**Submission for** http://share-psi.eu/cfp/

## Removing the roadblocks to a pan European market for Public Sector Information re-use

**by *Paola Di Maio***

> *According to a survey conducted by the European Commission in 2006 (MEPSIR study), the overall market size for PSI in the EU is estimated at 27 billion. Increase in the re-use of PSI generates new businesses and jobs and provides consumers with more choice and more value for money. (source)*

The many issues surrounding Public Sector Information Reuse can be portrayed primarily as business opportunities, but the impact of PSI policies extends beyond the economic and financial sphere, into the broad agenda of the future of civil society at large.

PSI is not necessarily a 'market' in an economic sense. The commercialization and privatization of public sector information, if not clearly paced and moderated by good old civic values, could even become a threat to the integrity and value of the future of PSI initiatives.

Economic and commercial opportunities offered by the exploitation of PSI are important and should not be missed out, but should be carefully weighted against non economic ones, for example the increased ability to make more informed decisions, more knowledgeable citizens, transparent justice, equal opportunities, to participate in the social and economic activities, increased visibility of decision making processes in the public administration, less manipulation of opinions, and so on.

This position paper identifies and analyses issues that can be considered 'roadblocks' to PSI reuse in the light of the considerations above, and proposes a systemic approach to detangle what is essentially a complex problem space. The paper is roughly organised as follows:

1) Accesss to PSI literature
2) Shared view, shared data schemas and a common language for PSI
3) Identifying dimensions and stakeholders
4) Identifying conflicts and tensions (have we been going around in circles?)
5) Metrics
6) Learning and applications
Conclusion: toward a systemic approach to PSI Reuse

---

**1) Access to PSI literature and historical evolution**

The opening up of PSI has been a long process that started decades ago . Yet younger generations stewarding current efforts may not necessarily be familiar with its historical evolution, and how it got this far in the first place. This is partly due to a fragmented and not always accessible PSI body of knowledge.(PSIBOK?), Much of the essential PSI bibliography is still accessible only to subscribers of selected journals. For example the seminal Janseen paper ' Towards a European. Framework for the Re-use of Public Sector. Information: a Long and. Winding Road' (Jansen Dumortier) published by Oxford University Press in 2003, which analysed some of the core issues identified almost a decade ago and which are still open issues, is only accessible to subscribers of the relevant journal..

That article, (uplodaded a fair use copy here) contains essential pointers to the background and history of PSI that the community of PSI stakeholders across Europe, and beyond, should at least know of. The 'Synergy Guidelines' published in 1989 does not seem to be known in the wider PSI community. A search for 'synergy guidelines 1989' on the EPSI Platform does not yield immediately meaningful and usable pointers to the document.

The lack of historical perspective on contemporary issues that pertain our Society is lamentable.

PSI discussions are addressing 'accessibility' in a rather vague notional form (some statistics are available, but little qualitative evaluation of the accessibility challenges), but to date no PSI initiative guarantees full access to PSI related body of knowledge to every citizen in the Union, who should or would like to make an appropriate contribution, in a language that they understand.

Without appropriate knowledge history and background of PSI, and its evolution, it is very difficult for contemporary stewards and citizens to make informed decisions.

EPSI and LAPSI projects more recently have attempted to make PSI bibliographical references available. Some examples can be found here and here . However these resources are not complete (some list references only back to 2006, no references to everything that was published before) , in some cases , bibliographical lists are still published as PDFs (not accessible via a web browser but only as downloadable document).

[ As a side note, but when I tried to download one of such pdfs the system message warned me:

This type of file can harm your computer. Are you sure you want to download PSIBibliography_V2_Section7_DocumentType.pdf? Not very encouraging. ]

By not making the history of PSI evolution available and accessible in accurate and easy to get form, contemporary citizens may be asked to to reinvent the wheel over and over. There is no clear user friendly directory of PSI knowledge and information

**Roadblocks under this perspective**: Lack of adequate PSI knowledge, limited BOK, limited accessibility to historical knowledge sources, poor knowledge organisation, permissions/digital rights/copyright and technology implementation (documents only available as PDFs? arghh)
lack of adequately knowledgeable personnel both in EU administration and in the public administration at large

