Perspectives on Open Data: Motivation, Benefits and Examples

Peter Conradie | Creating 010, Rotterdam University of Applied Sciences | p.d.conradie@hr.nl

Introduction

Recent years has seen the introduction of the term Open Data. This denotes data that is reusable and distributable without cost. Open Data could thus include any data object that can be used by third parties for free, with naming the source of the data as only requirement. While the term Open Data has gained popularity in recent years, it describes a phenomenon that is rooted in movements such as Open Source Software. This movement has transformed from being driven by individual enthusiasm to commercially viable products [9]. For public sector data, judicial frameworks that ensure personal data access by individuals have been instrumental in opening up government.

In this paper, we will introduce the concept of open data, from the point of view of government, private companies and individuals. It gives an oversight of its origins and motivations, looking at judicial, societal, cultural and financial aspects. We also highlight some examples of open data implementation, concluding with the future of open data.

Origins

Public Sector

Legislation that prescribes access to government information, known as public sector information, is now prevalent in many industrialized countries [15, 16]. From a government open data perspective, such legislation has been invaluable in supporting data release. Within these laws, it is described that citizens should be afforded access to data, in the broadest sense of the word. Access to this data has traditionally meant requesting physical copies of internal government information. However, many government processes also rely on digital systems to achieve tasks, which means that data is already digitally available, which could further support data release.

Private Sector

While the government has traditionally been one of the biggest owners of data, the private sector also controls vast amounts of information. Various tasks that were previously conducted solely by government organizations have been taken over by the private sector. Examples of this would be telecoms providers such at the Dutch KPN or German T-Mobile. Organization such as the post office, traditionally government run, has also been privatized. Due to their status as non-governmental organizations, these companies are not subjected to traditional freedom of information legislation. However, it has been argued that the private sector should also comply with some forms of information access legislation [17].

We find evidence of this in recent calls by the European Commission that could potentially enable private citizens much better access to personal data, as stored by commercial companies, such as Facebook [7]. While this relates to personal data, certain semi-public organizations such as the Dutch Railway (NS), have taken the initiative to offer access to their transport data. Recent movements from the scientific community has also indicated the wish to ensure better access to scientific results [14], by freeing research findings and allowing it to be distributed without charge.

From the point of innovation, more open approaches have also been advocated. From academia and commerce, arguments has been made for a more transparent innovation culture inside companies, whereby results are shared and published, in order to fuel innovation [1].

Individuals

Along with the public sector and companies, people are also increasingly sharing personal information. The ever-decreasing cost of sensor technology means that is becoming increasingly affordable for people to gather high quality data and share it online. Indicative of this are platforms such as Pachube[1] or SAWA [8]. These platforms allow people to submit sensor data to central servers. Other sources of personal open data could also be data stored by social networks such as twitter and Facebook. Strictly speaking, these data are not currently automatically classifiable as open data, but legislation such as proposed by the European Parliament [7] could result in better access to these datasets, which means that they could be used by ourselves, or application developers to develop innovative applications.

From a cultural point of view, the release of data is also relevant. Lessig [12] describes the shift of read-only culture (perhaps best embodied by television), to remix culture, where individuals are copying and remaking art. Results from such cultural remixes are subsequently also built upon by others, further creating value.

Value of Open Data

This illustrates that releasing data for use and re-use by third parties is gaining traction on various levels. It also begs the question what the motivation is behind such calls for data release. While the freedom of information legislation discussed earlier is perhaps a good argumentation why data should be made

[1] https://pachube.com/ accessed on 4 April 2012

available, we can identify two main reasons for supporting data release: innovation and transparency. These two are not necessarily mutually exclusive and can often be found overlapping. Increased transparency also relates to other benefits that open data could contribute to, namely increased participation in political life, stronger democracy or e-governance. These hopes have been mentioned by public sector decision makers on the highest level, as illustrated by United States President Obama [18], European Commissioner Kroes [4] and (former) Minister Donner [5]. While increased transparency relates in this context mostly to the government, it has been argued that it should also apply to the private sector, as mentioned before [17].

Innovation forms the second category of benefits associated with open data. While placing exact figures on the financial benefits of open data is hard, certain studies have shown that there is indeed a significant financial value of open data. For example, a study commissioned by the European Council remarks that 27 billion is a reasonable estimate for the size of the open data market in Europe [3]. Based on a calculation of the current Danish market, also commissioned by the European Council [21], the value of the open data market for the Netherlands is estimated at 200 million euro [19].

These calculations are based on estimated values of data that is currently in control of governmental organizations. It would not be far fetched to say that the private and semi-private sector, embodied by companies such as the Dutch railways, Twitter or Facebook, could also contribute to the monetary value of open data, further adding value.

Besides the directly calculable financial potential of open data, more access to public sector information could also have broader economic benefits [20]. For example, better information about public transport could result in increased efficiency for commuters. The benefits to researchers are also apparent: with better access to high quality data provided by public sector bodies, scientists can have a much better insight into certain subjects. Re-use of data is thus important from various perspectives. What follows is a review of some examples of how open data can be applied, further underscoring its potential.

