CSC

BUSINESS SOLUTIONS
TECHNOLOGY
OUTSOURCING

UNCHARTERED WATERS

THE STATE OF OPEN DATA IN EUROPE

UNCHARTERED WATERS

The State of Open Data in Europe

Alexander Schellong Ekaterina Stepanets

Opening up government data to the public has been part of the European policy agenda since the introduction of the PSI directive in 2003. European Member States continue to lean towards a cautious approach of making their data available to citizens. This is partly caused by conflicting legal frameworks, cultural norms and the idea to recover the costs of data production. At the same time and inspired by activities in the U.S. and UK, the open data movement has emerged in many countries around the globe. They have a simple demand: Government agencies should put as much of their data online as possible in a machine-readable format so that everyone can re-use it since they were paid for by taxes. This study analyses the current state of the open data policy ecosystem and open government data offerings in nine European Member States. Since none of the countries studied currently offers a national open data portal, this study compares the statistics offices' online data offerings. The analysis shows that they fulfill a number of open data principles but that there is still a lot of room for improvement. This study underlines that the development of data catalogues and portals should not be seen as means to an end.

Introduction

Government data seem to resemble the open seas in ancient times. They remain a vast unexplored space. Only a few have a basic grasp of them. Proper navigational equipment and utilization is also missing. Some, like the Greek setting sail to find new trade routes to India in 510 B.C., expect great economic potential beyond the horizon; others are in fear of falling of the edge of the world. As in the past, it is the same mix of human curiosity and technological advancements that now leads people to push existing boundaries and long-standing believes about government data.

The open data movement, geek and creative types—explorers of a new kind—who want to gain access to and make sense of public data in new ways through technology, has become a visible force internationally.[1] It is a part of a general trend towards open and transparent government, also coined Government 2.0 or Open Government.[2] Starting with projects in the United States and United Kingdom[3], governments around the globe are now working on opening up their "ocean of data" in machine-readable format.[4] Over the years, public bodies have created and accumulated vast amounts of information—ranging from scientific, economic and geospatial data to reports—available in a wide variety of structures and formats. With the diffusion of

The authors Dr. Alexander Schellong (aschellong@csc.com) and Ekaterina Stepanets would like to thank the external and internal sources who contributed to this study.

[1] The Economist (2010)
[2] Metzl (2001); Osimo (2008); Noveck (2009); Nabatchi/Mergel (2010); Müller (2010a); Lathrop/Ruma (2010)
[3] Launch of: www.data.gov and www.data.gov.uk
[4] Davies/Lithwick (2010)

technology in every branch of government, the proliferation data continues at an ever increasing speed.[5]

Opening up these data to the public promises to create public value: ensuring transparency[6] and accountability, encouraging innovation and economic growth, educating and influencing people, or improving efficiency of the government.[7] These values make public sector information (PSI) a strategic resource, potentially important for different public sector agencies, private businesses, academia, citizens and civic organizations.[8] Accordingly, the European Commission (EC) adopted the PSI directive in late 2003.[9] European governments followed by passing freedom of information (FOI) and data re-use legislations. Easy access and supply of open data has, however, been limited to a few European Member States (MS) and areas of government. In addition, many government agencies still treat their data as a commodity that can be utilized to generate additional revenue through sometimes distorted pricing models.

Following the Obama administration's open government directive[10] (2009), various MS announced open data activities such as plans to develop national data catalogues by the end of 2010. Open data is also a policy objective of the Malmö Ministerial Declaration on eGovernment[11] which sets the priorities of the EU in this area for the period 2010-2015, and in the recently adopted European Digital Agenda[12] that outlines Europe's information society policies until 2015.

Meanwhile, non-governmental organizations and individuals have not only been engaged in making government data more available to the public but also in shaping the open data discourse[13]. For example, the Sunlight Foundation, an advocate for greater transparency in the U.S. government, has already set the quasi-standard any open data offering is currently measured against. It defined 10 principles[14] which describe the accessibility, format, licensing and cost of open government data. Yet the understanding of what "open government data" (OGD) are varies. Some proponents of open data prioritize citizen's access to and control over their data held by public agencies and commercial entities.[15] With Wikileaks, an international online publishing services for whistle-blowers, dominating the news in late 2010 other proponents of open data inside and outside of government need to delineate their scope and understanding of open data.

---

[5] Schellong (2008)
[6] A common argument is that knowledge gained through transparency (e.g. data, consecutive analysis of the data) encourages alternatives and helps put alternatives into action.
[7] Thomson (1999); Uhlir (2004); Allan (2009); Report PSI Meeting (2010); Davies (2010); Deloitte Research (2010), Poikola (2010)
[8] Sharif (2009)
[9] European Commission (2003)
[10] www.whitehouse.gov/omb/assets/memoranda_2010/m10-06.pdf
[11] European Commission (2009)
[12] European Commission (2010a); European Commission (2010b)
[13] OKF (2006), Sunlight Foundation (2010)

[14] Sunlight Foundation (2010): Generally, open data are government data that are freely available. The latter implies provision of data in a digital form—via the Internet—completely free of charge and without any signed contracts. The adherents of the open data movement also stress the importance of raw data, "with the highest possible level of granularity", use of machine-readable formats, as well as timely and full provision of data.
[15] Parycek/Sachs (2010)

Scope

This study seeks to assess the state of open government data in nine European Member States (Germany, France, Denmark, Norway, Sweden, Finland, Netherlands, Italy and Spain). Besides the policy level, this study reviews existing data portals from an end-user perspective and checks their compliance with a selection of open data principles.

Open Data

The term “open data” (OD) is relatively new. It derives from similar roots as “open source”[16] and “open access”[17]. It is applied in a variety of domains.[18] Open data is a philosophy, its principles being outlined at a later stage in this study, and now a decade-old debate on making resources such as scientific data[19] freely available and reusable. The concept applies both to data in raw and processed form, including data as varied as genetic sequences, geographic information, electromagnetic emissions, images, public transport schedules, data from medical experiments, voting results, reports and so on.[20]

In general, definitions of “open data” do not offer insight into what data are, but rather on the issue of openness and re-use. Public data are commonly defined as “data that are not subject to valid privacy, security or privilege limitations”.[21] Therefore, what comprises a government data-set is an open discussion and seems to resemble the Heisenberg uncertainty principle. The knowledge management literature provides a rich discussion of the epistemological issues related to data, information, knowledge and their inter-relationships.[22] The conventional view is that there is a hierarchical relationship between data, information and knowledge (Figure 1). 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?).[23] Yet others argue that there is no raw data,

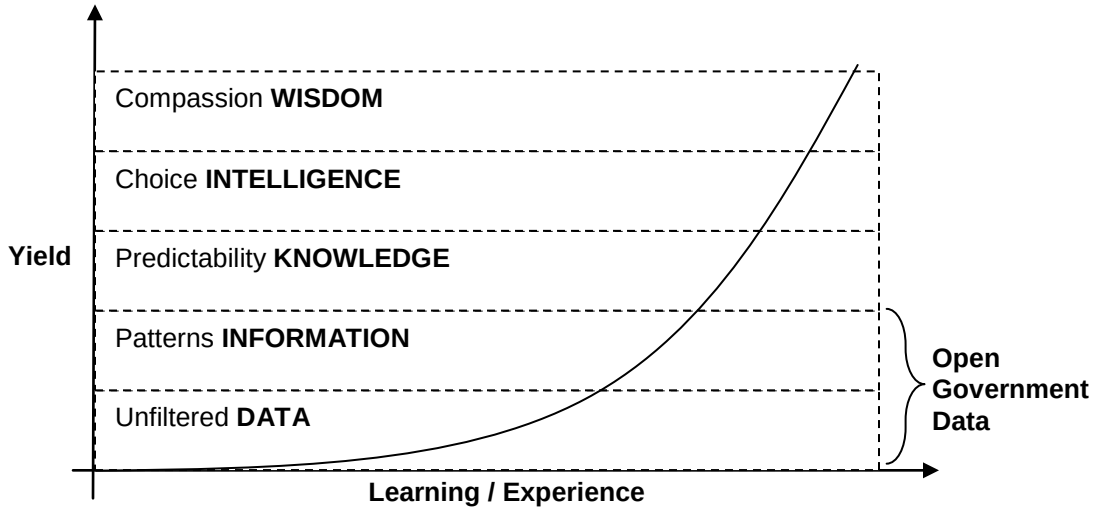

Figure 1: Data in the conventional view on the knowledge hierarchy (adapted from Tuomi (1999))

[16] Feller/Fitzgerald (2002)
[17] Soros (2001); Malakoff (2003)
[18] Murray-Rust (2008)
[19] Arzberger et. al. (2004)
[20] Miller/Doyle (2007); Internet & Gesellschaft Collaboratory (2010): 51
[21] Internet & Gesellschaft Collaboratory (2010): 51-52, Sunlight Foundation (2010); Graudenz et. al. (2010): 30-56.
[22] Quintas/Lefrere/Jones (1997); Alavi/Leidner (2001);
[23] Ackoff (1989); Tuomi (1999); Bellinger/Castro/Mills (2004)

since every measurable or collectable piece of fact has already been affected by knowledge.[24] As data and information artefacts (e.g. a report) may carry its own information representation, it's the combination of content and interpretation users find valuable. That's the idea of the curve in Figure 1. 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 be of philosophical nature but is necessary because OGD publications rarely provide guidance on the meaning of data. Terms are often used casually which leads to confusion in and outside of government, 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.[25]

Open data may also be categorized based on their format. File types such as XML, CSV, GML or RDF are strongly associated with open or raw data. Table 1 provides a list of potential data formats for OGD. Bennett (2009), however, recommends creating and publishing new documents in various standard formats is easier than fixing after the fact. Data may exist in digital and non-digital form (e.g. paper), the latter making re-use a challenge.[27] Government data are gathered or produced in various intervals by humans in and outside of government, or automatically through technology (e.g. sensors). Data may also be sold on the basis of different pricing models (e.g. flat vs. volume charge).