**2) A shared view of PSI, shared data schemas and a common language**

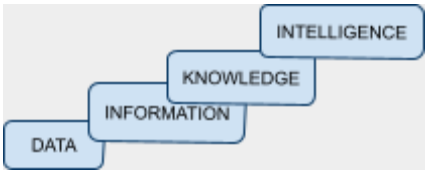
The choice of terminology and metaphors used in PSI discussions are bound to shape discourse. Yet some choices may seem either arbitrary, or the result of some bias. For example, just the wordings used in the title of the workshop: 'market' , where does it say that PSI is a market? Could it be a sector?

'Removing roadblocks' shows a bit of a brute force approach to problem solving. Formulating the statement differently, for example 'understanding and resolving conflicts of interests around PSI' may bring discussions to an entirely different level, and could lead to innovative ways of looking at the issues, no?

Where are the definitions of 'Public Sector Information'? What is the public sector? And what is 'reuse'? How do they impact people's lives? Are these terminology and concepts and their importance as initially defined in the directive still valide, and are they understood by citizens iin the EU? . Very few citizens outside specialised circles know how to answer these questions. (*try walking into any public authority building in your country and survey people at random. let us know the answers).

The language, metaphors and cognitive models chosen to represent and express what is PSI and its relevance to citizens may promote or hinder comprehension, and participation, in these important debates. The official language of the EU is English (I am personally grateful for that) but the average citizen in large portions of Europe is not conversant with the language at all. (*eurostats link? could not find relevant data.) Among citizens who may be interested to contribute to discussions, fewer still are entirely confident to argue eloquently in English.
Do we continue to translate all the PSI related documentation to all the EU language to engage citizens in critical decision making processes, or do we invest in more language learning opportunities to encourage participation ? Language and communication policies may pertain to another EU administrative jurisdiction, and contribute a long term strategy and vision and concern other policy areas, but they are strictly relevant here too.

Information is a very vague word, that cannot be easily defined. In research we rely on a distinction that helps us address the different levels of granularity. Below a diagrammatic rendering of DIKI (Data Information Knowledge Intelligence) (from ; Ontologies, Making the Business Case, Di Maio, 2007 Cutter Consortium)




Public sector '*information'* is in itself vaguely defined, and much of legal and policy instruments are currently only targeted at 'data' .The burden of this disambiguation currently falls on the administrative organs (the various FOI Officers and Commissions in the UK for example) who are trained to disambiguate bureaucratic and legal jargon to feasibly enforce the directive, but the lack of an appropriate language is a roadblock.

There is a lot of literature derived from years of EU publicly funded research in networks of excellence that defines 'knowledge' and 'knowledge schemas', but why is it not taken into account when designing PSI related instruments?

Better definitions and understanding of the various levels of granularity that contribute to define 'public sector information' can help us develop suitable wording

that can be adopted by technical and cognitive instruments for the coherent and fair handling of PSI as a whole.

How should PSI be conceptualized and defined, and what language should be adopted to address and resolve PSI issues?

Another reflective example comes from the word 'reuse' itself in the context of this discussion.

Reuse is not necessarily synonym with 'commercial exploitation', but without a clear disambiguation and a shared understanding of what reuse means, the whole discussion may inevitably be influenced by the arbitrary choices of interpretation of the language.

It is not my intention to advocate here a consensual PSI Ontology, but since some resources are already devoted to that effect (http://eprints.rclis.org/handle/10760/12819), it may become necessary to evaluate their effectiveness in the light of the challenges at hand.

Last but not least, there is no shared standard of what information/data/knowledge public authorities should store, nor what data schemas should be adopted to make it more reusable.The lack of harmonization across data schemas (an IT design and modelling task) is causing a lot of unnecessary extra costs to both the public and the public authorities to understand what data is held, and in what format. It is perceived that such lack of harmonisation and starndards to mandate proactive publication of information (what is held/where) is a deliberate barrier to reuse to increase red tape and create adminsitrative bottlenecks for access to public information, slow down processes and make it harder and more expensive for authorities to enforce.

