MINISTRY OF FINANCE

The Impact of open data – a preliminary study

Ministry of Finance publications– 15b/2015

Public Sector ICT

Description page

Publisher and date	Ministry of Finance, March 2015	
Author(s)	Heli Koski, Elinkeinoelämän tutkimuslaitos	
Title of publication	The Impact of open data – a preliminary study	
Publication series and number	Ministry of Finance publications 15b/2015	
Distribution and sale	The publication can be accessed in pdf-format in Finnish at www.vm.fi/julkaisut. There are also instructions for ordering a printed version of the publication.	
Printed by	Juvenes Print – Finland University Print Ltd, 2015	
ISBN 978-952-251-669-5 (print.) ISSN 1459-3394 (print.) ISBN 978-952-251-670-1 (PDF) ISSN 1797-9714 (PDF)	No. of pages 27	Language English

Abstract

Various countries have implemented the open (government) data strategy aiming at providing wide access to government data in machine-readable format such that it can be freely used, reused and redistributed by anyone. Reported ex ante evaluations have estimated that the potential benefits of opening up public data resources are substantial.

Very little is known about the underlying economic and organizational mechanisms and implications of open data use at the organizational level or at the level of economy as a whole. To my best knowledge, there is no reported comprehensive country-level ex-post impact assessment of opening up government data.

Currently, Finland is among the leading countries in opening up the government data, and it has also a chance to be among the most advanced ones in the impact assessment of open data. The impacts of opening up government data can be divided to economic impacts and to other social impacts. The prerequisite for this is a careful development of the monitoring and evaluation model for opening up government data as well as a systematic gathering of data for the impact assessment.

Furthermore, the usability and usefulness of different public data resources for consumers, firms and public sector organizations can be assessed via the users' own evaluations. In addition, it is important to assess appropriate means to disseminate and promote efficient utilization of information on the best practices of open data re-use in different organizations.

The impact of open data – a preliminary study

Heli Koski, The Research Institute of the Finnish Economy

12 January 2015

SUMMARY

Reported ex ante evaluations have estimated that the potential benefits of opening up public data resources are substantial. Consequently, various countries have implemented the open (government) data strategy aiming at providing wide access to government data in machine-readable format such that it can be freely used, reused and redistributed by anyone. In March 2011, the Finnish government made a resolution to open up government data. In May 2013, the Ministry of Finance launched a two-year Open Data Program to accelerate government data openings. In Finland, the target is to open up all major public databases by the year 2020.

Very little is known about the underlying economic and organizational mechanisms and implications of open data use at the organizational level or at the level of economy as a whole. Various reports have evaluated the economic impacts of open data ex ante but the literature lacks ex post analysis focusing on the materialized economic impacts. To my best knowledge, there is no reported comprehensive country-level ex-post impact assessment of opening up government data. The current research-based knowledge concerning the impacts of open data is limited only to narrow areas and largely based on case examples.

Currently, Finland is among the leading countries in opening up the government data, and it has also a chance to be among the most advanced ones in the impact assessment of open data. The prerequisite for this is a careful development of the monitoring and evaluation model for opening up government data as well as a systematic gathering of data for the impact assessment. This report provides guidelines for the ex-post evaluation of the impacts of open data. It summarizes contemporary knowledge concerning the impacts of open data, what methods and measures can be utilized in research assessing the socio-economic impacts of open data, what kind of data are currently available for such research and what kind of data are further needed. This report focuses on measurable, economic impacts but also other social impacts are briefly discussed.

The impacts of opening up government data can be divided to economic impacts and to other social impacts. Economic impacts can be assessed at the level of firms, citizens and households, public sector and the economy as a whole. The most notable potential economic impacts for firms are growth and increased productivity via the efficiency improvements and via the development of new products and services enabled by open data. Opening up government data is also expected to result in the establishment of new companies. For citizens, the most important economic benefits of open data are likely to arise from free access to data instead of using chargeable data resources and from the time savings. For public sector organizations, opening up data resources offers cost savings and an opportunity to improve efficiency of service provision. The economic impacts of opening up government data may also be observed at the level of economy as a whole if open data can be exploited sufficiently broadly across different sectors of the economy. Other potential social impacts of opening up government data concern, for instance, the transparency of government and decision making, education, health, citizens engagement, environmental impacts, sustainability and transportation.

Contemporarily gathered and available (statistical) data is not sufficient for the comprehensive assessment of the materialized socio-economic impacts of opening up government data. For such assessment, we first need to develop the monitoring and evaluation model and then systematically gather data over the years for the ex post assessment of the impacts of open data. This is a prerequisite for the quantitative impact assessment of open government data. After that, (partial) cost-benefit analysis can be applied for evaluating the ex post changes in the welfare followed by the opening up the government data. Furthermore, the impacts of open data can be qualitatively assessed focusing on those important (expected) social impacts of which measurement in euros is difficult.

Furthermore, the usability and usefulness of different public data resources for citizens, firms and public sector organizations can be assessed via the users' own evaluations. A natural platform for such evaluations and feedback is the open data service providing access to the government data. In addition, it is important to assess appropriate means to disseminate and promote efficient utilization of information on the best practices of open data re-use in different organizations.

1. Introduction

The importance of using data resources is constantly increasing both in the private and public sectors. At European level, the European Commission supports the opening up of public data as part of its Open Data Strategy, published in December 2011. It further supports the use of data resources in the Member States, for example through research programmes and infrastructure projects. Finland has designated the opening up of public data resources as one of the Government's priority projects. In May 2013, the Ministry of Finance launched a two-year Open Data Programme whose aim is to open up public data resources as extensively and quickly as possible. The goal is to open up all significant data resources maintained by the public administration by 2020, so that the data therein will be available and usable in machine-readable format, free of charge and under clear terms of use.

However, socioeconomic research of open data use has been relatively scarce to date, and very little is known about the underlying economic and organisational mechanisms and implications of data use at the level of organisations or the economy as a whole. While several reports have sought to evaluate the economic impacts of open data ex ante (see e.g. Vicker, 2011), ex post analysis focusing on the materialised economic impacts is still in its early stages. This is due to challenges within the research area (e.g. measurement difficulties, lack of systematically collected statistical data on the use of data) and the fact that the phenomenon is relatively new. Research-based knowledge is particularly needed on the preconditions and measures required to ensure that maximum socioeconomic benefits will be gained from open data use.