Text, Tables and Pictures	M	S	O	Geodata	M	S	O
Text (.txt)	+	+	+	Geography Markup Language (.gml)	+	+	+
Comma Separated Value (.csv)	+	+	+	GPS Exchange Format (.gpx)	+	+	+
Hypertext Markup Language (.html)	+	+	+	Keyhole Markup Language (.kml)	+	+	+
Extensible Markup Lanuage (.xml)	+	+	+	Drawing Interchange File Format (.dxf)	+	+	-
Resource Description Framework (.rdf)	+	+	+	Autodesk Drawing Format (.dwg)	+	-	-
Open Document Formats (.odt, .ods..)	+	+	+	ESRI Shapefile Format (.shp, .shx, .dbf)	+	+	-
Newsfeed / Webfeed Syndication (.rss)	+	+	+	Enhanced Compression Wavelet (.ecw)	+	-	-
Portable Document Format (.pdf)	-	+	+	MrSID Format (.sid)	+	-	-
Microsoft Word (.doc, .docx)	-	+	-	Normbas. Exchangeinterface (NAS)	+	+	-
Microsoft Excel (.xls, .xlsx)	+	+	-	Unified Database Interface (EDBS)	+	+	-
Microsoft Rich Text Format (.rtf)	+	+	-	BGRUND (Geospital agency BaWü)	+	+	-
Graphics Interchange Format (.gif)	-	+	-	WLDGE-Format (.wld)	+	+	-
JPEG Format (.jpg / .jpeg)	-	+	+				
Portable Networks Graphics (.png)	-	+	+				
Tagged Image File Format (.tiff, .tif)	-	+	-				
Geo-TIFF Format (.geotiff)	-	+	-				

M = machine-readable; S = specifications available; O = open format[26]

Table 1: Potential data formats for OGD (von Lucke/Geiger (2010): 9; Gray (2010): 10)

[24] Tuomi (1999)
[25] OECD (2005)
[26] Neither von Lucke/Geiger (2010) nor Gray (2010) provide clear guidance on the „open format" (OF). Any discussion of the meaning of OF requires further clarification. See Tauberer (2009) for some insights on OF
[27] Robinson et. al. (2009)

Finally, any discussion on open data also needs to consider the issue of "Linked Data" (LD), a critical element of the semantic web[28]. In summary, LD is about using the Web to create typed links between data—from different sources—it is linked to other external data.[29] There is, however, currently considerable ambiguity as to the exact nature of Linked Data.[30]

Open Government Data in Europe

The European public sector information platform (ePSIplatform)[31] keeps track of OGD developments in Europe. A number of country reports document the current state of affairs in the sphere of PSI re-use. Citizens and governments in MS continue to struggle with issues such as legislation, data production, data pricing, access to local and central data and ensuring cross-border interoperability. Furthermore, even though Norway, France, Sweden, the Netherlands and Italy announced plans to develop open data portals analogous to data.gov by the end of 2010, all failed in achieving that goal. A pan-European data portal is also being envisioned but discussions are ongoing.[32]

The legal framework is a critical aspect of facilitating PSI re-use[33] in MS. Sweden and Italy were the last MS to implement the PSI directive in 2010 five years after the deadline of July 1st 2005 set in the PSI directive. The reports indicate a confusing situation for end-users with regards to the existing translations of the PSI directive into national law. For example, in the case of Sweden and Germany this is caused by the fragmented and diverse legislation concerning privacy, re-use, and sometimes fees.[34] 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.[35] Controversial interpretations of the PSI legislations by the public bodies and, as a result, varying practices are also mentioned in the case of Norway[36] and the Netherlands[37]. Along these lines, none of Finland's PSI laws explicitly prevent opening up and re-using government data "but unfortunately […] does not support it either."[38] From the countries covered in this study, only Spain and Denmark have laws that, according to the authors of the PSI reports, encourage PSI re-use[39].

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.[40] In Germany,

[28] www.w3.org/standards/semanticweb/data
[29] Bizer/Heath/Berners-Lee (2009)
[30] Campbell/MacNeill (2010)
[31] www.epsiplus.net/
[32] PSI Platform (2010b)
[33] Fanning (2010)
[34] Fanning (2010); Sand (2010)
[35] Deutscher Bundestag (2010)
[36] Øvrebø (2010)
[37] Zijlstra (2010a)
[38] Poikola (2010)
[39] Aporta Project Team (2010); Siegumfeldt (2010)
[40] Fanning (2010)

five of its sixteen states still do not have a FOI legislation.[41] According to Fanning (2010: 15), Germany's FOI has overall been a weak driving force with respect to the reengineering of public sector records management practices. In contrast, the UK's the Freedom of Information Act 2000 (FOIA) "not only set up the general right of access to public sector information it also forced (in the sense that it was a statutory requirement) 100,000 public authorities of all sizes across the country to reengineer their own information holdings and to examine the conditions of their holding. […] In other words the FOIA provided a legal and regulatory framework requiring public authorities to reengineer their records."[42]

Government agencies produce data of different quality and take different positions on PSI re-use.[43] Many times their focus lies on the production of the information or documents rather than data which is just a process supporting the former.[44] Agencies also tend to treat their data as a commodity to generate revenue or recover the cost of data production.[45] While it is being widely noticed that data production, preparation, hosting, maintenance and preservation involves money and efforts[46], the cost recovery model is being challenged on a variety of grounds:[47]

- Economists argue that the benefits to the tax that accrue from corporate and individual taxes from secondary publishing and service activities stimulated by open access policies far exceed any revenues that might be generated through cost recover policies.[48]

- Cost recovery policies often mean that budgetary constraints prevent some government agencies from acquiring information that has already been created to be collected by another part of government, resulting in agencies either doing without or using inferior alternatives.

- No supplier can design all information products required to meet the emerging needs of all users in a modern information based economy. Private actors have repeatedly demonstrated that they are willing and able to build useful new tools and services on top of government data, even if they have to do a great deal of work to reverse engineer and recover the structured information that government bodies possess, but have not published.[49]

- Recognition that efforts to build transnational data sets are hampered by national agencies bent on preserving intellectual property to pursue local cost recover policies.

[41] Transparency International (2009)
[42] Fanning (2010)
[43] Øvrebø (2010); Zijlstra (2010a)
[44] Poikola (2010)
[45] European Commission (2000); Weiss (2004); Miller/Doyle (2007)
[46] Aichholzer/Burkert (2004); European Commission (2000): The study estimates that the investment value of PSI is EUR 9.5 bn vs. 40-60 bn in economic value.
[47] Weiss (2004): 138
[48] Consider that many eBusiness companies such as Google, Facebook, Amazon or eBay derive their competitive advantage and business model partly or fully on compiling masses of information on activities, preferences, relationships of individuals and aggregates of the former.
[49] Robinson et. al. (2009): 171

- A growing understanding of the wealth-creating possibilities that arise from a common information base.

That being said, authors notice that benefits of open data stretch far beyond its economic potential and that open data may also lead to adverse behavior of some actors.[50] In general, more research is needed to better understand this issue or measure the impact of the open data on wider social and political issues.[51]

Based on a review of the strategies to develop data catalogues in the U.S. and UK, Hogge (2010) finds that in both countries, a "three tiered drive was at play". Three groups of actors can be distinguished: civil society, mid- and top-level public servants. All actors must be engaged in order to ensure the success of an open data project.[52] Hogge's (2010) study also underlines that simply putting data online is not enough. The quality of data as well as measures to encourage re-use is equally important.

An assessment of how open data is used in practice (data.gov.uk) has been undertaken by Davis (2010):

- Search / Browse
- Extract
- Manipulate
- Visualize
- Contextualize
- Analyze (statistically)
- Clean
- Combine
- Filter
- Convert
- Integrate (into an existing product / service)

Davis' (2010) study also finds that the current OGD community in the UK is very technology focused. The author warns against focusing solely on an "idealized digital infrastructure" as it may lead to loosing the sight of the real end-user. Use-cases must be identified and taken into consideration for further OGD development. Human readable raw data-sets might be as valuable as machine-readable ones. Moreover, the availability of data may require greater government capacity to debate the meaning of data and find ways to use it in democratic debate.

Finally, a survey conducted by Forsa/SAS (2010) tried to shed greater light on the opinion of German citizens on the subject of open data and attitude towards opening up public data in general. 68% of the respondents supported the notion of making all government data that is not subject to privacy available to the public online. A more detailed question showed that 88% would be in favor of publishing government data such as the utilization of economic stimulus

[50] Milner (2000): 143-150; Aichholzer/Burkert (2004); Antelman (2004), Report PSI Group Meeting (2010); Davies (2010)
[51] Hogge (2010)
[52] Hogge (2010): 3

funds, respirable dust rates, or restaurant inspection reports. Another study by DIMAP[53] revealed that 45% of respondents thought that OGD would make the work of administration more transparent and effective, while 44% believed the opposite.

