



European Commission
Information Society and Media
Directorate-General

POPSIS

Pricing Of Public Sector Information Study

Open Data Portals (E)

Final Report

October 2011

Marc de Vries (Citadel Consulting)
Lionel Kapff (Deloitte Consulting)
Mar Negreiro Achiaga (Deloitte Consulting)
Patrick Wauters (Deloitte Consulting)
David Osimo (Tech4i2)
Paul Foley (Tech4i2)
Katarzyna Szkuta (Tech4i2)
Joanna O'Connor (Citadel Consulting)
Diane Whitehouse (The Castlegate Consultancy)



Acknowledgements

Many individuals and organizations have contributed to the success of this study. We wish to acknowledge some of the key contributors and thank them for their assistance.

First and foremost we would like to thank the team's Project Officer, Mr. Juan Pelegrin, his Head of Unit Mr. Richard Swetenham and his staff in DG INFSO. In addition, special thanks are due to the many individuals who consented to give interviews and who responded to our many questions. Finally, an expression of our gratitude is offered to Mr. Robbin te Velde, Mr. Raimondo Iemma and Mr. Graham Vickery who participated in the meeting validating our final draft report.

The study team

Marc de Vries (Citadel Consulting)

Lionel Kapff (Deloitte Consulting)

Mar Negreiro Achiaga (Deloitte Consulting)

Patrick Wauters (Deloitte Consulting)

David Osimo (Tech4i2)

Paul Foley (Tech4i2)

Katarzyna Szkuta (Tech4i2)

Joanna O'Connor (Citadel Consulting)

Diane Whitehouse (The Castlegate Consultancy)

The study was awarded in response to a request for services to undertake a study on the assessment of the different models of supply and charging for public sector information (SMART 2010/0046) under the Framework Contract Lot 2 DG INFSO SMART 2009/0042 to Deloitte Consulting CVBA located in Brussels, in cooperation with external experts.

Legal Notices:

The opinions expressed are the view of the authors and their sole responsibility and not necessarily those of the European Commission or any of its services. Neither the European Commission nor any person acting on behalf of the European Commission is responsible for the use that might be made of the following information.

TABLE OF CONTENTS

1	Executive summary	4
2	Introduction	5
2.1	<i>Overview.....</i>	<i>5</i>
2.2	<i>Thematic introduction</i>	<i>5</i>
3	Methodology.....	7
4	Key indicators to monitor the impact of Open Data Portals.....	10
5	Web survey results and case studies	19
5.1	<i>Portal structure and governing body.....</i>	<i>20</i>
5.2	<i>Rationale.....</i>	<i>21</i>
5.3	<i>Portals development and maintenance costs.....</i>	<i>21</i>
5.4	<i>Functionalities.....</i>	<i>22</i>
5.5	<i>Data availability.....</i>	<i>23</i>
5.6	<i>Pricing and licensing models.....</i>	<i>26</i>
5.7	<i>Accessibility.....</i>	<i>26</i>
5.8	<i>Take-up</i>	<i>27</i>
5.9	<i>Drivers and barriers</i>	<i>30</i>
5.10	<i>Impact.....</i>	<i>31</i>
6	Conclusions – key success factors of open data initiatives	35
7	Annexes.....	38
7.1	<i>Annex 1: Data.gov portals case studies.....</i>	<i>38</i>
7.1.1	<i>United Kingdom – DATA.GOV.UK.....</i>	<i>38</i>
7.1.2	<i>Denmark – DIGITALISER.DK</i>	<i>46</i>
7.1.3	<i>Spain, Basque Country – OPEN DATA EUSKADI.....</i>	<i>52</i>
7.1.4	<i>France – DATA.GOUV.FR</i>	<i>54</i>
7.2	<i>Annex 2. Web survey</i>	<i>57</i>
7.2.1	<i>Web survey summary</i>	<i>57</i>
7.2.2	<i>Web survey results</i>	<i>62</i>

1 Executive summary

In the last three years (2009-2011), the **number of Open Data Portals has grown at an increasingly rapid pace**. Today, the notion of an open data portal has become mainstream. This high-profile policy activity is being implemented by a range of government organizations as a key element of their different open government strategies.

There is a large **variability among supply, take-up and impact of Open Data Portals**. The datasets on portals vary in number from 6,200 to five, with huge differences in the type of data being made available. In the study's sample, the volume of datasets downloaded varies from 320,000 to 250 per month. There are also occasional portals which function as directories rather than as repositories.

All portals make data available for free, in **most** cases using a **Creative Commons license**. The portals do **not determine direct changes in the pricing policy** for datasets, but they do **help to support a shift to PSI** that is distributed free or at marginal cost.

The portals' budgets vary considerable from the \$US 9.2 million so far spent on US data.gov to a hundred thousand euros on average allocated for European national and regional initiatives. The animation activities organized by most portals – such as apps competitions – continue to be popular among developers as the 430 entries to the recent Open Data Challenge (held in June 2011) show. The **largest data portal initiatives are, however, currently encountering funding problems** due to budgetary cuts. **Smaller initiatives are less affected**.

Open Data Portals appear to offer an important step in pushing forward the open data agenda and delivering its policy impact. Their **impact on opening up high added-value datasets is modest** and their **direct short term economic effects** have been so far limited. Their **largest impacts** to date are **indirect**: the portals stimulate **creativity and innovation** and **pave the way to unanticipated value creation**. In this context, a **“start small” approach appears to be most effective**. Open Data are an effective way to kick-start a process of cultural change that ultimately leads to the application of these high-level policy goals.

2 Introduction

2.1 Overview

This document is the final report of the POPSIS study objective E. It identifies key indicators to monitor the impact of Open Data portal initiatives (data.govs) and undertakes a measurement of several portal initiatives based on those variables.

This final report for the POPSIS study objective E is structured as follows:

This section 2 provides an overview and introduction to the report on objective E.

Section 3 explains the methodological approach to POPSIS study objective E.

Section 4 discusses key indicators to monitor the impact of Open Data portals.

Section 5 presents the main web survey and case study results.

Section 6 ends with a set of conclusions on the key success factors of open data initiatives.

The annexes in section 7 contain the complete web survey and case study results.

2.2 Thematic introduction

In the US, President Obama's Open Government memo from January 21, 2009 called for a more transparent, participatory and collaborative government and stipulated that "executive departments and agencies should harness new technologies to put information about their operations and decisions online and readily available to the public." The use of social media and the actual participation of all federal departments and agencies were reinforced by the Office of Management and Budget director Peter Orzag's executive directive. It gave agencies a 60-day deadline to publish their open government plans and upload their first datasets to a dedicated website called data.gov. Even before the launch of data.gov, another website already existed – data.octo.dc.gov. Vivek Kundra, the CIO in the Obama government and the person behind the data.gov initiative, put in place around 300 local datasets, such as real-time crime information, the locations of potholes, and school test scores when he was still the Chief Technology Officer of Government of the District of Columbia.

A few months after Data.gov went live in May 2009, the UK followed by launching data.gov.uk in September 2009. The UK and the US Open Data Portal initiatives have been copied by many international, national and local public bodies, such as the Australian national government, the Basque country, the city of Rennes in France, and the World Bank. An interactive map presenting governmental and open data community put together by CTIC Foundation can be found at datos.fundacionctic.org/sandbox/catalog/faceted/.



Figure 1: Screenshot of the interactive map of open data portals (government and citizen initiatives)

Today, the notion of an “open data portal” has become a policy meme, a catchy policy action which is easy to understand and implement, and is being replicated by different government organizations as a key element in their open government strategies. Often, this is accompanied by animation or public relations initiatives such as the now omnipresent apps competitions.

This part of the study has as its main aim to assess the extent to which open data portals have an impact on, or are influenced by, the PSI pricing policy. To achieve this, the team identified key indicators to monitor the impact of the Open Data Portals initiatives. This exercise has involved examining critical success factors for data.gov initiatives in Europe (including the relation to national and PSB charging policies) in comparison with their non-European counterparts (especially the US data.gov). The team studied several examples of Open Data Portals initiatives in order to establish a list of barriers and drivers. The team members also examined the extent to which portals could impact the price and availability of PSI, for example, in terms of lower pricing policies and easier re-usability of data for both commercial and non-commercial use.

3 Methodology

As the data.gov initiatives are still at an early stage of development, there is no robust data on supply, usage and impact of their portals. The study team therefore implemented a dual approach:

- A traditional ICT measurement framework focusing on supply, take-up and impact;
- A qualitative analysis of the data at citizens' disposal (for example, range and rankings, availability, and any machine-readable formats) and other qualitative aspects of data portals (such as app registers, visualization techniques and awareness-raising activities).

The data gathering techniques included:

- A web survey of 15 portals complemented with an e-mail request for data that were not available on portals themselves;
- In-depth interviews with three existing Open Data Portals (two national and one local) and one project.

Web survey

The online data collection exercise entailed the development of a list of Open Data Portals. The portals are listed in the table below.

Table 1: List of portal initiatives

	EU	Non-EU
National	UK data.gov.uk Finland data.suomi.fi Denmark digitaliser.dk	New Zealand data.govt.nz Australia data.australia.gov.au US data.gov
Regional	Asturia, ES risp.asturias.es/catalogo Basque Country, ES opendata.euskadi.net Catalonia, ES dadesobertes.gencat.cat Piemonte Region, IT dati.piemonte.it	Utah, US utah.gov/data/

Local	Helsinki Region, FI www.hri.fi London, UK data.london.gov.uk/datastore Rennes, FR data.rennes-metropole.fr	NYC, US nyc.gov/html/datamine
--------------	--	---

In order to create international benchmarks, the web survey did not limit its scope to European initiatives but complemented its work by exploring a few portal examples from Australia, New Zealand and the US. The web survey examined a number of variables that are presented in the next section of this report.

The web survey was accompanied by an email with a request for information that could not be obtained via the website (such as the number of applications developed or any success indicators used by portals themselves).

In-depth interviews

The quantitative analysis was complemented by three in-depth case studies on three specific portal initiatives:

- Basque country data.gov initiative opendata.euskadi.net
- Denmark digitaliser.dk
- UK data.gov.uk

Moreover, due to the recent 2011 French government initiatives, another interview was conducted with representatives of the body responsible for the data.gouv.fr initiative.

The case studies were based on a thorough literature review of existing studies and reports and in-depth interviews with managers of Open Data Portals.

The interviews aimed to deepen the information collected through the web survey, and particularly looked at:

- Background and history of the portal creation and deployment (such as the initiative launch, number of datasets)
- Pricing and availability of data (charging policy, conditions of reuse, etc)
- Drivers and obstacles
- Impact (take up, number of apps developed).

Table 2: Interviewees for objective E

Interviewees	
data.gov.uk	James Forrester and Elyria Miller
data.gouv.fr	Séverin NAUDET, Director, Etalab, Secretary general of Government
digitaliser.dk	Cathrine Lippert, Special Advisor, National IT & Telecom Agency
Open Euskadi Project	Alberto Ortiz de Zarate, director of the Open Data Euskadi

The following figure summarizes the approach and planning for objective E:

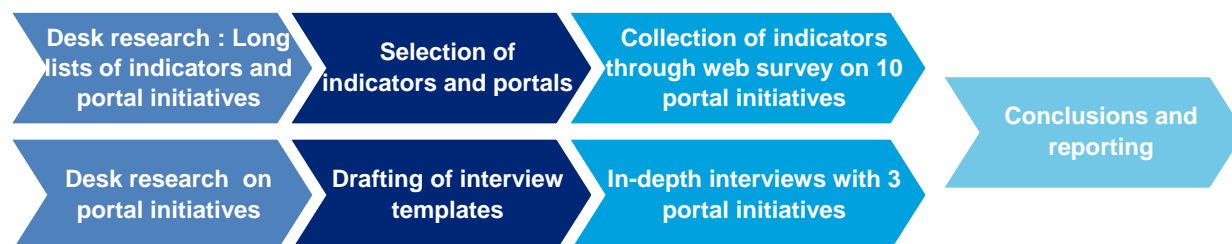


Figure 2: Approach for objective E

4 Key indicators to monitor the impact of Open Data Portals

The final list of indicators developed to measure the impact of Open Data Portals was established by the study team after a thorough literature review and was tested via a web survey.

Table 3: Web survey variables

Number of datasets available	
Timeliness of updates	
Data format	Machine readable Format: cvs, xml, rdf, Raw data vs processed data
Re-use conditions	License free or license bound, type of license
Pricing	Marginal cost, free or cost recovery
Institutional positioning of the portal governing body	
Accessibility	Anonymous or log-in access, accessible for everyone – accessible formats (proprietary or non-proprietary), different languages options
Take-up by citizens	Number of visits and unique visitors Number of data downloads
Take-up by app developers	Number of visits and unique visitors Number of data downloads
Number of applications developed	Apps competitions Number of app in app register

The analysis of the results of this exercise resulted in a **revised version of the key indicators** that is presented below:

Table 4: Open Data Portals impact – key indicators

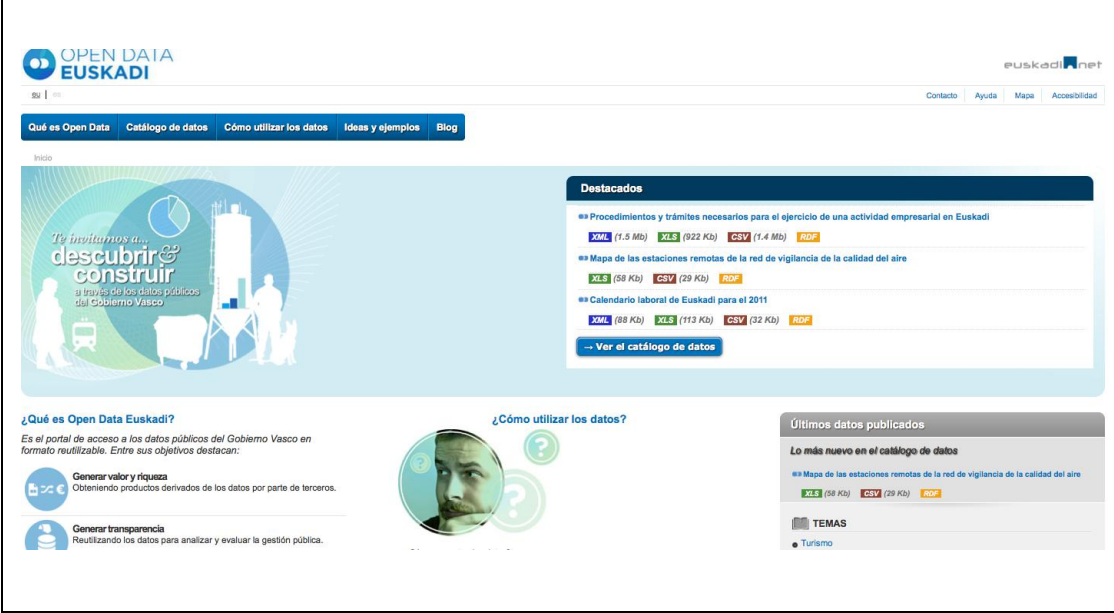
1. Number of datasets available or number of dataset listing pages	
2. Implementation of five star linked data system of Tim Berners-Lee¹	<ul style="list-style-type: none"> - data availability (★★make your stuff available on the web (whatever format)) - machine-readable data (★★★make it available as structured data (e.g. excel instead of image scan of a table)) - non-proprietary formats (★★★★ non-proprietary format (e.g. csv instead of excel)) - linked data (★★★★★use URLs to identify things, so that people can point at your stuff , ★★★★★ link your data to other people’s data to provide context)
4. Clear re-use conditions (license-free or license bound)	
5. Position of the governing body (high-level, low-level)	
6. Accessibility – anonymous access	
7. Take-up by citizens	<ul style="list-style-type: none"> • Number of visits and unique visitors • Number of data downloads • Conversion rate (click-throughs to data hosting websites) in case of web repositories
8. Take-up by application developers	<ul style="list-style-type: none"> • Number of apps • App contests entries numbers
9. Availability of take-up data on the website	
10. Range of government agencies contributing dataset listings	
11. Availability of ‘suggest a dataset’ form	
11. Presence of guidelines, events and other capacity building strategies	
12. Community engagement indicators	<ul style="list-style-type: none"> • Number of blog comments, tweets, activities on fora, ATOM feeds subscriptions, other feedback channels and metrics

The study team believes that this revised and final list of indicators can be used to measure the impact of existing Open Data Portal initiatives and benchmark them as well as help in creating new projects.

¹ <http://www.w3.org/DesignIssues/LinkedData.html>

Case studies overview

The case studies were based on interviews with managers of three existing Open Data portals – Basque Country, Denmark and UK – and further desk research. They are presented in the Annex. Nevertheless, this section introduces the highlights of each of the case studies.

