D. 3.2.1.2.

Open data in Finland - Public sector perspectives on open data

Author: Pauliina Lehtonen

Confidentiality: Public

Date and status: 22.06.2011; Version 1.0

This work was supported by TEKES as part of the next Media programme of TIVIT (Finnish Strategic Centre for Science, Technology and Innovation in the field of ICT)

Participants	Name	Organisation
	Pauliina Lehtonen	**UTA, JRDC**
	Esa Sirkkunen	**UTA, JRDC**

next Media
www.nextmedia.fi
www.tivit.fi

Executive Summary

The report provides a brief overview of the current development in the field of open data in Finland. The focus is on recognising the starting points and recent questions that have emerged in the public sector within the theme of opening up data. The report is based on expert interviews and a selection of open data initiatives.

The inspiration for the national development in the public sector has increased partly due to the international interest in this issue. Especially, the visible development of this field in the UK and US has worked as an incentive for the Finnish initiatives. The recently executed guiding principle of the Finnish Council of State encourages public sector actors to independently identify the information resources that would be useful for wide public utilisation and re-use.

Until now, the trend of opening up the public sector's data has proceeded rather slowly in Finland. Due to the fragmented nature of the development, there was no coherent view of the data resources that would already be opened for public use. The geographically referenced data that utilises map interfaces and displays was most often mentioned in the interviews as a typical example of open data.

Commonly, the escalating discussion of the open data phenomenon is connected to an emerging information ecosystem. In this ecosystem, information use and processing may engender new initiatives and partnerships between public and private sectors. The potential of the open data is seen to facilitate the production of new knowledge and collection of profit by generating new ways of gathering, producing, visualising and analysing information.

Typically, the understanding that considers the value of information to arise from its active use is emphasised. Thus, opening up public data may provide opportunities to enrich information and encourage a more open democratic system and decision making.

According to interviews, the field of open data is still plagued by several questions and restricting factors that demand greater attention. Questions deriving from *legislation*, from *established work cultures* and questions that discuss the *practices on data use* were distinguishable during the interviews. For instance, the question of how to deal with decreasing revenue from the sale of data was raised when talking about some public sector actors. This question concerned organisations such as Statistics Finland, the Population Register Centre and the Finnish Meteorological Institute, all of which produce, for instance, statistical or geographical data for sale.

From the point of view of the media, the interviewees considered the development of opening up data resources to provide opportunities for more information-efficient, transparent and trustworthy journalism. The role of journalism was also seen as a mediator between decision-makers and citizens.

This report also provides a list of 60 case examples that aim to illustrate the different genres of utilising data. The examples are divided into five groups: data catalogues, the visualisation and illustration of political information, public participation, solving everyday problems and citizen surveillance and transparency of public sector and politics.

Table of Contents

Table of Figures

1. Introduction

This report provides a brief overview of the current utilisation of open data in Finland, especially in the public sector. Attention is paid to the starting points and recent questions that have emerged within this theme. The report addresses possibilities and challenges that the increasing development of opening up public sector data might encourage. The research data for this report was gathered by expert interviews that were conducted in the spring of 2011 and by mapping the field by going through some relevant and current examples that have experimented with information that can be categorised as open data.

The number of the interviews is limited, but they depict a representative sample of the field of active actors within open data. The interviewees have recently been working on these topics that have related to questions of opening up, utilising, producing, visualising or interpreting information, thus having links to open data discussions. The aim of this short study is to build a brief overview of the Finnish situation and to develop ideas and initiatives of the possible public and private sector partnership in this field.

The interviewed experts present the following actors: Ministry of Transport and Communications, Ministry of Justice (OSY-project), City of Helsinki (talous- ja suunnittelukeskuksen tietotekniikkaosasto), Forum Virium (Helsinki Region Infoshare), Helsingin Sanomat, Fillarikanava and Vaalilupausarkisto.

2. The background of the development of open data

Lately, the interest in opening the public sector's information resources has increased massively. The topic has gained attention from various parts of society, such as from the government, business sector and active civic actors such as individual citizens and various developer communities or collectives that gather together activists who are interested in the issue of open data.

At the governmental level, the Ubiquitous Information Society Advisory Board (Arjen tietoyhteiskunnan neuvottelukunta) that works under the Ministry of Transport and Communications has raised the question of the accessibility of public information (e.g. the Ubiquitous Information Society Advisory Board 2008). In the new, national information society strategy, the role of information as a part of our basic infrastructure has been named as one of the strategy's focal points (Ministry of Transport and Communications 2010).[1] Also, this year, the Finnish Council of State executed a guiding principle that public sector information should be freely available.[2]

The inspiration for national interest has been the international development that has increased the on-going discussion in Finland. At the EU level, questions around opening up

[1] See also Paukku 2009 and Turkki 2009.
[2] See, http://www.lvm.fi/c/document_library/get_file?folderId=1591058&name=DLFE-11755.pdf.

administrative data were touched upon already in 2003 when the Public Sector Information directive (003/98/EY) was legislated. This public sector information (PSI) directive sets boundaries and guidelines for the commercial use of governmental information (see Poikola ym. 2010). The directive aims at increasing the availability of public information with low expenses. In addition to the PSI directive, the EU's Inspire directive (2007/2/EY) focuses on the use and availability of geographically referenced data (GIS data). The directive was executed in 2007, and it aims at developing the co-cooperativeness of national geographic information systems and their infrastructure.[3] National legislation that carries out the Inspire directive directs the opening up of geographically referenced data in Finland.[4] Also, the EU's digital strategy forms guidelines for the use of public sector information by focusing especially on the PSI and Inspire directives.

Besides the EU-level guidance, a strong influence comes from United States, especially within the public sector. The government of Barack Obama has paid a lot of attention to the issue. On the vanguard of opening data is also Great Britain, where the mandate for opening national information has also been executed at the political level. In Finland, some politicians have recently started to show interest towards opening public data. Most likely, the interest at the political level is about to strengthen in future.

All these examples point to the tendency of opening up the information resources of the public sector for free utilisation. In addition to the public sector, some media companies and individual journalists have started to show interest in the possibilities that emerge from open data, especially from the point of view of data-driven journalism[5]. However, concrete examples in the media field in Finland are still rather scarce. One initiative is the HS Open project of the leading national news paper, *Helsingin Sanomat* (HS), whose HS Open workshops are targeted at different actors interested in the open data phenomenon.[6] The newspaper has, for instance, opened the data of its electoral machine for wider public use. Also, the public broadcasting company YLE is starting to pay attention to the issue of open data.[7]

In addition to the public sector and the media, the themes of open data are taken forth in various developer communities and by individual citizens. The topic seems to gather interested civic actors and loose collectives (see part 4.3. "Key actors") together to discuss and experiment with issues that relate to the free availability of information, technological development and information visualising and processing. At the background of this ideology, democratic ideas exist for producing solutions to societal problems by making information freely available that can be later utilised as a source of societal innovations.