[53] Internet & Gesellschaft Collaboratory (2010)

Methodology

The research design (Figure 2) of this study entails two-levels of analysis, two methods and multi-lingual data collection for nine MS (Germany, France, the Netherlands, Finland, Denmark, Sweden, Norway, Italy, and Spain). Methods were entirely qualitative; and included web and document content analysis, and an email survey based on open-ended questions (see Appendix for details).

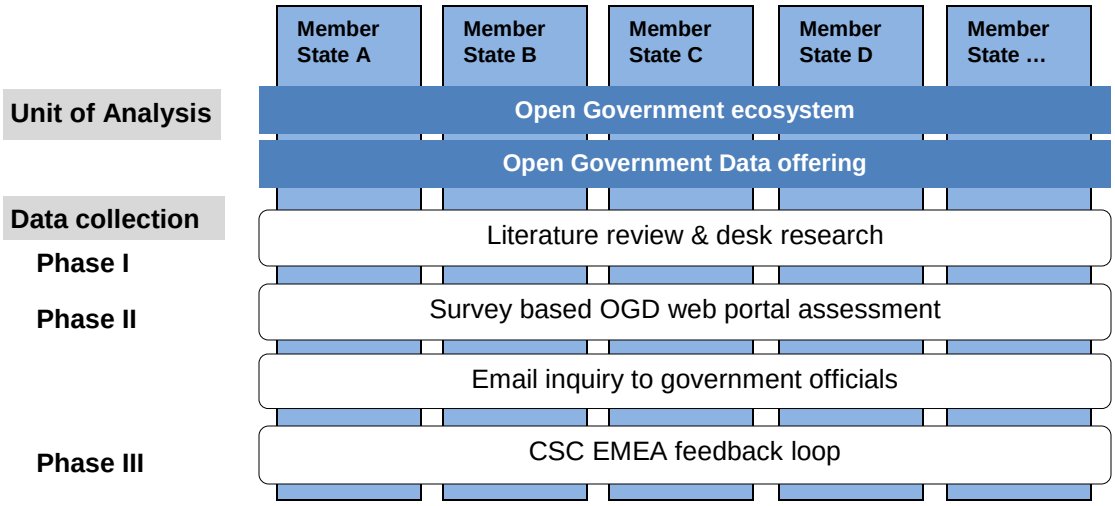

Figure 2: Research design

The unit of analysis is two-fold. The macro level analysis looked at the open data ecosystem: the policy domain related to open data and civil society initiatives in the nine MS. Assessments of the state of the legislative environment were not made but respective analysis and opinions were considered for the literature review of this study. The micro level analysis focused on data catalogues analogous to data.gov identified for each country.[54] These government web-portals needed to provide access to a broad and reasonably substantial set of government data not limited to data from one or two sectors such as only geographic data. The data on these portals must be accessible in electronic "raw data" formats such csv.[55] However, an initial review of MS activities showed the absence of such portals at the national level. Traditionally, statistics offices produce a lot of data and offer it to the public on their Websites. The statistics portals were thus chosen as a comparative basis for the assessment the existing government data offerings. They were analyzed along with the data.gov catalogues at the regional level.

A number of countries in the study, namely Denmark (Digitaliser.dk), Spain (Aporta.es) and Finland (data.suomi.fi) have developed web-portals that fall under the category "information portals" on the ePSIplatform. These web-resources are developed by the government and serve as a single access point to locate government data. They contain a listing of which government data are available and links to their source. As these portals do not provide direct access to raw government data in re-usable electronic format, they were excluded from this study. Since other

[54] The classification of data catalogues is influenced by: www.epsiplatform.eu/psi_data_catalogues
[55] www.epsiplatform.eu/psi_data_catalogues/category_1_public_sector_information_psi_data_catalogues_by_governments_direct_access_to_data

studies mentioned in the literature review (e.g. Davis 2010) provide greater details on the UK's open data portal than this study could have achieved in the given timeframe, it was also decided against including it from this study.

Web-portals were assessed on the basis of a questionnaire (Appendix). The questionnaire is divided into the three sections: general questions, questions that assess compliance with the selected open data principles[56] and questions that assess the usability of the portals from an end-user perspective. The last section contains a number of sub-sections that indicate different aspects of usability: layout of the web-site; statistics and metadata; ease of use for a new user; search engine; interactivity/ social media. Some questions refer to more than one section/ sub-section.

The principles of open data that guided this study are not mutually exclusive. The principles of open data were selected on the basis of feasibility of their evaluation within the scope of this study and included the following characteristics of open data: accessible, machine-readable, non-discriminatory, non-proprietary, license-free. Table 2 provides definitions of the principles, remarks on how to measure them and the reasons for abandoning some of them.[57]

The email inquiry to government officials contained six questions related to the open data ecosystem. Out of eight inquiries sent (the authors already possessed the information for one country), five answers were received.

Data collection took place September through November 2010. As a final step to ensure the quality, results were cross-checked with CSC public sector experts familiar with the various countries studied.

Principles		Remarks
1. Completeness	All public data are made available. Public data are data that are not subject to valid privacy, security or privilege limitations.	The review of the PSI directive has revealed that only a limited number of countries have developed PSI asset registries. It is thus impossible to measure this criterion in this study.
2. Primary	Data are collected at the source, with the finest possible level of granularity, not in aggregate or modified forms.	Public bodies have focused on the production of documents and information. The need to significantly increase the provision of primary data is identified in Europe.
3. Timeliness	Data are made available as quickly as possible to preserve the value of the data.	Many statistics portals have publishing calendars and claim to provide data in a timely fashion. Within the limits of this study, it is not feasible to check this criterion. It was, thus, abandoned.

[56] Sunlight Foundation (2010)

[57] Adopted from: www.opengovdata.org/home/8principles. We started with the now 10 open data principles released by the Sunlight Foundation (2010) but ended up with less than eight principles we could realistically utilize for the user oriented compliance check of the OGD Websites. We, therefore, decided to use the original 8 principles as a basis.

4. Accessible	Data are available to the widest range of users for the widest range of purposes.	This principle consists of two parts: availability of data on the Internet (so that all data is available online) and the ability of the existing tools to process it even if a dataset is large; accessibility of data for the disabled individuals (a compliance to the World Wide Web Consortium's Web Accessibility Initiative[58], for example). Use of multiple languages also increases the accessibility of data. Usage costs (fees) affect accessibility of data and may prevent the quantity of data usage.
5. Machine-readable	Data are reasonably structured to allow automated processing.	Provision of data in machine-readable formats, such as xml, cvs, rdf.
6. Non-discriminatory	Data are available to anyone, with no requirement of registration.	There should be anonymous access to the data. This principle is a reiteration of some of the goals of principle 4: accessibility.
7. Non-proprietary	Data are available in a format over which no entity has exclusive control.	Proprietary formats add unnecessary restrictions over who can use the data, how it can be used and shared, and whether the data will be usable in the future. While some proprietary formats are nearly ubiquitous, it is nevertheless not acceptable to use only proprietary formats. Likewise, the relevant non-proprietary formats may not reach a wide audience. In these cases, it may be necessary to make the data available in multiple formats.
8. License-free	Data are not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.	Government information is a mix of public records, personal information, copyrighted work and other non-open data. It is important to be clear about what data is available and what licensing, terms of service and legal restrictions apply. Data for which no restrictions apply should be marked clearly as being in the public domain.

Table 2: OGD principles

[58] www.w3.org/WAI/

Findings

Open Data Ecosystem

Throughout Europe there is an interest in the topic of open government data, both among politicians and citizens. Besides the European Commission[59], this study was able to identify 11 MS that have official open data initiatives: Germany, France, Spain, Italy, the Netherlands, the UK, Finland, Denmark, Norway, Greece and Estonia. It is important to notice that most of these initiatives, except maybe for the UK's data.gov.uk, are new, are beta-versions and represent the political and administrative will to move forward.

Open government data initiatives by civil society can be found in all countries of this study, except for Denmark (see Appendix for a detailed list of actors). 21 countries in Europe have such initiatives[60] which either provide access to government data (from any level of the government) or links to where government data can be accessed (in any format).[61] The shared-goal of these initiatives is demonstrating the benefits of open government data to government and the public.

Germany

The topic of open government and open data is part of the government's program "Networked and Transparent Government"[62] and has high-level political visibility through the support of Thomas de Mazière, Germany's minister of the interior.[63] The open government strategy must be implemented by 2013. The program pursues publishing PSI and open data as a progressive development, but does not abandon fees and calls for "a simple and unified" payment system for re-users. The Federal Ministry of the Interior is currently discussing the development of a data portal internally and with the public.[64] There are, however, no social media policies or guidelines at the federal level.

Germany has a vibrant open data community which includes actors such as the Open Data Network[65], Government 2.0 Netzwerk Deutschland[66], Politikdigital[67], ISPRAT[68], BITKOM[69] or D21[70]. The web-sites of the Open Data Network and Government 2.0 Netzwerk Deutschland serve as platforms to discuss and promote open data in Germany. Offenedaten.de is an initiative by Open Data Network supported by the Open Knowledge Foundation. It is a registry

[59] European Commission (2010a)
[60] Gray (2010)
[61] www.epsiplatform.eu/psi_data_catalogues/category_3_public_sector_information_psi_data_catalogues_by_civil_society_initiatives
[62] Bundesministerium des Innern (2010)
[63] www.e-konsultation.de/netzpolitik/mitreden
[64] Matzat (2010); Expert interviews
[65] opendata-network.org/,
[66] www.gov20.de/
[67] www.politik-digital.de
[68] www.isprat.net
[69] www.bitkom.org
[70] www.initiatived21.de

of open government and academic data in Germany. Another project, Offener Haushalt[71] ("open public budget"), provides a visualization of the federal budget.