Opendata.euskadi.net	Basque Country, Spain
Government of Basque Country, Open Data Euskadi Project	
	
<p>Launch date: April 2010</p>	
<p>Short description</p> <p>The Millennium Open Data Euskadi portal is aimed at generating added-value from PSI, promoting transparency in public administration, interoperability and efficiency of public agencies. It was the first non-Anglo-Saxon gov.data portal.</p>	
<p>Data</p> <p>The portal makes available 1,125 statistic datasets and 77 other datasets, i.e. total of 1,202 in machine-readable formats (xml, csv, wms, other).</p>	
<p>Pricing</p> <p>The data are made available on the condition of licensed recognition of authorship (set to a Creative Commons Attribution 3.0 Spain) or free licence and free re-use of public information.</p>	
<p>Apps</p> <p>Approximately 10-15 users have downloaded the data in order to develop an app, although a register of apps is not available.</p>	

Accessibility

The access to data is anonymous. The website is compliant with WCAG 2.0.

Impact

Between 6 April 2010 and 25 February 2011, the website has been visited by 25,362 unique visitors, had 33,682 visits and 157,520 pages were visited. Over 250 datasets are downloaded each month. The competition organized by the portal team attracted 30 working apps. The portal features also interesting applications based on PSI, such as infocarretera.com (traffic data), metroo.es (accident data), wikiloc (geo ortofoto), euroalert.com (tenders). Still, the data on re-use and end-user impact are rather scarce, since the portal does not require registration in order to maximize data re-use.

According to the interviewee, the launch has not had any impact on PSI pricing, as all the released data were already public although they were not easily accessible by citizens and re-users.

The portal launch had a high impact on internal public administration, as the largest re-users of the portal's data are government agencies. Moreover, it had impact on public administration interoperability as other government agencies on different levels are using the released data (e.g. weather widget in cities websites).

The portal had resulted in increased business productivity: the most requested dataset is the official calendar of public holidays, which are different in each municipality. This dataset helps companies to organize their work better.

Still, the greatest impact is overall indirect economic impact in terms of efficiency of the market.

Critical success factors

The portal owes its success to the strong Basque tradition of transparency and innovation. The engagement of the strong political leadership was also a crucial factor in the fast and effective portal launch. The adoption of a low-hanging fruit strategy helped with a rapid release of many datasets. The interviewee underlined the importance of the adoption of low cost solutions in the release of a large amount of data. Finally, the strong connection with PSI re-users and the use of existing open standards are matters of considerable importance.

Digitaliser.dk	Denmark
National IT and Telecom Agency, part of Ministry of Science, Technology & Innovation	
Launch date: Summer 2009	
Short description	
<p>The Data Catalogue (as Danish gov.portal is a web directory, data are located at the public authorities' own servers or chosen host servers) is hosted at the collaborative platform (digitaliser.dk is), created and maintained by the Danish National IT and Telecom Agency, which is part of Ministry of Science, Technology & Innovation. The platform is a collaboration tool open for public. The portal is available only in Danish.</p>	
Data	
<p>The catalogue lists approximately 900 public datasets (which were put together by screening the websites of all state agencies and ministries). It means that not all public data already available in Denmark are listed on the website. In December 2010, the list was updated. The directory links to data in machine-readable formats (such as GeoJSON, HTML, JSON, KML, XML, xls, html) and PDFs. The list consists mainly of central government datasets but local governments are being encouraged to participate (there are already some local municipalities datasets).</p>	

Pricing

The data are available either for free or at marginal cost depending on the dataset.

Apps

The portal launched an app contest in 2009 (from October 2009 to February 2010). Three application ideas were awarded a prize of 100.000 DK (these were given to two individual developers and one company). The contest comprised two parts, ideas gathering and solid deployment projects. Twenty-five applicants submitted app projects in the second round app contest (NGOs, private companies and individuals).

Impact

Until now, the only feedback, which may help to analyze the impact, was gathered through face-to-face meetings with re-users and users' comments on the platform. The platform functionalities unfortunately do not allow any tracking of the usage of the data catalogue. NITA is currently working on a reconfiguration of its statistics-gathering tool to be able to analyze some quantitative traffic data.

The Danish open data policy is very much targeted at the middle layer of three important agents i.e. professional public administrators mentioned by Tim Berners-Lee ("It has to start at the top, it has to start in the middle and it has to start at the bottom"²). This overall strategy of awareness raising, community building and putting in place designated online space for a digital catalogue (even in its present directory form), has already:

- encouraged the PSBs to publicize their online datasets to the wider public,
- centralized the information on PSI re-use in Denmark,
- encouraged the data holders to release more data .

Critical success factors

The overall digitalization of society, the arrival of Web 2.0 and the availability of interesting examples of public data use, drives the demand for this data in Denmark. Even though there is no legal obligation in this field, thanks to the bottom-up engagement of local enthusiasts working in public bodies, many datasets are being released and their supply is thus increased. The NITA agency has encouraged public bodies to adopt an agile approach starting with "low-hanging fruits", i.e. datasets which can be easily (and often virtually costlessly) published, and then moving on to publishing datasets that involve more workload

² Citation following Hogge, 2010

(and therefore costs) at a later stage. They are responding to the actual demand of re-users and citizens accessing data.

What is more, as the overall level of digitalization of public data in Denmark is very high, the data released in digital format often do not accrue high costs.

The economic crisis, globalization and crisis of welfare state, make decision-makers much more sensitive to the idea of innovative, efficient and co-produced services that might help to respond to the upcoming societal challenges.

The PSI re-use directive cannot be considered as an important driver owing to its generality. However, the EC is still seen as an important actor in the PSI re-use field through its showcasing of best practices and facilitation of knowledge exchange.

Data.gov.uk	UK
Transparency and Digital Engagement team in the Cabinet Office	
Launch date: 30 September 2009	
<p>Short description</p> <p>Data.gov.uk is still operating as a beta version. Data.gov.uk is a key component of the UK Government's Transparency Agenda, which is being led by the Transparency Board. Its implementation is the task of the Transparency and Digital Engagement team in the Cabinet Office, which is working across government departments to ensure that data are released in a timely and accessible way. It aims at displaying data in easy-to-find, easy-to-license and easy-to-re-use ways. The publishing team draws on the combined expertise and wisdom of the Transparency Board to publish government data in a manner consistent with the Public Data Principles.</p>	
<p>Data</p> <p>The portal makes available over 6,200 datasets in machine-readable formats (csv, xml, etc). The publishing team plans to convert datasets to use Linked Data standards.</p>	
<p>Pricing</p> <p>The data and information are licensed under the Open Government Licence, which enables commercial and non-commercial use of data (unless stated</p>	

otherwise, e.g. Met Office data are subject to the Met Office license).

Apps

The [Show Us a Better Way](#) contest was organized before the launch of the portal. On 22 April 2011, the site recorded that more than 140 applications that had been developed by users. Reviewers gave 40 of the applications a five star rating which is the highest possible rating. Five star applications include those relating to how to find post boxes, care homes and determine air quality. Other five star applications map road accidents, crime, landfill sites and pharmacies.

Impact

Interviewees suggested that development of the site is still at an early stage and use should increase in the future as it is further developed. There are considerable fluctuations in use as new datasets are added and there is media coverage of their release. For example, the crime mapping application received 18 million hits per hour on the day of its launch. It received more than 400 million hits in its first three months of operation. Data.gov gets about 80,000 unique users per month.

A number of events, app development competitions and data hacks have been organized to support the achievement of the open data and transparency agendas (e.g. the Open Knowledge Foundation held a two-day international workshop in 2010 at the University of London Union, LinkedGov held its first hack camp on the 9 and 10 April, 2011 in Shoreditch, London).

Critical success factors

Interviewees felt they had made the right choice in launching the portal quickly with as much data as possible rather than spending time cleaning data and delaying the launch until the data and site were pristine. It was also highlighted that the idea of linking the data with a non-transactional license was important.

Data visualization was thought to be an important factor in enabling access for citizens. The team are now looking for cost-effective solutions that will provide automatic generic visualization of any data.

5 Web survey results and case studies

This section of the report outlined the results of the web survey exercise that is based on the portals' set of characteristics (see detailed information on the portals in Annex 2). The characteristics described in the methodology section enabled a comparison to be made across different government levels, cross-country and between the European and US approach.

So far Europe has only one fully developed national Open Data Portal – the UK data.gov.uk. Other national initiatives have so far been limited to web directories listing data sets that are available on different government and agencies websites (Aporta in Spain and the two cases presented in this report – Denmark and Finland).

The Norwegian government has started to prepare for the launch of data.norge.no. So far, the website has had the role of a hub on information on open data which makes very few datasets available (11 mostly yearly updated datasets). Nevertheless, the new version of the website should be published in a few weeks following the submission of this report. It will most probably contain a significant number of datasets in comparison to the existing version. Furthermore, France has plans to build data.gouv.fr by the end of 2011 (for more information, see the case study in the annex).

The European regions are represented in this study by three Spanish examples – Asturia, Basque Country, Catalonia – and the Italian Piedmont region. These cases are complemented by an analysis of the US Utah state portal. The study has also investigated some local initiatives, such as London Data Store, Data.rennes.metropole.fr, the recently launched Helsinki Region Infoshare and the New York City DataMine.

Furthermore, this section also discusses information gathered thanks to four in-depth case studies of the UK, Danish and Basque Country data portals as well as the planned French portal initiative introduced already in the previous section. The case studies complemented the information gathered thanks to the web survey but also investigated the question of drivers and barriers of those initiatives.

The figure below presents the timeline of the development of all data portals in the sample.



Figure 3: Timeline of data.gov project from the sample

5.1 Portal structure and governing body

The institutional positioning of the portal governing body plays an important role in the launch and future developments of the Open Data Portals. Obviously, portals launched and managed at the highest political level (such as by a President or Prime Minister) are expected to have higher political priority and therefore higher potential impact.

The first portals – those of the US and UK – were leading initiatives started by their respective governments and were closely related to the political leads. In the US, the portal creation was led by the Chief Information Officer under the express mandate of President Obama and has been run by his team. It was by the President's decree that requested governmental agencies to publish at least three high-value datasets on the US data.gov portal. In the UK, the data.gov.uk portal was also an initiative from the Prime Minister's office that featured highly in the press and media. The team in charge of the portal is the Transparency and Digital Engagement team in the Cabinet Office. The Australian portal is run by Australian Government Information Management Office.

On the contrary, the positioning in the government hierarchy of other national portals under study is much lower – the Finnish portal is run by State Treasury's Government IT Shared Service Centre, the Danish portal by National IT and Telecom Agency is part of Ministry of Science, Technology & Innovation and New Zealand's portal is under the remit of the Government Technology Services at the Department of Internal Affairs.

The position of the portal's team may have obvious consequences when it comes to the sustainability and the influence of the initiative towards widening the scope of datasets or increasing budget for the website development. However, even flagship initiatives such as the US data.gov portal are currently facing budget cuts³. The running of the website was created through an executive action of the president and therefore it is not guaranteed federal funding. It may even face a shutdown.

Regional and local portals are much more often run by a dedicated project team in specific departments (such as the Catalonia Open Data Project, Open Data Euskadi Project in Basque Country and the Greater London Authority Digital Projects). The City of Rennes and New York chose to appoint their IT departments to be in charge of their portals.

This overview of the variety of different types of governing bodies that are in charge of portals shows that it is not easy to see clearly what is the most effective institutional level to be responsible for a portal initiative. Project teams may be more effective than the departments that take on another responsibility, but they risk struggling to maintain the funding once the project cycle of development and launch

³ http://www.washingtonpost.com/blogs/federal-eye/post/budget-could-close-the-door-on-open-government/2011/04/13/AFttdXbD_blog.html?wpisrc=nl_fedinsider

is over. The recent data.gov budget cuts prove that, even if an initiative is started and endorsed by the highest level of hierarchy, its sustainability may still be challenged.

The Open Data Portals under study differ substantially in structure. The UK portal is a data repository as are the US and Australian portals, and the Danish, Finnish and New Zealand portals are web directories with links. Finally, the French portal is planned as a mix of a data repository and index of data. The regional portals are more often data repositories: for example, the Basque country portal is a distributed repository with a single point of access, similar to cities that most often opt for a distributed repository if they list many datasets.

5.2 Rationale

The underlying rationale of the US initiative was the enhancement of transparency and democracy, therefore it had a focus on public spending. In the UK, the portal was launched in September 2009. It was initiated by the Labour Party when in government power and later taken over by the coalition government. It has changed its content slightly from the opening up of government, efficiency of public services and economic growth and social value stimulation to government and public bodies transparency. This change in its rationale was very visible with the establishment of the Transparency Board and publishing of government spending data and public officials salaries. In Denmark, the rationale was much more focused on the importance of innovative services and the unlocking the value of public data. The Basque country portal was built in order to enhance the openness of the government and transparency but also with focus on innovation. Finally, the French portal is being developed with a view to stimulate innovation, and enhance the transparency of the government and its agencies and move the openness principle forward. In summary, the main drivers of open data portals are better public services (such as traffic management), increased monitoring of government activities (such as government spending) and new business opportunities (such as new apps to be developed).

From these four case studies examples, it can be seen that the transparency and openness principle are often linked to a high level of political commitment, which reflects the examples of the portals that received the support of Prime Ministers and Presidents (as in the US, UK and France). The countries which lack national open data policies more often underline the tangible results of open data release – to foster innovative services and stimulate social value.

5.3 Portals development and maintenance costs

The budgets of all the portals under study vary considerably. According to the White House's IT Dashboard, Data.gov has cost US\$ 9.2 million by mid-2011 and USASpending.gov has cost US\$ 13.3 million to date. It cost around US\$ 1.3 M to

launch the portal and total planned costs until 2020 are estimated on US\$ 102⁴. Still, the portal faces already major cuts: the April 2011 Senate cuts to the Electronic Government Fund from \$US 35 million down to \$US 8 million cannot cover the costs of the three portals, the data.gov, spending.gov and IT dashboard portals.

The UK portal costs have been not yet presented to Parliament. The interviewees estimated the costs as much lower than the US data.gov since the core Cabinet Office team developing the portal consists of only six full time developers. The Danish catalogue which indexes data without providing hosting required a budget of EUR 67,000 (DKK 500,000) for the creation of four catalogues (data, open source software, XML schema and standards), which were subsequently added to the existing platform. The costs of maintenance are difficult to estimate since the catalogue is part of the platform. That solution reduced the costs considerably. Finally, the Basque Country website has a budget of EUR 100,000 for both technological development and animation (which constitutes the bulk of the costs). France plans to allocate EUR 350,000 to the data.gouv.fr launch.

5.4 Functionalities

The different rationales and positions of the governing body result in different citizen engagement strategies. On the one hand, even if the US and the UK portals are committed to community building activities (such as data camps, hackathons, apps registers), they are focused on making the data available to the public and businesses. The Danish portal team energy is, on the other hand, focused on apps developers' capacity-building and awareness-raising by publishing the business case, guidelines for civil servants and developers and organizing conferences and open data days. Still, all the interviewees underlined the importance of applications and data visualizations in showing the value of data to citizens and the market.

Most of the portals offer much more to the users than just an access to a data catalogue. Some of the Open Data Portals add supplementary functionality in order to raise awareness about the data and stimulate re-use (examples include the Basque Country, Piedmont Region, UK and US). Other portals provide an add-on to an existing portal on PSI or data digitalization (as it is in the case of Denmark and Finland). US data.gov also offers community areas (open data, health, semantic web and law) aimed at capacity-building of the PSI re-users, public sector officials interested in data re-use and PSI evangelists and software tools for developers. The UK data.gov.uk portal makes available blogs, forum, wiki and some resources on open data. The digitaliser.dk portal is built to support and stimulate the community (the data catalogue is only a small part of its strategy) and contains a community forum, resources on digital services and knowledge society and an open source software catalogue. Piedmont informs its users about planned PSI events. On registration, the Rennes portal gives access to a developers' corner that contains a forum for PSI re-users. Finally, some portals have elements of social media strategy, twitter account and blog (examples include Piedmont, Basque Country, Helsinki and

⁴ <http://it.usaspending.gov/investment/cost-summary/603>

Utah). Two of these portals offer automatic translation via Google translate (NYC, and Utah). For more details see additional features summary in Annex 2.