Universities and research organisations have also shown active interest towards open data. For instance, Aalto University is currently starting an "Open Aalto" initiative to bring together different actors, such as software developers and public sector actors. The issues of a semantic web and linked open data have been employed in the research projects of Aalto University.

[3] http://www.mmm.fi/fi/index/etusivu/maanmittaus_paikkatiedot/paikkatietojenyhteiskaytto/inspire.html.
[4] The law on geographically referenced data (421/2009) and regulation (725/2009).
[5] See http://datadrivenjournalism.net/.
[6] http://blogit.hs.fi/hsnext/aihe/hs-open
[7] With its open database of "Elävä arkisto", YLE can be seen to approach the theme of open data. See also YLE's application that provides data concerning weapon legislation in Finland, http://beta.yle.fi/aselaki/#kommentit. See also http://blogit.yle.fi/kehitys-kehittyy/yle-avaa-dataa.

For instance, “FinnONTO” projects[8] aim at contributing to the development of open infrastructure. In addition, open data was one of the research interests in the joint research project one of the joint research projects of VTT, Aalto University, University of Jyväskylä and University of Tampere. The project “Social media for citizens and public sector collaboration” (Somus) pondered the questions of openness of governmental data (Näkki et al. 2011). Also, the Finnish Centre for Open Source Solutions (COSS) is among the active agents in the fields of open source and open data. This national development agency just recently organised a workshop on open data[9] in Tampere.

3. Defining the concept of open data

Open data has quickly become a widely used concept, but the meaning of which remains rather vague. The term has been adopted for use on various occasions and by diverse actors without a clear definition of what it is used to refer to.[10] When talking about the definition in the expert interviews, two dimensions could be recognised: (1) the origin of information and features describing the openness, and (2) the practices of maintaining the information.

The origin of information
In the interviews, the talk of open data was connected to a fact that to be regarded as open, information should be *easy to find* and *to access*. It should also be available as a whole in the sense that it forms *a coherent entity*. It also needs to consist of raw data in a form that is *commonly identifiable*, which is often used to refer to data that is in a *machine-readable* form. As an example of a commonly identifiable form of data, one interviewee mentioned GoogleMaps, which has heavily increased map-based displays of local information. The feature of a known format was seen to help in increasing the general accessibility of the data.

The practices of maintaining information
When discussing the definition, it was often emphasised that information should be *free* and *legal* to use. This brings up questions regarding, for instance, personal information or copyright, i.e. the kind of information that can be regarded as too sensitive to be allowed free access to. Except sensitive data, there should be *no restrictions on the use of data* or *licenses* connected to it. The licenses or the rights attached to the use of data should be *equal, original and regularly updated*. The information should be *regularly updated* to secure the relevancy of the data. Also, it was brought up that *intelligibility* and *ease of use* are two key features of information.

Often, the term "open data" is used to refer to public sector information that is produced with taxpayers’ money. However, this definition is too broad, as, for instance, information that provides details of personal privacy cannot be made freely and openly available. Poikola et al.

[8] http://www.seco.tkk.fi/projects/finnonto/, http://www.seco.tkk.fi/. See also an example of semantic web application “Kulttuurisampo,” which is a national communal publishing tool for institutional memory organisations and citizens, http://www.kulttuurisampo.fi/.
[9] See http://www.coss.fi/avoimendatantalkoot.
[10] For instance, it can be discussed if Elävä Arkisto of YLE can be taken as an example of opening up data. In this case, the opened data consists of television and radio programmes and not, for example, statistical data that is often associated with the concept of open data because statistics are machine-readable. Another example would be the web broadcasts of the Helsinki City Council, which makes the processes of decision making more transparent, but are not in a machine-readable format.

(2010, 34) have distinguished criteria for open data based on the definition made by the Open Knowledge Foundation[11]. Poikola et al. categorise their definition into features that are considered the most crucial when defining openness:

- data has to be commonly available to be found,
- data has to form a coherent entity,
- licenses to use data need to be equal,
- data needs to be original and current,
- data needs to be legal and free to re-use,
- data needs to be free,
- data needs to be in a machine-readable format,
- data needs to be in an open format, and
- data needs to be described and documented thoroughly and clearly.

The mentioned definitions of open data stress mostly the ways how data is provided and acquired and, also, how data is used. However, as the definitions concentrate on its use, they pay less attention to the content of the data.

4. Current practices

4.1. Starting points and current situation

In the interviews conducted for this overview, making information better available, thereby increasing possibilities to utilise it, was one of the main aims guiding the public sector's initiatives. The general benefits that affect the background of making administrative data public can be divided into four groups. Opening up data is considered important because it can help in making *processes, services and communication of information* more *efficient* inside the government; it is meaningful from the point of view of *democracy and participation*, especially by aiming to develop the openness and transparency of the public sector; information is considered as a material that lays *a foundation for business* and innovation activities, which, hence, benefits society; and information creates *resources for research*. Opening up public data brings material for research, which can then produce new information and new knowledge for the utilisation of the public sector.

More specifically, the possibilities of the development of open data relate to the possibilities of *creating diverse ways of using information*. These new practices can

(1) produce new business possibilities based on the re-use and utilisation of information and data in new ways,
(2) necessitate better utilisation and re-use of current information, for instance by enabling tailored information or services for specialised target groups,

[11] One of the general definitions for open data is the one of the British Open Knowledge Foundation. It has coined a thorough definition for open, accessible and re-usable data. See http://www.opendefinition.org/okd/.

(3) help to concretise vast amounts of (abstract) information by creating illustrations and visualisations and utilise other versatile means of presenting information (e.g. tax tree, see Näkki et al. 2011), and
(4) be of use in discussing societal problems and issues (e.g. using visualisation to support public discussion and decision making).

Thus far, the issue of opening up public data has proceeded rather slowly in Finland, and none of the interviewees produced a clear picture of the general situation in the public sector. Due to the fragmented nature of the development of open data, there was no coherent view of the data resources that would already be open for public use and no clear view of the data that the public sector would be planning to open. The geographically referenced data that utilises map interfaces or displays was most often mentioned as data that has already been opened. This kind of data consists of, for example, location and service details (e.g. a service map of Helsinki) or navigation and public transport services (e.g. a journey planner). Other typical examples of open data that were mentioned were, for instance, various public registers, information on public purchases and public projects and information on decision making, such as voting behaviour of politicians[12].

An example of the emerging activities within open data is internationally popular data catalogues that have started to gather interest in Finland. Data catalogues are portals that gather information from different datasets and databases.[13] A current Finnish example is suomi.fi,[14] which provides a selection of public datasets of governmental services and information resources. Examples of possible datasets that should be opened according to the interviewees were, for instance, public revenue and expenditure, public purchases, information on decision-making procedures and various societal statistics and, in general, the various other data that the public sector possesses.

