

The Power of Open Statistics for Advancing the Smart City and Citizens Participation

Ari Jaakola*

City of Helsinki Urban Facts, Helsinki, FINLAND

e-mail: ari.jaakola@hel.fi

Abstract

The cities of Helsinki Metropolitan area (Helsinki, Espoo, Vantaa and Kauniainen) have determinedly cooperated in order to open administrative and statistical data for public use. The publicly available data can be found at the new web service www.hri.fi. The data can be utilized free of charge. The Helsinki Region Infoshare (www.hri.fi) web service makes public data from the Helsinki Region available for anyone to use. The 1000th dataset was recently opened.

The datasets made open cover for example data of living conditions, economy and welfare, employment and mobility. The benefits of open data are considered to be at least threefold. First, it enhances decision making and policy design of the cities. Secondly, it stimulates new economic activity and innovations. Thirdly, it will increase the citizens understanding of the operational environment and development of one's home town. Therefore, it is assumed that open data could activate the citizens and advance citizens participation in city development issues. The first experiences indicate that open statistics can act as fuel for the Smart City development. For example, open data can be used to build new web applications that benefit citizens. The data can be used to monitor, analyze and visualize social phenomena.

Key Words: Helsinki Region Infoshare, Open data, Smart City

1. Introduction

The cities of Helsinki Metropolitan Area (Helsinki, Espoo, Vantaa and Kauniainen) have determinedly cooperated in order to to open administrative and statistical data for public use. The cooperation has made under Helsinki Region Infoshare (HRI) project, which is funded by the cities and The Finnish Innovation Fund (SITRA). The Finnish Ministry of Finance has given the project a municipality cooperation grant. The operative execution of the project is managed by a directive board formed by the funding and executive parties of the project, as well as City of Helsinki Urban Facts and Forum Virium Helsinki, the organization for developing new digital services together with firms, the City of Helsinki and other public organizations and with citizens.

The HRI project was launced in June 2010, and it aims to make regional information quickly and easily accessible to all. The data may be used by citizens, businesses, universities, academies, research facilities or municipal administration. The data on offer is ready to be used freely at no cost.

The data published during the project is mainly statistical, giving a comprehensive and diverse outlook on different urban phenomena, such as living conditions, economics

and well-being, employment and transport. A good proportion of the data material offered by the project is GIS based. The project includes building a web service for fast and easy access to open data sources. Users can download information and use it in decision-making, utilise it in their applications, or develop entirely new services based on the information, to name just a few examples.

The goals of the HRI project can be summarised as follows:

- To develop an extensive network consisting of all those in the Helsinki region who are in possession of basic information pools and materials. The members of the network produce, maintain, share and develop the data pools in cooperation, following common guidelines.
- To open up the data pools of the network for everyone. The data in the pools will also be ready to be further processed or utilised in IT applications. Data is available free of charge and in a machine readable format.
- To build a web service through which the data can be easily found, downloaded and utilised. The web service is also used to encourage the producers and users of the data to cooperate in new ways.
- To pilot an open data activity model and its implications to both the producers and end users of the data. To learning by doing and sharing experiences, too.

The benefits of open data are considered to be at least threefold. First, it enhances decision making and policy design of the cities. Increasing data openness makes it easier to use for other government agencies, too. Harmonizing information management practices brings economies of scale and makes knowledge transfer between organizations easier. Open data may also help in finding practices where the potential of digital information hasn't yet been realized. Secondly, it stimulates new economic activity and innovations. By giving access to government data for free to companies, it breeds new markets and supports innovation. For example, new ways to use information have many times been found by people who have a different educational background compared to the usual users. Thirdly, it will increase the citizens understanding of the operational environment and development of one's home town. Open data supports active citizenship, research and journalism by increasing transparency. For example, it facilitates discussions in social media by making it easier to reference to government information. Therefore, it is assumed that open data could activate the citizens and advance citizens participation in city development issues.

A beta-version of the web service was issued ready for use in early 2011 and after this, the project advanced according to user-based testing and learning. In April 2013, web service has on offer over 1000 different data sets on the Helsinki region and its localities.

In this article I will describe the ways open statistics and open data can advance the smart city development and citizens participation by using the experiences and lessons learned from HRI project so far.

