OECD Working Papers on Public Governance No. 22

Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives

Barbara Ubaldi

https://dx.doi.org/10.1787/5k46bj4f03s7-en

OECD

WORKING PAPER

"OPEN GOVERNMENT DATA: TOWARDS EMPIRICAL ANALYSIS OF OPEN GOVERNMENT DATA INITIATIVES"

BARBARA UBALDI

ABSTRACT

Open Government Data (OGD) initiatives, and in particular the development of OGD portals, have proliferated since the mid-2000s both at central and local government levels in OECD and non OECD countries. Understanding the preconditions that enable the efficient and effective implementation of these initiatives is essential for achieving their overall objectives. This is especially true in terms of the role played by OGD in relation to Open Government policies in general.

This paper highlights the main principles, concepts and criteria framing open government data initiatives and the issues challenging their implementation. It underlines the opportunities that OGD and data analytics may offer policy makers, while providing a note of caution on the challenges this agenda poses for the public sector.

Finally, the overall analysis of key concepts and issues aims to pave the way for an empirical analysis of OGD initiatives. So far, little has been done to analyse and prove the impact and accrued value of these initiatives. The paper suggests a methodology comprising an analytical framework for OGD initiatives (to be applied to *ex post* and *ex ante* analysis of initiatives) and a related set of data to be collected across OECD countries. The application of the analytical framework and the collection of data would enable the acquisition of a solid body of evidence that could ultimately lead to mapping initiatives across OECD countries (*i.e.* a typography of initiatives) and developing a common set of metrics to consistently assess impact and value creation within and across countries.

TABLE OF CONTENTS

BACKGROUND 4
Why Open Government Data? 4
The "Right to Information" and "Open Government Data" Movements 4
Definitions and principles 5
Defining Public Sector Information 5
Defining Open Government Data 6
Defining Big Data 6
Setting the principles 7
BEYOND ACCESS: WHAT VALUE? 11
Creating value 11
Value for whom? 11
What value? 12
Improving government accountability, transparency, responsiveness and democratic control 12
Promoting citizens self-empowerment, social participation and engagement 13
Building the next generation of empowered civil servants 14
Fostering innovation, efficiency and effectiveness in government services 14
Creating value for the wider economy 15
Understanding the value chain of open government data 16
THE USE OF OPEN GOVERNMENT DATA 18
New types of actors, roles and relationships for open government data provision in the open OGD value chain 18
The changing role of governments 18
The expanding role of private actors 19
Citizens and civil societies' growing contribution 19
How is open government data being used? 20
ENABLING VALUE CREATION 22
Understanding prerequisites for data openness 22
Paving the way for value creation 22
Are data really open? 24
Conditions for use, re-use and value accrual 25
The role of ICTs and OGD portals 28
KEY DIMENSIONS FOR IMPLEMENTATION 30
Policy challenges 30
Technical challenges 31
Economic and financial challenges 32
Organisational challenges 34
Cultural challenges 35
Legal challenges 37
NEXT STEPS IN THE ANALYSIS 40
Overarching conceptual issues 40
Addressing the emerging needs for governments and public at large 41
Harmonising definitions and sorting out legal "conflicts" 42

Advancing empirical analysis and measurements 44
ENABLING AN EMPIRICAL ANALYSIS.....45
GLOSSARY OF TERMS.....53
REFERENCES.....55

Figures

Figure 1. Related definitions of government data.....7
Figure 2. The contribution of government data to the public service value chain.....23

Boxes

Box 1. The Open Government Data Principles.....8
Box 2. Vivek Kundra's 10 Principles for Improving Federal Transparency.....9
Box 3. United Kingdom's Public Data Principles 10
Box 4. Business Models Archetypes for value creation 16
Box 5. Tim Berners-Lee "Five Star Open-Data Scheme".....26
Box 6. The United Kingdom's "Open Data White Paper": Unleashing the Potential of OGD 31
Box 7. The Case of Regulations.gov.....32
Box 8. Landmark decision from the Netherlands 39

BACKGROUND

Why Open Government Data?

Many public organisations produce and collect a broad range of different types of data in order to perform their tasks. The extraordinary quantity and centrality of data collected by governments make these data particularly significant as a resource for increased public transparency. OGD can be used to help the public better understand what the government does and how well it performs, and to hold it accountable for wrongdoing or unachieved results. This is particularly true as a considerable amount of these government data are progressively becoming more easily accessible and can be used in conjunction with information from other sources (e.g. proprietary information). In addition to increasing government transparency and public awareness of government programmes and activities, opening up data can also help generate insights into how to improve government performance. Increased data transparency provides the basis for public participation and collaboration in the creation of innovative, value-added services. Additionally, data openness is eventually expected to improve the decision making of both governments and individuals. In particular, the public is expected to be able to use government data to make better decision and improve the quality of their lives, e.g. making specific databases easily accessible, such as through mobile apps, to better inform their choices; while governments are expected to be able to more easily access a wider range of datasets to foster evidence-based decision making. Finally, OGD is also seen as an important source of economic growth, new forms of entrepreneurships and social innovation.

However, OGD still remains an uncharted territory. Much untapped potential could be unleashed if government data are turned into OGD. This can only happen if clear boundaries and criteria are set e.g. to protect privacy, and if all restrictions to the access and re-use of authorised data – including legal, financial and technological ones –are overcome.

Furthermore, how OGD initiatives are implemented matters for fostering the public's effective use of government data. Data need to be relevant, easily accessible, usable and re-usable by all. It is important for governments to seek feedback from the public on the usefulness, relevance and accessibility of their data, in order to allow for continuous improvement. Enhanced data accessibility can enable higher collaboration within governments, as well as between government agencies and the wider society, including the private sector, civil society organisations and citizens. This is spurring a shift in public sector organisational culture not only towards openness, transparency and accountability but also to sharing, collaborating and higher public engagement.

The "Right to Information" and "Open Government Data" Movements

There are two main civil society movements that are campaigning for greater openness of information, documents and datasets held by public bodies. The first is the "Right to Information" movement, which promotes a public right of access to information from a human rights perspective[1]. The second is the "Open Government Data" movement, which uses predominantly social and economic arguments to encourage the opening up of government data[2]. The latter claims that putting such information into the public domain can benefit society by creating conditions for more social inclusive service delivery and for more participatory democracy. They also argue that it can stimulate the economy

by allowing the possibility for third parties (e.g. individuals, private enterprises, civil society organisations) to create new products and services using public data.

There is significant overlap between both movements, in that both aim to increase the transparency of government so that all members of society can enjoy the inherent social and economic value of information that has been generated and collected with public funds. Nevertheless, there are also differences in the approaches and strategies employed by each. One of the main differences is that right to information advocates place emphasis on access to qualitative as well as quantitative information, which is often stored in the form of documents, whereas open government data advocates focus on data that are held in government databases, and they are concerned with both the technical and the legal issues related to the access, use and re-use of these datasets. Therefore, while the right to information movement has put an emphasis on the obligation of public bodies to respond to requests for information, the open government data movement emphasises proactive release of large volumes of information in formats and under conditions that permit re-use.

Definitions and principles

From a conceptual point of view data can be seen as the lowest level of abstraction from which information and then knowledge are derived[3]. Information is considered as the "lifeblood of a robust democracy and productive economy" (Office of the Australian Government, 2013), and governments and societies have increasingly come to appreciate the relevance of accessible public sector information to sustain good governance and stimulate knowledge-driven economic growth. This situation has changed the way in which governments have started valuing, managing, using and sharing information both inside and outside their boundaries; and this shift has been supported by the upgrading of existing policies, structures and practices to produce, manage and use public sector information, or by the adoption of new ones. The term "Public Sector Information" (PSI) is the one that best captures the concept at the heart of this shift and that has inspired the "FOI" movement described in the previous section of this paper.

As the focus has increasingly been placed on the potential for economic and social value, which can be derived from PSI, the emphasis on the accessibility of the data underlying reports, analysis and information has augmented. First strongly supported in the private sector and market economy environment, the concept of Open Data has also been picked up by the public sector and has stirred the Open Government Data Movement described earlier in this paper. In general, adopted definitions of "open data", which apply both to data in raw and processed forms, do not offer full insights on what data are, but focus rather on the issue of openness and re-use. Therefore, what comprises a government data set is still an open discussion, and the understanding of what OGD are varies.

Big Data as a phenomenon has emerged as available datasets produced by various sources have grown larger and as the data users have become more aware of the value obtainable through the combination and linkage of different data sets produced by different actors, both private and public.

The sections below provide definitions of the terms "public sector information", "open government data", and "big data". Each of them is grounded on the relevance attributed to data, information and knowledge as a resource for the economic, social and democratic development of our societies.

Defining Public Sector Information

For the purposes of the OECD "Recommendation for enhanced access and more effective use of Public Sector Information (PSI)" (OECD, 2008) the latter is broadly defined as "information, including information products and services, generated, created, collected, processed, preserved, maintained,

disseminated, or funded by or for a government or public institution", taking into account the legal requirements and restrictions referred to in the last paragraph of the preamble of the Recommendation.

Defining Open Government Data

The term "Open Government Data" (OGD) has come into prominence relatively recently, becoming popular in 2008 after the publication of a set of open government data principles by advocates in the United States (see Box 1).

The two main elements of OGD are normally defined as follows:

- Government data: is any data and information produced or commissioned by public bodies.
- Open data: are data that can be freely used, re-used and distributed by anyone, only subject to (at the most) the requirement that users attribute the data and that they make their work available to be shared as well.

For the purpose of this paper, public datasets being considered as a reference for analysis concerning OGD initiatives include:

- Business information (including Chamber of Commerce information, official business);
- Registers, patent and trademark information and public tender databases;
- Geographic information (including address information, aerial photos, buildings, cadastral information, geodetic networks, geology, hydrographical data and topographic information);
- Legal information (including decisions of national, foreign and international courts, national, legislation and treaties);
- Meteorological information (including climate data and models and weather forecasts);
- Social data (including various types of statistics on economics, employment, health, population, public administration);
- Transport information (including information on traffic congestion, work on roads, public transport and vehicle registration).

Well into the 2000s, the concept of "open government" among public officials was still centred on fresh disclosures, rather than improved access to data that was already in the public domain. In the past decade there has been a shift from a commitment prevalently focusing on Open Government's goals of higher transparency and public accountability to an increasing pledge to OGD with an agenda strongly driven also by innovation, efficiency and flexibility in government (Robinson, Yu 2012). To date, governments, civil society organisations and private sector representatives, consider OGD as a building block for open government, as they see it as a key enabler of improved service delivery, transparency and public engagement and, as a result of better relations between governments and citizens.

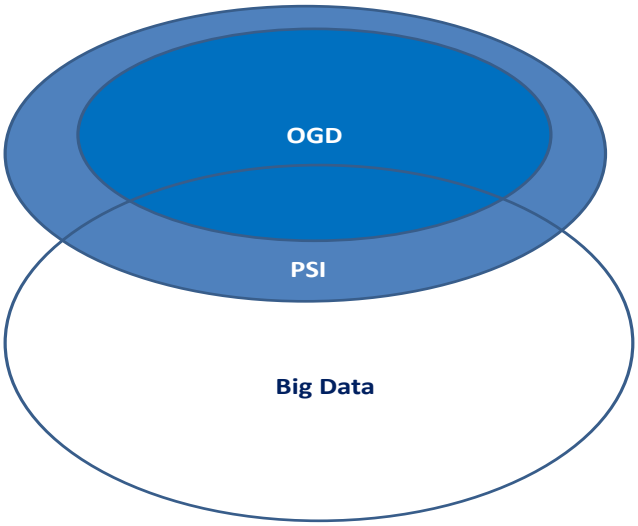
Defining Big Data

Governments are no exception in the modern world tendency to generating a staggering amount of data and larger datasets. Across the public sector, extraordinary quantities of data are amassed in the course

of delivering public services, from managing welfare payments and national health services through issuing passports and driving licences.[4] While economic and social activities have long revolved around the use of data, the significant volume, velocity and variety of data increasingly being used across the economy, and the important social and economic value of this data, signal a shift towards a data-driven socio-economic model, commonly referred to as "big data". In this model, data are a core asset for creating significant competitive advantages and for driving innovation, sustainable growth and development (OECD, 2012b).

"Big data" is a collection of datasets so large and complex that it is difficult to use on-hand database management tools, or traditional data processing applications, for their processing that includes capturing, storage, search, sharing, transfer, analysis, and visualization. The trend towards having and dealing with larger data sets produced by different actors is also due to the additional information derivable from the analysis of a single large set of related data, as compared to separate smaller sets with the same total amount of data, allowing correlations to be found to spot business trends, determine quality of research, prevent diseases, link legal citations, combat crime, and determine real-time roadway traffic conditions.

Figure 1. Related definitions of government data



Source: Author.

Setting the principles

In order to support and guide the development of OGD, different set of principles have been put forward by several non-governmental organisations. Many of these principles have been picked up by governments to shape and guide their OGD initiatives.

Eight Open Government Data Principles were defined and put forward for governments' consideration[5] in December 2007, during an Open Government Working Group Meeting held in Sebastopol (California, United States), which gathered 30 open government advocates and was organised by Public.Resource.Org, with sponsorship from the Sunlight Foundation, Google and Yahoo.

Box 1. The Open Government Data Principles

Government data shall be considered open if it is made public in a way that complies with the principles below:

1. **Complete:** All public data are made available. Public data are data that is not subject to valid privacy, security or privilege limitations.

2. **Primary**: Data are as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.

3. **Timely:** Data are made available as quickly as necessary to preserve the value of the data.

4. **Accessible**: Data are available to the widest range of users for the widest range of purposes.

5. **Machine processable**: Data are reasonably structured to allow automated processing.

6. **Non-discriminatory:** Data are available to anyone, with no requirement of registration.

7. **Non-proprietary**: Data are available in a format over which no entity has exclusive control.

8. **License-free**: Data are not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.

Source: http://resource.org/8_principles.html, adopted in December 2007.

The Sebastopol list has since then been updated by the Sunlight Foundation and ten principles have been identified that are meant to provide a lens to evaluate the extent to which government data are open and accessible to the public. The principles are completeness, primacy, timeliness, ease of physical and electronic access, machine readability, non-discrimination, use of commonly owned standards, licensing, permanence and usage costs.

On 14 July 2011, then United States Federal Chief Information Officer Vivek Kundra identified ten principles for improving federal transparency in his testimony before the House Committee on Oversight and the Government Reform Subcommittee.

Box 2. Vivek Kundra's 10 Principles for Improving Federal Transparency

1. **Build end-to-end digital processes**: Automate transfer of data between systems to increase productivity, protect data integrity, and speed data dissemination. Capitalize on game-changing technologies to increase transparency.

2. **Build once, use often**: Architect systems for re-use and share platforms to reduce costs, streamline systems and processes, reduce errors, and foster collaboration.

3. **Tap into golden sources of data**: Pull data directly from authoritative sources to improve data quality, shorten processes and protect data integrity.

4. **Release machine-readable data and encourage third party applications**: Make data machine-readable to allow the public to easily analyse, visualise and use government information.

5. **Use common data standards**: Develop and use uniform, unique identifiers and data standards to ease the flow of data and reduce system complexity.

6. **Validate data up front**: Correct errors during collection and at the point of entry to block bad data from ever entering the system.

7. **Release data in real time and preserve for future use**: Release data as quickly as feasible to enhance its relevance and utility while maintaining future accessibility.

8. **Reduce burden**: Collect data once and use it repeatedly. Pull from existing data sets to reduce costs and burden and to increase productivity and uniformity.

9. **Protect privacy and security**: Safeguard the release of information to increase public trust, participation, preserve privacy, and protect national security. Open Government doesn't mean vulnerable government.

10. **Provide equal access and incorporate user feedback**: Provide a common view of data to all stakeholders to foster collaboration. Incorporate user feedback to help identify high-value, meaningful data sets, set priorities, to continuously drive and improve future planning and processes.

Source : http://oversight.house.gov/hearing/transparency-and-federal-management-it-systems

In June 2012, the United Kingdom Public Sector Transparency Board published the Board's Public Data Principles that have been the foundation for guiding the public sector in publishing data.[6]

Box 3. United Kingdom's Public Data Principles

1. Public data policy and practice will be clearly driven by the public and businesses that want and use the data, including what data are released when and in what form.