5.5 Data availability

The quantity of data available is a rough indicator of the size and importance of these portal initiatives. However, the very notion of what constitutes a dataset is far from clearly defined, and is open to flexible interpretations that can significantly change its meaning. For example, one UK agency counted as two datasets what is in fact the same dataset but classified by gender. Thus, the methodology of counting datasets differs, some portals count every single dataset (e.g. every year, every variable) whereas others count datasets in clusters. It is important to refrain from a competitive approach towards benchmarking portals that "have the most datasets". In general, all Open Data Portals adopted the approach to start going public even with few datasets, and then increase the volume based on demand.

The portals differ very significantly in the number of datasets available. The UK portal makes over 6,200 datasets available in machine-readable formats whereas the US data.gov portal has around half of that number (3,200, complemented by 385,821 geospatial datasets). This difference in metrics methodology makes it difficult to establish which portal has more datasets. Other national Open Data Portals have released much less data – from a couple of hundred items of data (800-900 links to datasets on Danish portal, more than 500 in New Zealand and more than 200 in Australia) to 46 links on the Finnish portal.

The regional and local portals have an obviously narrower scope. The biggest regional data.gov in the Basque country has an impressive collection of 1,200 datasets. The other portals' datasets oscillate in number around a couple of hundreds (549 Helsinki Region, 429 - London, 361 – New York, 242 – Piedmont, 124 – Catalonia) or less (90 – Rennes, Utah – 31 and Asturia – 5).

Table 5: Portals' datasets numbers

Name of the portal	Number of datasets (raw data)
Data.gov.uk	6,200
Data.gov	3,200 (+385,821 geospatial data)
Opendata.euskadi.net	1202
Data.digitaliser.dk	800-900 (links)
Data.govt.nz	561 (links)
Helsinki Region Infoshare	549
Nyc.gov/html/datamine	361
Data.london.gov.uk/datastore	242
Data.piemonte.it	242
Data.australia.gov.au	223
Dadesobertes.gencat.cat	124
Data.rennes-metropole.fr	90
Data.suomi.fi	46 (links)
Utah.gov/data	31
Risp.asturias.es	5

However tempting it is to measure the development and impact of the portal by the number of its datasets (a quantity which to some extent may show the importance of the open data initiative on the political agenda), these numbers do not take into account the data characteristics and value.

The accuracy of the data plays an important role in the data quality. Timeliness of the data (especially for traffic or meteorological data) is also very important. The data.gov.uk team and Helsinki Region Infoshare plan to use Linked Data standards simplifying access to given information. The value of the data for the market, transparency activists and citizens vary considerably from data set to dataset. For example, after the Obama decree on 8 December 2009 that urged governmental agencies to publish at least three high-value datasets, the results were far from impressive. As Ellen Miller of the Sunlight Foundation boldly put it "it turns out that the government has some weird ideas about what counts as 'high value' information. The Department of the Interior seems to feel that population counts of

wild horses and burros are 'high value' but records of safety violations like the ones that seem to have led to the recent West Virginia mine disaster are not"⁵.

Navigation through vast data repositories is simplified as a result of such features as rating options, tags, linked data or the most downloaded data rankings. In order to facilitate the search for data, Australia, the US and UK introduced a rating feature which permit users to rank the datasets according to their value or perceived usefulness. For example, from the data.gov website it can be seen that the most accessed dataset on this portal was *Worldwide M1+ Earthquakes, Past 7 Days* which has been downloaded 139,071 so far and *U.S. Overseas Loans and Grants (Greenbook)* 52,478 times. Food and food calorie listings, drugs and child-related product recalls are all also high in the rankings. The most downloaded and highest ranked file on dati.piemonte is data related to primary and secondary school students' statistics (1,030 times). Still, it is not known if the download data can be attributed to unique users or are generated by applications that access the data often.

Moreover, to enhance the data usefulness, some portals introduced the options of suggesting new datasets to add to the portal (US, Australia, London, NYC) and a repository of data re-use ideas or inspirational uses (UK, London, Rennes). Throughout 2009 the US portal received suggestions for approximately 900 datasets. Out of this number, the agencies estimated that 16 percent of the data is already published on Data.gov (Already Published), 26 percent of the suggestions can be published in the near future (Actionable), 36 percent of the suggestions could be published at a later date (Potentially Actionable), and the remaining 22 percent of the suggestions cannot be published due primarily to security, privacy, or technology constraints (Not Actionable)⁶. These suggestion systems do not, however, reflect fully the value of datasets as they tend to favour "most popular" datasets while some datasets could be highly useful for a specific kind of re-user.

All the case studies examples pointed out that the portals have adopted a "low hanging fruit" strategy. They release data that are already available for free in different agencies and public bodies, although the UK and the US focus more on data value: they release data on public spending and health. In the US, before the launch of the portals, the data.gov team asked volunteers from agencies to suggest which data they thought have the biggest value for citizens and businesses. This resulted in the primary release of geospatial, Census Bureau, Center for Disease Control and Environmental Protection Agency Data. Still, according to our interviewees many civil servants fail to notice the importance of some of the datasets. In fact, the very nature of open data is based on the notion that "the coolest thing to do with your data will be thought by someone else" (Pollock). It is ineffective to let government select the datasets to release. By definition, open data enable unexpected uses of data: for example, the most downloaded dataset from the Basque Country is the

⁵ <http://sunlightfoundation.com/blog/2010/09/14/the-future-government-transparency-and-tracking-money-in-politics/>

⁶ <http://www.data.gov/suggestdataset>

calendar of public holidays in each municipality, which is used by companies to plan their work calendars. Therefore many portals include a *suggest a data set* feature to gather more insight on which datasets to release in near future.

5.6 Pricing and licensing models

Most of the data available on Open Data Portals are available for free without registration, in order to promote maximum re-use. Data.gov, Utah and the Basque Country do not apply any license to the data use and re-use. Most of other portals licensed their data under national Creative Commons license.

Still, some portals develop their own licenses adapted to the specificity of the public data re-use and access. The UK developed a dedicated license, [Open Government Licence](#), London DataStore currently on licensing terms with the advice from Open Knowledge Foundation and Rennes put forward the [Licence de réutilisation des données publiques - Rennes Métropole en accès libre](#). Furthermore, some of the portals request a contact with the data holder (Catalonia) or the portal (Rennes, New York) prior to the re-use of the data and/or include a disclaimer available from the portal. Finally, the web directories (e.g., the Danish catalogue) have not standardized the licensing schemes between the PSBs since the licensing remains part of the remit of agencies.

According to UK interviewees, the current database protection law in Europe demands a coverage of copyright and database rights in order to be exhaustive (which is not the case if a portal uses the Creative Commons CCBY license).

5.7 Accessibility

Access to the data on most of the portals is anonymous. The web directories (New Zealand, Denmark and Finland) provide anonymous access to datasets' lists but some of the agencies may require registration before data download can occur. Still, the Rennes portal offers access to API and developers' corner only on registration.

Data accessibility can be understood also as availability of machine-readable data in non-proprietary formats. All data portals make most of their data available in machine-readable formats, although some also contain PDFs. Use of non-proprietary formats is not always ensured: some data sets are available only xls (proprietary) and not in csv (non-proprietary format). The UK is the most advanced in making datasets available as "linked data" in RDFa format, although it is not clear to what extent this approach has led to higher re-use of data.

Finally, not all portals have an explicit web accessibility policy (although the US, UK, AU, Basque Country, Catalonia, Utah and London describe theirs in length) which enables non-discriminatory access for all. However, the majority of sites possess at least some element of web accessibility functionalities (e.g. text size change or access key browsing).

5.8 Take-up

Overall take-up is measured mostly by downloads times and numbers of visits. The most complex statistics are available on the US data.gov portal. They entail:

- number of datasets per agency and times of downloads per week
- data sets per agency per month
- datasets by category
- daily and monthly visitor statistics
- downloads by category
- monthly downloads trends
- hits in 12 months,
- countries and states of origin of visitors
- top countries by hits
- most downloaded datasets since the launch.

The second portal that makes available some statistics on the website (via Google Analytics tool) is that of the Piedmont region. It presents most downloaded datasets (also per theme or per organization) and number of visits and country of origin of visitors. The web directories have fewer possibilities of measuring take-up as they have no direct downloads, nevertheless, the New Zealand portal measures click-through to data holder websites.

Other proposed metrics deployed by data portals are:

- outgoing ATOM feeds subscriptions (for dataset listing updates)
- qualitative information gathered through comments by portal users
- quantitative and qualitative analysis of blog comments and tweets.

Table 6: Data portals take-up statistics

Name of the portal	Visits	Downloads
Data.gov.uk	80,000 unique users per month	No data on downloads available yet
Data.gov	Around 160.000 visits per month	320 000 per month
Data.govt.nz	5.377 page visit per month, 15% of all page visits – click-throughs to datasets	No possibility of measuring downloads
Opendata.euskadi.net	2305 unique visitors, 3062 visits and 14320 page views per month	250 datasets per month

Data.piemonte.it	3000 visits per month	1700 per month
Dadesobertes.gencat.cat	6500 visits per month, 18.000 page views	Do not track downloads
Data.london.gov.uk/data store	Over 40.000 visits per month and 200.000 page views	6000 downloads per month (of CSV files only) which gives 7GB of data a year
Nyc.gov/html/datamine	Visits 8,864, visitors - 4,086 and page views, 21,229 (data from February 2011)	Does not track downloads
Data.digitaliser.dk	No data on take-up yet	
Data.australia.gov.au		
Data.suomi.fi		
Helsinki Region Infoshare		
Utah.gov/data	No data on take-up received	
Risp.asturias.es		
Data.rennes- metropole.fr		

The data.gov portal leads in terms of visitors' statistics: it has 160,000 visits a month. In comparison with European national data portals, data.gov.uk registers 80,000 unique visitors per month. Since other national portals are in fact web directories, they do not measure the take-up of the data catalogue. The New Zealand portal lists more than 5,000 visits per month, which is comparable with the regional portal of Piedmont (3,000), the Basque Country (3,000) and Catalonia (6,500). What is interesting, but maybe not so surprising, given the number of apps developed for those cities and the number of inhabitants, the London and New York data catalogues attract much more traffic than the regional portals (40,000 and more than 8,000 respectively).

As far as number of downloads are concerned, these numbers are less tracked by data portals and therefore are not available for two-thirds of the portals surveyed, the data.gov registers an impressive number of 320,000 downloads per month and the London Datastore 6,000. The regional portals have between a couple of hundreds and 2,000 downloads a month.

Re-users take-up

As all the main portals give anonymous access to data, their only possibility of tracking the developers' take-up or to discern it from citizens' take-up is to follow the number of apps created which is known to them.

The data.gov portal mentions 934 government and 236 citizen-developed apps (including 44 mobile apps developed by government agencies⁷ at the end of March 2011). The portal also showcases apps on its website⁸. The UK data.gov.uk showcases 140 apps built with the use of its data. Forty of these applications received a five star rating. The most popular app (crime mapping) noted 400 million hits in the first three months since the portal launch. The Finnish portal showcases 50 apps. The Australian portal presents 2 apps so far, but encourages developers to submit other examples. Catalonia showcases 12 apps, including three created by citizens.

Other ways to measure impact but also to incite the interests of developers is to organize an app contests. Today, this has become a sort of must-have for data.gov initiatives. Some initiatives were organized before the portals were launched (such as the UK's [Show Us a Better Way](#) contest, and Apps for Democracy Finland organized before national and regional portals launch date). Other portals launched a contest shortly after their launch.

The Danish portal launched an app contest in 2009 (from October 2009 to February 2010), in which three application ideas were awarded a price of 100.000 DK (two individual developers and one company). The contest comprised of two parts, ideas-gathering and solid deployment projects. Twenty-five applicants submitted app projects in the second round app contest (NGOs, private companies and individuals). The Australian portal has run one apps contest, Mashup Australia (<http://mashupaustralia.org/>) organized by Government 2.0 Taskforce with 82 entries and announced another - [Libraryhack](#), a mashup and apps competition designed to encourage the creative and innovative reuse of library data and digital content (May 2011). The latter is being run by the National, State and Territory libraries of Australia and New Zealand. London portal showcases around 62 innovative uses of its data⁹. One application, Tube Deluxe, which shows live departure boards of the metro, the quickest routes and whether there are any travel delays, has been downloaded so far 350,000 times. The Rennes portal has organized recently an app contest Rennes Métropole en accès libre, which resulted in 43 apps competing for 50,000 EUR. The New York City portal has already organized an apps contest twice, i.e., the NYC BigApps Competition (<http://nycbigapps.com/>). There were 57 submissions in the second round, including 24 mobile apps.

Finally, the Basque country has participated in apps contest organized by third parties, and estimates that at least 10-15 users have downloaded the data in order

⁷ <http://apps.usa.gov/>

⁸ <http://www.data.gov/developers/showcase>

⁹ <http://data.london.gov.uk/datastore/inspirational-uses>

to develop an app, although a register of apps is not available. The Catalan portal received 70 requests for data that may help to estimate the number of prospective app developers.

In general, data contests define a total prize of less than € 100,000 and generate dozens of apps.

5.9 Drivers and barriers

According to interviewees, the main driver of the UK initiative was the Task Force Power of Information report, a situation similar to Australian portal. In the US, the main driver was the engagement of the President Obama, which had started earlier with his legal initiatives in the Senate and the choice of Vivek Kundra as CIO and data.gov project leader. The US and the UK have also subsequently inspired the Danish and Basque country portals. This once more underlines the importance of high-level government hierarchy commitment. This commitment is coupled with bottom-up citizen open data movements that put pressure on public bodies and government for the free data release and pave the way by launching citizen data portals.

The EU PSI directive was not mentioned as the primary driver of the Open Data Portals as most of the data released was already available for free although not yet centralized in a single directory.

Finally, the most important drivers were external to government such as advent of Web.2.0 and in consequence gov.2.0, digitisation of information (the low cost of information release) as well as open source and open data movement.

One of the important barriers of the data portals launch and maintenance is the budget cuts related to the economic crisis. The US portal has to face senate-level cuts. The UK interviewees also mentioned this factor as an important restraint in further releasing data.

The second important barrier is the lack of middle layer engagement – the civil servants' reserved attitude towards open data release. As the data holders rarely see the importance of the data for wider society and business or are locked in their roles as public body revenue protectors, it is difficult to persuade them to release the data. What is more there are rarely any incentives used to further data release or data preparation.

Furthermore, the PSBs have no budget for changing their data sets into accessible, non-proprietary formats and often do not know the standards for data release. Since all the portals applied the low hanging fruit strategy, the risk is that they focused much less on cleanliness of the data or citizen and developers needs. Therefore the data may be underexploited due to low levels of interest and incompatible data formats.

Although all the portals work on open data community building, there is still no

strong community of re-users to advocate more rapid data release. There is the risk that the open data movement remains confined to “geeks”.

Finally, the Danish interviewee underlined the need for a European analysis on the economic and societal impact of PSI and open data re-use in order to prioritize important data sets and gather more arguments to persuade the laggards in the public bodies.

5.10 Impact

It is difficult to measure the impact of the portals covered by the study, since there are no long-term economic or social changes to be observed yet. Nevertheless, the interviewees underlined the increasing traffic data and popularity of apps and visualizations. The UK portal noted more than 140 apps based on its data, some of them of very high interest – a crime mapping app attracted 400 million hits in its first three months. The largest interest resides in apps that are useful to daily activities, such as traffic, and those that monitor sensitive issues such as crime.

The simple fact that the data catalogues list available data in the given country or region enables the developers and citizens to look and find the given dataset. The high media attention paid to the first data.govs has also encouraged many other PSBs to release their data.

Nevertheless, the case studies showed that the portals have had no impact on the PSBs pricing strategies as the data made available on portals were already available or not subject of any pricing regime.

The Basque Country example shows that one of the important beneficiaries of the data portals is public administration itself as government agencies are often the largest data re-users and find the portal much easier to use.

Moreover, the same case study pointed out that the applications developed thanks to open data can be used by other government structures on different levels. For example, in the Basque Country many cities make use of the weather widget and central government may make use of some local data.

The impact on business is difficult to measure but once more the Basque Country case shows that the most requested dataset in the region is the official calendar of public holidays, which are different in each municipality, and which help companies to organize their work more effectively.

Therefore, the greatest impact overall is the indirect economic impact in terms of efficiency of the market disposing more and rapidly available information.

Summary

This overview of the national, regional and local portals shows the different approaches to structure, governing bodies, data release policy and licensing. There has been no prevailing model developed yet: data portals range from very small to

very big initiatives, with many or few sets of data, very valuable or very trivial applications, initiated with support of a Prime Minister or only by a small team in an agency. Therefore, it is very difficult to compare the data catalogues under study. Moreover, the first data portal – the data.gov in the US – has been operating for only two and a half years, hence its long-term effects cannot yet be captured.