However, traditional working practices of the public sector have often restrained the attempts to open up data. The public sector's logic towards information sharing has rested on the idea of controlling and possessing information instead of letting it be freely processed and utilised in society. The general atmosphere towards opening data resources and making them accessible for the public has lacked enthusiasm as the old bureaucratic attitudes are still often dominant. Now, the trend of valuing data openings has raised questions, especially among those public sector actors whose activity is based on selling their data (such as Statistics Finland, National Land Survey of Finland, Finnish Meteorological Institute, Population Register Centre).[15]

In the public sector, one crucial question within open data development seems to be the lack of clearly defined responsibilities. At the moment, according to the interviews, it is not clearly

[12] See also http://www.hri.fi/fi/ for the open public data of the Helsinki region.
[13] See, for instance, http://data.gov.uk/, http://data.london.gov.uk/, http://ckan.net/, http://www.datasf.org/.
[14] http://www.suomi.fi/suomifi/tyohuone/yhteiset_palvelut/avoin_data/index.html. See also a service map of Finnish administration ("valtionhallinnon palvelukartta"), http://www.suomi.fi/suomifi/suomi/palvelukartta/index.html.
[15] A working group in the Ministry of Finance is currently working on this question. The group is co-operating with the Ministry of Transport and Communications and with Ministry of Education and Culture who have also set up working groups to consider the opening up of public information resources (see the guiding principle of the Council of State 2010).

defined in the administration who in the organisational structure is responsible for following and guiding this development.

The guiding initiative of Council of State encourages the public sector to independently identify the kind of data that would be useful to open. Instead of centralised guidance, the government follows the British example in emphasising the active approach to data openings. In Britain, the government has, for instance, asked for suggestions and requests from the public for datasets that should be made available (e.g. http://data.gov.uk/blog/call-for-developers-to-request-datasets, see also http://www.data.gov/suggestdataset).

Based on the interviews and current examples, it seems that the development proceeds under the attention of certain enthusiastic actors. Also governmental guidance, such as the recent principle that was executed, sets frames for the national development. But the concrete development work seems currently to rest on a group of active actors, be it inside the public sector or the media or initiatives taken by individual citizens and civic entrepreneurs. These actors stress that the value of information comes from its availability and use, instead of restricting and controlling the ways in which information is being used.

4.2. Key actors

Based on the interviews, the actors who seem to have a relevant position in the discussions and development related to open data can be roughly divided into three groups: the owners of information, the processors of information and the interpreters of information. Often, the public sector is self-evidently regarded as the owner of information, but some of the interviewees also pointed out private sector actors to be as relevant agents and information owners as the public sector. Here, the problem often seems to be that there is no official and functioning system to map the field of open data within private actors. There is no checking of data that the private sector possesses and produces, which makes it difficult to ensure the origin of information or if the information is up to date.

The line between information owners and processors (usually software developers, such as small companies) and interpreters of information (for example, citizens and journalists) can be flexible. For instance, in the interviews, media companies were grouped into interpreters, but they can also be information owners and openers. Actually, it was suggested that the media could also open its data for wider re-use. Similarly, citizen-based data can prove to be important for the public sector, which then acts as an interpreter of information whose “owners” or producers are citizens.

Currently, according to the interviews, the main public sector actors are ministries, civil service departments and other governmental actors, such as the Ministry of Transport and Communications, Ministry of Finance, Ministry of Employment and the Economy, Ministry of Justice, Ministry of Education and Culture, State Treasury (Valtiokonttori), Statistics Finland (Tilastokeskus), Finnish Transport Safety Agency (Trafi), Population Register Centre (Väestörekisterikeskus), National Land Survey of Finland (Maanmittauslaitos) and Parliament of Finland. On the municipal level, some municipalities and cities have started to take action on the question of open data. For instance, the city of Helsinki is among the active cities in Finland. For instance, Forum Virium is currently implementing the “Helsinki region infoshare” development project of the cities in the Helsinki area (Espoo, Helsinki, Vantaa, Kauniainen).

Also, libraries have shown interest in participating in this development[16]. Other important actors are research and development organisations, such as Tekes, The Academy of Finland, Finnish universities (e.g. Aalto University, University of Tampere, University of Jyväskylä) and VTT. The interviewees also mentioned private sector actors, especially in the IT field, as interested in participating in this development. Organisations such as Gemilo, Flo Apps and Tieto were mentioned. Sanoma and YLE seem to be interested in the media field at the moment. Also, developer communities/collectives, such as the developer community that is formed around the "Osallistumisympäristö" project of the Ministry of Justice, the "Avoin demokratia" network and "Verkkodemokratiaseura", were brought up. For instance, Verkkodemokratiaseura is organising an Apps for democracy competition[17] in 2011.

4.3. Questions and restricting factors

The interviews addressed some questions and restrictions concerning the free utilisation of public sector information. They are divided here into three groups:

1) Questions of legislation that set frames for the information use:
- The adjustment of the principle of openness and legislative guidance on information
- The questions of sensitive data and individual privacy

Legislative questions concerned the adjustment of legislation that defines information use, sale and publicness to the principles and attempts of opening up the public sector's databases. This question has extreme relevance, especially in the legislation that forms the basis for the payment politics on public data. Attention must be paid to those actors who produce data for sale. They would no longer derive revenue if all data would be made open (e.g. Statistics Finland, Population Register Centre, Finnish Meteorological Institute).

Other legislative restrictions are based on privacy legislation, copyright legislation and legislation regarding the public nature of public sector data. For example, the legislation on the public nature of public sector data does not include definitions of the re-use of data. [18] It is also important to pay attention to politically sensitive data, for example on data that includes information of issues of like national defence. The legislation needs to be acknowledged when opening databases, especially if the opened data could infringe the limits of privacy.

2) Questions of work culture practices:
- Organisational responsibilities and tasks

[16] See, for example, http://labs.kirjastot.fi/. See also, http://www.digitalkoot.fi/fi/splash
[17] The competition is now organised for the third time. It was first started by the Somus research project as an initiative to produce new ideas on what could be done with governmental data (Näkki et al. 2011, 74–75).
[18] Also, it has been noticed that despite the guiding legislation, the authorities continuously ignore the legislation. See http://www.mtv3.fi/uutiset/kotimaa.shtml/2010/12/1244200/tutkija-viranomaisten-salailun-takana-tiedonpuute-ja-pelko.

- Established practices in the public sector
- Indicators and instruments for evaluation
- Effects on transparency

The question of organisational responsibilities noted that issues concerning data openings are not clearly defined in the public sector. Issues or tasks dealing with open data may be transferred from one unit or person to another because roles and tasks often stay undefined. Also, interviewees noted the lack of resources that are directed to the development of open data in the public sector.

The established practices in the public sector have restricted data openings, as the change of work culture attitudes takes time. Thus far, the public sector has more often relied on controlling information delivery than freely opening information. Working environments have been rather closed. Also, governmental processes proceed slowly, which can delay providing access to datasets. In addition, the interviewees mentioned hesitation that relates to the question of the incompleteness of information, which – if information would be freely available and re-used – might produce false interpretations and lead to misrepresentation of information. The hesitation was based on the fear that information would be used and presented in an incorrect way that would encourage false interpretation.

One important question addresses the issue of how to choose and create indicators and instruments to analyse the effects of opening data and to evaluate organisational activities. As mentioned before, this issue is relevant especially among those public sector actors who produce information for sale. Finally, the question of transparency was brought up. The interviewees asked if openness would decrease transparency and increase the covering up because openness might produce unwanted results, unwanted development and unwanted publicness.

3) Questions of processing the data:

- Developing common principles for guiding the data processing
- Data quality and verification
- Interpretation and evaluation of services and the authenticity of data

The guiding principles for dealing with the use of governmental data are still shaping up. The question as to whether data should be available and for use only as entities or also as segments was introduced in relation to the issue of privacy. The interviewees pondered if combining and adjusting different datasets or parts of them would generate privacy threats. How do we guarantee personal privacy if different databases that possess personal information are interlinked? Or, could vaster societal effects be generated by combining datasets? For instance, could opening and linking spatial data generate vaster effects on neighbourhoods' development or their public image or popularity (such as, if public information that presents ethnic origin, apartment prices, crimes, etc. of a certain neighbourhood would be linked together).[19]

[19] For instance, http://www.suburbantrends.com.au/

Also, the quality and reliability of data and data presentations were addressed; who guarantees that data is of good quality, and that it is verified and definite? This also stressed the maintenance of data; who is responsible for maintaining and updating data? Or, what happens when all data is free for everybody to process and experiment with? The questions of resources and expertise also concern the quality and maintenance of datasets, for instance if in the processing phase there are enough knowledgeable technological and other resources to process and interpret the data. One solution to the question of verifying data could be data clearinghouse/warehouse-type services.

When talking about the data quality, the harmonisation of data was raised to pay attention to various databases and systems that the public sector utilises. How do we harmonise data from various sources and from various information systems? It was also noted that in data presentation and visualization, it is important to acknowledge the limits and restrictions regarding the means of presentation. For instance, different visual choices vary and are illustrative in print and on the Net.

Finally, a critical view on the developed services was also called for. The trend of openness necessitates a reasonable amount of consideration from the users and interpreters of data; they need new capabilities to evaluate how the information in various services and applications is presented, organised and interpreted. Also, evaluating the origins of information and the authenticity of information was mentioned as an important question.

5. The private and public sector partnership

The guiding principle of the Council of State (2011) distinguishes three major areas that are paid special attention to when proceeding with the opening up of public data:

1) Politics of information
   -discussing payments and licenses to clarify legislation on public sector information, such as payment politics and rights attached to the use of data

2) Building of open information infrastructure
   -building supporting services, general practices, structures

3) Support and improvement of software and service development
   -creating environments to support creativity and innovativeness
   -advancing possibilities for research to produce information of the effects of the open data phenomenon for future guidelines and development

In these main development areas, the role of private-public partnership can be generated from the different roles connected to information use that the open data development can enable. In the future, profit comes from not only information possessing, but more likely from new roles and tasks that emerge from the information processing. The commercialisation of data/information is believed to increase; such as using data to collect profit by developing new ways and means of processing, analysing, producing, visualising or interpreting information.

Discussions around the issue focus largely on the idea of a new kind of information ecosystem where traditional roles around information ownership and information processing get new dimensions. According to the interviews, the public sector actors recognised the importance of private actors, namely from the point of view of a developing information ecosystem. They saw the potential of this ecosystem in creating innovative ideas and ways for the private sector to function in co-operation with the public sector, for example in different supporting or mediating tasks. Also, they believed the new ecosystem would transfer both economic and intellectual benefits from the private sector back to the public sector. It was this kind of feedback that some of the interviewees estimated to be the most prominent benefit of opening up data resources.

Nearly all of the interviewees shared a vision that business development around open data would be useful for the public sector, and typically for the national economy. They envisioned the business sector to produce new ideas and means for a more effective utilisation of societal information, which would then support the functions of society and national growth by producing taxable income.

One prospective trend within open data development was believed to be the international standardisation of data. One current example that came up was the journey planner of the Helsinki region that is already utilised in Spain. The data system of the journey planner has been adopted and modified to serve local Spanish geographies. This kind of standardisation, for instance at the EU level, was assumed to create new possibilities for public and private sectors.

5.1. The role of media

In the interviews, the role that was most often considered for media was one of a mediator or a monitoring actor in society. Media was situated to a position where it gathers and filters diverse information and then winnows out relevant information that needs to be made public. It was seen as a relevant actor in filtering the important content, making it public and organising it in a way that would make the information understandable. Generally, this kind of role as “a sense-maker” resembles the traditional working practices of (investigative) journalism. In the media field, a developing interest towards data-driven journalism can be distinguished. Data-driven journalism[20] is seen to support the functions of journalism in the new information environment.

[20] Wikipedia defines data-driven journalism as follows: “Data-driven journalism (also known as datajournalism or data journalism) is a journalistic process based on analyzing and filtering large data sets for the purpose of creating a new story. Unlike database journalism, which is mostly used for investigative reporting, data-driven journalism deals with open data that is freely available online and analyzed with open source tools. Another distinction is that data-driven journalism strives to reach new levels of service for the public, helping consumers, managers, politicians to understand patterns and make decisions based on the findings. As such, data driven journalism might help to put journalists into a role relevant for society in a new way.” (Wikipedia, checked 10.5.2011)

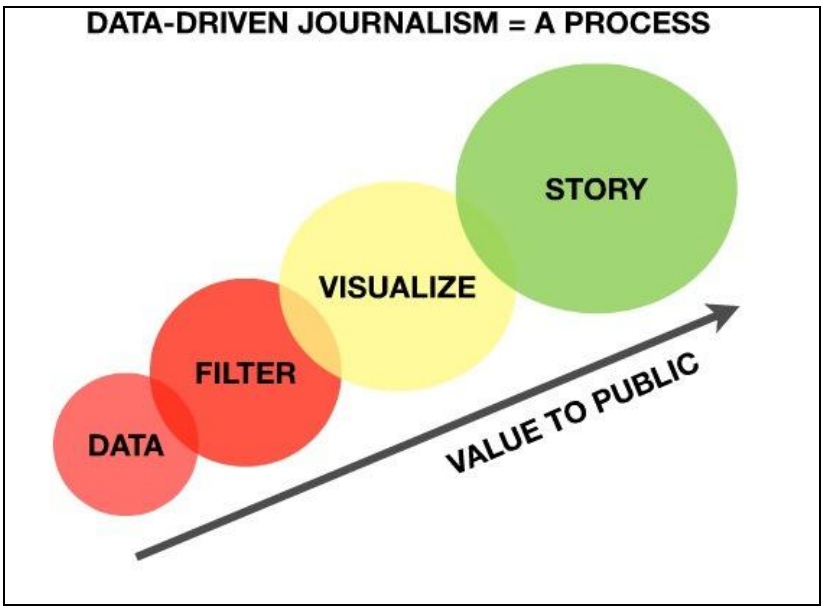

Figure 1. A working process of data-driven journalism (Lorentz 2010).

