

**PUBLIC GOVERNANCE AND TERRITORIAL DEVELOPMENT DIRECTORATE
PUBLIC GOVERNANCE COMMITTEE**

OECD E-Government Project

Open Government Data

Towards empirical analysis of open government data initiatives

**OECD Workshop on e-government indicators
10 December, Paris**

This report highlights main principles, arguments, concepts and criteria framing open government data initiatives and the issues challenging their implementation across OECD countries.

The aim is to underline the opportunities emerging for policy makers from OGD and data analytics, while providing a note of caution on the challenges this agenda poses for the public sector.

The overall analysis paves the way for further empirical work to be conducted by the Secretariat in 2013 to: map initiatives across OECD countries and provide a typography of the same, identify good experiences and develop and test a methodology to analyse OGD's implementation and impact.

Delegates are invited to comment and discuss this report during the workshop and in writing before 18 January 2013.

Delegates are particularly invited to:

- *Provide data and information on national OGD initiatives;*
- *Provide feedback on the suggested Analytical Framework for National Open Government Data Initiatives;*
- *Provide feedback on the relevance and feasibility of suggested data collection*

Barbara Ubaldi (Barbara.Ubaldi@oecd.org; +33 1 45 24 15 26).

JT03332347

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

TABLE OF CONTENTS

INTRODUCTION 3

BACKGROUND 3

 The “Right to Information” and “Open Government Data” Movements 3

 Definitions and principles 4

 Defining PSI 4

 Defining Open Government Data 4

 Defining Big Data 5

 Setting the principles 5

 Why Open Government Data? 7

BEYOND ACCESS: WHAT VALUE? 8

 Creating value 8

 Value for whom? 9

 What value? 11

 Improve government accountability, transparency, responsiveness and democratic control 11

 Self-empowerment, participation and engagement 11

 Fostering innovation, efficiency and effectiveness in government services 12

THE USE OF OPEN GOVERNMENT DATA 13

 New types of actors, roles and relationships for open government data provision 13

 How is OGD being used? 15

ENABLING OPEN GOVERNMENT DATA 16

 Is data really Open? 16

 Prerequisites for data openness 16

 Conditions for use and re-use 18

KEY DIMENSIONS FOR IMPLEMENTATION 19

 Policy challenges 19

 Technological challenges 20

 Economic and financial challenges 21

 Organisational challenges 22

 Cultural challenges 25

 Legal challenges 27

NEXT STEPS IN THE ANALYSIS 29

 Overarching conceptual issues 30

 Addressing the emerging needs for governments and public at large 30

 Harmonising definitions and sorting out legal “conflicts” 32

 Advancing the empirical analysis and measurements 33

ENABLING AN EMPIRICAL ANALYSIS 34

REFERENCES 42

INTRODUCTION

1. Open Government Data Initiatives and in particular the development of Open Government Data (OGD) portals have been proliferating in the past decade 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 and lead to the achievement of their overall objectives is essential. This is particularly true in consideration of the role played by OGD in relation to Open Government policies in a broader sense.

2. With this paper the OECD answers the request expressed by the member countries at the Meeting of the OECD Network on E-government (E-leaders 2012 “New ICT Solutions for Public Sector Agility”, Mexico City, 26-27 March 2012) to conduct exploratory work on Open Government Data to be able to better understand overarching issues commonly addressed across the OECD.

3. This paper highlights the main principles, concepts and criteria framing open government data initiatives and issues challenging their implementation with the idea of underlining the opportunities emerging for policy makers from OGD and data analytics while providing a note of caution on the challenges this agenda poses for the public sector.

4. Finally, the overall analysis of key concepts and issues aims to pave the way for empirical analysis of OGD initiatives. This is why the paper suggests an Analytical Framework for OGD initiatives (to be applied to ex post and ex ante analysis of open data initiatives) and a related set of data on national open government data initiatives to be collected across OECD countries. The application of the Analytical Framework and the collection of data would enable to acquire a solid body of evidence which can ultimately lead to: mapping initiatives across OECD countries (a typography of initiatives) and developing a common set of metrics to consistently assess impact and value within and across countries.

BACKGROUND

The “Right to Information” and “Open Government Data” Movements

5. There are two main civil society movements which 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. The second is the open government data movement, which uses social and economic arguments to encourage the opening up of government data, namely that putting such information into the public domain benefits society and can stimulate the economy by allowing the possibility for private enterprises to create new products and services using public data.

6. There is significant overlap between the aims of both movements, in that both aim to increase transparency of government so that all members of society enjoy the inherent social and economic value of information generated and collected with public funds. Nevertheless, there are also differences in the approaches and strategies employed by each to argue for more open government. One of the main differences is that the 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 the open government data advocates focus on data that is held in government databases, and they are concerned with both the technical and the legal issues related to the access to and 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 which permit reuse.

Definitions and principles

Defining PSI

7. For the purposes of the OECD “Recommendation for enhanced access and more effective use of Public Sector Information” (OECD, xxxx) the latter is broadly defined as “information, including information products and services, generated, created, collected, processed, preserved, maintained, disseminated, or funded by or for the 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

8. The term "open government data" 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 US (see box below).

9. The two main elements of open government data are normally defined as follows:

- Government data: is any data and information produced or commissioned by public bodies.
- Open Data is data that can be freely used, reused and distributed by anyone subject only – at the most – to the requirement to attribute and share alike.

10. In general, adopted definitions of “open data”, which apply both to data in raw and processed form, 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 “open government data” (OGD) are varies.

11. Public datasets being considered as a reference for analysis concerning OGD initiatives normally 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 (economic, employment, health, population, public administration, social);
- Transport information, including information on traffic congestion, work on roads, public transport, and vehicle registration.

12. However, well into the 2000s, the concept of “open government” among public officials was still centred on fresh disclosures, rather than improved access to already-public data. In the past decade we have assisted to a shift from a prevalent commitment to Open Government (focused on increasing transparency and public accountability) to an increasing commitment to Open Government Data with an agenda strongly driven also by innovation, efficiency and flexibility in government (Robinson, Yu 2012). To date, many practitioners consider OGD as 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 improved relations between governments and citizens.

Defining Big Data

13. The modern world generates a staggering quantity of data and the government is no exception. Across the public sector, extraordinary quantities of data are amassed in the course of delivering public services – from managing welfare payments and the National Health Service, through to issuing passports and driving licences¹. 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, but most importantly their social and economic *value*, 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 advantage and for driving innovation, sustainable growth and development (OECD 2012b).

Setting the principles

14. Different set of principles have been put forward by several non-governmental organisations, and eventually been picked up by Governments to guide the implementation of OGD initiatives.

15. Eight Open Government Data Principles were defined and put forward for governments’ consideration², in December 2007, during a Open Government Working Group Meeting held in Sebastopol (California), which gathered 30 open government advocates and organised by the 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 is made available. Public data is data that is not subject to valid privacy, security or privilege limitations.
2. **Primary:** Data is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.
3. **Timely:** Data is made available as quickly as necessary to preserve the value of the data.
4. **Accessible:** Data is available to the widest range of users for the widest range of purposes.
5. **Machine processable:** Data is reasonably structured to allow automated processing.
6. **Non-discriminatory:** Data is available to anyone, with no requirement of registration.
7. **Non-proprietary:** Data is available in a format over which no entity has exclusive control.
8. **License-free:** Data is 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.

16. The Sebastopol list has since then been updated by the [Sunlight Foundation](#) and ten principles have been identified which are meant to provide a lens to evaluate the extent to which government data is 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.