2. The concepts of the Smart City and Open Statistics

According to Caragliu et al. (2009) urban performance depends not only on the city's endowment of hard infrastructure ('physical capital'), but also, and increasingly so, on the availability and quality of knowledge communication and social infrastructure ('human and social capital'). The latter form of capital is decisive for urban competitiveness. Against this background, the concept of the 'smart city' has recently been introduced as a strategic device to encompass modern urban production factors in a common framework and, in particular, to highlight the importance of Information and Communication Technologies (ICTs) in the last 20 years for enhancing the competitive profile of a city. A city can be defined as 'smart' when investments in human and social capital and traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic development and a high quality of life, with a wise management of natural resources, through participatory action and engagement. (Caragliu et al. 2009)

You may say that open data and disposing hindrances for using data is needed for to enhance the smart city development. It is important that the citizens can get information about how the city works and developing. New visions and innovations require information and new ways of combine various information. Also, new tools for representing and delivering these ideas is needed. In this context, open data have a great potential.

In the HRI project open data refers to government information and public data which has been made available for free for anyone to use. However, also companies, other organizations and citizens publish open data. The openness of data means in practice that data has been made as easy as possible for anyone to use. The level of openness can vary in different aspects. The more the use of data is restricted, on purpose or accidentally, the less open it is. The following criteria for open data is set in the HRI project:

- Technical accessibility: data has been published in a format which computer programs can understand and which makes it easy to use in the development of new web services. Information in PDF documents or on most HTML web pages is often hard to use in derivative works. Publishing data in CSV or XLS formats or giving access to the original data source through an application programming interface (API) are better solutions.
- Free access: data can be accessed without costs. Free access is especially important for people exploring a data source for the first time. Free access enables experimentation with data without the need for budget bureaucracy.
- Reuse permitting licensing: the publisher permits reuse and clearly expresses it in a license accompanying the data. Otherwise finding out the terms of use can be so laborious that data is left unused.
- Findability: the existence and location of a data source should be publicly known. Publishing data in a public data catalog is a good way to improve findability.
- Understandability: the structure and the meaning of the data in question has been described and documented.

3. Helsinki Region Infoshare – The clearinghouse and distributor for open data

The Helsinki Region Infoshare project has developed four essential sectors: data production, data opening, data distribution and data utilisation (see fig.1).

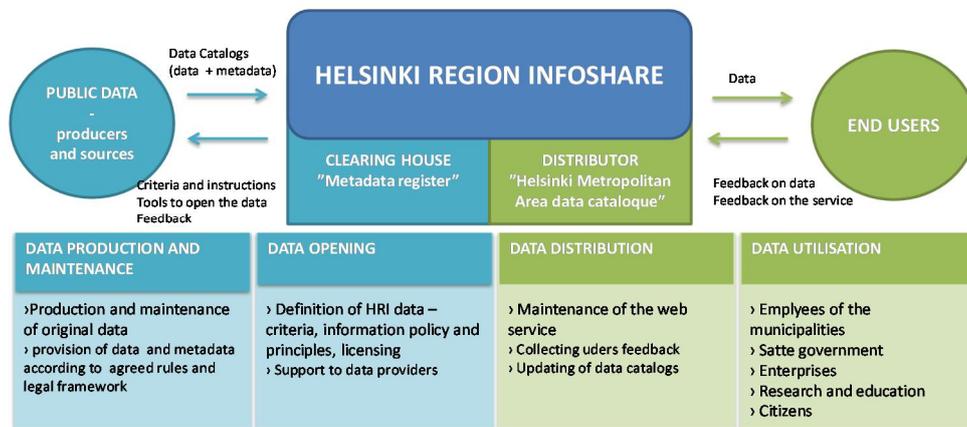


Fig 1. Helsinki Region Infoshare - Open data operational model

3.1 Data production

The project aim is to advance the opening up of existing municipal data for wider availability. Comparability, accessibility and usability of the data material requires cooperation between municipalities, or data producers. This cooperation takes the form of harmonising data as well as further development of material, among other things. It is crucial to develop data production so that it supports an open, web-based data distribution model. At the same time, end users of the data are taken into consideration.

3.2 Data opening

The project includes opening up a feature designed to assist and coordinate data accessibility. Called the Clearinghouse, this feature acts as a bridge between producers and users of data. Among other things, it prioritises data being made available, assist data producers in making their data accessible, look after quality control, and collect and distribute user feedback to the producers.

