Deliverable 4.2

Report on the third workshop

Open data priorities and engagement – identifying data sets for publication

Standards for Open Data and Public Sector Information

Table of Contents

Editor: Phil Archer, W3C/ERCIM

Submitted: 04/05/2015

1 Executive Summary

The third Share-PSI workshop was held in the west Romanian city of Timişoara, hosted by the West University. The event showed an evolution from previous workshops in the series:

- Samos was largely a traditional paper-presentation event.
- Lisbon was much more interactive with fewer presentations and more discussions.
- Timişoara, outwardly, was similar to Lisbon, but there was a much stronger focus on eliciting best practices that the project partners could codify and link to the PSI Directive.

Almost all sessions included discussion of best practices that followed from the shared experiences. These were:

Engage a broad community, including technical and non-technical people, in planning and executing open data policies.

Publishers should encourage and facilitate consumers' **reporting of usage of data** to encourage its continued provision.

Publishers should encourage and facilitate consumers' **corrections to data**, using tools such as GitHub, gamification techniques etc.

Publishers should clearly **define their role** in providing PSI whilst encouraging others to build on it.

Publishers should provide a **knowledge base** as part of a data portal.

Publishers should focus on providing **services** such as mixing and visualisation as much as data.

Machine translation technologies should be harnessed to offer data in multiple languages.

As a minimum, data should be available for **bulk download**.

Publishers should be **explicit about the rights** that consumers have in accessing and re-using data.

Publishers should unambiguously express and **communicate the quality level** of their data.

Datasets must **refer to locations consistently** following international standards.

Public authorities should use **common criteria for assessing the impact** of their PSI provision.

Inventories of available information, whether open or closed, should be generated through **scraping** of public authorities' websites.

Public authorities should follow the techniques developed in research publishing to **link reports and studies with the underlying data**.

2 Introduction

The third Share-PSI workshop took the theme of "Open Data Priorities and Engagement — Identifying data sets for publication" and was hosted by West University in the Romanian city of Timişoara. Following on from the success of events in Samos and Lisbon, the Timişoara workshop comprised a series of facilitated discussions with only a small number of presentation-based sessions. The aim of the project overall is to identify what works and what doesn't work as the public sector across Europe implements open data policies in the context of the revised PSI Directive. Recurring topics were the impact of relative studies and indices, the importance of gathering user feedback, the publishers' desire to know who is using their data and for what, and that the demand from citizens is not for data but for services that may be built on that data.

83 people registered for the workshop, a number that included participants from Serbia, Poland, Croatia, Bulgaria, the Czech Republic and, of course, Romania – countries that that had been absent or under-represented at previous Share-PSI events. The sessions in the main hall were streamed and recorded and the event generated a good amount of buzz on Twitter.

2.1   Plenary Talks

Figure 1 The opening session chaired by workshop host Dana Petcu with (from L to R) Chris Harding (Open Group), Benedikt Kotmel (Czech Ministry of Finance), Radu Puchiu (Secretary of State from the Chancellery of the Romanian Prime Minister), Marilen Pirtea, (Rector of West University)

After a welcome from Dr Marilen Pirtea, Rector of West University, the Share-PSI partners were honoured to be joined by the Secretary of State from the Chancellery of the Romanian Prime Minister, Radu Puchiu. During his remarks, he described the hackathons organised to publicise the datasets available on the national portal (data.gov.ro) and to engage the wider community - something that is seen as crucial. Mr Puchiu used the workshop to announce that a new

platform for public procurement will soon be established and its data will be exported and made available in standard formats. This and his comments on the Global Data Index proved highly relevant to later sessions.

The Czech Ministry of Finance's Benedikt Kotmel presented the situation is his country. His ministry is the first to establish an open data portal and five other ministries are following suit. This is in addition to portal.gov.cz, which is run by the Ministry of the Interior. Interoperability of the various catalogues therefore depends on inter-departmental communication, a situation made more complicated by government ministers being drawn from more than one political party. An internal directive was essential to begin the conversation. Mr Kotmel described the demand analysis carried out prior to releasing any data. This drew on several sources:

- FOI requests;
- Universities that were approached directly
- Non-profit organisations;
- Private sector companies.

The latter proved hard to engage but the other sources showed which data sets were most wanted. Further work was then done to see how feasible it was – technically and legally – to release those datasets, before seeking sign off by the minister. As with Romania, the importance of gathering feedback was emphasised. The result is that the Ministry of Finance's data catalogue is very well used, even though the number of datasets is a modest 25.

There was agreement between the Romanian and Czech speakers that a mechanism is needed to ensure continuous publication of data, including updates, and that this requires a change in culture. But there is some understandable frustration. First of all it is hard to know who is using published data and what for – not knowing this makes it hard to see an end result and to develop that culture change. Promoting open data to the tech community doesn't reach the citizens who have no interest in data but who might be interested in services. It is also frustrating to receive requests for data from people who are unaware that what they have asked for is already freely available.

The second day of the workshop began with a presentation of the situation in Poland from Jacek Wolszczak of the Ministry of Administration and Digitization. An important distinction for that country is between access and re-use. Individuals need to identify themselves when requesting data and access is free of charge, but the re-use may or may not be free of charge. This puts Poland outside the usual definition of open data[1] but it is within the PSI Directive. One eye-catching idea was that of a knowledge base associated with the data portal, that is, documents designed for humans to read rather than data for machines to process. The workshop felt that the provision of such a knowledge base could be regarded as a best practice. Indeed, participants emphasised that Public Sector Information includes such documents, not all of which would be thought of as 'open data' – there is an overlap, but there is also a difference. As with other speakers, Mr Wolszczak was keen to highlight the importance of receiving and acting on community feedback. The feedback received will be published in full.

---

[1] http://opendefinition.org

Branislav Dobrosavljevic presented the work of the Serbian Business Registers Agency. It operates as a one stop shop for many different and long established registers in Serbia and offers a range of services for internal and external users. The focus is very much on services rather than data. Such provision of eServices in Serbia is new as, until recently, the law demanded stamps on paper as part of the processes. Core services are now available from apr.gov.rs which includes free access to information about registered companies, although this is mediated via a search box with no option to download the data in bulk. The site is also not amenable to scraping and so the Serbian register is not available, for example, via OpenCorporates.

Figure 2 Partial screenshot from the Serbian Business Register showing the result of a search for a specific company

Mr Dobrosavljevic extended an invitation to collaborate internationally to ensure better interoperability, and ended by emphasising some key points:

- keep it simple;

- focus on front end services;

- have complete control on services, selling enriched data, not raw data;

- create a consistent market of data so that third party companies are able to run profitable services of their own;

- activities need to be covered by legislation, which is more important than technology.

The European Commission's Szymon Lewandowski gave an update on developments at DG CONNECT. The main issues arising from the revised PSI

Directive concern some requests for clarification of some of the details, including details of the original Directive, and the effect of the revision on running contracts. The launch of the Open Data Incubator (ODINE[2]) and the new contract for the publicdata.eu portal[3] are seen as important components of Europe's big data infrastructure. The latter is foreseen as a single gateway to reusable information with the aim of enabling the combination and visualisation of information held by various open data portals at various levels throughout the EU. The new portal will be a focus for services around open data and include a dedicated service infrastructure for language resources in order facilitate multi-lingual access. This point was picked up by the LIDER project who took part in a Share-PSI workshop for the second time. Mr Lewandowski reported that machine translation technologies will be included in the first version of the new portal at its launch in November 2015 at the European Data Forum although, of course, they will need testing and further development.

A hope for the new portal is that it will include wizards to guide publishers through licensing issues and allow you to combine different datasets with different licences. This implies that licences need to be at least partially machine readable.

The final plenary presentation was by Nicolas Loozen of PwC who carried out some work under the ISA Programme looking at the prioritisation of datasets for publication. That work suggested a series of factors to take into account (Figure 3).

The data owner's perspective	
Transparency	Does the publication of the dataset increase transparency and openness of the government towards its citizens?
Legal obligation	Is there a law that makes open publication mandatory or is there no specific obligation?
Relation to the public task	Is the data the direct result of the primary public task of government or is it a product of a non-essential activity?
Cost reduction	The availability and re-use of a dataset eliminates the need for duplication of data and effort, which reduces costs and increases interoperability.
A re-user's perspective	
Target audience	In terms of size and dynamics
Systems & services	The number of new and existing uses of the data.

