Deliverable 7.2
Stable Version of the Share-PSI 2.0 Best Practices

Share-PSI 2.0

Standards for Open Data and Public Sector Information

Contractual Date of Delivery to the EC:	31 July 2016
Actual Date of Delivery to the EC:	29 July 2016
Managing Editor:	Daniel Pop, West University of Timişoara
Editors:	Fatemeh Ahmadi Zeleti, Insight Centre for Data Analytics
	Emma Beer, Open Knowledge
	Lorenzo Canova, Politecnico di Torino
	Makx Dekkers, AMI Consult
	Chris Harding, The Open Group
	Johann Höchtl, Danube University, Krems
	Valentina Janev, Institute Mihajlo Pupin
	Jan Kučera, University of Economics, Prague
	Nancy Routzouni, The Hellenic Ministry of Administrative Reform and E-governance
	Noël Van Herreweghe, AGIV
	Peter Winstanley, Scottish Government
Contributors:	Martin Alvarez-Espinar, CTIC
	Phil Archer, W3C/ERCIM
	Heather Broomfield, Difi
	Yannis Charalabidis, University of the Aegean
	Mª Dolores Hernández Maroto, Ministry of Finances and Public Administrations, Spain
	Peter Krantz
	Nicolas Loozen, PwC
	András Micsik, MTA SZTAKI
	José Luis Roda-García, University of La Laguna, Tenerife
	Ingo Simonis, OGC

Table of Contents

1 Introduction

The Share-PSI 2.0 Best Practices[1] are designed to act as a source of guidance for public organisations fulfilling their duties under the European Commission's Revised PSI Directive[2]. It extends and contextualises the W3C Data on the Web Best Practices, which provides general advice for sharing information online in such a way as to maximise its value and potential for reuse. For example, the W3C Best Practice "Provide Metadata[3]" can readily be seen as a *technique for opening data*' within the context of the Directive. Policy-related issues are out of scope for technical standards bodies like the World Wide Web Consortium (W3C), The Open Group (TOG) or the Open Geospatial Consortium (OGC), and so a number of additional Best practices are provided here to give a broader coverage of the steps that can be taken.

This document describes the process through which the Share-PSI 2.0 Thematic Network is developing the set of best practices that relate to policies around the implementation of the revised PSI Directive (Sections 2, 3 and 4), the stable versions of the Best Practices (Sections 5) and current status of implementation or adoption of the Best Practices by partner countries (Section 7). W3C Data on the Web Best Practices are summarized in Chapter 6, while Chapter 8 concludes on the work of the Network for past 2.5 years. The annex is devoted to those recommendations that did not pass the selection process set out within the project, as we consider them valuable for organizations responsible for Revised PSI Directive implementation process.

1.1 Addressing the objectives of the project

The aim of Share-PSI 2.0 Thematic Network is to bring together a very broad range of stakeholders in the reuse of public sector information and to help them go beyond the text of the revised PSI Directive. This section describes how the objectives of the Share-PSI 2.0 Thematic Network are covered by the best practices derived so far. Section 3 provides more details about the process through which best practices are delivered.

O1: To identify best practice and provide guidance on technical standards to European Member States implementing an open data policy.

An extensive amount of work submitted to Share-PSI 2.0 workshops is related to Open Data publication and reuse, addressing from both technical and policy-related viewpoints. The W3C Data on the Web Best Practices, which provide general advice for sharing information online in such a way as to maximise its value and potential for reuse, are included in chapter 6 "W3C Best Practices" of this deliverable and cover the technical best practices part. As described in D7.1, the scope of the two groups is different, nevertheless, the best practices identified by W3C can be applied directly to the technical aspects of the implementation of the revised PSI Directive. All W3C Best Practices are now stable, i.e. in their final form.

The policy-related best practices for Member States implementing an open data policy have been derived from the work of the Share-PSI 2.0 network. Examples of such practices are "Categorise Openness of Data", "Dataset Criteria", "Re(use) federated tools", or "Develop an Open Data Publication Plan". These best practices are detailed in chapter 5 "Published Share-PSI 2.0 Best Practices".

---

[1] https://www.w3.org/2013/share-psi/bp/
[2] http://ec.europa.eu/digital-agenda/en/legal-rules
[3] https://www.w3.org/TR/dwbp/#ProvideMetadata

O2: To identify best practice and offer guidance on technical standards particularly in the implementation of the revised PSI Directive, to improve harmonisation and interoperability.

As it was stated above, the technical best practices are developed in close collaboration with the W3C Data on the Web Best Practice (DWBP) Working Group. Information collected during the Share-PSI 2.0 workshops serves as one on of the inputs based on which W3C best practices are built. In order to make the link between the W3C best practices and the challenges related to the implementation of the PSI Directive, W3C best practices have been classified with relevant PSI elements.

However, improving harmonisation and interoperability might require more than the technical best practices. Share-PSI 2.0 best practices are also relevant, even though they address non-technical issues such as policy, legislation or organisation. On the Share-PSI 2.0 website for each of the PSI elements users can always find a complete list of relevant best practices covering both Share-PSI 2.0 and W3C best practices.

Here is the list of Share-PSI 2.0 published best practices dealing with harmonisation and interoperability in the implementation of PSI Directive: “Develop and Implement a Cross Agency Strategy”, “Enable Feedback Channels for Improving the Quality of Existing Government Data”, “Enable Quality Assessment of Open Data”, “Encourage Crowdsourcing around PSI”, “Establish an Open Data Ecosystem”, “Establish an Open Government Portal for Data Sharing”, “Identifying what you already publish”, “Open Up Public Transport Data”, “Publish Overview of Managed Data”, “Standards for Geospatial Data”, “Support Open Data Start Ups”. These best practices are detailed in chapter “Published Share-PSI 2.0 Best Practices”.

O3: To ensure that globally agreed guidance and best practice on technical standards can be implemented by Member States within their respective legal and cultural framework.

This objective is pursued by work in task 7.2 (Publicise best practices and encourage uptake across Europe and beyond) and in task 7.3 (Localisation of the best practices country by country). The coverage of the best practices with respect to localised guides is given in deliverable D7.3 “Localised implementations guides for the best practice” and also made available on Share-PSI Website at https://www.w3.org/2013/share-psi/lg/ .

O4: To ensure that the commercial interests around the technical provision and use of PSI are given due regard as well as those of the public sector, in particular offering guidance on the calculation and charging of marginal cost.

This objective is addressed in “Provide PSI at zero charge” Best Practice, as well as in a couple of recommendations included in Annex: “Cost-Benefit Analysis of the Value of Information” and “Select High Value Datasets for Publication”.

