Digital Agenda Assembly

Brussels 16-17 June 2011

Report from the workshop

"Open Data and re-use of public sector information,
http://ec.europa.eu/information_society/events/cf/daa11/item-display.cfm?id=5963,#daa11psi"

(DAE Action 3 - Open up public data resources for re-use)
Workshop headlines

Headlines from the workshop "1. Open data and re-use of public sector information" (Digital Agenda action 3)

Purpose of this workshop
Bring together data holders, technology providers and potential re-users to promote open data/PSI as key driver to develop European content markets and deploy new data initiatives.

Status of progress and challenges reported in the workshop
- Access still critical: Make PSI available as open data in readable formats
- Licensing regimes: data and metadata, copyright and database rights
- Economics: don’t just look at PSI market, look at macroeconomic footprint

Stakeholders’ actions and commitments mentioned in the workshop
- lower the barriers of access to data: cost models, open standards, licensing
- decentralized innovation happens when data is available: loosen up!

Purpose and context of the workshop

One objective of the Digital Agenda for Europe (DAE) is to facilitate the wider deployment and more effective use of digital technologies to enable Europe to address its key challenges and provide Europeans with a better quality of life, for example, by opening up access to governmentally held information and content (PSI – Public sector information).

Governments are increasingly recognizing the benefits of making their data open and reusable. The wide availability of Public Sector Information can be a key driver to develop content markets in Europe, which can generate new businesses and jobs and provide consumers with more choice and more value for money. Governments will be more transparent and government services will improve and innovate much faster as more PSI becomes widely available. Severin Naudet summarized the meaning of PSI re-use and Open Data during the workshop: He stated that: “Open Data is a digital application of democracy, which will fuel a major share of growth and jobs creation”.

This workshop brought together data/PSI holders, technology providers and potential re-users to facilitate an exchange of views, to support the definition of requirements (technical interoperability, data management, strategy), to promote the pan-European added value of data portals, to deploy data initiatives and enhance synergies between them. The participants in the workshop aid the Commission
in developing its strategy for open data for the EU, taking into consideration the review of the Public Sector Information Directive (Digital Agenda action 3).

The workshop built on the results of the Share-PSI.eu meeting, a multi-stakeholder initiative initiated by W3C and ETSI that held a preliminary workshop in Brussels on 10-11 May 2011.

The workshop is also linked to Big Idea 3 (Beyond raw data: public sector information, done well).

**Related DAE action(s)**

By 2012, review the Directive on Re-Use of Public Sector Information, notably its scope and principles on charging for access and use.

**What is the problem? Turning public data to business opportunities**

Public authorities produce large amounts of data that could become the raw material for new, innovative cross-border applications and services. Examples of products and services based on the re-use of public sector information (PSI) are GPS, weather forecasts, financial and insurance services.

PSI is the single largest source of information in Europe. Its estimated market value is €27 billion. Re-used, this public data could generate new businesses and jobs and give consumers more choice and more value for money.

**Why is EU action needed? To unlock the public data potential across Europe**

Often the use of this information is restricted to governments or public bodies. Governments could stimulate content markets by making public sector information more widely available for re-use on transparent and effective terms.

**What has the European Commission done?**

- The Directive was adopted in 2003 which lays down basic principles of availability, transparency and cost limits to ensure a level playing field for commercial re-users of public sector information.

- In preparation for a review, the Commission has held a public consultation on the review of the PSI Directive and published an analysis of the responses.

**What will the European Commission do?**

- Drawn up an impact assessment of the various options and, if appropriate, propose a revision of the Directive.

**Ongoing actions:**

- Awareness-raising and promotional activities through seminars and workshops organised by Member States and stakeholders,

- Networking across Europe and in a wider international context (with Member States, PSI Alliance, national industrial associations, PSI platform, software developers and open data communities) to further stimulate action and monitor progress towards a stronger and more transparent environment for the growth of European PSI.
Results from previous online exchanges and activities

A number of previous online and offline exchanges and activities took place.

A blog post was written by Mr. Richard Swetenham [here](#), that was viewed 1059 times (June 21st 2011).

The [SharePSI](#) meeting took place in Brussels on may 10th and 11th. Notes were taken on the discussion in the SharePsi [http://sharepsi.okfnpad.org/workshopnotes](http://sharepsi.okfnpad.org/workshopnotes). The official report from the workshop is also available: [https://daa.co-ment.com/text/GXXnDNBIU68/view/](https://daa.co-ment.com/text/GXXnDNBIU68/view/). Ton Zijlstra made an annotated, visualization of all barriers that people encounter when working with PSI. All barriers are grouped in 7 clusters: Licensing, The transition process for government, From access to new business, Disruption of existing economic activity, Towards linked federated data, Privacy and Other issues. An example of one of these clusters (licensing) is shown below. These graphics give a good insight into the breath of the challenges the whole community is still facing. More information on the images and in introduction to them can be found [here](#), while the original overview of all clusters can be found [here](#).