Research of the topic will require systematic, long-term efforts. A prerequisite for research concerning the impacts of opening up government data would be the development of an assessment model depicting the impacts of open data. It is essential to first explore what is currently known about the impacts of open data, what research methods and indicators can be used for studying the socioeconomic impacts of open data, which types of usable data are already being collected and what are the additional data needs. This preliminary study aims at producing such knowledge. While the focus of this report is on measurable economic impacts, social impacts are also briefly addressed.

Chapter 2 describes the key actions taken in Finland to open up government data, and sheds light on Finland's position in the international ranking of openness of government data. Chapter 3 examines the estimated economic and other social impacts of open public data. Chapter 4 lays down a preliminary framework for the development of an assessment model depicting the impacts of open data. Chapter 5 provides a summary.

2. Opening up data resources – current situation in Finland and an international comparison

Public data to be opened up by 2020

Open public data refers to any data produced or maintained by a government organisation, saved in machine-readable format and available free of charge for anyone to use, reuse and redistribute for both private and commercial purposes. Private actors can also offer open data, but this report focuses merely on the open government data . Important public sector data, which is also widely used by firms, includes geographic, weather, demographic, firm, traffic and legislative data, for example.

In March 2011, the Finnish Government issued a resolution to increase re-use of public data as one of the key measures of the Government Programme. Prompted by this resolution, government organizations began opening up their data resources in 2012.[1] For example, the National Land Survey of Finland opened up its digital topographic data in May 2012. A year later, the Finnish Meteorological Institute opened up weather data via its online service. The goal is to open all significant data resources collected and maintained by the public administration by 2020, making them available and usable in machine-readable format, free of charge and under clear terms of use[2]. Currently, the Finnish legislation does not allow the opening up of all data resources; legal restrictions such as privacy protection, trade secrets, copyrights and national security must be taken into account when

[1] Important public data had been opened up before the resolution was adopted. Prior to 2011, the Finnish Environment Institute for example offered registered users environmental and geographic information through its OIVA service (https://wwwp2.ymparisto.fi/scripts/oiva.asp).
[2] http://valtioneuvosto.fi/etusivu/rakenneuudistus395285/tiedostot/ministerioiden-materiaalit-15112013/vm/VM-aineistot-6-15112013.pdf

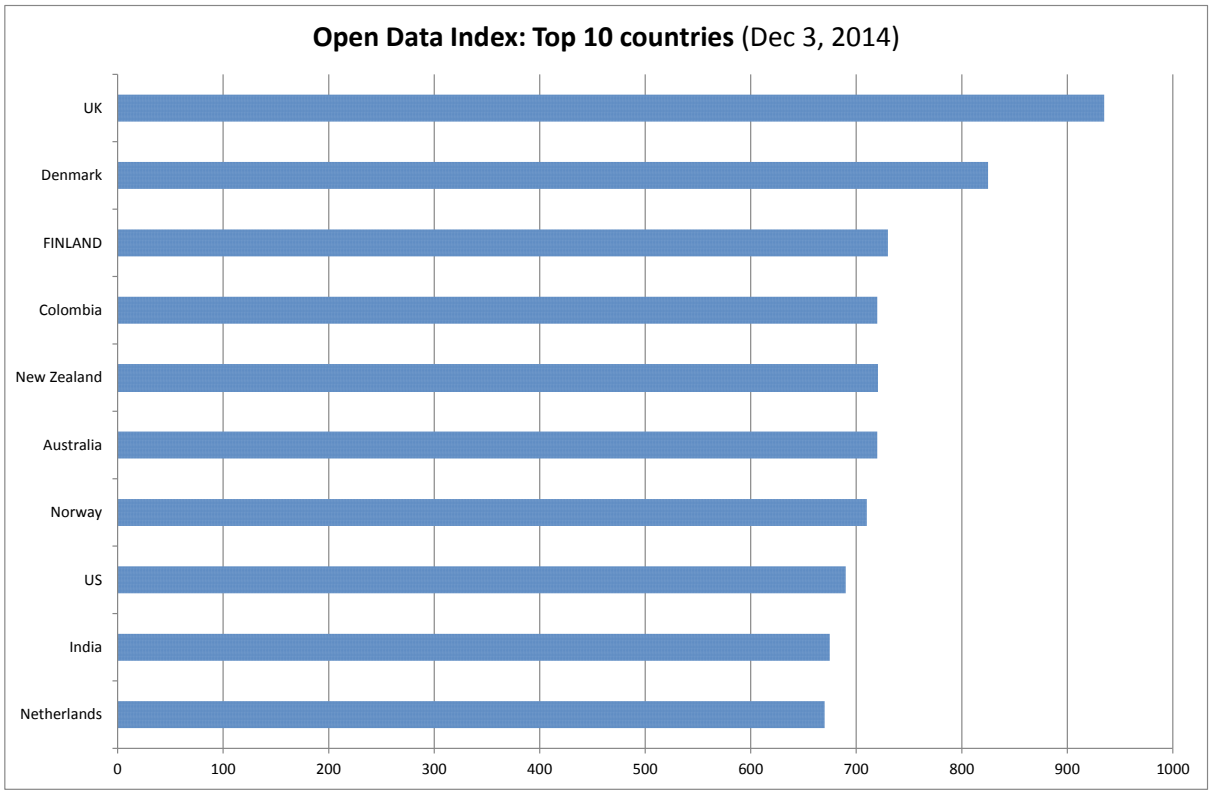
opening up data. All data resources that have been or will be opened are available through the public administration's joint open data service Avoindata.fi, launched in September 2014.

Extensive international cooperation is undertaken to increase the openness and transparency of public administration, with the opening up of data resources as one of its key aspects. The international Open Government Partnership (OGP) launched in 2011 aims to make governments more open, accountable and responsive to citizens. Finland was accepted as an OGP partner in April 2013. In November 2014, a total of 65 countries participated in the OGP, with the aim of achieving more open government (http://www.opengovpartnership.org/). Commitments to opening up public data have played a key role in the open government action plans published by the participating countries (Davies et al., 2013).