Figure 3 PwC's matrix for prioritising datasets for publication

---

[2] http://opendataincubator.eu/

[3] https://www.uk.capgemini.com/news/european-commission-awards-the-deployment-of-the-pan-european-open-data-portal-to-a-consortium

These factors were applied to the European Commission's Tenders Electronic Daily service (TED[4]), which meets all the criteria, and followed up by interviews and a questionnaire with TED users.

That work reinforced the comments made by other plenary speakers and throughout the workshop, that user engagement is an essential aspect of PSI provision. Mr Loozen further raised the issue of collaborative tools – a recurring theme in other sessions.

2.2 Data Quality

One of the best-attended sessions of the workshop was lead by Makx Dekkers. His work with the ISA Programme in the Open Data Support project identified 9 dimensions of quality that might be applied to data.

- **Accuracy:** is the data correctly representing the real-world entity or event?

- **Consistency:** Is the data not containing contradictions?

- **Availability:** Can the data be accessed now and over time?

- **Completeness:** Does the data include all data items representing the entity or event?

- **Conformance:** Is the data following accepted standards?

- **Credibility:** Is the data based on trustworthy sources?

- **Processability:** Is the data machine-readable?

- **Relevance:** Does the data include an appropriate amount of data?

- **Timeliness:** Is the data representing the actual situation and is it published soon enough?

This sparked a good deal of debate (captured more or less fully in the raw notes for this session[5]). One suggestion that found favour was the addition of **context**, that is, the reason the data was collected in the first place. Other topics for discussion were

- the usefulness of the 5 Stars of Linked Open Data scheme as a measure of processability (general agreement that it is useful);

- whether the methods by which the data was collected is relevant as a measure of quality (again, yes);

- the usefulness of schemes like the ODI's Certificates (useful);

- whether an indication that access is open or restricted, or of differences between access and re-use, are part of a quality assessment (dubious).

The topics discussed in this session could perhaps be the basis of a full two day workshop but the interim conclusion was that publishers should unambiguously express and communicate the quality level of their data. This allows potential

---

[4] http://ted.europa.eu
[5] https://www.w3.org/2013/share-psi/wiki/Timisoara/Scribe#How_good_is_good_enough.3F_A_common_language_for_quality.3F

users to make informed decisions on whether and how to re-use the data. A standard set of terms should be developed… which is being done in the W3C Data on the Web Best Practices Working Group of which Mr Dekkers is an active member.

Figure 4 Daniel Pop (for Valentina Dimulescu), Peter Winstanley and Vasile Crăciunescu prepare to give their 1 minute "Come To My Session" pitches, a feature of Share-PSI workshops. W3C's Phil Archer controls the stopwatch.

A specific aspect of quality and re-usability – how to identify locations - was discussed in the session *Free our Maps*, lead by Vasile Crăciunescu and Codrina Maria Ilie of the Technical University of Civil Engineering Bucharest. Licensing is an issue for Open Street Map usage, and few national mapping agencies make their data available for free. Many of the issues raised, such as the choice of spatial vocabulary, how to best represent spatial objects in RDF etc. are being

addressed in the Spatial Data on the Web Working Group[6] in which W3C is collaborating with the Open Geospatial Consortium to produce joint standards, including a best practice guide.

The session concluded that there is a need for a reference data set that covers at least the whole of Europe so that locations can be referred to consistently. The suggestion was that this should be created by the INSPIRE community as an authoritative dataset. The alternative is to use OSM and/or Geonames but this raises issues of quality, authority and licensing.

2.3 Collaboration

Valentina Dimulescu of the Romanian Academic Society lead a session discussing the Romanian Electronic Public Procurement System. As announced by Radu Puchiu, this is being replaced by a new system but the discussion raised several issues. The current system is hard to access and this is clearly deliberate as it is necessary to enter a CAPTCHA code at every step and search is only possible via NACE code[7], not company name. A bulk download is available from the government data portal but apparently the downloadable data is corrupted and unusable.

It was noteworthy that the data didn't shed any light on corruption. Many local businesses in Romania are publicly owned and there is a general policy to favour local businesses, so the fact that many local businesses have public contracts cannot be taken as a sign of corruption. However, private company shareholder information is not made public and so it's hard to make correlations between contracts with private businesses and public officials. In Albania, this information is available and those correlations are clear to see. It's also possible to correlate procurement contracts with companies that make donations to Albanian political parties. Another factor is that, according to Miss Dimulescu, most of the corruption that does happen, takes place *after* the contract has been awarded. A much more complete dataset covering the whole process would be needed to uncover such corruption.

During her work, Miss Dimulescu made many corrections to the data but it wasn't possible to feed those corrections back. One issue she dealt with was simply that the downloaded data was all in single text fields rather than separate ones for company name, address etc. The CKAN software used by the majority of data portals doesn't have a mechanism for providing cleaned up versions of datasets. It was suggested that contributing a software module to CKAN that would support this might be a good future project.

An alternative approach is adopted by the City of Chicago where some datasets are published on GitHub (Figure 5). This was presented by Peter Krantz in his session on crowd sourcing and was seen as an excellent method of gathering corrections and engaging the community.

Crowd sourcing is a prime example of community engagement – the community that wants the data helps to create and manage the data. The problems are usually legal; for example, the person behind a project to crowd source Swedish post codes quickly received a cease and desist notice. However, Chicago provides

---

[6] https://www.w3.org/2015/spatial/

[7] http://ec.europa.eu/competition/mergers/cases/index/nace_all.html

an example where crowd sourcing complements official sources to the benefit of all.
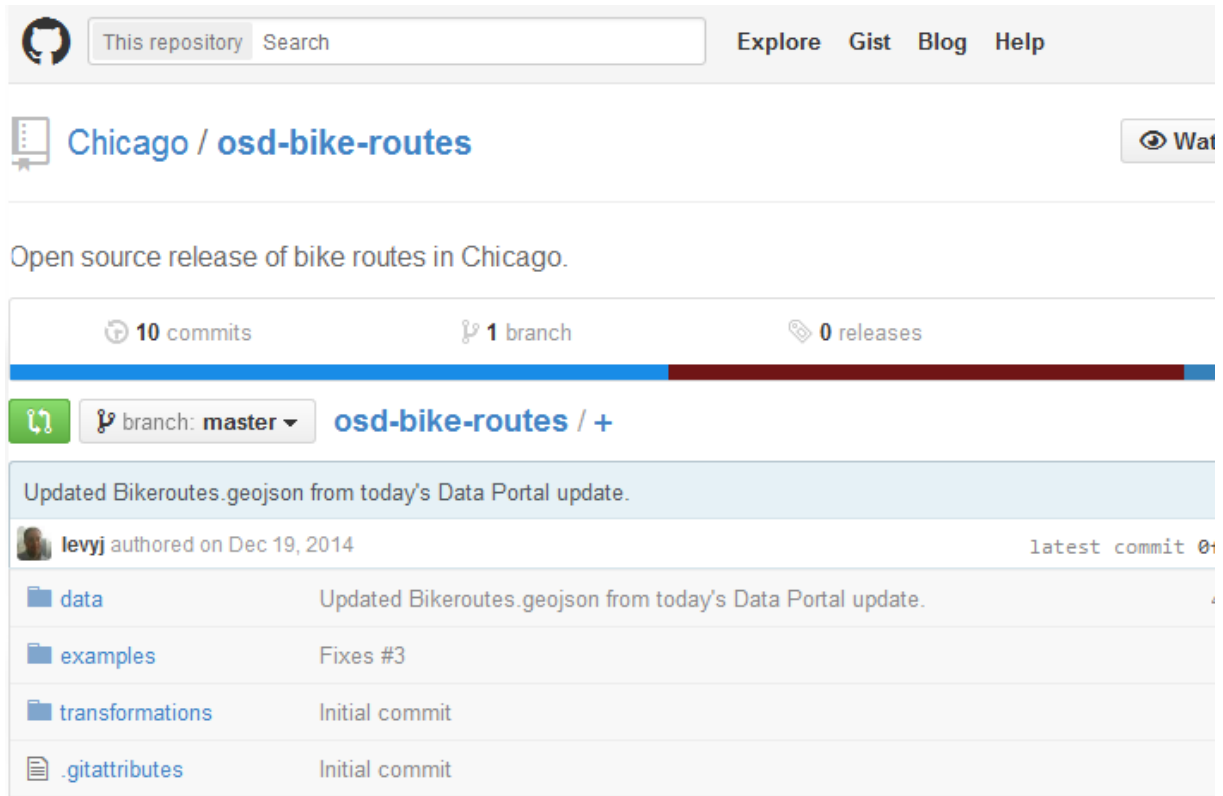
Figure 5 Partial screenshot of a City of Chicago dataset on GitHub with corrections committed by the community. See https://github.com/Chicago

The discussion around crowd sourcing lead to some concrete proposals for best practices.