17. On July 14, 2011 the then US Federal CIO Vivek Kundra identified these ten principles for improving federal transparency in his [testimony](#) before a [House Committee on Oversight and 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 reuse 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 3rd party applications – Make data machine-readable to allow the public to easily analyze, visualize 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>

18. In June 2010 the UK published the Board's Public Data Principles which have been the foundation guiding the public sector in publishing data³.

Why Open Government Data?

19. Many public organisations collect and produce a broad range of different types of data in order to perform their tasks. The quantity and centrality of data collected by governments make these data particularly significant. This remains still, however, an uncharted territory and there is a lot of un-tapped potential which can be unleashed if government data is really turned into open government data. This can only happen if clear boundaries and criteria are set, e.g. in order to protect privacy and/or because of compelling policy reasons, and if all restrictions (legal, financial and technological) to the access to and re-use of authorised data are overcome.

20. Public agencies' vast amounts of data are an important national resource which can be utilized to help the public better understand what the government does and how well it performs, and to hold it accountable for any wrongdoings or unachieved result. In addition to increasing governments'

transparency and public awareness of government programmes and activities, opening up data can also help to generate insights into how to improve government performance.

21. Increased data transparency provides the basis for the public to participate in and to collaborate on government work to spur innovation and create value-added services. Additionally, data openness is eventually expected to improve decision making of both governments and individuals. The public is expected to be able to use government data to make better decisions and improve the quality of their lives (for instance access to specific databases or use of mobile apps developed based on newly developed datasets can allow more informed decision on daily choices); while governments are expected to be able to access more easily a wider range of datasets to foster evidence-based decision making. To foster the public's effective use of government data, the data needs to be relevant, easily accessible and usable. Additionally, it is important for the agencies to seek feedback from the public on the usefulness, relevance and accessibility of their data for continuous improvement. As the public starts to engage in government work through accessing data, increased data transparency enables government agencies to begin to shift their organisational culture not only towards openness, transparency and accountability but also to sharing and collaboration.

BEYOND ACCESS: WHAT VALUE?

22. The section on "Why Open Government Data" has underlined a number of values to the benefit of a number of stakeholders, which the sections below describe more in details.

Creating value

23. As the amount of data in the Information Economy explodes (The Economist 2010/OECD Issue Paper), public agencies have increasingly been focusing on increasing transparency of government processes and performance by publishing relevant data online and sharing it with the public. Public data made available in machine readable linked datasets, which can also be searched and manipulated using standard tools, is increasingly being considered as a critical new resource to fuel changes in value creation, in economic, social and political terms.

24. The evidence-based economic and social case for open public data is being established with increasingly solid evidence from many national and local governments, e.g. many US cities including San Francisco, the UK. A European Commission study predicted that open data across Europe will increase direct business activity by up to €40 billion per year (0.3% of total GDP), whilst indirect business activity will push this to €200 billion per year (1.7% of total GDP)⁴. However, potential value creation is also expected in political and social terms.

25. In relation to this goal the two most important tasks 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 use by the various actors;

26. Creating value implies indeed a good understanding what data users need and how data is created, protected, shared and used. There are different sets of data that have proven of great value if made open:

- - 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 (e.g. data on mortality for a General practitioner or in public hospitals)
- - Personal data of public services' users
- - Public cultural information

27. Implementing 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 is the data located? Who owns it?), the role of various actors and assessment of the return on investment are all key factors to enable open government data. Finally, understanding and fostering data use is critical to create value, as well as creating an ecosystem that supports the data creation process.

28. Finally, it is now widely considered that the income that can be generated by commercialising government data is only one of the potential benefits of OGD. Recent economic analyses show that when information is provided to the public free of charge or at very low cost, then individuals and private enterprises can take that information and create added-value products which they can then market. This economic activity stimulates the economy and also provides revenue to the government in the form of taxes. In addition, the information now available free of charge from the public bodies can be used for other civic society projects. This is particularly true for data that has broader potential value include mapping, meteorological, legal, traffic, financial, and economic data. Much of this raw data could be used for, or integrated into, new products and services which we use on a daily basis, such as car navigation systems, weather forecasts, or financial and insurance services. According to a survey conducted by the European Commission in 2006, the overall market size for PSI in the EU is estimated at EURO 27 billion⁵.

29. In Spain, for instance 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 opening its government data currently enables up to EUR 0.5 billion in private sector turnover per year; those companies employ a total of around 4,000 people. 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)

Value for whom?

30. Public sector information is a strategic resource holding great potential for different public sector agencies, private businesses, the academia, citizens and civic organisations. In many countries, the open data community still appears to be, however, uncoordinated and not unified. It includes IT professionals, both small and large companies' 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 open data 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 secondary information-type products. Civil society organisations are keen in having a number of datasets that, if combined, may help improve life and service delivery to certain segments of the population or to certain neighbourhoods.

- **Government:** At a macro level, OGD provides the scope for new ways of conducting “government business” 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 deliver smarter, innovative and more personalised public services while improving the quality of interactions between the governments and the users. OGD equally enables increase in transparency that all in all contributes to strengthen accountability. Additionally, there is an important benefit in terms of legitimacy that governments gain by appearing to be more open.
- **Citizens:** open government data is expected to enable public participation in designing responses to public needs (e.g. co-production and delivery of services through newly developed apps); and to enable sourcing information and knowledge from the public to make more informed policy choices and drive efficiencies in government. The application of data, technology and analytics can cut paperwork, get questions answered more quickly, help people find and claim the benefits they are entitled to, and tune front-line services more closely to individual needs and behaviours. 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 UK and Chicago’s 311 Internet portal lever mobile government and open data. 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, increasing public participation in designing responses to public needs (e.g. co-production and delivery of services through newly developed apps); and sourcing information and knowledge from the public to make more informed policy choices and improve efficiencies in government. As such, OGD has the potential for helping improve citizens’ quality of life.
- **Civil Society:** OGD initiatives by civil society can be found across many OECD member countries. The shared goals of these initiatives include demonstrating the benefits of open government data to government and the public, although there is fragmentation in terms of goals they specifically target which range from increasing transparency, to improve service delivery particularly for vulnerable segments of the population, to protect the environment or sustain growth. And it has been observed that this fragmentation does not help to move items higher in national political agendas. Additionally, civil society organisations can play a pivotal role as intermediaries in the identification of key datasets that, if open, can 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 US, the Open Knowledge Foundation in Germany, the Open Rights Foundation in the UK, etc.
- **Wider economy, private sec and public service market place:** OGD can stimulate a competitive marketplace for public sector services for instance. The common belief is that innovators from outside governments are provided with the opportunity to develop modular services which are more agile than those developed in-house by the governments. When data is 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. 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?

31. The sections below highlight some of the main values targeted by OGD initiatives across OECD member countries.

Improve government accountability, transparency, responsiveness and democratic control

32. Strong supporters of Open Government Data as key enabler of Open Government believe there is a correlation between lack of open government data and the levels of corruption in a given country. For instance, a common assumption is that the lack of data in the public domain allows public servants to engage in corrupted behaviour with impunity. OGD can be a powerful force for public accountability by making existing information easier to analyse, process and combine, allowing a new level of public scrutiny. Supporters of Open Government Data (OGD) believe indeed that OGD can promote greater transparency, which can increase public accountability. This can raise the level of public trust and perceived responsiveness of government actions. The Open Government Declaration⁶ 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 Open Data can each exist without the other: 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 has indeed the potential to improve service delivery and citizens' quality of life, but may have little impact on political accountability.

