Unraveling the mystery of Open Government Data Apps

Nikolaos Loutas[1], Anastasia Varitimou[2], Vassilios Peristeras[3]

[1]DERI, NUI Galway, Ireland; nikos.loutas@deri.org
[2]International Hellenic University, Thessaloniki, Greece; a.varitimou@ihu.edu.gr
[3]European Commission, Directorate-General for Informatics, Interoperability Solutions for European Public Administrations; vassilios.peristeras@ec.europa.eu

1. Introduction

In the Digital Agenda for Europe, the European Commission (EC) [1] is emphasizing the importance of opening up data resources for reuse. The great potential of open public sector information is widely recognized and the impact this would have on transparency, innovation and the real economy is indisputable [2, 3].

Towards this direction, the EC encourage and facilitate Member States to open up government data and metadata [4,5]. Governments in turn have started taking vital steps towards opening up their data [6]. As the volume of data opened up all over the world through national or local initiatives as well as by organizations such as Eurostat and the World Bank is steadily growing, open government data (OGD) consumption and usage in different types of applications has become an index to justify this enormous effort.

In this vein, OGD providers often invite and encourage software developers to create applications that use OGD (e.g. through hackathlons and contests) and integrate into value-added applications. As a result, nowadays hundreds of OGD apps exist. However, a holistic view of the growing space of OGD apps is currently missing.

Towards this direction, this position paper emphasizes the need to explore OGD apps in order to identify their types and categories, the datasets they use, their delivery models, their intended audience and other aspects discussed later on. The survey commenced end of March and is still in progress. This paper discusses preliminary results and discusses future research directions.

2. Data Collection, Presentation & Analysis

The Web search started from the app catalogues and directories of national and local OGD portals maintained by OGD champions, such as data.gov, data.co.uk and Ottawa city. The search was expanded to cover also the journalist world that has also joined the open data initiative, e.g. the Guardian Open Platform. The open data portals of large non-governmental organizations that publish OGD, e.g. World Bank and Eurostat, were also examined. Finally, a general purpose Web search using Google Search Engine was also carried. The most frequently used search terms include: OGD, open government data, open data, app, application, linked government data, as well as their combinations.

Acknowledgement: The work presented in this paper has been funded in part by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion-2).

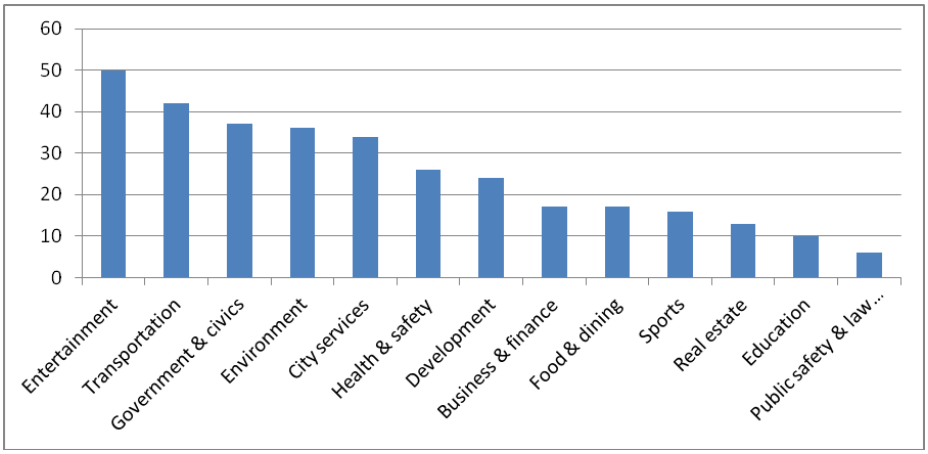
The metadata used for describing OGD apps (usually the ones published in catalogues) were also documented. The analysis of the metadata gave us a superset of common terms that we then use for documenting the apps. Hence, the following information is kept (wherever possible due to availability constraints) for each of the OGD apps discovered: Name, URL, Description, Publisher, Catalogue, Programming Language & Framework, Reuse & Pricing (e.g. open source, require fee), Delivery mode (e.g. mobile, web, desktop), Category, Datasets used, Vocabularies used (e.g. VoID [7] or dcat [8]), Use Linked Data, Usage info (e.g. downloads or likes), Date of creation, Date of last modification, Intended audience, and Status.