O5: To inform the development of relevant technical standards to best meet the needs of Member States, i.e. communicate demand for standards from the MSs to the relevant standards body, whether they are members of the network or not.

Representatives of standardisation bodies such as, W3C, The Open Group or OGC, participated on each of the five workshops. Following the 1st Share-PSI 2.0 workshop, in Samos, a use case was added to the relevant W3C document that added weight to several of the requirements already identified[4]. The one technical area highlighted in Samos that led

[4] More details are available here: https://www.w3.org/TR/dwbp-ucr/#UC-SharePSI

to a new requirement being added to the W3C work concerned location. Since that time, W3C and OGC have worked together to create an entire Working Group to develop best practices for spatial data on the Web[5]. It is noteworthy in this context that the Berlin workshop included location as a specific theme, with a track run by OGC and the EC Joint Research Centre (JRC).

[5] http://www.w3.org/2015/spatial/

2 Elements of the PSI Directive

The Directive on the reuse of public sector information (Directive 2003/98/EC, known as the 'PSI Directive') entered into force on 31 December 2003. It was revised by Directive 2013/37/EU[6] which entered into force on 17 July 2013.

It focuses on the economic aspects of reuse of information rather than on the access by citizens to information. It encourages the Member States to make as much information available for reuse as possible. It addresses material held by public sector bodies in the Member States, at national, regional and local levels, such as ministries, state agencies, municipalities, as well as organisations funded for the most part by or under the control of public authorities (e.g. meteorological institutes). Since 2013 content held by museums, libraries and archives falls within the scope of Directive as well. The Directive covers written texts, databases, audio files and film fragments; it does not apply to the educational, scientific, and broadcasting sectors.

The impetus for the Share PSI 2.0 Thematic Network is this revised European Directive on Public Sector Information. Share-PSI 2.0 is the European network for the exchange of experience and ideas around implementing these open data policies in the public sector. It brings together government departments, standards bodies, academic institutions, commercial organisations, trade associations and interest groups to identify what does and doesn't work, what is and isn't practical, what can and can't be expected of different stakeholders.

2.1 Derived elements

After analysis of the revised directive it was decided to focus on a particular set of sections from the directive. The decision to focus on these sections came from the following:

- In the preliminary discussions in preparation for the Share-PSI 2.0 workshops, it became obvious that most workshop participants could relate to, and saw certain sections in the revised directive as challenges which needed to be dealt with as priorities for the workshops;
- The identified Best Practices had to be technical and non-technical. Taken together, these two sets of Best Practices needed to underpin the sharing of PSI;
- As a result of the Share-PSI 2.0 thematic network's efforts, national, sectoral and community guidelines around Public Sector were going to be created, or updated. The information exchanged at those workshops needed to be consistent with the need for clarification on certain elements seen as challenges and needing more clarification;
- It was decided that the best practices developed by Share-PSI 2.0 needed to follow a similar structure to those developed by W3C provided that they were closely linked to the elements of the PSI Directive.

[6] http://ec.europa.eu/digital-agenda/en/legal-rules

2.2 References to specific sections of the Directive

This section presents the 13 derived elements, linking each to specific articles in the Directive and providing a short summary for each.

Policies and Legislation

Legal requirements, licenses etc. licensing of information, data and metadata

- DIRECTIVE 2003/98/EC: Chapter 1 - Art.1 :2, 3, 4, 5
- DIRECTIVE 2013/37/EU: Chapter 1 - Art.1 :2 (replaced), 3, 4, 5
- DIRECTIVE 2003/98/EC: Chapter 1 - Art.3: 1,2
- DIRECTIVE 2013/37/EU: Chapter 1 - Art.3: 1 (replaced), 2
- DIRECTIVE 2003/98/EC: Chapter 2 - Art.4: 1,2,3,4
- DIRECTIVE 2013/37/EU: Chapter 2 - Art.4: 1,2, 3 (replaced), 4 (‘replaced)
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.5: 1,2,3,
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.5: 1 (replaced),2 (replaced),3 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.6: 1,2, 3, 4
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.6: 1 (replaced), 2 (replaced), 3 (replaced), 4 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.7: 1,2, 3, 4
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.7: 1 (replaced), 2 (replaced), 3 (replaced), 4 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.8: 1,2
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.8: 1 (replaced), 2 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 4 - Art.10: 1,2
- DIRECTIVE 2013/37/EU: Chapter 4 - Art.10: 1,2
- DIRECTIVE 2003/98/EC: Chapter 4 - Art.11
- DIRECTIVE 2013/37/EU: Chapter 4 - Art.11 (amended)
- DIRECTIVE 2013/37/EU: Chapter 5 - Art.12
- DIRECTIVE 2013/37/EU: Chapter 5 - Art.12
- DIRECTIVE 2013/37/EU: Chapter 5 - Art.13: 1,2,3,
- DIRECTIVE 2013/37/EU: Chapter 5 - Art.13: 1 (replaced), 2 (replaced), 3 (replaced)

Governments have established various PSI reuse policies and legislation thus adopting various relevant approaches. However, all different approaches adopted by the European Union member states apply to the common rules set in the revised PSI Directive. Applying common rules eliminates the barriers to the cross-border offer of products and services, and enables comparable public data sets to be reused for Pan-European applications based on them.

Open data policies which encourage the wide availability and reuse of public sector information for private or commercial purposes, with minimal or no legal, technical or financial constraints, can play an important role in kick-starting the development of new services based on novel ways to combine and make use of such information, stimulating economic growth and promoting social engagement.

The aim of the Best Practices falling under this element is to present success stories, recommendations, achievements, effective processes and suggested actions illustrating how different Member States have formed and implemented governmental open data policies. Examples cover various aspects that support PSI policies (e.g. legislation, strategic actions, planning issues, organisational issues, policy positioning in public statements of intent).

These Best Practices would be of interest to policy makers and public sector representatives involved in forming or implementing the governmental open data policy, open data providers in Public Administrations, and also government data consumers.

Platforms

Open data platform(s), publication and deployment of information, data and metadata

- DIRECTIVE 2003/98/EC: Chapter 1 - Art.3: 2
- DIRECTIVE 2013/37/EU: Chapter 1 - Art.3: 2
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.5: 1
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.5: 1 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.9
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.9 (replaced)

As the length of the list of Best Practices suggests: there is more to sharing public sector information, making it available and encouraging its reuse than simply sticking a PDF on an organisation's Web site. Documents and data are the raw material for a community of providers and reusers, each with different needs. Those needs are most readily met through a dedicated platform that has a number of functions:

- a catalogue of documents and data, accessible by both humans and machines;
- a catalogue of uses of documents and data, again, accessible by humans and machines;
- a discussion forum;
- a dissemination channel;
- a feedback channel.