Self-empowerment, participation and engagement

33. A second assumption is that Open Government Data 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 for citizens' engagement in public life, e.g. in policy making and service design and delivery. The expanding use of new technologies, combined with the rising of the OGD movement are seen as key enablers and drivers of self-empowerment, higher e-participation and public engagement. Open Government Data initiatives, particularly as they are supported by Web 2.0 and social media applications, are creating an architecture of participation which enables users not only to be passive consumers of content and services, but also active contributors and designers of their own right.

34. This means that legitimate stakeholders are invited more openly into a participative and empowering relationship with government in terms of:

- workings arrangements of the public sector and public governance more widely
- planning and land use issues
- service design and delivery
- community building
- dispute and conflict resolution broader public policy and decision making as part of the overall democratic process.

35. Open Government Data, and in particular smart disclosure, empowers citizens also to take more informed decisions which can enhance the quality of their lives. For some of this to happen, governments need to enable users to have access to their own data and decide how to use it to improve lives (e.g. the Blue Button Initiative in the U.S. to give veterans complete control of their personal health record held by the public sector; or the Green Button, also in the U.S., which is a similar initiative around individuals' energy use data).

36. Equally important to empowering citizens is to empower the public sector workforce. Opening up government data can enable civil servants, many of whom are frontline professionals, to themselves participate in ensuring that government is open and participative, and to develop applications that better responds 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 which could improve service experience if they were given the data, tools and incentives to do so, for example by being enabled to participate in a professional capacity in citizens' social networks to offer advice and knowledge.

37. Moreover, many civil servants also see a blurring of their personal and professional lives in terms of the tools they use, which could improve both 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 requires also that civil servants be equipped with the necessary skills, tools and mechanisms (Millard 2012) and guidelines.

38. But for this to happen, strategies and programmes are needed to build the next generation of civil servants. New skills are needed, which are not only strictly IT related. They should indeed enable: 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 IT use to solve them (e.g. cybercrime investigation).

Fostering innovation, efficiency and effectiveness in government services

39. OGD can increase government efficiency, effectiveness and innovation in service delivery and internal public sector operations. Even though the release of data online can raise a number of substantive enquiries in terms of government activities, from a service delivery perspective its re-use can also lead to a significant reduction of the questions routinely received by public authorities, thus reducing work-load and costs. Additionally, the remaining questions concerning service delivery per se would be easier for civil servants to answer as it would be clearer where all the relevant information can be found. 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 and ultimately make government departments smaller, while encouraging external inputs and new sources of knowledge, possibly making them more innovative. In addition one could argue that the co-development of knowledge in this example increases not just the quality, but also the awareness of the Dutch public authority's work, thereby further increasing its value and relevance.

40. Moreover, OGD can help bringing down silos and foster collaboration across and within public agencies and departments. As common, or shared, datasets and/ or registers are being created collaboration and exchange on who owns which public information and for what purpose is needed, which provides an opportunity to also re-engineer and simplify internal procedures. Furthermore, as public resources are freed from having to maintain individual registers and datasets they can be reallocated to more productive tasks. Finally, OGD has provided a platform for innovating in service delivery. This has not only resulted from

the re-use of data by private sector actors or civil society organisations, but also thanks to the re-use of civil servants who have in several instances taken the initiative for example to develop new apps.

THE USE OF OPEN GOVERNMENT DATA

New types of actors, roles and relationships for open government data provision

41. The distinctions between professional, politician, practitioner, civil servant, expert, consumer and citizen, are blurring dramatically. These roles are still important but the (power) relationships between them are changing and any given 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 which were previously the preserve of the public sector or 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 PSI.

42. Governments, particularly at local level, are an important convener, facilitator, enabler, partner, participant, etc. Government agencies produce data of various types and qualities and many time their focus lies on the provision of the information or documents rather than data which supports the process of delivering the former. Agencies also tend to treat their data as a commodity to generate revenue or cover the cost of data production. As producer of PSI and data, governments will likely retain responsibility for overall quality standards. For example, public services are likely to continue to be the main funder of open government data initiatives, whether or not they themselves design and deliver them. They will remain the only legitimising organisations with democratic accountability to act fully in the interests of the whole of society and to manage the risk of providing poor quality or erroneous data.

43. However, in order for public data to benefit from the same innovation and dynamism that characterise private parties’ use of the Internet, it is critical for government to rethink its role as an information provider and realise that pushing out data is not enough to create value. As mentioned earlier, it is currently being argued by many that rather than struggling, as it currently does, to design sites that meet each end-user need, it would be more efficient for it to focus on:

- creating a simple, reliable and publicly accessible infrastructure that “exposes” the underlying data;
- identifying relevant datasets through public consultation;
- increasing data release in open formats.

44. There are many private actors, either non-profit or commercial, that are well suited to deliver government information to citizens in ways and forms that they want, and that can 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.

45. 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 third parties can qualify them and prove their identity.

46. 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 make available to the public at large. Based on the engineering principle of separating data from interaction, which is commonly used in constructing Web sites or Service Oriented Architectures⁴, many private sector players argue that private parties are best equipped to build Web sites 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 services such as advanced search, automated content analysis, cross indexing with other data sources, and data visualisation tools. These tools and opportunities for exploring new form of private-public partnerships and collaboration are promising, but it is far from obvious how best to combine them to maximize the public value of government data. Given this uncertainty, rather than deciding up front the best option, it would be wiser for governments to stimulate private parties' engineering of ideas to experiment and discover what works.

47. Meanwhile, 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 to great lengths to reassemble data that government bodies already possess but are not sharing in a machine-readable form. An important advantage which can be derived by involving private parties to present data to citizens arises from the fact of encouraging the emergence of more advanced features beyond the simple delivery of data, such as: advanced search, RSS feeds, links to information sources, mashups with other data sources, discussion for a and wikis, visualization, automated content and data analysis, collaborative filtering and crowd sourcing analysis

48. Citizens, communities, civil society organisations, as well as 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 for 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 and capture the relevant talent outside any organisation.

49. An important new role for government therefore is to act as an open participation and collaboration platform, both online and offline. It is important for it to pro-actively collaborate with companies including SMEs, civil society organisations, 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, arbitration is essential to produce value. Standardised modules for basic functionalities which 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 need increasingly to crowd source content, services and policies. In a fast changing context and in light of the continuing time 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); now also all and any partnerships and groups (many hands) are needed.

50. These developments which are largely beneficial, also give rise to threats and challenges, for example:

- Loss of control and blurred accountability (by whom to whom?)
- quality standards more difficult to determine and maintain and this data quality risk (outdated, inaccurate or incomplete data) needs to be envisaged and tackled
- privacy and data security
- danger of data and content mis-use, propaganda and subversion
- digital elite formation – new digital divides?
- information overload – or is this more a filter failure?
- crowding out other perhaps 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 update

How is OGD being used?

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

1. Data to fact – often underestimated in accounts of ‘data for developers’ – 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. Facts could be found through online interfaces, but also by browsing downloaded Excel spreadsheets.
2. Data to information – creating a static representation and interpretation of one or more data sources. Leading to visualisations, blog posts, infographics and written reports.
3. 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.
4. 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 API onto a dataset created, or an interface that makes it easy to download subsets of a large dataset.
5. Data to service – where OGD plays a ‘behind the scenes’ role in making some online or offline service function. 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 uses employ multiple processes.

52. 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 OPEN GOVERNMENT DATA

Is data really Open?

53. Making government data open requires a holistic approach as it involves the many stakeholders mentioned in the previous section. Lee and Kwak (2011) recommend that the implementation of open government data initiatives should be incremental because each stage is important. Accordingly, they propose a model that requires incremental implementation of the OGD initiative. The model presents four stages of implementation before full open government data is attained.