3.3 Data distribution

The project includes a web service that ensures the availability, accessibility and usability of the open data sources. This service could be described as a data source search engine which leads the end user to the correct data. The service is also utilised in collecting direct user feedback on the data sources and their applicability. Similar web services have already been implemented in e.g. United States and United Kingdom.

3.4 Data utilisation

Open data is meant for anyone interested in it: municipal administration, businesses, universities, academies, research facilities and citizens. The information can be applied in many ways. It can be used as such to aid decision-making, planning or research. Information can be added to existing or new applications. Or it can be combined and presented in an entirely brand-new way, too.

How the data may be used should not be predefined in any way. The main issue is to offer the data for a larger audience to use it as easily as possible in any way they want.

4. Experiences and lessons learned

The first experiences of the HRI project indicates that open statistics and open data is received with excitement. The number of users of HRI-data has grown rapidly. Especially the active and expanding community of data developers has taken new service in use. Finnish media has developed new ways of making journalism, called data journalism that is often based on open data. Another and fast growing field of utilisation of open data is data visualisation which has made possible to present many societal issues more understandable way for wider public.

Together with data users the number of open data sets has grown rapidly. The very first concrete data accessibility event of the HRI project was the July 2010 web launch of the Statistical Yearbook of the City of Helsinki in a downloadable package of statistical tables with more than 300 tables and charts covering fifteen different topics. In May 2013 the HRI project opened 1000th data.

The most popular data opened via HRI so far is the aerial photograph of Helsinki from the year 1943. It has downloaded over 1000 times since it was opened 8th March 2013. Geospatial data as a whole and historical maps especially has turned out to be the most popular type of data. Statistical data is another data type that has proved very popular. One of the most promising recently opened data set from the open government point of view is the agendas and minutes of the city council, city board and committees. These are distributed from the original data source through an application programming interface (API). The service is called Open Ahjo (<http://openhelsinki.hel.fi/files/>).

Beside the data journalism and visualisation, the development of different kind of applications based on open data tend to be very promising way of enhancing smart city development and interaction between the city government and the citizens. The application that shows the real time locations of snowplow cars is one example of this kind of application. The application that enables one to calculate his carbon dioxide emissions that caused by moving within the city by different mode of transport is another. Several applications that uses statistical data has been developed. It is expected that the developers will soon publish interesting applications based on Open Ahjo data, too.

Also the the data providers tends to be satisfied with the data opening. The use of government data has grown and many useful applications and services has developed, often with dispatch after data opened and without extra costs. City bureaus has got a

lot of good publicity for these services. Another gain that has turned out is the improved quality of data thanks the feedback from the data users.

Notwithstanding all good experiences the setting up the HRI open data service has been far from easy. There are many technical, legislative, administrative and also principled questions that had been resolved during the project. The most important lessons learned can be summarised as follows:

- The data providers should be committed to data opening. It is impossible to get the resources needed without the support of all levels of administration of the data providers.
- The data clearing house and distribution service should be developed with close interaction with data providers and end users. By that way it is possible to compose appropriate criteria and instructions for data producers and set up services that advance the use of data. This will also expedite getting applications based on data.
- It is advisable to start with data that is ready to open (i.e. meets the criteria of open data) and has demand from end users and developers.
- The georeferenced (maps and geocoded data) and statistical data has turned out to be popular as well as all kind of operational data that represents the actions and services of the city. The more detailed and online data the better for the developers.
- Developers prefer to use data from the original data source through an application programming interface (API). This model is more easy for data producers also in the long run though it requires more efforts at the beginning.
- Social media is an important mediator between the open data and end users. The information of open data and applications spreads most effectively via social media. Facebook proved to be the most important route to HRI web service. Sharing experiences enhances development effectively.

5. Conclusions

Open statistics and data have advanced the use of administrative data. Increasing use of open data has improved the quality and usability of that data. Visualizations and applications based on open data have made the actions and services of the city more understandable for citizens, advanced the transparency of the city administration and by that way enhanced citizens possibilities to participate city development.

Making administrative data open just begun. The city of Helsinki for example has around 1000 different information systems which includes plenty of information that can be opened and use more effectively. Still it is important to notice that the data only is not enough. In addition to this the active users who participate the smart city development is needed. Only then open data can become fuel of the smart city development

References

Caragliu, A., Del Bo, C. and Nijkamp, P. (2009) "Smart cities in Europe", *Serie Research Memoranda 0048*, VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics.

HRI web pages. (www.hri.fi), 30.4.2013.