Platforms may offer further functionality such as data conversion and visualisation.

Dataset criteria

Dataset criteria, priorities, value and scope

- DIRECTIVE 2003/98/EC: Chapter 1 - Art.1: 2,4 (replaced)
- DIRECTIVE 2013/37/EU: Chapter 1 - Art.1: 2, 4
- DIRECTIVE 2003/98/EC: Chapter 1 - Art.5: 1
- DIRECTIVE 2013/37/EU: Chapter 1 - Art.5: 1 (replaced)

Aims, priorities and scope should be set for an open data initiative. Planning, executing and monitoring such an initiative requires datasets to be assessed according to various criteria that would allow balancing value of the provided datasets to users against the costs and risk

to the publishers. Best practices in this domain aim at providing the publishers with a set of criteria for assessment of datasets in various situations. The emphasis is on criteria enabling identification of high value datasets and on reflecting the views and needs of various stakeholders in the open data ecosystem. Best practices in this domain complement the best practices aimed at selection of datasets to be made available for reuse because they can be used to prioritise release of the selected datasets.

Charging

Charging issues and proposals

- DIRECTIVE 2003/98/EC: Chapter 3 - Art.6: 1,2, 3, 4
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.6: 1 (replaced), 2 (replaced), 3 (replaced), 4 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 4 - Art.10: 2
- DIRECTIVE 2013/37/EU: Chapter 4 - Art.10: 2

The subject of charging is a key element to the whole Open Government Data (OGD) and PSI strategy. In (Open data) theory public sector bodies release their data at no charge so providing scope for Transparency and the development of information products by existing and new companies, which in turn leads to new jobs and economic growth. The resulting taxation gains and social security contributions should outweigh the potential gains that would be made if the public sector bodies charged for access to the information.

The overall principle of pricing for public sector information touches on the following aspects:

- Studies such as that by Rufus Pollock[7] (PDF) prove high price elasticity of PSI demand. That means that if pricing changes, the result is disproportionate changes in PSI demand.
- Pricing influences the enforcement of other provisions. As soon as a public sector body releases data at no charge and under the terms of an open licence, there ceases to be a cost of license management and enforcement
- Pricing influences downstream market in the data value chain. When information is too costly, no aggregators and processors will use it.

These OGD and PSI basics have not been implemented in an adequate way everywhere. Up to now many public sector bodies, like trading funds, are reliant upon the income resulting from access fees to public sector information. Benefits from taxation gains and social security contributions accrue to the general budget and not directly to the relevant public sector bodies. The unwillingness of some central governments to compensate the public sector bodies for revenue reduction due to the release of public sector data means that they are reluctant to publish their data at an OGD basis.

Therefore, the issue of charging was intensely disputed in 2003 during the introduction of the first PSI directive. From the implementation of the 2003 directive up to the 2013 revision relevant public sector bodies could charge a maximum of *the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment (Article 6)* or the reuse of public sector data. Inspired by the open data developments in recent years, the 2011 proposal for an amendment of the PSI directive by the European Commission included the marginal costs principle. The total amount of any charges made should be limited to the marginal costs incurred for the reproduction and dissemination of the

[7] http://rufuspollock.org/papers/economics_of_psi.pdf

information asset. In the trialogue negotiations for the directive 2013/37/EC charges were a major topic again and the compromise reached led to a fragmentation of Article 6:

Article 6 now advocates the marginal cost principle, but in exceptional cases a public sector body may apply the provisions regarding charging as laid down in the 2003 PSI Directive. These exceptional cases are:

a. public sector bodies that are required to generate revenue to cover a substantial part of their costs relating to the performance of their public tasks;
b. by way of exception, documents for which the public sector body concerned is required to generate sufficient revenue to cover a substantial part of the costs relating to their collection, production, reproduction and dissemination.
c. libraries, including university libraries, museums and archives.

The European Commission's guidelines on recommended standard licences, datasets and charging for the reuse of documents[8] are a primary source for the definition of the marginal cost elements. According to these guidelines in an online environment, however, total charges could be limited to the costs relating directly to the maintenance and functioning of the infrastructure (electronic database), subject to what is necessary for reproducing the documents and providing them to one more reuser. Given that average database running costs are low and falling, the figure is likely to be close to zero.

Techniques

Techniques for opening data, technical requirements and tools

- DIRECTIVE 2003/98/EC: Chapter 1 - Art.3: 2
- DIRECTIVE 2013/37/EU: Chapter 1 - Art.3: 2
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.5: 1
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.5: 1 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.9
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.9 (replaced)

The directive mandates EU member states to open-up public sector information previously inaccessible to citizens. An increasing number of governmental, non-governmental or private organizations therefore publish sets of data in public spaces, covering domains such as healthcare, education, cultural heritage, financial, transportation etc.

Once on the Web, data become the bridge between data owners / publishers and the data consumers, which raises several technical challenges on both sides, such as: how to make the published data searchable, discoverable, machine-readable or accessible? What formats are appropriate for publishing data? How can data owners increase the interoperability of published data? What techniques / tools exists to cope with modifications of data already published?

To address these issues, we seek to provide guidance to stakeholders, through a set of good practices that apply to multiple technologies that aim to improve the consistency in the way data is published and consumed on the Web. Providing structured metadata, using well-established ontologies, implementing data versioning, providing bulk data downloads and so on are the types of recommendations covered by this section.

[8] http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_.2014.240.01.0001.01.ENG

Organisation

How to organise PSI sharing, necessary functions and communications

Applies to all content of both the Directives (communication, sharing, creating awareness etc.)

Implementation of the PSI directive is not a trivial task. Delivery of the full economic potential of public sector information might require legal, financial, structural, technical and cultural change within public sector bodies to efficiently and effectively deliver open, re-usable information assets. It also requires the development of the sector in the same way that public sector bodies support, nurture and encourage other industrial and economic sectors such as fisheries, engineering or tourism. The Organisational BPs focus on the areas of policy and governance, sectorial support, stakeholder involvement, and so on that underpin the development of a robust and sustainable ecosystem for the PSI industry.

Formats

Dataset structures, formats, APIs

- DIRECTIVE 2003/98/EC: Chapter 1 - Art.3: 2
- DIRECTIVE 2013/37/EU: Chapter 1 - Art.3: 2
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.5: 1
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.5: 1 (replaced)

Format is considered as those specifications and means for representing information. Within the PSI reuse is usually considered to depend on formats that are electronic and processable by computers. Format often corresponds with the type of the document used to present the information —e.g., XML, HTML. Structure and storage of the information depends on the format, but also on the mechanism or protocol of data delivery and access. For instance, web services and Application Programming Interfaces (APIs) may enable the information to be distributed in different final formats —SOAP, WFS, WMS, etc.