6. Stage 0: Getting a view of what government data exists – the data catalogue!
7. Stage One: Increasing Data Transparency
8. Stage Two: Improving Open Participation
9. Stage Three: Enhancing Open Collaboration
10. Stage Four: Realising Ubiquitous Engagement

54. 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 open government data projects in succession as they move towards higher levels of engagement and civic participation through the use of ICT.

55. It is evident that ICT plays a pivotal role in the development of OGD. ICT enables OGD by hosting and publishing content and exposing standards for data exchange that encourage free use, or embedding, of information in other packages for example radio, web, SMS etc. There are benefits in using ICT which range from improved timeliness of reporting, improved data quality by decreasing the manual interactions with the captured data, increased quality through the business intelligence inbuilt in the systems, among others. These capabilities are possible if the technologies are appropriately utilised to return maximum benefit.

Prerequisites for data openness

56. Reaping the benefits of the change brought about by OGD implies good understanding on how to create the preconditions for making citizens' engagement valuable. To this end, proactive release of data is only one of them. For instance, it is essential to engage the civil society to understand which data is needed.

57. Today, many governments focus on the development of the national open government data portal as if it were a higher priority than developing technical infrastructures that open up their data for others to use. Understanding the preconditions for effective open government data, e.g. opening up data and providing reusable data is just as important, if not more important, than setting up Web sites, and lies at the core of the government data publishing responsibility.⁷

58. 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 data portals over frameworks and infrastructures that will open up their data, the pace of change and real value will be retarded. Much of the current criticism on national open government data portals is 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.

59. The status quo has its virtues. 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 use (if not from home, then from a library or other public facility) without the need to understand the inner workings of technology.

60. There is also a certain economy to the current situation: Under the current system, the costs of developing 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 to move SEC data online⁸). The status quo also has marked drawbacks. The institutional workings of government make it systematically incapable of adapting and improving Web sites as fast as technology itself progresses. No one site can meet as many different needs as well as a range of privately provided options can.

61. It is widely recognised that in order to be considered open, government data need to be available to anyone who wants to access and use it and should be provided with highest details available and in a timely manner. To achieve this, the following elements need to be met:

a) **Availability and accessibility:** the entire data set must be available, preferably for free (or at no more than a reasonable reproduction cost), and be downloadable over the internet. The data must also be available in a convenient and modifiable form. Specific criteria include:

1. Data is easy discoverable and findable

1. government data catalogues [US, UK, Australia], information asset registers (IARs) [US, UK, Australia, Canada], citizen-driven catalogues [Canada, Germany]
2. technical issues related to discoverability (there are a number of format and tagging issues which make information easier to process with machines and hence easier both to locate and reuse, e.g. "metadata", "microformats").

2. Data is easily accessible

1. information is available in disaggregated forms
2. Data is available through bulk downloads (enabling access not just to one or two pieces of government data, but to entire databases/full datasets as programmers can use these to develop applications that make the most of publicly generated data).

3. information is released in a timely fashion: rapid disclosure allows builders of apps to have access to the very latest data
 4. information is interlinked: 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.
3. Data is affordable: it should be made available for free or at the marginal cost for distribution (free by web channels).
- b) universal participation:** 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.
- c) Reuse and redistribution:** the data must be provided under terms that permit reuse and redistribution including intermixing with other datasets. The core of a “commons” of data is indeed 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:
1. Data is available in electronic format and the right to access data in electronic format needs to be recognised
 2. Data is in machine readable format (‘screen scarping’ can be time consuming⁹) [e.g. unlike open/machine readable formats (e.g. XML, XSLT), PDF files are not machine readable (i.e. data cannot be processed on PCs using databases or spreadsheets software)]
 3. Data is released in open file formats (specifications have been made public, no need to buy a specific software to use the information) which are machine-readable. Even though no AIL gives the right to access to information in open formats, most OGD initiatives are starting to be accompanied by policy documents stipulating that official information must be available in open file format.
 4. Users have the right to re-use data

Conditions for use and re-use

62. The real value of OGD is realised when there is local capacity to re-use the data. ‘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 open data. When apps developed during these events use dummy data it does not help one to appreciate the value of the application.

63. Reusable data needs to be available. The challenge is that often when data is available it is not in a format that enables re-use. Many times this data is shared in closed formats like PDF and this does not enable re-use. The cornerstone of Open Government Data Initiatives is to enable government data to be accessible, usable and re-usable. This requires structures in place that will ensure that data is in a format that is acceptable for the three attributes. Having data that is 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. There have been instances where government officials have witnessed tremendous value out of the data they possess. In order to assess the degree of datasets re-usability Tim Berners-Lee proposed in 2010 a Five Star Scheme¹⁰. The UK Government expressed its intention to adopt such scheme as a measure of the usability of its Open Data¹¹.

64. OECD member countries have adopted FOIA to ensure public right to access information. Freedom of 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 has not given an automatic right of re-using information. FOIA have therefore undergone, in many instances, amendments when needed in order to require release of dataset in a form that can be used and re-used by specifying the licence it can be used under, and in a re-usable format. Such FOIA are meant to create the conditions for accessing raw un-manipulated datasets.

65. In order to give the public an enhanced right to data, some countries such as the UK – which is generally recognised by peers as being at the forefront of promoting re-use of PSI – plan to expand the 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 government plans to offer the opportunity to shape the guidance using a crowd sourced wiki to be launched on the government open data portal data.gov.uk.

Key dimensions for implementation

66. Even when the right legal framework is in place, OGD initiatives can be undermined by problems with implementation, technical challenges and administrative delays. Technological, legal, financial restrictions, among others, may limit data access and reusability (e.g. making it difficult funding data or finding valuable ways to reuse data). A number of challenges may be associated with the implementation of open government data initiatives which, if not properly tackled, might obstruct or restrict the capturing of benefits of national efforts aimed at spurring OGD. Addressing various challenges related to technology, financing, organisation, culture, policy, and legal frameworks is essential to create an ecosystem, and build sustainable business models for OGD initiatives that can bear the desired benefits.

Policy challenges

- **Disclosure policies** may limit data transparency and copyrights may cause lack of clarity on who owns government's data thus restricting the right of the public to use government data (e.g. info is sold or comes with restrictive copyright licenses which prevent 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 from copyrights and changes for commercial use. In particular, policies should enable smart disclosure, i.e. the release of data in a way that enables better choices in individuals' lives (e.g. the health and consumer service departments in the US 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).
- 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) can compromise the quality of the data and eventually the output of OGD initiatives.

Technological challenges

67. Government data often is not harmonised as every unit has its own set of data. This can make it difficult from the user perspective to know which piece of data is valid or should be trusted. Critical to access is to know the source of what you are searching, and in many instances where to start searching is a challenge. Accessibility can also be limited if PSI cannot be re-used, and data transparency might be hindered if data is not simple to be accessed due to its formats. Additional technology related shortcomings include the need to: improve information technology infrastructure, enhance privacy and information security, integrate open data government tools and applications.

68. A second 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. Therefore, a lot of emphasis is place 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. 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 3. The Case of Regulations.gov

Regulations.gov is a government-wide docket publishing system created in the US in response to the E-government Act of 2002 and launched in 2003. It is used today by most US Departments and Agencies¹² and the policy of the Office of Management and Budget (OMB) not only requires its use but also precludes the agencies from using “ancillary and duplicative” docketing and rulemaking systems of their own design¹³.¹⁵ This exclusivity rule, combined with the difficult interagency politics involved in honing system features, is 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.

