

PSI Portals: Overview of Progress (Part 1)

Rob Davies, Co-ordinator European Public Sector Information (PSI) Platform

Abstract

This Topic Report describes current progress in the establishment of PSI 'Asset Registries' in the context of the provisions of the PSI Directive, taking into account recent developments in the area of open data catalogues which are relevant to PSI re-use. It goes on to discuss the need to satisfy demand for easier discovery of PSI across Europe for re-users wishing to create cross-border services and products, describing the possible advantages and disadvantages of different models for approaching this issue. It will be followed by a second Topic Report covering current best practice in the implementation of PSI portals in greater depth.

Keywords

Access registries, cross-border data discovery, linked data, open data, portals, PSI data catalogues.

About the Author

Rob Davies is the Co-ordinator of the ePSIplatform. He has worked in European information society activities since 1993 and previously led the ePSIplus Thematic Network which provided recommendations to the Commission on the review of the Directive in 2008, following a 2 year programme of analysis and awareness raising in every member state. Prior to that, he was a participant in the PSINet Preparatory Action (2001) which succeeded the 'PIRA Report' on Commercial Exploitation of PSI and the ePSINet Accompanying Measure. Rob is Chief Executive of MDR Partners.

© 2010 European PSI Platform - This document and all material therein has been compiled with great care; however, the author, editor and/or publisher and/or any party within the European PSI Platform or its predecessor projects the ePSIplus Network project or ePSINet consortium cannot be held liable in any way for the consequences of using the content of this document and/or any material referenced therein. This report has been published under the auspices of the European Public Sector Information Platform.

The report may be reproduced providing acknowledgement is made to the European Public Sector Information (PSI) Platform.

The [European Public Sector Information \(PSI\) Platform](#) is funded under the European Commission [eContentplus programme](#).

Introduction

The EU PSI Directive 2003/98/EC¹ came into force on the 31 January 2003. Article 9 (Practical Arrangements) states that: '*Member States shall ensure that practical arrangements are in place that facilitate the search for documents available for reuse, such as assets lists, accessible preferably online, of main documents, and portal sites that are linked to decentralized assets lists*'.

Following a public consultation, the European Commission on 8 May 2009 also published the *Communication on the Re-use of Public Sector Information – Review of Directive 2003/98/EC*². This document pointed to the creation of web portals on available PSI as a tool for finding, using and trading information in Slovenia and the United Kingdom as evidence of impact. However, progress in many other member states elsewhere has been limited. The communication provided further encouragement to public sector bodies to '*identify and make their information resources readily and promptly available in stable formats*' and emphasized that '*Information asset registries and national PSI portals are important tools in this respect*'.

Current developments

Since this Communication, a significant and potentially high impact development has taken place in the general context of the movement towards 'open data', which is gaining momentum in both e-government and civil society contexts. Specific linkages have been established between open data and the availability and re-use of PSI as a means of adding both economic and social value through opening up the data. At the ePSIplatform meeting on 'Realising the Value of Public Sector Information'³, held in Madrid on 8 June, 2010 the preference of re-users for access to 'raw data' created by the public sector was also stated repeatedly.

Open Data is a philosophy and practice requiring that certain data are freely available to everyone, without restrictions from copyright, patents or other mechanisms of control. It has a similar ethos to a number of other "Open" movements and communities such as open source and open access⁴. The Open Knowledge Foundation has also produced a wider definition of open knowledge covering data, content and services⁵

A number of countries, including several member states, have created or are now in the process of creating open data catalogues of PSI. This approach is also beginning to be reflected in innovative municipalities and regions. The ePSIplatform has created a page on its web portal⁶ which endeavours to track major developments of this kind. In doing so, an attempt has been made to categorise these developments, which are somewhat heterogeneous in their objectives and approach, as follows.