Finland is approaching the front line of open data

International comparisons have been conducted to assess the progress made in opening up public data in different countries. The Open Data Index is an indicator used by the Open Knowledge Foundation to describe the degree of openness of public sector data in various countries. The opening up of public sector data is monitored in ten areas: public transport schedules, state budget, government expenditure, voting results, business registers, national map database, each country's key economic and demographic statistics, legislation, postal codes and environmental emissions. The bodies responsible for updating the Open Data Index vary between countries, and in some countries reporting on the status of open data depends on volunteers. Therefore, the international comparison of openness of government data offered by the Open Data Index should be considered indicative.

According to the Open Data Index, the United Kingdom, Denmark and Finland were the three leading countries in terms of opening up public data in December 2014. Finland had risen three places from the previous year, while the United States fell from second position to eighth. In Finland, the greatest deficiencies in the opening up of public data resources concerned business registers. In addition, legislative data were not completely open.



Data source: Open Data Index 2014, Open Knowledge Foundation. https://index.okfn.org/country/

Countries that were forerunners in opening up government data (e.g. United Kingdom and the United States) initiated the major policy measures for opening up public data slightly earlier than Finland. Published in December 2009, the key objectives of the United Kingdom's open data strategy "Putting the Frontline First: Smarter Government"[3] were to support economic growth and innovation and to strengthen the role of citizens and civic society by opening up data resources and increasing transparency. In the United States, President Barack Obama published a memorandum in January 2009 stressing the importance of opening up public data, leading to the launch of the Open Government programme in December in the same year. One of the focal aims of the programme was to increase the transparency of political decision-making by launching online databases to enable citizens to monitor the actions of the US government. Strengthening law enforcement via the use of data-based applications has been one of the reasons for the opening up of data in countries such as the United Kingdom and the United States. This has been done in order to both inform citi-

[3] Chief Secreatry to the Treasury (2009). Putting the Fronline First: Smarter Government. London: HMGovernment.

zens and to increase their participation in law enforcement (e.g. through applications that provide crime information).

Several countries have opened up public data resources as part of their open and transparent government strategies. The national strategies of several European countries in particular (e.g. Denmark[4]) cite the economic benefits of open data, such as the emergence of new products, services and firms, reduced costs, greater efficiency and economic growth, as important factors that work in favour of open public data. Several countries refer to various socio-economic benefits as the underlying reason for opening up public data.

3. The impact of open data

Opening up data increases its use

There is only little research-based knowledge concerning the realised socio-economic impacts of open data use. The Open Data Barometer project reported that, in 2013, none of the 77 countries assessed had conducted a comprehensive assessment of the impacts of opening up data.[5] A major reason for this is the lack of systematic data collection and compilation of statistics concerning data use. As only limited knowledge is available based on empirical analysis, theoretical speculation and case examples are the primary information sources that can be used for shedding light on the impacts of open data.

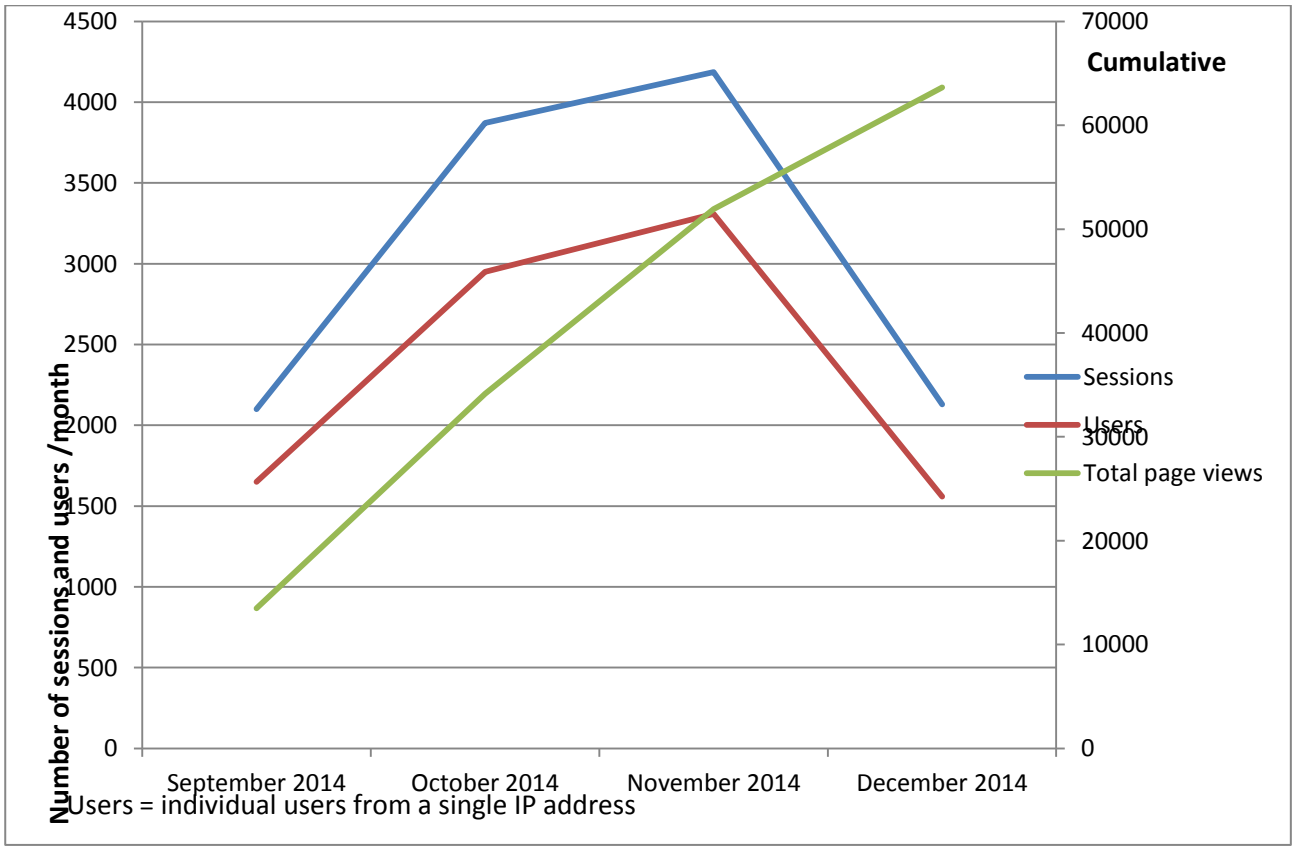
Various examples suggest that opening up government data significantly increases the use of data resources. For example, in March 2003 Spain founded the OVC ("Virtual Office of Cadastre") in order to offer comprehensive geographic data free of charge. The process of opening up data continued in June 2004, when the launch of the IDEE geoportal further improved the availability of geographic data. The geoportal made a significant amount of geographic data available to anyone

