGAGE Platform Functionality

In this section we present the ENGAGE platform functionality divided into two main categories: the first includes the classical ones, mainly data publication-upload and modelling (metadata) for the data providers, and data search, visualization and download for the data users (see Table 1 for more details), while the second includes the novel ones based on the capabilities provided by Web 2.0 social media. These novel functionalities includes users’ groups formation and extensive communication and collaboration within them, data processing, enhanced data modelling (flat, contextual and detailed metadata), commenting existing datasets and expressing needs for new datasets, datasets quality rating, data linking, publication/upload of new versions of existing datasets, advanced data visualization (see Table 2 for more details).

Table 1: Classical Functionalities

	Functionality	Stakeholder	Description
1	Data Publication/Upload	Provider	Support for publication/upload of datasets by providers
2	Data Modelling	Provider	Capabilities of flat metadata descriptions (based on a specific metadata models)
3	Data Search	User	Simple search via keywords, resource format, publisher, topic categories and countries
4	Data Visualisation	User	Simple visualisation techniques on specific datasets (maps, charts)
5	Data Download	User	Data and metadata downloading capabilities – also provision of API for this purpose

Table 2: Novel Web 2.0 Functionalities

	Functionality	Description

1	Grouping and Interaction	Capabilities for: (a) searching for users/providers having similar interests with us in order to have knowledge exchange – cooperation, (b) forming groups with other users/providers having similar interests with us in order to have knowledge exchange – cooperation, (c) maintaining datasets/working on datasets within one group, (d) communicating with other users/providers through messages in order to exchange knowledge, (e) getting immediately updated about the upload of new versions and enrichments of datasets maintained/worked on within the group, or new relevant items (e.g. publications, visualizations, etc.).
2	Data Processing	Capabilities for: (a) data enrichment (e.g. adding new elements – fields), (b) for metadata enrichment (e.g. filling missing fields), (c) for data cleansing (e.g. detecting and correcting ubiquities in a dataset, matching text names to database IDs (keys) etc.), (d) converting datasets to another format, (e) submitting various types of items (e.g. visualisations, publications) related to a dataset, (f) datasets combination and mash ups.
3	Enhanced Data Modeling	Capabilities for description of flat, contextual and detailed metadata of any metadata/vocabulary model.
4	Feedback and Collaboration	Capabilities for: (a) communicating our own thoughts and ideas on the datasets to the other users and the providers of them through comments on them, (b) reading interesting thoughts and ideas of other users on the datasets expressed through comments they enter on them, (c) expressing our own needs for additional datasets that would be interesting and useful to us, (d) getting informed about the needs of other users for additional datasets, (e) getting informed about datasets extensions and revisions.
5	Data Quality Rating	Capabilities to (a) communicate to the other users and the providers the level of quality of the datasets that I perceive, (b) get informed on the level of quality of the datasets perceived by other users through their ratings.
6	Data Linking	Capabilities of data and metadata linking to other ontologies in the Linked Open Data Cloud.

		Capabilities of querying data and metadata through Sparql Endpoints.
7	Data New Versions Publication	Support for publication/upload of new versions of the existing datasets, and connection with previous ones and initial datasets.
8	Data Visualisation	Advanced visualisation techniques on specific datasets and/or datasets mashups (maps, charts, plots and other).

3. Usage Scenario

In this section we present a typical usage scenario that utilizes the functional architecture described in section 3, which illustrates how the ENGAGE platform can be used, focusing on its Web 2.0 social media oriented features (although as mentioned above the 'classical' features are also available, such as capabilities for uploading and downloading datasets). The usage scenario presented in this paper is an ideal one, and includes a number of steps that can be taken by a researcher who uses this open government data (OGD) platform:

- First, the researcher goes to the OGD platform, registers with one of his/her social media accounts (e.g. Facebook or LinkedIn), searches for a particular dataset, views this dataset and has a look at how other users of the OGD platform have used this dataset and assessed its quality.
- Taking into account what other users of the OGD platform wrote about they have (or can) use this dataset for similar purposes with the ones of the researcher (which is a very useful knowledge transfer), he/she proceeds with the analysis of the dataset, by visualising it in tables and charts, and possibly on a map, and finally drawing conclusions from these analyses and visualizations.
- Subsequently, the researcher participates in a discussion about the above dataset, by providing feedback or discussing the dataset, posting the conclusions that he/has drawn from the use of it (in a discussion area accessible while viewing the dataset); all other users of the platform can see the messages that the researcher has posted and can respond to them by adding a message themselves, and in this way getting involved in the discussion. All these posted conclusions provide an overview of what has been learnt from the analysis of the dataset. This information is visible to any user of the OGD platform, including the providers of these data, who can also view this and assess how the data that they have opened are used. This allows them to gain a better understanding of the value (commercial, social, political, scientific) generated from the datasets they have opened, and provide them guidance and support for making decisions as to what kinds of data they should open next, or required adaptations of their data publication processes.
- As a next step the researcher connects the original dataset with results of the use he/she has made of this dataset (i.e. connects the raw data to the processed data, and also to publications, visualisations, applications and results based on these data).
- Then the researcher rates the quality of the dataset and writes for which purposes the data quality was good enough, and possibly mentions other purposes for which it is not of satisfactory quality, and proposes possible improvements.
- Next the researcher disseminates his/her findings by sharing them via social media (e.g. Twitter, Facebook, LinkedIn), linking the accounts used for this to the OGD platform; this helps him/her to make this research more visible to other researchers.