1

http://www.epsiplatform.eu/psi_library/reports/european_directive_on_the_re_use_of_public_sector_information_psi

2

http://www.epsiplatform.eu/psi_library/reports/european_commission_communication_on_re_use_of_psi_directive_may_2009

³ http://www.epsiplatform.eu/news/events/meeting_realising_the_value_of_public_sector_information

⁴ http://en.wikipedia.org/wiki/Open_Data

⁵ <http://www.opendefinition.org/>

⁶ http://www.epsiplatform.eu/psi_data_catalogues

Catalogues by Governments – data.gov style catalogues (with access to raw data)

Those which: are an undertaking by the public sector (government) and not a civil society initiative; provide access to a broad and reasonably substantial set of government data and are not limited to data from one or two sectors such as geographic data; make data accessible in electronic 'raw' data formats, such as XLS, XML, CSV or others (not just textual information), making it possible to work with or re-use the data.

Examples are identified in the UK at both central (data.gov.uk listing over 3000 data sets)⁷ and local government level (five examples) and in the Piemonte region of Italy and the Asturias and Basque regions of Spain. Developments outside Europe in countries such as Australia, Canada, New Zealand and the USA are, if anything, better established at present.

Catalogues by Governments – Information Portals (no or limited access to raw data)

These entries describe 'Information Portals' put online by governments. These do not provide direct access to raw government data in re-usable electronic formats. They provide a listing of what government data is available and links to the source where the data can be located. These information catalogues: provide a single access point to locate government data and make it easier to locate; provide descriptive information about the data and sources which is useful for the re-user to know what is available; and have the potential to evolve and provide more direct access to raw government data in re-usable electronic file formats.

Inclusion criteria are that: an 'Information Portals' is an undertaking by a government at one level or another; and that they provide a listing of reasonably substantial government data sources and are not only limited to data from one or two sectors such as geographic data. At the time of writing developments at central government level in two EU countries are listed: Digitalisér.dk (Denmark) and Proyecto Aporta (Spain).

Catalogues by Civil Society Initiatives

This listing includes European 'Civil Society' initiated portals, as well as those from non European countries at a central government level or by an international organization, that is: those catalogues put online not governments but rather by individuals, digital activists or non-governmental organisations.

The inclusion criterion is that the civil society initiated portal provides access to government data (from any level of government) or links to where government data can be accessed (in any format). Some of these portals also provide descriptive information about the data available. There is a relatively rich emergence of such initiatives in a number of countries such as Denmark, Finland, Netherlands, Norway, Spain, Sweden and the United Kingdom. Internationally, initiatives are cited in Canada, New Zealand, Russia, USA and by the World Bank.

Future developments

It is evident that such initiatives are 'catching on' and hold considerable promise for the re-use of data created in individual member states. Discussions and planning towards the creation of open data catalogues by government are also reported on the ePSIplatform site in countries such as France and Norway.

⁷ <http://data.gov.uk/>

Evident challenges across Europe include: public sector retrenchment in times of economic hardship and a probably greater than ever need to demonstrate value for public investments. It remains unclear to what extent civil society will be able to 'take up the slack' across Europe, perhaps especially in its weaker economies; the need to contain the resource costs of development by sharing best practice in areas such as governance, public-private cooperation and on technical implementation (e.g. tools for metadata management). To this end, this Topic Report (Part 1) will be enhanced through a second Topic Report (Part 2) covering current best practice in the implementation of PSI portals in greater depth.

Joining up PSI discovery

The PSI Directive refers at several points to the need to support cross-border exploitation of public sector information. Recital (5) states as follows:

'One of the principal aims of the establishment of an internal market is the creation of conditions conducive to the development of Community-wide services. Public sector information is an important primary material for digital content products and services..... Broad cross-border geographical coverage will also be essential in this context. Wider possibilities of re-using public sector information should inter alia allow European companies to exploit its potential and contribute to economic growth and job creation'.