[4] The Danish government/local government Denmark, October 2012: "The eGovernment strategy 2011-2015. Good basic data for everyone – a driver for growth and efficiency." (http://uk.fm.dk/publications/2012/good-basic-data-for-everyone/~/media/Publikationer/Imported/2012/Gode%20grunddata%20til%20alle/BasicData_UK_web_2012.10.08.ashx.)

[5] Open Data Barometer. 2013 Global Report. Source:http:// www.opendataresearch.org.

free of charge. This resulted in a considerable increase in demand for geographic data. For example, map use increased by some 700% from 2004 to 2005 (from over a million uses to more than 41 million). In 2010, the corresponding figure was more than 124 million, meaning that demand was a hundred times greater than in 2004. Many other, similar examples exist. Geographic data was opened up in Finland in 2012, making its use more frequent and versatile. Ahonen-Rainio et al. (2014) report that new users have emerged among private individuals and SMEs in particular. Use of the open government data portal Avoindata.fi has also rapidly increased since it became available. After the service was launched, the site was viewed more than 63,000 times in the first 3.5 months.

The use of Avoindata.fi service between 15 September and 31 December 2014



The observed prominent increase in the use of data after opening up of public data does not in itself reveal anything about the socio-economics impacts of open data use. The impacts of opening up government data can be categorised into economic and other social impacts. The economic impacts

of open data can be assessed at the level of firms, citizens (and households), the public sector and the national economy as a whole:

i) *impacts on firms and industries*, such as impacts on competitiveness, innovation, productivity and competition and market dynamics (companies' entry into and exit from the market, entrepreneurship)

ii) *impacts on citizens*, such as impacts on consumption, income distribution and the status of various population groups (e.g. wage earners, the recipients of different benefits)

iii) *impacts on the public sector*, such as impacts on central government and municipal finances and the authorities' use of resources, impacts on employment and productivity in the public sector, and impacts on public services

iv) *economy-level impacts*, such as impacts on productivity and growth at the level of the national economy, and impacts on the labour markets.

Other potential social impacts of opening up data resources concern the transparency of governmental decision-making, education, health, citizens' status and social engagement, environmental impacts, sustainable development and transport.

Innovation, growth and productivity gains for firms using open data?

The ability of a firm to benefit from open data use greatly depends on how information-intensive its operations are. The information-intensiveness of different sectors has been measured by various means such as by the amount of data per organisation and the number of data analysts employed by the organisation in relation to its total number of employees. According to a 2011 report by McKinsey Global Institute, the most information-intensive sectors include financial services, information and communications, and electricity, gas and heating supply. The OECD (2013) has reported that use of open data to stimulate innovation and increase profitability in the private sector offers most potential in the health care sector, advertising, the production of infrastructure services (e.g. gas, electricity, water) and transport and logistics.

Based on the above, the impacts of open data on business operations are likely to vary significantly from sector to sector. Companies whose current operations are only slightly based on data use may

begin to use various kinds of data more extensively, as more data become available and more knowledge is gained on data-driven business operations. Companies may use data not only as raw material for new products and services, but also more extensively to support their decision-making, for instance in construction management, agricultural production management and in business location choices. In the long run, the benefits for firms from open data greatly depend on how rapidly and extensively firms in different sectors are able to implement data-based operations (e.g. data-based decision-making systems). An empirical analysis by Brynjolfsson et al. (2011) showed that firms using data-driven decision-making models were 5–6% more productive and profitable than their competitors.

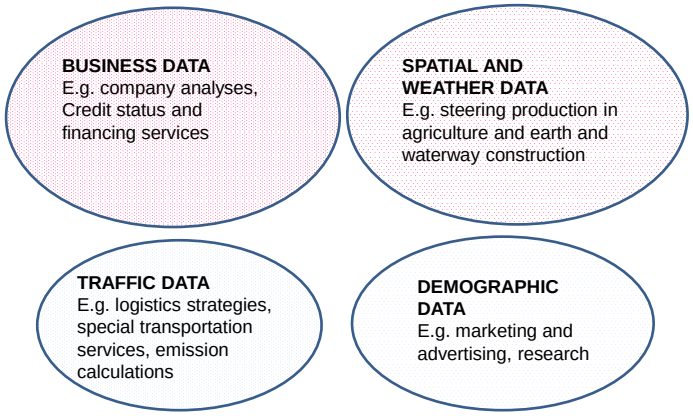
Contemporary literature on the firm-level impacts of open data use focuses, by and large, on the firms' use of spatial data. Ahonen-Rainio et al. (2014) assess the impact of open spatial data. Their report sheds light on the ways of using, importance and benefits of geographic data for users but it does not seek to assess its economic impacts in euro terms. Their survey reveals that open spatial data brings financial benefits to companies by enabling the development of new services and products, making operations more efficient and providing companies with a competitive edge both in Finland and abroad.

Empirical research conducted by Koski (2011) indicates that, for firms in architectural and engineering activities and related technical consultancy sector (i.e. Standard Industrial Classification 7420) in 15 countries, the opening up of geographic data and a shift towards using marginal cost pricing was connected to significantly faster sales growth. In the countries where geographic data was available at marginal cost prices or free of charge, firms' sales grew by an average of 15% faster than the sales of firms in other countries. The opening up of geographic data did not have a significant impact on the growth of large enterprises, however.

The literature also provides only scarce research-based knowledge concerning the use of open data in firms' innovation activities. Based on survey responses from 531 Finnish companies, Koski's (2012) research examined the generation of new data-based products and services in Finnish companies. The results suggest that companies engaged in professional, scientific or technical activities,

real estate or construction were more likely to produce data-based product and service innovations than companies in other sectors. Interviews conducted by Kiuru et al. (2012) revealed that that particularly small firms are able to benefic quickly from newly opened public data in their products and services.

Examples of the use of open data in companies



Various studies suggest that the opening up of public data may facilitate innovation and growth, particularly in SMEs operating in sectors where data is a key component of business operations. The literature also implies that the size and growth of certain data-based markets in various countries is closely linked to the openness of the data used in these sectors (see e.g. Pettifer, 2009). In the United States – unlike in Europe – meteorological data has long been open. The U.S. meteorologial market has grown markedly faster than its European counterpart: in 2006, the total size of the meteorological products and services market was USD 1.4 billion in the United States, while the corresponding figure for Europe was only USD 372 million. The average annual growth rates of these markets were also quite different: in 2000–2006, this was around 17% in the United States, but only 1.2% in Europe.

The use of open data enables the establishment of new companies the adoption of new business models. It further contributes to the emergence of new sectors whose operations are based on the use of open data. For example, in Spain in 2012 more than 150 companies operated exclusively in the "infomediary" sector producing value-added services using open data. The sector employed some 4,000 people at the time, and the annual value of these services was estimated to be EUR 330–500 million.

The Spanish "infomediary" sector is an example of the new business opportunities created by open data, but we do not currently know to what extent the reuse of raw public data has led to the emergence of new companies in different sectors. The literature mainly reports on individual success stories. In the United States for example, the Weather Channel and Garmin, a leading company in the GPS market, were founded using raw public data. Garmin's market value was more than USD 7 billion in 2013 (Capgemini Consulting, 2013; The Open Data Economy). Another example of harnessing raw public data for commercial use is provided by Climate corp., founded in 2006. It initially compiled weather and crop forecast data for weather insurance purposes.. The company later expanded its use of such data to support sowing plans. Climate corp. was sold in the autumn of 2013 for USD 930 million.

Open data for citizens – savings, better decisions and participatory democracy?

Data resources collected and managed by the public sector are public goods funded by taxes, insofar as legislation does not limit their public status. Therefore, citizens are entitled to have access to government data. From the viewpoint of citizens, one of the most notable economic benefits of public data is the way in which it enables more efficient time management. Easy-to-use applications providing real-time information on estimated travel times using different means of transport are an example of the economic benefits that can be offered to consumers (e.g. time savings when rush hours are avoided).[6] Free access to data that previously had a price tag also provides users with financial benefits.

[6] An example of such an application is the Journey Planner of Helsinki Region Transport: www.reittiopas.fi.

Opening up public data may help citizens to make better decisions in their personal lives. The availability of data on the local environment and applications for using such data may support decision-making and improve the quality of people's lives. For example, precise local weather forecasts may influence an individual's transport-related decisions (e.g. the time and mode of transport to another location) and further improve traffic safety. In several countries, such as the United States, applications using open data have been developed to map local crime incidents (e.g. on a specific street), in order to improve the safety of citizens and prevent crime. Such applications providing localized data may also influence people's choices concerning their neighbourhood or area of living.

The openness of data on political decision-making makes policy-making more transparent and easier to monitor, and can further support participatory democracy. When citizens can monitor government and decision-making, they are able to vote for decision-makers that better represent them. This may improve welfare, at least in principle. It is not known, though, to what extent citizens use open data to monitor governmental decision-making processes.

More efficient public services through open data?

The public sector is one of the most information-intensive sectors of the national economy. It is an important producer, maintainer and user of data. The OECD (2013) estimates that the public sector is one of the most likely sectors to highly benefit from the use of data resources. Open data may provide the public sector with higher data revenues due to the establishment of new (open) data-based companies. Pettifer's (2009) calculations suggest that the opening up of meteorological data in Europe would have increased annual tax revenue from the meteorological sector by about EUR 300 million. Opening up the public administration's internal data resources across borders could also increase the public sector's tax revenue, because international openness of taxation data can reduce international tax evasion.

Meanwhile, open data resources enable public sector organisations to improve the efficiency of their services and cut costs. Organisations managing closed data resources need to manage a significantly larger flow of database-related queries from their customers than organisations whose data

resources are open. In other words, opening up data can reduce the respective public organisations' transaction costs. For example, San Francisco has been able to reduce its number of incoming telephone queries by about 22% and generate annual savings of more than USD 1 million by opening up data on public transport[7].

Increasing the transparency of public organisations' activities can provide means of increasing the efficiency of decision-making and operations. Open data can be used for planning in public administration and to bring greater operational cost-efficiency. It can also be used to analyse cost-saving potential and the amount of savings that can be attained by taking, or encouraging others to take, various measures. The National Health Service (NHS) in the United Kingdom would have saved more than GBP 200 million in 2011 if each doctor had prescribed a similar generic drug for cardiovascular diseases rather than a brand-name drug[8].

The public sector can use open data as an internal incentive mechanism for improving the efficiency and quality of public services. When the United Kingdom opened up the infection statistics of all hospitals and listed the hospitals with the worst infection status via its open data portal (data.gov.uk), this encouraged hospitals to share best practices (Gapgemini Consulting, 2013). Capgemini's investigation revealed that the annual number of infected patients fell from about 5,000 patients to fewer than 1,200. Open data can also be used to improve the efficiency and quality of public transport.[9] Data on passenger numbers at different times and in different places can be used to plan bus and train timetables that better meet customers' needs.

Opening up public data resources can also support education and research. Making data available improves the preconditions for studying various social and economic phenomena. Major organisations can avoid duplicate work by sharing their data and research reports based on it. Similarly, in

[7] Source: http://www.resetsanfrancisco.org/better-government/opendata-and-real-time-information-saves-san-francisco-over-1-million/.
[8] Source: http://www.economist.com/news/britain/21567980-how-scrutiny-freely-available-data-might-save-nhs-money-beggar-thy-neighbour.
[9] The example set by Helsinki Region Transport illustrates that opening up data and developing open interfaces for application developers can also contribute to cost-efficient development of the quality of public transport services. http://www.hri.fi/fi/ajankohtaista/datan-avaaminen/hsln-joukkoliikennedata/

the field of education it is possible to reduce duplicate work and enhance education by making existing data and teaching materials openly available.