Based on web survey data, it is difficult to establish on which hierarchy level the Open Data Portal should be created – whether government, agency or project team – in order to ensure its effective and efficient running. The big initiatives have benefited from a large budget and salient high-level political support. Still, these impressive initial endeavours are more susceptible to policy change and budget cuts than smaller, less expensive portals.

Furthermore, the variety of additional functionality features proves that the data catalogues are perceived as a part of a wider strategy of raising awareness of open data re-use and community-building activities and the social media channels use is also an important part of this engagement strategy.

As far as the number of datasets is concerned, this method of establishing the importance and success of the portal is in fact erroneous. First of all, the portals do not apply the same methodology of counting the data sets. Second, which is much more important, the data sets do not have equal value (both for the public and for the market). As most of the portals apply the low-hanging fruit strategy (releasing what is already public and easy to publish), many data sets are not downloaded very often and are much less re-used. Third, it is very rare for the portals to apply the five star linked data system¹⁰. This five star system is as follows:

- ★make stuff available on the web (whatever the format)
- ★★make it available as structured data (e.g., excel instead of image scan of a table)
- ★★★non-proprietary format (e.g. csv instead of excel)
- ★★★★use URLs to identify things, so that people can point at stuff
- ★★★★★link data to other people’s data to provide context.

The fourth and fifth levels of this ranking are rarely achieved and even non-proprietary formats (third level) are not yet a standard as the portals rarely have a sufficiently large budget or are not capable of putting sufficient pressure on data holders to change data sets in proprietary formats (such as, for example, xls) into a non-proprietary (csv).

The most useful features enabling browsing through the cornucopia of datasets, are ranking and rating features, collaborative tagging and statistics related to downloads. What is also useful to both portal owners and re-users is to *suggest a datasets* option enabling an understanding of what kinds of datasets are actually sought.

¹⁰ <http://www.w3.org/DesignIssues/LinkedData.html>

Most of the portals do not require users to register to use the portal (primarily to ensure there are as few barriers as possible to use) but this does make development of performance metrics and usage or download levels more difficult. The most popular licensing strategy is the national creative common policy applications (this applies to most portals in Europe apart from UK and France). There are also portals that do not apply any license to the use of their datasets (data.gov, Utah portal) and those with bespoke licenses (data.gov.uk, London DataStore, Rennes). Finally, not all the portals are compliant with the web accessibility standards and, if they are, the standards are implemented on different levels of Web Content Accessibility Guidelines (WCAG) requirements.

To measure the take-up of the portal, the data.gov administrators use rather simple metrics of page views, unique visitors, visits and downloads. The existing data do not enable a robust analysis but show that the take-up numbers are somewhat linked to the number of datasets (high take-up on data.gov, data.gov.uk or Basque Country and the Piedmont region) but also associated with the data quality or usefulness of datasets (very high take-up of London Data Store which has only 240 datasets but a lot of applications based on its traffic data). We should also take into account the fact that cities releasing traffic data will drive much more of the public interest compared with regional data that are much less related to the everyday life of citizens.

Finally, take-up by developers can be measured only by the number of the apps reported on the website, since the portals rarely require registration (Rennes) or notification of the portal in case of re-use (New York). Moreover, not all app developers submit their apps to app registers on the websites and not all portals have an app register or showcase of app and visualizations. Other indicators may be the success of the app contest organized by the portal. However, in many cases the app contest precedes or coincides with the development of the portal (New York, data.gov, Australia, Helsinki Region).

In general, data contests proved to be the main driver of data re-use and generate an average of 40 apps with average total prizes of about 50,000 Euros. These numbers show clearly the high return of investment of open data, especially when compared to existing procurement of government web services.

To sum up, with regard to the success factors for portal development the presence of high-level political commitment is very important. This high-level commitment is enhanced in some countries but local agendas that utilise data access. For example the transparency agenda in Anglo-Saxon countries and innovation push in the Scandinavia were cited as important drivers for development of open portals. US and UK portals have inspired other governments to create open data portals. The study team also has to underline the importance of open source and open data movement citizens' initiatives that were also a significant stimulus for government projects. The PSI re-use directive has not proven to be an influential driver for the portals under study. Nevertheless, the interviews showed that there are expectations that the

future revision of the directive will put more pressure on European governments to release more datasets and to reinforce the initiatives already carried out. Furthermore, app contests often triggered the national portal initiatives. Last but not least, the salient factor was the advent of web.2.0, mobile apps and data visualization tools.

Several interviewees thought that the biggest barrier to further data portal development will be the budget cuts due to financial crisis and the financing of the data.gov initiatives' projects. The lack of civil servants hampers rapid release of further datasets. What is more, the PSBs' data quality and lack of a budget to clean the data also have a negative impact on the data re-use (although this does leave space for the commercial services). The lack of an analysis of the socio-economic benefits of different data types makes it difficult to prioritize the data sets to be released and advocate further data release. As a final point, the impact so far is only measured by apps and traffic. The risk is that an open data portal appears to be a useful tool for developers, while the real beneficiaries are end-users (citizens and companies).

6 Conclusions – key success factors of open data initiatives

Open data portals have moved over the course of two years from the pioneering stage to the mainstream: they can be considered today as a recognized flagship initiative of government technology policy. More and more governments are creating portals: at the recent Digital Agenda Assembly, EC Commissioner Kroes announced the EU Open Data Portal.

The findings show that in general, the momentum for Open Data Portals is growing. More initiatives are launched, more and more datasets are being published, with an increasing focus on high-value data and improving the quality of data. Developers show continuous and growing interest in downloading data and developing apps, also through competitions, and in some cases develop added-value services which are highly used by citizens, as well as (more rarely) generate revenues.

While it is still too early to assess properly the overall impact of these Open Data portals, it is fair to say that the available evidence suggest that its short term direct impact is limited. This is nothing new expected: many years of technology policy analysis indicate that impact is visible in the long-term and comes from the combination of technological and organizational innovation.

Based on the study findings, we rather identify a set of *success factors* about “how” Open Data portals can be implemented.

Use of open data for solving problems and pursuing higher-level policy priorities

It might seem self-evident, but open data are a means, not a goal. Only exceptionally does transparency prove to be a sufficient policy driver for sustainable commitment. Successful initiatives, instead, pursue the goals of improving public services and promoting web-based innovation, and thereby covering the policy priorities of good governance, service delivery and innovation.

Not all datasets are equally important: a dataset is not useful in itself, but because it helps to solve a problem. As an approach, it is important to start from low-hanging fruits and from high-value datasets, such as transport data, which address the concrete needs of citizens, re-users and public administration. Many successful initiatives launch *ad hoc* awareness raising initiatives that concentrate on specific issues, such as healthcamp.org or transportationcamp.org (that focus on health and transportation issues).

It is also important to reach out to final users, beyond the “open data community”. Open Data Portals are at risk of self-referencing by focussing exclusively on developers while forgetting the needs of the end-users. As Prof. Nigel Shadbolt points out referring to the UK Open Data portal, “we need to increase its usability

and utility to a much wider audience.”¹¹ Even in the US, 65% of citizens are not aware of open data initiatives, and when asked, they largely prefer to explore and visualize data online rather than downloading excel sheets: portals, however, currently focus on publishing data as xls or csv. In short, the answer to the long-standing question: “should government visualize?” is probably yes, contrarily to what open data advocates state.¹²

Engagement of both the grassroots developers community and top policy level

In order to solve problems through open data, governments are learning to engage with independent developers and make use of their skills. Governments start to realize that the immense and diverse needs of citizens cannot be met by services invented and managed in-house. There is a large scope to reach out and involve the community intelligence, with specific initiatives such as contests, challenges, and open innovation services. The US government has launched challenge.gov, a platform to make use of collective intelligence to solve real problems; there are now dedicated providers for open innovation in public services, such as simpl.co. But beyond the tools, a general culture of openness is required to successfully involve developers, who care for personal gratification and fulfilment, more than funding.

At the same time, successful initiatives require top political commitment: Open Data need time to deliver impact, and a critical mass of published high-value data are needed before any difference is visible. Opening high-value data requires political strength and stable, sustained investment (although not necessarily high).

Follow-up and animation, well beyond simple data publication

Just as for many other technology policy issues, “build it and they will come” is not an effective approach. In order to ensure impact, a momentum needs to be created and sustained. Successful initiatives carefully design their websites and revisit the design after publication; consult with the public on the most needed data; provide examples and stories about data re-use; and organize contests for applications. Some datasets are just too complicated to be reused: the UK COINS database on public spending proved to be too obscure and complex to be genuinely useful. A continuous effort to improve both the raw data quality and the presentation of the data is needed. Furthermore, continuous technological evolution calls for more sophisticated solutions such as Linked data, real-time data, and open API: however, the cost-effectiveness of these solutions is yet to be proven, as they require more substantial investment.

Small start, internal growth and coordination with other initiatives

Especially in this period of public spending restrictions, we should closely look at examples of portals that are able to deliver quick results while requiring small investment. Open data requires a long-term commitment to generate change, and the more investment is required, the more it is necessary to “prove the economic

¹¹ <http://www.guardian.co.uk/news/datablog/2011/jan/21/data-gov-nigel-shadbolt-government>

¹² <http://sunlightlabs.com/blog/2009/should-datagov-visualize-probably-not/>

case” of open data, which is difficult and premature. It is necessary to join forces and funding in order to achieve critical mass: there is the risk that Open Data Portals become a “me-too” policy approach. As a positive example, the Basque Country did not organize an independent competition for Open Data applications, but instead co-funded the Spanish and the European competitions. At the technological layer of interoperability, there is a need for stronger and more systematic coordination between Open Data initiatives, which are currently fragmented and not interoperable: paradoxically, if not coordinated, open data could create barriers to the “apps single market”.

The “start small” approach also allows the creation and cultivation of internal competence in government with regard to Open Data management. Successful initiatives are managed largely in-house both at the technological and managerial level. Because Open Data Portals are much more than a technological solution, and they are continuously and fast evolving, they call for strong internal thought leadership and institutional learning. In short, Open Data Portals cannot be completely outsourced.

Open Data Portals are a *cost-effective tool for cultural change*, of government, developers and citizens. It delivers medium-term impact by acting on specific incentives (e.g. transparency by default) and delivering short-term benefits (e.g. better apps for public services). Government start to understand what it means to open up by default to external input and to deliver, as a standard rule rather than on request, information in re-usable format because other people will make the best use of it. Developers understand that they can concretely make a difference in the quality of life of citizens as well as launch entrepreneurial ventures. Citizens learn to scrutinize government and take more responsibility for the co-delivery of services. In conclusion, Open Data Portals are useful to kick-start the development of a “Public System of Innovation”: they are the beginning, not the end.

7 Annexes

7.1 Annex 1: Data.gov portals case studies

7.1.1 United Kingdom – DATA.GOV.UK

Background and history

The catalyst for data.gov.uk and the transparency agenda in the UK was probably Ed Mayo and Tom Steinberg's *Power of Information Task Force Report*¹³, published in 2007. This raised government thinking about the value and role of information in a world in which most people regularly use the Internet.

The creation of the Power of Information Taskforce in 2008 was part of a series of positive responses by Government to the 2007 report. The taskforce brought together representatives from government, industry and the third sector to develop the agenda set out in the report. All taskforce members contributed to it in a personal capacity rather than on behalf of their respective organizations and this independence of approach was reflected in the conclusions reached in the *Power of Information Taskforce Report*¹⁴ published in February 2009. Data.gov.uk interviewees suggested this report was the primary foundation stone for the data.gov portal.

Government published its response to the Taskforce Report – *Digital engagement: Update on power of information*¹⁵ – in May 2009. Highlights of the government response included the following four recommendations:

- Accept the taskforce's recommendation that geospatial data produced by the Ordnance Survey should be opened up and made more widely available in the economy and society (Recommendation 7).
- Appoint the Office of Public Sector Information (OPSI), which is part of the National Archives¹⁶, to develop a new licence model to 'take the licensing of government content to the next level'. Users will not need to register and apply for a licence (Recommendation 8).
- Make public information available at marginal cost, which in practice means for free online (Recommendation 10).

¹³ <http://webarchive.nationalarchives.gov.uk/20100413152047/http://poit.cabinetoffice.gov.uk/poit/>

¹⁴ http://webarchive.nationalarchives.gov.uk/20100413152047/http://www.cabinetoffice.gov.uk/reports/power_of_information.aspx

¹⁵ <http://www.epsiplus.net/content/download/23528/311309/version/1/file/Digital+Enlargement+Update+on+Power+of+Information.pdf>

¹⁶ <http://www.nationalarchives.gov.uk/>

- Create a data service through the Office of Public Sector Information and the Chief Information Officers Council, akin to proposed data.gov.uk, which would expose government's data feeds in a well ordered and useful way. "It will provide a focal point for development using Government information. A basic service will go live in the next three months, with iterative improvements after that". (Recommendation 14).

Activities to develop a portal began in earnest following this report. In June 2009 Sir Tim Berners-Lee, inventor of the world wide web, together with Professor Nigel Shadbolt from Southampton University, an expert on open data, became involved with the development of the project. When it was first developed, the aim of the site was to enable users to hold public service providers to account for their activities. The primary audience during development was thought to be small businesses and campaign organizations. Subsequently, greater emphasis has been placed on the needs of 'geeks' and citizens.

The UK portal was opened on September 30, 2009 and had 2,300 data sets. The US data.gov portal was opened four months previously, but only had 47 data sets. Development of the UK portal was largely undertaken without reference to other countries, since no other European portal had been developed at the time.

Feedback was sought on the UK portal after it was opened. It was officially launched by Sir Tim Berners-Lee on 21st January 2010 (in September 2010, the portal was enhanced). At the launch MP Stephen Timms, minister for Digital Britain, noted that the government had "agreed to an open licence for free reuse of the data. We have agreed to release data on weather, transport and public finance. We have started public consultation to open up data from the Ordnance Survey"¹⁷. Mr. Timms noted that the initiative would:

- Open up government;
- Improve public services;
- Increase economic growth and social value.

Following the UK general election on May 6, 2010 a new government was elected. As a result, the open data agenda took on a much greater focus on transparency. The new government set out the need for greater transparency across its operations to enable the public to hold public bodies and politicians to account. This included commitments relating to public expenditure that were intended to help achieve better value for money. The transparency agenda as a whole is being overseen by the Public Sector Transparency Board, which was established in the Cabinet Office¹⁸. The Board is chaired by Mr. Francis Maude, the Minister for the Cabinet Office. The

¹⁷ <http://www.guardian.co.uk/technology/blog/2010/jan/21/timbernerslee-government-data>

¹⁸ http://www.ogc.gov.uk/policy_and_standards_framework_transparency.asp

Board's role includes providing support to departments as they deliver the government's transparency commitments.

The new transparency agenda was based on a letter published on May 31, 2010 from Prime Minister David Cameron and sent to government departments on plans to open up government data. The letter set the tone for the new agenda. Key commitments included:

- Her Majesty's Treasury's COINS database of public spending data to be published online in June 2010.
- Names, grades, job titles and annual pay rates for most senior civil servants with salaries above £150,000 to be published in June 2010.
- All new central government ICT contracts to be published online from July 2010.
- All new central government tender documents for contracts over £10,000 to be published on a single website from September 2010, with this information to be made available to the public free of charge.
- New items of local government spending over £500 to be published on a council-by-council basis from January 2011.
- All new central government contracts to be published in full from January 2011.
- Crime data to be published at a level that allows members of the public to see what is happening on their streets from January 2011.

These requirements were to be applied to all central government departments including their agencies, all non-departmental public bodies, National Health Service bodies and trading funds. Departments are responsible for meeting these requirements. Separate requirements are in place for the wider public sector such as local authorities. The Cabinet Office's Efficiency and Reform Group (ERG) has produced a range of practical instructions and supporting resources to assist departments to meet these requirements as part of a consistent approach¹⁹.

In June 2010, a team of experts was invited to form a Transparency Board. The aim of the Transparency Board is to drive forward the government's transparency agenda, making it a core part of all government business. In addition, it is responsible for establishing public data principles across the public sector and making datasets available for potential development and re-use.

The Transparency Board is chaired by Mr Francis Maude. The other members of the Transparency Board are Tim Berners-Lee, Nigel Shadbolt, Tom Steinberg, founder of mySociety (mysociety.org), Andrew Stott, former Director of Digital Engagement at the Cabinet Office, and Rufus Pollock from Cambridge University, an economist who

¹⁹ http://www.ogc.gov.uk/policy_and_standards_framework_transparency.asp

was one of the founders of the Open Knowledge Foundation. The board has given input to the creation of the data.gov.uk portal. Several members of the board also contributed to earlier development work.