2. Public data will be published in re-usable, machine-readable form.

3. Public data will be released under the same open licence which enables free re-use, including commercial re-use.

4. Public data will be available and easy to find through a single, easy-to-use, online access point (*www.data.gov.uk*).

5. Public data will be published using open standards, and following relevant recommendations of the World Wide Web Consortium (W3C).

6. Public data from different departments about the same subject will be published in the same, standard formats and with the same definitions.

7. Public data underlying the Government's own websites will be published in re-usable form.

8. Public data will be timely and fine-grained.

9. Release data quickly, and then work to make sure that it is available in open standard formats, including linked data forms.

10. Public data will be freely available to use in any lawful way.

11. Public data will be available without application or registration, and without requiring details of the user.

12. Public bodies should actively encourage the re-use of their public data.

13. Public bodies should maintain and publish inventories of their data holdings.

14. Public bodies should publish relevant metadata about their datasets and this should be available through a single online access point; and they should publish supporting descriptions of the format provenance and meaning of the data.

Source : *http://data.gov.uk/library/public-data-principles*.

BEYOND ACCESS: WHAT VALUE?

Creating value

As the amount of data available in today's Information Economy explodes (The Economist 2010; OECD 2012d), public agencies are trying to increase the transparency of government processes and performance by publishing relevant data online and sharing it with the public. Government data, made available in machine-readable, linked datasets that can also be searched and manipulated using standard tools, is a critical new resource for fuelling changes in value creation (economic, social and political).

The economic and social case for open government data are being established with increasingly solid evidence from many national and local governments, *e.g.* the United Kingdom and many United States cities including San Francisco. However, potential value creation is also expected in political and social terms.

In dealing with the creation of value, it is essential to identify the main beneficiaries of the value being created.

Value for whom?

Public sector information is a strategic resource, holding great potential for a number of beneficiaries including public sector agencies, private businesses, the academia, citizens and civic organisations. In many countries, the open government data community still appears to be unco-ordinated. It includes IT professionals, both small and large companies, and entrepreneurs as well as developers, government employees, civil society organisations and individual citizens active on the national or international level. Their motivations, level of understanding of government processes and structures, scope and priorities in advocating for OGD differ. Entrepreneurs stress clear conditions for re-use and reliable licensing. Programmers demand raw data. Transparency activists want access to internal government documents. Individual citizens might not be interested in data *per se*, but in secondary information-type products (*e.g.* new services and mobile apps). Civil society organisations are keen to have a number of datasets that, if combined, may help improve life and service delivery to certain segments of the population or to certain neighbourhoods.

The following is a list of commonly recognised main beneficiaries of OGD:

Government: At a macro level, OGD provides the scope for new ways of conducting "government business", taking decisions and allocating resources, in order to improve the overall efficiency of government operations (*e.g.* accelerate efforts to reduce fraud and error, make further inroads into the tax gaps) and more effectively and efficiently deliver smarter, innovative and more personalised public services, while improving the quality of interactions between the governments and the users. OGD can equally enable an increase in transparency that --can strengthen accountability. Additionally, by being more open, governments gain in terms of legitimacy vis-à-vis citizens and civil society.

Citizens: Open government data are expected to enable public participation and social engagement in designing responses to public needs (*e.g.* co-development and co-production of services through newly developed apps); and to allow the sourcing information and knowledge from more diverse sources (e.g. crowd sourcing). Innovative service delivery is emerging from "mashing up" sets of data that originate from various sources, and by various parties. "Fix My Street" in the United Kingdom and Chicago's "311" Internet portal illustrate the intersection of mobile government and OGD. One is built by citizens, the other by government. Opening up public sector data (*e.g.* crime rates, gas emissions, teachers per student in city schools) is geared also towards allowing citizens to make more informed personal choices. As such, OGD has the potential for helping improve citizens' quality of life.

Civil Society: Civil society initiatives that build on OGD can be found across many OECD member countries. The shared goals of these initiatives include demonstrating the benefits of OGD to government and to the public. However, there is fragmentation in terms of specifically targeted goals that range from increasing transparency to improving service delivery – particularly for vulnerable segments of the population – to protecting the environment or sustaining growth; and it has been observed that this fragmentation does not help to move items higher up national political agendas. Additionally, civil society organisations can play a pivotal role as intermediaries in the identification of key datasets that, if open, could produce high value. This is particularly true for vulnerable segments of the population they engage with. Widely known examples include the Sunlight Foundation and the Open Forum Foundation in the United States, the Open Knowledge Foundation in Germany, the Open Rights Foundation in the United Kingdom, etc.

Wider economy, private sector and public service market place: OGD can stimulate a competitive marketplace, for instance for public sector services. The common belief is that innovators from outside governments are provided with the opportunity to develop modular services which are more agile and targeted to users' needs than those developed in-house by governments. When government data are open, however, access to data per se does not provide a competitive advantage to firms with exclusive data-access agreements. Competitive advantage has to come from offering innovative value-added services on top of data, and providing opportunities for business start-ups. The private sector (technology developers) is expected to be amongst the primary users of datasets to pursue commercial exploitation of OGD. A profit incentive can help to drive innovation and experimentation, while one would expect the best ideas to be emulated and improved upon as no one service provider has the monopoly on data.

What value?

The sections below highlight some of the main values targeted by OGD initiatives across OECD member countries, and which may simultaneously benefit several of the beneficiaries identified in the previous section. The potential benefits are not only envisaged in monetary and economic terms, but also from social and good governance perspectives. Understanding the different values is essential to identify which enablers and type of data to prioritise in order to achieve the targeted value creation goals. Different values require different types of data. For instance, while many accountability and good governance objectives can be served by releasing aggregated data, boosting economic growth may require specific datasets to be released to the business community or apps developers at a more granular level data, in a timely manner and updated regularly to diffuse them vastly and rapidly in order to create business opportunities.

Improving government accountability, transparency, responsiveness and democratic control

Strong supporters of OGD argue that it is a key enabler of Open Government, promoting transparency and public accountability. For instance, a common assumption is that the absence of data in the public domain allows public servants to engage in corrupt behaviour with impunity. OGD can be a powerful force

for public accountability by making existing information easier to analyse, process and combine, allowing for a new level of public scrutiny. This can raise the level of public trust and perceived responsiveness of government actions. The Open Government Declaration[7] is considered to have situated the use of ICTs, and new technologies in particular, to spur data sharing in the context of political accountability, thus blurring the distinction between the technology of open data and the politics of open government.

However, it is important to underline that open government and OGD can each exist without the other, and increased transparency is not an automatic driver of greater accountability. A government can be open, in the sense of being transparent, even if it does not embrace new technology, and a government can provide open data and still remain deeply opaque and unaccountable (Robinson, Yu, 2012). Making public sector data available in machine-readable format can improve service delivery and citizens' quality of life, while having little impact on political accountability.

Promoting citizens self-empowerment, social participation and engagement

A second assumption is that OGD enables individuals to make better decisions in their lives and increases participation in public affairs. Normally, e-participation is part of a government's broader e-government policy aiming at harnessing IT use for openness, transparency and collaboration within the public sector, but also to increase citizen engagement in public life, *e.g.* in policy making and service design and delivery. The expanding use of new technologies, combined with the emergence of the OGD movement, are becoming key enablers and drivers of citizen self-empowerment, higher e-participation and public engagement.

OGD initiatives, particularly as they are supported by Web 2.0 and social media applications, are creating an architecture for participation that enables users to not only be passive consumers of content and services, but also active contributors and designers in their own right. This means that stakeholders are invited more openly into a participative and empowering relationship with the government in relation to a number of matters which are part of the overall democratic process, such as:

- working arrangements of the public sector, as well as of public governance more widely;
- planning and land use issues;
- service design and delivery;
- community building; and
- dispute and conflict resolution, broader public policy and decision making.

OGD, and in particular data smart disclosure (*i.e.* the timely release of data in standardised, machine-readable formats in ways that enable users to make better decisions about finance, healthcare, energy or other personal contexts) empowers citizens to take more informed decisions that can enhance the quality of their lives. For example, the health and consumer service departments in the United States have pushed for the smart disclosure of data on flights operated by national airlines to enable people to make informed choices on the airline company selection. To support this, governments need to enable users to have access to their own data and decide how to use it to improve lives. Another example is the Blue Button Initiative in the U.S. which gives veterans complete control of their personal health record held by the public sector; and the Green Button, also in the U.S., which is a similar initiative around individuals' energy use data.

Building the next generation of empowered civil servants

Equally important to empowering citizens is empowering the public sector workforce. Opening up government data can enable civil servants, many of whom are frontline professionals, to participate directly in ensuring that government is open and participative, and to develop applications that better respond to users' needs. Many civil servants see the real time performance and impact of public services and public policies on citizens, and would be able to generate appropriate data and other inputs, or use available ones, to improve service experience if they were given the tools and incentives to do so (for example, by being enabled to participate in a professional capacity in online social networks to offer advice and knowledge to the public).

Moreover, many civil servants also see a blurring of their personal and professional lives in terms of the tools they use. This fact could improve their performance through the two-way exchange of experience and skills. Sensible structures are needed to ensure that civil servants are empowered in this way while they are also able to retain impartiality and a position of trust both from the government itself as well as from citizens. This also requires that civil servants be equipped with the necessary skills, tools, mechanisms and guidelines (Millard, 2012).

Empowering civil servants with OGD requires strategies and programmes to build the next generation of civil servants. New skills are needed, which are not only strictly IT related. They should include: data science; predictive analytics to identify patterns and create models; a better knowledge on how to use web 2.0 technologies for social engagement and to negotiate and connect to people; and a finer understanding of emerging problems and of the use of IT to solve them (*e.g.* cybercrime investigation). As an example, the Dutch police force, in collaboration with Deloitte and a forensic consultant (ForensicPlaza), co-produced a programme known as "Awareness & Digitalization". This programme is intended to provide a flexible and innovative way to raise the police force's awareness of risks and opportunities in the cyber environment and to develop their skills in dealing with the emerging challenges of an increasingly digitisalised society.

Fostering innovation, efficiency and effectiveness in government services

OGD has the potential to increase government efficiency, effectiveness and innovation in service delivery and internal public sector operations. Even though the release of government data online can raise a number of substantive enquiries in terms of government activities that require time to be addressed, from a service delivery perspective data re-use can also lead to a significant decrease of the questions routinely received by public authorities, or enable questions to be answered more quickly. Additionally, the remaining questions concerning service delivery *per se* would be easier for civil servants to answer as the relevant information would be easier to find. Subsequent benefits include reduced workload, a reduction in paperwork and lower transitional costs. Services are also improved as people more easily find and claim the benefits they are entitled to, and public sectors can tune front-line services more closely to individual needs and behaviours. The Dutch department for cultural heritage is, for instance, actively releasing their data and collaborating with amateur historical societies and groups such as the Wikimedia Foundation in order to execute their own tasks more effectively. This can result in improvements in the quality of data, while encouraging external inputs and new sources of knowledge, possibly making them more innovative and comprehensive. In addition, one could argue that the co-development of knowledge in this case increases not just the quality, but also the awareness of the Dutch public authority's work, thereby further increasing its value and relevance. Similarly, the Bristol City Council reduced the typical service transaction cost (15 times less expensive) by introducing the Open Government Data Catalogue[8].

OGD can also help foster collaboration across and within public agencies and departments. As shared datasets and/or registers are being created, collaboration and exchange on who owns what public

information and for what purpose is needed, which provides an opportunity to also re-engineer and simplify internal procedures, and/or automate processes and as a result eliminate redundant expenditures or reduce internal transactional costs. Furthermore, as public resources are freed from having to maintain individual registers and datasets they can be reallocated to more productive tasks.

Moreover, OGD provides a platform for innovative service delivery. This can not only result from the re-use of data by private sector actors or by civil society organisations, but also thanks to the re-use of data by civil servants, who have in several instances taken the initiative to develop new mobile applications. The use of technology and data analytics within the public sector, and the integration of analysis in policy making and design of public service, can boost more integrated and innovative service delivery.

Finally, improved service delivery can also emerge as a result of initiatives originally driven by a government's push for greater transparency around data. In the UK, for instance, discussions between the Department of Transport and Trade Association for Train Operating Companies (ATOC) led to free access to the fares databases in early 2013. This enabled the development of price comparison websites and mobile phone apps for the benefit of passengers who can access accurate information and find the ticket they need at the best price[9].

Creating value for the wider economy

It is now widely considered that one important potential benefit of OGD is the income that can be generated by commercialising government data. According to some, there is evidence that reduced pricing (*e.g.* allowing non-commercial re-use at zero cost, and reducing the charges for commercial use) significantly increases the use of OGD and its direct or indirect application across the economy, which may lead to an increase in the revenues generated from it, or in job creation in new sectors such as infomediaries (Capgemini Consulting, 2013) or thanks to increased innovation.

According to a survey conducted by the European Commission (EC) in 2006, the overall market size for Public Sector Information in the EU is estimated to range from EUR 10 to EUR 48 billion, with a mean value around EUR 27 billion and these amount to 0.25% of the total aggregated GDP for the EU and Norway[10].

Recent economic analyses have additionally shown that when information is provided to the public free of charge or at very low cost, individuals, developers and private enterprises are more likely to take that information and create added-value products that they can then market[11]. This can increase the volume of private sector activities which can stimulate the national economy and also provide revenue to the government in the form of taxes.

In addition, data made available free of charge from public bodies can be used for civil society projects. This is particularly true for data that have broader potential value include mapping, meteorological, legal, traffic, financial, and economic data. Much of this raw data are used for, or integrated into new data-enabled products, apps and services that individuals use on a daily basis, such as car navigation systems, weather forecasts, or financial and insurance services[12]. According to a Finnish Study, businesses that re-use geographical data grew 15% more per annum in countries where governments released such information freely, compared with countries that price such information in order to recover production costs[13].

In Spain, the Aporta project encourages the re-use of public sector information by providing access to over 650 government datasets in reusable formats through a public data portal, datos.gob.es. Spain estimates that there are over 150 companies that work solely on the infomediary sector and employ around 4 000 people generating 330-550 million Euros annually that can be directly attributed to open data re-

use[14]. This is an important stimulus for Spain from the ICT sector, which is a bit of a beacon in otherwise dire economic times (OECD, 2010; OECD, 2012).

Although the potential economic benefits which may emerge arise from OGD are evident, these are still estimates and there is little empirical data. Much still needs to be proven to show economic impact as OGD implementation is still relatively new and difficult to measure[15]. Despite the challenges, it is essential to shed light on this aspect in order to assist governments in weighing investments against future benefits.

Box 4. Business Models Archetypes for value creation

Based on the analysis of a number of companies, Deloitte identifies different business models to use data to generate revenue, which can be clustered around five archetypes:

- **Suppliers** that publish data, including the public sector;
- **Aggregators** that pool publicly available data and combine it to produce useful insights to be used by the various users;
- Apps **developers** that enable users to make more informed decisions (*e.g.* apps building on crime data, transport data);
- **Enrichers** that are large and established businesses producing significant quantities of open data and combine it with their own proprietary sources to provide services (*e.g.* insurers, retailers); and
- **Enablers** that are organisations that don't make money out of open data but provide platforms and technologies that others can use (e.g. websites that enable data sources of all types to make subsets of their data available to seek solutions from the public.

Source : "Open growth. Stimulating demand for open data in the UK."

Understanding the value chain of open government data

Understanding the OGD value chain is essential to grasp the role of the main actors, and to understand how the speed at which the context is evolving and causing significant change in the roles and boundaries of the respective responsibilities and is affecting the actors'involvement in various data related activities.

Similarly to the value chain of public sector information identified in relation to the European Commission PSI Directive,[16] four main phases can be identified in the OGD value chain:

1. **data generation**: this covers the generation of public data, which is normally done by public sector entities even if this function may increasingly become shared with other publicly funded data sources (e.g. social statistics, aerial data).