The need for services with wide European coverage in a mobile and cohesive Europe appears self-evident. The Digital Agenda for Europe⁸, one of the EU's 7 flagship programmes points to the continuing problem of fragmented digital markets:

'Europe is still a patchwork of national online markets and Europeans are prevented by solvable problems from enjoying the benefits of a digital single market. Commercial and cultural content and services need to flow across borders'.

Evidence presented at the recent ePSIplatform meeting in Madrid, suggests that many would-be re-users find the task of securing access to data across European borders a formidable task which is deterring the creation of new products and services. The ePSIplatform, building on the work of its predecessor projects, ePSINet and ePSIplus has sought to keep track of the development of cross-border products and services through a Directory⁹. Evidence of the abundant development of 'web 2.0' style services based on open data may in future render this task difficult or redundant.¹⁰

Previous project-based endeavours to this end, such as PSI Navigator¹¹, funded under the eContent programme, have largely foundered on barriers such as unsustainable technology choices, cost, workload, governance and copyright issues.

The questions therefore remain: how might such a need be satisfied on a realistic and affordable basis? What more do we need to know? What action might next be taken?

Initial discussion in the ePSIplatform has drawn attention to the need to identify ways of enabling 'cross-border', even 'trans-European', products by easing discovery of PSI data for re-use. Put simply, how might we best begin in terms of technical approach to assist discovery from multiple European PSI sources?

⁸ http://ec.europa.eu/information_society/digital-agenda/documents/digital-agenda-communication-en.pdf

⁹ <http://www.epsiplatform.eu/examples/directory>

¹⁰ <http://www.telegraph.co.uk/technology/news/7044147/Data.gov.uk-Top-Ten-Apps-so-far.html>

¹¹ http://www.epsiplatform.eu/psi_library/reports/psi_navigator_crosses_the_borders

The cultural sector provides one possible model through the major Europeana initiative¹² which has gained a high level of political support from the Commission and at Member State level and achieved significant funding support. This model requires the development of a Europe-wide infrastructure of data aggregations from which metadata is 'harvested' into a central index upon which 'portal' services and APIs are developed. At the time of writing, metadata for approaching 10 million cultural objects has been harvested and normalized by this and similar methods and is now available for the application of semantic web applications such as Linked Data (see Topic Report 7)¹³. In this model, the data (or 'digital objects') themselves, stay on their original site.

This model involves considerable political will and substantial cost. Uncertainties remain as to whether sustainable business models for data aggregators can be established. Progress is being made in multilingual discovery technologies, but a complete solution remains to be achieved. Differences in nature between digital 'cultural objects' and the raw data preferred by many PSI re-users also render it somewhat uncertain whether such a model could be 'read across' to the broader government data sector, although the emergence of data.gov initiatives does promise to provide one element of a PSI aggregations infrastructure. The results of the INSPIRE initiative¹⁴ are also relevant, although more work is needed to coherently relate the goals of this to PSI re-use in general.

The Linked Data approach, proposed by Tim Berners-Lee and others¹⁵, may possibly provide a different, less 'centralised' approach to thinking about the issue of pan-European PSI discovery, drawing on a behaviorally-driven linking of all kinds of named data. An advantage pointed to by proponents of a Linked Data approach is that it does not require cross-border agreement on each schema, tag, term or attribute before progress can be made. Instead, national, regional, local, and domain-specific initiatives can gradually be 'stitched together' as and when required. Those with an interest in combining a specific set of resources from different European locations could - presuming that they have made the data openly available in the first place - map key attributes to one another at the point of need and enable their users to move reasonably painlessly from one resource to another.

This would not obviate the value of identifying some core terminologies and bringing pressure to bear upon those responsible for maintaining them in order to have them made freely available in a manner that complies with Berners-Lee's four principles. This would make it far easier for data providers to refer unambiguously to a set of shared concepts, without the need to duplicate prior effort or map redundantly from one set of terms to another. There would however be no need to wait for those core terminologies to become available before doing anything: a degree of duplication or mapping to get things moving may be preferable to waiting for a long time for the 'right' solution to come along.