The initial search of this survey yielded more than 350 apps up to the time when this position paper was submitted. The majority of them were found in government and city catalogues. 46 of them came from the Guardian Open platform, 10 from the World Bank and 17 apps were found in independent sites.

We observe that by now numerous cities, including Ottawa, Toronto, New York, Seattle, Colorado, Dublin, London, Rennes, Berlin, Stockholm and Singapore, have carried out an OGD apps contest of some type to bootstrap the use of their data. This paradigm has also been followed by other types of OGD providers, such as the World Bank.

We have identified 13 categories of OGD apps, namely health & safety, entertainment, sports, transportation, city services, real estate, environment, education, public safety & law enforcement, food & dining, development, business & finance, government & civics. This classification was created by integrating and generalizing classifications of apps found in different catalogues. Apparently, different classifications are also possible, namely according to other metadata elements used for documenting the apps.

Graph 1 illustrates the number of OGD apps per category reviewed so far. We observed that local initiatives such as hackathlons and OGD app challenges organized by cities delivered apps mostly related to entertainment and transportation. This justifies the fact that these two categories, followed by environment and government & civics, are the most popular ones.



Graph 1: Number of OGD apps per category

The vast majority of OGD apps have been developed by individual developers, freelancers and research institutes. This finding indicates that the business community is not yet convinced about the business and economic potential of OGD, and are thus reluctant to experiment with apps which could produce revenues and increase their firm's visibility in a rapidly evolving community.

Currently over 90% of OGD apps are available for free. Few of them (found in App Store and in Android Market) require a monthly subscription fee.

Most of the apps are Web-based followed by apps developed on proprietary mobile frameworks, mainly iOS, Android and Windows Mobile. Some apps can be deployed on different platforms.

OGD providers develop usually support apps and services (e.g. publish/subscribe mechanisms and rss feeds), focused mainly on facilitating the access to OGD and on communicating changes (e.g. release/update/deprecation of a dataset).

OGD apps are all about data visualization. Different ways of visualizing data are come across, e.g. using graphs, maps, tables etc. For example, apps classified under health & safety use graphs and charts to visualize information related to obesity statistics, smoking and non-smoking venues, infant mortalities and the spread of diseases. Along the same lines, Government & civics apps use visual methods to bridge the gap between citizens and central government and deliver PSI in an easy to digest manner. For example, they visualize the votes of elected representatives on specific bills, and the spending of both public bodies and individual elected representatives.

Approximately 90% of the apps surveyed combine OGD with maps, mainly Google Maps and Open Street Map, in order to visualize data and provide value-added location-based services. Many of these apps are also context-aware, i.e. they identify the exact location of the user either by using GPS coordinates or by searching based on postcodes, thus offering a richer and personalized user experience. For example, a significant part of the apps under health & safety locate hospitals, health facilities, emergency facilities (e.g. firehouses), pharmacists and care homes and guide to the user to the closest one. In order to select the closest facility they either use GPS coordinates for finding the user's location or support search based on postcodes. Likewise, entertainment includes a wide variety of apps, e.g. for finding parks, special events, theatres, gyms, points of interest, shops and all sorts of recreational activities, while city services location-aware apps allow citizens to find the closest recycling point or provide local employment market information.

Most of the OGD apps rely on static datasets. However, few of them consume real-time OGD. These are found under transportation, environment and public safety & law enforcement. For example, many apps under transportation visualize traffic information on city maps or inform the user on expected arrival times of busses and trains. Other apps under environment consume real-time UV and pollution data to warn the user in cases of emergency, while apps classified under public safety & law enforcement use real-time OGD to inform the user on crimes, accidents or other types of emergencies that have occurred within a specific distance and/or timeframe.