2. **data collection, aggregation and processing:** data need to be collected, gathered to enable access, sharing and re-use. Most government data also need to be aggregated, linked, and or manipulated once open in order to add value for the majority of users (*e.g.* to support personal decisions). Many users would in fact not be able to understand and make use of the data 'as it is'. This applies in particular to non commercial re-use of government data.

3. **data distribution and delivery:** data need to be distributed to the potential users to enable access and re-use.

4. **final data use:** open government data need to be re-used by a whole list of different users to sustain public value creation.

The next section of the paper underlines how the roles of the various actors are changing. Understanding these changes and the new dynamics is essential as they can have important policy implications and can impact implementation (*e.g.* on sharing and integrating data sources, data usability, stimulating demand and use, and creating the right ecosystem).

THE USE OF OPEN GOVERNMENT DATA

New types of actors, roles and relationships for open government data provision in the open OGD value chain

The changing role of governments

The interests and roles of the different OGD actors can be better understood when conceived as part of a value chain of public value creation. Rather than assuming that greater openness automatically creates value, this approach focuses on the steps leading to the creation of public value, including the preconditions to unlock such value. Government, particularly at the local level, have a critical role to play as convener, facilitator, enabler, partner, and participant of OGD. Public agencies produce data of various types and quality, although many times their focus lies on increasing the openness and provision of information or documents rather than on opening up the data underlying the production of information, documents, etc. In many countries, agencies also tend to treat their data as a commodity to generate revenue or to cover the cost of data production. Typically data production is at the core of public tasks and it is often backed up by a legal monopoly (Dekkers, G. *et al* 2006). The public sector is likely to continue to be the main funder of OGD initiatives, whether or not it itself designs and delivers them. Government remains the only legitimising organisation with democratic accountability to act fully in the interests of the whole of society and to manage the risk of providing data of poor or erroneous quality. As the main producer of public sector information and data, governments will likely retain responsibility for overall quality standards.

Nevertheless, the distinctions between professional, politician, practitioner, civil servant, expert, consumer and citizen are blurring dramatically. These roles are still important but the relationships between them are changing, and any individual may play several roles at the same time. In relation to government, this means that many stakeholders can and are becoming involved in areas of competence that were previously the preserve of the public sector or of specific agencies alone. Instead of always being the sole actor, the public sector is increasingly becoming just one player in a new form of "open-source governance" in which it may often only play the role of arbiter, co-ordinator, funder and regulator for the activities of others in delivering public value through the use of public sector information and data.

Therefore, in order for government data to benefit from the same innovation and dynamism that characterise private party use of the Internet, it is critical for government to rethink its role as data provider and realise that pushing out data is not enough to create value. As mentioned earlier, many are currently arguing that it would be more efficient for the public sector to focus on the following rather to struggle, as it currently does, on designing sites that meet each end-user need:

- Create a simple, reliable and publicly accessible infrastructure that "exposes" the underlying data;
- Identify relevant datasets through public consultation;
- Increase data release in formats that enable and foster re-use.

The expanding role of private actors

Many private actors, either non-profit or commercial, are well suited to deliver government information to citizens in ways and forms that they want and that create value for them. These private actors have the capacity to take the government data "as is" and to provide it to final users in formats which enable a valuable use. These private sector entities constantly create and reshape the tools that individuals use to find and leverage public data. Their role can range from "repackaging" data to proposing various forms of value-added analysis in response to citizen demand.

Based on the engineering principle of separating data from interaction (*e.g.* commonly used in constructing websites or Service Oriented Architectures[4]), many private sector players argue that private parties are best equipped to build websites providing interactive access for the public. This approach is especially important given recent advances in interaction, which go far beyond merely offering data for viewing, to providing interactive services such as: advanced search, RSS feeds, links to information sources, cross indexing and mash-ups with other data sources, automated content and data analysis, discussion forums and wikis, collaborative filtering and crowd sourcing analysis.

These tools and opportunities for exploring new forms of private-public partnerships and collaboration are promising, but it is far from obvious how best to combine them to maximise the public value of government data and to tackle the risk related to data quality. Given this uncertainty, rather than deciding on the best option up front, governments can stimulate private party interest to engineer ideas and experiments to discover what works in different situations.

An important way to ensure that the government allows private parties to compete on equal terms in the provision of government data is to require that national open data portals themselves use the same open systems for accessing the underlying data as they do to make them available to the public at large.

Some people might want government to continue to present data because they want access to "genuine" data, unmediated by any private party. As long as there is vigorous competition between third party sites, however, expectations are that most citizens will be able to find a site provider they trust. Where it is necessary for a third party site to convince a user that a unit of government data is genuine, this can be accomplished in various ways, *e.g.* by using digital signatures for data so that third parties can qualify themselves as authorised government data providers and prove their identity.

Private actors – including businesses and non-profit organisations – have demonstrated a remarkably strong desire and ability to make government data more available and useful for citizens often by going a long way to reassemble data that government bodies already possess, but that are not sharing in a machine-readable form. An important advantage that can be derived by involving private parties to present data to citizens may arise from encouraging the emergence of more advanced features beyond the simple delivery of data to stimulate use, social engagement and collaboration.

Citizens and civil societies' growing contribution

Citizens, communities, civil society organisations, and businesses are themselves changing from passive consumers to active producers of data. For example, citizens share and consult more and more with each other rather than with the public sector on advice to organise and improve their lives, as well as tackle specific daily problems. It is therefore important for governments to recognise the value of crowd sourcing to enable real time data and information sharing, and to engage relevant actors outside public organisations to use them to create value.

An important new role for governments is therefore is to act as an open participation and collaboration platform, both online and offline, allowing stakeholders to contextualise data provision and thus increase the potential for value creation. Governments can promote collaboration by pro-actively working with companies including SMEs, civil society organisations, local communities, groups, people and hackers. Resources provided by the government itself, and that could be elicited from other stakeholders, include data, applications, knowledge, content, capacity and service building blocks. Encouraging their use through discussion forums, blogs, consultation, support and advice, brokerage, good practices, and arbitration, is essential to produce value. Standardised modules for basic functionalities that are cross institutional so that users do not have to contend with unnecessary differences, but which can easily be used, re-used and combined in new ways to address specific needs, are key. Governments thus increasingly need to crowd source content, services and policies. In a fast-changing world, and in light of continuing times of austerity and fiscal consolidation, complex societal problems can no longer be solved by the government alone (the visible hand), or by the market alone (the invisible hand); rather, partnerships and groups (many hands) are needed (Felten E. W. *et al*, 2009).

These developments, although largely beneficial, also give rise to threats and challenges that need to be tackled, for example:

- loss of control and blurred accountability (data by whom, to whom?)
- quality standards that are more difficult to determine and maintain and an increased risk for data quality (risk of (outdated, inaccurate or incomplete data)
- privacy and data security
- danger of data and content mis-use, propaganda and subversion
- digital elite formation (creation of new digital divides?)
- information overload (or is this more a filter failure?)
- crowding out of other, more relevant, channels
- inadequate capacities within the public sector, and the society at large, to conduct data analytics and exploit data potential
- timely provision of data and updates to data.

How is open government data being used?

One can identify five distinct processes of OGD use (Davies, 2010):

- **Data to fact**: often underestimated individuals may seek out specific facts in a newly open dataset. These facts may support their engagement in civic or bureaucratic processes, or in business planning, or may inform personal choices. Facts could be found through online interfaces, but also by browsing downloaded Excel spreadsheets.
- **Data to information**: creating a static representation and interpretation of one or more data sources, information can take the form of visualisations, blog posts, info-graphics or written reports.

- **Data to interface**: creating a means to interactively access and explore one or more datasets. For example, creating a searchable mapping mash-up, or providing a tool to browse a large dataset and crowd source feedback or scrutiny. Interfaces often also include "static" interpretations of data (data to information), showing particular summary statistics or algorithmically derived assessments of underlying data.

- **Data to data**: sharing derived data (either simply an original dataset in a new format, or data that is augmented, combined with other data, or manipulated in some way). A whole dataset may be shared; an Application Programming Interface (API) onto a dataset created; or an interface that makes it easy to download subsets of a large dataset. API may also enable the sending of data updates to users in real time, avoiding the need to re-download entire files or to search for new entries.

- **Data to service**: services where OGD plays a "behind the scenes" role either online or offline. For example, the use of boundary data to route messages reporting potholes to the responsible authority.

These processes of OGD use are not mutually exclusive, and many OGD users employ multiple processes.

By mapping out the path from original OGD to end-use in different cases, certain trends, such as the frequent caching of bulk datasets, become visible, raising questions about how well current OGD infrastructures and patterns of use will cope with updates to original datasets.

ENABLING VALUE CREATION

Understanding prerequisites for data openness

Making government data truly open is needed to produce the desired value; and doing it requires a holistic approach as it involves the many stakeholders mentioned in the previous section. Lee and Kwak (2011) recommend that the implementation of OGD initiatives should be incremental because each stage is important. Accordingly, they propose a model that requires incremental implementation of OGD initiatives. The model presents four stages of implementation before government data is made fully open.

1. Stage Zero: Getting a view of what government data exists;
2. Stage One: Increasing data transparency;
3. Stage Two: Improving open participation;
4. Stage Three: Enhancing open collaboration;
5. Stage Four: Realising ubiquitous engagement.

At each of these stages, the level of engagement and participation increases as one progresses to the next stage. Fundamental to this model is data transparency as the essential stage for OGD; the other stages are dependent on available data in formats that enable the realisation of the subsequent stages. This seems to point to the need for countries to implement comprehensive OGD projects in succession as they move towards higher levels of engagement and civic participation through the use of ICT. Successful implementation of OGD can only happen if supported by an appropriate approach, *i.e.* one based on a good understanding and awareness of the OGD requirement and implications. For example, recognising the relevance to engage all affected stakeholders and to create the ecosystems to spur collaboration with various communities as preconditions for reaping the potential benefits of OGD, is a prerequisite for the progression from “data transparency” to “open participation” and “open collaboration”.

Paving the way for value creation

In order to ensure that OGD actually creates value, the most important tasks facing governments appear to be:

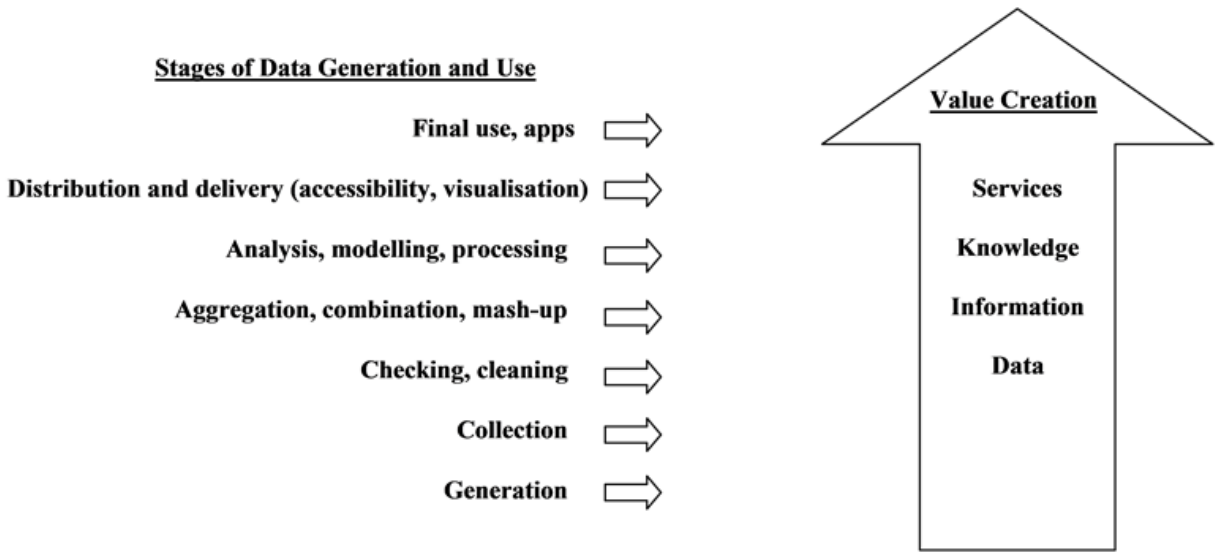
- identifying high-value, high-impact data for the public;
- improving and assuring data quality in terms of accuracy, consistency, and timeliness;
- fostering data demand and use by the various actors.

Creating value out of OGD implies a good understanding of both data users' need – in terms of content and format – as well as how data are created, protected, shared and used. Different sets of data that have proven of great value if made open include:

- public data (*e.g.* trade data, fiscal data, health data, education data, transport data, weather data, census data, map data/geographic data, crime data);
- micro-statistics: data the state needs to function itself;
- performance data on government programmes and spending (*e.g.* data on mortality for a doctor or in public hospitals);
- personal data of public services' users; and
- public cultural information.

Creating value from OGD means, however, also addressing the issue of escalating the quantity of public digital records (*e.g.* through digitisation of public records and e-archiving). Similarly, exploring and understanding the process barriers to data release at the various levels of government, data ownership (*e.g.* where are the data located? Who owns it?), the role of various actors and assessment of the return on investment are all key factors to enable value creation through OGD. Finally, understanding and fostering data use is critical to create value, as well as creating an ecosystem that supports a collaborative data creation process.

Figure 2. The contribution of government data to the public service value chain



Source: Author and based on OECD (2006), Digital Broadband content: Public Sector Information and content.

Are data really open?

It is widely recognised that in order to be considered open, government data need to be available to anyone who wants to access and re-use it and should be provided with the most complete detail available, and in a timely manner. To achieve this, data need to be available and accessible, on the one hand, and authorised for reuse and redistribution, on the other.

In order to be **available and accessible**, the entire data set must be offered, preferably for free (or at no more than a reasonable reproduction and distribution cost), and be downloadable over the internet. Specific criteria for availability and accessibility include:

- Data are easily ***accessible***, *e.g.* it is available in disaggregated forms and in electronic format, and the right to access data in electronic format is recognised.

- Data are available in a ***convenient*** and ***modifiable*** form.

- Data are easy ***discoverable*** and ***findable***. For example, government data catalogues (*e.g.* in the United States, the United Kingdom, Australia), information asset registers (IARs) (*e.g.* in the United States, United Kingdom, Australia, Canada), citizen-driven catalogues (*e.g.* in Canada and Germany). Formats and tagging also make data easier to locate and re-use, *e.g.* "metadata", "microformats".

In order to ensure **re-use and redistribution,** the data must be provided under terms that permit re-use and redistribution including intermixing with other datasets. The core of a "commons" of data is indeed the fact that one piece of "open" material can be freely intermixed with other "open" material. The ability to combine different datasets together is one of the key conditions that allow the development of more and better products and services. New combinations of data can create new knowledge and insights, which can lead to whole new fields of applications. In relation to this aspect, specific criteria include:

- Data are in ***machine-readable format*** – as 'screen scraping' can be time consuming[17]. For example, PDF files are not machine-readable (i.e. data cannot be processed on PCs using databases or spread sheets software), whereas formats such as XML, XSLT are machine-readable.

- Data are released in ***open formats*** (specifications have been made public and there is no need of having a specific software to use the information) which are machine-readable. Even though no access to information legislation grants the right to access information in open formats, most OGD initiatives are starting to be accompanied by policy documents stipulating that official information must be available in open format.

- Data are ***available through bulk downloads*** thus enabling access not just to one or two pieces of government data, but to full datasets as programmers can use these to develop applications that make the most of publicly generated data, and for the highest number of users.

- Data are released in a ***timely fashion***: rapid disclosure allows builders of apps to have access to the very latest data.

- Data are ***linked***: more sophisticated user queries require the creation of structured relationships between government databases enabled by semantic web technologies that convert large quantities of data to linked data formats.

- Users have the right to ***re-use data without discrimination***.

In order to enable exploitation of OGD to develop new business and create new economic opportunities it is imperative to provide data unstructured and in the most easily accessible formats to all businesses regardless of their size. This is a key precondition to avoid creating new forms of economic exclusions.