It appears however inescapable that some form of central promotion of such an initiative and the provision of a structural framework within which activity can be generated and monitored would be needed. This appears to be the kind of work which the European Commission seeks to stimulate and co-ordinate through its programmes in other, related areas.

¹² http://ec.europa.eu/information_society/activities/digital_libraries/index_en.htm

¹³ http://www.epsiplatform.eu/topic_reports/topic_report_no_7_linked_data_and_government

¹⁴ <http://inspire.jrc.ec.europa.eu/>

¹⁵ http://en.wikipedia.org/wiki/Linked_Data

Conclusions

Current progress towards a network of open data initiatives at central and local level across Europe provides considerable optimism for a resolution of the difficulties (licensing, charging etc) which re-users have experienced in the past in gaining access to PSI, especially raw data. It is a development which should continue to be encouraged and supported right across Europe.

It is still often problematic and laborious for re-users to access data which is sourced in more than a single member state, even if that country has made it 'open'. The time is perhaps ripe to consider by what means best to address this question and who should take action in this regard, in this interests of expanding the market for cross-border, pan-European and international approaches.

Related Selection of European Public Sector Information Platform References

ePSIplatform Public Sector Information (PSI) Data Catalogues – Web Portal Page

http://www.epsiplatform.eu/psi_data_catalogues

Brunetta Promises an Italian Style data.gov within a Year

http://www.epsiplatform.eu/news/news/brunetta_promises_an_italian_style_data_gov_within_a_year

Norwegian Government is serious about a data.norge.no

http://www.epsiplatform.eu/news/news/norwegian_government_is_serious_about_a_data_norge_no

Spanish Government continues as European leader on PSI

http://www.epsiplatform.eu/news/news/spanish_government_continues_as_european_leader_on_psi

Spanish Regional Open Data Catalogues

http://www.epsiplatform.eu/news/news/spanish_regional_open_data_catalogues

Brest métropole Océane adopts open data motion on geographic data

http://www.epsiplus.net/news/news/brest_metropole_oceane_adopts_open_data_motion_on_geographic_data

Rennes Métropole: 1st French Community to strategically open up government transport data

http://www.epsiplus.net/news/news/rennes_metropole_1st_french_community_to_strategically_open_up_government_transport_data

Basque Government puts #OpenData and #oGov into Practice

http://www.epsiplatform.eu/news/news/basque_government_puts_opendata_and_ogov_into_practice

Asturias frees local PSI

http://www.epsiplatform.eu/news/news/asturias_frees_local_psi

French PSI Portal Coming Soon!

http://www.epsiplatform.eu/news/news/french_psi_portal_coming_soon

We need more national PSI project champions!

http://www.epsiplatform.eu/news/news/european_psi_portals

UK data.gov + 6 days

http://www.epsiplatform.eu/news/news/uk_data_gov_6_days

UK data.gov onestop-shop now open!

http://www.epsiplatform.eu/news/news/uk_data_gov_onestop_shop_now_open

European PSI Portals

http://www.epsiplatform.eu/news/news/european_psi_portals

Locating Danish PSI becomes easier!

http://www.epsiplatform.eu/news/news/locating_danish_psi_becomes_easier

Related ePSIplatform Topic Reports

Topic Report No. 7: Linked Data and Government

http://www.epsiplatform.eu/topic_reports/topic_report_no_7_linked_data_and_government

Topic Report No. 6: State of Play: PSI Re-use in Slovenia

http://www.epsiplatform.eu/topic_reports/topic_report_no_6_state_of_play_psi_re_use_in_slovenia

Topic Report No. 5: Momentum building for open government data in Norway

http://www.epsiplatform.eu/topic_reports/topic_report_no_5_momentum_building_for_open_government_data_in_norway