France

APIE (Agence du Patrimoine Immatériel de l'Etat) is responsible for open data policy making and selling in France.[72] The prime minister has created a working group (Comité d'Orientation de l'Edition Publique et de l'Information Administrative, COEPIA) which is in charge of development of a strategy to make PSI more widely accessible with transparent and understandable conditions for re-use. During the last three years, APIE has been actively working on the subject of open data and creating standard licences, organized information sessions for administrations and encouraged them to make their data available on the portals.

Currently, ministries are working on their individual catalogues to make PSI more broadly accessible. Half of them have already edited their catalogues, including the Ministry of Justice and the Ministry of Economy. APIE is developing a national meta-portal.[73] The technical specifications have been determined and the portal is planned to be launched in the first quarter of 2011. This project is part of the government policy "France Numérique 2012"[74] which aims to develop the digital economy in France. No social media policy could be identified.

Regards Citoyens[75] and LiberTic[76] are civil society initiatives that promote open data in France. Data-publica[77], a French data catalogue, has been a private initiative which is partly publicly funding. It has the goal of creating a market that allows data-holders to publish metadata. The Webiste lists a number of datasets and will list tools for using the data.[78]

Denmark

The National IT and Telecom Agency (ITST) is the leading office for all open government data initiatives in Denmark. There is no open data policy, but the agency runs the 'Open Data Innovation Strategy' initiative which was launched in the spring of 2009. The approach used is "very agile and light, experimenting and collaborative", "primarily bottom-up as there is little or no political focus on PSI in Denmark."[79] It mostly involves working with data re-users and interested public authorities to demonstrate its potential and to stimulate demand.

A Danish government official explained that, "different public agencies publish selected data for a variety of individual reasons, often by initiative of enthusiastic civil servants or as part of their communications strategy, creating a very varied and fragmented overall picture. However, due

[71] bund.offenerhaushalt.de/
[72] Martinez (2010)
[73] www.apiefrance.com
[74] www.francenumerique2012.fr/
[75] www.regardscitoyens.org/
[76] www.liber-tic.com/
[77] www.data-publica.com/
[78] www.epsiplatform.eu/psi_data_catalogues/category_3_public_sector_information_psi_data_catalogues_by_civil_society_initiatives
[79] Siegumfeldt (2010)

to the EU INSPIRE Directive[80], the National Survey and Cadastre has put large amounts of data online."[81]

According to a Twitter message by Cathrine Lippert[82], a member of ITST, a new version of the Danish government data catalogue is currently being developed. The new version will be part of the digitaliser.dk collaborative platform which allows citizens and government bodies to interact around themes connected to digitalization. The digitaliser.dk platform also reports plans of a re-design of the Website to improve usability and to include the new data catalogue, which is now available in Danish only.

There is no central social media policy within the Danish government. The National IT & Telecom Agency uses social media as a natural part of its communication efforts. A number of individual agencies have developed their own guidelines or are in the process of doing so.

Sweden

Freedom of information legislation in Scandinavian countries, for instance Sweden, has been in place since the 18th century.[83] In March 2010 the Swedish eGovernment Delegation which was previously instructed to coordinate the government agencies' IT projects received an additional mandate on public information sharing and social media. While each agency is responsible for provision of data, the eDelegation is instructed to promote and coordinate the agencies' efforts to improve the conditions for the re-use of data. Many agencies in Sweden charge for access to raw data.

According to an eDelegation's survey, 44 per cent of Swedish agencies are using social media– primarily Facebook, Twitter, YouTube and blogs. Only five percent of the agencies, however, have a social media policy.[84] The eGovernment Delegation is currently working on general guidelines for agencies' use of social media. The guidelines were published in late December 2010.[85]

Frustrated by the state of existing Swedish government Websites, Peter Kranz, a member of the Swedish eDelegation team, started a private project called opengov.se as an initiative to highlight available public datasets in Sweden. Roughly 15% of datasets are available with an open license and in at least one open format (RDF/XML).[86]

Norway

The Ministry of Government Administration is currently developing a national data portal which will be reachable online at data.norge.no.[87] At the time of the data collection, the Website was

[80] European Parliament and Council (2007)
[81] Email reply by Danish government official
[82] twitter.com/cathrinelippert/status/29123895338
[83] Sand (2010): 3 - The first freedom of information act came into force in 1766 in Sweden
[84] Swedish Government eDelegation (2010): 6
[85] www.edelegationen.se/sites/default/files/imce/filer/publikationer/Riktlinjer_sociala_medier.pdf
[86] CTPR (2010): 21
[87] Øvrebø (2010): 3

mainly used a platform to discuss OGD questions. It now includes a small data catalogue in beta phase. Two reports on public data re-use have been published in 2010, raising awareness for the issue of OGD. The first was written by a group at the University of Bergen.[88] The second was produced by an expert group put together by the Norwegian Board of Technology, a consultative office to Parliament.[89] In 2010, the government launched a two month long idea competition called Nettskap 2.0, a Norwegian version of the Apps for Democracy. It encouraged OGD re-use. The competition resulted in 135 Apps. The winner won a price of 2,5 million Norwegian Krona (approx. EUR 320,000).

On November 17th, 2010, all national government agencies were instructed to publish data in machine readable format. This applies to data (1) that is valuable to the society, (2) that can be re-used, (3) that is not confidential or bound by other regulations prohibiting release, and (4) that is presumed to be inexpensive to publish.[90]. There is no central social media policy within the Norwegian government.

Finland

Finland does not have an open data or open government policy. According to a government official contacted for this study, guidelines and recommendations may be forthcoming in 2011 but that is still very much in the open. Moreover, a single social media policy does not exist and there are no plans to change this. The person also pointed out that Finland's public administration publishes a lot of data online. For example, Statistics Finland is about to release a subsection to their web service where they intend to publish all their data they have managed to turn into open data. In general, the topic report on Finland notices that progress in identifying PSI resources, opening new datasets and promoting re-use has been moderate and mostly related to the implementation of the INSPIRE directive. Other initiatives in the OGD domain include Apps for Democracy-type contests and ongoing data catalogues projects, such as the development of the national data catalogue data.suomi.fi. The latter is an attempt to collect open data sources available in Finland. The data listing is, however, very incomplete.

There is an open data portal in the works for the 14 cities and municipalities in the greater Helsinki metro region. The Helsinki Region Infoshare[91] project may become a national benchmark for central government open data solutions in the future. In March 2010 the Finnish Open Government Data Guidebook ("Julkinen data – johdatus tietovarantojen avaamiseen" / "Public data – an introduction to opening the information resources") was published. It was commissioned by the Finnish Ministry of Transportation and Communications and is part of the Ministry's work on developing a national information strategy and efforts to enhance the use of public sector information.

Opengov.fi is a civil society initiative that provides access to Finland's central government data.

[88] voxpublica.no/2010/01/open-government-data-in-norway-project-report-summary/
[89] Øvrebø (2010)
[90] www.epsiplatform.eu/news/news/norway_executive_decision_raw_data_now
[91] www.hri.fi/en/

The Netherlands

The directorate for service provision, simplification and information policy within the Dutch Ministry of the Interior and Kingdom Relations is responsible for open data policy. Currently, there is no open data policy but a social media policy in the Netherlands. The portal data.overheid.nl is in preparation but will not be online until political discussions on OGD have ended. In the meantime, a list of OGD sets can be found online.[92] The list is not complete as the Dutch government at this point does not know all of its data sources.[93]

The Ministry of Interior organized a competition ("Dat zou handig zijn!") while the Ministry of Education offered money to those with ideas around educational re-use of its data. Zijlstra (2010a) notices the central government stimulates bottom-up innovation. Close collaboration between individual civil servants and public sector bodies with civil society is typical for the Netherlands. There is an online network called "civil servants 2.0" (Ambtenaar 2.0), the Waag Society and initiatives that run the data catalogue overheidsfeeds.nl or organized a BarCamp on Open Government.

Italy

The Italian Minister of Public Administration promised to start an Italian data.gov-type portal by the end of 2010. The Website is not yet available for constituents. The public sector information project EVPSI[94] supported by the Piemonte regional government held a public launch on the 26th of March 2010 in Torino and will run until 2012. It aims at developing the regulatory and legal framework for OGD. The Italian Ministry for Public Administration and Innovation also started the MiaPA (my government) platform.[95] It combines social media, OGD and mobile capabilities. Citizens can download an app for their smart phone or use the service on the Internet. Citizens can find the nearest government office, comment on its services quality and review opinions of others. It is the first use of an open data license in Italy, granting free access to data, to extract, to reproduce and to re-use the data by any party for any purpose. The Italian open data license (version 0.1) is open for feedback and is compliant with Creative Commons 2.5 and the Data Commons.[96]

Social media policies or guidelines could not be identified for the Italian government. The government has a broader eGovernment Strategy (e-Gov 2012 Plan) which includes elements of open government and open data.[97] Spaghetti Open Data is an initiative by individual citizens to collect Italian OGD.[98]

Spain

The Aporta Project, initiated by the Spanish Ministry of Industry, Tourism and Commerce and the Ministry of the Presidency, is supposed to raise transparency in Spanish government and

[92] www.dialogiconderzoek.nl/opendata
[93] Expert interviews
[94] Extracting Value from Public Sector Information: Legal Framework and Regional Policies
[95] 95.110.228.151:8080/elgg/mod/westorDownloadCounter/download.php?file_guid=747
[96] www.epsiplatform.eu/news/news/italian_open_data_license
[97] OECD (2010b)
[98] www.spaghettiopendata.org/

promote government data re-use. It is a repository and search engine similar to the United States' data.gov portal. Aporta has a budget of 1 million, allocated to the construction and maintenance of the online portal. Government organizations can upload data and share it with citizens, business and public sector organizations. The Aporta Project has been launched within the framework of the national Information Society strategy Plan Avanza II[99] (2011-2015). Plan Avanza II includes open government as one of its strategic objectives, however, progress in other area is still in an early stage.[100] At the time of the assessment, the Spanish government had launched a public consultation on its law 37/2007 to allow re-use of government data. Social media policies or guidelines could not be identified for the Spanish government.

On November 10th 2010, the Catalan regional government also launched an open government data portal. Datos Públicos[101] is a public wiki to collect open government date sources in Spain.

[99] www.planavanza.es
[100] OECD (2010a)
[101] datospublicos.jottit.com

Assessment of the Web data-portals

None of the MS analyzed in this study has a national data portal analogous to data.gov or data.gov.uk an. The Asturias[102], Basque[103] and Catalonia[104] region in Spain, the Rennes metropolitan region[105] in France and the Piemonte region[106] in Italy have, however, developed data catalogues. Consequently, this section includes the assessment of national-level statistics offices' portals.

Compliance with the selected open government data principles

In summary, data on the statistics offices' portals in the survey are non-proprietary, are non-discriminatory, are provided at no cost and in machine-readable formats. However, the accessibility principle is not met as most of the portals lack an inclusive design for the disabled. Data are generally subject to copyright protection and are not provided license free. Table 3 and 4 provide an overview of the results which are followed by detailed findings for each MS.

In the majority of countries the information provided on the statistics portals falls under the copyright protection. Reproduction, distribution and re-use of the information are allowed provided that the integrity of the information is maintained and the source is mentioned.

All the portals studied provide data/information free of charge. Exceptions are personalized data-sets, large-volume and micro data which may be subject for additional fees.

All MS offer data in different formats. All statistics portals offer data in xls and machine-readable formats, such as csv. The use of the .pdf format is rare and limited to two countries.

All web-portals have an English language interface and provide at least some data in English. In particular, Scandinavian countries have made an effort to provide all data in English, such as Norway and Denmark, or explicitly show what data are available in English and which not, as in the case of Sweden.

Only three statistics portals in the study (Denmark, Sweden, Spain) have undertaken measures to facilitate access to data for the users with disabilities.

Web-portals provide data without registration. The registration is only required to use additional features and services, especially when users want to buy micro data. All statistics offices also provide additional services like tailored data, publications, special statistics software and so on.

[102] www.asturias.es/
[103] opendata.euskadi.net
[104] dadesobertes.gencat.cat/
[105] www.data.rennes-metropole.fr/
[106] www.dati.piemonte.it/

Country	Selected OGD criteria A	B	C	D	E	F	G	Data formats
Germany	+	+	-	+	-	2	2	xls, csv, html
France	-	NA	-	+	-	2	2	xls, csv, dbf, pdf, txt
Denmark	+	-	-	+	+	2	2	xls, csv, dbf, txt, tsd, asb, px, sas
Sweden	+	-	-	+	+	2	2	xls, csv, dbf, pdf, txt, scb, px, doc
Norway	-	NA	-	+	-	2	2	xls, csv, html, pdf, sdv
Finland	+	+	+	+	-	3	3	xls, csv, html, dbf, txt, xml, px, wk1
Netherlands	-	NA	-	+	-	2	2	xls, csv, html, spss
Italy	-	NA	-	+	-	2	2	xls, csv
Spain	-	NA	-	+	+	2	2	xls, csv, px

+ = compliant / - = non-compliant / NA = not available
A. Is a registration possible?
B. Is there a fee for registration?
C. Is there a fee for obtaining data?
D. Is there a data license?
E. Is there a functionality for disabled people?
F. Number of languages available on the interface
G. Number of languages of the data

Table 3: Compliance of MS statistics portals with selected OGD factors

Country / Region	Selected OGD criteria A	B	C	D	E	F	G	File formats
Italy/ Piemonte region	-	NA	-	+	-	1	1	csv, xml
Spain/ Basque region	-	NA	-	+	+	2	2	csv, xml, wms

+ = compliant / - = non-compliant / NA = not available
A. Is a registration possible?
B. Is there a fee for registration?
C. Is there a fee for obtaining data?
D. Is there a data license?
E. Is there a functionality for disabled people?
F. Number of languages available on the interface
G. Number of languages of the data

Table 4: Compliance of regional data.gov style portals with selected OGD factors

Germany

The Federal Statistics Office provides data online through its database GENESIS[107] which can generally be used free of charge. The data fall under copyright protection. Reproduction and distribution, also of parts, are permitted provided that the source is mentioned. Registration is only necessary for those who want to access additional personalized services such as permanent storage of table structures for retrieval, large volumes of data or web services. Those services are available on a chargeable basis only. The regular annual flat rate is EUR 50. There is also a premium flat rate of EUR 500 which is primarily aimed at businesses and institutions that process large volumes of the statistics office's data in their own systems.[108] Data are generally divided into nine categories. There is a tag cloud and a top five list of requested statistics. The website includes English and German parts. Some of the data are also available in English. The data are provided in different file formats.

France

In France, official statistics are produced by many agencies. The National Institute for Statistics and Economic Studies (INSEE) compiles statistics on population and businesses. Statistical departments of ministries and other public bodies broaden this information with specialized statistics, published on paper or online according to their area of expertise (e.g. agriculture, education, industry, tourism).

The web portal of official statistics[109] has been established upon the request of the French National Council for Statistical Information (CNIS) to enable a unified access to the statistics produced by different authorities. The French official statistics portal has no right of intellectual property over the data which can be reached through its site. The conditions of re-use are mentioned on the web-sites of different authorities.

INSEE, for example, has a number of databases.[110] The data made available on this site may be viewed and downloaded free of charge. Unless otherwise stated, they may be re-used, including for commercial purposes, without a license and without payment of royalties. Re-use is however conditional upon respect of the integrity of the information and data and a specific mention of source. There is no registration. The languages of the interface are French and English. Some data are also available in English. Data are provided in multiple file formats.

[107] www-genesis.destatis.de/genesis/online
[108] genesis.destatis.de/genesis/online;jsessionid=FEF3A5C8B26D051CA61612644691C29B.tomcat_GO_1_2?Menu=Registrierung
[109] www.statistique-publique.fr/
[110] www.insee.fr/en/bases-de-donnees/

Denmark

All data produced by Statistics Denmark[111] are available online free of charge. Reproduction and distribution, also of parts, are permitted provided that the source is mentioned. A registration is free of charge and is only required for large data-set inquiries, e.g. extracting tables with up to 100,000 cases. Similar to the offering of the German statistics office, a user profile also allows to save tables for later use. Tables containing up to 1,000 cases can be downloaded without prior registration. The languages of the interface are Danish and English. Data are available in both Danish and English; and different file formats. The web-site has a text-enlarge option which facilitates the use of the resource by disabled individuals.

Sweden

Statistics Sweden[112] is a government agency that produces statistics. Its data are available online and free of charge. All statistics on the website may be copied freely, duplicated and further distributed with the requirement to state the source of data. Registration is free of charge and is only required for large data-sets: to retrieve a table with more than 10,000 cases. The data and interface are in Swedish and English with an explicit separation of which data are available in each language. The data are provided in different file formats. The portal contains the “BrowseAloud” functionality for individuals with disabilities.

Norway

Statistics Norway[113] provides all of its data online and free of charge. The content and layout of the Website are subject to copyright. The materials of the web-site may be stored electronically, printed, copied or forwarded to other parties provided that reference is made to the source. All data are available in Norwegian and English and in different file formats.

Finland

Statistics Finland[114] provides vast amount of data free of charge. Chargeable databases contain longer time series, analytical indicators and more detailed information of certain topics. The price of the time series database Astika for one user per month is 404 euro + VAT 23%, for one user per year is EUR 2,020 + VAT 23%, for one organization (3 or more users) per year is EUR 6,055 euro + VAT 23%. The data can be used provided that the source is mentioned. The data and interface are provided in three languages, Finnish, Swedish and English. It is not clear what which data are available in English. The registration is required for the chargeable services. The data are provided in different file formats.

[111] www.statbank.dk
[112] www.ssd.scb.se
[113] ssb.no
[114] stat.fi/

The Netherlands

Statistics Netherlands[115] provides its data free of charge. Reproduction is permitted provided that Statistics Netherlands is quoted as the source. The interface is both in English and in Dutch. Some data are available in English. There is no registration functionality. The website specifically states that, if certain types of data are not available, the agency will help with finding it within certain timeframes. If the information is readily available, the user will receive data within two working days, free of charge. If the information is not readily available, the agency will contact the user as soon as possible. Data are generally provided in different file formats. The portal has sections for different target groups: media, policymakers, the public, researchers, survey respondents and a list of tables in the database. There is also an iPhone-App to access the database.

Italy

Statistics Italy[116] provides data online free of charge. Only CD-ROMs and personalized data analysis are subject to a fee in order to cover production costs. The data fall under copyright protection. There is no registration functionality. The languages of the interface are Italian and English. Some of the data are available in English. Data are provided in two formats. The Website has a mobile version.

The data portal of the Piemonte region provides access to raw data from the regional government free of charge including for commercial purposes. It also contains a blog with updates on European (PSI) and international open government data developments. A "testimonials" page allows users to share their experience with OGD re-use. The language of the Website is Italian.

Spain

The National Institute of Statistics (INE)[117] provides visitors to its Website with data from various sources. The INE makes this material available to users on an individual basis as a single end user license. Any commercialization of the data is therefore forbidden. There is no registration functionality. The languages of the interface are Spanish and English. Some of the data are available in English. Data are provided in different file formats. The web-site has a text-enlarge option which facilitates the use of the resource by disabled individuals.

The Basques government was the first Spanish regional administration to commit to opening up the data.[118] Open Data Euskadi (opendata.euskadi.net) is a data catalogue that provides data in machine-readable re-usable formats under open property licenses allowing re-distribution and re-use for commercial purposes.[119] It also serves as an information source on OGD. The design of the web-site follows WAI guidelines (Web Accessibility Initiative) Consortium (W3C World Wide Web Consortium) in order to facilitate the access for users with disabilities.

[115] www.cbs.nl
[116] en.istat.it/
[117] www. ine.es
[118] www.epsiplatform.eu/news/news/basque_government_puts_opendata_and_ogov_into_practice
[119] creativecommons.org/licenses/by/3.0/es/

User-centricity

In general, portals of national statistics offices are user friendly (Table 5/6). Almost all portals have “FAQ” (frequently asked questions) or “Help” sections and contain some instructions on how to use the portal. All portals provide an opportunity to send feedback on the website and receive assistance via different means (phone, email, mail, e-form). However, more than a half do not have a feature to enable the provision of feedback on data and only a few explicitly encourage it. Most have metadata, but its structure varies.

None of the portals provides statistics about its users (although statistics authorities collect these data for the internal use) or most popular sections. Only Germany’s statistics office features a “top five” list of frequently requested statistics.

	User centricity factors								
Country[120]	**FAQ**	**Help**	**Metadata**	**UFW**	**UFD**	**SocM**	**RSS**	**Forum**	**Apps**
Germany	+	+	+	+	-	-	+	-	-
France	+	+	+	+	-	-	+	-	-
Denmark	-	+	+	+	-	-	+	-	-
Sweden	+	-	+	+	-	+	+	-	-
Norway	+	+	-	+	-	-	+	-	-
Finland	-	-	+	+	+	-	+	-	-
Netherlands	+	+	-	+	+	+	+	-	-
Italy	+	-	+	+	+	-	+	-	-
Spain	+	+	+	+	+	-	+	-	-

+ = identified / - = not identified
UFW = User feedback on Website
UFD = User feedback on data
SocM = Social Media integration (e.g. Twitter, Facebook)
Apps = Users can create and posts apps

Table 5: User-centricity of national statistics offices’ portals

All portals provide a number of search options: There’s always a simple text search and search by theme, topic, and so on. Only a few have a browsing mechanism, such as a list of datasets/tables that are available or tag cloud (this is important because it allows the users to understand, what information is available on the web-site).

Most of the statistics portals are not of interactive nature and try to facilitate user to agency or user to user interaction such as feedback on data quality, usage or suggestions for additional datasets. Only a few portals (Sweden, the Netherlands and Italy’s Piemonte region) are linked to social media platforms. One has a mobile version (Italy) and one an iPhone App (the Netherlands). Almost all include RSS subscription and newsletters functionality.

[120] see Appendix or the following pages for the details on the URL‘s

	User centricity factors								
Country	**FAQ**	**Help**	**Metadata**	**UFW**	**UFD**	**SocM**	**RSS**	**Forum**	**Apps**
Italy (PMR)	+	-	-	+	+	+	-	-	-
Spain (BR)	-	+	+	+	+	-	+	-	-

Table 6: User centricity on regional data.gov style portals

Conclusion

As identified in the literature review, FOI and PSI legislation as well as clear licensing guidelines are a cornerstone of OGD. Even though this research did not look further into the issue, the various reports on the EU PSI underline that these areas continue to deserve great attention in all MS with regard to putting open data (re-use) ideas into practice. FOI laws and those pieces of legislation containing exemptions from the access principle are an indicator of a societies appreciation and interest in balancing transparency and secrecy.[121] This attention is not restricted to information in the public sector. Transparency obligations generally turn into compensation mechanisms for other asymmetries in the access to resources of power. Hence, it is only a question of time when there are greater demands for disclosure in the private sector.

Social media and open data policies are only components of a coherent open government strategy. Open data initiatives have a social component and offer the opportunity to engage with constituents. Web 2.0 technologies provide the tools for that but are equally important as thinking about processes.[122] In reality, these two policies are not interconnected, if present at all.

Metadata, the data about data, are very important as well. Currently most governments do not have a comprehensive overview of the data in their possession. This was highlighted both in the literature and the answers that we received from government authorities throughout Europe. The creation of data assets lists stipulated by the EU PSI Directive is highly desirable. Information engineering programs, like the one that took place in the UK and forced more 100,000 authorities to reengineer their records, seem to be essential for the success of open data initiatives. The implementation of the EU INSPIRE directive might add additional momentum for OGD. In any case, a conception of which data are available is an essential step for any government's OGD strategy.

The value proposition for OGD sometimes misses that public bodies are consumers of data themselves. Interoperability is a huge and many times unresolved issue in eGovernment. Dealing with OGD in general and open data file formats in particular facilitate IT system interoperability in government. It should be noted that the interoperability of data catalogues and creation of a pan-European data catalogue are a big challenge already looming on the horizon.

[121] Mayer-Schönberger/Lazer (2007)
[122] DiMaio (2010)

The open data community is uncoordinated and not unified. It includes IT professionals, government employees, associations and individual citizens active on the national or international level. Their motivations, level of understanding of government processes and structures, scope and priorities of 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.[123] This makes a comprehensive assessment of the open data movement a challenge. Interestingly, given the relative small size of the open-data activists—Julian Assange, Wikileaks' spokesperson, is not considered part of this spectrum—they are having quite an impact on government representatives. They are being heard and taken seriously on the national and European level. The principles and demands of the open data community still leave a lot of room for interpretation that can be filled through an open dialogue between government and the citizenry. For example, one may ask whether all of the principles lead to the expected value in practice. Such an open question might be directly countered by pointing to the estimated 4,000% return on investment (ROI) in a 30-day time period of the Apps for Democracy contents for the government of Washington, D.C (47 iPhone, Facebook and Web Apps).[124] These numbers need to be treated with caution due to a lack of research. Moreover, the numbers neither cover the issue of sustainability of the Apps nor do provide insights on the value on the consumer side.[125]

While we believe in the value of OGD, we recognize that open is not always free, especially in times of government austerity. Opening up government data is associated with costs for its discovery, preparation, provision and maintenance. The creation of an open data catalogue and portal is just the beginning. Ensuring the quality of data, their timeless, permanence (elimination of broken links and vanishing content) and potentially certifying supreme products developed by the community will require as much work as the initial step to make a decision to create a catalogue on the policy-level. Citizens could be informed about the limits to government resources available for OGD and consulted on where to set priorities.

Along these lines, a distinction between the technical characteristics of open data, such as requirement of the proper formats, and their transformational potential should be drawn.[126] The study has clearly shown that except for a national OGD portal—which we find highly desirable even in federal systems—and the license principle, the data provided by the statistics offices meets most of of the open data principles: data are provided for free, in machine-readable formats, they are non-proprietary and non-discriminatory. Random checks of some agency websites revealed that there is already OGD out there. Finding it is a different issue. What is also missing are participatory and collaborative elements, the completeness of data and raw data. The latter represents much more of a challenge, since these are not merely technical issues as statistics often contain person-related data and therefore cannot be made open because of the privacy issues. They imply a fundamental cultural change in the functioning of

[123] Matzat (2010)
[124] Nichols (2010) - Washington D.C. has decided to discontinue the contest
[125] Müller (2010b)
[126] Stephenson (2009); Schuppan (2010)

the public authorities: from disclosure to proactive disclosure; from provision of information to provision of data—the loss of interpretational sovereignty.[127]

Government officials are concerned about the quality of the data and unintended consequences of its use. A study conducted by The Sunlight Foundation might prove them right. A critical review of the open government efforts of the US administration showed how flawed its published data was.[128] The study identified over 1.3 trillion U.S. dollars of broken reporting in 2009 alone on the USASpending.gov, the Website launched to track federal government spending. Any media reaction to the above might further weaken the support of public servants and politicians for OGD. A government official also told us that there is simply the culture of "these are my data" in government which are at least in some parts a prerogative to be in existence.[129] He also pointed out that there is a struggle between the technical units responsible for the production of data and those responsible for publishing the data.[130]

The media are equally responsible for the success of open data. If they demand OGD and a flexible approach to implementation but then relentlessly punish any failure of a government's OGD project, OGD will lose its momentum.

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 chose. These guidelines cover technical and legal issues, economics as well as communication strategies. Several MS are already working on such guidelines (Norway) or have recently published them (Spain[131], France[132], Denmark[133]). Benchmarking activities, as a form of soft governance, could keep the momentum alive and flank other measures taken by governments. Future government IT system projects or upgrade activities might include vendor requirements to develop OGD interfaces in anticipation of a central OGD portal or export functionality. It is, however, important to understand for any government official that increasingly civil society has the means and sometimes motivation to bring government data online if government is not moving fast enough. Technology is not the obstacle anymore. Building a simple solution and letting it evolve[134] is a strategy that deserves to be tested among European governments.