Publication of information in open formats — formats defined by clear specifications with free access for anyone — guarantees the universal accessibility and non-discrimination principle. Open formats often are those formats that can be accessed, and be processed through software tools that are themselves available free of charge and under an open license.

Selection of formats determines the options the reusers have in order to access and process the exposed information, as well as the tools they can use to do it. Formats could be a barrier to access the information (complexity of the distribution, or the cost of the tools to process it). Because of this, selection of the appropriate formats for the publication of specific datasets is a crucial element in fostering PSI reuse.

Reuse

Reuse challenges and encouraging (commercial) reuse

- DIRECTIVE 2003/98/EC: Chapter 1 - Art.3: 2
- DIRECTIVE 2013/37/EU: Chapter 1 - Art.3: 2
- DIRECTIVE 2003/98/EC: Chapter 2 - Art.4: 1,2, 3, 4
- DIRECTIVE 2013/37/EU: Chapter 2 - Art.4: 1,2, 3 (replaced), 4 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.8: 1
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.3: 1 (replaced)

- DIRECTIVE 2003/98/EC: Chapter 4 - Art.10: 1, 2
- DIRECTIVE 2013/37/EU: Chapter 4 - Art.10: 1, 2
- DIRECTIVE 2003/98/EC: Chapter 4 - Art.11
- DIRECTIVE 2013/37/EU: Chapter 4 - Art.11 (amended)

The benefit of data appears when it is used. Simply making data available to the public isn't enough. When government encourages data reuse actively, the likelihood of useful end user services increase and ultimately provides value for society as a whole. Commercial reuse depends on knowing of which data can be reused, providing low barriers to entry (e.g. clear terms of reuse) and there being a supportive legal framework. While many public sector bodies in Europe now make public data available, the development of Pan-European services and reuse by commercial vendors has probably not reached its full potential.

The revised PSI directive says "*Open data policies which encourage the wide availability and reuse of public sector information for private or commercial purposes, with minimal or no legal technical or financial constraints, and which promote the circulation of information not only for economic operators but also for the public, can play an important role in kick-starting the development of new services based on novel ways to combine and make use of such information, stimulate economic growth and promote social engagement.*"

Best practices in this area aim to lower the barriers for reuse.

Persistence

Persistence and maintenance of data and metadata

- DIRECTIVE 2003/98/EC: Chapter 3 - Art.5: 1, 2, 3
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.5: 1 (replaced), 2 (replaced), 3 (replaced)

When records of an official nature need to be used as part of process management protocols, a standard method of referring to a particular "snapshot" of the data elements grouped together and forming such records is required. There are three main needs which emerge from the ability to refer to and gain access to information in a consistent and predictable manner are:

- the latest (current) version of a record
- a specific version of the record in a specific time-frame
- a specific language manifestation of such records which is published as a "translation" of a previously assigned "original" record

Persistence of records should be based on the use of a Uniform Resource Identifier (URI) which is created by using a strategy which ensures that such identifiers can be created and continue to be used while keeping the same meaning over a long period of time which goes beyond the normal life-cycle of a record and well beyond the "disposal" or "archival" stages. The methods for the identifiers used for the persistence of records should be technology-neutral and allow both human and machine use of the information and its associated metadata.

Quality

Data quality issues and solutions, quality assurance, feedback channels and evaluation

- DIRECTIVE 2003/98/EC: Chapter 3 - Art.5: 1
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.5: 1 (replaced)

Quality doesn't feature as such in the directives. There were however many questions during Share-PSI 2.0 workshops regarding data quality. This issue was therefore put on the agenda because it impacts the efficient and effective use and reuse of our data and information.

Data quality has been one of the major concern of reusers. It is one of the most difficult tasks to achieve for public bodies that intend to publish existing datasets as fast as possible to ensure that they are shared with the community. However, the quality of data is critical to ensure reuse and support the whole ecosystem of service and apps creation. The revised PSI directive states that "*To facilitate reuse, public sector bodies should, where possible and appropriate, make documents available through open and machine-readable formats and together with their metadata, at the best level of precision and granularity, in a format that ensures interoperability*".

It therefore defines quality criteria for the data that is to be made available following the PSI directive implementation. Those criteria are related to formats, as well as metadata associated.

The "Guidelines on recommended standard licences, datasets and charging for the reuse of documents (2014/C 240/01)"[9] following the revised PSI directive state that "*In order to maximise the intended benefits of […] 'high-demand' datasets, particular attention should be paid to ensuring their availability, quality, usability and interoperability.*" Quality in this document is in particular related to update, granularity, persistence and referentiability, the presence of metadata, and data formats. However, it also highlights the importance of involving reusers in the maintenance of data quality over time. While more datasets have been published and reused, data quality has emerged as a major concern for both data publishers and data consumers. The W3C is working on a Data Quality vocabulary[10], while the European Commission has funded the Open Data Support project[11] to define quality of Open Data and associated metadata.

The characteristics of data that are recommended are illustrated in a set of best practices that take into consideration feedback of reusers and potential reusers to improve the data, the creation of metadata that support the discovery and reuse of data, the provision of versioning information, the unambiguous and persistent identification of datasets and finally the definition of quality criteria for the datasets.

Documentation

Documentation of information/data, creation of metadata

- DIRECTIVE 2003/98/EC: Chapter 3 - Art.9
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.9 (replaced)
- DIRECTIVE 2003/98/EC: Chapter 5 - Art.12
- DIRECTIVE 2013/37/EU: Chapter 5 - Art.12

Documentation is an essential aspect of information sharing. When any information is created, it is created for a specific purpose and usually within a specific context - to report on an activity, to monitor a change etc. When that information is shared, as envisaged by the revised PSI Directive, that context is likely to be lost. Think of it as a set of answers without the questions.

---

[9] https://ec.europa.eu/digital-agenda/en/news/commission-notice-guidelines-recommended-standard-licences-datasets-and-charging-reuse

[10] http://www.w3.org/TR/vocab-dqv/

[11] https://joinup.ec.europa.eu/sites/default/files/d2.1.2_training_module_2.2_open_data_quality_v1.00_en.pdf

To be useful outside the original context, potential users will want to know things like why was information gathered/created? How was it done? What assumptions were made? and so on.

