

Open data portals as part of the open data ecosystem? Lessons learned from geoportal research¹

Bastiaan van Loenen, TU Delft, Knowledge Centre Open Data²

Joep Cromptoets, KU Leuven, Instituut voor de Overheid

Alenka Poplin, Iowa State University, Department of Community and Regional Planning

Many countries and also cities have their own open data portal which provide geographic data that can be used even by citizens. One of the current challenges is to satisfy user needs to ensure that the data that is provided through the portal is actually used. This paper provides insights in the findability of datasets through of a special kind of portal: the geoportal. It presents the main findings of research accomplished on the findability, attainability and usability of geoportals through an assessment of the transaction costs involved.

According to Ellen Bates (2002), 107 billion dollars a year is spent by American companies on employees that spend their time trying to find the required information through the internet. This is mainly because the impulsive growth of the internet did not allow for standardization of the information put on the web, resulting into a lack of legislations and almost no consistency between the different websites. The internet has become like a whirlpool of all kinds of information, both relevant and irrelevant, with an enormous network of links between (sometimes total irrelevant) websites, from which the user has to filter the needed information. The internet gives the “illusion of an easy access” (Bates 2002), it might seem easy to find information and access the website while it is in fact quite opposite. Such complexity can discourage potential users or buyers of a particular type of product. This directly influences providers of products; i.e. the more difficult it is for users to find their products, the less likely the provider) are willing to invest in the product assortment (for instance product quality).

In the context of geoinformation, no much empirical evidence is available supporting the claims of Ellen Bates advocates. The transaction cost theory and its application to geoinformation might contribute to the assessment from a user perspective that may provide empirical data on the use of geoportals and the internet supporting to find data readily.

Backx (2003) has captured the concepts related to geoinformation search and acquisition in Figure 1. He explains that before a data set can be used, the user must pass through the two outer rings as shown in figure 1. A user should first be aware of the existence (the outer Known ring) of a data set in order to be able to obtain it (the middle Attainable ring). Transaction cost theory as applied to geoinformation by Krek (2009a; 2009b; Poplin 2010) quantifies all three rings including the fitness of use test. Geoinformation that is difficult to find will result in a thick Known ring which in terms of

¹ Position paper submitted to “Open Data for Open Cities Re-use and discovery level – AGILE 2017 Workshop”

² Paper is based on: Loenen, B. van, Cromptoets, J. and A. Poplin. 2010. Assessing geoportals from a user perspective, in the proceedings of the GeoValue Workshop, Sept. 30 – Oct. 2, 2010, Hamburg, Germany.

transaction cost theory means that the measurement geoinformation transaction cost will rise (an user has to spend more time in searching for the data). Measurement geoinformation transaction cost is the cost related to search for an appropriate geoinformation and geoinformation provider, to verification of the geoinformation quality and possible transformation to the needed format. A data set which is easy to find (thin Known ring) but difficult to obtain (thick Attainable ring) will result in low measurement transaction cost but a high enforcement transaction cost. The enforcement cost of geoinformation is the cost related to negotiating for the conditions of trade such as price of the geoinformation, enforcing agreements, protecting copyright, and defining the right to use and distribute the acquired geoinformation product or service (Poplin 2010).



Figure 1: Accessibility model of Backx (2003).

The geoinformation transaction cost appears also on the supplier's side. The supply geoinformation transaction cost is the cost imposed on the geo-information provider. Supply geoinformation transaction cost is related to explaining the complex rules about the acquisition of geoinformation, the use of data and its copyright issues. The communication happens either via email or phone and can be very costly for the providing institution. However, here we focus on measuring the potential user's transaction cost being aware that the cost exists on the supplier's side as well.

In order to be able to assess the value of geoportals from the user's perspective we designed a series of experiments.³ MSc students of the Geomatics curriculum of Delft University of Technology were assigned to assess geoportals. This was performed through several tasks:

(1) Each student conducted a literature study on both (geo)portal theory and transaction cost theory. Based on the literature study, the requirements for geoportals were developed. These criteria were discussed, and a final list of assessment criteria was agreed.

(2) Each student applied this list of criteria to assess 2 portals of their choice, a list of existing portals was provided for some guidance. The geoportals were assessed on a scale 1 (very poor) – 5 (excellent).

³ This research was described in more detail in: Loenen, B. van, Cromptvoets, J. and A. Poplin. 2010. Assessing geoportals from a user perspective, in the proceedings of the GeoValue Workshop, Sept. 30 – Oct. 2, 2010, Hamburg, Germany.

After the theoretical assessment, the students assessed the geoportals from a transaction cost perspective. They formed groups of two. One of those two would assess a geoportal by finding, assessing and accessing a dataset at their choice (Scenario 1). The other student was tasked to do the same experiment, but without using the geoportal (Scenario 2), which meant that he or she was searching for datasets through web search. Students were required to keep track of the time spent on each transaction cost stage: i.e. finding, assessing and accessing the dataset. Further they were asked to perform as if they were students that need the data for a special assignment; thus non-commercial use.

From the research it appeared that the portals sometimes, but certainly not always, are supportive to finding a dataset. In several instances datasets and/ or services could more easily be found through a general search engine search than through the myriads of portals nationally and internationally available. Moreover, datasets that were found only through a portal were not available otherwise and could only be acquired after registration (Van Loenen, Cromptvoets and Poplin 2010).

The research shows that the publication of a dataset in a portal may serve the user well. However, if the data is also properly published on the internet, which might be through the open data portal, more users will be able to find the data and in the end use the data. Since the value of data comes from its use (Onsrud and Rushton 1995), only then portals, including open data portals, will be of additional benefit to users and the open data ecosystem and worthwhile investing in.

References

- Backx, Merijn, (2003). *Gebouwgegevens redden levens [Building information saves lives]*, MSc thesis. TU Delft.
- Bates, M-E., (2002). November 2002, "Making Search Choices: Fee vs. Free" Web Search University, Dallas, TX
- Krek, A., (2009a). In Krek, Rumor, Zlatanova & Fendel (Ed.), *Quantifying transaction costs of geoinformation: Experiments in national information structures in Sweden and Germany*. Paper presented at the 27th Urban and Regional Data Management (UDMS), June 24-26, Ljubljana, Slovenia. Taylor & Francis Group.
- Krek, A. (2009b). *Measuring Transaction Costs in Spatial Data Infrastructures: Examples of Sweden and Germany*, GEOWS '09, The International Conference on Advanced Geographic Information Systems & Web Services, February 1-7, 2009, Cancun, Mexico.
- Loenen, B. van, Cromptvoets, J. and A. Poplin. 2010. Assessing geoportals from a user perspective, in the proceedings of the GeoValue Workshop, Sept. 30 – Oct. 2, 2010, Hamburg, Germany.
- Onsrud H.J. and G. Rushton (eds.) (1995). *Sharing Geographic Information*. New Brunswick, NJ (Centre for Urban Policy Research).
- Poplin, A. (2010). *Methodology for Measuring the Demand Geoinformation Transaction Costs: Based on Experiments in Berlin, Vienna and Zurich*, *Journal of Spatial Data Infrastructures Research*, Vol (5), 168-193.