Conditions for use, re-use and value accrual

As underlined earlier, reaping the benefits of the change brought about by OGD implies an awareness of the importance of citizen and civil society engagement and a good understanding of how to create the preconditions for making the engagement valuable and to spur data re-use. To this end, proactive release of data is only one step. Stakeholder involvement can allow for a better understanding of the context of public demand (*e.g.* users' needs in terms of data, timeliness of updates, formats) and to craft appropriate strategies that support higher use given the specific context (e.g. prioritise data based on the desired objective, adjust data supply to the needs). As a result, higher and more sustainable value creation can be generated. Capturing users' feedback may result in value creation as this feedback enable new features, new lines of business, new markets, new competencies and new tools. Similarly, users can engage in spotting anomalies and mistakes in government data. Developers at the cutting edge of technology can be kept up to date on new datasets being released, and governments can find help in doing things differently and in more agile ways.

The real value of OGD is realised when there is interest and the capacity in re-using data exist. This requires a knowledge of the capacity of a specific context to be a "prepared" recipient of OGD (*e.g.* technical preparedness, understanding and awareness of open government data in the society), as well as strategies to stimulate the active demand of useful data and to motivate data usage. "Hackathon" events or "hack-days" are one way to engage potential users and to develop new uses for public data. An often commonly cited challenge for such events is poor access to OGD. When apps developed during these events use dummy government data it does not help one to appreciate the real value of the application.

Several initiatives are emerging in OECD countries. The UK government is working on a "Government Developer Engagement Strategy", setting out principles for how individual government departments are expected to engage with the development community. Several government initiatives launched competitions with the intention to encourage OGD re-use. For example, the Apps for Democracy, run for a 30-day time period by the US government, apparently led to an estimated 4 000% return on investment. Or a similar Norwegian initiative, Nettskap 2.0, resulted in the development of 135 apps.

Other governments have fostered close collaboration between individual civil servants and public sector bodies with civil society via apps. For example, in the Netherlands the online network "civil servants 2.0" (Ambetnaar 2.0) was developed together with initiatives sustaining a community-based and collaborative approach such as running the data catalogue, overheidsfeeds.nl, or events such as BarCamps[18] on open government.. Similarly, an EU Data Challenge competition conducted in April 2011, over a period of 60 days, resulted in 430 entries from 24 EU countries on ideas, applications, visualisations and datasets[19].

Re-usable data need to be available. The challenge is that often when data are available it is not in a format that enables re-use. Many times this data are shared in closed formats like PDF which does not enable re-use. The previous section of this paper emphasised that the cornerstone of OGD initiatives is to enable government data to be accessible, usable and re-usable. This requires structures in place that will ensure that data are in a format that is acceptable for the three attributes. Having data that are accessible,

usable and can be re-used allows third-party entities to generate useful information, services, products that were not originally intended by those that generated the data.

In order to assess the degree of dataset re-usability Tim Berners-Lee proposed a Five Star Scheme.[20] In practical terms, at the One-star level, data are accessible on the Internet under an open license (such as PDDL, ODC-by or CC0). However, the data are locked up in a document. Other than writing a custom scraper, it would be hard to get the data out of the document. At the Two-star level, the data are accessible on the Web in a structured way (*i.e.* machine-readable); however, the data are still locked up in a document. To get the data out of the document, users depend on proprietary software. At the Three-star level, the data are not only available via the Internet but everyone can use it easily as well. On the other hand, it still remains data *on* the internet and not data *in* the internet. As Jelliffe explained in 2002, if data are "on" the Web, it does not have any links pointing to it, if a resource is "in" the Web, it has links from other resources to it[21]. At the Four-star level, data are in the Internet, and at the Five-star level data are in the Web and linked to other data thus increasing the user benefits[22].

In its "Open Data White Paper", the UK Government expressed its intention to adopt such a scheme as a measure of the usability of its Open Data.[23]

Box 5. Tim Berners-Lee "Five Star Open-Data Scheme"

★	make your stuff available on the Web (whatever format) under an open license
★★	make it available as structured data (e.g., Excel instead of image scan of a table)
★★★	use non-proprietary formats (e.g., CSV instead of Excel)
★★★★	use URIs to identify things, so that people can point at your stuff
★★★★★	link your data to other data to provide context

Source: http://5stardata.info/

Many OECD member countries have adopted freedom of information (FOI) legislation to ensure the public's right to access government information. FOI legislation is an important cornerstone of PSI re-use because the latter can only take place when there is a right to access government information. Traditionally, providing access to information does not given an automatic right to re-use information. FOI Acts (FOIAs) passed in OECD countries have therefore undergone, in many instances, amendments when needed in order to require the release of datasets in a form that can be used and re-used. These amendments usually specify the licence under which the data can be used and require that data must be provided in a re-usable format. Such FOIA are meant to create the conditions for accessing raw, un-manipulated datasets.

In order to give the public an enhanced right to data, some countries such as the United Kingdom have planned to expand their FOIA by accompanying it with a Code of Practice to provide guidance to applicants and public authorities on a number of administrative issues related to making datasets available for re-use, *e.g.* giving permission for datasets, releasing datasets in electronic form enabling re-use. In order to open up the development process of the Code to the public, the UK government offered the

opportunity to shape the guidance using a crowd sourced wiki launched on the government open data portal, data.gov.uk. The Code of Practice was issued under Section 45 of the Freedom of Information Act (2000) in April 2012, and the draft publication is in consultation until early 2013.[24]

Finally, universal participation is essential to reap the value of OGD. Everyone should be able to use, re-use and re-distribute without discrimination against fields of endeavour, persons or groups as only these conditions enable real universal participation. For example, "non-commercial" restrictions that would prevent "commercial" use or restrictions of use for certain purposes (*e.g.* only in education) limit the openness of information.

Creating an Open Government Data ecosystem

Building an ecosystem that responds to specific demands asking for the provision of open government data are a necessity to create value out of OGD. Several scholars have framed the discussion on Open Government value creation using the metaphor of the "ecosystem"; and the ecosystem idea was mentioned by O'Reilly in one of the first discussions of open government (Harrison, T.M et al, 2012). Aneesh Chopra, the former US Chief Technology Officer, differentiated Data.gov from a "repository of data" by describing it as way "to foster a thriving ecosystem that creates opportunities in research and development" (Hendler, C., 2010).

The concept of ecosystem in this context applies specifically to the provision and use of Open Government Data. Establishing the right ecosystem is not so evident, and its creation implies identification and involvement of various categories of actors as well as the provision of the right business case to spur usage. Additionally, it entails the adoption of policies to enable the establishment of the ecosystem built around issues that are universal, the nurturing of a culture of public sector interaction with the actors and the reaching out to some of them that in some countries might traditionally be less actively involved in public affairs (e.g. civil society organisations operating in geographically remote areas and more aware of data that might be needed to develop targeted services that would better respond to local community needs). At least three categories of ecosystems can be identified:

- **Ecosystem of data producers**: government data are produced by actors of very different nature *e.g.* public sector, academia, media, private sector. As the relevance of linked data and data crowd-sourcing increases in relation to value creation, the interactions between data producers and networks of data producers will become more complex.

- **Ecosystem of infomediaries**: as intermediate consumers of data, infomediaries (*e.g.* media, developers, civil society) play an essential role in making sense of, and creating value out of, raw data. Media can tell interesting stories based on such data; developers can develop apps using them; civil society organisations can spot the relevance of certain data for specific segments of the population (e.g. charities in remote areas), can play a critical role to build capacities at the community level, and can create a culture that appreciates the relevance of the data.

- **Ecosystem of users**: communities need to use data and engage based on data in order to get the most out of OGD initiatives. Libraries play a key role in relation as facilitator of accessibility to data, particularly in remote areas, thus enhancing the cost- effectiveness of access, and in relation to data mining. Libraries generate, accumulate, and disseminate information to its readers and users in the desired format and at the desired location. Information in a library is of two kinds — there is the content in books, journals, etc.; and there is information about that content, i.e. the metadata. Metadata concerns information about where things are located, how they relate to other things, how often they circulate, etc. The role of libraries in relation to data mining refers to the process of using library data more effectively, for example starting by discovering ways to

connect the sources of data created by libraries. Data mining can assist communities in understanding the information and data needs of its members, to better respond to their needs.

Interaction among all actors is key, and understanding each category is important, as it helps to grasp what value can be created for the community and how this can be achieved[25]. The key questions are for instance: Who are the members? Who leads the interaction with them? What are the expected outcomes of this interaction and how can these be measured?

Good examples exist at local government level. The City of San Francisco, for instance, is characterised by a culture based on a strong sense of community, with a relatively large number of citizens and ICT activists forming a dynamic ecosystem supporting a strong bottom-up innovative context. The City can also count on the open-minded and collaborative attitude of the City authorities as a real driver for value creation. San Francisco is not a unique example and presents many elements that typify several OECD large- and medium-sized cities and municipalities. A way to replicate this positive experience might be through the adoption of a strategy that levers these conditions where they exist, or fosters their development where they are lacking. The establishment of collaboration frameworks may also help ensure the involvement of different actors (*e.g.* also SMEs that may be important incubators for innovation but are often still too little aware of opportunities brought about by OGD).

The role of ICTs and OGD portals

In understanding the prerequisites for data openness, it appears evident that ICT plays a pivotal role in the development of OGD. ICT enables OGD by hosting and publishing content and by divulging standards for data exchange that encourage free use or embedding of information in other packages for example radio, web, SMS etc. Benefits of using ICT include improved timeliness of reporting, improved data quality by decreasing the manual interactions with the captured data, increased quality through business intelligence inbuilt in the systems, among others. These capabilities are possible if the technologies are appropriately utilised to return maximum benefit.

However today, many governments focus on the development of a national OGD portal as if it were a higher priority than developing technical infrastructures to open up public data for others to use. Understanding the preconditions for effective OGD in a specific context is essential to set up websites that enable value creation, and lies at the core of the government data publishing responsibility.[26]

Many governments' current steps toward releasing reusable data and making them available on a website are valuable and important. However, in many instances these efforts are still seen and prioritised as afterthoughts to the finished sites. As long as governments prioritise the development of open government data portals over frameworks and infrastructures that will open up their data, the pace of change and real value will be held back. Much of current criticism on national OGD portals is based on the fact that governmental interest appears to be on presenting data in a particular fashion, which distracts from, and thereby limits, the increasing provision to users of data that they are really interested in using for their own purposes.

However, there are also arguments for the status quo focus on OGD portals. They are visible initiatives and are easy to explain and communicate. As long as government portals are the priority, there is no risk that a lack of interest by private parties will limit citizens' access to government data. Instead, the government creates a system that every citizen can easily use from home, from a library or other public facilities without the need to understand the inner workings of technology.

There is also a certain economy to the current situation: under the current system, the costs to develop an open infrastructure for third party access are typically incurred in response to specific interest by

citizens in accessing particular data (for example Carl Malamud's campaign in 1993 to move SEC data online based on the idea that public access to US Security and Exchange Commission data was important[27]).

The status quo approach, however, also has marked drawbacks. The institutional workings of government make it challenging to systematically adapt and improve websites as fast as technology itself progresses. As a result, the general assumption is that no single public website can easily meet as many different needs as well as a range of privately provided options can.

KEY DIMENSIONS FOR IMPLEMENTATION

Establishing a legal and regulatory framework is essential to enable OGD; however, even when such a framework is in place, OGD initiatives can be undermined by problems with implementation, technical challenges and administrative delays. Technical, legal and financial restrictions, among others, may limit data accessibility and reusability (*e.g.* making it difficult to find data or valuable ways to re-use data). A number of challenges may be associated with the implementation of OGD initiatives which, if not properly tackled, might obstruct or restrict the capture of benefits of national efforts aimed at spurring OGD. Addressing various challenges related to policy, technology, financing, organisation, culture, and legal frameworks is essential to create an ecosystem, and build sustainable business models for OGD initiatives that can generate the desired benefits.

Policy challenges

Various policy challenges may represent important obstacles to the development of OGD initiatives. Disclosure policies may, for instance, limit data transparency and copyrights may result in lack of clarity over who owns government data. This in turn, can restrict the public's right to use government data (*e.g.* info is sold or comes with restrictive copyright licenses, preventing re-use). There is also an unresolved conflict between the right to access information as an inherent part of the right to freedom of expression, versus the limitation on re-use arising from copyrights and changes for commercial use.

In addition, policies should enable smart disclosure; and they should address the lack of procedures and standards on how to deal with open data in governments (*e.g.* lack of tools available to make data open, of validation structures and guidelines, guidelines on data collection that can compromise the quality of the data and eventually the output of OGD initiatives).

Finally, although the vast majority of OECD member countries have engaged in the development of OGD initiatives, few have specific strategies and related policies. Long-term strategies which take into account the context within the society and in the public administration, and comprehensively address technical, economic, social, legal, institutional and implementation-related aspects are crucial. These are important to ensure that the focus is not on the means or platforms being used to foster OGD (*e.g.* the development of an OGD portal), but on the aim and results envisioned. For instance, many countries seem to adopt, strategies that look for "quick wins" in the first few years, focusing on a fast release of open government datasets. Over time, however, the availability of relevant open government data diminishes in terms of content and adequacy of format. This may result in a loss of public interest and use of OGD, as well as missed value creation resulting from the use (Centre for Technology in Government, 2013).

Box 6. The United Kingdom's "Open Data White Paper": Unleashing the Potential of OGD

In June 2012, the UK Cabinet published the Open Data White paper, which set out how the Government intended to put data and transparency at the heart of government and public services. The document underlines the intention of the Central government to make it easier to access public data; easier for data publishers to release data in standardised, open formats; and engraining a 'presumption to publish' unless specific reasons exist (such as privacy or national security). These objectives are indicated as critical in relation to the full commitment to making Open Data an effective engine of economic growth, social wellbeing, political accountability and public service improvement in the United Kingdom. In order to frame a feasible public sector implementation of OGD, the paper highlights that after two years of having had the centre of government leading the initiative, government departments were expected to take a greater role in driving forward the efforts. Therefore, alongside the White Paper, each government department published their first Open Data Strategy. Each strategy contains a department's commitments for proactively publishing data over the next two years and will complement their existing statutory publication schemes. These strategies represent an important step forward in the way the UK is making data readily and systematically accessible and are a core requirement of each department's activity.

Source : *www.cabinetoffice.gov.uk*

Technical challenges

Government data are often un-harmonised as every public agency has its own set of data, formats and standards. This can make it difficult from the user perspective to know which piece of data is valid or should be trusted. A critical element of accessibility is to know the source of what the user is searching for, and in many instances, where to begin searching can pose a challenge.

OGD value can also be limited if PSI cannot be re-used, and data transparency might be hindered if data are not simple to access due to its formats. Additional technology-related shortcomings include the need to: improve information technology infrastructure, enhance privacy and information security, and integrate OGD tools and applications.

An additional layer of challenges can emerge when the federal government seeks to impose co-ordination or consistency across the broad range of rulemaking processes, data and portals enabling access to government data. Even though the establishment of a single OGD portal should not be the goal, and is far from being the best advisable solution for implementing OGD, a single portal can certainly ensure integration and shared data input from various sectors of government, and can greatly enhance accessibility. This is the reason why emphasis is often placed on the establishment of a single portal – as underlined earlier – and not on the wide spectrum of relevant issues. However, to meet government-wide needs in terms of data management, when the decision is taken to create a single portal, it is important that this is developed through a collaborative approach, creating ownership and sustainability. Poor data management practices can increase the cost and efforts for the government to open data and make it available in machine-readable formats (Centre for Technology in Government, 2013). The trade-off between standardisation and experimentation, and the concerns about incomplete or inaccurate data in centralised government repositories are inherently difficult problems that most governments are currently dealing with.

Box 7. The Case of Regulations.gov

Regulations.gov is a government-wide docket publishing system created in the United States in response to the E-government Act of 2002 and launched in 2003. It is used today by most US departments and agencies[28] and the policy of the Office of Management and Budget (OMB) not only requires its use, but also precludes agencies from using "ancillary and duplicative" docketing and rulemaking systems of their own design.[29]This exclusivity rule, has been considered by many to have led to a bare-bones approach that leaves out the agency-tailored functionality found in many of the systems it replaced.