![Diagram of licensing issues](image)

All key issues and ideas from the workshop are collected in the ‘uservoice’ tool as well. A number of people have also voted on the issues and ideas collected [there](#). The key issues for the workshop are (in order of popularity) The Share-PSI workshop was central in the build-up to the Digital Agenda Assembly Workshop 1 (Open Data and PSI re-use). 80 people were present and debated on the ‘barriers’ existing for PSI re-use and Open Data, as well as on the solutions to overcome these barriers. Margot Dor presented [the results](#) of the SharePSI workshop at the DAA Ws. 1.

On the DAA site, three people have commented on the workshop. Richard Blad of the ministry of Economic Affairs, Agriculture and Innovation of the Netherlands stated he “look[ed] forward towards the workshop”. Antti Eskola from the Ministry of Employment and the Economy of Finland said afterwards that the workshop “gives us ideas on policy measures. Nice apps competition results”. Francesca di Massimo of Microsoft (Italy) congratulated the EC on a “really very good” session.
Status of progress and challenges reported in the workshop

Access still critical

Access to PSI is still the main obstacle for a thriving PSI re-use practice in Europe. Tommy Kaas talked about fire and police departments, stating “Every department should publish data”, indicating that they do not yet, do so.

Even when there is data available, there is much fragmentation and most data sources are not yet exposed. Available data sources are often hard to find, with unclear provenance and inconsistencies in structure and format. Luis Jose Marin said that the greatest challenge is “Unlocking the data, in machine readable formats”.

Licensing regimes

Much has been said during the discussion on licensing. Hodgson concluded that “The PSI Directive itself is “is subject to a number of important exceptions eg 3rd party IPR, education & research establishments … The issue of the re-use of marine environmental data remains subject to the individual data policies of individual data centres which in turn have received differing messages including as regards the issue of use/re-use”.

Jose Luis Marin (Euroalert, Spain) also stated that licensing is one of the big issues for re-use. Jim Wrathem (National Archives, UK) said that the licensing issue is the single biggest problem in the field.

The problem in licensing is that there either are no licenses, that PSI holders release data under (too) strict licenses, or that commercial re-use is not allowed by the PSI Holder. Re-users feel that licenses should not limit them in their possibilities to create new products with the data (or as minimal as possible), while PSI holders like to prevent misuse and or budget problems within their own organisations.

Economics

Panel member Graham Vickery showed that the direct annual turnover in the PSI Sector is estimated to be around EUR 28 Billion, with annual growth of around 8%. This makes the PSI Sector one of the faster growing sectors of the EU economy. He also stated that the aggregated footprint of the PSI industry is EUR 140 Billion. This means that the ‘spill’ or overflow from the PSI industry is even bigger then the sector itself. Conversely, there still aren’t enough examples of PSB’s that used Open Data and PSI Re-use to positively influence their own balance sheets.

We might conclude that even though the macro business case for Open Data and PSI re-use is very positive, there is not enough evidence for individual companies that the micro business case is also positive for them.

Profile of speakers and audience: constituencies present/ missing, geographical balance, any other relevant aspect

The speakers at the workshop were from all over Europe. Speakers from the UK (4), Sweden (1), the Netherlands (2), France (3), Spain (2), Denmark (1), Iceland (1) and Canada (1) were present. 24 nationalities from all over Europe were present, but most from Belgium (52), France (22) and the Netherlands (19). There were few people present from the eastern European member states.

The illustration below shows how many delegates came from all countries.
Key points from the discussion

Keep it Simple – Small steps!

In their presentations, Margot Dor (SharePSI), Simon Chignard and many others emphasised the necessity to focus on improving the PSI sector by taking small steps, and refrain from . Dor said “the way forward is incremental rather then a ‘big bang’.”

Mr. Nigel Shadbolt (UK) referred to the toilet maps of both the UK and Denmark that have been drawing a lot of attention recently. “Even though many people don’t realise it, these applications are very empowering for large groups of people who have a condition that requires them to be quite close to a bathroom (at all times)”, he said. “Even in starting with a small project like this, you can have an incredible impact!”

Making small steps also means that PSB’s shouldn’t be exerting too much control over the datasets they release. As Thomas Roessler paraphrases it, “PSB’s should loosen Up!” and see what results come from it. The recommendation is to refrain from first building a huge (legal, technical or other) framework around the data, but to “Just do it!”

A number of people emphasise the importance of sharing and showing the ‘success stories’ of the PSI and Open Data sector. Thus we all can encourage each other and new stakeholders to find better ways to use the PSI or Open Data, open up more datasets and have more impact in the world we are living in! A few of the success stories are the winners of the Hack4Europe contest and the OpenDataChallenge. They all are impressing examples of the power of PSI. Winners can be found here (Open Data Challenge) and here (Hack4Europe).

Data formats and linked Data?

When it comes to the question of the shape of the data, data in any machine readable format is a good starting point. Luis Jose Marin says it’s the single biggest task in PSI to make more data available in machine readable formats. Hjalmar Gislason adds, “if you have data in a format that
makes sense to you, it will make sense to other smart people too!”, encouraging everyone to open up their data in their own lay-out, but in machine readable formats.