Selection

Selection of information/data to be published according to various criteria

- DIRECTIVE 2003/98/EC: Chapter 1 - Art.2, 4
- DIRECTIVE 2013/37/EU: Chapter 1 - Art. 2 (replaced), 4
- DIRECTIVE 2003/98/EC: Chapter 2 - Art.1
- DIRECTIVE 2013/37/EU: Chapter 2 - Art. 12 (replaced)

Organisations often hold significant amounts of data, but they operate under various constraints that might make selection of datasets for publication challenging. Missing or incomplete inventories of managed datasets, lack of feedback, limited resources, technical constraints or constraints that result from the legislative framework (i.e. privacy protection) are examples of the common challenges in this domain. The aim of the best practices falling under the selection element is to help publishers to identify and select suitable datasets to be made available for reuse. These best practices should help data publishers to deal with questions such as where to start when selecting datasets for publication, how to make the community aware of existing datasets and thus enable the community to provide meaningful feedback about data that is in demand or what risks should be taken into account when selecting datasets that should be made available for reuse.

Resource and other constraints might prevent publishers from publishing all the selected datasets at once. In such cases, a release roadmap is usually developed where multiple iterations of the data release are planned over a specified period of time. When developing such a roadmap, best practices aimed at dataset criteria might be used to prioritise the dataset release.

Discoverability

- DIRECTIVE 2003/98/EC: Chapter 1 - Art.3, 2
- DIRECTIVE 2013/37/EU: Chapter 1 - Art. 3, 2
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.1
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.1
- DIRECTIVE 2003/98/EC: Chapter 3 - Art.9
- DIRECTIVE 2013/37/EU: Chapter 3 - Art.9 (replaced)

It is self-evident that for information to be reused, it must be discovered. However, the scale of the Web - the sheer volume of information published online every minute - means that publishers need to make an effort to help that discovery. Publishers of all kinds depend on search engines to catalogue and classify information, a task that can only be achieved by using extremely sophisticated computing, but search engines do not work by magic. It's important to help this process of automated classification if your information is to be discovered. Two key factors aid discovery:

- the information is well organised, self-contained and can be understood in isolation;
- the information is described using appropriate metadata.

3 The process of using the thematic network and the workshops to deliver knowledge

The main value of the Share-PSI 2.0 project is the knowledge gathered and exchanged between the thematic network members but also all the domain experts and entrepreneurs that have interacted and exchanged their expertise during the project workshops. Common challenges, problems that persist but also new opportunities in the domain of PSI have triggered a dynamic interaction which mainly focused on the implementation of the revised PSI directive.

The project also managed to spread this knowledge, mainly through the following means:

- **Knowledge exchange between project partners:**

During the workshops but also through bilateral cooperation between members of the thematic network and workshops participants, many experts in the domain have exchanged information, knowledge and expertise to face common problems and meet common challenges. Success stories, both in technical and policy-related aspects, have been exchanged and reused during this process. Such stories are gathered in the project wiki at https://www.w3.org/2013/share-psi/wiki/Stories

- **Knowledge transfer to workshops participants:**

An extensive amount of knowledge, addressing both technical and policy-related issues, is gathered and delivered to the participants of the project workshops through the presentation of the partners' case studies and the selected papers.

- **Knowledge delivered to all interested parties and the public through a list of Best Practices that derive from the gathered stories:**

Knowledge delivered to all interested parties and the public through the Share-PSI Best Practices[12] and W3C Best Practices[13] (practices developed by the W3C Data on the Web Best Practices Work Group) in the light of the project workshops and other discussions. These are extensively presented in chapters 5 "Published Share-PSI 2.0 Best Practices" and 6 "W3C Best Practices", respectively.

- **Knowledge delivered to the public through the localised implementation guides:**

Localised implementation guides listed on the project website at https://www.w3.org/2013/share-psi/lg/ provide advice that is consistent with the Best Practices developed by the Share-PSI 2.0 network. Therefore these guides represent another mean through which knowledge gathered by the Share-PSI 2.0 project could be delivered to the public.

In this section we describe the workshops organised by the Share-PSI 2.0 project and illustrate their impact. We also explain how knowledge was gathered during the workshops and we describe the process of delivering the Best Practices, i.e. how the gathered

[12] https://www.w3.org/2013/share-psi/bp/
[13] https://www.w3.org/TR/dwbp/

knowledge is translated into Best Practices, how we ensure that the Best Practices related to the project objectives and how the consensus on Best Practices is achieved.

In the context of the Share-PSI 2.0 project, a Best Practice is viewed as a commonly accepted method, technique or group of tasks that could be used, applied or followed to overcome some PSI or open data related challenge or that optimizes efficiency or effectiveness of publication or reuse of PSI or open data. Over time, best practice could evolve, be improved or be replaced with other best practice as more knowledge is gathered or new findings become evident.

The process of distilling best practices is based on the process that was outlined in the Description of Work and depicted in Figure 1 Share-PSI 2.0 Information Flow.

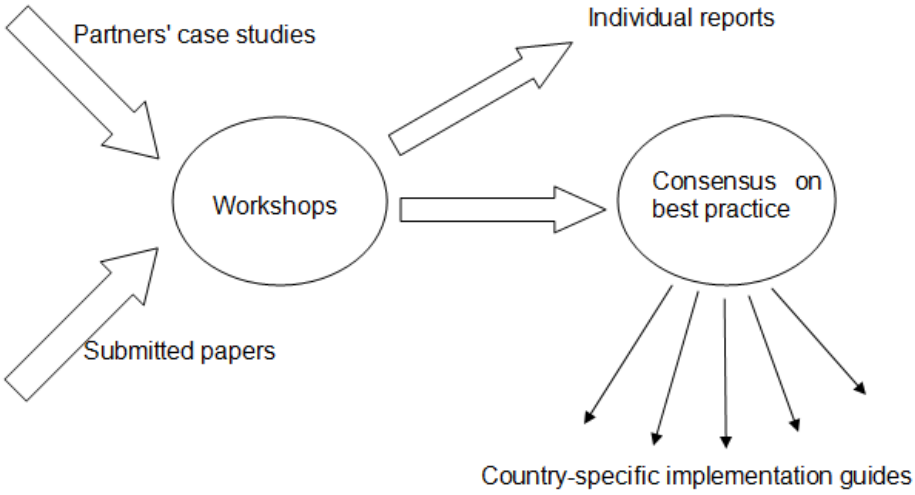

Figure 1 Information flow within Share-PSI 2.0

The starting point is the content of the workshops in the form of case studies contributed by the network partners and additional experience and suggestions contributed in submitted papers. This initial knowledge has been translated into best practices through the following process:

Step 1: Collect stories from the Share-PSI 2.0 workshops, keep notes and produce reports about important discussions.