The costs of developing the portal are yet to be released to Parliament, and so public disclosure of this information at this stage is not possible. Nevertheless, the costs are known to be relatively small; the core Cabinet Office team developing the portal consists of six full-time developers. In comparison, it was suggested during interviews that the US portal has had a budget of US\$8.5 million.

The focus of the portal is on making data available. Cabinet Office personnel are not developing applications to enable access or data visualization. The UK National Archives are developing a number of interface and visualization tools to enable easier access to information. For example, the National Archives website has user-friendly access to a catalogue of all its datasets and direct access to information such as army records, census data, divorce records, legislation and regulations.

Scope

The data.gov.uk site states that the portal had 6,200 datasets on April 22, 2011. The portal provides access to COINS – Her Majesty's Treasury's database of public spending. Other leading government departments that are releasing information include the Department of Health (1,059 datasets), Communities and Local Government (783), UK Statistics Authority (715) and Environment, Food and Rural Affairs (406). The most popular tags for data reflect the input of the government departments that have provided the most data. The most popular tag is 'health' (2,294). Other tags include 'care' (1,628), 'transparency' (1,530), 'communities' (1,318) and 'child' (1,208).

Interviewees noted that defining what constitutes a dataset is open to interpretation. Some portals have been a little over-zealous with regard to counting the constituent elements of a dataset separately. For example, some portals have counted datasets provided by gender as two datasets (for male and female) rather than as a single data set distinguished by a gender variable.

A key focus of attention for data.gov.uk is the development of linked data. Linked data is data in which real-world items and activities are given Uniform Resource Identifier (URI) addresses on the web, and data about them is published in machine-readable formats at those locations. Other datasets can then point to these items by using their URIs: as a result, people using the data can find out more about a particular item without information being copied into the original dataset. The data.gov.uk portal has a page²⁰ that lists the sectors for which it is currently

²⁰ <http://data.gov.uk/linked-data>

publishing linked data and some additional resources that will help users to use the data.

Interviews established that, when data.gov.uk was created, its primary focus was to provide data as quickly as possible. Less emphasis was placed on the cleanliness of the data. The majority of government data was thus 'fit for the purpose' for which it was collected. However, this purpose might not always align with the needs of others users who want to re-use the data. Interviewees suggested that sometimes it is necessary to emphasize that datasets are only intended for their original purpose: hence, difficulties in re-use of the data for other purposes might be encountered.

It was emphasized that the basic level of data accuracy is the responsibility of the publishing organization. The government is unlikely to use additional resources to ensure data is enhanced in some way to make it more usable.

It was suggested that the LinkedGov²¹ initiative has a focus on obtaining data, cleaning it and re-using it. The goals of LinkedGov are:

- Promoting economic growth and innovation through greater access to immediately usable public data;
- Improving government transparency for citizens, researchers and the media;
- Promoting efficiencies through effective use of data in the public sector;
- Supporting public sector workers in their own goals and decision-making processes by providing them with greater access to information about their own organizations and similar organizations across the same sector;
- Encouraging all sections of government and the public sector to publish high quality, standardized data.

Pricing and availability

All data provided by data.gov.uk is free. With regard to wider government data access, the government's response to the Power of Information Taskforce Report was an agreement that public information should be available at marginal cost: in practice, this means for free online.

The Transparency Board published a set of transparency principles that underpin demand and access to data, these include:

- Public data policy and practice will be clearly driven by members of the public and businesses who want and use the data, including what data is released when and in what form;
- Public data will be published in reusable, machine- readable form;

²¹ <http://linkedgov.org>

- Public data will be released under the same open licence that enables free re-use, including commercial re-use;
- Public data will be available and easy to find through a single easy-to-use online access point (data.gov.uk);
- Public data will be published using open standards, and following relevant recommendations of the World Wide Web Consortium (w3.org);
- Public data will be freely available to use in any lawful way – public bodies should actively encourage the re-use of their public data and should maintain and publish inventories of their data.

The UK government pioneered online licensing with the launch of the Click-Use Licence in 2001. This licensing model was highly successful, and more than 20,000 licences were issued. This initiative had the effect of opening up PSI re-use to a global audience. To support the release of datasets under data.gov.uk and the UK government's transparency agenda it was felt that licensing should be taken to yet another level of openness and transparency.

The UK Government Licensing Framework²² provides a policy and legal overview for the licensing of PSI, both in central government and the wider public sector. The framework has been created to meet the needs of:

- The public, including community groups and social organizations;
- The information re-user community, both in the public and private sector;
- The public data developer community.

The UK was able to draw on the work of public sector colleagues in Australia and New Zealand in terms of open data. Both countries had launched policies designed to open up government and make PSI more readily available for re-use. They did this through the adoption of Creative Commons model licences. The UK, however, decided to develop a new licence – the Open Government Licence. The main reason for this was that none of the existing Creative Commons licences extended to the licensing of works protected by the database right²³.

The Open Government Licence is at the heart of the UK Government Licensing Framework. The challenge in developing the Open Government Licence was to create a licence that would be:

- Interoperable with other internationally recognized models such as Creative Commons and Open Data Commons (opendatacommons.org);
- Feasible to overcome copyright and database rights;

²² The National Archives. 2010. The United Kingdom Report on the Re-use of Public Sector Information: Unlocking PSI potential. <http://www.nationalarchives.gov.uk/documents/psi-report.pdf>

²³ Databases enjoy separate protection in Europe.

- More enabling (as a result of avoiding the need for re-users to register and apply for a licence);
- Machine-readable;
- Simple and legally robust;
- Sufficiently flexible so that it could be adopted across the public sector by local government and health bodies.

The UK Government Licensing Framework, incorporating the Open Government Licence, was launched on September 30, 2010. It has been welcomed as an important innovation by people within both government and the re-use community.

Drivers and obstacles

Interviewees suggested that the May 31, 2010 letter from the Prime Minister acted as a catalyst and incentive to persuading and cajoling government departments and local government organizations to release data. However, they also acknowledged that making data available in an economic downturn could be politically problematic. Reluctance to move in this direction has occasionally been encountered.

The data.gov.uk portal has a wiki, blogs and forum to increase awareness and support for public officials, individuals and organizations requiring and re-using information. It was noted earlier that the Transparency Board had published transparency principles: the most notable among these is that that public data policy and practice will be clearly driven by the public and businesses who want and use the data, including what data is released, when, and in what form. A number of the forum posts on the portal appear to be supportive in assisting users to obtain data that was not previously available.

A further catalyst to enhance the release of data could turn out to be the Protection of Freedoms Bill²⁴ that was first placed before the British House of Commons in February 2011. The bill proposes to modify the existing Freedom of Information Act 2000, by extending the scope of the Act to request the release of data as well as information. Section 45 of the Act could be amended to ensure that any data provided by Freedom of Information requests should be placed on data.gov.uk, with the possibility of providing regular data updates. This would represent a significant change from current UK information practices.

Impact

Interviewees suggested that development of the portal is still at an early stage. Use should increase in the future as it is developed further. There are considerable fluctuations in use as new datasets are added and there is media coverage of the

²⁴ <http://services.parliament.uk/bills/2010-11/protectionoffreedoms/documents.html>

site. For example, a crime mapping application received 18 million hits an hour on the day of its launch. It continued to receive more than 400 million hits in its first three months of existence. data.gov has about 80,000 unique users per month.

On April 22, 2011 the portal recorded having more than 140 applications that had been developed by users. Reviewers gave 40 of the applications a five-star rating (the highest possible). The five-star applications include those about how to find post boxes, care homes, local air quality, road accidents, high-crime areas, landfill sites and pharmacies.

A number of events, app development competitions and data hacks have been developed to support the UK government goals of the open data and transparency agendas.

The Open Knowledge Foundation organized a two-day international workshop on November 18/19, 2010 for people interested in open government which took place at the University of London Union. The event was supported by data.gov.uk and Francis Maude, the Minister for the Cabinet Office and chairman of the Transparency Board, made a major speech on the topic of the open data during the event.

LinkedGov held its first hack camp on the April 9/10, 2011 in Shoreditch (in east London). Over the weekend, 19 hacks were submitted. Winners included an application that permits British citizens abroad to call a single UK number from anywhere in the world (the service will read them the Foreign and Commonwealth Office travel advice for that country based on the CallerID of the country they are calling from²⁵). Another application analyses the Business Link data feed and republishes relevant announcements about tender opportunities for SMEs: the information is provided by Twitter and by RSS updates²⁶.

Lessons and reflections

Interviewees felt they had made the right choice in launching the data.gov.uk portal quickly with as much data as possible (rather than spending time cleaning the data and delaying the launch until the data and site were pristine). They felt that a lot of value could be extracted from the data that was available at an early stage. Interviewees also highlighted that the idea of linking the data with a non-transactional licence was important.

Data visualization was thought to be an important factor in enabling access for citizens. The data.gov team acknowledged that, if more work had been undertaken to enable data visualization, greater value to citizens and a higher profile for the site might have been achieved. The team is now looking for cost-effective solutions that will

²⁵ <http://linkedgov.hackcamp.org.uk/hacks/ukphonehome>

²⁶ <http://linkedgov.hackcamp.org.uk/hacks/at-ukgovcontracts>

provide automatic generic visualization of any data. However, team members are also aware that there are already people within government active in this area and they do not want to undermine what is already being done by others. Therefore they have adopted an inclusive approach within government and are engaging with existing teams active in this area.

7.1.2 Denmark – DIGITALISER.DK

Background and rationale

Digitaliser.dk portal is a part of the Danish Open Data Innovation Strategy initiative (ODIS) that aims at facilitating access to raw public data for the private sector in order to develop innovative digital products and services, useful analyses, data visualization and data journalism. Both the initiative and the portal are led by the National IT and Telecom Agency (NITA), which is part of the Ministry of Science, Technology and Innovation. NITA was created in 2006 with the aim of ensuring that Denmark has the best conditions for growth as a knowledge society. The agency's vision is to create the groundwork for digitalization of Denmark. It employs around 300 people.

The ODIS initiative focuses on encouraging public sector bodies in opening up data by demonstrating potential PSI data and collaborating with data re-users to stimulate demand. For example, NITA works on developing a uniform framework and common approach to accession and re-use of PSI. It also prepares guidelines and software tools for publishing and re-using data.

Even though Denmark complies with the PSI Directive and the Danish PSI act, its legal acts do not oblige public agencies to release data. Due to the lack of a legal framework, the activities undertaken so far have mostly been focused on awareness-building strategies. The ODIS strategy makes use of the collaborative platform Digitaliser.dk portal, which was also created by NITA. Its functionalities include a community forum and a listing of open data catalogues in Denmark hosted on this portal.

The Digitaliser.dk portal was launched in 2009²⁷. It is a collaborative, social networking platform targeted at PSBs, civil servants, ICT industry and citizens who want to discuss and enhance the creation of digital services and knowledge society.

²⁷ Digitaliser.dk

The portal facilitates collaboration in making use of web2.0 tools. It is also a place for public discussion on data digitalization processes. Even if the website is administered by a public institution, the content is co-created by its community of users. What is more, Digitaliser.dk was released in full as open source software and therefore is available for re-use. After summer 2009, the portal has offered a data catalogue in a form of web directory.

As NITA has no legal power to compel agencies to publish data, it has developed a set of incentives and works on emphasizing awareness-building strategies. In summer 2010, the agency established a small unit of “Data Hunters”. The unit is responsible for locating interesting public data, helping public authorities to publish their data, and assisting PSI re-users to request data from public authorities. NITA also showcases the participating actors. It organizes different meetings (e.g. datacamps) to bring together public servants and members of the re-use community to show how much can be done with datasets that are already available. The agency has published a business case in the beginning of 2011 (based on interviews with industry representatives) that points to the economic potential of PSI for selected industries in Denmark. It has also published simple and easy-to-use technical, practical and legal guidelines on how to publish public data to help public authorities initiate open data projects.

The data catalogue idea was inspired by apps contests in the US (Apps for democracy) and the UK (Show us the better way) together with the data release by both countries through dedicated portals (data.gov and data.gov.uk). Nevertheless, according to the interviewee, the initiative rationale was slightly different. Whereas data.gov founders underlined the need for democracy, transparency and the enhancement of trust in government, and the UK advocated the need for efficiency of public services, the Danish agency focused primarily on the stimulation of innovation and digital markets and only secondly on obtaining less expensive public services through co-delivery. This rationale was pinpointed by a report by the Gartner Group commissioned by NITA that estimates that business re-use of public data in Denmark could be worth at least DKK 600 million (more than EUR 80 million) a year.

The portal

Since the open data policy has not been harmonized in Denmark, and ministries or agencies decide which strategy they apply regarding the data release, the data catalogue remains only a list of available data that are located on ministries and agencies' websites (i.e. data directory). Moreover, in Denmark, it was decided to retain the decentralized maintenance of datasets and not to develop a central data.gov portal.

Therefore data is located on the public authorities' own servers or chosen host servers. To list available public data, a consultancy was commissioned to screen the websites of all Danish public bodies on a national level. Nevertheless, according to the interviewee, the 900 datasets listed on the website are only a subset of data which has already been digitalized and can be made available if a decision to open up data is made at a higher level of government.

The open data and PSI re-use policy in Denmark plays a coordination and awareness-raising role that has the form of a fully-fledged top-down political strategy.

History and budget

The first version of web directory, as a subpage of digitaliser.dk portal, was launched in mid 2009. The revised version was released in December 2010. As the collaborative portal already existed, it is difficult to estimate the portal budget for deployment and maintenance of the data catalogue. The project encompassed the development of four individual catalogues (i.e. data, open source software, XML schema, and standards), hence the specific cost of the data catalogue cannot be defined. Nonetheless, a budget of DKK 500,000 (EUR 67,000) was allocated for the creation of all four catalogues in 2009, which were subsequently added to the existing platform, i.e., general platform functionality and design had already been in place.

Resources and functionalities

So far, the web repository lists 900 datasets. The directory links to data in machine-readable formats (such as GeoJSON, HTML, JSON, KML, XML, xls, html and pdfs). The data is available either for free or at marginal cost depending on a dataset. The list consists mainly of central government datasets, but local governments are also being encouraged to participate (e.g., there are already some local municipalities datasets). As the data catalogue is part of the collaborative platform, it offers much more than access to PSI data – it also has a community forum, resources on digital

services and knowledge society, and open source software catalogue.

App competition

The Open Data Innovation Strategy (*Offentlige Data I Spil*) competition was launched in October 2009. Three winners were announced in February 2010. The authors of the successful applications authors were given an award of DKK 100,000 aimed at developing their application proposal. One of the successful applications – <http://www.husetsweb.dk/> – based on the Danish Building and Housing Register (BBR) was coupled with Totalkredit and Rockwool (a provider of insulating materials). The application offers a user-friendly tool for homeowners to assess and optimize their energy consumption. It prompted the creation of the Husets Web ApS company. Another application, Web of the House, generates a profit by charging municipalities for the use of the calculation tool website as part of local energy conservation campaigns.

Another example of a service developed thanks to the use of PSI data is the mobile application, DMI Byvejret (<http://byvejret.dmi.woerk.dk/>). The application uses the meteorological data from the DMI and smartphone GPS data to locate the user. Digitaliser.dk offers another solution by matching DMI weather data and geo-location data from the Geoservicen portal (<http://geo.oiorest.dk/>): developers and citizens can receive official geocoded data via a web API. The mobile web application (<http://vejret.oiorest.dk/>) establishes the phone owner's location with the help of government address database and searches for the most appropriate weather forecast to offer them based on their physical location.

Critical success factors

In Denmark, the major hindrance to the open data strategy development is the lack of political commitment on the national level, i.e., the lack of a national policy regarding access to and re-use of PSI. According to the interviewee, the government's commitment to data release does not imply a political decision to open up all the already-digitalized public data. Rather it implies a need for the development of a unified legal framework on a national level.

In response to the public consultation on the evaluation of the PSI Directive (concluded in March 2011), a Danish government position paper called for an in-depth study of barriers and challenges as well as socio-economic benefits of PSI re-

use on a European level targeted at specific sectors.

Data holders in Denmark very often do not realize the value of PSI and the importance of the role that they could play in stimulating its re-use and, hence, the innovative services that can be created based on this data. On the one hand, this is because of the absence of a more widely known analysis on the economic and societal impact of PSI re-use (hence the Gartner report commissioned by the agency, and plans to develop a business case for strategic sectors). On the other hand, the benefits are far more likely to be applied to end-users and re-users than public bodies directly. This is due to the public sector bodies' mindset: i.e., civil servants may see data re-use as an additional task which increases their workload. The modification of civil servants' approaches to the re-use is at the centre of the NITA strategy (for example, through the organization of data camps which bring together administration representatives and app developers).