1. Identify the need first and then seek groups able to support solving that need via crowd sourcing.

2. Think of crowd sourcing as another tool to create/improve data sets and think about the phases of your data collection project and where crowd sourcing could best fit in.

3. Involve stakeholders who could benefit from a free source of certain data sets and have them provide funding in order to sustain crowd sourcing efforts.

4. Minimise the size of each task.

5. Use a gamification approach.

6. Consider using crowdsourcing without the users' knowledge, for example by using CAPTCHA systems.

The last of these is well known in the cultural heritage community where CAPTCHAs can be used to gather human reading of text from scanned documents that OCR software cannot read[8].

---

[8] http://en.wikipedia.org/wiki/ReCAPTCHA

Many of these ideas were reflected and emphasised in the session *Raising awareness and engaging citizens in re-using PSI* lead by Daniel Pop of West University of Timişoara, and Yannis Charalabidis of the University of the Aegean. Engaging end users is essential to ensure that the data made available is the data people want and that it is worth the effort of publishing. The point was made again that end users are not interested in data – but they might be interested in data-driven services, and public authorities need to know if someone is going to do something with the data to justify the effort made.

Engaging citizens requires effort – it is a job in itself to reach out to different members of the community and to respond to requests. One method of doing this that was highlighted is the Karlsruhe City Wiki[9] which is run entirely by the community. Professor Charalabidis offered 5 ways for a community manager to engage users:

- Provide a home - offer the ability to citizens / users to create a profile and login via social media.

- Create an open data marketplace – citizens can put in a request – that is public for everyone to see (this draws on gamification principles).

- Allow users to be publishers - allow for upload of datasets by users.

- Allow working on datasets, i.e. make users curators.

- Provide incentives such as:

    o publishing the popularity of the users;

    o free tickets to community events, free parking etc.

    o 'Datathons' (longer competitions);

    o data journalism competitions.

2.4 Indices

Crowd sourcing is the technique used by both Open Knowledge to create the Global Open Data Index[10] and by ePSI Platform to create the PSI Scoreboard[11]. As Emma Beer from Open Knowledge and Martin Alvarez from ePSI Platform described in their session *How benchmarking tools can stimulate government departments to open up their data*, data submitted by volunteers is then curated and reviewed in a documented process. Some common problems faced were in helping contributors understand the questions they are tasked to answer and subsequently in generating publicity. One aspect of the former problem is multilingualism. Some of the data received is translated using Google Translate – it's a very manual process. On the second point 'UK still top of the table' is not a news story, although France's position as the 'most improved' did generate a lot of coverage.

---

[9] http://ka.stadtwiki.net/

[10] http://index.okfn.org/

[11] http://www.epsiplatform.eu/content/european-psi-scoreboard

Rank	Place	Transport Timetables	Government Budget	Government Spending	Election Results	Company Register	National Map	National Statistics	Legislation	Postcodes / Zipcodes	Pollutant Emissions	Score
1	United Kingdom											97%
2	Denmark											83%
3	France											80%
4	Finland											73%
5	Australia											72%
5	New Zealand											72%
7	Norway											71%

Figure 6 The Global Open Data Index, 2014
http://index.okfn.org/place/

Most of the discussion focussed on the impact that the Index and Scoreboard have. In his opening remarks, Radu Puchiu said that Romania had been pleased to be ranked joint 16th alongside the Netherlands and Iceland and had a target to be in the top 10 in 2015. That's a clear case where the Index is having a positive effect. Silviu Vert of Open Knowledge Romania said that being able to show that the openness of budget and spending data is internationally benchmarked helps make the case against the usual excuses for not publishing. Anne Kauhanen-Simanainen from the Finnish Ministry of Finance said that they were considering what indicators they should use to measure the impact of their open data policies. It was suggested that they look carefully at the indicators used in the Global Index, the PSI Scoreboard and the Web Foundation's Barometer[12] so that comparisons could be made easily.
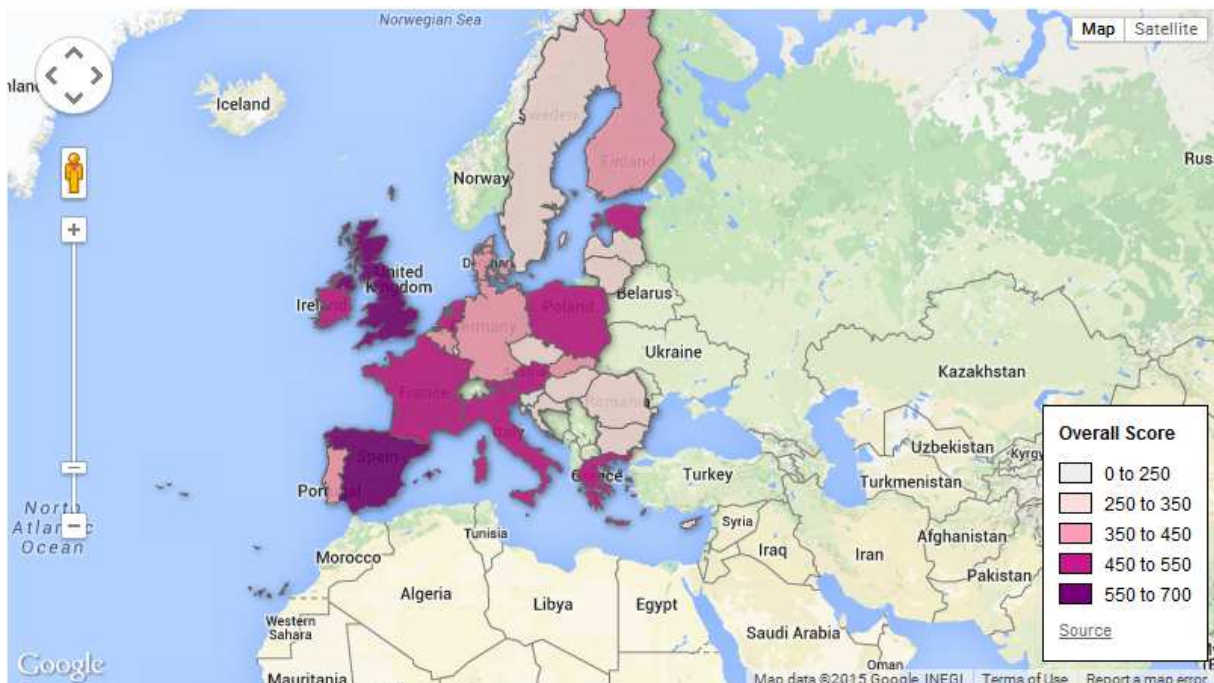

Figure 7 ePSI Scoreboard, 2014
http://www.epsiplatform.eu/content/european-psi-scoreboard

---

[12] http://barometer.opendataresearch.org/

If success in the Global Index is helpful, and if the EC uses the PSI Scoreboard to measure progress across Europe, what of countries at the bottom of the list? Martin Alvarez reported that his attempts to contact governments at the bottom of the Scoreboard had been unsuccessful; with the exception of some individuals the governments seem simply not to care. The low score given to Belgium threatened the continuation of the excellent work done in Flanders since regional efforts are, perhaps unfairly, not reflected in the indices[13].

One way to tackle this, and to address the 'no change is not news' story, would be to increase the number of available comparisons. In particular, countries often judge themselves against their geographical or cultural neighbours more than more distant territories. It's also worth highlighting specific areas in which countries do well. Greece, for example, is one of only a handful of countries to make its spending data available at transaction level and are a clear leader in this regard.

