Pathologies of Open Data Platforms and Desired Transparency-Related Affordances for Future Platforms

Adegboyega Ojo Lukasz Porwol Mohammad Waqar Edobor Osagie
Arkadiusz Stasiewicz Fatemeh Ahmadi Zeleti
Insight Centre for Data Analytics
National University of Ireland, Galway
name.surname@insight-centre.org

Michael Hogan
National University of Ireland, Galway
michael.hogan@nuigalway.ie

Owen Harney
National University of Ireland, Galway
owen.harney@gmail.com

Pauline Riordan
Dublin City Council
pauline.riordan@dublincity.ie

ABSTRACT
The increasing volumes of datasets published on open data platforms have had little impact on the public use of open data and perceived transparency of respective governments. At the same time, the innovation potentials of these datasets are far from realized due to many factors including poor quality of datasets. While past studies have attempted to catalog barriers to open data exploitation and use; few studies have focused on the role of the available open data platforms in tackling this problem. In addressing this gap, this research work examines the problems (or pathologies) associated with the use of current generation of open data platforms and perspectives of stakeholders on desirable features and affordances. Results from our analysis of existing platforms and stakeholders' views show several limiting factors on available platforms. Findings also provide insights into three categories of platform affordances that could spur greater use of open data published on these platforms and enhanced transparency of respective governments.

CCS Concepts
Information systems, Data Management Systems, General and reference, Cross-computing tools and techniques, Evaluation

Keywords
Open Data Platforms, Open data and Transparency; Platform Affordances, Transparency Qualities, Future Open Data Platform

1. INTRODUCTION
Opening up government data to the public has been recognized to have a significant impact on enhancing transparency and openness of public sector entities while promoting new forms of accountability and improving citizens' trust in governments [2]. In response to the European Public Sector Information (PSI) directive, many European Union (EU) member states have launched their Open Data initiatives [3] with over 8,000 datasets available on the EU Open Data Portal. However, due to persistent barriers such as limited access and use of open data by citizens and third-parties, limited budget and resources on the part of government agencies to publish new datasets of high value, and weak legislative framework to enable ethical reuse of available datasets [5], the high expectations have not been met. Few studies like [6] have highlighted the centrality and importance of open data platforms in enabling greater transparency through access, use and interaction around of open data. In this work, we argue that the shortcomings of open data platforms and, more importantly, their affordances can have a significant impact on the transparency-related outcomes of open data initiatives. Thus, a thorough understanding of both the limitations and required transparency-related affordances of open data platforms is imperative for progress in this domain.

2. METHOD
Four data gathering methods were adopted in the study in order to investigate the pathologies of current open data platforms and desired affordances of future open data platforms. The first source is extant literature on open data platforms. The second is the survey of selected open data platforms through hand-on use of these platforms and review of accompanying documentations. The third source of information is through expert interviews. Four open data expert stakeholders (publisher, data intermediaries or wrangler, platform developer and end-user) were interviewed. This fourth source of data is through a Collective Intelligence Workshop organized in collaboration with Dublin City Council in April 2015 to discuss perceived barriers and problems with using the current set of open data platforms and desired features in future platforms.

3. RESULTS
3.1 Features of existing platforms
We investigated the features available on 11 state-of-the-art open data platforms. These platforms were selected depending on their popularity with respect to their installed bases. The platforms reviewed include: CKAN, DKAN, Socrata, PublishMyData, Information Workbench, Enigma, Junar, DataTank, OpenDataSoft, Callimachus, DataTank and Semantic MediaWiki. Overall, while features like the use of social media channels, customisation and personalisation of platform features are common place in state-of-the-art platforms, support for metadata schema adaptation, options for Visualisation of datasets and accessibility (including at granular level) to datasets are limited. Features like availability of publishing pipelines or workflows are visualisation still relatively limited on existing platforms. Whereas, personalisation and customisation feature are very common features of platforms. However, it must be noted that in terms of social media integration, these platforms only allow a link to social media accounts. Personalisation in the context of this evaluation is only limited to end-user ability to change the behaviour of the platform based on preferences and does not

extend to the aspects like the recommendations of datasets to end-users based on relationships with other users or preferences.

3.2 Platform Pathologies

This section presents a summary of the data obtained from interviews and workshop sessions. Our analysis showed that the most common barrier to the use of open data platforms and open data (visible as a spike in Figure 1) is *perceived the poor quality of open data available on the platforms*. Poor data quality according to stakeholders is associated with poor metadata, failure to use the right format for different audience and difficulty in locating data of interest. Other barriers identified are related to non-relevancy of available datasets, the usability of platforms and data available on the platform and lack of example of the prior use of available datasets.

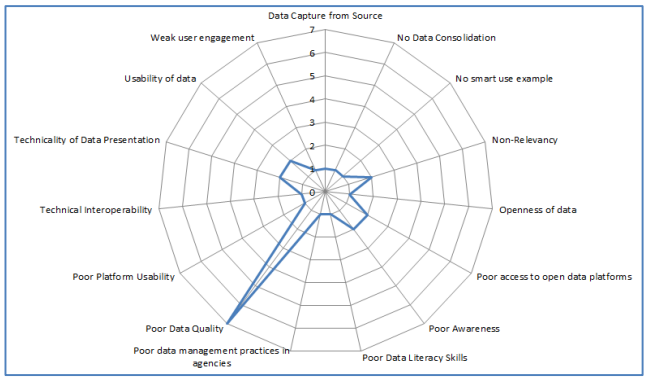

Figure 1: Perceived Barriers to Use and Adoption Open Data Platforms

In Figure 2 we present the associated transparency issues that are related to the above obstacles:

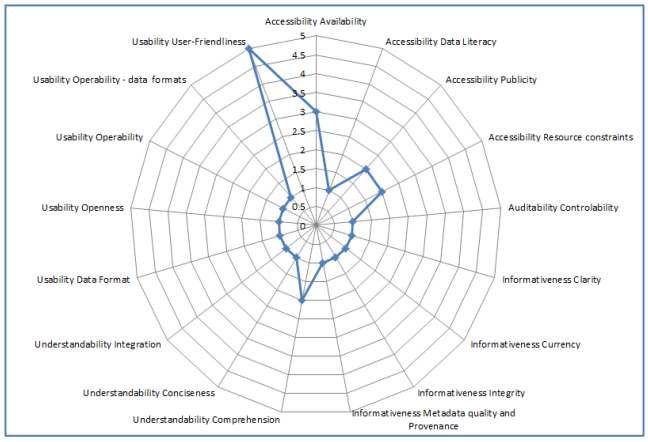

Figure 2: Data Transparency attributes related to the Perceived Barriers

3.3 Desired Affordances

The desired features contributed by stakeholders for next generation open data platforms were captured under three categories: 1) Information needs, 2) Social and Collaboration, and 3) Understandability, Usability and Decision making needs. Regarding information need, stakeholders wanted the platform to provide access to datasets about their immediate communities like crime statistics, public health data and data about their environment. Dataset rating and feedback on datasets, Wall style feedback, collaborative curation of datasets, prioritization and voting on dataset requests, reward system and gamification are some of the features expressed under the social and collaborative needs. To enable better understandability, usability and better decision making on next generation platforms, users requested for customizable dashboards, data mining tools and custom visualization tools, support for linked data and map-based search as well as a question and answering features.

4. DISCUSSION

Our study showed that regular end-users like members of the public *require significantly friendlier or more usable platform to further transparency goals*. Drawing from sound practices in the e-government domain where citizen-centric and one-stop service design are imperative for uptake and use of e-services*; we argue that open data portals must offer the public a one-stop access to "data services"*. Regardless of the transparency context considered, social interaction among members of the community is important. *Consequently, any platform support for the use of open data in a transparency context should include support for social interaction. Interviewed experts and workshop participants requested features for sharing and discussing datasets*. The integration of open data and social media platform was the subject of research reported in [1]. An interesting finding is the *demand for anonymity in the use of open data by end-users*. This raises a point that citizens still treat open data portals like a government website. Our results are consistent with the Open Data Barometer report [4] indicating lack of data quality, trustworthiness and relevant data and the need for more timely data a need for sustained ways to maintain open data and to keep it up to date are preventing open data initiatives from achieving their ends. In transitioning into next generation open data platforms; it is reasonable to expect that future platforms will be built upon existing ones. Our review of the existing platforms shows some of the current platforms could have an open architecture and could be extended to accommodate new features.

5. CONCLUSIONS

We have sought in our work to provide a better understanding of the shortcomings of the current open data platforms and desirable affordances for next generation ones. This work complements existing research as it focuses on the evaluation of the platform from perspectives of open data transparency. Other current reports have focused primarily on the technical aspects of the platforms. In addition, the complementary analyses of the stakeholders input on platform pathologies and desired affordances provide a practical context for the technical evaluation.

6. REFERENCES

[1] Alexopoulos, C., Zuiderwijk, A., Charapabidis, Y., Loukis, E., and Janssen, M. Designing a second generation of open data platforms: integrating open data and social media. In *Electronic Government*. Springer, 2014, 230–241.

[2] Bonsón, E., Torres, L., Royo, S., and Flores, F. Local e-government 2.0: Social media and corporate transparency in municipalities. *Government information quarterly 29*, 2 (2012), 123–132.

[3] Colpaert, P., Dimou, A., Sande, M. Vander, et al. A three-level data publishing portal. (2014).

[4] Davies, T., Sharif, R.M., and Alonso, J.M. Open Data Barometer Global Report. *World Wide Web Foundation*, (2015).

[5] Janssen, M., Charalabidis, Y., and Zuiderwijk, A. Benefits, adoption barriers and myths of open data and open government. *Information Systems Management 29*, 4 (2012), 258–268.

[6] Peled, A. Effective Openness – The Role of Open Data 2 . 0 in a Wider Transparency Program. *3rd Global Conference on Transparency Research*, (2012), 44–46.