Regulations.gov was launched with a limited search engine and no browsing capability, so that only those who already knew the terms of art used to categorize rulemaking documents were able to use it effectively. Five years later, a re-launched version of the site offered up a limited inventory of computer-readable data directly to the public, using a single RSS feed that allowed any interested person or group to create an alternative, enhanced version of the Web site. This has permitted the creation of OpenRegulations.org, which competes with Regulations.gov by offering "paired down, simple-to-navigate listings of new agency dockets" and a more sensible set of RSS feeds, one for each individual agency.[30]

Source: Author.

Lastly, interoperability remains an unresolved issue in e-government, and can potentially have an impact on OGD development as well. Dealing with OGD in general, and open data file formats in particular, can facilitate IT system interoperability in government. Interoperability is a major concern for policy makers working on the implementation of OGD. For example, ensuring the interoperability of data catalogues, or the creation of a pan-European data catalogue, is a big challenge faced by EU policy makers at the moment.

Economic and financial challenges

A number of economic and financial challenges are hindering the fast-paced development of OGD initiatives in several OECD countries. These mainly involve how to develop a new business case and financing model for collecting, converting, and diffusing public data in light of the accessibility principle that open government data should be free or provided at cost. Specific costs that need to be financed include:

1. Collection and provision of data

Governments have concerns about the cost of opening up government data, although such cost – as well as the cost of data production – has not been sufficiently appraised so far. The common assumption is that data that can be made open are just a product of what happens already inside the public sector. Open is not always free, especially in times of government austerity, as there are undoubtedly some potential costs associated with the production and presentation of open data that need to be considered and accounted for. Such matters raise the need to be mindful about implementing OGD in pragmatic and affordable ways, which do not add unnecessary financial burdens and potential loss of revenues.

There is a substantial commitment and investment on the part of public agencies as they need to acquire new skills, train employees, purchase technologies, and upgrade network infrastructure, which need to be accounted for. There are also human-resource costs associated with organising and preparing information to be put online, particularly if the decision is taken to develop a special portal that may require an IT and design team and to ensure timely publishing and update of data, or coherent production of high quality data. These are normally held by each agency, which often relies on existing charging frameworks to get compensation for the costs incurred. This status quo is changing and several governments such as the Dutch and Danish government are looking into the development of a business

case with alternative funding and financial models. Some countries have chosen an intermediate path to recover costs allowing non-commercial reuse at zero cost and charging for commercial re-use on the basis of license cost per sentence[31].

Another interesting example of recent national efforts to revisit existing charging models comes from the United Kingdom. The Government launched a formal consultation on 4 August 2011 to look in detail at key questions on the data policy and charging framework for a Public Data Corporation known as the Public Data Group (PDG). The consultation questions focused on: charging, licensing and regulatory oversight.

2. Converting large volumes of data into re-useable formats

Converting public data can have cost implications, particularly if there is a high level use of proprietary software. Initiatives such as converting government data to semantic web and linked data formats can be time consuming and therefore costly, similarly to enabling partial access to large volumes of data (*e.g.* a field in a database that contains personal data such as the e-mail addresses of private individuals that can be removed before the remainder of the information is released in order to protect personal privacy while respecting the right of access to information). These additional costs can lead to some reticence on the part of government bodies to convert public data. This in turn, can result in the refusal of even partial access to a requested database. Another reason is that a huge database may contain many bits of data, which takes time to identify and remove. However, to comply with the right of access to information, public bodies often have no option but to take the time to remove the sensitive data and to grant access. A cost-benefit analysis often shows the value of taking the time to facilitate access.[32]

Designing databases with the right of public access in mind is seen as longer term solution, even as it appears to be increasingly easy to achieve, at least from a technical perspective. It is possible, for example, to build a database that performs one-way encryption. This permits e-mail addresses to be included in a database, but in another table that is linked via a hash value so that when the data are shared, the e-mail addresses can be separated, thus ensuring that any personal data are kept secure. One way to think about this is to take the analogy of a bank: it is possible to enter a bank and look around, but part of the bank, *e.g.* the vault where the money is kept, is locked and secure.

3. Posting full datasets in open source formats on government websites

Users can more easily find the specific information that they are looking for when OGD is properly tagged with metadata. This is one of many solutions to releasing information that come at a very low cost. It would be advisable to consider such activities as part of the day-to-day activity of public bodies.

4. Designing a new data framework to assess costs and benefits of OGD

It is difficult to develop an OGD business case when there is a lack of both cost clarity and a clear measure of the gains from opening up government data. When government provides reusable data, the practical costs of re-use, adaptation, and innovation by third parties are significantly reduced. It is reasonable to expect that the low costs of entry will lead to a flourishing of third-party sites extending and enhancing government data in a range of areas – rulemaking, procurement, and registered intellectual property, for example. This approach could be adopted by those governments that decide to shift their online focus from developing finished OGD websites to the infrastructure that allows new sites to be created. If the creation of infrastructure results in better third party alternatives then the government entity can cut costs by limiting the creation of its own websites.

This underlines the need for better data framework to assess the costs and benefits of OGD, as well as a clear OGD strategy that provides incentives to public officials to invest in OGD activities. Such an

approach would more clearly frame co-ordinated and efficient decisions on government IT and information architecture, and could secure alignment with a government-wide IT procurement strategy. If, on the other hand, third party alternatives to the government site do not satisfactorily emerge then the public site can be maintained. The overall picture is that government IT costs will decline in those areas where private actors have the greatest interest in helping to leverage the underlying data, while government IT costs will increase in those areas where, for whatever reason, there is no private actor willing to step forward and create a compelling website based on the data. Governments are keen to collect evidence on recent initiatives which show that putting raw data online demonstrates that it can be considerably cheaper than presenting the data to the user via a custom web interface.

Organisational challenges

Governments are still struggling with a number of key organisational challenges. These are strategic elements which need to be understood and addressed by national strategies and plans. They include:

1. **Ensuring accountability, quality of data and responsibility in a context of collaboration**

Given the complexity and cross-cutting nature of OGD, governments need to put in place the appropriate institutional structures. Tasking a government body, often the Centre of Government (*e.g.* the Prime Minister's Office), with championing, co-ordinating and providing support and leadership on OGD initiatives and programmes, has been seen by many OECD countries as way to bring the various stakeholders on board. Having a ministry or other dedicated body in charge of soliciting the various datasets from government agencies that will then be made public has been also considered as a way to i) sustain collective work to strengthen data integration across different parts of the public sector, ii) help building better capacities across government to deal with emerging concerns (*e.g.* privacy, transparency) and iii) ensure that those making decisions about the release of data do so in a rigorous and consistent fashion. The US Government has decided to take a very innovative approach, creating a new position - "Evangelist" - for its website, Data.gov, in order to facilitate OGD development and boost data use.[33]

2. **Balancing autonomy and control**

Empowering independent oversight bodies to demand and to publish information on budgets, procurement and expenditures is considered crucial to ensure data transparency. Several governments, *e.g.* the United Kingdom, are considering the establishment of independent ethics and governance groups to oversee policies and procedures for improving the use of administrative data.

3. **Ensuring sustainable change through "ecosystems" of key actors**

The creation of the right ecosystem – i.e. a community of key actors - is essential not only to reap the economic benefits, but also to generate the value of OGD initiatives in social and political terms. As indicated earlier in this paper, open data use by third parties, as well as the uptake of the apps based on open data, are essential factors to make OGD initiatives sustainable and to create value. The case of the app "Asthmopolis" in the U.S. is an excellent example of an application developed thanks to an ecosystem that has brought social value and improved quality of life to a vulnerable segment of the population: people with asthma. Public data and data provided by people affected by the disease have been merged into the app to enable the identification of highly dangerous spots in the U.S. for asthmatic people. Hospitals have recorded a decrease of 25% of the incidents since the app was created.

4. **Engaging with the wider community in a two-way conversation to build capacities and find agile solutions**

Pushing data out is not sufficient to create value. Robust engagement models and strategies also need to be in place to allow two-way dialogue to take place between the public sector and the users of government data (*e.g.* individual developers, SMEs, citizens, civil society organisations, academics and large companies). This is key for governments to focus on user need and for users to provide feedback on the datasets they would like to see released as a priority which they consider of greater value or more likely to be used by the community.

Social media can play an important role in inspiring or enabling OGD usage, and in involving communities of practice, formed by people who engage in a process of collective learning related to OGD to sustain relevant initiatives and help create a network of actors. The use of social media to capture user feedback can also help create a need for use, *i.e.* get the data where people really need them. However, engaging users requires resources. In order to capture the views of Open Data users', the United Kingdom established a group in the Cabinet Office that comprises 14 officially selected volunteers from the civil society and the private sector who play an advisory role on the data the government should release.

5. **Revisiting internal processes to support data release workflows**

The actual implementation of OGD portals implies the establishment of adequate workflows for data gathering, integration, validation, release, approval granting, update and re-use promotion. In some instances the process of online data release is supported by an organisational culture already oriented towards data sharing and re-use, which facilitates the process reengineering needed. In other cases the internal culture of public sector institutions is not immediately conducive to data sharing and requires additional efforts.

Cultural challenges

Legislation, IT platforms and codes need to be matched by a culture within the public sector that supports the presumption to publish/release data. Below are some of the most pressing issues from this perspective.

1. **Increasing public interest and preparedness**

Raising the capacity and awareness of civil servants, citizens and the private sector on their rights to access and re-use of public data is important for society as a whole to fully capture the benefits of OGD initiatives. Government departments, in partnership with civil society groups, can for instance create awareness on the available legislation and policies that empower citizens to access information, such as the Access to Information or Freedom of Information Acts in OECD countries. Additionally, undertaking research to establish citizen's information needs and barriers to information use and re-use, or seeking public-private partnerships to encourage open government data use for public service innovation, can lead to ventures for the worthwhile re-use, re-distribution and universal participation in OGD such as applications development and provision of e-government services.

2. **Recognising the value of crowd sourcing**

Of critical importance for governments is to recognise the value of crowd sourcing to find the 'talent' outside the public sector to use, exploit and create value from data. This is not necessarily easy as successful crowd sourcing also depends on a sufficient scale and representativeness of participation to achieve valuable results. To date, only a limited number of governments are well embarked down this path

and even fewer local and regional governments where the benefits are likely to be the greatest. U.S. cities and the U.S. Federal Government, the United Kingdom, Australia and France, as well as a handful of others, have been leading the way in this respect.[34] Companies and SMEs in some countries are exploiting such data to expand business and create jobs, whilst a few governments are using such data to encourage co-creation events *(e.g.* innovation camps, "hackathons", code sprints and apps challenges) to create apps, services and enable various stakeholders to contribute to policy making, public value creation and social innovation.

3. Shifting the culture of the public sector, providing incentives and building new capabilities

Missing participatory and collaborative elements, incompleteness of data and lack of raw data represent much more than technological challenges. These are not merely technical issues and solving these matters implies a fundamental cultural change in the approach of public authorities: from disclosure to proactive and smart disclosure; from provision of information to provision of data, abandoning the idea of the public sector's interpretational sovereignty. The belief that making data public disempowers public officials, or makes them more vulnerable as they risk unveiling faults, can at times create an environment among civil servants, or even policy makers, which does not fully support implementation of OGD initiatives. Governments are for instance increasingly considering the development of programmes that help change the attitudes of public officials on making information available to the public and on improving its sharing with peers. Many governments are realising that cultural and administrative barriers to data sharing can be best addressed through engaging with, and crowd sourcing the experiences of, civil servants working with data both in the front line and in central governments.

Additionally, governments should have the capability to conduct, interpret and consume the outputs of data analysis to inform better decisions. This also includes the capacity to debate the meaning of data and to find ways to use it in democratic debates. This is only partly about cutting edge IT and data science skills, as it is also about ensuring that public sector managers and policy makers are confident in combining data with sound judgement, and are aware of the need to encourage the pursuit of the OGD agenda with strong ethics and integrity.

Furthermore, even though awareness of available data is an essential step for any government's OGD strategy, most governments do not currently have a comprehensive overview of the data in their possession. The UK Government organised for instance information engineering programmes that required more than 100,000 authorities to reengineer their records. This exercise is considered to have been essential for the success of open data initiatives. However, the cultural context matters, and the forceful approach that may have worked to make OGD initiatives successful in one country may not have the same rate of success in another.

Finally, particular efforts have been undertaken by many governments engaged in the establishment of OGD portals to encourage the use of linked data, e.g. data.gov.uk. The skills and experience for working with linked data are still at a relatively early stage of development, but advocates of linked data approaches believe it has the potential to enable a revolution in how data are accessed and utilised (Davis 2010). There is, however, much that governments need to do in this sense. Data.gov.uk appears to be one of the few OGD initiatives where links between the different datasets have been created (Kalampokis et. al, 2012).

4. Ensuring the "buy in" of all stakeholders

Initiating dialogue among various stakeholders about the importance of sharing information and its benefits to the public can help secure their participation and ensure buy in. Current and potential re-use

initiatives by the private sector, civil society organisations and individuals can be publicised to increase the awareness on the benefits of opening up data.

Legal challenges

The legal landscape surrounding data sharing and opening is undoubtedly complex.[35] Having a consistent legal framework in place is critical to facilitate government data accessibility and re-use, and to improve secure data sharing between public authorities and the wider community to improve insights, results, impact and inform better policy making. Certain legal requirements, as well as fragmented and diverse legislation concerning privacy (*e.g.* UK Data Protection Act, re-use and sometimes fees (*e.g.* in Sweden and Germany[36]) can create confusion for end-users, or can make it more difficult for governments to make data available. FOI and PSI legislation as well as clear licensing guidelines are a cornerstone of OGD. In order to facilitate and co-ordinate the work of agencies in their transition towards OGD provision, guidelines and handbooks are among the useful measures a government may choose to adopt. These guidelines cover technical and legal issues, economics as well as communication strategies. Several countries are already working on the development of such guidelines (*e.g.* Norway) or have recently published them (*e.g.* Spain[37], France[38] and Denmark[39]).

Some of the current legal challenges include:

1. The scope of right of access to information

In principle the right of access to information applies to all information held by public bodies, and hence should apply to databases. But in some countries databases are excluded from the scope of the right and in others the law is not clear, while practices vary across countries. Similarly, not all countries establish a right of access to information in electronic format wherever possible, and many FOIAs do not make reference to machine-readable or open formats. The definition of information in most access to information laws typically refers to all information recorded in any format that should include databases. However, there is often no explicit reference to right of access to databases, except for laws such as in Finland and Norway, which expressly permit access to databases. On the other hand, in Sweden such access is provided, but only in printed format, while in the Netherlands and Denmark, databases are specifically excluded from the scope of the law. This is predominantly a problem with older access to information laws. In the majority of countries where there is no specific exclusion for databases, access to information and OGD advocates can use the wording of the national access to information law to argue that the right applies to databases.

2. Legal exceptions to openness

There are a number of circumstances in which information held by public bodies may – rightfully from a legal perspective – not be completely open to information seekers. The first is where the information falls under legal exceptions on grounds such as national security or protection of privacy, and is therefore not released to the public, even when someone files an information request. The second is when the public body rules that the information can be commercialised by being sold to for-profit companies which produce added-value products, *e.g.* for commercial re-use. The information will therefore be released to members of the public or to private companies, only upon payment of a fee. From a public sector perspective, these exemptions are necessary to reassure users that the right data are protected, and the challenge emerges to ensure that the right criteria are explained to third parties and applied consistently.

3. Complexities of the various national legal frameworks for copyright and related rights

One additional legal area that particularly lacks clarity and that has implications on government data "openness", is the question of who owns government information. Many 'access to information laws' presume that all public information, except for the confidential ones, is to be accessible as they consider the general public as the legitimate owner of this information. However, in many countries it is still the case that public bodies assert intellectual property rights such as copyright and database rights (*i.e.* intellectual property rights generated by the compilation of a database even from pre-existing material) over the information they have generated or collected. Even where intellectual property rights are not asserted, public bodies tend to assume that they are the exclusive owners of the information, and this is why their economic model sometimes includes selling the information for profit.