2.5 New Discoveries

There were several sessions in Timişoara that, in one way or another, tackled the issue of data discovery.

In his session, *Site scraping techniques to identify and showcase information in closed formats - How do organisations find out what they already publish?*, Peter Winstanley of the Scottish Government considered the large amounts of data published within documents and websites designed for human readership. In the same way that search engines are able to make sense of unstructured web pages (to a greater or lesser degree), scraping can be used to create at least an inventory of what an organisation has. The Scottish Government Data Labs provides an example of this[14] (Figure 8). It dynamically scrapes the organisation's website to generate lists of various types of document, including keywords etc.

Similar exercises have been carried out elsewhere but the participants agreed that this was only a first step. Such lists don't include licence data, for example, and the manual effort may still be substantial. Martin Alvarez used FOCA[15] to create an inventory of data although that inventory isn't available publicly. It was a way to show the public authorities what they already had. It also showed the value in publishing structured metadata for published documents. In all cases, generating lists through site scraping must be seen as a first step or staging area and not as a substitute for publishing datasets explicitly.

---

[13] http://www.openknowledge.be/2014/12/09/belgium-scores-slightly-higher-on-the-global-open-data-index-big-expectations-for-2015/

[14] http://labs.data.scotland.gov.uk/

[15] https://www.elevenpaths.com/labstools/foca/

TEXT SEARCH:

Search

3065 Items

TOPICS CLOUD:

5-14 Curriculum A-levels and NVQs
Aberdeen City Abuse Accessibility
Accidents Additional Support for Learning
Adoption Advanced Higher grades
Advertising Agricultural land
Agriculture Alcohol use and abuse Anti-
social behaviour and disorder Aquaculture
Architecture Area regeneration Arts
Benefits Biodiversity Biotechnology
Building and construction Buses
Business & Industry Business and
industry Cabinet Cancer Care & Social
Work Care and Social Work Careers
and career development Carers Central
government Central Scotland Police
Cereals Child abuse Child care Child
protection Children Children in care Civil
Justice Civil Law Climate Change

sorted by: types; then by… • ○ grouped as sorted

1 • 2 • 3 … 154   Next »

1. Getting the best from our land A draft land use strategy for Scotland Environmental Report
   Draft land use strategy SEA Environmental Report
   Consultation http://www.berr.gov.uk/files/file48643.xls [http://www.scotland.gov.uk/ID3/190750
   http://www.scotland.gov.uk/Publications/2010/09/30105117/12]
2. BRIAN PACK OBE DOING BETTER INITIATIVE TO REDUCE RED TAPE IN AGRICULTURE

   Consultation https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/205358/auk-
   chapter03-06jun13.xls [http://www.scotland.gov.uk/ID3/290798
   http://www.scotland.gov.uk/Publications/2013/12/4967/12]
3. Energy in Scotland: A Compendium of Scottish Energy Statistics and Information

   Consultation http://www.decc.gov.uk/assets/decc/statistics/source/renewables/et6_2.xls [http://
   www.scotland.gov.uk/ID3/257525
   http://www.scotland.gov.uk/Publications/2012/03/2818/8]
4. Consultation on Draft Technical Guidance for Noise Management Areas, under Strategic Noise Action Plans
   Consultation Document on Draft Technical Guidacne for Noise Management Areas

Figure 8 Partial screenshot of the Scottish Government's Data Labs showing the dynamically generated list of available CSV files

The session on scraping concluded with some concrete proposals for how to proceed:

1. Identify the information assets that are already published on the website by the institution, e.g. by scraping, harvesting, crawling.

2. Identify how the information assets are published (closed formats, open formats), e.g. extracting information from the header, extracting information from RDF representations.

3. Establish usage in a user interface over this retrieved information to create a "staging area."

4. Use the staging area to pre-fill the production-ready catalogue. Use staging area to identify and to monitor the progress of work on information assets that need improvement to have them added to a production-ready catalogue.

Two sessions looked at storage and discovery of scientific research data. The session led by Tamás Gyulai of the Regional Innovation Agency in Szeged, *The Role of Open Data in Research Institutions with International Significance* and Robert Ulrich's bar camp on re3data.org both considered similar issues. Researchers are being encouraged, in some cases forced, to publish the data that underpins their work. In some disciplines, such as astronomy and biochemistry, this is already part of the culture but in others, such as the social sciences, publishing data goes against that culture. There is no separate finance available for publishing data and careers are built on published papers, not published datasets – at least, that is the situation today. Changing the landscape so that the incentives and rewards for publishing data are equal to those for publishing papers will be an important change in the culture among researchers in all disciplines.

One engine of change is that funders are increasingly asking for descriptions of how researchers plan to publish their data to be included in the proposal/grant application. Hungary's SZTAKI is experimenting with journal publications that

include the paper, the data and the algorithm used so that experiments can be re-run and results reproduced.

Increasing the number of people with the skills necessary to publish research data in a re-usable manner must be an important target within education and policy. Alongside this, infrastructures for publishing and archiving research data need to be established.

re3data.org is an effort to provide information about the rapidly growing number of data repositories. It publishes information about more than 1,000 such repositories, making it easier for researchers to identify a repository suitable for their own work. Initially established through a collaboration of several German institutions, re3data.org will be managed as part of DataCite by the end of 2015. That organisation (DataCite[16]) is part of the ecosystem around journals, researchers and citations that is already established among large sections of the academic community.

2.6 Legal Matters

The Open Science Link project[17] is also working on new models for publishing scientific information including papers and their associated data. As part of this project, Freyja van den Boom of KU Leuven lead a session discussing the European Database Directive. There is no copyright on facts and so collections of facts – databases – are not protected. However, the 1996 Database Directive recognises the investment necessary to create databases. These are the *sui generis* rights. The problem is that the Directive is applied inconsistently across the EU. In one case in Germany, 40 people employed to maintain the database was sufficient evidence that the investment was substantial, in another case, 500 workers was not.

The situation is very unclear, especially in relation to access versus bulk download, individual versus institutional re-use, state-funded but privately created data and so on. The Database Directive makes no distinction between publicly owned and privately owned data so there are some cases where the PSI Directive overrides the Database Directive (this is true if the public authority owns the *sui generis* rights). Georg Hittmair of Compass described how his company had created an electronic business register from paper records for many years. When the Austrian government created their own electronic register in 1999, Compass copied that data (which it already had, it was simply a different way of accessing it). The Database Directive meant that Compass was no longer able to resell the data. It went all the way to the European Court – and went in the Austrian government's favour.

In Finland, if a public body owns a database and makes it available for free, without having to register, even anonymously, then they have effectively given permission to re-use. This is seen by many as an effective 'right to scrape' in Finland.

The conclusion of the session was that harmonisation is necessary across the EU, both in how the Database Directive is implemented and the relationship with the PSI Directive. In the meantime, agreement on licensing would be helpful, and in

---

[16] https://www.datacite.org/
[17] http://opensciencelink.eu/

that regard, Creative Commons licences are a good option. Machine readable licences and rights statements would also help.

The topic of machine readable licences came up yet again in the session on multilingual data. An aspect of data is the language in which it is expressed and multilingualism must be a part of any European data infrastructure. Felix Sasaki presented his work in the LIDER project and advocated the use of Linguistic Linked Data as a bridge to reach a global audience. Agreed vocabularies, standardised APIs and links to other resources are all important building blocks but an outcome of LIDER will be Linguistic Linked Licensed Data (3LD). This is designed to produce language resources using standard data models along with something called repeatedly throughout the Timişoara workshop – machine readable licences.

Taken together, these techniques can lead to making data available in the language in which the potential users want it, even if it is not the original language.

Figure 9 Participants had plenty of time for networking and were well looked after.

3 Conclusion

The event in Timişoara was successful in engaging stakeholders in the PSI and open data landscape from across Europe, including countries that are often under represented in these discussions. The city of Timişoara is a leader in this field within Romania, and the participation of the country's top civil servant responsible for open data policy, validated the choice of location.

The main conclusions of the workshop were:

- Citizen engagement is essential in creating an ecosystem around PSI publication and use.

- The technical community is important, but not the only important community with which to engage.