In current discussions, the increased interest in data-driven journalism is connected to the changes experienced in journalistic work. Open data development can provide new opportunities for journalism to cope with the current changes, such as to strengthen its role in the strongly user-driven environment (e.g. Lorentz 2010).

Journalism as a mediator was seen in the interviews from two angles. First, this development was considered to provide opportunities for a more *information efficient and transparent journalism* because it widens the information resources that journalists can utilise. Open data was believed to improve the general trust towards journalism because the source material of news stories would be public. In general, the main points within journalism were linked in the interviews to journalistic practices, such as being able to produce better and more reliable stories and to enhance the information acquisition of journalists.

Second, according to the interviews, the role of journalism as *a mediator between administration and citizens* was suggested to increase in the future. This resembles a general development trend that is already visible in the on-going discussions of hyperlocal journalism. The insight from hyperlocal journalism touches upon the media's role in encouraging and empowering active citizenship. Both angles (the one of journalistic practices focusing on developing more transparent journalism and the other of the role of journalism as a mediator between citizens and administration) that are evident in the interviews can be seen in the idea of improving societal transparency in journalism, administration and decision-making.

The open data development may provide resources for the functions of hyperlocal journalism. In general, the importance of getting access to vaster data sources is connected to everyday news work. Open data sources might not provide concrete news directly, but they could provide interesting topics, findings or questions that can be processed and formulated into

news.[21] For instance, local services such as “Fix my street” or “Fillarikanava” might generate some interesting questions or observations that could lead to local news pieces.

The interviewees appreciated the emerging co-operation between the public sector and the media. They viewed the co-operative relationship as producing benefits for both parties. For example, from the point of view of hyperlocal journalism, media could develop a service that would collect local citizen-based information (for instance, on a map display), filter the relevant content (e.g. the most frequently mentioned problems) and then forward it for use by the public sector. The public sector could then utilise the information for fixing, for instance, local flaws such as broken traffic lights or building zebra crossings. This kind of service could bring benefit to the media by producing possible news stories and supporting its position as a trustworthy public actor. Also, the public sector could perhaps provide some datasets that would present local data for the media to utilise as background material for news articles or for service development. In addition, one of the interviewees proposed that media could start to collect and maintain open datasets from private actors, which are currently scattered and not necessarily actively updated.

In addition to the effects on journalistic practices and roles, the change in the information environment was suggested to possibly transform the basic concept and the form of news, as one of the interviewees proposed.[22] For instance, could news be a simple visual presentation that would be frequently updated?

Open data has the potential to develop a broader view for journalism, which is addressed in the role of functioning as a mediating institution in society. This would also serve the view of hyperlocal journalism by reserving it for a role in developing local services. Thereby, open data enables the development of tailored, focused services and news for different target audiences (see the case examples in this report). The possibilities of producing tailored services are unlimited in different fields in society, such as in social services (services for families, pensioners, students, etc.), in planning (for residents, associations, NGOs), in leisure facilities (for culture, sport activities, etc.) or in tourism.

6. Conclusions: increasing transparency and innovation in the new ecosystem

In general, the discussion of the open data phenomenon relates to the current development of a new information ecosystem that is argued to be emerging (see, e.g., Aitamurto 2011; Mitchell 2010; also Benkler 2006). In this ecosystem, the collaboration between the public and private sectors is believed to facilitate new services for information use. In the background, the assumption of increasing broader societal transparency is also distinguishable. However, according to the interviews, the development of open data is confronted with restrictions and problems stemming from (1) legislation, (2) established work cultures and practices and (3) from questions that address practices on data use and processing.

[21] See, for instance, http://open.blogs.nytimes.com/2009/01/08/introducing-the-congress-api/.
[22] See, for instance, http://projects.washingtonpost.com/fallen/ as an example of news that updates regularly online.

The development of open data is so recent that estimations about future directions or developments are difficult to make. Many of the interviewees emphasised the importance of pilot projects. They can generate further development by showing how data could be used and for what purposes, and what would be useful and functioning applications and services.

Generally, from an organisational viewpoint, the aim of providing assets for organisations that are making their data open has been one of the goals in the development of open data. However, the advantages of opening data are not always immediately visible or do not necessarily emerge directly in organisations. More often, the advantages or the profits for organisations are indirect and become concrete in the functions of the information ecosystem that develops around the open data (Poikola ym. 2010, 31). These indirect profits can affect the activities of organisations by making them more well-known or by increasing their popularity or customer satisfaction.

From the point of view of the media, the possibilities of the open data phenomenon were focused on increasing the reliability of journalism in the interviews. Open data was believed, at first, to help in producing better and more transparent news stories, and second, in strengthening the mediating or monitoring role that journalism plays in society. Both of these functions were figured to fuel the general trust in journalism and media.

According to the interviews, the development of open data revolves around the idea of producing new knowledge and collecting profit by developing new ways of gathering, producing, visualising and analysing information. The interviewees called for better utilisation of information, as much of the current information that is produced and possessed, especially in the public sector, still remains largely un-used. The aim of the public-private partnership focuses on producing benefits for both parties, which is the result of utilising information resources more efficiently than before. Here, features of crowdsourcing are also visible, as the loose social network around open data developers in Finland share similar interests. Already now, they have a substantial role in enriching the open data by creating additional modules and services on top of public sector data.

The figure on the next page summarises the general phases that are related to the process of data processing. It also illustrates the dimensions that form currently recognisable tasks relating to information utilisation within the new information ecosystem. In the data utilisation process, the first phase is data filtering/mining that goes through raw data that is derived from databases. In the phase of organising, data is arranged to highlight the important content. Visualisation aims to illustrate information. Now data is more processed than in the first phase when it is collected from datasets. In the interpretation and production phase, data is even more cultivated than in the previous phases. Often, in this phase, data is processed, presented and visualised to serve certain aims or functions, for instance, to produce information for the grounds of decision making.

The means of data utilisation	Data filtering/ Data mining	Data organising	Data visualisation	Data interpretation and production
Purpose	Accessing data and going through datasets; seeking out and filtering the essential info. in the data	To arrange data in a clear form; to organise data in a way that highlights the important content, such as the grounds for public decisions	To support the interpretation of data by visualising it in a comprehensible manner, such as illustrating relations by utilising acreages	Creating new knowledge for problem solving, innovation, decision making, etc.
Example	http://data.gov.uk/	http://beta.vaalilupausarkisto.fi/	http://wheredoesmymoneygo.org/	http://kerrokartalla.hel.fi/

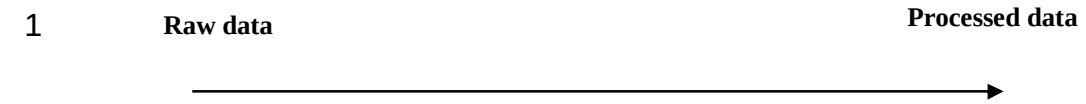

Figure 2. The process of data utilisation.