4. Legal requirements for Web masters

Compared to technologists in the private sector, national webmasters in the public sector face a daunting array of additional challenges and requirements that are often not technology-related in and of themselves, but are in fact legal challenges with technological implications. In the United States, for instance, an online compliance checklist for designers of federal websites identifies about twenty-four different regulatory regimes with which all public federal websites must comply.[40] These range from privacy and usability to FOIA, compliance to the demands of the Paperwork Reduction Act and the Government Paperwork Elimination Act. Each of these requirements alone is justified by federal mandate and reflects the considered judgment of political process, informed by the understanding of information technology that was available when it was written. However, the cumulative effect of these requirements, taken together, is to place federal web designers in a compliance minefield that makes it hard for them to avoid breaking the rules – while diverting energy from innovation towards compliance. These problems are not unique to the United States and are faced by public sector webmasters in many countries.

5. Extent of flexibility in existing regimes

Updating policies and rules is essential to properly address issues related to putting government information online. A number of recently adopted laws that explicitly address the issues of putting government data online pose for instance a question of construal: Does an Internet location that contains machine-readable XML, which can be displayed directly in a web browser and deciphered by humans, but is designed to be used as input into a presentation system or engine, count as a "website"? If not, these statutory requirements may require government bodies to continue maintaining their own sites. It could be argued that XML pages are not web pages because they cannot be conveniently understood without suitable software to "parse" them and create a human-friendly display.

Adopting the required regulations that allow Access to Information Acts to be operational is important. Furthermore with Access to Information Acts, the government is expected to promote accessibility to open government data for minority groups to avoid creating new forms of digital divides and to increase inclusion, *e.g.* language options for content and access for the disabled including the hearing and vision impaired. Inconsistent laws, such as the Official Secrets Act in the United States, if not amended to be brought in line with the requirements of increased transparency and openness by public bodies, can hinder the full-fledged development of OGD initiatives and the enforcement of the supporting legislation.

Box 8. Landmark decision from the Netherlands

In April 2009 the Judicial Division of the Dutch Council of State (*Raad van State*), the highest Administrative Court in the Netherlands, placed limits on the possibility for public bodies to charge for access to databases they have created when it ruled that a public authority could not assert database rights over, nor charge for, data collected with public funds as part of its regular activities. The case was taken to the Court by Landmark Nederland, a large supplier of land and property search information, which in 2006 brought together a national dataset of environmental risks such as contaminated land from a range of sources including Dutch council records. These reports were part of a portfolio of products to be sold to home buyers via estate agency brokers. The City of Amsterdam sought compensation for supplying the data and also wanted to limit its re-use, arguing a substantial investment had been made in compiling the original dataset.

The Court rejected the appeal lodged by the City of Amsterdam for compensation costs for supplying information that would be sold for profit. The Court ruled that, while the data could be considered to form a database because there had been a substantial investment in its collection, the City of Amsterdam had not borne the risk of this substantial investment, and was therefore not a producer of the database for the purposes of asserting database rights. Consequently the City was not entitled to attach financial conditions or other limitations on the use of this data by Landmark Netherland.

Source : Based on material published on the EPSIPlatform website,
www.epsiplatform.eu/examples/cases/landmark_nederland_bv_v_amsterdam_city_council

NEXT STEPS IN THE ANALYSIS

The Executive Order released in 9 May 2013 by the White House [41]makes "open and machine readable" the new default for the release of government information. A Memorandum was also published by the White House establishing a framework to institutionalize the treatment of government information as an asset. The new executive order is more legally binding and specific than any measure previously adopted by the White House to establish a clear and sound legal environment for open government data. The White House has also taken steps to operationalize these principles by:

- Mandating that when an agency procures a new computer or system that collects data, those data must be exportable. That won't address digitizing existing government documents and data but will create a default setting going forward.

- Planning to re-launch data.gov in a format compatible with dozens of other open-data platforms around the world.

- Requiring agencies to catalogue what data they have. Understanding what you have is fundamental to managing information as an asset, although an open data policy that requires creating and maintaining an enterprise data inventory won't be without cost. Creating a public list of agency data assets based upon audits is one of the most important aspects of the new open data policy.

Some key issues still remain unresolved and are worth being debated and addressed. These issues are crucial aspects for the conceptualisation, development and implementation of OGD initiatives as well as for the definition of the underlying policies. Further research, analysis and exchange could help governments clarify their position and approaches, and ensure long term sustainability of the initiatives.

Overarching conceptual issues

1. **Bridging the distance between access to information and OGD movements.** Advocates of access to information and those sustaining OGD have been pushing for their own agendas individually and in a way that has been, at times, disconnected from country initiatives. It is important to see how the rights-based approach of the access to information movement could complement the arguments on the economic and social benefits of releasing government data employed by OGD advocates.

2. **Resolving the conflict between the right of access to information, as an inherent part of the right to freedom of expression, and versus the limitation on re-use arising from copyrights and charges for commercial use.** In spite of the fact that in many countries governments assert rights over public sector information, including intellectual property rights, there is need for serious debate as to whether or not this is appropriate in modern democracies, given that the information has been created with tax-payers' money. This debate could take as a point of departure, the principle that public bodies are mere custodians of the information created for, and on behalf of, members of the public (Open Knowledge Foundation, 2011). These are issues that

need to be addressed in the long term, as part of the on-going democratic challenge to promote fully open and accountable government.

3. **Adopting an Open Government data strategy.** Many countries do not have a dedicated OGD strategy – either a standalone strategy or one embedded in a wider Open Government strategy. An OGD Strategy can help establish what data to open, for what purpose and how the government is planning to stimulate a market for its use. The strategy would normally contain a public organisation's commitment to proactively publish data and a timeline for implementing such a plan. On this basis, departments could be required to report on progress made on their commitments (*e.g.* alongside the Open Data White Paper published by each government department in the United Kingdom in June 2012, the UK Government also published its first Open Data strategy).

Addressing the emerging needs for governments and public at large

1. **Enhanced ability to combine different datasets together can help develop additional, more innovative and better products and services.** Mixing public data with commercial, civil society and citizen input data, and pooling and sharing with those produced by other public agencies and/or cities – *i.e.* data sharing for developing shared content, services and policies between cities – holds considerable potential for public value creation. These aggregated city-to-city data sets could in principle also be opened up to companies and civil groups as well as individual citizens[42]. Authorities points to the need in the future, not just for "big data" drawing on citizen inputs and facilitating data analytics, for example to develop and simulate public policies and better target services, but also for a more qualitative approach including ethnographic surveys. A need is thus foreseen for both big quantitative data crunching to provide explicit codified evidence for public sector activities, on the one hand, as well as more qualitative survey data to contextualise "big data" to provide the necessary implicit and un-codified evidence. Countries are still struggling with the development of the skills the public sector needs to conduct data analytics and make the best use of data analysis, as well as to cross link data and sources. This is essential to spur open government data use by the public sector that drives better decisions, informs policies, supports the development of data driven processes and services, and delivers more innovative services.

2. **The value for good and participative governance to experiment and to be innovative.** Learning from the private sector is essential. Over the last decade the private sector has increasingly used data analytics to target the delivery of goods and services. There is much governments can learn from the private sector on how to combine the use of data and the latest technology to achieve the delivery of modern and personalised services targeted to the needs of users. As an example, Facebook's model builds on personal interaction through messaging. The lesson for the non-commercial world could be, for instance, that it is important to **establish** channels that are basically push mechanisms for information, or retrieval mechanisms for complaints and comments. This can help to improve government performance. What's missing, however, would be the dialogue, not so much between people themselves, which often happens anyway, but between individual citizens and governments. This requires structuring, tracking, tracing and personalising answers to the input received by local officials at the right level in the government rather than by an anonymous agency or ministry. This requires time and effort, but provides potential wins for citizens, as well as for the government. It can move governments from one size fits all to segmentation and finally to personalisation.

3. **Collective learning and collective intelligence**. Encouraging the emergence of more advanced features, beyond simple delivery of data, can foster a collective learning process. For those desiring to build interactive sites, the barriers to entry are remarkably low once government data is

conveniently available. Web hosting is inexpensive, software building blocks are often free and open source, and new sites can iterate their designs rapidly. Successes thus far *(e.g.* the Govtrack.us site built by Joshua Tauberer[43]), show that significant resources are not required to enter this space.

4. **Avoid creating new divides and inequalities and focus on user centricity and universal access.** Greater access to government data is an essential part of any government strategy for making the most effective use of data. Data that are trustworthy, easy and free to access, and easy to use (*e.g.* through better search facilities, simpler ways to access information, advanced GIS data search) is key. Unequal distribution of skills, resources and time have a big impact on who is making use of OGD and the reach they can have using OGD. For instance, there is a risk that a focus on data for developers, and the expectation that all data will be accessed through online interfaces, can lead to a neglect of mechanisms for direct OGD access. The current OGD community, for instance, appears to be very technology focused in many countries. As Davies points out, focusing excessively on an "idealised digital infrastructure" and an "idealised fully digital-savvy society" may risk losing sight of the real end-user (Davies, 2010). For instance, human readable raw data sets might be as valuable as machine-readable ones. It is impossible to predict which data provide the greatest value to society in the near and long future. What seems of little value today may be an important component of a combined data set tomorrow. Equally, what is popular today might lose its attraction tomorrow. As in any sector that democratises, more diversity will exist and demand will spread out over more data products.

5. **Enhanced access to data requires changes in the procurement and delivery of IT at the strategic level.** Government data are often locked into inflexible IT systems and retrieving the data are frequently a costly exercise requiring a detailed business case or contractual amendments. The analysis and multi-stakeholder dialogue on these matters need to be on-going.

6. **Understanding the context of data demand.** Users can play a key role in helping governments create value out of OGD. The involvement of different stakeholders in the various steps of the value chain, and particularly in identifying relevant data and the most adequate format given the use they intend to make of the data, can significantly increase the value of data. This can ensure that data supply meets the demand, generating the needed level of uptake to create value. However, understanding data users' needs and nurturing their engagement to drive value creation still does not appear to be a priority for many governments. Many still invest more in the technical developments rather than focusing on improving the set of conditions that support value creation in a specific context starting from understanding its overall characteristics.

7. **Grasping the value of open government data for the public sector:** There is a growing sense of awareness of the potential value for the citizens and for the wider economy which may be derived from opening up government data and enabling wide re-use. However, there does not seem to be strong focus yet on understanding the value for the public sector emerging from OGD.

Harmonising definitions and sorting out legal "conflicts"

1. **Clearly defining terms and clarifying the relationship between data, information and knowledge and its implications is critical.** The knowledge management literature provides a rich discussion of the epistemological issues related to data, information, knowledge and their inter-relationships (Alavi, 2011). The conventional view is that there is a hierarchical relationship between data, information and knowledge. Thus data would correspond to observations of the world or facts that have not yet been interpreted. Information is data with meaning and purpose (*e.g.* Who? What? Where? When?), and knowledge is able to assign meaning to information based

on beliefs, perspectives, expectations or judgments (*e.g.* How? Why?). Yet others argue that there is no raw data since every measurable or collectable piece of fact has already been affected by knowledge (Tuomi, 2009). As data are becoming increasingly refined, the yield of the various forms of data, information and knowledge for an individual or society increases. This discussion may seem to be of philosophical nature, but is necessary as OGD publications rarely provide guidance on the meaning of data. Terms are often used casually, which leads to confusion in and outside of governments, *e.g.* when discussing the meaning of raw data. The same applies to PSI, which is often used as an umbrella term for all content produced by public bodies (OECD 2005). Additionally, some countries underline how the absence of commonly agreed upon conceptual tools used to analyse the meaning of "data" wrongly leads to the interpretation of its relevance and nature through concepts of documents, files, etc.

2. **Striking the balance between transparency and privacy.** Sorting out how to balance both the benefits brought about by OGD in terms of public trust and confidentiality, with the need to protect the formulation of government policy or commercial interest or to protect public interest or personal privacy (*e.g.* health and care records/safe use vs. benefit to the wider society), remains an unresolved issue. Governments realise that releasing greater quantities of anonymised data raises complex questions about how to guarantee that privacy is protected; particularly as increasing the availability of anonymised data has the potential to increase the possibility of identity disclosure through a mosaic effect. Critics of OGD initiatives underline the potential risks for privacy and fraud. This requires further thinking that may also lead to the adoption of sector-specific approaches to transparency given the variety of sector-specific issues. For instance, in the United Kingdom and the United States, the government has implemented pilot programmes for giving patients and citizens direct access to their own data that have proven quite popular (*e.g.* Blue and Green Button Initiatives).

3. **Building greater trust in public data and managing risks.** The success of creating a government data marketplace is very much related to the public sector's ability to safeguard people's data from misuse and protect the public's rights to privacy, but also with the level of overall public trust in government. Opaque ownership of government data used to deliver services can for instance hinder their uptake. Ontology and data architectures can support ethical consumption of OGDOGC. Trust in companies' use of public data, third parties' certification and the involvement of trusted infomediaries can play an instrumental role towards higher public trust in government data. This is why knowing who controls data and avoiding manipulation that may endanger quality is essential to maintain public trust. Data cleaning, regular updates and management of risks involved with release of data that can be used to triangulate on individuals are all key factors that may affect public trust.

4. **A further set of issues surround the relationship between the European Union's PSI Directive[44] and access to information laws.** These are issues that are particularly relevant in Europe, but that carry cautionary lessons for other parts of the world where there is still a culture of charging for the right to use large datasets from the administrative, legislative and judicial branches of government. A first problem is the need to harmonise definitions between rules on the use of public sector information and access to information laws in order to avoid creating an artificial division as to which rules apply to which kind of information, or worse, separating the right of access from the right to re-use. In France for instance, a chapter on the re-use of PSI was introduced into the 1978 Law on Access to Administrative Documents. In doing so, limits were created on the use of information obtained under access to information requests, with the need to seek permission to re-use the information and the possible threat of a fine for not doing so. Hence, although the law defines PSI as information contained in administrative documents, this leaves some doubt about what information is being referred to, particularly with respect to databases[45].

5. **A potentially more serious problem is that government data re-use rules can set up a dual charging regime**. On the one hand, there is the principle that access to information requests should always be free of charge because the information generated by public bodies is created using taxpayers' money. The only permissible charges are copies of information (photocopies, copies on discs) and postage or other costs that are actually incurred in delivering the data. On the other hand, the licence fees for data released under the PSI Directive can run to millions of Euros. A 2009 study commissioned by the EU found that a single piece of geographic data – an aerial photo of 10 km2 of land – can range in price from free of charge to EUR 292 with the average price being EUR62. And with any serious re-user, many such units would be required. The same study found that 27 holders of geographic information from 24 EU countries had an income in 2007 for their products of a total of EUR 356 million. Whilst geographic data seem to be a particularly lucrative market (so prices may be different in other sectors), these figures do indicate the potentially very high costs of obtaining public sector information[46].

Advancing empirical analysis and measurements

1. **Empirical analysis and measurements are pivotal to better understand and measure both cost and impact of OGD in wide social and economic terms, and with regards to political issues, in order to inform decisions around the design and implementation of OGD.** Expectation that developers will make widespread use of OGD in civic contexts requires further justification and the connections between OGD and democratic empowerment or public sector reform need to be outlined in more detail. Understanding the different ways in which OGD is being used is essential groundwork to sustain the evaluation of its social political and economic value. The democratic value of OGD can be assessed intrinsically on the basis of other values such as freedom or the equality that it promotes; or it can be valued instrumentally for producing better outcomes in terms of laws and public services. It is theoretically possible for public services to be improved in the absence of, or in contradiction with, improvements in democracy. Furthermore, governments have concerns about the cost of opening up government data although such costs – as well as the cost of data production – have not been sufficiently analysed and appraised so far. This is an area worthy of additional work. Existing assessment exercises evaluating and assessing OGD initiatives are each providing only a partial view of the full scale issues of OGD initiatives.

2. **A good understanding of the relationship between the supply and use of OGD can allow a deeper evaluation of the OGD value chain.** Understanding the full public value chain of OGD development and use can inform future investments concerning OGD supply. Impact measurement of policies and new knowledge from combined data sources and patterns in large volumes would also be critical. Additionally, how far innovation can move from impacting predominantly digital and informational services to impacting wider areas of public sector reform is yet to be assessed.

3. **Adopting solid "cost and benefit analysis" frameworks and clear business models to boost the desired impact and obtain the desired results.** OGD users motivated by innovation, creating applications and/or websites without a clear business model have little incentive to move from proof-of-concept to production uses of data. Additionally, many unresolved issues concerning to the financial implications of OGD (see section on economic challenges) further underline the relevance of having a business case to resolve these uses and provide incentives to the various agencies, as well as to provide a financing model for future data collection under a scenario in which much OGD is made available for free. Additionally, this could help creating drivers for OGD initiatives beyond the political push of a very committed leader (top-down approach), and ensure the wide sense of ownership and buy-in which is needed to secure sustainable results in the long-run (bottom-up approach).

ENABLING AN EMPIRICAL ANALYSIS

This section provides a suggestion for an analytical framework, and data collection as a complement[47], to gather key data on OGD Initiatives: supporting strategies, implementation frameworks and impact and value creation assessment. The analysis of background information collected through interviews and desk research, and structured around the main components of the Analytical Framework, to be completed by data collection, will enable the mapping of initiatives in OECD countries, identify typographies of initiatives and approaches and move towards a common understanding of what matters for outcomes and a unique methodology to understand correlations and underlying causes and thus assess the efficiency of implementation and achievements of impact. This exercise is ultimately expected to assist countries in keeping up the promises of open government data initiatives as a lever for public sector reform which requires their full scale implementation backed up by a comprehensive set of measures.

The proposed data collection includes:

1. **The questions** to be sent to the countries based on the framework. The questions could, for instance, address some of the following items:

   – Design (strategy): link with Open Government wider strategy [transparency, participation/ engagement, accountability + innovation, efficiency, collaboration, effectiveness].

   – Implementation: local vs. central, online (websites) vs. offline initiatives, role of intermediaries

   – Sustainability: creation of an ecosystem, financial model behind them, system of incentives, etc.

   – Evaluation: proven impact in terms of economic (quantified) and social value, uptake.

2. **Conference calls** with national counterparts working on OGD initiatives (e.g. a sample call with the Netherlands enabled the OECD to capture important knowledge and insights which could not be captured otherwise)

3. **Other empirical data** from questionnaires and interviews, but also from existing reports and country data.

The tables below puts forward a suggested analytical framework for national (or sub national) OGD portals and supporting initiatives, based on the analysis contained in this paper (Table 1); and a proposed set of metrics on OGD (Table 2).

Table 1. Analytical framework for national (and sub-national) OGD portals and supporting initiatives

Component/dimension		Focus of analysis
Overarching issues	***Overall vision***	This section aims to provide an understanding of: - Overall strategy setting the vision, objectives (expected results and overall outcomes) and the political priorities underlying OGD initiatives. - Strategic alignment and coherence between OGD and OG vision/strategy - the existing political leadership and support and foreseen actions to ensure sustainability in the long run and across political mandates. - Balance of OGD initiatives between central and local level (open data portals at the central level – single or multiple – and at the local level) will also form object of this section.
	Governance / institutional framework	This section focuses on: - Institutional arrangements and framework supporting data development, provision, distribution, cleaning, release approval, and promotion; and ensuring ownership, sustainability; balancing autonomy and control; and managing risks. - Accountability and responsibility frameworks. - Overall culture and practices of off-line information sharing and data re-use in the public administration.
Implementation	***Legal framework and policy environment***	Key objectives of this section are the understanding of: - Main relevant pieces of legislation (e.g. Access to Information Act, Freedom of Information Act, Open Government Act, Data Protection Bill) - Complexity of national legal frameworks (laws on copyright and as they apply to government data and information) - Extent of flexibility in existing regimes and regularity of updates of current legal and regulatory frameworks. - Disclosure policies (facilitating or impeding data transparency, sustaining smart disclosure) - Existence of restrictive copyright licenses; - Existing procedures and standards to deal with open data in government, and clear conditions on information re-use; legal exceptions to openness, etc.
	Technical issues and matters pertaining specially to data	This section will focus on technical matters that sustain or limit real data openness, i.e. data accessibility, availability, reusability. It will also provide an overview of current situation concerning: - Data quality - Data up-take - Interoperability (e.g. interoperability of data catalogues) - Workflow for data release and approval - Dataset storage - Data cataloguing and metadata

	Economic and financial	To complete a cost and benefits analysis, the main focus of this section is on: - Business case model adopted to clarify who bears the costs and benefits, to provide incentives to public agencies to open up their data, etc. - Financing mechanisms and costing model to sustain the OGD portal. - Establishment of an ecosystem ensuring value creation for the whole economy and society.
	Organisational	This section focuses on the measure taken to enable and foster the changes required in the public sector: - Measures and collaboration mechanisms in place to ensure accountability, quality of data, etc. - Measures to shift the culture of the public sector towards OGD, provide incentives to civil servants and build new capabilities. - Initiatives to ensure "buying in" of all stakeholders within the public sector.
	Communication and interaction	In light of the need to establish an ecosystem, this section analyses models for increasing public interest and for enabling and encouraging engagement of the wider community, specific segments of the population, civil society organisations, private sector, etc. The capturing of feedback on relevance of datasets, or perceived usefulness of existing engagement tools, to better comprehend the level of public satisfaction with the interaction with governments, will also be addressed in this section.
Impact		This section will analyse: - Main characteristics of the existing measures and mechanisms to appraise the real / concrete impact of OGD initiatives on: economic, political and social value creation. - Tools to establish and measure the correlation between OGD and service delivery innovation, as well as on enhanced collaboration and integration within t eh public sector. - Capacity to produce and use statistics on datasets usage and users to monitor user satisfaction, and understand where and how value can be created based on analysis of trends and patterns of up-take, e.g. considering data download data and observed uses of the provided government data and apps developed, analyse download and API statistics, key initiatives observed developed building on the OGD, discuss any existing data on tendencies in requests for access to government data - Comparison of the available OGD with international estimates of growth potentials in different public and private sectors and industries. - Conditions for ensuring impact for different users' groups (citizens, business, civil society organisations, government agencies) - Percentage of datasets released for a specific purpose (e.g. public service innovation, smart disclosure, transparency). - Exogenous cultural and societal factors (e.g. low ICT readiness of society at large, lack of trust towards public institutions, sense of opaque ownership of data, low interaction with public institutions) that may hinder use and consequently limit impact.

Table 2. Towards a set of metrics (indicators) on open government data

Dimension /component	Theme	Questions/data/metrics	Source
Policies and laws	- strategic alignment and consistency btw OGD and OG - strategic alignment btw national and local initiatives - AOIL - FOIL - OGD laws/legal provision - disclosure policies - standards - legislation on privacy, re-use fees, etc.	- How does OGD fit/support Open Government in your strategic view? - Do you have an OGD strategy/policy? - How many initiatives exist at the central and local level of government? What is the specific focus of the initiatives? - Right of access to information is included in the law - Right of access to information in electronic format is included in the law - Right of access to databases is included in the law - Right to access to information in machine-readable and/or open formats is included in the law - Is there a fee for obtaining data or part it (applied within and outside government)? - Does information come with restrictive copyright licenses (which prevent re-use)? - Are the conditions on re-use of information clear? - Have interoperable data-standards been developed?	OECD Survey 2013 [some data already available in G@G]
Technical	Accessibility availability open and machine readable format reusability	- Availability a. making data available of the web as downloadable files in well-known formats such as PDF, Excel, CSV, KML, XML, JSON etc. b. making data available of the web as Linked Data through RESTful APIs and/or SPARQL search interfaces. - Is there a search engine? What kind of search does the engine offer?	OECD survey 2013
Governance model*/institutional framework supporting data provision models	data provision (direct vs. indirect) Institutional framework Workflow for release and approval	- How is data provided? a. direct data provision, where data belonging to various public agencies is published by the one-stop government data portal. Subset of data_ was the portal developed through a collaborative approach to create ownership and sustainability? Subset of data_ Which authority hosts the portal? Subset of data_Which government agency/authority have published most of the data online? b. indirect data provision, where data belonging to various public agencies is published in a decentralized manner by these agencies (usually in their website)	OECD survey 2013

		while the portal provides some kind of linking mechanism and/or metadata for the identification of the actual dataset. Subset of data_ how many ministries/central government agencies have OGD portals? - What is the institutional framework in place to support OGD Initiatives? - What is the financial model in place to sustain OGD portal? - What is the business case model adopted to clarify who bears the costs and benefits, provide incentives to public agencies to open up their datasets? - What are the financial arrangement sin place to sustain OGD initiatives/central portal - What is the process for approval, cleaning and release of data?	
		Quantitative data - number of interagency collaborations (data on increase) - number of private-public collaborations (data on increase) - number of citizens-government collaborations (data on increase) - data on diversity of external partners - number of Ministries that have appointed Information Officers	
		Perception questions: *- level of integration of OGD enabling processes and emerging services*	
Organisational		- Are there specific programs that aim to foster cultural change/change of attitude of civil servants on availing information to the public - Are there programmes in place to enable public officers to change their attitude towards OGD? - Are there programmes to build civil servants capabilities and provide them with incentives to use data to create new services and value?	OECD Survey 2013
		Perception questions *- Cultural change in government agencies* *- Was a dialogue initiated among various stakeholders on the importance of sharing data and its benefits to the public?*	

Communication and interaction	- interaction : existence of web 2.0 embedded platforms to gather feedback and comments	Quantitative data - data on the number of forums - data on the average number of replies - data on the posts - data on the topics on which replies are provided - requests on datasets to be opened- number of communication channels - time duration of webpage view	Web crawl and log files, OECD 2013 survey
		Perception questions: *- overall user experience* *- perceived usefulness of public engagement tools and applications* *- public satisfaction with interactions with government*	
		Qualitative questions: - What is the mechanism in place to involve the various stakeholders'? - What ecosystem was created to make the OGD initiative sustainable? - Does the government promote accessibility to OGD for minorities (e.g. language options for content and access for the disabled including the hearing and vision impaired?) - Has research been undertaken to establish citizen's information needs and barriers to information use and re-use? - Are public-private partnerships actively sought to encourage innovation? - What are the programmes in place to encourage and enable community engagement in using OGD to participate in public affairs, develop new services? - Are there mechanisms in place to structure, track, trace and personalize the input received on OGD portal/initiative provided by civil servants at central and local level? - Does the OGD portal (or portals) include a "help" functionality? - Is there a FAQ section? Is there an opportunity to provide feedback on the data? - Is the portal connected to a social media platform?	
Political priorities	Transparency and accountability, participation, collaboration, public sector efficiency and effectiveness, citizens' quality of life, service delivery, innovation		OECD Survey 2013
Impact **	- on innovation - economic value and productivity	Qualitative questions: - What is the impact (social and economic) of the OGD initiative and how is this	OECD Survey 2013

	- social value - political - user centricity	appraised and measured? - What are the concrete accrued economic and social benefits of current OGD initiatives? Have these been proven? What are the measures in place to monitor and prove them? - Are current and potential re-use initiatives by third parties publicised to make stronger the benefits of opening up data? - How is the impact on collaboration and innovation measured? - Are statistics on users collected and analysed? - Are there statistics about dataset use (most popular) in place to understand use and how are they are being used? - What are the measures in place to encourage the emergence of more advanced features (beyond data delivery?)	
Data (availability. quality, uptake, re-use)			
Data (questions on national OGD portal)	- data availability (supply) Datasets available Datasets downloadable Datasets_API_callable	- Percentage of public datasets/ PSI available in electronic format - Total number of datasets available on the portal (published) - Subset of dataset available: numbers of datasets available in the following areas: trade data, fiscal data, health data, education data, transport data, census data, map data/geographic data, crime data. - Subsets of dataset available: number of dataset that can be downloaded - Subset of datasets available: number of datasets that can be accessed using APIs - Increase in the number of shared datasets - Metadata availability	Web crawl and log files, OECD 2013 survey
	- data uptake (demand) Views Downloads API_calls	- Total number of online views for all datasets. This data is typically available within governments - Total number of downloads for all datasets. This data is typically available within governments. - Total number of API calls for all datasets. This data is typically available internally. - number of total and unique visitors - percentage of repeated visitors - increase in the number of users - number of requests received by users for the release of	OECD Survey 2013

		new datasets in the past 6 months	
	- quality (data accuracy, consistency, update timeliness) and affordability	- frequency of data updates - information on the process and costs to maintain the data quality - data on the fees applicable for data access (by public authorities, the private sector, citizens).	
	- re-use	- number of apps developed based on the PSI/data published - subset of data_ apps developed by the civil servants - subset of data _apps developed by the private sector - subset of data_ apps developed by civil society organisations - subset of data_apps developed by individuals - data on data fields/areas used to develop the apps - number of new services created based on (built on top of) government published data - data on the fields where they were created (such as health, education, security, budgets) - number of apps downloaded and information on the fields	OECD Survey 2013 (STI?)

GLOSSARY OF TERMS

This Glossary was compiled for the purpose of this Working Paper and describes how the following terms are used in this document:

Application programming interface (API): A protocol intended to be used as an interface by software components to communicate with each other. Such interface helps developers extend reach of their apps and/or services.

Apps: computer software designed to help the user to perform specific tasks using smart phones.

Blog: A discussion, or informational site, published on the World Wide Web and consisting of a number of entries ("posts") typically displayed in reverse chronological order (the most recent post appears first).

Crowd-sourcing: the practice of obtaining needed services, ideas, or content by soliciting contributions from a large group of people, rather than from traditional employees or suppliers. This process can occur both online and offline. It combines the efforts of crowds of self-identified volunteers, part-time workers, etc., where each one on their own initiative adds a small portion that combines into a greater result. Crowd-sourcing differs from an ordinary outsourcing since it is a task or problem that is outsourced to an undefined public rather than to a specific named group.

Datasets: a collection of data, usually presented in tabular form.

Data analytics: a process of inspecting, cleaning, transforming, and modelling data with the goal of highlighting useful information, suggesting conclusions, and supporting decision making and achieved through multiple facets and approaches, encompassing diverse techniques under a variety of names, in different business, science, and social science domains.

Data mining: the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data processing and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

Geographic Information Systems (GIS): In a general sense, geospatial tools are information systems that integrate, store, edit, analyse, share and display geographic information for decision making. GIS applications are tools that allow users to create interactive queries (user-created searches), analyse spatial information, edit data in maps, and present the results of all these operations.

Linked data: A term coined by Tim Berners-Lee that describes a method of publishing structured data so that it can be interlinked thus becoming more useful. It builds upon standard Web technologies such as HTTP and URIs, but rather than using them to serve web pages for human readers, it extends them to share information in a way that can be read automatically by computers. This enables data from different sources to be connected and queried.

Mashup: using and combining data, presentation or functionality from two or more sources to create new services through a web page, or web application. The term implies easy, fast integration, frequently using open application programming interfaces (API) and data sources to produce enriched results that were not necessarily the original reason for producing the raw source data. The main characteristics of a mashup are combination, visualization, and aggregation. It is important to make existing data more useful, for personal and professional use. To be able to permanently access the data of other services, mashups are generally client applications or hosted online. In the past years, more and more Web applications have published APIs that enable software developers to easily integrate data and functions instead of building them by themselves. Mashups can be considered to have an active role in the evolution of social software and Web 2.0. Mashup composition tools are usually simple enough to be used by end-users. They generally do not require programming skills; therefore, these tools contribute to a new vision of the Web, where users are able to contribute.

Metadata: Metadata attribute information to content and are related to the contents of works such as periods, authors and descriptions, and the information related to right holders and conditions for use. They are of big importance for content preservation because they give information on existing collections. Their harmonisation facilitates common access to, and search for, information as well as re-use of information. Meta-data helps to standardise data definitions and improve comparability.

Open source: In IT production and development, open source is a philosophy, or pragmatic methodology that promotes free redistribution and access to an end-product's design and implementation details.

Screen scraping: The act of capturing data from a system or program by snooping the contents of some display that is not actually intended for data transport or inspection by programs.

Search function: A search function that searches a Website/portal offers users a way to find content/data. Users can locate content by searching for specific words or phrases, without needing to understand or navigate through the structure of the Web site/portal. This can be a quicker or easier way to find content, particularly on large sites.