- For end users, services are what counts, not data.

- The release of further data is best incentivised by seeing use of what's already available.

- The legal landscape around databases, crowd sourcing, access and re-use is far from clear.

- There is a need for machine readable licences/rights statements.

- Users of data should be empowered to curate and correct datasets to the benefit of all.

- There is a need to describe the quality of data in a consistent manner if potential consumers are to make informed choices.

Annex 1 - Agenda

Opening Plenary

Introduced by: Dana Petcu, West University, Timişoara.

Welcome: Prof. Univ. Dr. Marilen Pirtea, Rector of West University of Timişoara

Radu Puchiu, Secretary of State, Chancellery of the Prime-Minister (Romania)

Experiences of identifying datasets for sharing, Benedikt Kotmel, Ministry of Finance (Czech Republic)

Capturing Best Practices, Chris Harding, The Open Group (Chris will outline what we need to capture from each session)

Parallel Sessions A

Share-PSI Track

Site scraping techniques to identify and showcase information in closed formats - How do organisations find out what they already publish?

Facilitator: Peter Winstanley, Scottish Government

This session addresses the question of how organisations that already publish considerable amounts of information on their website but in non-interoperable formats such as Excel and PDF might 'discover' what they are publishing and present it in various helpful ways to end users (including the organisation's own staff) as part of the engagement to discover priorities for open data publication. Illustrations from site scraping of Scottish Government and NHS Scotland will be presented.

Many government and public sector bodies already publish a considerable amount of information including data and reports on their websites. However, this is frequently done under a distributed management process and using content management systems, both of which tend to militate against being able to present in a quick and flexible way the assets of any particular publication format. As a consequence, organisations might find it challenging to know where to start when establishing a programme of converting existing information resources that are not in open formats (1-2 stars on the 5 star model) to more open formats.

Open Track

The Electronic Public Procurement System, open data and story telling in Romania

Facilitator: Valentina Dimulescu, Romanian Academic Society

Gaining access and managing public procurement information in Romania by third parties is a strenuous activity. Although the Romanian Government created the online portal under the European Open Data initiative - Digital Agenda for Europe, which includes a section on public procurement, upon accessing public

datasets the question arises whether the information provided are affected by human error or malice.

After continuous failed attempts to acquire a database containing complete information on various types of public procurement contracts, The Romanian Academic Society (RAS), a Romanian think tank, concluded that the only way to get systematic access to this type of data is to connect directly to the Romanian Public Procurement Electronic System's (SEAP) server so as to copy the available information. The authors have encountered two major challenges:

1. to assemble all the data in a consistent database;

2. matching the errata notices to award notices.

Both information collected directly from SEAP and those from CSV files provided by the Government under open data rules contain obvious errors which refer to, among others, the economic agent's country field or absurdly low or high prices. The only manner in which these errors can be corrected is to connect the so-called "errata notices" to its respective award notice. SEAP errata notices containing modifications of errors have not been applied to public procurement open data.

The authors recommend that the newly envisaged online public procurement system (SICAP), financed through European Union development funds, should assume an export module and incorporate standardized errata information in order to correctly export the data base. Still, the publication of this specific Public Sector Information is sensitive, as many corruption cases arise from public procurement contracts.

Open Track

Free Our Maps

Facilitator: Vasile Crăciunescu, Codrina Maria Ilie, Technical University of Civil Engineering Bucharest.

Our session is dedicated to the importance of releasing public geodata over the Internet, under an open license and in a reusable format. Geodata is a broad term that refers to data that has a spatial component, defined through various methods, such as pairs of coordinates, name of location, address identifiers and so on. Its usage is wide spread over various domains. Even though world leading business, such as Google, Yahoo, Nokia, Apple and more, have developed services and products that have ultimately and permanently changed the way in which geodata is perceived by the wider community, such as Google Maps; even though the community itself stepped up, building an international network that, in a collaborative, volunteer and open manner, continuously works to build an open map of the world, OpenStreetMap, we do consider that, that there is an immense untapped resource of geospatial information.

That resource is represented by the databases of national agencies and institutions that have produced and collected data within national monitoring networks and research projects for an extensive period of time. For the society, to harvest in the most productive way the benefits of open public geodata, some matters need to be discussed, such as: quality and relevance, different angles of open geodata understanding: public sector, private sector and academia, bridging community driven data with public data, the impact of INSPIRE Directive

to open geodata movement, geodata licenses interoperability and technical issues on releasing public geodata as open data.

Day 2 Plenary

Chair: Heather Broomfield, Difi

Jacek Wolszczak, Ministry of Administration and Digitization (Poland)

Branislav Dobrosavljevic, Business Registers Agency (Serbia)

Szymon Lewandowski, European Commission

Good practices for identifying high value datasets and engaging with re-users: the case of public tendering data, Nicolas Loozen, PwC EU Services

---

Parallel Sessions B

Share-PSI Track

How good is good enough? A common language for quality?

Facilitator: Makx Dekkers, AMI Consult

This session will look at the requirements and possible solutions for defining, measuring, expressing and communicating quality of published Public Sector Information. It is the intention that the outcome of the session will be submitted to W3C's Data on the Web Best Practices Working Group to inform the development of the Data Quality vocabulary.

There is a lot of talk about the need to publish "high-quality" PSI. While it is certainly important that data has sufficient quality to make it useful and usable for users and re-users, we currently lack a common or standard way to express what the quality of data is. The question is whether it is necessary to have such a common way, and, if so, what a "quality vocabulary" could look like.

Open Track

The European Database Directive

Facilitator: Freyja van den Boom, KU Leuven

The European Database Directive is the key legal instrument when dealing with various databases of open scientific and raw data. This Directive harmonises the treatment of databases under copyright law and creates a new sui generis right for the creators of databases which do not qualify for copyright.

According to Article 3 of the Database Directive, for a database to receive legal protection, it must be 'original', i.e. the author's 'own intellectual creation' by reason of the selection or arrangement of the contents.1 This level of 'originality' is the same as in Article 1 (3) of the Software Directive and Article 6 of the Terms of Protection Directive. Considerable variety exists in the national Courts' approaches to the requirement of originality. Whether collections of scientific research data will meet the criterion of 'originality' is a question that will be dealt

with on a case-by-case basis. It depends on the interpretation of each national Court.

If the database qualifies for copyright protection under the Directive, the copyright-holder will hold 'exclusive rights' in respect to that data.

Article 5 of the Directive enumerates those 'exclusive rights'. The author of the work shall have the exclusive right to carry out or authorise:

A. Temporary or permanent reproduction by any means, in any form, in whole or in part;

B. Rights of adaptation, translation, arrangement and any other alteration;

C. Any form of distribution to the public of the database or of copies thereof (subject to Community exhaustion); and

D. Any communication to the public, display or performance to the public;

E. Any reproduction, distribution, communication, display or performance to the public of the results of the acts referred to in (b).

In all Member States of the Union, an exception exists for "all acts, which are necessary to obtain access to the contents of the database and to obtain normal use of the contents by the lawful user". This also applies to a part of the contents of the database.

Member States are also free to apply four exhaustive other exceptions to the 'exclusive rights' listed above. The possible exceptions are listed in article 6 (2):

• Reproduction for private purposes of a non-electronic database;

• Illustrative uses for teaching or scientific purposes as long as there is proper attribution and justification for this purpose;

• Public security, administrative or judicial procedure; and

• Other exceptions traditionally authorised in the Member State.

Note that unauthorized copying for private purposes is not permitted for digital databases.

How has the SGRDdirective been implemented in the different member states: share experiences.

How do these national differences affect the ability to (crossborder) re-use PSI

What would be best practices with respect to licensing and disclaimers?

These are some of the questions I hope to address during this session.

Open Track

Role of Open Data in Research Institutions with International Significance

Facilitator: Tamás Gyulai, Regional Innovation Agency

Szeged has been known in Hungary as a central location for open software development and utilisation: the municipality of Szeged was among the first town administrations in Hungary to use open software in large applications and the University of Szeged is an acknowledged development center of open software solutions in Hungary.