The prevailing tendency that assumes that the open data phenomenon will only produce positive outcomes must not go un-noticed. This extremely positive view seems to be dominant in the current discussions in Finland. However, some criticism has been internationally raised due to viewpoints that open data would produce firmly positive consequences. This critique points out that the shifts in information, power and responsibilities are still difficult to foresee (Kaplan 2010). Also, the issue of commercialisation of data that is first produced with taxpayer money and then made public for business interests seems to remain un-problematised within the current discussions.

It needs also to be noted that the question of illustrating and making data/information more usable and more understandable is not a totally new issue within discussions related to information and communication technology. Similar topics have been previously employed in discussions, for example in the field of (e-) democracy and citizen participation (e.g. Coleman & Blumber 2009; Hacker & van Dijk 2000; Papacharissi 2002) or within discussions in the area of the semantic web[23]. These discussions have broached themes similar to the current open data debate. Although, in the discussions on (e-) democracy and citizen participation, information has been considered from a broader perspective; the aim has been to search for ways and tools for communicating and presenting information to provide more open and participatory decision making. However, in the discussions going on about the open data, the focus is more on defining the origin or format of the data and developing the norms that guide the data use.

[23] See for instance http://www.seco.tkk.fi/projects/finnonto/index.fi.php, and http://www.digitoday.fi/page.php?page_id=66&news_id=200411119

References

Aitamurto, Tanja (2011) Kehityssuuntia Yhdysvalloissa: sisältötehtaat, tutkivat tukut ja yrittäjyys. Journalismikritiikin vuosikirja. Media & Viestintä, 34 (1), 37–45.
Benkler, Yochai (2006) The Wealth of Networks: How Social Production Transforms Markets and Freedom. New Haven and London: Yale University Press.
Coleman, Stephen & Blumler, Jay G. (2009) The Internet and Democratic Citizenship. Theory, practice and policy. Cambridge: Cambridge University Press.
Hacker, Kenneth L. & van Dijk, Jan (2000) Digital democracy: issues of theory and practice. Sage.
Kaplan, Daniel (2011) Open public data: Then what? Part 1. http://blog.okfn.org/2011/01/28/open-public-data-then-what-part-1/
Lorentz, Mirko (2010) Data driven journalism. Five Ws and (and one H). http://datadrivenjournalism.net/slideshows.html
Ministry of Transport and Communications (2010) Tuottava ja uudistuva Suomi - Digitaalinen agenda vuosille 2011–2020. http://www.lvm.fi/c/document_library/get_file?folderId=1551286&name=DLFE-11771.pdf&title=Tuottava%20ja%20uudistuva%20Suomi.%20Digitaalinen%20agenda%202011-2020
Mitchell, Bill (2010) Rusbridger: Openness, Collaboration Key to New Information Ecosystem. Poynter.org. http://www.poynter.org/latest-news/106389/rusbridger-openness-collaboration-key-to-new-information-ecosystem/
Näkki, Pirjo; Bäck, Asta; Ropponen, Teemu; Kronqvist, Juha; Hintikka, Kari A.; Harju, Auli; Pöyhtäri, Reeta & Kola, Petri (2011) Social media for citizen participation. Report on the Somus project. VTT Publications 755.
Paukku, Paul (2009) Julkiset sähköiset palvelut/kehittämisen pullonkaulat. Selvitysmiehen raportti 12.6.2009. http://www.lvm.fi/c/document_library/get_file?folderId=191434&name=DLFE-7917.pdf&title=Selvitysmiehen%20raportti.%20Julkiset%20s%C3%A4hk%C3%B6iset%20palvelut.%20Kehitt%C3%A4misen%20pullonkaulat
Papacharissi, Zizi (2002) The Virtual Sphere: The Internet as the Public Sphere. New Media & Society, 4(1), 5–23.
Poikola, Antti; Kola, Petri; Hintikka, Kari A. (2010) Julkinen data. Johdatus tietovarantojen avaamiseen. Liikenne- ja viestintäministeriö.
The Ubiquitous Information Society Advisory Board (2008) Ubiquitous Information Society. Action Programme 2008–2011. http://www.arjentietoyhteiskunta.fi/files/73/Esite_englanniksi.pdf
Turkki, Teppo (2009) Nykyaikaa etsimässä. Suomen digitaalinen tulevaisuus. EVA. Taloustieto oy.
Valtioneuvoston periaatepäätös julkisen sektorin digitaalisten tietoaineistojen saatavuuden parantamisesta ja uudelleenkäytön edistämisestä (2011) http://www.lvm.fi/c/document_library/get_file?folderId=1591058&name=DLFE-11755.pdf

Appendixes

Appendix I: Case examples

The following list gives examples of services and applications that utilise various data, which is often derived from the public sector. The listing is by no means extensive, as the amount of data applications and services is continuously mounting. Rather, the aim is to give a brief overview of the different genres that are recognised with the use of (open) data. Not all of the examples can be regarded as open data, but they are included in the list to produce insight into the possible functions of how data could be utilised to benefit society and societal purposes, such as the interaction between citizens and public administration. The genres are partly overlapping, although presented here in separate categories. The 60 listed examples present both international and national initiatives.