Concerns about cost-sharing have also led the system to omit even features whose usefulness and desirability is a matter of broad consensus¹⁴. 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 its limited inventory of computer-readable data directly to the public (in this case, using a single RSS feed) which 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 [sic] down, simple-to-navigate listings of new agency dockets” and a more sensible set of RSS feeds, one for each individual agency¹⁵.

Source : OECD.

69. Lastly, interoperability remains an unresolved issue in e-government; and dealing with OGD in general, and open data file formats in particular, can facilitate IT system interoperability in government. Interoperability is a concern high in the agenda of policy makers tasked with the implementation of OGD. For instance, interoperability of data catalogues, or the creation of a pan-European data catalogue, are for instance a big challenges EU policy makers are facing at the moment.

Economic and financial challenges

70. Several economic and financial challenges are hindering a fast-paced development of OGD initiatives in several OECD countries. Key aspects are highlighted below.

- Revisiting financing and costing models and establishing the right ecosystem. 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 indeed that data which 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 which need to be considered. This is combined with the difficulty of establishing a clear measure of the gains emerging from opening up government data. Such matters raise the need to be mindful about implementing OGD in pragmatic and affordable ways, which do not add unnecessary burdens and potential loss of revenues. There is a substantial commitment and investment on the part of agencies as they need to acquire new skills, train employees, purchase technologies, and upgrade network infrastructure, which need to be accounted for. There are indeed human-resource costs associated with organising and preparing information to be put online - particularly if the decision is taken to develop a special portal which may require an IT and design team and also with ensuring timely update of data. The additional costs to timely publish or coherently produce high quality are normally held by each agency although governments such as the Dutch and Danish government are looking into the development of a business case with funding and alternative financial models.
- Converting large volumes of data into re-useable formats can have cost implications, particularly if there is a high level of use of proprietary software. Initiatives such as converting government data to semantic web and linked data formats can be time consuming and therefore costly as well as enabling partial access to large volumes of data (e.g. a field in a database which contains personal data such as the e-mail addresses of private individuals which 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). Due to these additional costs there is some reticence on the part of government bodies, which can result in refusal of even partial access to a requested database. One 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.
- In the longer term, one solution is to design databases with the right of public access in mind, which appears to be increasingly easy, at least from a technical perspective. It is possible, for example, to build a database which 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 is shared, the e-mail addresses can be separated, thus ensuring that any personal data is 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 – the vault where the money is kept – is locked and secure.
- Similarly, there are many solutions to releasing information which come at a very low cost and it would be advisable that these were seen as part of the day-to-day activity of public bodies, such as posting full datasets in open source formats on government websites, properly tagged with metadata so that the information can be found, but with no other special formatting or presentation.
- The business model for OGD also needs to take into account where potential benefits may accrue and how to align funding and incentives. When government provides reusable data, the practical

costs of reuse, 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 Open Government Data Web sites to the infrastructure that allows new sites to be created. If the creation of infrastructure causes better third party alternatives to emerge then the government entity can cut costs by limiting its own. This underlines the need for better data on 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 wide government IT procurement strategy.

Box 4. The Evangelist for US Data.gov

On August 10, 2013 a position was posted on the for “Evangelist for Data.gov Open Government”. The job description indicated that the candidate to the role of Evangelist for Data.gov was required to show four very different capabilities: (1) extensive outreach and communications skill 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) managing a complex data and information environment that encompasses public to classified data. The job description also indicated the importance of the extensive work of the Evangelist with multiple parts of the Government thus underlining the importance of understanding the multitude of policy issues inherent in the release of information key to Data.gov. Hence, the role required knowledge of, and access to, an extensive network of people organizations and experience leading many areas of public outreach and engagement that can be brought to bear for 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 announcement clarified also that 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.

Source : <http://www.fbo.gov>

71. 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 the government IT costs will decline in those areas where private actors have the greatest interest in helping to leverage the underlying data, while the 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 Web site based on the data. Governments are keen to collect evidence on recent initiatives showing 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

- **Ensuring accountability, quality of data and responsibility in a context of collaboration.** Given the complexity and cross-cutting nature of OGD, government need to put in place the appropriate institutional structures. Tasking a government body, often the Centre of Government (e.g. Prime Minister’s Office), with championing, co-ordinating providing support and leadership on open government data initiatives and programmes has been seen as way to bring the various stakeholders on board. Having a ministry (dedicated body) in charge for soliciting from governments the various datasets that will then be made public has been additionally considered as a way to ensure timely and full compliance with the national strategic directions. This can sustain collective work to strengthen data integration across different parts of the public sector, help building better capacities across government to deal with emerging concerns (e.g.

privacy/transparency) and ensure that those making decisions about the release of data do so in a rigorous and consistent fashion.

- **Balancing autonomy and control.** Empowering independent oversight bodies to demand and to publish information on budgets, procurement and expenditures is considered as crucial to ensure data transparency. Several governments, *e.g.* the UK, are considering establishing independent ethics and governance groups to oversee policies and procedures for improving the use of administrative data.
- **Ensuring sustainable change through ecosystems.** The creation of the right ecosystem is essential not only to reap the economic benefits, but also to create 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 developed based on them, are essential factors to make OGD initiatives sustainable and to create value. The development of the app “Asthmopolis” in the U.S. is an excellent example of an app developed thanks to the ecosystem and which 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 in 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. However, establishing the right ecosystem is not so evident, and their creation implies involvement of all actors and the provision of the right business case to spur usage. Establishing the right ecosystem entails identification of the various categories of actors, the adoption of policies enabling 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 which might normally be less actively involved in public affairs (*e.g.* civil society organisations operating in geographically remote areas and as such more aware on data that might be needed to develop target services that would better serve the local community).

72. We can identify at least three categories of ecosystems:

- a) **Ecosystem of data producers** (*i.e.* public sector, academia, media, private sector).
- b) **Ecosystem of infomediaries** (*i.e.* media, developers, civil society) which is essential to make sense of and create value out of the 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) and can play a critical role to build capacities at the community level and create a culture that appreciates the relevance of the data.
- c) **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 to data mining and as facilitator of accessibility to data, particularly in remote areas of the country thus enhancing cost – effectiveness of access.

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

74. 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 a real driver being the open-minded and collaborative attitude of the City authorities. Examples of places such San Francisco are not unique. It presents indeed many elements which typify several OECD medium-sized cities, as well as large municipalities. A way to replicate the positive experience might therefore be adopting a strategy that leveres these conditions where they exist, or fosters their development where they are lacking. The establishment of collaboration frameworks may also help ensuring involvement of different actors (e.g. also SMEs which may be important incubators for innovation but still little aware of opportunities brought about by OGD).

- **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 also need to be in place to allow two-way dialogue to take place between the public sector and the users of government data (*i.e.* individual developers, SMEs, citizens, civil society organisations, academics and large companies). This is key for governments to show a focus on user need and for users to provide feedback on the datasets they would like to see released as a priority as they considered them of a greater value or more likely to be used by the community. Capturing of feedback may result in value creation as they 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.

75. The UK government is for instance working on a “Government Developer Engagement Strategy, setting out principles for how individual government departments are expected to engage with the development community. Several governments’ initiatives launched competitions with the intention to encourage OGD reuse (*e.g.* the Apps for Democracy run for a 30-day time period by the US government – which apparently led to an estimated 4, 000% return on investment - or the similar Norwegian initiative Nettskap 2.0 which resulted in the development of 135 apps) or have fostered close collaboration between individual civil servants and public sector bodies with civil society (*e.g.* 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 the running of the data catalogue overheidsfeeds.nl or the event BarCamp on Open Government).