The Regional Innovation Agency (RIA) of the South Great Plain region has been a promoter of innovation in the key thematic areas of the region, including IT development, as well. Several cluster initiatives have been implemented in the course of the years in cooperation with partners in the neighbouring regions, including also Timisoara and especially Tehimpuls Association and its professional partners. One of the most successful initiatives was the Cluster2Success project where Romanian and Hungarian IT companies met several times with the objective to work out new innovative solutions together.

One of the actual challenges that the activities of the RIA have focus on is the development of IT background of the Extreme Light Infrastructure (ELI) project as it will be a key infrastructure for research and development not only in Szeged but also in the wider region. The experiments that the researchers will conduct at the ELI facility will produce data in enormous quantities that shall be analysed and processed by international teams of researchers therefore the newest and most advanced „big data" software and hardware solutions shall be used here.

As the ELI is co-financed by the European Union, the research facility will be a public institution and therefore they shall have an open policy of information. On the other hand, some of the experiments might lead to patented inventions therefore the management of information about the research activites shall respect also the intellectual property rigths (IPR) considerations, as well.

The main local stakeholders in Szeged are all committed to the successful implementation and operation of the ELI as a key element of the scientific and economic life in the town. The cooperation among them shall be extended also to the sharing of data with the objective to make an open system that is accessible also to foreign partners. Consequently, the Share-PSI workhop can be an excellent event for meeting professional people that have already been confronted with similar challenges. It might lead to common solutions that can be designed and can be tested later in real life in Szeged.

LIDER Track

Linguistic Linked Data as a bridge to reach a global audience

Presentation: Asunción Gómez-Pérez, Universidad Politécnica de Madrid and coordinador of LIDER project

This presentation will introduce the notion of Linguistic Linked Data (LLD): linked data sets that can play a crucial role in making data on the Web multilingual. LLD can help PSI providers to engage directly with users around the world.

We will discuss what LLD data sets are already available, which ones should have a high priority for you, and what needs to happen to make your data multilingual.

Parallel Sessions C

Crowd sourcing alternatives to government data – how should governments respond?

Facilitator: Peter Krantz

This session addresses the question of how public authorities can/should respond to community efforts to crowd source data that replicates official data that is not open (e.g. post code and address data). The session will start with a brief case study of how crowd sourcing initiatives of post code and address data in Sweden evolved, and the response by agencies. The session will also touch on data quality aspects of crowd sourcing initiatives.

In many areas governments have a monopoly on high quality PSI, typically by regulation for its creation, maintenance and distribution. For types of data that are used in many scenarios, e.g. geodata, there may be a sufficient number of potential users that are excluded by expensive access to government data. In these areas crowd sourcing initiatives may be able to create alternative datasets that compete with those provided by governments. There are already several initiatives, e.g. OpenStreetMap, that are good enough to make even large companies stop buying government data. The outcome of these initiatives may disturb the market for government data while at the same time contribute to lower quality services given the data is not of the same quality as that from government agencies. Governments need to find a way to deal with these issues in a way that serves society, but responses typically include legal action.

Raising awareness and engaging citizens in re-using PSI

Facilitator: Daniel Pop, West University of Timisoara, Yannis Charalabidis, University of the Aegean.

Governments have been investing in publishing considerable amount of data and in modernisation of administration through e-Government services for, in some cases, more than 5 years. A legitimate question is "What is the impact on citizens, or more generally speaking, in reusing available electronic data and service?" There are different actions, initiatives, platforms that can be used to raise citizens (reusers) awareness on existing PSI and engage them in usage.

For example, open data hackathons are a widely spread tool to raise awareness on data published in data.gov.* portals. Semantically enriched platforms, such as ENGAGE, enables not only reusage but feedback collection as well. What other alternative for raising awareness on open data repositories have you been using in your case? What are preferred feedback channels (e.g. social media) in your case?

Public Sector Advertising is also frequently used to raise awareness and engage citizens in re-using PSI. For example, local networks of (interactive) devices (public displays, Smart TVs, Info kiosks etc.) have been deployed by local/regional governments to cover 'hot points' at city/region level. Information on these networks is either managed by means of on-premises (locally installed) software packages or they can be operated by Cloud based, Web-enabled platforms, such as SEED.

We are planning to start our discussion by sharing our experiences and outcomes emerged out of two EC-funded projects (ENGAGE and SEED) and we'll be happy to hear from you what initiatives did you put in place and how these levelled your expectations.

The session will address the following questions:

A. How can public bodies engage the potential reusers of their data and/or services? Methodologies, channels, technical platforms have been used?

B. What methods are available to reusers of your data to send feedback about published data?

C. How do you handle feedback received so as to improve your data?

<div style="background-color:#d84315;color:white;padding:4px">Share-PSI Track</div>

How benchmarking tools can stimulate government departments to open up their data

Facilitator: Emma Beer, Open Knowledge, Martin Alvarez, ePSI Platform Advisory Board

Open Knowledge published the 2014 Global Open Data Index which shows that whilst there has been some progress, most governments are still not providing key information in an accessible form to their citizens and businesses. With recent estimates from McKinsey and others putting the potential benefits of open data at over $1 trillion, slow progress risks the loss of a major opportunity.

The Index ranks countries based on the availability and accessibility of information in ten key areas, including government spending, election results, transport timetables, and pollution levels. The UK topped the 2014 Index retaining its pole position with an overall score of 96%, closely followed by Denmark and then France at number 3 up from 12th last year.

Francis Maude, Minister for the UK Cabinet Office and responsible for the UK open data agenda, said:

> *We have called for people to hold our feet to the fire and the Open Data Index is a great tool for doing just that.*

In this session the project manager of the Index for 2014 will share some of the successes in stimulating governments' to take action to open up further datasets.

After the discussion, and in order to complement this session, Martin Alvarez (Advisory Board at ePSI Platform) will introduce the PSI Scoreboard. This scoreboard is a 'crowdsourced' web tool published on ePSI Platform used by the European Commission as reference for their metrics. This is 'yet another index' to measure the status of Open Data and PSI re-use throughout the EU. It does NOT monitor government policies, but aims to assess the overall PSI re-use situation in the EU28, including the open data community's activities.

This PSI Scoreboard could be enhanced including new indicators, one of them could be the level of openness published by the Global Open Data Index in each specific country. Attendees can decide if this is interesting or not, as well as the feasible technical mechanisms to do it (automatically).

Your requirements for reaching a global audience with PSI data

Facilitator: Asunción Gómez-Pérez, Universidad Politécnica de Madrid and coordinador of LIDER project

In this session we will discuss a goal everybody has: your data wants to reach a global audience. If prioritization of data sets takes the current state into account, this aim fails: most of PSI data sets are monolingual.

The aim of this session is to understand your priorities: what data sets do you want to be multilingual? What (technical, organizational, other) obstacles do you see in achieving multilingual data sets? What business value and usage scenarios are of high priority for you that would benefit from multilingual PSI?

The outcome of this session will feed directly into activities of the LIDER project, which is building a community around linguistic linked data - an important ingredient for making your data multilingual.

Bar Camp

Time keeper: Noël Van Herreweghe

Pitch your idea for an afternoon session in 60 seconds or less, then take your group to an available space. Remember to appoint a scribe. Please let Phil Archer know the title of your session as soon as convenient.