Moreover, at present in Denmark, there are no internal incentives used by the PSBs or the state and no allocated budget for data release. Hence, the process has been very much slowed down and is mostly driven by engaged individuals that take on the initiative in certain PSBs. There is also no strong community of re-users who could advocate for more rapid release.

Nevertheless, there are several drivers that are moving forward NITA's PSI re-use initiative. Firstly, it is the overall, the digitalization of society, advent of Web 2.0 and the availability of interesting examples of public data use that is driving the demand for this data forward in Denmark.

Secondly, the fact that the Danish data release process is driven extremely bottom-up, due to local enthusiasts in public bodies, may be also considered as a success factor. Thanks to this engagement, even though there is no legal obligation to release data many of the datasets are being released and are increasing the supply of public data.

Thirdly, the agency has encouraged public bodies to adopt an agile approach that starts with "low-hanging fruit", i.e. datasets which can be easily published and often at no cost. Therefore, they would move on to publishing datasets that involve more workload (and therefore costs) later on, acting in response to the actual demand of

re-users and citizens accessing data. What is more, as the overall level of digitalization of the public data in Denmark is very high, data release in the digital format often does not accrue high costs.

Finally, the economic crisis, globalization and crisis of the welfare state, make decision-makers much more sensitive to the idea of innovative, efficient and co-produced services that might help them to face the upcoming societal challenges.

In the interviewee's view, the PSI re-use directive cannot be considered as an important driver in Denmark owing to its generality. The EC is, however, seen as an important actor in the PSI re-use field because it aims to showcase the best practices and facilitate knowledge exchange. In the public consultation on the evaluation of the PSI Directive, the Danish government's official standpoint which was stated by the Ministry of Science, Technology and Innovation was to alter article 6 concerning the charging models by setting marginal costs as an upper limit (i.e. the Danish law already sets the maximum charging model to marginal costs). This standpoint focuses on the assumption that high charges are a barrier to the re-use of PSI and can hamper the innovation potential of SMEs. Once more, this underlines the Danish data release strategy focus on innovation.

According to the *Uncharted Waters, The State of Open Data in Europe* report authors, based on an interview with a Danish government representative, as a result of the transposition of the EU INSPIRE directive, the Danish National Survey and Cadaster has already put large amounts of data online (CSC, 2011, p.14).

Impact

Until now, the only feedback, which may help to analyze the impact of this activity in Denmark, was gathered through face-to-face meetings with re-users and users and their comments on the platform. Unfortunately, the platform's functionalities do not permit tracking of the usage of the data catalogue. NITA is currently working on a reconfiguration of its statistics-gathering tool so as to be able to analyze quantitative traffic data.

Given the short time that elapsed between the creation of the data catalogue and the agency's lead of a broader strategy aimed at promoting data release and re-use, it is very hard to establish what has been the impact of the web directory itself.

The Danish open data policy is very much targeted at the middle layer of the three important agents mentioned by Tim Berners-Lee (“It has to start at the top, it has to start in the middle and it has to start at the bottom”²⁸), i.e. through professional public administrators. The Danes' overall strategy of awareness-raising, community-building and putting in place a designated online space for a digital catalogue (even in its present directory form), has already:

- encouraged the PSBs to publicize their online datasets to the wider public,
- centralized the information on PSI re-use in Denmark
- encouraged the data holders to release more data.

It is difficult to establish whether digitaliser.dk has had any impact on the PSI re-use pricing as the strategy is focused on making more data available rather than changing the Danish PSBs' pricing models.

Sources:

Phone interview with Cathrine Lippert from digitaliser.dk, Special Adviser at the Danish National IT and Telecom Agency

European Public Sector Information Platform Topic Report No. 20 PSI Reuse in Denmark by Cathrine Lippert

Public Consultation on the evaluation of the PSI Directive - Danish government position paper

Schellong A., Stepanets E. (2011), *Uncharted Waters, The State of Open Data in Europe*, CSC, Public Sector Study Series

7.1.3 Spain, Basque Country – OPEN DATA EUSKADI

Background and rationale:

- building on UK and US momentum
- following Spanish implementation of PSI directive
- after change of government in the Basque country, there is a need to push transparency to enhance trust and to foster innovation in view of the crisis
- policy framework is openness to citizens.

Description:

- 1,000+ datasets from very different sources

²⁸ Citation after Hogge, 2010

- all free (marginal costs), no registration
- Creative Commons 3.0 licence allowing for commercial re-use
- Distributed repository with single point of access
- 100K budget for technological development and animation
- participation in existing competitions, such as FICOD and Open Data Challenge
- strong involvement to foster regional innovation
- forthcoming datasets: location of hospitals, voting records, job opportunities in public sector.

Drivers

- strong Basque tradition of transparency (e.g. free orto-photos) and innovation (web players)
- strong political leadership
- release of fast to reap low-hanging fruit
- strong connection with PSI re-users
- use existing open standards
- use low costs solutions and deploy large amount of data, rather than going for highly sophisticated solutions.

Impact

- high usage: 250 datasets downloads per month, 30,000 unique users in one year
- competition attracted 30 working apps, the winner was from the Basque country: this application is being further developed in the mobile environment
- apps based on PSI: infocarretera.com (traffic data), metroo.es (accident data), wikiloc (geo ortofoto), euroalert.com (tenders)
- low data on impact: no requests for registration to maximize data re-use
- no impact on pricing: all data were already public, but not easily accessible
- low revenues for apps developer, most are "side projects"
- high impact on internal public administration: government agencies are the largest re-users
- high impact on per annum interoperability: other government levels are using the data (e.g. weather widget in cities websites)
- impact on research: researchers are using climate data to carry out bioclimatic studies for rural areas, which is then useful for the tourism sector
- high impact on business productivity: the most requested dataset is the official calendar of public holidays, which is different in each municipality. This helps companies to better organize their work
- the greatest impact is overall **indirect economic impact** in terms of efficiency of the market.

Source:

Interview with Alberto Ortiz de Zarate, director of Open Euskadi Project

7.1.4 France – DATA.GOUV.FR

Data.gouv.fr portal is still in the design process. However, it is an interesting example of a top-down PSI policy, as the strongest driver for a rapid implementation of the portal comes from the political level, from the French Prime Minister.

Background and rationale

The 2008 report *‘France numérique 2012 - Plan de développement de l'économie numérique'*²⁹ by M. Eric Besson, Minister for the development of the digital sector, called for the creation of a unique e-government innovation portal for France. Only recently, however, in November 2010, did the French Council of Ministers decided to implement the data.gouv.fr portal by the end of 2011. This decision was followed by the decree n° 2011-194 of February 21, 2011. The French Prime Minister officially established 'Etalab' and entrusted it with the mission of creating a unique inter-ministerial French data.gov portal. Etalab currently has six FTEs and is part of the Secretary General of the French government (the office of the Prime Minister).

Etalab was created in order to direct and enforce the French PSI policy and to coordinate its implementation by PSBs. According to a press release by the French Prime Minister (on February 22, 2011), there is a political will to make “public sector data freely available online” and its re-use “as easy and as widespread as possible”. This means that PSI will be made available free-of-charge for re-use (whether it is of a commercial nature or not), as a general principle. However, existing charging schemes will not be overhauled and new PSI re-use charging models will still be possible, but only on a duly-justified and exceptional basis.

This new policy has resulted in the decision to create a national data.gov portal. The PSBs will be obliged to make their PSI data available on data.gouv.fr (if the PSB applies the new free-of-charge policy).

According to the interviewees, the new French PSI policy has the intention to put France in the position of a European leader in the field of PSI re-use and open data.

The underlying rationale of the portal is the stimulation of innovation, enhanced transparency of the government and its agencies, and the openness principle. What is more, since one of the objectives of the French state modernization strategy is to reduce the number of state websites from 567 to 30-60, the future data.gouv.fr portal is considered to be a milestone on this trajectory towards website reduction.

The portal

²⁹ <http://www.ladocumentationfrancaise.fr/rapports-publics/084000664/index.shtml>

The new portal will contain PSI that is newly released by central administration (or any other PSB on a regional or local level) that is willing to participate in the project and use the platform to release its data.

A first PSI data audit found that around 630,000 tabular data files are already freely available for access and re-use in France. Nevertheless, the databases of most economic value and/or potential are available on payment of licence fees.

A beta version of the portal is foreseen to be released at the beginning of December 2011.

Budget

The final budget for the portal development and maintenance was estimated at least at around 350,000€ in its first year. The Direction de l'information légale et administrative DILA (the PSB in charge of legal and administrative PSI) will act as the logistic and technical support for the implementation of the portal (it will build on support by a private contractor for the technical work). DILA will also finance the portal implementation.

DILA is a financially autonomous PSB that finances itself with revenues from legal information sales and fees for legal announcements by private sector bodies. It does not receive any state funding.

Resources and functionalities

The beta version of the French data.gouv.fr portal will contain a first collection of PSI already available free of charge. The ambition behind the portal is for it to continuously evolve to a more and more complete collection of French PSI.

The portal is envisaged as a mix of data repository and index of data with hyperlinks. The technical solutions will be chosen in a way that releases the maximum amount of PSI in a short period of time.

The data.gouv.fr portal will make raw data PSI available in an exploitable format. Metadata will also be made freely available. The proposed data formats will build on established open standards (such as XML, CSV) and the 'Référentiel Général d'Interopérabilité' (RGI in France) that gives guidelines and best practices for data held by PSBs.

A licence based on the zero cost principle developed by Etalab will apply to all PSI available on the portal. Other PSI data, to which rights to re-use are restricted (e.g. under cost recovery pricing), will be made available under *ad hoc* licences.

Although the plans foresee making the data freely available without authentication (where legally possible), an optional registration process will be proposed to

implement customized services such as 'alerts on new datasets by keywords', 'producers', and 'ask a question to a data holder'.

Apart from the data catalogue and above-mentioned functionalities, it will also contain additional features, namely a FAQ for re-users, awareness trainings for PSBs and re-users, application contests and other animation activities for the re-use community.

Etatlab is currently working on different indicators to assess the impact of the future data.gouv.fr portal (such as the number of its users, downloads and indirect traffic) that will be also made freely available on the portal.

Impact foreseen

The data.gouv.fr portal is implemented with several aims in mind: to foster innovation and to develop the digital sector and economic growth, notably in the French ICT industry. It aims at making new data available, facilitating access and re-use data and providing a reliable catalogue of already available data. Apart from the PSI-related goals, the portal authors strive to foster innovation and contribute to the development of the digital sector in France.

The portal also aims also at instigating important economies of scale for the French state administration.

Sources

Interview with Etatlab team - Séverin NAUDET, Director, Etatlab; Romain LACOMBE, Chargé de mission, Etatlab; Alexandre QUINTARD KAIGRE, Chargé de mission, Etatlab

7.2 Annex 2. Web survey

7.2.1 Web survey summary

Data catalogue	data.gov	data.gov.uk/	*digitaliser.dk	*data.suomi.fi	data.australia.gov.au	*data.govt.nz
Number of datasets	379,943 (raw and geo data sets)	6,200	800-900 links to datasets	46 links to datasets	223	561 links to databases
Data format	xml, csv/txt, kml/kmz, shapefile, rdf, other	csv and many others	geojson, html, json, kml, xml, xls, html, ods, pdf,	xls, pdf, kml, xml	xml, csv/txt, xls, kml/kmz, esri	pdf, csv, xml, html table, xls
Data source	federal government and agencies	national and agencies (some municipalities)	national and agencies (some municipalities)	national and agencies	national government	national and local government
Timeliness	depending on dataset	depending on dataset	depending on dataset	depending on dataset	depending on dataset	depending on dataset
Reuse conditions	Free (no licence)	Open Government Licence excluding the Met Office	Data holder licence	Data holder licence	Creative Commons	Creative Commons + licences of agencies
Pricing	free	mostly free	mostly free	free	free	free
Accessibility	anonymous access	anonymous access	depending on the dataset	anonymous access	anonymous access	anonymous access
Take-up	1,941,453 visits last year, 80,000 downloads per week	80,000 unique users per month.	no data available	no data available	recently released as a production version, does not collect any	click-throughs (or referrals) to data hosting websites

					statistics yet	(15% of all page visits). 5,377 page visit per month
Take-up by app developers	934 government and 236 citizen-developed apps	140 apps showcased	no data available	50 apps showcased	area on website where the apps can be submitted (only one so far)	not available
App contests	Apps for Democracy, although it is not organized by the portal	Show us a Better Way (before the portal launch)	App contest, October 2009-February 2010, 25 full ideas of apps, three awards	Apps for Democracy Finland	mash up Australia (82 entries), libraryhack	November 2010, another planned for August 2011

* web directory

Summary

Data catalogue	risp.asturias.es	opendata.euskadi.net	dadesobertes.gencat.cat	www.dati.piemonte.it/	data.london.gov.uk/datastore	data.rennes-metropole.fr	www.hri.fi	nyc.gov/html/datamine/	utah.gov/data/
Number of datasets	5	1,202	124	242	438	90	549	361	31
Data format	html, sparql/sml, sparql/json, rdf/xml, rdf/n3/ turtle	xml, csv, wms, other	tmx, zip, pdf, csv, kml-kmz, doc, xls, xml, json, rdf-xml, shp, sparql	cvs, pdf	csv, xml and others	csv, xml, jpg, dxf, shp, gfts, rss, api, ics	xls, xml, px, json, kml	xls, xml, csv, and rss	csv, kml and xls
Data source	Regional	Regional	Regional	Regional	municipal	municipal	municipal	municipal data	regional data
Timeliness	depending on dataset	depending on dataset	depending on dataset	depending on dataset	depending on dataset	depending on dataset	depending on dataset	depending on dataset	depending on dataset
Reuse conditions	Creative Commons	Creative Commons or free use	per dataset, mostly Creative Commons	Creative Commons	still working on a licence with the Open Knowledge Foundation	Licence de réutilisation des données publiques - Rennes Métropole en accès libre (prepared	HRI data licence	re-users should notify the City and include a disclaimer	no information on re-use conditions

						with APIE)			
Pricing	free	free	free	free	free	free	free	free	free
Accessibility	anonymous access	anonymous access	anonymous access	anonymous access	anonymous access	user registration	anonymous access	anonymous access	anonymous access
Take-up	no data available from portal team	February-April 2011 25,362 users / 33,682 visits / 157,520 pages visited	November 2010 - March 2011 Pages viewed: 83,689, Visits: 29,900	around 200 per day	0.5 million visits and 2.4 million page views in the first year. File downloads: 79,000 downloads of CSV files, 7GB of data	no data available from portal team	no data available yet	2011: visits 22,328, unique visitors 9,326 page views 51,261	no data available
Take-up by app developers	no data available	no data available	70 inquiries from users, 9 apps featured on the website	8 apps showcased	62 inspirational uses showcased	43 apps showcased, 2.6 M request for API	no data available yet	not available	no data available
App contests	no contest organized yet	no contest organized	no contest organized	plans to open apps.piemonte.it portal	no contest organized	Rennes Métropole en accès libre	Apps for Democracy – Finland	NYC BigApps 2.0 Competition	none

						43 entries	before portal launch		
--	--	--	--	--	--	------------	-------------------------	--	--

7.2.2 Web survey results

Data.gov	US
General Services Administration Agency	
<p>Launch date</p> <p>May 2009</p>	
<p>Short description</p> <p>The portal aims to increase public access to high value, machine-readable datasets and expand creative use of it by encouraging innovative ideas. It wants to make government more transparent and is committed to creating an unprecedented level of openness in government. It is one of the priorities of the Open Government Initiative of President Obama's administration. The data are generated by the Executive Branch of the Federal Government.</p>	

Data

In the beginning of April 2011, the portal made available 379,943 raw and geospatial datasets (including 3,263 raw datasets). The portal splits the catalogue into geodata and a raw data catalogue. The data are available in machine-readable non-proprietary formats - xml, csv/txt, kml/kmz, shapefile, rdf and others.

Pricing

All data are available for free. The data provided have to meet the agency's Information Quality Guidelines and are free to re-use without any licence.

Additional features

The portal has a rating feature and a communities area (that includes open data, health, semantic web and law). It provides metadata, information about how to access the datasets, and tools that leverage government datasets.

Apps

The portal showcases apps at <http://www.data.gov/developers/showcase>. It mentions that 934 government and 236 citizen-developed apps (including 44 mobile apps developed by government agencies³⁰) have been developed by the end of March 2011.

Accessibility

The access is anonymous. The portal is compliant with Web Accessibility Initiative Guidelines for W3C standards.

Impact

In one week the total downloads times oscillates around 80 thousand in number (data accessed April 5, 2011). The portal displays statistics regarding the most popular datasets and visitors (number of visits, country and state of origin) on the website. It had 144,201 visits in February 2011 and total of nearly 2M (1,941,453) over the last 12 months.