Before each Share-PSI 2.0 workshop, the consortium assigned responsible persons for moderating sessions and taking notes of the discussion i.e. notes from the oral presentations, workshop sessions, bar camps etc. Starting from the second workshop, in order to better capture the experiences behind each presented use case, each session facilitator was asked to focus on three key questions:

1. **What X** is the thing that should be done to publish or reuse PSI?
2. **Why does X** facilitate the publication or reuse of PSI?
3. **How can one achieve X** and how can you measure or test it?

Step 2: Rewrite the stories to fit the best practices template.

Consortium members were responsible for analysing the stories and the notes kept from the previous workshops and for writing down the first draft of the best practices. In addition, they were asked to link them to the elements from the PSI Directive.

Step 3: Internal Peer Review of the first draft set of the best practices by an Editorial Board formed by a few members of the consortium.

The Editorial Board worked during Summer 2015 and after a thorough analysis of the raw material it created a collection of proposed Atomic Best Practices (deleting duplicates, interlinking the stories, enhancing the writing) to be taken into consideration in the subsequent consensus-building step. The term 'Atomic Best Practice' was used to indicate a difference from the stories, and to highlight the fact that each best practice should have a single focus and be capable of standing alone (although it may refer to others).

Step 4: Consensus building process on the elaborated Atomic Best Practices for D7.1.1 deliverable

The process of determining the best practices is based on consensus of the network partners as shown in Figure 1. The main criterion is the opinion of and the consensus among the network members. This consensus was reached using a voting process. The rationale for using the voting among the network members was that the network brings together a community of main experts from a majority of countries in Europe and their opinion should be leading in identifying the best practices as was the intention expressed in the Description of Work. Thus, in September 2015, the coordinator set up an online form that listed all the atomic best practices and for each one asked partners to indicate which of the following applied (they could check more than one box):

- I agree this is a good practice
- I agree this is good practice and we already offer advice consistent with it.
- I agree this is good practice and will cite it directly in our guide.
- I do not think this is good practice but am open to persuasion.
- I do not think this is a good practice
- I have some other comment

All atomic best practices that received a support higher than 80% have been categorised as published (agreed) best practices and they were included in D7.1.1, while those not matching the threshold were presented in the same deliverable as “Candidate Share-PSI 2.0 Best Practices” and are also available on the Share-PSI 2.0 Wiki[14].

Step 5: Consensus building for stable versions of Best Practices

Best practices derived from the Berlin workshop were developed and added to the collection of candidate Best Practices included in D7.1.1. In order to reach consensus on the final list of stable Best Practices, two additional activities were organized:

a. A final project meeting in Zagreb

b. A second round of voting

A final project meeting was held on 15 and 16 March Zagreb hosted by the University of Zagreb. The main subjects on the agenda were (a) discussion on the further development of the Best Practices and (b) coordination of the work on developing and referencing localised guidelines.

a. Zagreb meeting

The Zagreb meeting was organised in conjunction with W3C’s Data on the Web Best Practices (DWBP) Working Group who have produced a number of more technical Best Practices that are integrated with the Share-PSI 2.0 Best Practices in accordance with the

[14] https://www.w3.org/2013/share-psi/wiki/BP_2

Description of Work of the Share-PSI 2.0 network. The Best Practices from the DWBP WG are marked with the W3C logo at Share-PSI Best Practices Webpage[15].

While each of the two groups held their own meetings, a joint meeting was held on 15 March where cross-reading and cross-fertilisation between the two groups took place.

Various candidate BPs were extensively discussed in Zagreb, including Enable quality assessment of open data, Publish Statistical Data In Linked Data Format, Holistic Metrics and Open Data Business Models & Value Disciplines.

The discussions were detailed and the proposer of each Best Practice faced intense questioning, and as the result of the discussion and agreements several Best Practices were updated and improved.

It was also agreed that Open Up Research Data should be included with the caveat that scientific research data is not within the scope of the revised PSI Directive but that it is a closely related topic that raises many of the same issues.

A candidate Best Practice Use simple and distributed tools was not accepted but it was agreed that it should be integrated with (Re)use federated tools which was already in the accepted collection.

b. Second survey

The updates agreed in the meeting in Zagreb were finalised by 8 May 2016. Since that date, further editorial changes have been made to improve the text of several of them. All changes can be tracked back through the 'previous version' links on the Best Practices.

In order to verify the changes and new Best Practices that were added since the Berlin workshop, a second survey was conducted in May with the same set of six options as for the first survey. In this second survey each BP is given a weighted score. If the score was 25 or above the BP was accepted and included in the final list of 21 stable Share-PSI Best Practices presented in chapter 5 "Published Share-PSI Best Practices" of this deliverable and are publicly visible at https://www.w3.org/2013/share-psi/bp/.

From the analysis presented above, it is obvious that the published Share-PSI 2.0 best practices have collected evidence from more than one workshop. Table 2 Evidence for Best Practices: From Workshop to Recommendation extensively details what workshops contributed to which best practices.

A series of 5 workshops have been conducted, each one covering a theme of particular interest to the partners under the overall theme of PSI Directive implementation. Each partner of the network (bar one) presented at least one use case, most of partners contributing multiple use cases. There were also joint sessions organised by two partners, and the participation of external contributors to each workshop was also considerable.

[15] https://www.w3.org/2013/share-psi/bp/

Table 1 summarises the five workshops. Detailed reports per workshop are available under `https://www.w3.org/2013/share-psi/workshop/<location>/report`, where `<location>` should be replaced by `samos`, `lisbon`, `timisoara`, `krems` and `berlin` respectively.

Table 1 Share-PSI 2.0 Workshops

WS	Date	Location	Topic	Subtopics
WS1[16]	30 June - 1 July 2014 Registrations: 85, Papers/sessions: 25	Samos	Uses of Open Data Within Government for Innovation and Efficiency	The risks of open data; mobilising the public sector to publish its data; open data for policy modelling; collaboration between different communities; the open data feedback loop - communication between organisations that publish data and users of the data;
WS2[17]	3-4 December 2014 Registrations: 237, Papers/sessions: 24	Lisbon	Encouraging open data usage by commercial developers	Including youth and student entrepreneurship; collaboration between different communities
WS3[18]	16-17 March 2015 Registrations: 83, Papers/sessions: 13	Timisoara	Open data priorities and engagement — identifying data sets for publication	The risks of open data; mobilising the public sector to publish its data
WS4[19]	20-21 May 2015 Registrations: 87, Papers/sessions: 26	Krems	A Self Sustaining Business Model for Open Data	Mobilising the public sector to publish its data; moving from open data (datasets) to open services; collaboration between different communities; dedicated services for making open data truly usable; data catalogue interoperability and discoverability for better aggregation and federation; lessons learned by business representatives making use of PSI and open data; sharing of ideas for business models and obstacles

[16] Uses of open data within government for innovation and efficiency, Samos, 30 June to 1 July, 2014
[17] Encouraging commercial use of open data, Lisbon 3 to 4 December, 2014
[18] Open Data Priorities and Engagement, Timişoara:16 to 17 March, 2015
[19] A self-sustaining business model for open data, Krems: 20 to 21 May, 2015

WS	Date	Location	Topic	Subtopics
WS5[20]	25-26 November 2015 Registrations: 145, Papers/sessions: 34	Berlin	Maximising interoperability — core vocabularies, location-aware data and more	Consistency in quality assessment as well as structure and description; preparing data for publication including data cleaning, summarising, anonymising; tooling for efficient data publication; the use of INSPIRE; compliance of data and services

The table below highlights how the contributions (sessions) of each workshop underpins the Share-PSI 2.0 best practices. All the collected stories are also documented on the Share-PSI 2.0 Wiki[21].

Table 2 Evidence for Best Practices: From Workshop to Recommendation

Best practice	Contributing workshops (evidence from stories)
PUBLISHED BEST PRACTICES	
Categorise openness of data	WS1:https://www.w3.org/2013/share-psi/wiki/images/d/d2/NorwegianPublicSectorSharePSISamos.pdf
Dataset criteria	WS3:https://www.w3.org/2013/share-psi/wiki/images/3/31/Share-PSI_Submission_Paper-PwC_v0.03.pdf WS3:https://www.w3.org/2013/share-psi/wiki/images/3/3e/AMI_proposal_Share-PSI_Timisoara_How_good_is_good_enough.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/OpenDataFactors_CeDEM
Develop an Open Data publication plan	WS1:https://www.w3.org/2013/share-psi/wiki/Best_Practices/Open_Data_Publication_Plan WS3:https://www.w3.org/2013/share-psi/workshop/Timisoara/report (WS3 title “Open data priorities and engagement — identifying data sets for publication”)
Develop and implement a cross agency strategy	WS1:https://www.w3.org/2013/share-psi/wiki/images/a/a0/TheFlemishInnovationProjects.pdf
Enable feedback channels for improving the quality of existing government data	WS1:https://www.w3.org/2013/share-psi/wiki/images/d/d2/NorwegianPublicSectorSharePSISamos.pdf WS1:http://www.w3.org/2013/share-psi/workshop/samos/supervizor WS1:http://www.w3.org/2013/share-psi/workshop/samos/OpenCoesioneMonithon WS3:http://www.w3.org/2013/share-psi/workshop/Timisoara/kotmel

[20] Maximising interoperability — core vocabularies, location-aware data and more, Berlin: 25 to 26 November
[21] https://www.w3.org/2013/share-psi/wiki/Best_Practices

Best practice	Contributing workshops (evidence from stories)
	WS3:http://www.w3.org/2013/share-psi/workshop/Timisoara/Jacek WS4:https://www.w3.org/2013/share-psi/wiki/Krems/Scribe#Publishing_and_improving_the_quality_of_open_data_with_Open_Data_Certificates.3B_Amanda_Smith_.26_Sumika_Sakanishi.3B_ODI (ODI releases Open Data Certificates which also automatically assess minimum quality levels)
Enable quality assessment of open data	WS1:http://www.w3.org/2013/share-psi/workshop/samos/OpenCoesioneMonithon WS3:https://www.w3.org/2013/share-psi/wiki/images/3/3e/AMI_proposal_Share-PSI_Timisoara_How_good_is_good_enough.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/OpenDataCertificates
Encourage crowdsourcing around PSI	WS3:https://www.w3.org/2013/share-psi/wiki/Timisoara/Scribe#Crowd_sourcing_alternatives_to_government_data_.E2.80.93_how_should_governments_respond.3F WS4:https://www.w3.org/2013/share-psi/wiki/Krems/Scribe#Towards_A_Sustainable_Austrian_Data_Market.3B_Michela_Vignoli.3B_AIT (workshop on “Towards A Sustainable Austrian Data Market” mentioned opportunities and challenges imposed by Open Innovation)
Establish an Open Data ecosystem	WS1:https://www.w3.org/2013/share-psi/wiki/images/f/fc/Samos_SharePSI_Austria_UptakeandImpact_fin.pdf WS1:https://www.w3.org/2013/share-psi/wiki/images/a/a0/TheFlemishInnovationProjects.pdf WS1: https://www.w3.org/2013/share-psi/wiki/images/6/63/ODA-UtilizationCases-SharePSI-Workshop.pdf WS2:https://www.w3.org/2013/share-psi/wiki/images/1/11/StimulatingOpenDataReuse.pdf WS2:https://www.w3.org/2013/share-psi/wiki/images/d/df/ISMB.pdf WS2:https://www.w3.org/2013/share-psi/wiki/images/0/0f/Spanishinfomediary.pdf WS3:https://www.w3.org/2013/share-psi/workshop/Timisoara/Lewandowski WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/AnOngoingOpenDialog WS5:https://www.w3.org/2013/share-psi/workshop/berlin/agenda#al5
Establish an Open Government portal for data sharing	WS1:http://www.w3.org/2013/share-psi/wiki/images/4/40/WorkshopSamosJun2014-ULL-Tourism.pdf WS1:https://www.w3.org/2013/share-psi/wiki/images/8/89/Share-PSI_FederationTool_v01_en_paper.pdf WS1:https://www.w3.org/2013/share-psi/wiki/images/e/eb/Feroz.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/ItalianNationalGuidelines WS5:https://www.w3.org/2013/share-psi/workshop/berlin/EuropeanInteroperabilityTheISACoreVocabularies WS5:https://www.w3.org/2013/share-