Are the benefits of open data realised at the level of the national economy?

If open data can be harnessed extensively enough for the benefit of different economic sectors, its impacts may also materialise at the level of the national economy. The growth of existing companies and the creation of new ones could also contribute to economic growth at national level. Increased public sector efficiency and higher private sector productivity could further generate aggregate productivity growth.

Several reports present *ex ante* estimates of the impacts of open data at the national and even global economic levels. Such estimates consider the economic potential of open data to be remarkable. However, measurement difficulties and other challenges related to *ex ante* impact assessment are reflected in the fact that the reported estimates tend to vary greatly in size. Dekkers et al. (2006) conducted a survey for maintainers and re-users of public sector data resources in different European countries, revealing considerable differences in the respondents' opinions and estimates concerning the value of the public data market in their respective countries. The value of the European open data markets was estimated to be between EUR 10 and 48 billion. Vickery's (2011) calculations suggest that the value of the markets created by harnessing public data resources was about EUR 28 billion in the European Union in 2008, while the annual value of the markets created by opening up data resources was about EUR 40 billion. McKinsey (2014) estimates that, at global level, open data (including public sector data and privately produced machine-readable data) could generate over USD 3 trillion in additional value on an annual basis in the seven key economic sectors in which such data plays the most pivotal role[10].

In other words, existing knowledge does not permit a detailed assessment of the realised and future impacts of the opening up of public data resources on employment, productivity and growth at the

[10] McKinsey (2014) estimated the impacts of open data in the following economic sectors: education, transport, consumer goods, electricity, oil and gas, health care and financing.

level of the national economy. In this context it should be noted, though, that only a few innovations (e.g. electricity, steam, information and communication technology) have been observed to have significant macroeconomic impacts.[11] Even then, these impacts have typically been realised decades after the initial adoption of the innovation.

The risks and risk management of the Open Data Strategy

Discussion of the impacts of open data has, by and large, revolved around an assessment of the related benefits (and sometimes costs). When the potential problems or risks associated with opening up public data are addressed, typical issues include the constitutional right of citizens to privacy protection and the resulting restrictions on opening up public data. For various data (e.g. patient records), there is a clear need for legal protection, so that they cannot be opened without the consent of the interested parties.[12] The situation is more complicated for instance in the case of geographic data. Data can only be opened insofar as personal-level information remains anonymous. The question of what constitutes anonymous information must be resolved on a case-specific basis though. For the Office of the Data Protection Ombudsman, "*According to the data protection working group, whether data can be used to identify a person or considered anonymous depends on the situation. Each case must be addressed separately, taking consideration of the likelihood of reasonable measures being taken to identify the person in question.*"[13]

Several studies have estimated that substantial potential benefits lie in the use of open public data. However, these benefits may not be realised in full. Public organisations' insufficient capacity to utilise data is a potential obstacle for doing so. If large amount of open data is not organised and presented in a sufficiently accessible format, it may be difficult for a user to find the information (s)he needs. From the viewpoint of data users, it is important that, once opened up, data is kept up-to-date and its quality is maintained. If public organisations that collect and maintain data do not

[11] Technologies that are widely used across sectors and contribute to macro-level economic growth are known as general purpose technologies.
[12] Personal data can be opened up with the consent of the person in question. For example, people suffering from serious diseases may want to open up their health information in support of medical research (Poikola et al., 2014).
[13] Source: http://www.tietosuoja.fi/fi/index/ratkaisut/matka-aikatiedontuottaminenmatkapuhelinpaikannuksenavulla-paikkatietohenkilotietojavalillinentunnistettavuus.html

have sufficient know-how and incentives to maintain and update open data[14], this may result in outdated data use and non-optimal decisions, and the reuse and redistribution of outdated data. If the public administration's information systems were adapted to offer open interfaces to third parties, the problem of maintaining open data would be solved and maintenance costs would fall. Data submitted by the municipalities to the state could also be easily made public using such interfaces. However, high-quality open data may not be enough: lack of know-how on the users' part can lead to false interpretations and decisions and less-than-optimal use of data resources in both households and companies.

The literature also addresses the emergence of a digital elite and increasing social exclusion among the risks associated with open data. Not every citizen has the same capabilities and skills to exploit open data resources. Without separate measures, open data will not therefore automatically enhance participatory democracy.

The risks involved in opening up and using public data can be minimised by carefully planning the measures (e.g. updates) to be taken after the opening up of public data resources. It would be important to train both data providers and citizens (e.g. information literacy and source criticism). Huijboom and van der Broek (2011) provide a list of instruments for implementing open data strategies. These instruments – many of which are also suitable for the management of risks involved in the opening up of public data – fall into four main categories:

1) education and training (e.g. platforms to exchange experiences/ideas on open data projects, guidelines, conferences, workshops, etc.),
2) voluntary approaches at national level to support the opening up of public data
3) economic instruments (e.g. competitions, application development contests, financing open data portals) and
4) legislation and control (e.g. legislation on the data resources of the public sector, technical standards to enhance compatibility, monitoring of open data projects)

[14] Peled (2011) reports that in May 2011, nearly 30% of the 169 organisations that provided data through the open data portal data.gov had not updated their data for at least a year.

4. Development of a monitoring and assessment model for the impacts of open data

Currently, there is no such data available that would enable a comprehensive assessment of the realised socio-economic impacts of the opening up of public data. Such an assessment would require the development of a monitoring model and indicators, followed by systematic data collection in order to measure the realised impacts. This chapter sheds light on potential indicators for impact assessment, useful data already being collected and additional data required. This will provide the basis for for the detailed development of a monitoring model and a set of indicators. It mainly covers measurable, economic impacts, but can be complemented by a qualitative assessment.

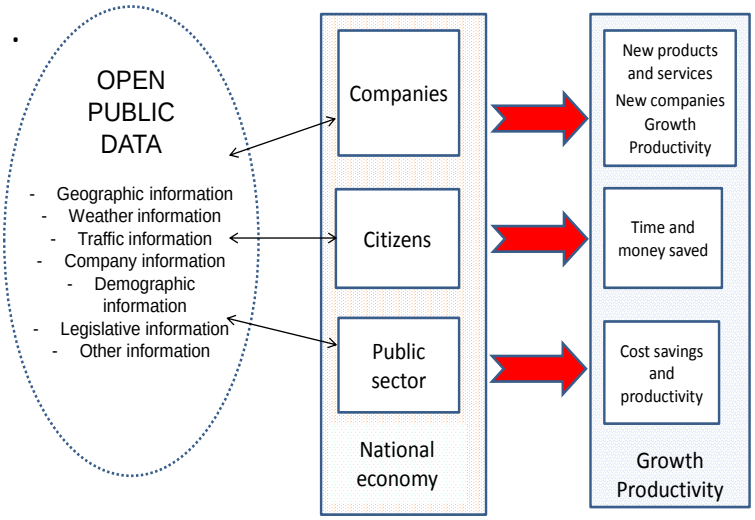
After indicators have been developed and data collected, a (partial) cost-benefit analysis can be conducted to assess the welfare implications of the opening up of public data. The aim of the cost-benefit analysis is to define a monetary value for all impacts of open data and use this to calculate their total value and profitability for society. An actual cost-benefit analysis could, in all likelihood, only be conducted for key impacts, but could be used to produce an estimate of the "minimum threshold" of the impacts of open data. In addition, the impacts of open data use can be assessed qualitatively[15]. A qualitative impact assessment is also likely to focus on the most prominent impacts from the society's point of view (e.g. increase in civic activity).

One of the most notable expected firm-level impacts of open data concerns new products and services. Successful innovations are an essential source of productivity and economic growth. The assessment of expected firm-level impacts would reasonably focus on the impacts of open data use on the innovation, growth and productivity of firms. Firm growth is often measured by the increase in the company's turnover and number of employees. Total factor productivity measures output in relation to inputs, or, in the case of companies, the quantities of products and services produced by labour and capital. Firm-level productivity is typically measured by labour productivity, which is

[15] For example, a qualitative assessment of the impacts of open data has been conducted by the Ministry of Transport and Communications (2013).

calculated by dividing added value by the number of employees or hours worked. Annually collected financial data on companies can be used to measure their growth and profitability.

Open data: the key expected economic impacts



Assessing the impacts of the use of public data on citizens and households in euro is challenging. In cases where data that used to be subject to a charge is now offered for free, it would basically be possible to examine use rates in order to form a rough estimate of the monetary savings attained by citizens when they use such data. In practice, this information could only be attained via a survey, since the user statistics of online open data services do not differentiate between citizens and other types of users. It would probably be too difficult for someone to assess the time and money (s)he has saved through the use of open data, though. However, surveys could be conducted in order to shed light on the use of open data, the types of data used by citizens and the impact of the use of public data resources on households.

Public organisations form an important sector in which the impact of open data should be assessed. The opening up of public data has impacts on both the organisations opening up such data and those using it. The potential costs savings and productivity gains that may be achieved through the use of

open data are particularly interesting. The productivity of the public administration is measured by the quantity of services produced relative to the inputs used for producing them. In order to assess such impacts, there is perhaps a need to develop a technical solution for measuring, as cost-effectively as possible, the offering and use of data and other services by public administration organisations (at least key ones), and the resources they use in providing such services.

The Ministry of Finance conducts an annual survey whose results are published in the Finnish Government ICT Review. The potential usability of this survey data in assessing the impacts of open data should be investigated. Furthermore, consideration should be given to whether additional survey questions should be added in order to measure the impacts of open data. It would be particularly important to evaluate the amount of additional work needed for, or the time that is saved by, the opening up of data. The average labour costs per hour of the occupational groups could then be used to calculate the price of the additional work caused or the time saved by open data. Statistics Finland conducts an annual survey for compiling the "Statistics on central government productivity" publication that includes information on volume changes in the output, labour input and total input of state accounting offices and institutions. This information is further used for calculating changes in labour productivity and total factor productivity, also reported in the publication. Consideration should also be given to using these productivity statistics for impact assessment.

Economic impact assessment of open data: indicators for expected impacts

	Expected impact	Indicators	Potential sources of information
Companies	New products and services New companies	The quantity of new products and services based on open data The share of income from open data services/products of total turnover	Surveys, adding questions to e.g. the "Use of information technology in enterprises" survey of Statistics Finland and Eurostat's Community Innovation Survey).
	Growth	Changes in turnover and number of employees	The Business Register of Statistics Finland Firm-level financial statements database of Suomen Asiakastieto Oy

	Productivity	Added value per employee or per hour of labour	Firm-level financial statements database of Suomen Asiakastieto Oy
Citizens	Time and money saved	Amount of free data used x earlier price Estimated time savings	Surveys, adding questions to e.g. the ”Use of information and communications technology” survey of Statistics Finland
Public sector	Cost savings Productivity	Changes in labour hours x labour costs per hour Service output in relation to the resources used	Automatic monitoring of information and service offering Surveys, the Finnish Government ICT Review, adding questions therein Statistics on central government productivity
National economy	Growth Productivity	Changes in added value i.e. GDP GDP per capita	Statistics Finland's statistical databases

The expected key impacts and indicators for systematic impact monitoring and assessment of open data are summarised in the table above. In addition to this, indicators are required for the measurement of the use of open data in different economic sectors. At the moment, data is not being systematically collected on the use of open data and its role in firms, households and the public sector. Surveys are the main method of gathering information on the use of open data. For example, citizens' use of open data could be mapped as part of the annual “Use of information and communications technology” survey of Statistics Finland, targeted at households.

Statistics Finland conducts regular statutory firm-level surveys that could be complemented by additional questions in support of the systematic monitoring of the impacts of open data use. Such surveys include the annual “Use of information technology in enterprises” survey and the biennial Community Innovation Survey (CIS). The key issues of interest include the extent and purpose (e.g. innovation, planning of logistics) for which firms use open data and the types of data (e.g. geographic data, company information) firms use in their operations. Surveys could also be employed to assess the role of open data in the establishment of new companies, their operations and job creation.