We observe that a significant portion of OGD apps rely on a single dataset, while few of them integrate more than one datasets. In our view though, the power and the real value of OGD can only be released through the integration of complementary datasets. We expect OGD integration in apps to gain popularity as linked data and

semantic technologies become more mature. Currently, however, only few apps use these technologies.

Some apps integrate OGD with data coming from the Social Web, e.g. user's opinions or even data from Wikipedia. For example, real estate apps combine OGD with user's preferences or engage citizens in discussions around urban development. Likewise, food & dining apps combine OGD with user's ratings in order to rank the quality and the cleanliness of restaurants.

Finally, a number of apps classified under city services and government & civics use OGD and visualization technologies in order to facilitate the communication between (local and/or central) government and citizens and to engage citizens in politics. For example, citizens can use OGD apps to report to the city issues related to graffiti, potholes, excessive garbage, street problems or street lighting. Other apps enable citizen engagement and participation by facilitating citizens to express opinions, interact with elected representatives, raise issues for discussion and comment on the activities of the (local and/or central) authorities.

3. Open Questions & Future Research Directions

This position paper reported the preliminary results of our survey on the OGD apps. The outlook of OGD apps is certainly positive; however certain factors that will impact their future growth should be investigated. We identify these factors, each of which constitutes a separate future research strand.

Open data initiatives are expected to promote transparency, foster collaboration across government and beyond, allow the creation of new, innovative, added-value services, and improve the quality of decision-making. To what extent do existing OGD apps contribute towards these objectives and how can their contribution be strengthened?

Integration of datasets is currently not the prevalent trend in OGD apps. However, we argue that data integration will unveil the real value of OGD. Additionally OGD is often made available in non-machine-readable formats, which also hampers ODG use and integration. How can semantic technologies and linked data revolutionize the OGD initiative and OGD app development in particular?

Currently, finding an OGD app that will fulfill the needs of an individual is not a trivial task. It may be published in an OGD app catalogue or it may be available anywhere on the Web just as well. How would a common metadata model for describing OGD apps and lightweight semantics facilitate OGD app discovery?

OGD apps can then be made available following different exploitations roots (from commercial applications to free and/or open-source ones. But can traditional business model fulfil the peculiarities of OGD apps?

Finally, by exploring existing apps in detail, one can collect valuable information related to the most used datasets, the most popular types of apps and the lifecycle of an app.

References

1. European Commission (2010). Digital Agenda for Europe 2010 - 2020, Retrieved from: http://ec.europa.eu/information_society/digital-agenda/index_en.htm. [Accessed May 2012].
2. Advisory Board on Public Sector Information (2006). Realising the Value of Public Sector Information. Retrieved from: http://www.nationalarchives.gov.uk/documents/appsi-annual-report-2006.pdf. [Accessed May 2012].
3. Australian Government, Office of the Australian Information Commissioner (2011). Issues Paper 2: Understanding the value of public sector information in Australia. Retrieved from: http://www.oaic.gov.au/publications/papers/issues_paper2_understanding_value_public_sector_information_in_australia.html. [Accessed May 2012].
4. European Commission (2008). Results of the online consultation of stakeholders 'Review of the PSI directive'. Retrieved from: http://ec.europa.eu/information_society/policy/psi/docs/pdfs/online_consultation/report_psi_online_consultaion_stakeholders.pdf. [Accessed May 2012].
5. European Commission, DG INFSO (2010). eGovernment Action Plan 2011 – 2015. Retrieved from: http://ec.europa.eu/information_society/activities/egovernment/ action_plan_2011_2015/index_en.htm. [Accessed May 2012].
6. Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. In: Government Information Quarterly 28, 446–456.
7. Alexander, K., Cyganiak, R., Hausenblas, M., Zhao, J. (2011). Describing Linked Datasets with the VoID Vocabulary. W3C Interest Group Note. Retrieved from: http://www.w3.org/TR/void/. [Accessed May 2012].
8. Maali, F., Cyganiak, R., Peristeras, V. (2010). Enabling Interoperability of Government Data Catalogues. In Proceedings of the Electronic Government 10th International Conference, EGOV 2010, Lausanne, Switzerland.