76. Social media can play an important role in inspiring or enabling many OGD uses thus underlining the relevance of involving communities of practices to sustain OGD initiatives and help the creation of a network of actors. Capturing the users’ feedback can also help creating a need for use, *i.e.* get the data where people really need them. However, engaging users requires resources. In order to ensure the capturing of Open Data users’ view the UK have established a group in the UK Cabinet Office which comprises 14 officially selected volunteers from the civil society, the private sector, etc who play an advisory role on the data the government should release.

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

Box 5. The UK Open Data White Paper: Unleashing the Potential

In June 2012, the UK Cabinet published the Open Data White paper, which set out how the Government intends 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 (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 UK. In order to frame a feasible public sector implementation of Open Government Data, 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 : <http://www.cabinetoffice.gov.uk>.

Cultural challenges

77. Legislation, IT platforms and codes need to be matched by a culture within the public service that supports a presumption to publish/release data.

- **Increasing public interest and engagement.** Raising capacity and awareness of civil servants, citizens and the private sector (SMEs) on their rights is important for the society as a whole to fully capture the benefits of Open Government Data 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. Additionally, undertaking research to establish citizen's information needs and barriers to information use and re-use, or seeking public-private partnerships to encourage 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.
- **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 data, create value and exploit it. This is not necessarily easy as successful crowd sourcing also depends on a sufficient scale and representativeness of participation to get valuable results. A critical new resource to fuel such changes is public data made available in machine readable linked datasets which can also be searched and manipulated using standard tools. To date, there are still only a limited number of governments which have substantially embarked down this path, and even fewer local and regional governments where the benefits are likely to be greatest. U.S. cities and the U.S. Federal government, the UK, Australia and France, as well as a handful of others, have been leading the way in this aspect¹⁶. Companies and SME start-ups in some countries are exploiting such data to expand business and create jobs, whilst a few governments are using such data to encourage innovation camps, 'hackathons', and other competitions to create apps, services and contribute to policy making.
- **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 represents much more than technological challenges. They are indeed not merely technical issues

and solving these matters implies therefore 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 interpretational sovereignty. Belief that making data public dis-empowers 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 to change the attitudes of public officers on availing information 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 on the front line and in central governments.

78. Additionally, governments must have the capability to conduct, interpret and consume the outputs of data and analytics work intelligently. This includes also the government capacity to debate the meaning of data and find ways to use it in democratic debate. 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.

79. Furthermore, even though a conception of which data are available is an essential step for any government's OGD strategy, currently most governments do not have a comprehensive overview of the data in their possession. The UK Government organised for instance information engineering programmes which forced more than 100,000 authorities to reengineer their records and which is considered having 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 one.

80. Finally, particular efforts have been undertaken by many governments engaged in the establishment of OGD portals (e.g. this is the case with data.gov.uk) to encourage use of linked data. The skills, experience for working with linked data are still at relatively early stages of development, but advocates of linked data approaches believe it has the potential to enable a revolution in how data is accessed and utilized (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). The following figure is an example of how the linking can be depicted. More specifically, data from the Department of Education describing schools is linked to data from the office for National Statistics. The 'joint point' of these datasets is the Local Learning Skills Council (LLSC) that is responsible for the specific school.

- **Ensuring the “buy in” of all stakeholders.** Initiate dialogue among various stakeholders about the importance of sharing information and its benefits to the public can help securing their participation and ensuring their 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.

Figure 1. Screen-shot showing the linking of data from the Department of Education to data from the Statistics Office

The screenshot displays the HM Government Linked Data API search results for 'Herne Hill School'. The page features a dark header with the HM Government logo and the text 'Linked Data API'. Below the header, a 'Search Results' section is visible, with a 'Show Search Form' button in the top right corner. The search results are presented in a table format with the following data:

Herne Hill School	
easting	532186
northing	174572
lat	51.45471
long	-0.09888
establishment number	6375
unique reference number	100866
LLSC	London Central

To the right of the table, there is a map showing the location of Herne Hill School, with a red pin and a green square indicating the school's location. The map is labeled 'Herne Hill' and includes a search icon and navigation arrows.

Source: <http://data.gov.uk>.

Legal challenges

81. The legal landscape surrounding data sharing and opening is undoubtedly complex¹⁷. Having a consistent legal framework in place is critical to facilitate PSI accessibility and re-use and to improve secure data sharing between public authorities and with the wider community to improve insights, results, impact and inform better policy making. Fragmented and diverse legislation concerning privacy (e.g. Data Protection Bill, re-use and sometimes fees (e.g. in Sweden and Germany¹⁸) can create confusion for end-users. FOI and PSI legislation as well as clear licensing guidelines are a cornerstone of OGD. In order to facilitate and coordinate 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. Spain¹⁹, France²⁰, and Denmark²¹).

- **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 practice varies across countries. Similarly, not all countries establish a right of access

to information in electronic format wherever possible, and many ‘access to information laws’ 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 which should include databases. However, there is often not an explicit reference to a right of access to databases, except for laws such as in Finland and Norway, which do 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 a problem predominantly with older access to information laws. In the majority of countries where there is not specific exclusion for databases, access to information and open government data advocates can use the wording of the national access to information law to argue that the right applies to databases.

- **Legal exceptions to openness.** There are a number of ways in which information held by public bodies may – rightfully - not be completely open to information seekers, from a legal perspective. The first is that 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 assesses that the information can be commercialised by being sold to for-profit companies, which produce added-value products. The information will therefore be released to members of the public or to private companies only upon payment of a fee. These exemptions are actually necessary to reassure users that the right data is protected, and the challenge emerges from ensuring that the right criteria are explained to third parties and applied consistently.
- Complexities of the various national legal frameworks for copyright and related rights as they apply to government information. One additional legal area which particularly lacks clarity and which has implications on public sector information and data “openness”, is the question of who owns government information. Many ‘accesses to information laws’ presume that public information is to be accessible and in that sense these laws consider the general public as the legitimate owner of public information. However, in many countries it is still the case that public bodies assert intellectual property rights such as copyright and database rights (intellectual property rights generated by 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 their economic model sometimes includes selling the information for profit.
- Compared to technologists in the private sector, national Web masters in the public sector face a daunting array of additional challenges and requirements which often are not technology-related in and of themselves, e.g. legal challenges, but still have an impact on technological matters. In the US, for instance, an online compliance checklist for designers of federal Web sites identifies about twenty-four different regulatory regimes with which all public federal Web sites must comply²². These range from privacy and usability to Freedom of Information Act (FOIA), compliance to the demands of the Paperwork Reduction Act and, separately, 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. But 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 into compliance. These problems are not unique to the US and are faced by public Web sites in many countries.

- **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 which explicitly address the issues of putting government information online open up 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 “Web site”? 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-facing 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 data for minorities to avoid creating new forms of digital divides and 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 US, 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 enforcement of the supporting legislation.

Box 6. 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 reuse, 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 which would be sold on 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.

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

NEXT STEPS IN THE ANALYSIS

82. There appear to be some “unresolved” issues around setting and implementing OGD policies. Further research, analysis and exchange could help governments clarify their position and approaches. The main are listed below.

Overarching conceptual issues

83. Bridging the distance between access to information and open government data movements. Advocates of access to information and those sustaining open government data have been pushing for their own agendas individually and in a way 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 about the economic and social benefits of releasing government data employed by open government data advocates.

- **Clearly defining terms and clarifying the relationship between data, information and knowledge and its implications for approaching OGD (initiatives) is critical.** The knowledge management literature provides a rich discussion of the epistemological issues related to data, information, knowledge and their inter-relationships (Alavi/Leidner 2011). The conventional view is that there is a hierarchical relationship between data, information and knowledge. Thus data would be observations of the world or facts which 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 is being 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 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).