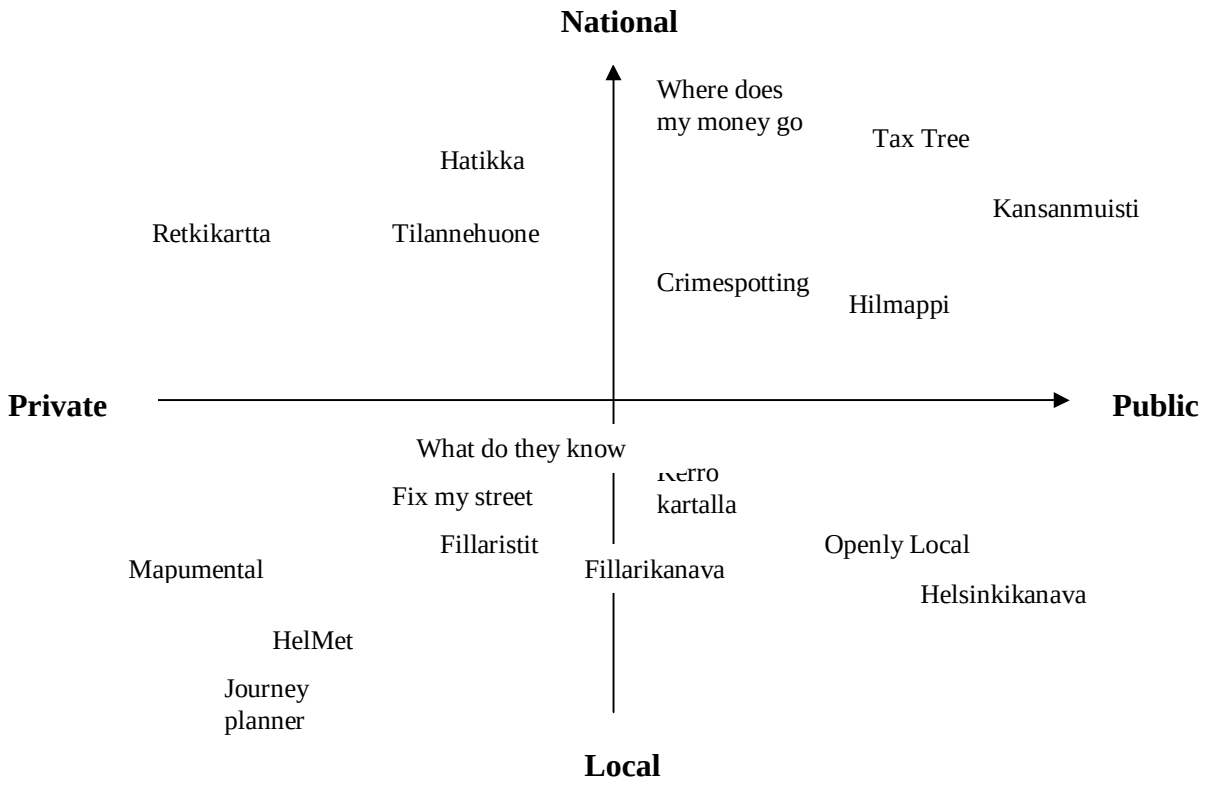

Figure 3. The framework of the open data examples

In the figure above, some of the introduced examples have been situated loosely on two axes. This is a modest attempt to visualise and illustrate the aspects of private-public and local-national that are recognised in the initiatives, and show how the examples can be located based on their functions and levels in this kind of framework. For instance, Fillaristit service provides information that is useful for local cyclists for planning their cycling routes in daily life. Or, Helsinkikanava makes it possible to follow the local decision making through web broadcasts.

Data catalogues

Data catalogues collect lists of public datasets that are free for re-use, analysis and processing.
- Suomi.fi data catalogue collects open data from Finnish public sector; see data.suomi.fi and http://www.suomi.fi/suomifi/tyohuone/yhteiset_palvelut/avoin_data/

- Open.gov.fi collects datasets and ideas for the utilisation of information; http://opengov.fi/

- Helsinki region infoshare provides public data from the Helsinki region; http://www.hri.fi/fi/data-haku/

- San Francisco data catalogue; http://www.datasf.org/

- London data catalogue; http://data.london.gov.uk/

- US data catalogue; http://www.data.gov/

- UK data catalogue; http://data.gov.uk/

- Australia data catalogue; http://data.gov.au/

- Open knowledge foundation's data catalogue; http://ckan.net/

- World Map of Open Government Data Initiatives shows those public and private initiatives on a map, which provides open government data or running challenges. http://maps.google.com/maps/ms?ie=UTF8&hl=de&msa=0&msid=105833408128032902805.00048bfbba4ecb314e822&ll=22.755921,-86.660156&spn=111.015734,270.527344&z=3&iwloc=00048c867b5e6a0a59178

- Infochimps; http://www.infochimps.com/

- Junar; http://www.junar.com/portal/HomeManager/actionQuery

The visualisation and illustration of political information

The services and applications aimed at making (governmental) data more understandable and clearer by utilising illustrations and visualisations. Visualisation can also facilitate the filtering of the essential information.

- "Puoluekenttä" application is based on the iniatiatives from the first HS Open. The application presents information on the values of the parliamentary representatives and places them on an axis; http://puoluekentta.tstm.info/ See also http://www.hs.fi/politiikka/artikkeli/HS-arvokartta+Soini+johtaa+vasemmistopuoluetta/1135265548291

- "Tax Tree" service makes public sector expenditure visible to the public. Income and expenditure of public bodies are represented in a tree-like diagram. The expenditure of any

department or sub-organisation is represented by the width of the branches. See http://www.mindtrek.org/2009/node/127 (service in a designing phase)

- “Where does my money go” visualises how taxes are spent in different areas; http://wheredoesmymoneygo.org/

- Various examples of the visualisation of information, http://www.informationisbeautiful.net/

- “Map ATL" visualises crime in Atlanta; http://crime.mapatl.com/

- “Crimespotting" visualises crime in San Francisco; http://sanfrancisco.crimespotting.org/

Public participation

Participatory services and applications aim at enhancing the interaction between citizens and administration. These services are often utilised in, for instance, urban planning where they are used to collect citizen-based information for the basis and utilisation of planning. [24]

- "Pitäiskö fiksata” service at the Omakaupunki platform pays attention to local problems; http://omakaupunki.hs.fi/paakaupunkiseutu/uutiset/pitaisiko_fiksata/

- “Fillarikanava” provides a service for cyclists by enabling them to mark problematic places on a map, to read and comment on other posts and have a public discussion; http://fillarikanava.hel.fi/

- “Kerro kartalla” is a service developed by the city of Helsinki that enables civic commenting, discussion, asking of questions, etc. relating to issues of everyday life. Local authorities can use the service to implement queries that use map displays combined with public commenting; http://kerrokartalla.hel.fi/

-“Hatikka” database collects information from citizens that have made observations of wild animals that are injured, ill or dead. The service aims to map how animal diseases, for instance, diffuse (also a possibility to go through collected information); http://vanha.hatikka.fi/public_victim_info-fin.php

-Service for reporting broken street lamps in Tampere provides citizen-based information directly to city organisation; http://katuvaloviat.tampere.fi/ (although the sent information is not visible to other users)

- “What do they know” service lets you choose the public authority that you would like information from, then directs to write a brief note describing what you want to know. The service then sends the request to the public authority. Any response is published on the website; http://www.whatdotheyknow.com/

[24] These services do not necessarily include the features of open data, but can offer insight for the possible services and innovations regarding the future use of open data. For instance, the so-called soft-GIS applications (“pehmo-GIS”) could offer interesting viewpoints concerning open information.

- "WriteToThem" is a British site that helps to work out who your politicians are and helps to send them messages; http://www.mysociety.org/projects/writetothem/

- "Fix my street," a service for public problem reports; http://www.fixmystreet.com/

- "Report empty homes" highlights the waste of empty property and tries proposing sustainable solutions to bring empty property back into use. The details are sent to the right person in the local council to take action; http://reportemptyhomes.com/

- The UK government wants to hear citizens' ideas for new products that could improve the way in which public information is communicated; http://webarchive.nationalarchives.gov.uk/20100807004350/http://showusabetterway.co.uk

- "Watchdog" service gathers information about politics (e.g. votes, lobbying records, campaign finance reports), and also includes developing tools to get active, start own causes and campaigns, invite friends to join, etc.; http://watchdog.net/

- "Mapa de la inseguridad" provides interactive service for citizens to mark places where they have witnessed crimes; http://www.mapadelainseguridad.com/

Solving everyday problems

Services and applications that help to find answers and solutions to problems that are based at a local level. Often, they might also be used improve the citizens' co-operation with the public sector.