Open Data Use and Re-use Illustrated

Through pilot research and case studies, the value of open data can be illustrated. For example, in [13], by involving various stakeholders in an open innovation process, data release can be explored. Below we introduce a some projects and cases that highlight how data from various sources can be re-used.

A very obvious choice for data re-use is transport information. The public transport authority of London, Transport for London, makes extensive use of data to support its service delivery, whether in the form of trains running on time, or by the use the Oyster card, with which commuters can pay for journeys. Some of this data has been released for re-use, resulting, amongst other things, in travel planners for smart phones. However, in addition to these obvious applications, this data has also been used in more playful ways. Chromaroma[2], offers such an example. By using the released data, Chromaroma has managed to build an interactive, place-based game. The game visualizes personal travel data in an innovative way, offering commuters insight into traveling behaviour and transforming the commuting experience (also see [2]).

FixMyStreet[3] is illustrative of how private initiatives can support local council objectives, using open data. This system is one the first citizen driven public service improvement [11]. It allows citizens of the United Kingdom to report issues in the public space, such as illegal garbage dumping or broken streetlights. The reported problems are submitted to the appropriate local council by the service. At the time of writing, FixMyStreet has reported more than 2700 problems fixed in the past month. Data reports are also published as open data by FixMyStreet, which enables other developers to build applications based on the gathered information.

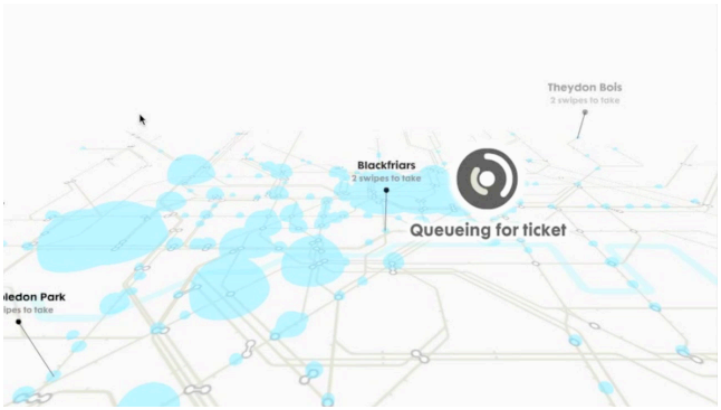

Figure 1: Visualization of a journey by Chromaroma. Image taken from http://vimeo.com/22023369 (accessed on 10 April 2012)

The price decrease of sensor technology has also led to various innovative ways of data gathering by citizens. Openstreetmap.org[4], is highly illustrative of this. By submitting GPS coordinates, volunteers have

[2] http://www.chromaroma.com/ accessed 4 April 2012
[3] http://www.fixmystreet.com/ accessed 4 April 2012
[4] www.openstreetmap.org accessed 4 April 2012

managed to develop detailed street map. This map was not only generated by enthusiastic individuals, but the source data is available freely for download on the site [10].
The most prominent example of decreased sensor technology comes in the form of the popularity of smartphones. Embedded with many sensors and able to run sophisticated applications that measure various metrics, smartphones has been implemented for scientific and community goals. For example, NoiseTube[5], a research project for participatory noise measurement [6], allows users to download an application to measure noise pollution. Data is available for download free of charge from their service. This gives people the tool to measure noise pollution for free and post results. These results can subsequently be compared with other cities.
One of the most interesting aspects of data release by the public sector, is that certain online service

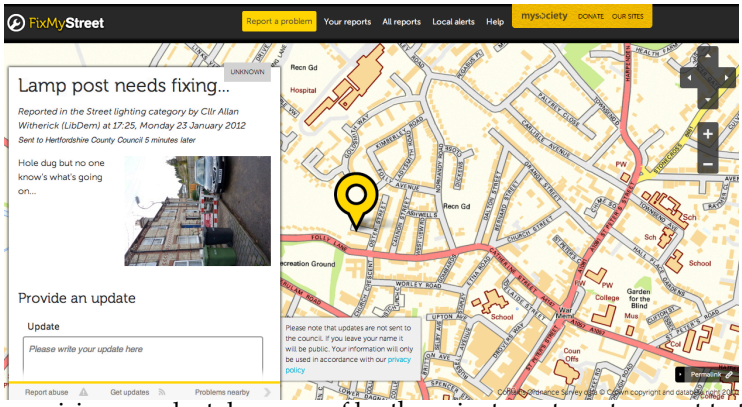

Figure 2: An example from a report on FixMyStreet about public maintenance. Taken from http://www.fixmystreet.com/report/226619 (accessed on 10 April 2012)

provisions can be taken care of by the private sector, at no cost to the taxpayer. A demonstrative example is the re-use of data supplied by various public sector bodies that manage education in The Netherlands. The data is re-used in 10000Scholen.nl[6], where schools can be rated and an overview of school statistics are given, such as number of students, teachers and on occasion, school performance.