Source: web survey.

³⁰ <http://apps.usa.gov/>

Data.gov.uk	UK
Transparency and Digital Engagement team in the Cabinet Office	

Launch date

September 30, 2009

Short description

The data.gov.uk is still running in a beta version. Data.gov.uk is a key part of the government's work on transparency which is being led by the Transparency Board. Data.gov.uk implementation is being led by the Transparency and Digital Engagement team in the Cabinet Office. It works across government departments to ensure that data is released in a timely and accessible way. It aims at displaying data in an easy-to-find, easy-to-licence and easy-to-re-use way. The publishing team draws on the combined expertise and wisdom of the Transparency Board to publish government data in a way that is consistent with the Public Data Principles.

Data

The portal makes available over 6,200 datasets in machine-readable formats (csv, xml, etc). The publishing team plans to convert datasets to use Linked Data standards.

Pricing

The data and information are licensed under the [Open Government Licence](#),

which enables commercial and non-commercial use of data (unless stated otherwise, e.g. Met Office data are subject of the Met Office licence).

Additional features

Tags for datasets, ranking feature, linked data, ideas on data use, forum, wiki, blogs, resources on open data

Apps

It showcases more than 100 apps (140 so far) built thanks to the PSI data. The [Show Us a Better Way](#) contest was organized before the launch of the portal. The most popular app (crime mapping) is noted for its 400 million hits in the first three months since the portal launch.

Accessibility

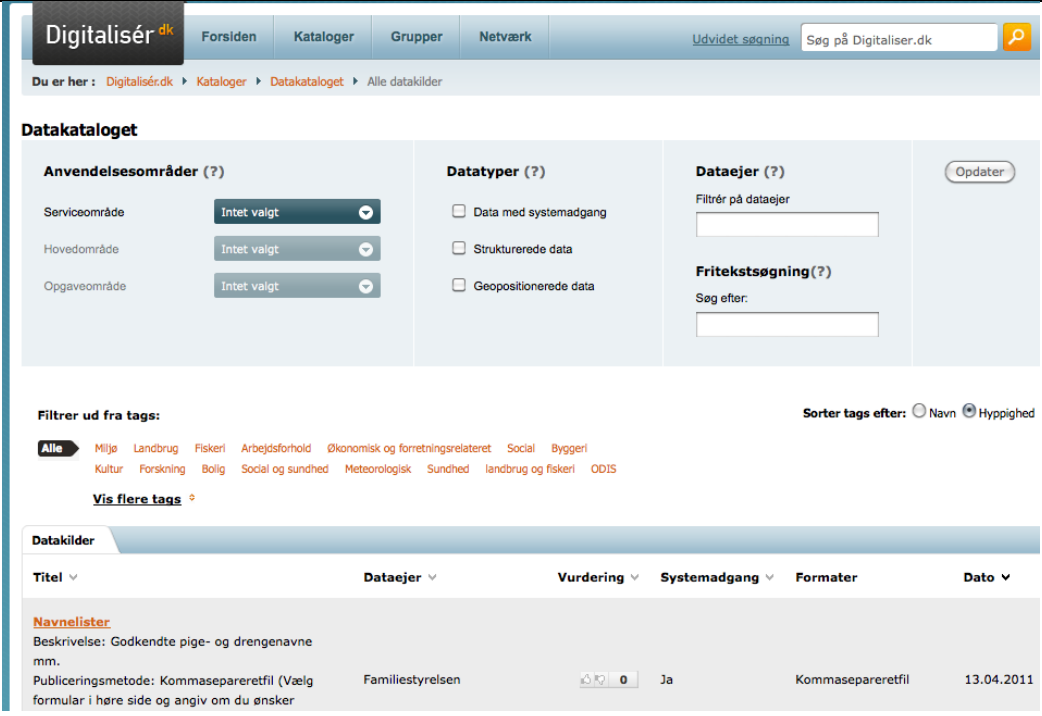
The portal complies with W3C accessibility guidelines. The access to the data is anonymous (the UK Met Office data are available on registration).

Impact

80,000 unique users per month.

40 of the applications have received a five-star rating.

Source: web survey, case study interview.

Digitaliser.dk	Denmark
National IT and Telecom Agency, part of Ministry of Science, Technology and Innovation	
 <p>Launch date Summer 2009</p> <p>Short description The Danish gov.portal is a web directory, and data are located at the public authorities' own servers or chosen host servers. The Data Catalogue is hosted at a collaborative platform (digitaliser.dk is), and was created and is maintained by the Danish National IT and Telecom Agency, which is part of the Ministry of Science, Technology and Innovation. The platform is a collaboration tool open to the public. The portal is available in Danish only.</p> <p>Data The catalogue lists around 00 public datasets (which were put together by screening the websites of all state agencies and ministries). It means that not all the public data already available in Denmark is listed on the website. In December 2010, the list was updated. The directory links to data in machine-readable formats (such as GeoJSON, HTML, JSON, KML, XML, xls, html) and pdfs. The list consists mainly of central government datasets but local governments are being encouraged to participate (there are already some local municipalities datasets). Open data portals are an increasing trend, and should therefore be carefully assessed. showing continuing positive results albeit not dramatic. While the strategy is certainly commendable, a number of risks have to be taken into</p>	

account:

- managing expectations:
- start from real needs
- focus on valuable datasets
- design for citizens, not only for developers

Pricing

The data are available either for free or at marginal cost depending on the dataset.

Additional features

Community forum, resources on digital services and knowledge society, open source software catalogue.

Apps

The portal launched an app contest in 2009 (which ran from October 2009 to February 2010). Three application ideas were awarded a price of 100,000 DK (two individual developers and one company). The contest comprised of two parts, ideas gathering and solid deployment projects. Twenty-five applicants submitted app projects in the second round app contest (they consisted of NGOs, private companies and individuals).

Accessibility

Access is either anonymous or by registration depending on the dataset. The portal has no explicit web accessibility strategy (such as W3C guidelines).

Impact

The portal has not collected any take-up data so far. Additional qualitative information is gathered through comments from the portal's users.

Source: web survey, information from Cathrine Lippert, Special Adviser at the Danish National IT and Telecom Agency and from her report (European Public Sector Information Platform Topic Report No. 20 PSI Reuse in Denmark).

Data.suomi.fi	Finland
State Treasury, Government IT Shared Service Centre	
<p>Launch date</p> <p>November 2009</p>	<p>Short description</p> <p>Suomi.fi is the Finnish national public services portal for citizens run by the State Treasury, IT division. The portal announces that its “primary mission is to offer a one-stop service for e-services and printable forms from both central and local government, and to provide information on citizens’ rights and duties and the role of different authorities in service production”. Similarly to the Danish portal it contains a small data catalogue – a web directory linked to public freely available datasets. Very recently in the beginning of 2011 the Finnish government announced that publicly funded information and raw data must be open, freely available (if not classified for some valid reason) and machine-readable. A new law on public information management has been passed by the Parliament, which might result in creation of a data.gov portal modelled on data.gov.uk. Once more, similarly to the Danish portal, its aim is to encourage discussion on PSI re-use and facilitate public agencies' publication of data.</p>

<p>Data</p> <p>The data catalogue has been in operation a year and a half but there are no plans to transform or expand its services. It contains 46 links to open data of different PSBs (national and local). The formats differ from machine-readable (xml, cvs) to pdfs.</p>
<p>Pricing</p> <p>All data listed in the catalogue are available free of charge.</p>
<p>Additional features</p> <p>Possibility of RSS feeds.</p>
<p>Apps</p> <p>The Apps for Democracy Finland competition co-organized by the portal governing body resulted in 23 submissions, out of which 12 were working prototypes. The portal showcases 50 apps.</p>
<p>Accessibility</p> <p>Anonymous access. The portal facilitates access for people with visual impairments.</p>
<p>Impact</p> <p>No indicators have been gathered so far.</p>
<p>Source</p> <p>web survey, information from Marko Latvanen, web editor, Suomi.fi team, Government IT Shared Service Centre, State Treasury.</p>

Data.australia.gov.au	Australia
Australian Government Information Management Office	
	
<p>Launch date</p> <p>September 2009</p>	<p>Short description</p> <p>Data.australia.gov.au was released in beta version. It was recently re-launched as final production version in the beginning of 2011. No user statistics have been gathered on the previous version. It hosts Australian government public information datasets and in some cases provides links to other datasets. It is run by the Australian Government Information Management Office. It was created following the Government’s Declaration of Open Government and as a response to the Government 2.0 Taskforce report³¹.</p>
<p>Data</p> <p>Currently, there are 223 datasets available but, according to the portal team, this number will grow as “a large number of Australian Government agencies are getting datasets ready to contribute”. The data are provided in machine-readable, proprietary and non-proprietary formats, such as XML, CSV/TXT, XLS, KML/KMZ, ESRI. The portal is not responsible for the quality or timeliness of the data provided.</p>	

³¹ <http://www.finance.gov.au/publications/govresponse20report/index.html>

<p>Pricing</p> <p>The data are released for free and in most cases are bound to Creative Commons licence.</p>
<p>Additional features</p> <p>The portal has a data ranking option and a suggestion page to propose new datasets.</p>
<p>Apps</p> <p>The portal has an area for developers where they can submit apps they have created. It has run one apps contest, Mashup Australia (http://mashupaustralia.org/) organized by Government 2.0 Taskforce with 82 entries and announced another - Libraryhack, a mashup and apps competition designed to encourage the creative and innovative reuse of library data and digital content (May 2011). The latter is being run by the National, State and Territory libraries of Australia and New Zealand.</p>
<p>Accessibility</p> <p>Accessible anonymously. The website is also compliant to Level A of the (WCAG 2.0 standard).</p>
<p>Impact</p> <p>No other indicators accept for apps registry has been introduced yet.</p>
<p>Source: web survey, information from the portal team.</p>

Launch date

November 2009

Short description

Data.govt.nz is a web repository, i.e. a directory of links to the New Zealand government datasets. The portal is run by Government Technology Services at the Department of Internal Affairs. It was inspired by citizen-led initiatives Digital New Zealand, the Open Data Catalogue and other open data projects.

Data

It links to more than 500 datasets (561 so far) and listings (each listing page refers to one or more datasets) and focuses on machine-readable formats xls, csv, xml but has also some html tables and pdf data formats.

Pricing

As the website do not host any datasets, the [Creative Commons Attribution 3.0 New Zealand licence](#) applies only to the website, and each government agency may impose their own licensing requirement on the data provided.

Additional features

The discussion forum enables community building. Moreover, the portal runs a user

survey.

Apps

One app contest has been organised so far - mixandmash.org.nz (in November 2010). The next contest is planned for August 2011. There are also plans to deploy an apps register in the near future.

Accessibility

The access to the web directory is anonymous (the further availability of datasets is defined by each data holder). The website has an accessibility policy (access key site navigation).

Impact

As datasets are not hosted on data.govt.nz, the team cannot monitor dataset downloads. Instead they monitor click-throughs (or referrals) to data hosting websites. This currently averages at 15% of all page visits, and is trending upwards. There are on average 5,377 page visit per month, and this is also following an upwards trend. There was a 68% increase in page visits in February 2011 compared to January 2011. The February 2011 click-through rate was 17%.

To measure the initiative's impact the portal uses the following measures:

- number of dataset listing pages
- number of unique visitors
- conversion rate (click-throughs to data hosting websites)
- range of government agencies contributing dataset listings
- proportion of datasets provided in machine-readable file formats
- outgoing ATOM feeds subscriptions (for dataset listing updates).

Source

web survey, information from Nadia Webster, Senior Web Specialist. Government Information Services.

risp.asturias.es

Spain, Asturia

Government of the Principality of Asturias



Catálogo de Datos de Asturias

En la siguiente tabla se enumeran los conjuntos de Datos de Asturias, los cuales están representados por filas. Cada conjunto de datos puede estar representado en los diversos formatos que aparecen en la tabla y se puede obtener sobre cada uno de ellos (detalles técnicos, licencia, etc.) en las fichas asociadas a cada uno.

Consulta más sobre los [formatos disponibles](#) o sobre los [términos de uso](#) de los datos.

Conjuntos de datos y los formatos en los que aparecen

Nombre	Actualización	HTML	SPARQL/XML	SPARQL/JSON	RDF/XML	RDF/N3	TURTLE
Recursos turísticos del Principado de Asturias	10/01/2011	✓	✓	✓	✓	✓	✓
Inversiones en infraestructuras de telecomunicaciones e impacto en la sociedad asturiana.	01/02/2011	✓	✓	✓	✓	✓	✓
Datos estadísticos del INE y datos de inversiones en infraestructuras y telecomunicaciones.	01/02/2011	✓	✓	✓	✓	✓	✓
Oferta Formativa del Servicio Público de Empleo del Principado de Asturias	10/01/2011	✓	✓	✓	✓	✓	✓
Organigrama de las Consejerías del Principado de Asturias	10/01/2011	✓	✓	✓	✓	✓	✓
Edificios y localizaciones de la Administración del Principado de Asturias	10/01/2011	✓	✓	✓	✓	✓	✓

Launch date

January 2010

Short description

Catálogo de Datos de Asturias is a part of the portal of the Government of the Principality of Asturias that lists publicly available PSI. This initiative aims to promote the use and re-use of PSI.

Data

The catalogue lists currently five datasets in following formats html, sparql/sml, sparql/json, rdf/xml, rdf/n3/ turtle, with different update timelines.

Pricing

The data are licensed under Creative Commons (CC 3.0) and the re-users are obliged to acknowledge the authorship and cite the Government of the Principality of Asturias as the source of the data.

Additional features

None.

Apps

No apps featured on the website.

Accessibility

Five-star linked data system.

Impact

No information available from portal team.

Source

web survey.

Government of Basque Country, Open Data Euskadi Project

¿Qué es Open Data Euskadi?
Es el portal de acceso a los datos públicos del Gobierno Vasco en formato reutilizable. Entre sus objetivos destacan:

- Generar valor y riqueza**
Obteniendo productos derivados de los datos por parte de terceros.
- Generar transparencia**
Reutilizando los datos para analizar y evaluar la gestión pública.

¿Cómo utilizar los datos?

Destacados

- Procedimientos y trámites necesarios para el ejercicio de una actividad empresarial en Euskadi
XML (1.5 Mb) XLS (922 Kb) CSV (1.4 Mb) RDF
- Mapa de las estaciones remotas de la red de vigilancia de la calidad del aire
XLS (58 Kb) CSV (29 Kb) RDF
- Calendario laboral de Euskadi para el 2011
XML (88 Kb) XLS (113 Kb) CSV (32 Kb) RDF

[Ver el catálogo de datos](#)

Últimos datos publicados

Lo más nuevo en el catálogo de datos

- Mapa de las estaciones remotas de la red de vigilancia de la calidad del aire
XLS (58 Kb) CSV (29 Kb) RDF

TEMAS

- Turismo

Launch date

April 2010

Short description

The Millennium Open Data Euskadi portal is aimed at generating added-value from PSI, promoting transparency in public administration, interoperability and efficiency of public agencies. It was the first non-Anglo-Saxon gov.data portal.

Data

The portal makes available 1,125 statistic datasets and 77 other datasets. There are a total of 1,202 datasets available in machine-readable formats (xml, csv, wms, other).

Pricing

The data are made available on the condition of licensed recognition of authorship (set to a Creative Commons Attribution 3.0 Spain) or a free licence and free re-use of public information.

Additional features

Blog, twitter.

Apps

Approximately 10-15 users have downloaded the data in order to develop an app, although a register of apps is not available.

Accessibility

The access to data is anonymous. The website is compliant with WCAG 2.0.

Impact

Between April 6, 2010 and February 25, 2011 the website was visited by 25,362 unique visitors, had 33,682 visits and 157,520 pages were visited. Over 250 datasets are downloaded each month.

Apart from web traffic information, the portal analyses blog comments

(<http://opendata.blog.euskadi.net/blog-en/>) and tweets

(<http://twitter.com/opendataeuskadi>).

Source

web survey, information from Imanol Argüeso Epelde

Open Data Euskadi Project

Government of Catalonia, Open Data project

Dades obertes gencat
Projecte d'obertura de dades públiques (open data) de la Generalitat de Catalunya

Catàleg de dades obertes

- **Imatges de Sant Jordi**
Fitxer que conté 196 fotografies en alta resolució relacionades amb la diada i el Palau de la Generalitat, cedides pels fotògrafs de la Generalitat.
- **Estadístiques sobre convocatòries en d'R+D+I desagregades per sexe**
Indicadors sobre les convocatòries públiques de beques i ajuts a la recerca finançades per la Generalitat de Catalunya i destinades a recursos humans des de l'any 2004.
- **Predicció meteorològica comarca a comarca**
Accediu a la predicció en format xml de cada una de les 41 comarques catalanes per a demà i demà passat.
- **Calendari laboral 2011**
Calendari oficial de festes laborals a Catalunya per a l'any 2011.
- **El municipi en xifres**
Estadístiques bàsiques de cadascun dels 947 municipis i 41 comarques de Catalunya.