As this is only the first star in the Tim Berners-Lee ‘5 stars of Linked Data’ scheme, both Gislason and others would like to see ‘more star’ datasets. In its best shape (5 stars) these datasets are much easier to re-use, thus spurring more re-use activity and more positive impact from the PSI (e.g. Jose M. Alonso).

However, Alonso also observes that the initial investment necessary to make the data ‘linked’ is high. He says that the mantra should be “Raw data now, better data later” as getting access to more datasets in machine readable formats is already a formidable task.

**Licensing and pricing models**

Jim Wrathem (National Archives, UK) is quite adamant in his statement: "Licensing is the biggest issue when working with PSI." Stephen Hodgson says that "the legal framework for access is available, but it’s still missing for re-use".

It is generally recognized that the legal framework surrounding PSI re-use is far from done as it contains elements of so many types of legislation (IPR, Access/FOI law, Database rights, Competition law etc.). This means that the concepts central to re-use are maybe not defined well enough in legal terms.

In a question to the panel, Federico Morando, suggested that empirical evidence shows that in the countries that do well in the fields of PSI re-use and open data, Creative Commons licenses (CC-0, CC-BY etc) are being used. Morando proposed that the commission would issue guidelines on licensing with the review of the directive. It was also suggested from the audience that the EC recommend member states to adopt the CC-0 license, but in the twitter stream Catherine Lippert (Dk) tweeted about an article that discusses why the license CC-0 might not be a good license for this goal.

**Actions and commitments mentioned in the workshop; future steps**

Severin Naudet (Director of the ETALAB, France) has announced that in his country PSB’s will have to release data for free. Any PSB can file for exemption of this rule (and charge for information) with the prime ministers office, but they will likely be very strict. France will thus create a ‘comply – or ask permission’ policy for its Public Sector Bodies. The prime ministers office will maintain a list of the exempted PSB’s.

The Europeana team is working towards permitting commercial use of metadata. Mr Harry Verwaijen of Europeana has said that over 40% of museums affiliated with Europeana now think that the metadata on their art should be available as widely as possible. He even said that “Metadata, is a business model”.
Norway is currently implementing a policy promoting access to and publication of PSI. Every letter that contains a ‘funding or allocation decision’ (Basically the contract between the government and said funded body) will contain a clause that requires the access to and publication of PSI as a prerequisite for the funding.

The commission itself will also promote the release of PSI within their own offices. Mr Rouhana (DG Information Society and Media), said that the European Commission will present their own data portal in 2012 (http://data.gov.eu) with a number of datasets from the EC, as well as a close relationship to EU sites that already have big datasets as Eurostat, Eur-Lex and CORDIS. In 2013 this data portal will either be supplemented with or extended to also include relationships to national and local dataportals and the commission will focus on making this a sustainable portal.

One of questions of the audience (“Is there any framework that connects all the different parts of the legal framework – IPR, FOIA/Access and Re-use?”) prompted Nigel Shadbolt to state that a recent project in the UK has been making progress, but “we’re not there yet”, he said.

Gohar Sargsyan of Logica (The Netherlands) has a report in the making that will answer many of the questions posed here. So will publicly announce it’s availability when the report is published.

Many of the speakers and audience members that took the stage have been emphasising the need for small steps that are taken as soon as possible. A great number of participants have indicated implicitly and explicitly they are working on resolving problems, generating new products and ideas, and opening up more datasets.
Annexes

Workshop programme

10:00 Short introduction of policy background - Khalil Rouhana | Download presentation

10:10 Show cases

- Results from Open Data Challenge http://opendatachallenge.org
- Results from Hack4Europe! http://version1.europeana.eu/web/api/hackathons - David Haskyia
- Local data initiative: Open Cities - Katalin Gallyas | Download presentation
- Regional data initiative: Rennes - Simon Chignard | Download presentation
- Marine Environmental Data & PSI - Stephen Hodgson MRAG | Download presentation
- European data initiative: EuroAlert - Jose Luis Marin | Download presentation
- Deriving news and patterns from complex data: A case for data-driven journalism - Tommy Kaas | Download presentation

11:00 Feedback from May Share.PSI Workshop and online discussion - Margot Dor https://daa.com ment.com/text/GXXnDNBIU68/view

11:15 Panel: Statements reacting to showcases

- Panel Chair - Prof. Nigel Shadbolt | Download presentation
- Perspective from the PSI Community - Séverin Naudet | Download presentation
- Technology perspective - José Manuel Alonso | Download presentation
- Perspective of re-users, service developers - Hjalmar Gislason | Download presentation
- Legal perspective - Lucie Guibault | Download presentation
- Economic perspective - Graham Vickery | Download presentation
- Business model open cultural data - Harry Verwayen | Download presentation

11:45 Open discussion including audience; brief contributions from panellists

13:00 Summary of the discussion

13:30 End of workshop
Presentations

All presentations can be found here

http://ec.europa.eu/information_society/events/cf/daa11/item-display.cfm?id=5963

List of Workshop participants

A list of workshop participants can be found here: