

If You Can't Find It, You Can't Re-use It

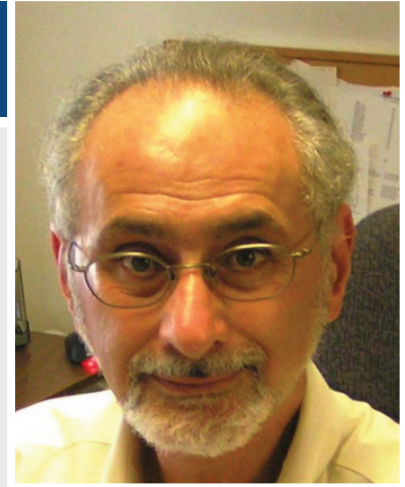
These days, thanks to Google and others, we take it for granted that we can find information at the click of a mouse button. The re-use of PSI is dependent on knowing what is available, where it can be found and on what terms it can be re-used and it would be great if one of the popular search engines could lead re-users to that information. There are two main problems.

The first is that, unfortunately, much PSI is unpublished and not publicly available on the web. The second, familiar to all users of search engines, is the imprecision of their results and the massive number of hits for each search (a Google search on "company information" returns 43.5 million hits). One way to overcome these barriers is for public sector information holders to provide PSI asset registers, including unpublished as well as published material, with sufficiently standard metadata to allow registers to allow federated search of these registers on a pan-European scale. In addition to facilitating PSI re-use, this would be a valuable asset for public sector bodies themselves, providing better knowledge of their own information assets and providing opportunities for increased

sharing and comparing data with their counterparts.

Participants at the recent [Standards Thematic Meeting](#) in Madrid devoted to the subject of PSI asset registers, ([reported in this issue](#)), concluded that three key requirements for such lists were the gathering of metadata requirements from potential PSI re-users, the creation of a minimum metadata set and the development of tools to help PSI holders provide that metadata.

[John Sheridan's article](#), in this issue, argues the need for Information Asset Registers but makes a case for use of semantic web technologies such as RDFa for enabling interoperability. The semantic web offers great opportunities in this area and ePSIplus is watching development with keen interest. However, not all European PSI holders have the necessary expertise to implement these technologies and there remains a pressing need to agree, in collaboration with potential re-users, a pan-European core set of mandatory metadata and to provide incentives, even compulsion, for PSI holders to provide it.



Brian Green, ePSIplus Analyst

The [INSPIRE Directive](#) provides an example of this that that could usefully be adopted in the area of PSI. INSPIRE offers rules for the metadata to be created, a degree of compulsion ("Member States shall ensure that metadata are created...") and is developing tools to facilitate the creation and gathering of such metadata.

Also in this issue, a report from the [CEN/ISSS eGov-Share Workshop](#) describes their exploration of some of the semantic and cultural issues involved in the federation of e-Government metadata. If we are to succeed in our aim to make PSI more available for re-use, we need to work hand in hand with e-Government, Freedom of Information and other initiatives such as INSPIRE, all of which are concerned with making European public sector information more discoverable, retrievable and usable.

Brian Green, ePSIplus Analyst

► Read more on the ePSIplus Forum about...

[Asset Lists](#)

[PSI Standards – Do they Impact on Re-use?](#)

Contact Us

Rob Davies

ePSIplus Network Co-ordinator
Rob.Davies@epsiplus.net

Mary Gianoli

ePSIplus Project Manager
Mary.Gianoli@epsiplus.net

Inside this issue

If You Can't Find It, You Can't Re-use It	1
'PSI Asset Registers and Metadata'- Information Management Standards Third Meeting	2
Rewiring Information Asset Registers	3
The Belgian Transposition of the PSI Directive and INSPIRE	4
CEN/ISSS eGov-Share: A Standardisation Initiative to Facilitate Discovery of and Access to eGovernment Resources	5
The Grace Period Ends on 31 December 2008	6
Spain Unravels the Value Chain of Legal PSI	7
Upcoming ePSIplus Meetings and published Country Reports	7

'PSI Asset Registers and Metadata' Information Management Standards 3rd Meeting

Brian Green, ePSIplus Analyst

The third thematic meeting on [Information Management Standards and Data Quality](#) was held in Madrid on 12 September 2008, hosted by the Facultad de Ciencias de la Documentación of the Universidad Complutense de Madrid.

The meeting, which was attended by 40 participants from 16 countries, focussed on the issue of PSI Asset Registers and Metadata. The meeting started with three keynote presentations. Don Lehane from Lendac Data systems in Ireland discussed the [problems facing PSI re-users](#) in attempting to find what data is available, from whom and under what terms and conditions, concluding that information asset registers were essential for this purpose. Jo Ellis of the UK Office of Public Sector Information (OPSI) argued that, rather than try to create a standard metadata schema for Information Asset Registers across Europe, and expecting Public Sector Information Holders to provide OAI-PMH compliant metadata, they should be encouraged to use semantic markup such as RDFa (Resource Description Framework-in-attributes) to facilitate access, use and re-use of data (see [John Sheridan's article on Rewiring Information Asset Registers in this newsletter](#)). Sheena Basset then outlined [lessons learned from the PSI Navigator Project](#), where Information Asset metadata had to be collected by project workers rather than by PSI Holders, arguing that more sticks and carrots were required to make PSI holders provide information about their data holdings.

In the next session, some important related projects were presented. Max Craglia and Rob Walker spoke about the INSPIRE initiative on geo-spatial data and its implications for PSI asset registers. INSPIRE actually has requirements for member countries

to provide metadata and lays down rules for the metadata elements to be collected and search criteria. Since there is a considerable overlap between PSI and geo-spatial data, INSPIRE should have a positive effect in kick-starting government organisations to produce asset registers and standard metadata. Karl Wessbrandt, Makx Dekkers and Marc Kunster described the [e-Gov Share project](#) which aims to make it easier for government departments and municipalities to share e-government resources, and the mechanisms and metadata required to support this. Rob Davies then talked about a number of [projects which are working together to build Europeana](#) (the European Digital Library) and the issues around metadata for cultural content.

The formal presentations ended with [three papers describing specific Information Asset Register initiatives](#) – from Ainars Brūvelis of Lursoft IT in Latvia, Andrea Muraca of CSI-Piemonte in Italy and Rodrigo Sánchez Jiménez of our hosts, UCM, in Spain. Rodrigo also provided a 'straw man' [metadata schema for PSI](#).

Lively discussion followed during which the participants agreed that a pan-European metadata standard for Information Asset Registers was required but expressed different views on the process for creating

such a standard. It was agreed that, in principle, the starting point should be an analysis of the requirements of potential re-users but, since this would be a hard and lengthy task, many felt that it would be best to put a very basic minimum metadata set in place as quickly as possible and be prepared to expand it if required.

Despite its limitations, Dublin Core had been used in most of the initiatives that had been described, so it was probably a good starting point. Formats such as Excel should be recommended as carriers.

It was also agreed that the problems were more cultural than technical and that many PSI holders were not prepared to put effort into providing metadata and Information Asset Registers as they did not view it as a priority. INSPIRE required them to do so for geo-spatial data and it was felt that a degree of compulsion may be necessary in the PSI legislation in order to persuade public sector organisations, especially in local government, to provide the data. It was also agreed that, since even Dublin Core was unknown to most PSI holders, let alone RDF and semantic web standards, some simple spreadsheet-based software tools were necessary to help them build asset registers. ePSIplus offered to set up an [open forum](#) on their website to facilitate further discussion.

Comprehensive meeting papers and presentations on the 3rd Standards Meeting are available on the ePSIplus Web Portal.

[Click Here to Access.](#)

- ▶ [Access the Report on the 2nd Information Management Standards and Data Quality Meeting](#)
- ▶ [Access the Report on the 1st Information Management Standards and Data Quality Meeting](#)

Rewiring Information Asset Registers

John Sheridan, OPSI

This article is based on a [discussion paper](#) tabled at the [ePSIplus 3rd Standards Meeting on IARs, 12 September 2008, in Madrid](#). John Sheridan is Head of e-Services at the [OPSI \(Office of Public Sector Information\)](#). He is also a co-chair of the [World Wide Web Consortium's e-Government Interest Group](#).

Government departments and agencies create information assets as part of their public task. These information assets, typically created for internal administrative purposes, have not been developed with onward re-use in mind. The government then needs to make decisions about what information to make available. To do this it says to re-users, "tell us what you want, and we'll tell you if we've got it", to which re-users reply, "tell us what you have, and we'll tell you what we want". It was to break this catch 22 that Information Asset Registers (IARs) were devised.

An IAR is a structured metadata catalogue of the key information assets that a public sector organisation holds. From a re-use perspective the primary purpose of an IAR is for resource discovery – identifying resources that may be available for re-use.

IARs have existed for a decade or more. Over that time the issue has been to justify ongoing the effort for their creation and maintenance. The case for IARs based narrowly on the re-use agenda can often lack traction with public sector information holders. One lesson is clear: for many officials the most compelling case that can be made for IARs is in terms of immediate benefit to public sector information holders themselves.

An IAR gives an organisation corporate knowledge of what information assets they hold. Importantly, the UK government's recent Data Handling Review published in June 2008 requires that departments identify their key information assets and information asset owners, as part of a strategic approach to managing information risk. As the issues surrounding data and information management have moved up the agenda in government, so has interest in IARs – not to facilitate re-use, but as a risk management tool.

IARs and metadata

Information needs to be described in order for it to be useful. That's the role of metadata. However, the nature and type of metadata is changing as the nature of digital content changes. In the world of linked data, content becomes more fragmented and more distributed, the distinctions between data and metadata become blurred.

Our view of metadata is increasingly far richer than looking narrowly at bibliographic style descriptions provided by Dublin Core. The descriptions of data will be different, depending on what the information is and how it can be used. Metadata can also enable us to present information in more re-usable forms. An important area is the use of semantic mark-up, embedding metadata at the fragment level, within web pages, to enrich semi-structured and unstructured data.

One of the advantages of an IAR is that it consists of standard metadata for each of the information assets it contains. Departments are encouraged to publish their IARs on the web. OPSI offers a web based search facility which allows users to search and access all central government department's individual IAR's, through a federated search facility. For such a federated search to work requires some level of standardisation and interoperability between IARs.

There are two approaches to enabling this type of data exchange - the development and adoption of agreed common standards, or, where standardisation is inappropriate, the use of other approaches to enable systems to exchange information. This is the realm of semantic interoperability. It is not always necessary to standardise – diversity can be right and good, especially when things are different by nature.

What level of standardisation for IARs?

It would be possible to devise a standard set of metadata elements for IARs across Europe. However, past experience teaches us that standardised metadata can lead some public sector organisations to add information simply because it is part of the standard not because they can usefully populate a given element. Such 'box ticking' is wholly counter-productive. Bad, incorrect or overly generalised metadata is worse than no metadata.

Rather than formally standardise metadata elements in IARs across Europe, there should be a common approach for surfacing IARs to the web, cognisant of how the web is changing. What is needed is an approach that the majority of public sector organisations can readily use.

Our focus should be on making IARs interoperable and re-usable in an entirely web centric way. It is not realistic to expect many public sector information holders to create OAI-PMH compliant metadata repositories. However most can surface their IAR to the web and it is here that a level of interoperability can be achieved. With new Semantic Web standards it is possible to markup textual information inside documents, in effect turning a set of traditional web pages, like an IAR, into an API. At OPSI we have been exploring the use of semantic markup inside XHTML documents, using a standard called RDFa, in order to facilitate access, use and re-use of data.

RDFa is a new W3C recommendation which is both pragmatic and fairly easy to implement. Combined with another approach, called GRDDL transformations, it's possible to serve RDFa enabled data in a number of re-usable data formats on the web.

continued...

...continued

We are looking carefully at RDFa and how it might enable re-use, particularly of structured and semi-structured data such as IARs. RDFa is ideal for IARs as existing metadata schemes such as Dublin Core can be easily and explicitly used within it (everyone uses Dublin Core for the "title" element for example). An XHTML template for an IAR record could be produced fairly easily, showing how an IAR record should be marked-up using RDFa.

There are a number of benefits to using RDFa for IARs:

- Existing web based IARs can be 'tweaked' to support RDFa, without creating separate web services so interoperability can be achieved at relatively little cost
- Existing ontology can be re-used, but public sector information holders are not entirely constrained. Additional elements can easily be added, by ontology created at European, national or local level.
- Lowly screen-scraping applications benefit from a well-specified and extensible method for embedding semantics, parsers can quickly extract the data whilst more advanced Semantic Web applications benefit from the full power of RDF
- This approach maintains a close link between semantic information and the provenance of that information

In this view of the world, individual public sector information holders would be able to use a combination of standard ontologies, and where needed either terms from a common ontology or their own individual ontology.

What is particularly appealing about RDFa from a government perspective is that existing website based IARs can be more or less 'tweaked' to serve RDF data, simply by adding attributes to the mark-up. In principle, it should be as easy to add appropriate attributes to surface the semantics of the data displayed in the browser, so that these web pages can be parsed and the data simply extracted.

The extent of the tweaking varies depending upon your starting point, but the typical webmaster in a department, agency or local authority knows how to implement a particular mark-up for a type of information. Here then is potentially an easy route for semantically enabling IARs and a design pattern for web based, interoperable IARs that others can adopt and use.

The Belgian Transposition of the PSI Directive and INSPIRE

Katleen Janssen, ICRI (Interdisciplinary Centre for Law and ICT) - K.U.Leuven –IBBT



The [Belgian transposition of the PSI directive](#) was only completed on 6 March 2008, when the Brussels Ordonnance on the re-use of PSI was adopted. It took seven laws and decrees to fully transpose the directive. Unfortunately, there was no co-ordination and little consultation between the federal and the regional governments during the transposition process, which has led to some differences between the different legal texts. The main differences exist between the federal Law of 7 March 2007 and the Flemish Decree of 27 April 2007. The other regional legislation closely follows the federal Law.

The basic principle remains the same: the public sector bodies can decide themselves whether or not they make their documents available for re-use. The Flemish government has decided for its own departments and agencies without legal personality that their documents can be re-used.

There are two main differences between the Flemish approach and the other texts. First, the Flemish decree is the only one

that does not explicitly require personal data to be anonymised before they can be re-used, based on the idea that the existing legislation on the processing of personal data remains fully applicable and is sufficient to deal with the issue of personal data. The federal law and the other decrees do require personal data to be made anonymous. This could actually jeopardize the re-use of e.g. the register of companies and undertakings, as this contains a lot of personal data. The second difference deals with charging. The Flemish decree follows the PSI directive, while the federal law and the other decrees resolutely draw the marginal cost card. They only charge the costs of reproduction and dissemination, except when extra treatments, such as anonymising, are necessary. Next to these two differences, there are some minor elements with regard to timing, licensing conditions, means of redress, etc.

The implementation of [INSPIRE](#) in Belgium also promises to be challenging, as this also will need legislation at the

federal and regional level.

The federal, Flemish and Walloon competent authorities agreed that the INSPIRE national contact point would come from the Flemish Region, the Walloon region would deliver the representative for the INSPIRE Committee and the federal level would be in charge of setting up a National INSPIRE Forum, where all issues regarding the implementation of INSPIRE could be discussed.

In addition, a temporary federal working group has been set up to draft the legal text for transposition, and to prepare the co-operation agreements between the governments. Until recently, however, it is not clear whether the INSPIRE forum is actually working or who the National Contact Point will be. In the meantime, the Flemish government is finishing its first version of a draft decree on the Flemish Spatial Data Infrastructure, which should be offered to the Flemish Parliament shortly.

CEN/ISSS eGov-Share

A Standardisation Initiative to Facilitate Discovery of and Access to eGovernment Resources

eGovernment programmes and projects across Europe build systems and applications to serve citizens, business, and administration agencies. One of the challenges across those programmes and projects is to unify information that is maintained in different places under different mandates. The intent is to offer that information in an integrated way to various audiences. Providing eGovernment services involves a substantial amount of co-operation and sharing.

The results of the [CEN/ISSS eGov-Share Workshop](#) will consist of specifications and guidelines to assist designers and developers of systems and services to consolidate multi-source information.