It is impossible to predict which data provides the greatest value to society. 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

[127] Langkabel (2010)
[128] www.epsiplatform.eu/news/news/open_government_score_card
[129] This opinion is supported by Hernández-Ros (2010), head of unit of PSI within EU DG INFSO
[130] Poikola (2010) - In many cases, the decision of publishing a dataset openly falls in between the traditional roles and responsibilities of officials within public sector organizations. Those who technically understand opening data sources (the IT department) are typically not authorized to make decisions about the content, those who are usually responsible for publishing information (the communications department) may not understand at the firsthand how publishing data in legally and technically re-usable form differs from publishing documents on the organization's online portal."
[131] www.aporta.es/web/guest/guia_reutilizacion
[132] www.gfii.asso.fr/article.php3?id_article=3278
[133] digitaliser.dk/resource/559456
[134] O'Reilly (2010): 18-19

be more diversity and demand should be spread out over more data products. This is the “long-tail” of OG, a statistical distribution that describes how data low in demand collectively make up a volume that rivals or exceeds that of the few but most popular data, if the supply and distribution channel is large enough. The long tail underlines that governments can transfer the responsibility of determining the value of its data to users.

In short, open government data remain unchartered waters that are yet to be fully explored by European Member States and their citizens. As the development of data catalogues and portals is underway it is important to ensure that development of a government data portal does not turn into the goal in itself; a blind and unreflected following of the OGD trend with the desire to look progressive on the national level while for example missing the European developments.

References

Ackoff, R. L. (1989) “From data to wisdom“, Journal of Applied Systems Analysis, 16, 3-9.

Aichholzer, G.; Burkert, H. (Ed.) (2004) “Public Sector Information in the Digital Age: Between Markets, Public Management and Citizens’ Rights”, Edward Elgar Publishing.

Alavi, M.; Leidner, D. E. (2001) „Review: Knowledge Management and Knowledge Management Systems: Conceptual foundations and research issues“, MIS Quarterly Review, 25, 1, 107-136.

Allan, R. (2009) “The Power of Government Information”. In Gøtze J., Pedersen C. B., eds. “State of the eUnion: Government 2.0 and Onwards”, AuthorHouse.

Anderson, C. (2004) “The Long Tail”, Wired, 12.10. URL: www.wired.com/wired/archive/12.10/tail.html

Antelman, K. (2004) „Do Open-Access Articles have a greater research impact?“, College & Research Libraries News, 65, 5, 372-382

Aporta Project Team (2010) “State of Play: PSI Re-use in Spain - Aporta Project”, European Public Sector Information Platform, Topic Report, 14.

Arzberger, P.; Schroeder, P.; Beaulieu, A.; Bowker, G.; Casey, K.; Laaksonen, L.; Morrman; D.; Uhlir, P.; Wouters, P. (2004) „Promoting Access to Public Research Data for Scientific, Economic and Social Development“, Data Science Journal, 3, 135-152.

Bellinger, G.; Castro, D.; Mills, A. (2004) “Data, Information, Knowledge, and Wisdom”, URL: http://www.systems-thinking.org/dikw/dikw.htm

Bennett, D., Harvey A. (2009) “Publishing Open Government Data”, W3C Working Draft. URL: http://www.w3.org/TR/gov-data/

Bizer, C.; Heath, T.; Berners-Lee, T. (2009) “Linked data – the story so far”, International Journal on Semantic Web and Information Systems, 5, 3, 1-22.

Bundesministerium des Innern (2010) “Vernetzte und transparente Verwaltung”, Berlin.

Campbell, L. M.; MacNeill, S. (2010) “The Semantic Web, Linked and Open Data”, briefing paper, JISC CETIS.

CTPR – Center for Technology Policy Research (2010) “Open Government – some next steps for the UK”, May, London.

Davies, R. (2010) “PSI Portals: Overview of Progress (Part 1)”, European Public Sector Information Platform Topic, Report, 8.

Davies, T. (2010) “Open Data, Democracy and Public Sector Reform. A Look at Open Government Data Use from data.gov.uk”, based on MSc Dissertation, URL: http://practicalparticipation.co.uk/odi/report/

Davies, A., Lithwick, D. (2010) “Government 2.0 and Access to Information”, 2010-14-E, Library of Parliament, Canada.

Deloitte Research (2010) „Unlocking government - How data transforms democracy“.

Deutscher Bundestag (2010) “Tätigkeitsberichts zur Informationsfreiheit 2008 und 2009”, 17/1350, Berlin.

DiMaio, A. (2010) “Why Do Governments Separate Open Data and Social Media Strategies?” URL: http://blogs.gartner.com/andrea_dimaio/2010/02/19/why-do-governments-separate-open-data-and-social-media-strategies/

European Commission (2000) „Commercial exploitation of Europe's Public Sector Information”.

_____ (2003) “Communication on the Re-use of Public Sector Information – Review of Directive 2003/98/EC”.

_____ (2006) “The Measuring European Public Sector Information Resources Study (MEPSIR)”.

_____ (2009a). “Creating impact for an eUnion 2015 – "The Visby Declaration".

_____ (2009b). “Malmö Ministerial Declaration on E-Government”.

_____ (2010a). “EU Digital Agenda”.

_____ (2010b). “EU E-Government Action Plan 2011-15”.

European Parliament and Council (2003). “Directive 2003/98/EC on the Re-Use of Public Sector Information”.

_____ (2007). “Directive 2007/2/EC on establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)“.

EPSI Platform (2010a) “Realizing the Value of Public Sector Information”, Madrid, Spain, 09/06/2010.

_____ (2010b) “Technical workshop on the goals and requirements for a pan-European data portal”, report, Brussels.

Fanning, M. (2010) “Recognizing the Road to data.gov.de. An assessment of the European and National Regulatory Framework Impacting PSI re-use in Germany”, European Public Sector Information Platform, Topic Report, 11.

Feller, J; Fitzgerald, B. (2002) “Understanding open source software development”, Addison-Wesley, Boston, MA.

Graudenz, D.; Krug, B.; Hoffmann, C.; Schulz, S. E.; Warnecke, T.; Klessman, J. (2010) „Vom Open Government zur Digitalen Agora“, ISPRAT Whitepaper, ISPRAT e.V., Hamburg.

Gray, J. (2010) “Open Government Data in Europe”, Open Knowledge Foundation, ICT2010 Brussels, 28/09/2010. URL: http://www.w3.org/2010/09/egov-session-ict2010/JonathanGray_ICT2010.pdf.

Gøtze ,J.; Pedersen, C. B. (2009) “State of the eUnion. Government 2.0 and Onwards”, AuthorHouse.

Hernández-Ros, J. (2010) “Realizing the Value of Public Sector Information”, 13th PSI Group Meeting, Luxembourg, 24/06/2010.

Hogge, B. (2010) “Open Data Study”, Open Society Institute.

Internet & Gesellschaft Collaboratory (2010) “Offene Staatskunst“, Bericht, URL: www.collaboratory.de.

Lathrop, D.; Ruma, L. (2010) “Open Government: Collaboration, Transparency, and Participation in Practice”, O'Reilly Media.

Langkabel, T. (2010) “Open Data und Verantwortung”, Blog post, 12/11/2010, URL: http://www.langkabel.de/?tag=open-data

Malakoff, D. (2003) „Opening the Books on Open Access“, Science, 302, 5645, 550-554.

Martinez, R. (2010) “PSI Re-use in France: Overview and Recent Developments”, European Public Sector Information Platform, Topic Report, 10.

Matzat, L. (2010) “Mit Rohdaten allein wird der Bürger nicht glücklich”, Zeit Online, 12/11/2010, URL: http://blog.zeit.de/open-data/2010/11/12/innenministerium-rohdaten-lohmann/

Mayer-Schoenberger, V.; Lazer, D. (2007) “Governance and Information Technology“, MIT Press, Cambridge, MA.

Metzl, J. (2001) “Perils of Secrecy in an Information Age”, Carnegie IMP Magazine, July.

Miller, D. R.; Doyle, J. D. (2007) „Understanding the open access data movement“, Canadian Journal of Anesthesia, 54, 11, 949-950.

Milner, E. M. (2000) “Managing information and knowledge in the public sector”, Routledge.

Müller, P. (2010a),“Government 2.0: Eine Logik für Politik und Verwaltungshandeln”, Keynote, Messe Effizienter Staat, Berlin, 28/04/2010. URL: http://www.daten.effizienterstaat.eu/2010/mueller.pdf.

Müller, P. (2010b). “Ignoring the ROI of Openness”, blog post, 19/7/2010, URL: http://www.philippmueller.de/ignoring-the-roi-of-openness/.

Murray-Rust, P. (2008) “Open Data in Science”, naturepreceedings, 10.1016/j.serrev.2008.01.001.

Nabatchi, T.; Mergel, I. (2010) “Participation 2.0: Using Internet and Social Media Technologies to Promote Distributed Democracy and Create Digital Neighborhoods” in: Svara, J.H.; Denhardt, J. V. (Eds.), White Paper: Promoting Citizen Engagement and Community Building. Phoenix, AZ: Alliance for Innovation.

Nichols, R. (2010) “Do Apps for Democracy and Other Contests create sustainable Applications?”, Government Technology, 07/11/2010, URL: http://www.govtech.com/e-government/Do-Apps-for-Democracy-and-Other.html?id=765522&story_pg=1&full=1

Noveck, B. (2009) “Wiki Government”, Brookings, Washington, D.C.

OECD (2005), “Digital Broadband Content: Public Sector Information and Content”, Paris.

_____ (2010a) “Good Governance for Digital Policies: How to Get the Most Out of ICT - The Case of Spain’s Plan Avanza”, Information Society Reviews, Paris.

_____ (2010b) “Modernising the Public Administration: A Study of Italy”, Paris.

Open Knowledge Foundation (OKF) (2006). Open Knowledge Definition. URL: http://www.opendefinition.org/.

O'Reilly, T. (2010) „"Government As a Platform", 11-39 in: Lathrop, D.; Ruma, L. (2010) "Open Government: Collaboration, Transparency, and Participation in Practice", O'Reilly Media.

Osimo, D. (2008) "Web 2.0 in Government - Why and How?", JRC Scientific Technical Reports, Seville.

Øvrebø, O. A. (2010) "Momentum building for open government data in Norway", European Public Sector Information Platform Topic Report, 5.

Poikola A. (2010) "Open data in Finland - bottom up and middle out, but not yet from top down", European Public Sector Information Platform, Topic Report, 12.

Parycek, P.; Sachs, M. (2010) "Open Government – Information Flow in Web 2.0", European Journal of ePractice, 9.

Pizzicannella, R. (2010) "Co-production and open data: the right mix for public service effectiveness?" In draft papers,10th European Conference on eGovernment, Limerick, Ireland. URL: http://pizzican.files.wordpress.com/2010/05/eceg2010_paper.pdf.

Pollock, R. (2009) "The Economics of Public Sector Information". URL: http://econpapers.repec.org/paper/camcamdae/0920.htm.

Quintas, P.; Lefrere, P.; Jones, G. (2003) "Knowledge management: A strategic agenda", Long Range Planing, 30, 3, 385-391.

Robinson, D.; Yu, H.; Zeller, W. P.; Felten, E. W. (2008) "Government data and the invisible hand", Yale Journal of Law and Technology, 11, 160-175.

SAS/ Forsa (2010) "Open Data – Open Government MONITOR2010: Wünschen Bürger mehr Transparenz?" URL: http://www.sas-deutschland.de/content/dokumente/studien/e623/views624/?gclid=COee64yCmaUCFQo65Qod2A2rJg

Sand F. (2010) "PSI in Sweden: from infringement to enforcement?", European Public Sector Information Platform, Topic Report, 9.

Schellong, A. (2008) "eGovernment: data growth and growing concerns" in: "The EU and the emergence of the "Big Brother" society", Challenge Europe series, European Policy Centre, Brussels.

Schuppan, T. (20010) "E-Government: Unaufgregt forschen statt weitere Technikhypes schüren", blog post, URL: http://www.government2020.de/blog/?p=242

Sharif, R. M. (2009) "Maximizing the Value of Public Sector Information for Scientific and Socioeconomic Development in Africa", KMAfrica 2009 conference, paper, 4-7/5/2009, Dakar, Senegal.

Siegumfeldt F. (2010) "Re-use of PSI in Denmark", 13th PSI Group Meeting, Luxembourg, 24/06/2010.

Soros (2001) „Budapest Open Access Initiative", URL: http://www.soros.org/openaccess

Stephenson, D.W. (2009) "Data Triggers Transformation". In Gøtze J., Pedersen C. B., eds. "State of the eUnion: Government 2.0 and Onwards", AuthorHouse.

Sunlight Foundation (2010) "Ten Principles of Open Data", URL: sunlightfoundation.com/policy/documents/ten-open-data-principles/

Swedish Government eDelegation (2010) “As simple as possible for as many as possible. Under construction - future eGovernment”, SOU, 62.

Tauberer, J. (2009) “Open Data is Civic Capital: Best Practices for Open Government Data”, URL: http://razor.occams.info/pubdocs/opendataciviccapital.html

The Economist (2010) „Of governments and geeks“, Feb 4th.

Transparency International (2009) “Informationsfreiheitsgesetz der Länder”, 12/1, URL: http://www.transparency.de/fileadmin/pdfs/Themen/Verwaltung/TI_Informationsfreiheit_Bundeslaender_Zusammenstellung_Jan.pdf

Tuomi; I (1999) “Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organizational memory”, Journal of Management Information Systems, 16, 3, 107-121.

Uhlir, P. (2004) “Policy Guidelines for the Development and Promotion of Governmental Public Domain Information”, UNESCO, Paris.

von Luck, J.; Geiger, C. (2010) „Open Government Data – Frei verfügbare Daten des öffentlichen Sektors“, Gutachten, Deutsche Telekom Institute for Connected Cities, Zeppelin University, Friedrichshafen.

Weiss, P. (2002) “Borders in Cyberspace: Conflicting Public Sector Information Policies and Their Economic Impacts”. In Aichholzer, G., Burkert, H. (2004) “Public Sector Information in the Digital Age: Between Markets, Public Management and Citizens’ Rights”, Edward Elgar Publishing.

Zijlstra T. (2010a) “State of Play: PSI in the Netherlands”, European Public Sector Information Platform, Topic Report, 17.

Zijlstra T. (2010b) “A Community Perspective on PSI”, 13th PSI Group Meeting, Luxembourg, 24/06/2010.

Appendix

Survey

A. GENERAL QUESTIONS
1. [Do they have an open government policy?]
2. [Do they have a social media policy?]
3. [Do they have an open data policy?]
4. Is there a data.gov style portal?
5. Is there a statistics portal?
6. Which authority offers the data?
7. When was the portal created?
8. How many data sets are there?
B. COMPLIANCE WITH THE SELECETED OPEN DATA PRINCIPLES
9. What types of file formats does it offer?
10. Is there a registration?
11. Is the registration free of charge?
12. Is there a fee for obtaining data or part of it?
13. Are data provided license-free?
14. Can the disabled individuals use the portal?
15. How many languages does the interface have?
16. How many languages are data presented on?
C. USABILITY
Layout of the web-site
17. Does the portal have sections/ headings/ tags to make the navigation easier?
18. Are there any services?
Statistics and metadata
19. Is there statistics about users?
20. Is there statistics about sections?
21. Is there information about most popular data sets?
22. Are there metadata?
Easiness of Use for a New User
23. Does the portal have “Help” functionality (e.g. tutorial)?
24. Is there an FAQ section?
25. Is there an email/ telephone contact info for questions and feedback on the web-site?
26. Is there an opportunity to provide Feedback on the data?
Search Engine
27. Is there a search engine?
28. What kind of search does the search engine offer?
29. How many criteria one can choose?
Interactivity/ Social Media
30. Is the portal connected to a social media platform?
31. Is there RSS?
32. Is there a forum?
33. Are users allowed to create and post apps?

Email inquiry to government entities

1. Do you have an "open data" policy?
-. If yes, please provide me with a copy or link
- If no, are you working on an "open data" policy?

2. Which government agency is responsible for the "open data" policy making?

3. Do you have a meta-portal such as data.gov that provides your citizens access to government data?

4. Which government agencies have put most of their data online?

5. Can users suggest government data-sets that should be made available to the public?

6. Nowadays, social media (e.g. facebook, blogs, twitter, wikis) are used by many people which sometimes leads to various challenges for organizations. Has your government decided to develop a social media policy as a guidance for government employees?

List of the reviewed Internet OGD portals

Germany:	www-genesis.destatis.de
France:	www.statistique-publique.fr, http://www.data-publica.com/en
Denmark:	www.statbank.dk
Sweden:	www.ssd.scb.se
Norway:	ssb.no
Finland:	stat.fi/
The Netherlands:	www.cbs.nl
Italy:	dati.piemonte.it, en.istat.it/
Spain:	opendata.euskadi.net, www.ine.es

List of the identified OGD civil society initiatives

Germany:	offenedaten.de, offenerhaushalt.de, gov20.de
France:	www.data-publica.com
Sweden:	www.opengov.se/
Norway:	datakilder.no
Finland:	opengov.fi
The Netherlands:	overheidsfeeds.nl; www.waag.org
Italy:	www.spaghettiopendata.org/en
Spain:	datospublicos.jottit.com/

The views expressed in this publication are those of the authors and do not necessarily reflect the views of CSC.

CSC
Abraham-Lincoln-Park 1
65189 Wiesbaden
Telefon: +49.611.142.22222
E-Mail: deutschland@csc.com

www.csc.com/de

Worldwide CSC Headquarters
The Americas
3170 Fairview Park Drive
Falls Church, Virginia 22042
United States
+1.703.876.1000

Europe, Middle East, Africa
Royal Pavilion
Wellesley Road
Aldershot, Hampshire GU11 1PZ
United Kingdom
+44(0)1252.534000

Australia
26 Talavera Road
Macquarie Park, NSW 2113
Australia
+61(0)29034.3000

Asia
139 Cecil Street
#08-00 Cecil House
Singapore 069539
Republic of Singapore
+65.6221.9095

About CSC
The mission of CSC is to be a global leader in providing technology-enabled business solutions and services.

With the broadest range of capabilities, CSC offers clients the solutions they need to manage complexity, focus on core businesses, collaborate with partners and clients,and improve operations.

CSC makes a special point of understanding its clients and provides experts with real-world experience to work with them. CSC is vendor-independent, delivering solutions that best meet each client's unique requirements.

For 50 years, clients in industries and governments worldwide have trusted CSC with their business process and information systems outsourcing, systems integration and consulting needs.

The company trades on the New York Stock Exchange under the symbol "CSC".

www.csc.com