1. Robert Ulrich, re3data.org - making research data repositories discoverable
2. The Pan European Data Portal - Early Wireframes, Philip Millard (via Skype) and Jens Klessmann
3. The Critical Success Factors Taxonomy for Open Data, Yannis Charalabidis
4. Government as a developer (to identify and open data), André Lapa
5. What do you want from the W3C Data on the Web Best Practices? Phil Archer

Annex 2 – workshop participants list

1. Ira Alanko, Ministry of Finance, Finland
2. Martin Alvarez-Espinar, CTIC
3. Phil Archer, W3C/ERCIM
4. Øystein Åsnes, Difi
5. Jospeh Azzopardi, MITA
6. Adina Barbulescu, Faculty of Economics and Business Administration, West University of Timişoara
7. Kai Barkowsky,]init[
8. Emma Beer, Open Knowledge
9. Angela Benga, Counsellor, Chancellery of the Prime-Minister
10. Petya Bozhkova, Balkan Services
11. Aranita Brahaj, AIS
12. Heather Broomfield, Difi
13. Dan Bugariu, Smart City Association
14. Lorenzo Canova, Politecnico di Torino
15. Edgars Celms, IMCS
16. Yannis Charalabidis, University of the Aegean
17. Gabriele Ciasullo, AgID
18. Adela Suzana Colţa, „Tibiscus" University of Timișoara
19. Vasile Crăciunescu, Technical University of Civil Engineering, Bucharest
20. Doina Danaiata, West University of Timişoara
21. Makx Dekkers, AMI Consult
22. Ivaylo Dobrev, Balkan Services
23. Branislav Dobrosavljević (Data services manager), Serbian Business Registers Agency
24. Tamás Gyulai, Regional Innovation Agency
25. Asunción Gómez Pérez, UPM/LIDER
26. Chris Harding, The Open Group
27. Dolores Hernandez, MINHAP
28. Georg Hittmair, Compass/PSI Alliance
29. Johann Höchtl, Danube University Krems
30. Codrina Ilie, Groundwater Engineering Research Center, CCIAS
31. Gabriel Iuhasz, West University of Timişoara
32. Valentina Janev, The Mihajlo Pupin Institute
33. Benedikt Kämpgen, KIT

34. Anne Kauhanen-Simanainen, Ministry of Finance, Finland

35. Jens Klessmann, Fraunhofer FOKUS

36. Pekka Koponen, Forum Virium Helsinki

37. Benedikt Kotmel, Ministry of Finance, Czech Republic

38. László Kovács, SZTAKI

39. Peter Krantz

40. Jan Kučera, University of Econsomics, Prague

41. Denis Kurunczi, MSI

42. Lado Langof, Commission for the Prevention of Corruption, Slovenia

43. Ira Alanko, Ministry of Finance, Finland

44. Martin Alvarez-Espinar, CTIC

45. Phil Archer, W3C/ERCIM

46. Øystein Åsnes, Difi

47. Jospeh Azzopardi, MITA

48. Adina Barbulescu, Faculty of Economics and Business Administration, West University of Timişoara

49. Kai Barkowsky,]init[

50. Emma Beer, Open Knowledge

51. Angela Benga, Counsellor, Chancellery of the Prime-Minister

52. Petya Bozhkova, Balkan Services

53. Aranita Brahaj, AIS

54. Heather Broomfield, Difi

55. Dan Bugariu, Smart City Association

56. Lorenzo Canova, Politecnico di Torino

57. Edgars Celms, IMCS

58. Yannis Charalabidis, University of the Aegean

59. Gabriele Ciasullo, AgID

60. Adela Suzana Colţa, „Tibiscus" University of Timișoara

61. Vasile Crăciunescu, Technical University of Civil Engineering, Bucharest

62. Doina Danaiata, West University of Timişoara

63. Makx Dekkers, AMI Consult

64. Ivaylo Dobrev, Balkan Services

65. Branislav Dobrosavljević (Data services manager), Serbian Business Registers Agency

66. Tamás Gyulai, Regional Innovation Agency

67. Asunción Gómez Pérez, UPM/LIDER

68. Chris Harding, The Open Group

69.Dolores Hernandez, MINHAP

70.Georg Hittmair, Compass/PSI Alliance

71.Johann Höchtl, Danube University Krems

72.Codrina Ilie, Groundwater Engineering Research Center, CCIAS

73.Gabriel Iuhasz, West University of Timişoara

74.Valentina Janev, The Mihajlo Pupin Institute

75.Benedikt Kämpgen, KIT

76.Anne Kauhanen-Simanainen, Ministry of Finance, Finland

77.Jens Klessmann, Fraunhofer FOKUS

78.Pekka Koponen, Forum Virium Helsinki

79.Benedikt Kotmel, Ministry of Finance, Czech Republic

80.László Kovács, SZTAKI

81.Peter Krantz

82.Jan Kučera, University of Econsomics, Prague

83.Denis Kurunczi, MSI

84.Lado Langof, Commission for the Prevention of Corruption, Slovenia

85.André Lapa, AMA

86.Szymon Lewandowski, European Commission

87.Renārs Liepiņš, IMCS

88.Oană Liviu, West University of Timişoara

89.Oana-Ramona Lobont, West University of Timişoara

90.Nicolas Loozen, PwC Belgium

91.Andrew McKenzie, Digital Birmingham

92.Mihai Adrian Maghiar, West University of Timişoara - FMI

93.Camelia Margea, West University of Timişoara

94.Matjaž Mešnjak, Commission for the Prevention of Corruption, Slovenia

95.András Micsik, SZTAKI

96.Vuk Mijović, The Mihajlo Pupin Institute

97.Uros Milošević, The Mihajlo Pupin Institute

98.Marian Neagul, Institute eAustria

99.Marcel Torok-Oance, West University of Timişoara

100.   Andrei Nicoară, Counsellor, Chancellery of the Prime-Minister

101.   Larisa Panait, Counsellor, Chancellery of the Prime-Minister

102.   Thodoris Papadopoulos, MAREG

103.   Dana Petcu, West University of Timişoara

104.   Daniel Pop, West University of Timişoara

105.   Catalina Popa, Association for Quality of Life, Timişoara

106.   Milan Popadic, Cube Risk Management Solutions

107.   Gaetan Poulin, Ministère des Transports du Québec

108.   Radu Puchiu, Secretary of State, Chancellery of the Prime-Minister

109.   Marko Radojicic, Cube Risk Management Solutions

110.   José-Luis Roda, University of La Laguna

111.   Julien Rossi, University of Szeged

112.   Danica Saponja, Ministry of Public Administration, Slovenia

113.   Felix Sasaki, DFKI/W3C/LIDER

114.   Ingo Simonis, OGC

115.   Margit Suurhasko, Ministry of Finance, Finland

116.   Džiugas Tornau, UALB/Graphity

117.   Slim Turki, Henri Tudor Research

118.   Robert Ulrich, KIT

119.   Alejandro Valdivia, Pentium SA

120.   Silviu Vert, Politehnica University Timişoara

121.   Noël Van Herreweghe, CORVe

122.   Neven Vrček, University of Zagreb

123.   Tomáš Vyhnánek, Ministry of Finance, Czech Republic

124.   Peter Winstanley, Scottish Government

125.   Jacek Wolszczak, Ministry of Administration and Digitization

Annex 3 - Dissemination Activities for Timişoara Workshop

Date	Action taken	Partner
2015-02-24	Email announcement to OGC Membership	OGCE
2014-12-09	[Presentation to Polish Government + stakeholders in Warsaw](#)	CTIC
2014-12-10	[Presentation to Lithuanian Government in Vilnius](#)	CTIC
2014-12-12	[Presentation to Estonian Government in Tallinn](#)	CTIC
2014-12-18	[Added to W3C Endorsed Events listing](#)	W3C
2014-12-20	[Published on ePSI Platform](#)	CTIC
2015-01-07	[Tweet from Share-PSI account](#)	W3C
2015-01-12	[Tweet](#)	DUK
2015-01-12	Promoted (retweeted) [Twitter message](#) originally by DUK about workshop call for participation	IMCS
2015-01-12	Promoted (retweeted) [Twitter message](#) originally by DUK about workshop call for participation	CTIC
2015-01-20	[OKFN-CZ mailing list](#)	UEP
2015-01-20	[COMSODE project](#) mailing list (private)	UEP
2015-01-27	[Tweet from Share-PSI account](#)	W3C
2015-01-28	Tweet from OKFN account	OKFN
2015-01-28	[Tweet](#)	DUK
2015-01-28	Promoted (retweeted) [Twitter message](#) originally by OKFN about workshop submission deadline	IMCS
2015-01-28	Post to LinkedIn Groups EGOV researcher community, European Data Forum, Future of Government, Major Cities of Europe IT Users Group, Open Data Research Network, Open Data Support, Open Government Europe, PSI4PROFIT, SEMIC, Share-PSI 2.0, W3C eGovernment Interest Group, eGov Community	DUK
2014-12-25	Promotion of the event at Serbian government institutions (Serbian Business Registers Agency, Serbian Statistical Office)	IMP
2015-1- 2015-2-	Mails to government and municipal contacts in Latvia	IMCS
2015-1-	Promotion of the event through the mailing lists (lod2, GeoKnow), Retweeted DUK	IMP
2015-2-	Mails to professional contacts in Bulgaria, Bosna and	IMP