Addressing the emerging needs for governments and public at large

- **Enhanced ability to combine different databases together can help develop more and better products and services.** Mixing public data with commercial, civil society and citizen input data, and pooling and sharing with other 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, civil groups as well as individual citizens²³. 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 these ‘big data’ to provide the necessary implicit and uncoded evidence.
- **The value for good and participative governance to experiment and 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 use of data and of the latest technology to unlock the potential insight of various data to achieve the delivery of modern and personalised services targeted to the needs of the users. Additionally, 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 which 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 is happening anyway, but between individual citizens and government. This means structuring, tracking, tracing and personalising 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 potentially provides wins for citizens, as well as for the government. It can move governments from one size fits all to segmentation and finally to personalisation. Current experiments, for instance in San Francisco and elsewhere are at an early stage but are likely to show the way in this important area.

- **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 cheap, software building blocks are often free and open source, and new sites can iterate their designs rapidly. Successes thus far (e.g. Govtrack.us site built by Joshua Tauberer²⁴), show that significant resources are not required to enter this space.
- **Adopting an Open Data Strategy.** The Strategy should set out 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 government department's commitment to proactively publish data and a timeline for implementing such plan. On this basis, the departments are normally required to report on progress made on their commitments (e.g. alongside the Open Data White Paper each government department in the UK has published in June 2012 its first Open Data Strategy).
- **Avoid creating new divides and inequalities and focus on user centricity.** Greater access to public data is an essential part of any government strategy for making the most effective use of data. Data that is trustworthy and easy to use (better search facilities, simpler ways to access information, advanced GIS data search) are key. Unequal distribution of skills, dedicated resources and time have a big impact on who is making use of OGD and the reach they can have with OGD use. For instance, there is a risk that a focus on data-for developers, and expectation that all data will be accessed through online interfaces, can lead to a neglect of mechanisms for direct access OGD access. The current OGD community, for instance, appears to be very technology focused in many countries (Davis 2010). Focusing excessively on an "idealized digital infrastructure" may cause the risk of losing the sight of the real end-user. For instance, human readable raw data-sets might be as valuable as machine-readable ones. It is impossible to predict which data provides 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 democratizes, there will be more diversity and demand would be spread out over more data products.
- **Enhanced access to data requires changes in the procurement and delivery of IT at the strategic level.** Government data is often locked into inflexible IT systems and retrieving the data is frequently a costly exercise requiring a detailed business case or contractual amendments. The analysis and multi-stakeholder dialogue on these matters need to be ongoing.

- **Adopting clear business models to boost the desired impact and obtain the desired results.** OGD users motivated by innovation, creating applications/websites without a clear business models have little incentive to move from proof-of-concept to production uses of data. Additionally, many unresolved issues pertaining to the financial implications of OGD (see section on economic challenges) further underline the relevance of having a business case to solve them and provide incentives within the various agencies. 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 results' sustainability in the long-run (bottom-up approach).

Harmonising definitions and sorting out legal “conflicts”

- **Striking the balance between transparency and privacy.** Sorting out how to balance between the benefits brought about by OGD in terms of public trust and confidentiality, both when the need is to protect the formulation of government policy or commercial interest or whether it is a matter of protecting 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 potential risks for privacy and fraud. This requires further thinking which may also lead to the decision of adopting ‘sector’ specific’ approaches to transparency given the variety of sector-specific issues. For instance, in the UK and the US the government has implemented pilot programmes for giving patients and students direct access to their own data which have proven quite popular. [Blue Button].
- **Building greater trust in public data and managing risks.** The success of creating an information marketplace is very much related with the public sector ability to safeguard people’s data from misuse and protect public’s rights to privacy, but also with the level of public trust. Opaque ownership of government data used to deliver services can for instance hinder their uptake. Ontologies and data architectures can support ethical consumption of open government data. 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 avoid manipulation which may endanger quality is essential to main 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.
- Unresolved conflict between the right of access to information, as an inherent part of the right to freedom of expression versus the limitation on re-use 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. According to many, this debate could take a point of departure in the principle that public bodies are mere custodians of the information created for, and on behalf of, members of the public. These are issues which need to be addressed in the long term, as part of the ongoing democratic challenge to promote fully open and accountable government.
- **A further set of issues surround the relationship between the PSI Directive²⁵ and access to information laws.** These are issues which are particularly relevant in Europe, but which 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 reuse. In France for instance a chapter on the reuse 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 reuse 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 databases²⁶.

84. A second, and potentially more serious problem, is that public sector information reuse rules can set up a dual charging regime. On the one hand, there is the principle that access to information under 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 which 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 around Europe a single piece of geographic data – an aerial photo of 10 km² of land – can range in price from free of charge to €292 with the average price being €62; of course for any serious reuser, 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 €356 million. Whilst geographic data seems 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 information²⁷.

Advancing the empirical analysis and measurements

- Better understanding and measuring the cost and impact of open data on wider social and economic terms, and with regards to political issues. 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 and economic value. Democratic value of OGD can be assessed intrinsically on the basis of some other values such as freedom or 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.

85. A good understanding of the relationship between supply and use of OGD can allow a deeper evaluation of OGD value-chain. Understanding the full civic value-chain of OGD 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 impact wider areas of public sector reform is yet to be assessed.

86. Finally, 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 analysed so far. This is an area worth 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.

ENABLING AN EMPIRICAL ANALYSIS

87. This section provides a suggestion for an Analytical Framework, and data collection as its complement, to gather key data on Open Government Data Initiatives (portals). 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 the data, will enable to map initiatives in OECD countries, identify typographies of initiatives and approaches and move towards a common methodology for assessing efficiency of implementation and achievements of impact.

88. Data collection includes:

- The questions to be sent to the countries based on the Analytical 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.
- 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)
- Empirical data we suggest to collect.

89. 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 open government data (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	<p>This section aims to provide an understanding of:</p> <ul style="list-style-type: none"> - Overall strategy setting the vision, objectives (expected results and overall outcomes) and the political priorities underlying Open Government Data 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 Open Government Data 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	<p>This section focuses on:</p> <ul style="list-style-type: none"> - 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	<p>Key objectives of this section are the understanding of:</p> <ul style="list-style-type: none"> - 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	<p>This section will focus on technical matters which sustain or limit real data openness, i.e. data accessibility, availability, reusability. It will also provide an overview of current situation concerning:</p> <ul style="list-style-type: none"> - Data quality - Data up-take - Interoperability (e.g. interoperability of data catalogues) - Workflow for data release and approval - Dataset storage - Data cataloguing and metadata

	<p>Economic and financial</p>	<p>To complete a cost and benefits analysis, the main focus of this section is on:</p> <ul style="list-style-type: none"> - 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.
	<p>Organisational</p>	<p>This section focuses on the measure taken to enable and foster the changes required in the public sector:</p> <ul style="list-style-type: none"> - 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.
	<p>Communication and interaction</p>	<p>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.</p>
<p>Impact</p>	<p>This section will analyse:</p> <ul style="list-style-type: none"> - 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 the 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 open government data, discuss any existing data on tendencies in requests for access to government data - Comparison of the available open government data 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) which 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	<ul style="list-style-type: none"> - 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 	<ul style="list-style-type: none"> - How does Open Government Data 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	<ul style="list-style-type: none"> - Availability <ul style="list-style-type: none"> 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	<ul style="list-style-type: none"> - How is data provided? <ul style="list-style-type: none"> a. direct data provision, where data belonging to various public agencies is published by the one-stop government data portal. <ul style="list-style-type: none"> 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