Open Data 2020

The current state of Open Data in Europe is mixed. Governments are slowly starting to implement infrastructure that will enable data release. Through research and experimentation, barriers to data release is being categorized. Legislation as proposed by the European Commission are slowly being implemented by member states. The recent introduction of new, open data directives from Brussels will force governments on all levels to allow better access to data. It remains to be seen to what extend calls for increased data release by the private sector will lead to actual release, however, it seems that proposed legislation in the EU could at least have the potential effect of supporting much greater access to private data which could lead to innovative reuses of data. Additionally, as innovation theories such as Open Innovation illustrate, there are also financial benefits to more transparency in innovation by large companies [1]. Such an approach bodes well for data release.
However, the most promising signs for more data gathering and release come from the increased availability of sensor technology. As elaborated on earlier, private partners, thus citizens, are increasingly given the tools to collect data about their direct environment, using sensor technology that was only available to a privileged few as little as a decade ago. The systems that allow these data to be centrally captured and be reused by third parties have already appeared in the form of start-ups, but are also enjoying attention in research. This movement gives individuals more power when it comes to information provision and can be fertile ground for further open data innovation.

Bibliography

[1] Chesbrough, H., Vanhaverbeke, W. and West, J. 2006. *Open Innovation: A New Paradigm for Understanding Industrial Innovation*. Oxford University Press.

[2] Chromaroma and the onward march of gameification | Technology | guardian.co.uk: 2010. *http://www.guardian.co.uk/technology/2010/dec/04/chromaroma-gameification-mudlark-london*. Accessed: 2012-04-03.

[3] Dekkers, M., Polman, F., te Velde, R. and de Vries, M. 2006. *MEPSIR: Measuring European Public Sector Information Resources*.

[4] Digital Agenda Commissioner – Neelie Kroes - European Commission: 2011. *http://blogs.ec.europa.eu/neelie-kroes/public-data-for-all-–-opening-up-europes-public-sector/*. Accessed: 2012-04-02.

[5] Donner 2011. Betreft Hergebruik en Open Data: naar betere vindbaarheid en herbruikbaarheid van overheidsinformatie. Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.

[6] D'Hondt, E. and Stevens, M. 2011. Participatory noise mapping. *Pervasive 2011* (San Francisco, CA, USA, 2011).

[7] European Parliament and Council 2012. Proposal With Regard to the Processing of Personal Data by Competent Authorities for the Purposes of Prevention, Investigation, Detection or Prosecution of Criminal Penalties and the free movement of such data. *EC Directive Proposal - COM(2012) 10 final*. European Commission. 1-55.

[8] Faria, S. and Kostakos, V. 2012. A scalable sensor middleware for social end-user programming. *Mobile Context Awareness*. T. Lovett and E. O'Neill, eds. Springer.

[9] Fitzgerald, B. 2006. The transformation of open source software. *Mis Quarterly*. 30, 3 (2006), 587-598.

[5] http://noisetube.net/ accessed 5 April 2012
[6] http://10000scholen.nl/ accessed 5 April 2012

[10] Haklay, M. 2008. Openstreetmap: User-generated street maps. *Pervasive Computing, IEEE*. 7, 4 (Oct. 2008), 12-18.
[11] King, S. and Brown, P. 2007. Fix my street or else: using the internet to voice local public service concerns. *Proceedings of the International Conference on Theory and Practice of Electronic Governance* (Macau, 2007), 72-80.
[12] Lessig, L. 2008. *Remix: Making Art and Commerce Thrive in the Hybrid Economy*. Penguin Press HC, The.
[13] Mulder, I., Conradie, P. and Choenni, S. 2012. Rotterdam Open Data: Exploring the release of public sector information through co-creation. *ICE 2012: International Conference On Engineering, Technology And Innovation* (Munich, 2012), to appear.
[14] New York Times 2012. "Open Science" Challenges Journal Tradition With Web Collaboration - NYTimes.com. *New York Times*.
[15] Privacy International 2006. *Freedom of Information Around the World 2006: A Global Survey of Access to Government Information Laws*.
[16] Relly, J.E. and Sabharwal, M. 2009. Perceptions of transparency of government policymaking: A cross-national study. *Government Information Quarterly*. 26, 1 (Jan. 2009), 148-157.
[17] Siraj, M. 2010. Exclusion of Private Sector from Freedom of Information Laws: Implications from a Human Rights Perspective. *Journal of Alternative Perspectives in the Social*. 2, 1 (2010), 211-226.
[18] Transparency and Open Government | The White House: 2009. *http://www.whitehouse.gov/the_press_office/Transparency_and_Open_Government*. Accessed: 2012-04-02.
[19] Tweede Kamer der Staten-Generaal 2012. Nieuwe Commissievoorstellen en initiatieven van de lidstaten van de Europese Unie; Brief regering; Fiche: Richtlijn hergebruik van overheidsinformatie, mededeling Open gegevens en het besluit over hergebruik van documenten van de Commissie.
[20] Velde, R. 2009. Public Sector Information: Why Bother? *The Socioeconomic Effects of Public Sector Information on Digital Networks: Toward a Better Understanding of Different Access and Reuse Policies*. (2009).
[21] Vickery, G. 2011. *Review of recent studies on PSI re-use and related market developments*.