**Roadblocks:** lack of understanding of the breadth of the problem space, terminologies and concepts, lack of disambiguation and shared understanding of core terms and issues, possible bias in choice of language, lack of shared data schemas of the public sector repositories, lack of proactive publishing of public sector information, deepening the digital divide between those who *know* and those who *know not*

**3) Dimensions and stakeholders**

To capture and address a complex scenario such as vast as PSI overall challenges, it can be useful to identify different 'dimensions' the problem spans across, and for each dimension identify the relevant group of stakeholders, to make sure these are adequately represented, included and consulted in decision making processes. To ensure a balanced and rational approach to problem scoping, as many dimensions as possible should be identified (via brainstorms) and as many stakeholders representatives as possible should be represented and included. For the purpose of this position paper, the following dimensions are identified purely as an example , each dimension can be associated with different communities of stakeholders and roles

| **DIMENSION** | **STAKEHOLDER** |
|---|---|
| Public good | every citizen |
| participatory governance | citizens that want to be involved in public administration |
| academia | researchers, students, academic, funding bodies |
| private sector, commercial exploitation, | entrepreneurs, traditional industries and industrial consortia |
| transcendental dimensions | philosophers, truth seekers, utopians |

**Roadblocks:** the lack of a clear and comprehensive multidimensional representation of the PSI problem space, and lack of a clearly identified stakeholder basis.

**4) Conflicts and tensions**

A plurality of views and intents is a beneficial and healthy attitude, there are some systemic dangers in promoting conflicting efforts
The most obvious example is the tension between **knowledge sharing culture** which appears to be advocated with great publicity and supported by public funding, and the **culture of competition** , instigated by stimulating competitive behaviour at the expenses of collaboration .
The first and most critical inhibitor of knowledge sharing for both individuals and organisations, is the desire to be perceived as as 'champion'. Knowledge which is not shared is the single most valuable assets that ensures economic and political supremacy of an individual or organisation over another.
As long as the EU and the National Funding bodies continue to encourage and reward competition at the expenses of cooperation, the efforts spent to promote knowledge sharing culture will be inhibited, and even nullified to some extent.
Besides, there are contractual agreements in place for example between research funding agencies and consortial partners in each country that restrict outright what knowledge /information can be publicly shared due to IP considerations. Some of these contractual agreements may be in direct contravention of PSI directive and related national legislation but there are no comprehensive monitoring processes in place for compliance, just spot checks.

The conflict between 'knowledge sharing' and 'lets be the best' whether in business or academia, is causing massive amounts of resources allocated to the respective efforts to zero each other out. One possible approach to resolve these conflicts is to attempt to establish some **cohesion** and **unity of purpose** among different stakeholders and some plan that can in turn aim to satisfy all of their requirements. Demonstrated knowledge sharing should be a factor in recognizing excellence, the more knowledge one shares, the more compliant the knowledge sharing is to good practices, the more it should contribute to 'excellence' There can be no excellence without compliance to good practices.

Economic benefits can spread evenly across stakeholders, and do not have to be achieved at the expense of non economic benefits, such as better public administration.
The public sector is financed by public money, raised mainly by taxes, therefore there should be no doubt that public sector main purpose should be to further the public good. What the PSI sector does not have, is a shared understanding of what the public good is. See roadblock 2.
**Roadblock:** Conflicting interests, tensions among stakeholders, lack of cohesion and unity of purpose among the stakeholder community, contractual agreements that restrict and limit knowledge sharing

**5) Metrics**

What are effective metrics of PSI reuse? Apart from monetary valuation of commercialisation, and the costs to governments, how is society monitoring the use and reuse of PSI?
**Roadblock:** lack of metrics for PSI reuse effectiveness (costs vs benefits, not just monetary).

**6) Learning and Applications**

Public data repositories are being designed and developed without the public as end user in mind, but are producing bunch of data that people dont know what to use and how to use it.Very few initiatives are taking place to inform and educate the public about the directive, its implications, and how to benefit from it.
**Roadblock:** publishing massive volumes of public sector data without appropriate application layer is no use, no appropriate application layer exists that permit its querying and manipulation for the purpose of acquiring new knowledge and support decision making, limited learning and traning opportunities exist for citizens

**Conclusion: Toward a systemic approach to PSI Reuse'**

A systemic approach is not generally an 'off the shelf' solution or silver bullet in any way. Each complex scenario requires an ad hoc plan, however some general principles apply, for example capturing the interactions: it is essential to try to capture not only each individual issue (say roadblocks identified along some dimensions as illustrated above) but also how these issues relate and impact each other. For example, a systemic approach would encourage viewing and exploring the technical and legal issues of PSI at their interaction point, rather than separately. A data format (say XML or RDF) needs to be attached to some policy and or organisational institutional process to be useful in practice, and the challenges to the practical usefulness of a data format need to be recorded and fed back into development with a feedback loop. A technology design (for example how dublin core or other data/knowledge schema for public data is defined, will impact the cost of implementation and monitoring of the relevant policies). It is essential that the interrelation of the various dimensions that pertain to PSI are discovered and addressed in the respective decisions making processes by all stakeholders.

A systemic approach would also cross reference the many challenges of PSI Reuse in the light of the principles presented by Commissioner Kroes earlier this year, reported below. For this analysis I remand to future occasions, due to the constrained scope and resources of this position paper.

- Transparency of rules affecting trade in ICT and ICT services
- Open networks for consumers to access and distribute information, applications and services of their choice
- Cross-border flows of information
- No requirement to use local infrastructure for ICT services
- Governments should allow full foreign participation in their ICT services sector, through establishment or other means
- Efficient and maximised use of radio spectrum
- Independence of regulatory authorities overseeing ICT services
- Simple authorisation of competitive telecommunications services
- ICT service suppliers must have the right to interconnect with other service providers for access to publicly available telecommunications networks and services. Public telecom services suppliers should be able to negotiate and obtain interconnection with major suppliers at cost-oriented, non-discriminatory and transparent rates.
- International cooperation with a view to increasing the level of digital literacy in third countries and reducing the 'digital divide'.

As more and more information and data become available, and better and more pervasive systems support their proliferation, individuals, organisations and institutions are becoming more involved, change rate faster, higher risks, higher opportunities. PSI, like other sectors, is increasingly complex and its dimensions entangled and interdependent, and requires a systemic approach to adequately ensure that technical and scientific progress serves the development of intellectual, social and human aspects of society, and not viceversa.

**Biographical Note:** Paola Di Maio is an international research scientist based in Edinburgh, the founder of ISTCS.org (Institute for Socio Technical Complex Systems Research), has a background in investigative reporting, and speaks fluently four European Languages. She studied and worked in most European countries, as well as Asia and Usa, she serves as expert and advisor and evaluator of technical and scientific committees. She has an active interest in participatory and sustainable initiatives.

References


http://www.slideshare.net/ccbuhr/open-data-eu-policies-and-activities

Bountouri, L., Papatheodorou, C., Soulikias, V., et al. Metadata interoperability in public sector information, 2009. In Journal of Information Science. SAGE. pp.204-231. (Published) [Journal Article (Print/Paginated)].

UK PUBLIC SECTOR INFORMATION AND RE-USE POLICY – A 2008 ANALYSIS Stephen Saxby
http://ses.library.usyd.edu.au/bitstream/2123/6562/1/PSI_vol1_chapter9.pdf

Di Maio, P 2007, From Data to Intelligence in:
Ontology: Making the Business Case | Cutter Consortium
1 Jul 2007 ... Ontology: Making the Business Case. by Paola Di Maio. Ontology has become an "in" thing for business intelligence professionals. ...
www.cutter.com/bia/fulltext/reports/2007/07/index.html

Borders in Cyberspace: Conflicting Public Sector Information Policies and their Economic Impacts
Yvette Pluijmers (PwC Consulting) Peter Weiss (United States National Weather Servic
http://www.spatial.maine.edu/~onsrud/Courses/SIE525/PluijmersWeiss.pdf


Public sector information
Renaissance, Reformation and the Age of Reason

1. Mike Clark1

**http://tinyurl.com/3n4nlnh**