Smart Disclosure: the practice of expanding access to data in machine-readable formats so that innovators can create interactive services and tools that allow consumers to make more informed important choices in sectors such as health care, education, finance, energy, transportation, etc.; and enables them to benefit from new products and services powered by data.

Social Media: The means of interactions among people in which they create, share, and exchange information and ideas in virtual communities and networks.

Social network platforms/ service: An online service, platform, or site that focuses on facilitating the building of social networks or social relations among people who, for example, share interests, activities, backgrounds, or real-life connections. A social network service consists of a representation of each user (a profile), his/her social links, and a variety of additional services. Most social network services (e.g. Facebook, Twitter, Google+) are web-based and provide means for users to interact over the Internet, such as e-mail and instant messaging to share ideas, activities, events, and interests within their individual networks.

REFERENCES

Access Info Europe, the Open Knowledge Foundation (2011), "Beyond Access: Open Government Data and the Right to (Re)use Public Information", 7 January 2011.

Alavi, M., D.E. Leidner (2001) "Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues", in MIS Quarterly, Vol. 25, No. 1., pp. 107-136, March.

Danish Government, Local Government Denmark (2012), "Good Basic Data for Everyone – A driver for Growth and Efficiency", October.

Davies, T. (2010), "Open Data, Democracy and Public Sector Reform: A look at Open Government Data use at data.gov.uk", August, *www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf.*

Capgemini Consulting (2013), "The Open Data Economy. Unlocking the Economic Value by Opening Government and Public Data", *www.capgemini.com/resources/the-open-data-economy-unlocking-economic-value-by-opening-government-and-public-data.*

Center for Technology on Government (2012), "The Dynamics of Opening Government Data", *White Paper*, University at Albany, *www.ctg.albany.edu/publications/reports/opendata/opendata.pdf.*

Deloitte Analytics (2013), "Open growth. Stimulating demand for open data in the UK", *www.deloitte.com/view/en_gb/uk/market-insights/deloitte-analytics/bfb570a79416b310VgnVCM1000003256f70aRCRD.htm.*

Dekkers, G. et al (2006), "Measuring European Public Sector Information Resources", European Commission, *http://ec.europa.eu/information_society/policy/psi/docs/pdfs/mepsir/final_report.pdf.*

The Economist (2010), "Data, data everywhere", *The Economist*, 25 February.

European Commission (2011), "Pricing of Public Sector Information Study", *http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/11_2012/summary.pdf.*

Felten, E.W. et al (2009), "Government Data and the Invisible Hand", *Yale Journal of Law & Technology*, Vol. 11, p. 160, *http://ssrn.com/abstract=1138083.*

Harrison, T. et al (2012), "Creating Open Government Ecosystems: A Research and Development Agenda", Future Internet 2012, 4, 900-928.

Hendler, C. (2010), "Report card: Obama's marks at Transparency", *Columbia University Journalism Review*.

Kalambokis, E., E. Tambouris and K. Tarabanis (2012), "A classification scheme for open government data: towards linking decentralised data", *International Journal of Web Engineering and Technology*, Vol. 6/3, p. 266-285, June.

Lee, G. and Y. Kwak (2011), "An Open Government Implementation Model: Moving to Increased Public Engagement", IBM Center for the Business of Government, *www.businessofgovernment.org/sites/default/files/An%20Open%20Government%20Implementation%20Model.pdf.*

The National Cross Industry Working Group on Open Data (2012), "Open Data Ireland", A briefing paper, February, *http://data.fingal.ie/media/OpenDataIreland.pdf.*

Noveck, B. (2012), "Open Data – The Democratic Imperative", *http://crookedtimber.org/2012/07/05/open-data-the-democratic-imperative/.*

Office of the Australian Information Commissioner, Government of Australia (2013), "Open Public Sector Information: from principles to practice", February.

OECD (2012a), "New ICT solutions for public sector agility", report from the OECD E-Leaders meeting, 26-27 March, Mexico, *www.oecd.org/governance/eleaders.*

OECD (2012b), "Exploring Data-driven Innovation as a New Source of Growth", unpublished report prepared by the Science Technology and Industry Directorate for the "Technology Foresight Forum: Big Data", Paris, October.

OECD (2012c), "Exploring 'Big Data' for policy", 22 May, STD/CSTAT(2012)2.

OECD (2012d), "The Role of New Technologies for Strategic and Agile Public Governance", unpublished Issue Paper prepared by the Public Governance and Territorial Development Directorate for the "OECD E-Leaders meeting", 26-27 March, Mexico.

OECD (2011), "Measuring the Economics of 'Big Data'", DSTI/ICCP/IIS(2011)4.

OECD (2008), "Recommendation of the Council for enhanced access and more effective use of Public Sector Information", *www.oecd.org/internet/ieconomy/40826024.pdf.*

OECD (2006), Digital Broadband content: Public Sector Information and content. Available at: *www.oecd.org/dataoecd/10/2236481524.pdf.*

O'Reilly, T. (2013), "Government as a platform", *Open government: Collaboration, Transparency and Participation in Practice,* O'Reilly Media, Inc.

Robinson, D., H. Yu (2012), "The New Ambiguity of Open Government", *UCLA Law Review Discourse*.

Pollock, R. (2012), "Managing Expectations II: Open Data, Technology and Government 2.0 – What Should We, And Should We Not Expect", *http://blog.okfn.org/2012/09/13/managing-expectations-ii-open-data-technology-and-government-2-0/.*

Schellong, A., E. Stepanets (2011), "Unchartered Waters: The State of Open Data in Europe", Public Sector Studies, Business Solutions Technology Outsourcing, *http://assets1.csc.com/de/downloads/CSC_policy_paper_series_01_2011_unchartered_waters_state_of_open_data_europe_English_2.pdf.*

Slee, T. (2012), Open Data Movement Redux: Tribes and Contradictions, *http://whimsley.typepad.com/whimsley/2012/05/open-data-movement-redux-tribes-and-contradictions.html.*

Davies, T. (2012), "Supporting open data use through active engagement", Position paper: W3C Using Open Data Workshop, June, *www.w3.org/2012/06/pmod/pmod2012_submission_5.pdf.*

Tuomi, I. (2009), "Theories of Open Innovation",*www.meaningprocessing.com/personalPages/tuomi*.

UK Cabinet Office (2012), "Open Data White Paper: Unleashing the Potential", June.

Vickery, G. (2012), "Review of Recent Studies on PSI-Reuse and Related Market developments", Information Economics Paris.

Wonderlich, J. (2012), " Open Data Creates Accountability", *http://sunlightfoundation.com/blog/2012/07/06/open-data-creates-accountability/.*

Yiu, C. (2011), "The Big Data Opportunity. Making Government faster, smarter and more personal", *Policy Exchange.org.uk*, June, *http://policyexchange.org.uk/images/publications/the%20big%20data%20opportunity.pdf.*

White, T, (2013), *Hadoop: The Definitive Guide*, O'Reilly Media, Inc.

World Economic Forum (2011), "Personal Data: The Emergence of a New Asset Class", January, *www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf.*

---

[1] For example Civicus, Citizen's Watch, Open Society Institute.

[2] For example the Sunlight Foundation, the Open Knowledge Foundation, Open Forum Foundation, OrgPedia Project, etc.

[3] Raw data, i.e., unprocessed data, refers to a collection of numbers, characters and is a relative term; data processing commonly occurs by stages, and the "processed data" from one stage may be considered the "raw data" of the next. Field data refers to raw data collected in an uncontrolled in situ environment. Experimental data refers to data generated within the context of a scientific investigation by observation and recording. For more see: http://en.wikipedia.org/wiki/Data.

[4] Yiu. C, The Big Data Opportunity – Making Government faster, smarter and more personal, June 2011.

[5] http://resource.org/8_principles.html adopted in December 2007 (last visited on 24 July 2012)

[6] http://data.gov.uk/blog/new--public-sector-transparency-board-and-public-data-transparency-principles (last visit on 13 March 2012).

[7] Open Government Declaration, Open Government Partnership (September 2011), http://www.opengovernmentpartnership.org/sites/www.opengovernmentpartnership.org/files/page_files/OGP_Declaration.pdf. signed by the United States and seven other countries in September 2011.

[8] "Bristol Open Data Innovation";*www.connectingbristol.org/2010/06/07/b-open/ in* "The Open Data Economy. Unlocking the Economic Value by Opening Government and Public Data", Capgemini Consulting, 2013.

[9] See *www.gov.uk/government/news/* (last accessed on 2 March 2013).

[10] For more detailed information on this see the study on Measuring European Public Sector Information Resources, available from http://ec.europa.eu/information_society/policy/psi/docs/pdfs/mepsir/final_report.pdf

[11] See also "A Review of recent studies on PSI re-use and related market developments" available at: *http://ec.europa.eu/information_society/policy/psi/revision_directive/index_en.htm.*

[12] "CloudMade Apps"" ; *http://cloudmate.com/about* (last visited 3 March 2013).

[13] Ref "Does Marginal Cost Pricing of Public Sector Information Spur Firm Growth", Research Institute of Finnish Economy in "The Open Data Economy. Unlocking the Economic Value by Opening Government and Public Data", Capgemini Consulting, 2013.

[14] Spanish Open Data Portal Annual Report, "Characterization Study of the Infomediary Sector", July 2012.

[15] See also "Governments looking for economic ROI must focus on open data with business value", A. Howard's interview with H. Lewis" at *http://radar.oreilly.com* (last visited March 3, 2013)

[16] "Dekkers, G. et al (2006),"Measuring European Public Sector Information Resources" .

[17] Governments have adopted policies explicitly calling for public bodies to make information available in machine-readable formats. For example in December 2009 the UK government published a set of "public data principles" as part of the "Smarter Government" initiative. One of these principles states: "Public data will be published in reusable, machine-readable form". Similarly, the US Open Government Directive, also published in December 2009, explicitly states that Executive Departments and Agencies should take steps to make information available in machine-readable formats. The new New Zealand policy launched in July 2010 specifically requires public bodies to ―release information in formats which make the data easy to use, taking into account the wishes of likely users.

[18] BarCamps are events bringing together thinkers from government, academia and industry to share information on Government 2.0 initiatives that are already in process and collaborate on the use of social media tools and Web 2.0 technologies to create a more effective, efficient and collaborative government 2.0 an all levels (local, state, and federal).

[19] Open Data Challenge Website *http://opendatachallenge.org/* (last accessed 2 March 2013).

[20] *www.w3.org/DesignIssues/LinkedData.html.*(lats visited on 24 July 2012)

[21]For more on this check an online discussion on the topic at *http://lists.xml.org/archives/xml-dev/200211/msg01290.html* (last visit on 13 March 2013).

[22] For more information and detailed explanations on the data quality and enabled uses at the various stars' level visit: http://5stardata.info (last visit on 14 March 2013).

[23] Open Data White Paper – Unleashing the Potential, HM UK Government, June 2012

[24] *http://data.gov.uk/library/draft-code-of-practice-datasets*

[25]To read more on how to build the "Open (Data) Ecosystem" read a post by Rufus Pollock, co-Founder of the Open Knowledge Foundation, at . *http://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem.*

[26] Barack Obama and Joe Biden: Technology, *http://www.barackobama.com/issues/technology/* (last visited Dec. 2, 2008).

[27] Posting of Taxpayer Assets, tap@essential.org, to listserver@essential.org, SEC's EDGAR on Net, What Happened and Why (Nov. 30, 1993, 10:36:34 EST), available at *http://w2.eff.org/Activism/ edgar_grant.announce*.

[28] Regulations.gov, What Is on This Site, *www.regulations.gov/ search/this_site.jsp* (last visited Dec. 2, 2008).

[29] *See* Office of management and budget, Executive Office of the President, Expanding E-Government: Partnering for a results oriented government 4 (2004), *available at www.whitehouse.gov/omb/ budintegration/expanding_egov12-2004.pdf.*

[30] OpenRegulations.org, About This Site, *www.openregulations.org/ about/* (last visited Dec. 2, 2008).

[31] European Commission, "Pricing of Public Sector Information", October 2011.

[32] See, for example, costs and benefits analysed by Tim Berners-Lee at http://5stardata.info.

[33] On 13 August 2010 a position was posted on the for "Evangelist for Data.gov Open Government". Candidates were asked to show four very different capabilities: (1) extensive outreach and communications skills and experience; (2) extensive experience in designing and implementing open government systems; (3) a proven research record for identifying and developing new technologies; and (4) experience managing a complex data and information environment that encompasses both public and classified data. The job description also indicated that the Evangelist was to work with multiple parts of the government, thus underlining the importance of understanding the myriad policy issues inherent in the release of information key to Data.gov. The role was established also with the intention to spur knowledge dissemination and evangelization in relation to the development and use of Data.gov to gain a greater involvement of agencies and other stakeholders such as the open government community and the mash-up programmer communities. The Evangelist was expected to create excitement and drive around the program to facilitate practical field application of leading edge technology issues with important stakeholders. *http://www.data.gov/blogs (*last visited 11 May 2013).

[34] Most of these countries provide open data via participation and collaboration platforms: US: www.data.gov ; UK: data.gov.uk; Australia: data.gov.au; France: www.data.gouv.fr(last accessed 30 July 2012).

[35]In order to establish a framework for fair, proportionate and non-discriminatory conditions for re-use of information held by public sector bodies in the European Union, the European Commission adopted the "Directive 2003/98/EC which states in Article 1 as its main objective: to establish "a minimum set of rules governing the re-use and the practical means of facilitating re-use of existing documents held by public sector bodies of the Member States". This objective should be placed in the context of the wider goal of facilitating access to knowledge for citizens and business promoting the emergence of Community-wide services as an important part of the internal market. *http://ec.europa.eu/information_society/policy/psi/index_en.htm*.

[36] The German Law on the re-use of information for public bodies ("Informationsweiterverwendungsgesetz") implemented in December 2006 reflects the aims and goals of the EU PSI Directive. However, it neither includes elements to proactively provide government data to the public nor does it create the right of access to government information, while the application of the law assumes such a right is already in place. As a result, the decision as to whether official information may be re-used and the details of that use are subject to the discretion of the public authority concerned (source Unchartered Waters – The State of Open Data in Europe, Business Solutions Technology Outsourcing, 2011).

[37] *www.aporta.es/web/guest/guia_reutilizacion*

[38] *www.gfii.asso.fr/article.php3?id_article=3278*

[39] Digitaliser.dk/resources/559456

[40] Web Content Managers Advisory Council, Requirements Checklist for Government Web Managers, *www.usa.gov/webcontent/reqs_bestpractices/reqs_checklist.shtml* (last visited Dec. 2, 2008).

[41] Executive Order "Making Open and Machine Readable the New Default for Government Information" accessible at http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-.(last visited on 23 May 2013) . See also http://www.slate.com/articles/technology/future_tense/2013/05/open_data_executive_order_is_the_best_thing_obama_s_done_this_month.html (last visited on 23 May 2013)

[42] Interview with Shannon Spanhake, Deputy Chief Innovation Officer, City of San Francisco, April 2012.

[43] About Govtrack.us, *www.govtrack.us/about.xpd* (last visited Dec. 2, 2008).

[44] Public Sector Information Directive available at: *http://ec.europa.eu/information_society/policy/psi/actions_eu/policy_actions/index_en.htm*

[45] See France's Law of 17 July 1978 on Access to Administrative Documents, available at: *www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068643&dateTexte=20100720*.

[46] MICUS Management Consulting report on "Assessment of the Re-use of Public Sector Information (PSI) in the Geographical Information, Meteorological Information and Legal Information Sectors", published March 2009, available at: *www.micus.de/pdf/MICUS-Studie_PSI_EU_March_2009.pdf*.

[47] The suggested data collection takes into consideration existing open government data measuring exercises conducted by other international organisations, academic institutions, consultancy firms. These include: The EU (the MEPSIR Study), the World Bank (Open Data Toolkit), Socrata (2010 Open Government Data Benchmark Study), Listpoint (Inaugural Public Sector Benchmark survey on Open Data 2013), The WorlWebFoundation (The WebIndex 2012), Kalambokis, E and E.Tambouris ("A classification scheme for open government data: towards linking decentralised data"). Business Solutions Technology Outsourcing ("Unchartered Waters: The State of Open Data in Europe, 2011).