In addition to developing a proposal for an eGovernment ontology, a metadata set and a standard for the exchange and syndication of information, the Workshop pays attention to different terminologies that are being used across agencies and countries, and to cultural elements that play a role in the use of information in different cultural contexts. The workshop provides hands-on examples on how to realize and to use those specifications.

The Workshop acknowledges that approaches may differ in various systems; therefore, there are no one-size-fits-all solutions. As such, the work concentrates on enabling exchange of information, without attempting to standardize local, regional, and national solutions across Europe.

Furthermore, the focus is on a federation of resources, enabling access to the description of services and other resources that are managed in different locations. However, the decisions to exchange information, and what information to

share, are not within the scope of the project. In many cases these decisions are based on bilateral agreements, or governed by European Directives or Regulations, such as the PSI and INSPIRE Directives.

As a consensus platform, the Workshop solicits participation and contributions from the community of program managers, designers, and developers of eGovernment systems and services. As recent examples, eGov-Share and ePSIplus co-located their September meetings in Madrid, and the October meeting was held in connection with the ESIG2 summit in Rome, thus facilitating the exchange of ideas and synergies in common areas of interest. Widespread participation ensures that the specifications, guidelines and tools meet the needs of the community. As part of the Workshop approach, open and workshop meetings are being organized. The next draft report will be available from mid November for a two month public comment period.

The final meeting before publishing of the workshop results is scheduled for 3 February 2009 at the new CEN premises in Brussels. All ePSIplus newsletter readers are welcome to read and comment the report and to register and attend the next meeting.

Contact

Karl Wessbrandt,
Swedish Administrative Development Agency
Chairman, CEN/ISSS Workshop eGov-Share
Tel. +46 8 550 557 12
e-mail: karl.wessbrandt@verva.se

eGov-Share Secretariat: Tim Kniep, NEN
Tel. +31 15 2690 100
e-mail: tim.kniep@nen.nl

The Grace Period Ends on 31 December 2008

Chris Corbin, ePSIplus Analyst

Five years is a long time but this is the period of grace that was given to the European Union Member States to comply with Article 11 – *Prohibition of exclusive arrangements* of the Directive 2003/98/EC.

The European Commission has over recent years constantly reminded Member States and public bodies that they need to comply with the Directive with respect to Article 11 and that the European Commission will take action from January 2009 against Member States that have failed to comply.

Mr. Horst Forster the Director responsible for Digital Content and Cognitive Systems within the Information Society and Media Directorate within the European Commission on the 4th June 2007 at the International Symposium titled *Public Data on the Private Market – New regulations on the Re-Use of Public Sector Information* that was held in Potsdam said in the conclusions of his presentation:

“The PSI Directive sets minimum standards, building on two basic pillars of the internal market: transparency and non discrimination. It is already levelling the playing field, but its impact depends on the willingness of the public sector bodies concerned to critically look at established practices and change them where necessary. In this respect the end of the grace period for phasing out exclusive arrangements between public sector bodies and private players is an important test. The grace period ends on 31 December 2008.”

A year later on the 13th June 2008 at the **ePSIplus Conference** titled *PSI Re-use – Who Takes Action Next?* **Mr. Horst Forster** stated:

“One area that the Commission will particularly scrutinise in the coming months concerns exclusive arrangements, the general prohibition of which becomes fully applicable from the beginning of next year.”

So have Member States and public bodies taken note and phased out the existing exclusive arrangements? Five years is a long time so the reasons to justify the continued existence of exclusive arrangements under the exceptions permitted in Article 11 are likely to be challenged.

Have Member States considered the risk of the European Commission commencing infringement proceedings against them under Article 11 of the Directive from January 2009 onwards? That is has each Member State undertaken an assessment (risk analysis) as to first whether public bodies have declared the existence of exclusive arrangements and second taken the appropriate action to either end them or open them up to other potential re-users?

In the run up to the European Commission’s review of the Directive there have been loud and consistent calls for the Directive to be made stronger and clearer and for action to be taken by the European Commission. Some of these calls have originated from the public sector themselves! Article 11 of the Directive is one area where the Commission has constantly flagged that it will take action from January 2009 onwards.