- The service map of Helsinki is an open forum for information regarding the departments and services of the City of Helsinki. The map provides the most up-to-date information on the services provided by the city and their locations. Users can also give feedback and engage in direct discussions with the people who are in charge of various departments and services; http://www.hel.fi/palvelukartta/

- Journey planner of Helsinki region provides timetables and routes for public transport; http://www.reittiopas.fi/

- "Fillaristit" is a service where cyclists can share and mark information about routes and places, among other things (e.g. parks, traffic lights, dangerous crossings). They can also comment on other users' messages, send photos, etc. The service is connected to journey planners of cycling in, for instance, Helsinki, Lahti and Tampere. http://fillaristit.fi/ ; see also http://kevytliikenne.tampere.fi/

- Trekking map (Retkikartta, Metsähallitus) presents hunting and fishing places and hiking facilities (routes, services); http://www.retkikartta.fi/retkikartta.php

- "HelMet" bar code reader is a service connected to libraries, where a bar code reader can be used in mobile phones to find out in what library a certain text is located;

http://www.apps4finland.fi/fi/kilpailutyot/yksityisten-sovellussarja/helmet-viivakoodinlukija (not active at the moment?)

-"Tilannehuone" produces information on various accidents (traffic, fires, etc.) and their locations. It also provides links to the news stories of the accidents; http://www.tilannehuone.fi/

- "Fwix," GeoData platform shows local data (places, news, events, reviews, photos, and more) to help find out what's happening in localities; http://fwix.com/

- "CycleStreets" is a UK-wide cycle journey planner system that lets you plan routes from A to B by bike; http://www.cyclestreets.net/

- "PlayaroundNYC" was created to help New Yorkers to see how well their neighbourhoods are supported by playgrounds and to see how this support is tempered by nearby conditions. It aims at creating a better understanding of and encourages the improvement of the quality of neighbourhood playgrounds; http://www.playaroundnyc.com/

- Mapumental mapping tool helps to work out where to live, where to find a job or where to go on holiday; http://www.mysociety.org/projects/mapumental/ (now in a testing phase, but see demo)

- Mapnificent shows the area that can be reached via public transport from any point in a given time. It is available in major cities in the US and globally; http://www.mapnificent.net/

Citizen surveillance & transparency of public sector and politics

Services and applications that provide tools for increasing transparency in politics and in the public sector in general. Transparency is believed to facilitate more open governance and decision making.

- "Kansan muisti" service provides information on, for instance, how politicians have voted, gives information of their campaign funding and shows petitions representatives have given, etc.; http://www.kansanmuisti.fi/about/main/

- "Vaalilupausarkisto" publishes information about the promises politicians have made during elections and how their action corresponds to the promises made. It also presents their public action, for instance in social media, http://beta.vaalilupausarkisto.fi/

- "Helsinkikanava.fi" provides web broadcasts of the meetings of the Helsinki City Council. The meetings are available as direct broadcasts and recordings. The service also shows members who are present in the meetings and provides agendas of the meetings; http://www.helsinkikanava.fi/fi

- "Aselaki" is an application of YLE, the Finnish Public Broadcasting Company that aims at opening the discussion and material that representatives (members of the parliament) use as grounds for their decisions/suggestions on the topic of weapon legislation. The service provides expert statements, the comments of the elected officials and news reports on the

topic, etc.; http://beta.yle.fi/aselaki/#kommentit
(See also beta.yle.fi)

- “HARE hankerekisteri," this service provides data of projects in the public sector, for instance ministries, parliament, civil service departments. Provides information on new projects, requests for comments and on-going governmental procedures; http://www.hare.vn.fi/mAloitussivu.asp

- “Paikkatietoikkuna” (Finnish geoportal) presents geographically referenced data (although not possible to access datasets?). This national portal presents the spatial data produced and exploited in Finland; http://www.paikkatietoikkuna.fi/web/en/

- “Hilmappi” presents map-based information on the purchases of administration. http://www.hilmappi.fi/ See also the administrative site, which produces the data for the purchases http://www.hankintailmoitukset.fi/fi/

- “Opensecrets” service tracks money in U.S. politics and illustrates its effect on elections and public policy; http://www.opensecrets.org/index.php

- “Earmark Watch” investigates spending measures inserted by members of Congress into bills that direct taxpayer dollars to their pet projects; http://earmarkwatch.org/

- “UK public spending” shows what government spending consists of; http://www.ukpublicspending.co.uk/

- Poligraft website creates an enhanced view of the people, organisations and relationships described within political news stories, blog posts and press releases. Graphs for politicians represent received campaign contributions, while graphs for organisations represent aggregate campaign contributions made; http://poligraft.com/about

- “Openly Local” project develops open and unified way of accessing local government information (gives access to data describing the workings of local councils); http://openlylocal.com/; see, for instance, http://openlylocal.com/councils/spending

- “KildareStreet” is a non-partisan website that aims to help people to follow their elected representatives; http://www.kildarestreet.com/

- “Fixing D.C.’s schools” service provides information on the schools in Washington (for instance, crime, health, safety issues); http://www.washingtonpost.com/wp-srv/metro/interactives/dcschools/scorecard.html

- “MapLight” service provides transparency tools to follow the influence of money on politics. Service makes money/vote connections transparent; http://maplight.org/

- Voting aid applications, such as http://www.vaalikone.fi/, let users compare their own views to politicians’ views on topical issues

- "Ennakoinnin tietopalvelu Ensti" of Finnish National Board of Education provides statistics and future guidelines/predictions in the area of education; http://www.oph.fi/tietopalvelut/ennakointi/ennakoinnin_sahkoinen_tietopalvelu_ensti

- A collaborative mapping project to build a database of bloggers who have been threatened, arrested or killed for speaking out online and to draw attention to the campaigns to free them, http://threatened.globalvoicesonline.org/

- "HowSFVotes" is a tool for exploring historical election results in San Francisco precinct by precinct, http://www.howsfvotes.com/

- "Suburban Trends" service compares Australian neighbourhoods, such as socio-economic standing, education levels and perceived safety levels; http://www.suburbantrends.com.au/

Appendix II: The list of interviews

Meloni, Ville (Forum Virium, HRI), 23.3.2011
Muurinen, Raimo (Vaalilupausarkisto), 19.5.2011
Mäkinen, Esa (Helsingin Sanomat), 11.5.2011
Poikola, Antti (Fillarikanava), 10.3.2011
Rantanen, Heli (Helsingin kaupunki), 15.3.2011
Rastas, Taru (Liikenne- ja viestintäministeriö), 30.3.2011
Ropponen, Teemu (Oikeusministeriö), 8