- The researcher also uses the Wiki forum in the OGD platform to discuss general data use possibilities and issues.
- Furthermore the researcher can write a tutorial on how certain data analysis and visualisation tools that he is familiar with can be used.
- Finally, the researcher has a look at the postings of all other users who have registered in the platform. He decides to send a private message to another platform user to discuss whether this other user would be interested in writing a joint paper in the future; if the other user agrees, they can create and open or closed group to work on the data processing together. They may also consider requesting the community to provide them with another additional dataset that the researcher wants to use in order to be able to write the publication. The OGD platform enables posting data requests, which can be answered by government organizations or other users of the platform.

The above scenario shows how such a Web 2.0 OGD platform can be used in order to stimulate user interaction and collaboration, and finally value generation from them.

4. Conclusions

In this paper we initially discussed to what extent existing open data infrastructures provide feedback mechanisms (from open data users to providers), and then describe how the ENGAGE open data infrastructure contributes to filling the identified gaps in bi-directional open data users – providers communication. An initial survey we have conducted showed that most respondents (who were highly knowledgeable professional open data users) did not know at least one open data infrastructure that provided various types of discussion and feedback mechanisms. The survey showed that much can still be done in order to improve feedback and discussion in open data infrastructures. In this direction we described an open data infrastructure which contributes to filling this gap: the so-called ENGAGE open data infrastructure provides functionalities to close the feedback loop and return information from open data users to public authorities that can assist in improving open data publication, as well as in establishing communication channels between all open data stakeholders. In this direction the ENGAGE open data platform provides both 'classical' first generation open data functionalities as well as a comprehensive set of additional Web 2.0 social media oriented capabilities; the latter can be grouped into nine categories: 1) grouping and interaction, 2) data processing, 3) enhanced data modeling, 4) feedback and collaboration, 5) data quality rating, 6) data linking, 7) data new versions publication and 8) data visualisation. These additional novel functionalities may lead to the stimulation and facilitation of value generation from open data.

Further research is required in this direction, including the development of more advanced and mature versions of the platform, and its evaluation by various categories of 'professional' users (e.g. researchers, journalists, politicians, value added services and application developers), in order to assess better the value it provides to each category, and identify possible weaknesses and required improvements. In particular, it would be quite interesting to investigate to what extent and how such a novel Web 2.0 oriented OGD platform can support the four types of OGD value generation mechanisms proposed by Jetzek et al. (2013): i) government efficiency gains (public sector organizations through OGD generate economic value by increasing their internal efficiency and effectiveness), ii) government transparency improvement (public sector organizations generate social value by offering increased transparency into government actions, which reduces 'information 'asymmetry' between government officials and citizens, which can lead to misuse of public power for private benefits and corruption), iii) private sector innovation (private sector firms generate economic value through the creation of new products and services using open government data (possibly combining various types of them, from various government sources, and also other private sector data as well)), iv) private participation and collaboration (private sector firms generate social value through participating in the formulation of government policies and in general collaboration with government).

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.

. (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.

Jetzek, T., Avital, M., Bjorn-Andersen, N. (2013). The Generative Mechanismsof Open Government Data. In Proceedings of of the 21st European Conference on Information Systems (ECIS) 2013, Utrecht, The Netherlands.

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

Pelet, A. (2013) Re-Designing Open Data 2.0. Journal of eDemocracy and Open Government 5(2): 187-199

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., 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., (2014). Innovation with open data: Essential elements of open data ecosystems. Information Polity, 19(1), 17-33, IOS Press

About the Authors

Charalampos Alexopoulos is a Researcher at the Information Systems Laboratory of the University of the Aegean working on high-level policy making, research and pilot application in FP7. He is a computer science graduate from the University of Peloponnese with MSc in Information Systems Management from the University of the Aegean. His research interest lies in the areas of open data, interoperability and IS evaluation.

Euripidis Loukis is Associate Professor of Information and Decision Support Systems at the Department of Information and Communication Systems Engineering, University of the Aegean, Greece. He has also lectured in both undergraduta and postgraduate courses at the National Technical University of Athens, the National Academy of Public Administration and the University of Thessaly. He has extensive experience of national and international research cooperations. His research interests are e-government/e-participation, information systems impact and business value, medical decision support *systems.*

Yannis Charalabidis is Tenured Assistant Professor of e-Government Information Systems at the Department of Information and Communication Systems Engineering, University of the Aegean, Greece. He also serves as the Scientific Director at the Greek Interoperability Centre, hosted at Decision Support Systems Laboratory of the National Technical University of Athens, delivering high quality research in the area of interoperability. He writes Downloaded by Aegean University At 02:06 and teaches on government service systems, enterprise interoperability, government transformation and citizen participation.