The ePSIplus Thematic Network through its wide range of activities has found little evidence that Member States and public sector bodies have taken note. The evidence from the existing re-users also indicate that exclusive arrangements still exist in many Member States in a number of thematic sectors but in particular that of legal information.

It is clear that Member States are running out of time to act and ensure compliance after that it is over to the public sector information re-users and their associations to flag up to the European Commission where exclusive arrangements that still exist and that do not fall under the Article 11 exceptions.

For Member States and public bodies that have called for the Directive to be strengthened and firmer action to be taken – beware! Ensure that action has been taken the risk and financial penalties are likely to far exceed the value of any one exclusive arrangement that is in existence!

For re-users of public sector information that have called for the Directive to be strengthened and firmer action to be taken and find that they are limited by the existence of exclusive arrangements then as of 1 January 2009 write to the Commission and provide the evidence. This action is not time consuming, not expensive, is easy and simple. In conclusion everyone has been warned – it is time to take the appropriate action.

Read Related News and Events on the ePSIplus Web Portal

[Click on links below to read...](#)

- ▶ **PSI Remains Underexploited**
- ▶ **ePSIplus Conference: ‘PSI Re-use – Who Takes Action Next?’**
- ▶ **Opening Address to the ePSIplus Conference, 13 June 2008, “Re-use of Public Sector Information: Change, Growth and Transformation” by Horst Forster, Director, Content, European Commission’s Directorate General for Information Society and Media, European Commission**

Spain Unravels the Value Chain of Legal PSI

Marc de Vries, ePSIplus Analyst



Opening up the national legal content resources (legislation, court decisions), Spain serves as a perfect example of how clear policy decisions impact PSI re-use potential. This was one of the main conclusions of the forum discussions held in the context of the [Spanish national meeting "Public sector information re-use in Spain"](#), on 30 June 2008 in Barcelona, kindly hosted by the [ePSIplus national representative for Spain](#) Mr Agustí Cerrillo Martínez, Professor at the Universitat Oberta de Catalunya in further collaboration with Ministerio de Industria, Comercio y Turismo and el Institut Cartogràfic de Catalunya.

Bringing together the full value chain of legal PSI - policy makers, public sector producers, private sector re-users and private sector secondary re-users – there was a general feeling that Spain serves as a best practice in this field. The basis for this success is that fact that the Spanish government has taken its responsibility towards the re-users, taking clear and consistent political decisions over time, thus creating certainty as to what the public sector data holders will offer and what not - thus creating a clear distinction between the public task and the market activities - and what the conditions will put up. Consequently, re-users have been able to rely on this and allowing them to take thought through business (development) decisions.

- ▶ Read Related News on the ePSIplus Web Portal **Madrid conclusions**
- ▶ Read about the **Spanish National Meeting** Comprehensive meeting papers, reports and presentations are available on the ePSIplus Web Portal

However, it also became apparent that there are still some tensions between the various players in the value chain, in particular as to the pricing of PSI. Mr Lluís Faus - CEO of vLex, an innovative legal publisher - stressed the importance of constant monitoring and assessment of price setting by the government, as it will significantly impact the chances for SMEs to enter the market: "The Spanish government has taken the first steps, which has been of paramount importance; now it needs to guard that the potential can be exploited by the market".

Upcoming ePSIplus Meetings

Registration is Open for the Following ePSIplus Meetings














Click on the links below to view details

- ▶ **3rd ePSIplus Financial Impact Meeting**
20-21 November 2008
Athens, Greece
- ▶ **Denmark Hosting ePSIplus National Meeting**
27 November 2008
Copenhagen, Denmark
- ▶ **Sweden Hosting ePSIplus National Meeting**
11 December 2008
Stockholm, Sweden

Read ePSIplus Country Reports

Comprehensive reports (countries represented in the ePSIplus Network) based on national meeting discussions focusing on Public Sector Information (PSI) re-use themes (Law and Regulation; Public Sector Culture; Re-use Business; Financial Impact; Data Standards).

Click on a country below to access a report

 Belgium	 Iceland
 Cyprus	 Ireland
 Estonia	 Latvia
 Finland	 Lithuania
 France	 The Netherlands
 Germany	 United Kingdom
 Hungary	