		<p>while the portal provides some kind of linking mechanism and/or metadata for the identification of the actual dataset.</p> <p>Subset of data_ how many ministries/central government agencies have OGD portals?</p> <ul style="list-style-type: none"> - 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? <p>Quantitative data</p> <ul style="list-style-type: none"> - 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 <p><i>Perception questions:</i></p> <ul style="list-style-type: none"> - <i>level of integration of OGD enabling processes and emerging services</i> 	
<p>Organisational</p>		<ul style="list-style-type: none"> - 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? <p><i>Perception questions</i></p> <ul style="list-style-type: none"> - <i>Cultural change in government agencies</i> - <i>Was a dialogue initiated among various stakeholders on the importance of sharing data and its benefits to the public?</i> 	<p>OECD Survey 2013</p>

<p>Communication and interaction</p>	<p>- interaction : existence of web 2.0 embedded platforms to gather feedback and comments</p>	<p>Quantitative data</p> <ul style="list-style-type: none"> - 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 <p><i>Perception questions:</i></p> <ul style="list-style-type: none"> - overall user experience - perceived usefulness of public engagement tools and applications - public satisfaction with interactions with government <p>Qualitative questions:</p> <ul style="list-style-type: none"> - 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 open government data 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? 	<p>Web crawl and log files, OECD 2013 survey</p>
<p>Political priorities</p>	<p>Transparency and accountability, participation, collaboration, public sector efficiency and effectiveness, citizens’ quality of life, service delivery, innovation</p>		<p>OECD Survey 2013</p>
<p>Impact **</p>	<p>- on innovation - economic value and productivity</p>	<p>Qualitative questions:</p> <ul style="list-style-type: none"> - What is the impact (social and economic) of the OGD initiative and how is this 	<p>OECD Survey 2013</p>

	<ul style="list-style-type: none"> - social value - political - user centricity 	<p>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?</p> <ul style="list-style-type: none"> - 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)	<ul style="list-style-type: none"> - data availability (supply) Datasets available Datasets downloadable Datasets_API_callable 	<ul style="list-style-type: none"> - 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
	<ul style="list-style-type: none"> - data uptake (demand) Views Downloads API_calls 	<ul style="list-style-type: none"> - 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

	<ul style="list-style-type: none"> - quality (data accuracy, consistency, update timeliness) and affordability 	<p>new datasets in the past 6 months</p> <ul style="list-style-type: none"> - 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). 	
	<ul style="list-style-type: none"> - reuse 	<ul style="list-style-type: none"> - 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 	<p>OECD Survey 2013 (STI?)</p>

REFERENCES

- Access Info Europe, the Open Knowledge Foundation (2011), “Beyond Access: Open Government Data and the Right to (Re)use Public Information”.
- Danish Government, Local Government Denmark (2012), “Good Basic Data for Everyone – A driver for Growth and Efficiency”, October.
- T. Davies (2010), “Open Data, Democracy and Public Sector Reform: A look at Open Government Data use at data.gov.uk”.
- European Commission (2006), “Measuring European Public Sector Information Resources”.
- E. W. Felten, D. Robinson, H. Yu, W. P. Zeller (2009), “Government Data and the Invisible Hand”.
- E. Kalambokis, E. Tambouris, K. Tarabanis (2012), “A classification scheme for open government data: towards linking decentralised data”.
- G. Lee, Y. Kwak (2011), “An Open Government Implementation Model: Moving to Increased Public Engagement”, IBM Center for the Business of Government.
- The National Cross Industry Working Group on Open Data (2012), "Open Data Ireland", A briefing paper, February.
- B. Noveck, “Open Data – The Democratic Imperative”, 5 July 2012, <http://crookedtimber.org/2012/07/05/open-data-the-democratic-imperative/>.
- 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 (2011), “Measuring the Economics of ‘Big Data’”, [DSTI/ICCP/IIS\(2011\)4](#).
- D. Robinson, H. Yu (2012), “The New Ambiguity of Open Government”, UCLA Law Review Discourse.
- R. Pollock (2012), “[Managing Expectations II: Open Data, Technology and Government 2.0 – What Should We, And Should We Not Expect](#)”.
- A. Schellong, E. Stepanets (2011), “Uncharted Waters: The State of Open Data in Europe”, Public Sector Studies, Business Solutions Technology Outsourcing.

T. Slee (2012), “ [Open Data Movement Redux: Tribes and Contradictions](#)”.

Supporting open data use through active engagement Position paper: W3C Using Open Data Workshop (June 2012).

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

J. Wonderlich (2012), “ [Open Data Creates Accountability](#)”.

C. Yiu (2011), “The Big Data Opportunity. Making Government faster, smarter and more personal”, June.

¹ Yiu.C, The Big Data Opportunity – Making Government faster, smarter and more personal. June 2011.

² http://resource.org/8_principles.html adopted in December 2007 (last visited on 24 July 2012)

³ <http://data.gov.uk/blog/new--public-sector-transparency-board-and-public-data-transparency-principles> (last visit on 20 July 2012).

⁴ http://ec.europa.eu/information_society/policy/psi/index_en.htm.

⁵ 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

⁶ Open Government Declaration, OPEN GOVERNMENT PARTNERSHIP (Sep 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.

⁷ Barack Obama and Joe Biden: Technology, <http://www.barackobama.com/issues/technology/> (last visited Dec. 2, 2008).

⁸ 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.

⁹ 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.

¹⁰ www.w3.org/DesignIssues/LinkedData.html. (last visited on 24 July 2012)

¹¹ Open Data White Paper – Unleashing the Potential, HM UK Government, June 2012

¹² Regulations.gov, What Is on This Site, http://www.regulations.gov/search/this_site.jsp (last visited Dec. 2, 2008).

¹³ See Office of management and budget, Executive Office of the President, Expanding E-Government: Partnering for a results oriented government 4 (2004), available at http://www.whitehouse.gov/omb/budintegration/expanding_egov12-2004.pdf.

¹⁴ Cynthia Farina et. al, *Achieving the Potential: The Future of Federal e-Rulemaking*, 2008 SEC. ADMIN. L. & REG. PRAC. AM. BAR ASS’N 1, available at <http://ceri.law.cornell.edu/erm-comm.php>.

¹⁵ OpenRegulations.org, About This Site, <http://www.openregulations.org/about/> (last visited Dec. 2, 2008).

¹⁶ 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; accessed 30 July 2012.

¹⁷In 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.

¹⁸ 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).

¹⁹ www.aporta.es/web/guest/guia_reutilizacion

²⁰ www.gfii.asso.fr/article.php?id_article=3278

²¹ Digitaliser.dk/resources/559456

²² Web Content Managers Advisory Council, Requirements Checklist for Government Web Managers, http://www.usa.gov/webcontent/reqs_bestpractices/reqs_checklist.shtml (last visited Dec. 2, 2008).

²³ Interview with Shannon Spanhake, Deputy Chief Innovation Officer, City of San Francisco, April 2012.

²⁴ About Govtrack.us, <http://www.govtrack.us/about.xpd> (last visited Dec. 2, 2008).

²⁵ Public Sector Information Directive available at: http://ec.europa.eu/information_society/policy/psi/actions_eu/policy_actions/index_en.htm

²⁶ See France’s Law of 17 July 1978 on Access to Administrative Documents, available at: <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068643&dateTexte=20100720>.

²⁷ 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: http://www.micus.de/pdf/MICUS-Studie_PSI_EU_March_2009.pdf.