Open data is often both offered and used by the same public organisation. One way to measure the use of data resources in the public sector would be to examine the amount of data shared by public

administration organisations. The impact of open data could also be assessed by monitoring, as automatically as possible, the extent of data use and the data flows between public organisations. At the level of the national economy, the offering and use of open data could, for example, be measured by investigating the quantity of opened-up data and the number of users. These aggregate indicators would enable monitoring of the opening up of data resources and provide further information on what types of data are the most important to users. The number of new applications and services based on open data also reflects the benefits of open data. However, systematically gathering information on this would be challenging.

Survey data on the use of open data can be used for impact assessment, by combining it with various statistics such as the Business Register of Statistics Finland and the firm-level financial statements database of Suomen Asiakastieto Oy. Long-term monitoring and the compilation of statistics on open data use in various sectors would enable a quantitative analysis, using advanced statistical methods, of the economic impacts of such data. Surveys could also be used to assess the other impacts on society of open data. Qualitative research methods best suit for assessing various other social impacts of open data.

5. Summary

Ex ante evaluations suggest that the opening up of government data has significant potential benefits. Consequently, it has become one of the central policy themes in various countries. Finland aims to open up all important data resources collected and maintained by the public administration by 2020. Research concerning the socio-economic impacts of opening up public data is still in its early stages. Contemporary knowledge of the impacts and best practices of open data stems from individual sectors and is mainly based on case examples. To the best of my knowledge, no countries have so far conducted a comprehensive assessment of the *realised* impacts and benefits of open data. The key reasons for this are measurement difficulties and a lack of statistical or systematically collected information on the use of public data.

Finland is currently among the forerunners in opening up public data, and also has the opportunity to become one of the more advanced countries in terms of research concerning the impacts of open data. However, this will require the careful development of a monitoring and assessment model for the impacts of open data. Furthermore, systematic collection of data is necessary for the empirical investigation of the impacts and best practices of open data use. This report presents tentative guidelines for the assessment of the economic impacts of open data.

In addition to an empirical analysis using the collected impact indicators, user experiences could be gathered from citizens, firms and public sector organisations in order to assess the usability and usefulness of different public data resources. A natural environment for user-driven assessment and feedback would be the open data portal in which public data are made available to users. Additionally, means should be developed of generating information on best practices related to the use of open data, particularly within public sector organisations. Furthermore, there is a need for means of efficiently disseminating such best practices for use within different organisations.

References

Ahonen-Rainio, P., Mäkelä, J. and Virrantaus, K. (2014) Menetelmä avoimen maastotiedon vaikuttavuuden arvioimiseksi. Aalto-yliopisto, Maankäyttötieteiden laitos.

Bargiotti, L., De Keyzer, M., Goedertier, S. and Loutas, N. (2014). Value based prioritisation of open government data investments. European Public Sector Information Platform, Topic Report No. 2014/08.

Brynjolfsson, E., Hitt, L. M. and Kim, H.H. (2011). Strength in Numbers: How does data-driven decision-making affect firm performance? Working Paper Available at SSRN: http://ssrn.com/abstract=1819486.

Capgemini Consulting, 2013; The Open Data Economy. Unlocking Economic Value by Unlocking Government and Public Data.

Davies, T. Perini, F. and Alonso; J. M. (2013). Researching the emerging impacts of open data. ODDC conceptual framework. World Wide Web Foundation.

Dekkers, M., Polman, F., te Velde, R., and de Vries, M. (2006). MEPSIR. Measuring European Public Sector Information Resources. June 2006.)

Fornefeld, M. Boele-Keimer, G.Recher, S. and Fanning, M. (2009). Assessment of the re-use of public sector information (PSI) in the geographical information, meteorological information and legal information services. Micus Management Consulting.

Huijboom, N. and Van der Broek, T. (2011). Open data: an international comparison of strategies. *European Journal of ePractice*, March/April 2011, 1-13.

Kiuru, P., Mäkelä, J. and Huvio, P. (2012). Avoimen julkisen tiedon hyödyntämisen potentiaalista suomalaisissa yrityksissä. Aalto-yliopiston raportti.

Koski, H. (2011). Does marginal cost pricing of public sector information spur firm growth? *Etla Discussion Paper* No. 1260.

Koski, H. (2012). The role of data and knowledge in firms' service and product innovation. *Etla Discussion Paper* No. 1272.

Liikenne- ja viestintäministeriö (2013). Liikenteen ja viestinnän avoin tieto. Työryhmän raportti. *Liikenne- ja viestintäministeriön julkaisuja 10/2013.*
http://www.lvm.fi/c/document_library/get_file?folderId=2497123&name=DLFE-19419.pdf&title=Julkaisuja%2010-2013

McKinsey Global Institute (2011). Big data. The next frontier for innovation, competition and productivity. http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

OECD (2013). *Supporting investment in knowledge capital, growth and innovation.* OECD Publishing. doi: 10.1787/9789264193307-en

Peled, A. (2011). When transparency and collaboration collade: The USA open data program. *Journal of the American Society for Information Science and Technology* 62, 2085-2094.

Pettifer, R.E.W. (2009). PSI in European meteorology – An unfilled potential. Mimeo. 25.4.2009.

Poikola, A, Kuikkaniemi, K. and Kuittinen, O. (2014). My data – johdatus ihmiskeskeiseen henkilötiedon hyödyntämiseen. Liikenne- ja viestintäministeriö. http://www.lvm.fi/julkaisu/4420389/my-data-johdatus-ihmiskeskeiseen-henkilotiedon-hyodyntamiseen.

Pollock, R. (2008). The economics of public sector information. University of Cambridge. Mimeo. December 2, 2008.

Salokannel, M. (2012). Opening-up public data – a spur for democracy and information economy. IPRinfo 2/2012.

Ubaldi, B. (2013) Open government data: Towards empirical analysis of open government data initiatives. *OECD Working Papers on Public Governance*, No. 22, OECD Publishing.

Verhulst, S., Noveck, B.S, Caplan, R., Brown, K and Paz, C. (2014). The open data era in health and social care. The governance lab, NHS, England.

Vickery, G. (2011). Review of recent studies on PSI re-use and related market developments. Information Economics, Paris. https://ec.europa.eu/digital-agenda/en/news/review-recent-studies-psi-reuse-and-related-market-developments