Herzegovina, Serbia

2015-3-09	Promotion of the event at the ICIST conference, Kopaonik, Serbia	IMP
2015-02-05	[West University of Timisoara Newsletter](#)	WUT
2015-02-06	[Tweet from Share-PSI account](#)	W3C
2015-2-17	Promoted event at semantic wiki community [1]	KIT
2015-2-17	Promoted event at Publishing Statistical Data community [2]	KIT
2015-2-17	Promoted Lisbon workshop and Share-PSI to people from re3data.org project [3]. The project fosters the publication of research data which fits well with Share-PSI objectives.	KIT
2015-02-19	[Tweet from Share-PSI account](#)	W3C
2015-02-24	[Blog post](#)	CTIC
2015-02-24	[Tweet on registration](#)	CTIC
2015-02-24	[Tweet](#)	CTIC
2015-02-24	[Tweet from SEMICeu account](#)	PwC
2015-02-24	[Publication on Joinup](#)	PwC
2015-02-24	[Slovenian Open Data portal](#)	MNZ
2015-02-25	Open Data Coalition Romania mailing list	WUT
2015-02-25	West University of Timisoara internal mailing list, SEED project mailing list	WUT
2015-02-27	Tweet from OKFN account	OKFN
2015-03-03	Tweet from OKFN account	OKFN
2015-03-04	Share PSI 2.0 will be mentioned within a presentation for international patent information experts at the European Patent Office Raw Data Day in Vienna on 18/03/2015, http://www.epo.org/learning-events/events/conferences/raw-data/programme_de.html	]init[
2015-02-18	Invitation during presentation at [BOK meetup](#)	SZTAKI
2015-02-26	Personal contact to Ministry of National Development	SZTAKI
2015-03-05	News in the website of the Agency for Digital Italy [4]	AgID
2015-03-06	Retweets by W3C Hungarian Office and Andras Micsik [5]	SZTAKI
2015-03-06	Facebook event in [opendata.hu group](#)	SZTAKI
2015-03-06	Press release translation as news item on w3c.hu [6]	SZTAKI

2015-03-06	Tweet from the RNDT account [7]	AgID
2015-03-06	News in the Italian Catalogue for Spatial Data (RNDT) [8]	AgID
2015-03-06	[http://www.agid.gov.it/notizie/agid-al-workshop-europeo-sullopen-data-share-psi-20-0 AgID's institutional web site	AgID
2015-03-06	posted a note on our Open Platform 3.0 Forum web page, sent mail to Forum members, and re-tweeted the tweet	TOG
2015-03-09	Invitation post to members of BOK meetup	SZTAKI
2015-03-07	"Open Data and Big Data–The Impact on Digital Society and Smart Cities" [9]	ULL
2015-03-01	Retweets [10]	ULL
2015-03-12	Post on Italian Open Data community - Spaghetti Open Data [11]	POLITO
2015-03-12	Facebook Master GIS	WUT
2015-03-16	Press release on Radio Timsioara	WUT
2015-03-16	Press release on Press Alert	WUT
2015-03-16	Press release on Tion	WUT
2015-03-16	Press release on Ora de Timis	WUT
2015-03-16	Press release on Sursa de Vest	WUT
2015-03-16	Press release on Pentru Timsioara	WUT
2015-03-16	Press release on Banatul meu	WUT
2015-03-16	Press release on Ziua de Vest	WUT
2015-03-16	Press release on Timis Plus	WUT
2015-03-16	Press release on Timsioara Online	WUT
2015-03-16	Press release on Radio West	WUT
2015-03	Published on Interoperability Solutions for European Public Administration Platform	PwC
2015-03	Spanish Government's PAe	CTIC
2015-03	Rede Comum de Conhecimento	AMA
2015-03-16 & 2015-03-17	The link to the workshop live streaming was shared on Twitter and on LinkedIn. The tweets related to the event were favorited and/or retweeted to increase their visibility.	TUDOR
2015-03-17	Tweet from the SEMICeu account	PwC

Annex 4 – Press Release

A press release was published in both English and Romanian, served via content negotiation at http://www.w3.org/2013/share-psi/workshop/Timisoara/press

English

Government data experts from across Europe gather in Timişoara

6 March 2015 — Government representatives from Romania, the Czech Republic, Poland and Serbia will be among those discussing Open Data Priorities and Engagement at the West University of Timişoara, Romania, 16 – 17 March. The event is the third in a workshop series organised by the EC-funded Share-PSI Network that brings together governments, academics, citizens, groups, commercial companies and standards bodies from across Europe developing best practice guidance on the sharing of public sector information and open data.

As initiatives like the Open Knowledge's Global Open Data Index, the Web Foundation's Open Data Barometer and the ePSI Platform Scoreboard show, many European countries are at the forefront of the global trend to make the information held by governments more readily available. It boosts public sector efficiency, increases transparency and drives innovation. However, the current economic climate means that governments must ensure that tax payers' money is spent as efficiently as possible. This workshop will gather experiences as data holders prioritise their datasets for publication and engage directly with users, whether those users are private individuals, commercial companies or public sector bodies, part of the same organisation or external to it.

The lead of the Share-PSI team within the Department of Computer Science - Faculty of Mathematics and Computer Science, West University of Timisoara - and local chair of the workshop, Prof. Dr. Dana Petcu emphasizes "Timisoara is a centre of academic excellence and business development and this event will show how those skills can be applied in a rapidly growing and exciting market across Europe."

As with all Share-PSI workshops, the aim is to identify best practices in these areas with a focus on Public Sector Information, although cultural heritage, commercial and scientific data are also relevant. Those best practices are being codified by the project itself as well as through international standards bodies including the World Wide Web Consortium (W3C), Open Geospatial Consortium (OGC) and the Open Group.

Română

6 martie 2015 — Reprezentanți guvernamentali din România, Cehia, Polonia și Serbia se vor număra printre participanții care vor discuta despre Open Data Priorities and Engagement (Prioritizarea și angajamentul datelor deschise) la Universitatea de Vest din Timișoara în perioada 16-17 martie 2015. Evenimentul este cel de-al treilea workshop, dintr-o serie de cinci, organizat în cadrul proiectului Share-PSI Network, proiect finanțat de către Comisia Europeană și care își propune să aducă împreună guvernele, lumea academică, societatea civilă, companiile comerciale și organizațiile de standardizare din întreaga Europă pentru a dezvolta împreună ghiduri de bune practici pentru partajarea informațiilor publice și publicarea datelor deschise.

După cum se arată în mai multe rapoarte recente, cum ar fi Global Open Data Index publicat de către Open Knowledge, Open Data Barometer al Web Foundation sau ePSI Platform Scoreboard, multe state europene se numără printre pionierii tendinței globale de a face informațiile deținute de către admistrația central sau locală disponibile publicului larg. Aceasta impulsionează eficiența în sectorul public, sporește transparența și impulsionează inovația la nivelul societății. Climatul economic actual obligă guvernele să asigure că banul public este cheltuit cât se poate de eficient. Workshopul timișorean va culege experiențele în prioritizarea deschiderii seturilor de date de către cei care le dețin, precum și cele legate de comunicarea cu utilizatorii acestora, indiferent dacă vorbim despre utilizatori privați, comerciali sau alte organizații publice.

Cordonatorul echipei Departamentului de Informatica - de la Facultatea de Matematică și Informatică a Universității de Vest din Timișoara - in cadrul proiectului SHARE - PSI si chair local al evenimentului, Prof. dr. Dana Petcu, subliniază faptul că "Timișoara este un centru al excelenței academice si al mediului de afaceri și acest eveniment va demonstra cum această experiză poate, și trebuie, să contribuie la dezvoltarea acestei piețe atrăgătoare și într-o susținută expansiune la nivel european.

La fel ca și în celelalte workshop-uri organizate în cadrul proiectului Share-PSI, scopul celui de la Timișora este de a identifica bune practici în acest domeniu, cu precădere legate de informațiile publice, dar acoperind și patrimonial cultural și datele științifice și comerciale. Aceste bune practici vor fi definite atât în cadrul proiectului, dar și de către organizații de standardizare precum World Wide Web Consortium (W3C), Open Geospatial Consortium (OGC) și Open Group.