Ajuts de la Fundació puntCAT a projectes basats en dades públiques obertes
Consulteu la llista de portals que ofereixen dades obertes en català.

Tràmits Ajuts Estadístiques
Indicadors Mapes Equipaments
Ortofotografies Imatges Multimèdia Dades
Lingüístiques Terminologia
Meteorologia Predicció Trànsit
Transport públic Actualitat

Aplicacions fetes

- Indicadors dels municipis catalans a la plataforma de dades públiques de Google
- Cercador d'equipaments de Catalunya
- Oficina Virtual de Tràmits
- Guia de Catalunya

Launch date: November 2010

Short description

Gencat Open Data is a recent project of Government of Catalonia's portal (November 2010).

Data

The portal make available the initial database of the 26,000 official facilities of Catalonia, the 1,400 procedures handled in the government offices and some of its multimedia archives. It makes available 124 datasets in machine readable formats (tmx, zip, pdf, csv, kml-kmz, doc, xls, xml, json, rdf-xml, shp, sparql).

Pricing

The re-users should contact the data holder and recognize its authorship of the data. Most of the datasets are released under a Creative Commons Attribution 3.0 Spain licence (four re-use licences recommended in the *Reuse Guide of the Aporta Project* released by the Ministry for Science and Technology (Avanza Plan): www.aporta.es/web/guest/guia_reutilizacion). The data are available free of charge.

Additional features

Section on ideas for apps and general information on open data.

Apps

The portal showcases 11 apps on its website. The portal received 70 requests for data that may help to estimate the number of prospective app developers.

Accessibility


Anonymous access. Compliant with WCAG rules.

Impact

The portal does not yet track the number of datasets downloaded, but since its launch in November 2010 until mid-March 2011 there have been 29,000 visits with 83,689 page views.

Source

web survey, information from Eva Gumà, Open Data project of the Government of Catalonia.

dati.piemonte.it	Piedmont, Italy
Piedmont regional government	
	
Launch date May 2010	
Short description In May 2010 the Piedmont Region launched an online beta data portal. It provides access to raw data from the regional government free of charge including for commercial purposes, as well as gives access to online resources on open data.	
Data The portal offers 242 datasets with data from the Piedmont region. The data are available in various formats (proprietary and non-proprietary, not all are machine-readable, e.g. many pdfs).	
Pricing The data is licensed under the Creative Commons Attribution (CC 2.5) and open to re-use for free.	
Additional features Features: most downloaded, latest added, top rated function, list of PSI events.	
Apps It showcases nine visualizations and applications built with the use of data available on the portal (http://dati.piemonte.it/casi-duso.html). The portal plans to	

organize an app contest.

Accessibility

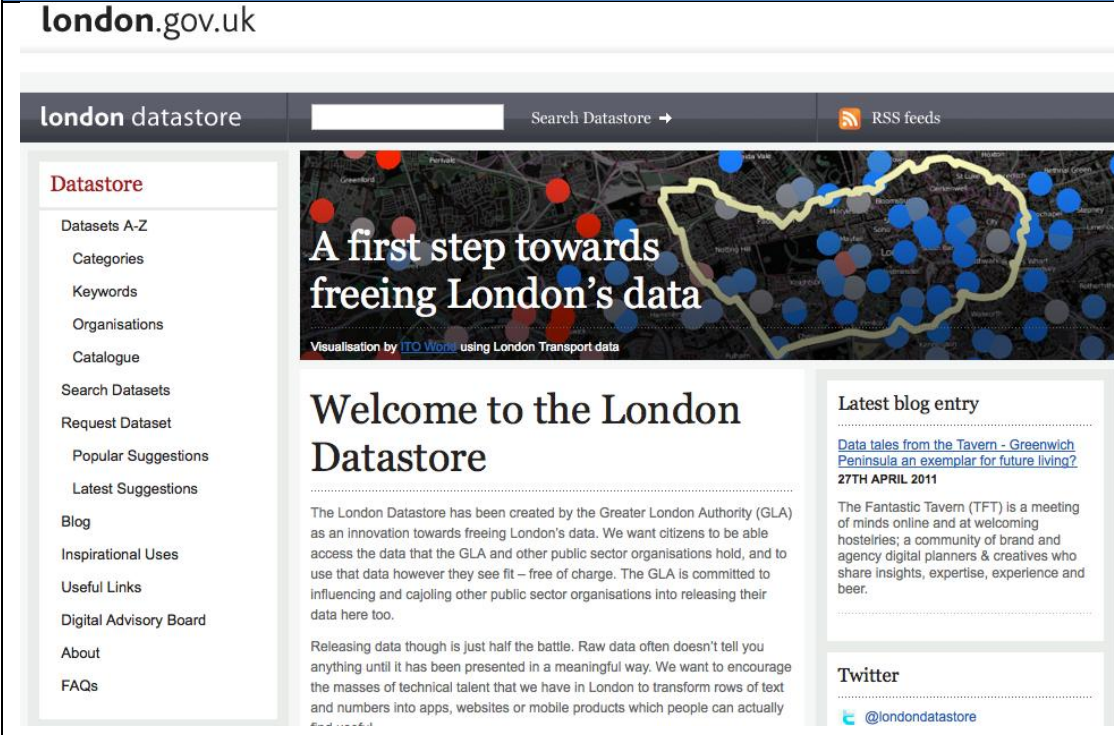
The access to data is anonymous. The website does not have an explicit web accessibility strategy.

Impact

The portal makes available several statistics on the website including the most downloaded datasets (either according to theme or organization) and number of visits. The portal monthly registers around 3,000 visits and 1,700 downloads.

Source

web survey, information from Silvia Bianco, Department of Innovation, Research and University, Piedmont Region.

London Datastore	UK, London
Greater London Authority, Digital Projects	
	
Launch date January 29, 2010	
<p>Short description</p> <p>The London Datastore was created by the Greater London Authority. Its Digital Advisory Board is chaired by the Mayor's Economic Advisor Anthony Browne, and includes a number of social innovators who have already engaged and supported the work of the London Datastore.</p>	
<p>Data</p> <p>The data range from air quality, Tube timetables, voting patterns, hospital performance, fires to road accidents. Currently, the portal makes available 438 datasets that are provided in machine-readable and are mostly in a non-proprietary format (csv and xml).</p>	
<p>Pricing</p> <p>The re-users should explicitly state that the Greater London Authority cannot warrant the quality or accuracy of the data. The Greater London Authority is currently working on licensing terms with the advice from Open Knowledge Foundation. Data are provided free of charge.</p>	
<p>Additional features</p> <p>Blog, inspirational uses, request a dataset, RSS feeds.</p>	

Apps

The portal showcases around 60 inspirational uses of its data <http://data.london.gov.uk/datastore/inspirational-uses>.

No app contest organised so far.

One application, Tube Deluxe, shows live departures of Underground trains, the quickest routes and whether there are any delays: it has been downloaded 350,000 times so far.

Accessibility


This website has been created to be as accessible as possible and in adherence with the Greater London Authority's current guidelines on web accessibility. The site is reviewed on a regular basis and will further improve its accessibility where possible. Wherever possible, datasets are provided in CSV and XML formats for ease of extraction and use. Where this has not been possible, data are provided in the most appropriate and widely used commercial format.

Impact

The first year: 0.5 million visits and 2.4 million page views. 79,000 downloads of CSV files, equating to around 7GB of data (these numbers do not include downloads from external sites such as the TfL Developer Area).

Source

web survey, information from Gareth Baker, Geographic Information Manager, Intelligence Unit, Greater London Authority.

Data.rennes.metropole.fr	FR, Rennes
Digital Innovation Service, Communication and Information Department	
	
Launch date: April 2010	
<p>Short description</p> <p>The data Rennes municipality portal is run by the Digital Innovation Service, which is the part of the city's Communication and Information Department. It was the first initiative of this kind in France to offer free access to transport, geographic and geo-location data.</p>	
<p>Data</p> <p>It makes available around 90 datasets with different time of updates in machine-readable format (jpg, dxf, shp, gfts, rss, api, ics, xml, csv, kml).</p>	
<p>Pricing</p> <p>The data are available for free re-use according to a « Licence de réutilisation des données publiques - Rennes Métropole en accès libre » which permits non-commercial and commercial re-use.</p>	
<p>Additional features</p> <p>Developers' corner, Ideas forum.</p>	
<p>Apps</p> <p>The portal has organised recently an app contest which is called Rennes Métropole en accès libre. It resulted in 43 apps competing for 50,000 EUR in awards. The App Register features no apps yet and will soon be complemented with apps developed by the contest winners.</p>	

Availability



The datasets are available anonymously but access to developer's corner requires registration. The portal has no explicit web accessibility policy.

Impact

No information available from portal team.

Source

web survey.

Helsinki Region Infoshare	Helsinki Region, Finland
City of Helsinki Urban Facts, Forum Virium Helsinki	
<div style="display: flex; justify-content: space-between; align-items: center;"> <div data-bbox="264 315 496 387">  </div> <div data-bbox="1066 353 1278 371"> Suomeksi / In English / På svenska </div> </div> <div style="display: flex; justify-content: space-between; align-items: center; margin-top: 10px;"> <div data-bbox="272 416 647 434"> Data Ajankohtaista Keskustelut Lisätietoja </div> <div data-bbox="1066 409 1278 430" style="border: 1px solid #ccc; padding: 2px;"> <input type="text" value="Hae koko HRI:stä"/> Search </div> </div> <div style="display: flex; justify-content: space-between; margin-top: 20px;"> <div data-bbox="268 486 544 602" style="width: 30%;"> <p>Hae dataa</p> <div style="border: 1px solid #ccc; padding: 5px; display: flex; align-items: center;"> <input style="width: 80%;" type="text" value="Hakusana"/> <input style="width: 15%; background-color: #92d050; color: white; border: none;" type="button" value="Haku"/> </div> <p style="font-size: small; color: #0070c0;">Tarkennettu haku</p> </div> <div data-bbox="600 486 871 629" style="width: 35%;"> <p>Yleisimmät avainsanat</p> <ul style="list-style-type: none"> ikä (103) sukupuoli (74) väkiluku (50) väestö (35) toimiala (33) suurpiiri (28) työpaikat (26) alueella-työssäkäyvät (26) ruotsinkieliset (26) miehet (24) </div> <div data-bbox="959 486 1278 636" style="width: 30%;"> <p>Viimeisin data</p> <p>Helsingin seudun sairastavuusindeksi 2009</p> <p style="font-size: x-small;">Sairastavuusindeksi lasketaan kolmen väestön sairastavuutta kuvaavan rekisterimuuttujan avulla. Nämä ovat työkyvyttömyyseläkkeellä olevien osuus</p> <p style="background-color: #92d050; color: white; padding: 2px; text-align: center; font-size: x-small;">LUE LISÄÄ →</p> </div> </div> <div style="margin-top: 20px;"> <h2 style="color: #0070c0;">Avointa seudullista tietoa</h2> <p style="font-size: x-small; color: #0070c0;">www.hri.fi -verkkopalvelu tarjoaa helpon tavan löytää, saada ja hyödyntää Helsingin seudun julkisia tietovarantoja avoimena datana. Tarjolla oleva tieto on pääosin tilastodataa, mutta palvelun kautta löytää myös muuta seudun avointa dataa.</p> <p style="font-size: x-small; color: #0070c0;">Verkkopalvelu on kehitysversio. Sen toiminnallisuuksia parannetaan ja avoimen datan saatavuutta edistetään käyttäjiä kuunnellen. Palvelu tarjoaa kanavan toiveiden ja palauteen antamiseen sekä aiheeseen liittyvään keskusteluun.</p> <p style="background-color: #92d050; color: white; padding: 2px; text-align: center; font-size: x-small;">LUE LISÄÄ →</p> </div> <div style="text-align: right; margin-top: 20px;">  </div>	
Launch date: March 2011	
<p>Short description</p> <p>The Helsinki Region Infoshare project is an initiative of the cities of Helsinki, Espoo, Vantaa and Kauniainen, as well as Forum Virium Helsinki and the Finnish Innovation Fund Sitra. The project is coordinated by City of Helsinki Urban Facts. The overall budget for the project is estimated to be around one 1 million EUR. The project started in 2010 and is supposed to be finished by the end of 2012. It was officially launched in March 2011 in beta version. It foresees web service for an easy and fast access to open data sources.</p>	
<p>Data</p> <p>The project offers access to mostly statistical but also to some other open data regarding the Helsinki area and its localities (living conditions, economics and well-being, employment and transport). A good proportion of the data material offered by the project is GIS based. At its launch, it offered free access to 549 data sets. However, the number of data sets is gradually increasing. The first set of release is planned for the end of 2012 and from this date onwards the data catalogue will be complemented with new datasets. The information will typically be updated annually, quarterly or monthly.</p>	
<p>Pricing</p> <p>Free of charge on complying with HRI data license.</p>	

Additional features

News page, tags, RSS feeds, linked data are yet planned to be published.

Apps

Apps for Democracy – Finland contest was co-organised by project partner - Forum Virium Helsinki in 2009 (before the project launch).

Accessibility

The portal is accessible in Finnish and Swedish and has some information in English. The data can be accessed anonymously. The website has no information on web accessibility policy.

Impact

The project has been of very short duration to date (as per June 2011). No data are available yet.

Source

web survey, information from Asta Manninen, Director of City of Helsinki Urban Facts

The screenshot shows the 'utah.gov/data/' website. At the top, there's a blue navigation bar with the Utah logo and 'US, Utah'. Below that, a grey banner reads 'THE OFFICIAL WEBSITE FOR THE STATE OF UTAH' with utility links like 'Subscribers', '24/7 Help', and 'Font Size: A A A'. A secondary navigation bar lists categories: GOVERNMENT, RESIDENTS, VISITORS, BUSINESS, EMPLOYMENT, EDUCATION, ABOUT, ONLINE SERVICES. The main content area has a dark blue header with 'DATA' and 'RELATED LINKS & RESOURCES'. The main body is a grid of six data source categories, each with an icon, title, description, and a 'View' link. A right sidebar provides additional links and a feedback form.

Launch date
 Summer 2009

Short description
 Utah was one of the first states to open a data.gov portal after the launch of federal data.gov in the US. The portal is linked to federal data portal, states and local data sources.

Data
 It makes available 31 state and local datasets which have different timelines (some datasets date from 2005 e.g., the Average Newborn Hospital Charges, others from 2006 or 2008). The data are accessible in machine-readable formats (CSV, KML and XLS).

Pricing
 There is no information on re-use conditions. The data are accessible for free.

Additional features
 It has a twitter account and translation feature (via Google Translate).

Apps
 No apps are featured on the website.

Accessibility
 The portal complies with accessibility standards established by the state government. The access to the data is anonymous.

Impact

No information on take-up was available.

Source

web survey.

Launch date

October 2009

Short description

NYC DataMine catalogue datasets of public data produced by municipal agencies and organizations. It is a part of an initiative to improve the accessibility, transparency, and accountability of municipal government.

Data

The data sets are available in a variety of machine-readable formats and are refreshed when new data becomes available. Data are presented by category, city agency, or other city organization. Metadata are available. There are 217 raw and 144 geospatial datasets. The datasets are presented in the machine-readable formats XLS, XML, CSV, and RSS, SHP and KML and others. Historical datasets are not retained. The frequency of update is indicated per dataset.

Pricing

The agencies submitting data are responsible for data quality and retain version control of datasets and feeds accessed on the site. All data are available free of charge. App developers have to notify the city and include a disclaimer available from the portal.

Additional features

The website has such features as a 'suggest new data set', e-mail to a friend, and automatic page translation.

Apps

The portal has already organized two apps contests, i.e. NYC BigApps Competition (<http://nycbigapps.com/>). There were 57 submissions to the second round of the contest, including 24 mobile apps.

Accessibility

All data are available anonymously. The website contains some elements of web accessibility features – text size changes.

Impact

The portal tracks visits, unique visitors and page views. From August 2010 to February 2011, the portal has had 57,403 visits, 25,479 unique visitors and 147,599 page views. The monthly statistics (for February 2011) are as follows: visits 8,864 (compared to 144,201 visits of US data.gov portal), visitors - 4,086 and page views - 21,229.

Source

web survey, information from NYC data support team.