Best practice	Contributing workshops (evidence from stories)
	psi/workshop/berlin/EuropeanDataPortalArchitecture
High level support	WS1: The Flemish Open Data Program
Holistic metrics	WS1:http://www.w3.org/2013/share-psi/workshop/samos/report#difi WS4:https://www.w3.org/2013/share-psi/wiki/Krems/Scribe#Current_achievements_and_suggested_actions_on_planning_and_implementing_a_Government_open_data_strategy.3B_Nancy_Routzouni_.26_Thodoris_Papadopoulos.3B_MAREG
Identifying what you already publish	WS3:https://www.w3.org/2013/share-psi/wiki/images/f/f0/SCOT_timisoara.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/imcs
Open Data business models & value disciplines	WS2:https://www.w3.org/2013/share-psi/wiki/images/d/dd/Ahmadi.pdf WS2:https://www.w3.org/2013/share-psi/wiki/File:GOV4ALL.pdf WS2:https://www.w3.org/2013/share-psi/wiki/images/d/dd/Ahmadi.pdf WS2:https://www.w3.org/2013/share-psi/wiki/File:GOV4ALL.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/LinkedOpenGovernmentDataBusinessModel WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/LinkedDataBusinessCube WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/LinkedOpenGovernmentDataBusinessModel WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/LinkedDataBusinessCube
Open up public transport data	WS1:https://www.w3.org/2013/share-psi/wiki/images/8/88/SharePSI-ODF-Samos0714-V1.pdf WS1:https://www.w3.org/2013/share-psi/wiki/images/f/f5/Transport_gijon_sharepsi.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/OpenMove
Open up research data	WS1: MTMT: The Hungarian Scientific Bibliography WS3: Role of Open Data in Research Institutions with International Significance (notes) WS3: Making research data repositories discoverable WS4: re3data.org - Making research data visible and discoverable WS4: Open Science & Technology
Provide PSI at zero cost	WS2:https://www.w3.org/2013/share-psi/wiki/images/0/0f/Spanishinfomediary.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/OpenMove
Publish overview of	WS1:https://www.w3.org/2013/share-

Best practice	Contributing workshops (evidence from stories)
managed data	psi/wiki/images/d/d2/NorwegianPublicSectorSharePSISamos.pdf WS3:https://www.w3.org/2013/share-psi/workshop/Timisoara/kotmel
Publishing statistical data in Linked Data format	WS1:https://www.w3.org/2013/share-psi/wiki/images/e/e2/LinkedStatistics_SharePSI2.0.pdf WS1:https://www.w3.org/2013/share-psi/wiki/images/6/65/Samos_Workshop_2014_-_IMP_submission.pdf
(Re)use federated tools	WS1:https://www.w3.org/2013/share-psi/wiki/images/8/89/Share-PSI_FederationTool_v01_en_paper.pdf
Standards for Geospatial Data	WS3:https://www.w3.org/2013/share-psi/wiki/images/b/b5/Abstract_free_our_maps.pdf WS5:https://www.w3.org/2013/share-psi/wiki/Berlin/Scribe#Location_Track
Support Open Data start ups	WS2:https://www.w3.org/2013/share-psi/wiki/images/c/ce/SharePSI-LisbonDecember-Startupsworkshop-1.pdf WS2:https://www.w3.org/2013/share-psi/wiki/images/d/d5/Alvarez.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/WS_Yannis_Academic_Business_Accelerators WS5:https://www.w3.org/2013/share-psi/wiki/Berlin/Scribe#An_Intelligent_Fire_Risk_Monitor_Based_On_Linked_Open_Data.2C_Nicky_van_Oorschot.2C_netage.nl
RECOMMENDATIONS	
Catalogues and indexes for reference	WS1:http://www.w3.org/2013/share-psi/wiki/images/4/40/WorkshopSamosJun2014-ULL-Tourism.pdf WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/DataBanks
Citizens participation to improve Open Data portal productivity and efficiency	WS1:https://www.w3.org/2013/share-psi/wiki/images/6/63/ODA-UtilizationCases-SharePSI-Workshop.pdf WS1:http://www.w3.org/2013/share-psi/wiki/images/4/4a/OpenCoesioneAndMonithon-Samos-Final.pdf WS3:https://www.w3.org/2013/share-psi/wiki/images/5/59/Concept_Note_PSI.pdf
Cost-benefit analysis of the value of information	WS1:https://www.w3.org/2013/share-psi/workshop/samos/ WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/RDB WS4:https://www.w3.org/2013/share-psi/wiki/Krems/Scribe#Towards_A_Sustainable_Austrian_Data_Market.3B_Michela_Vignoli.3B_AIT
Discover by location	WS2:https://www.w3.org/2013/share-psi/wiki/Lisbon/Scribe#Session:_The_Central_Role_of_Location WS5:https://www.w3.org/2013/share-

Best practice	Contributing workshops (evidence from stories)
	psi/wiki/Berlin/Scribe#Location_Track
Maintain records of stakeholders' rights and interests	WS1:https://www.w3.org/2013/share-psi/wiki/images/f/fc/Samos_SharePSI_Austria_UptakeandImpact_fin.pdf WS2:https://www.w3.org/2013/share-psi/workshop/lisbon/redress WS4:https://www.w3.org/2013/share-psi/workshop/krems/papers/mareg
Select high value datasets for publication	WS1:https://www.w3.org/2013/share-psi/wiki/Best_Practices/Opening_Up_Public_Transport_Information_to_Save_Costs WS1:https://www.w3.org/2013/share-psi/wiki/Best_Practices/Supervizor_-_An_Indispensable_Open_Government_Application_%28Transparency_Of_Public_Spending%29 WS3:https://www.w3.org/2013/share-psi/wiki/Best_Practices/Free_our_maps WS3:https://www.w3.org/2013/share-psi/workshop/Timisoara/Loozen WS3:https://docs.google.com/document/d/1Tq9VPq8SQY5y0QBq5ch3wxtaZUkU3dDpgLLOPgyrsvI/edit
Study the companies that build on PSI at national level	WS2:https://www.w3.org/2013/share-psi/wiki/images/0/0f/Spanishinfomediary.pdf
Understand demand for data	WS1:https://www.w3.org/2013/share-psi/wiki/images/6/63/ODA-UtilizationCases-SharePSI-Workshop.pdf WS2:https://www.w3.org/2013/share-psi/workshop/lisbon/OpenDataLifeCycleBarCampNotes WS3:https://www.w3.org/2013/share-psi/wiki/Timisoara/Scribe#Crowd_sourcing_alternatives_to_government_data_.E2.80.93_how_should_governments_respond.3F
Understand your internal needs and priorities	WS2:https://www.w3.org/2013/share-psi/wiki/Best_Practices/Open_Data_2.0_-_Changing_Perspectives

4 Overview of the Share-PSI 2.0 Best Practices

This section gives an overview of the Share-PSI 2.0 Best Practices. It shows how they correspond to the elements of the Revised PSI Directive, and describes the template used for their presentation. The best practices that are referenced have been finalized to their stable version in the last period of the project.

The **Published (Agreed)** Share-PSI 2.0 Best Practices are practices agreed by the project partners to be Best Practices. The selection process is detailed in Chapter 3, and they are all included in the next chapter.

The Best Practices developed by the **W3C Data on the Web** Work Group are summarized in Chapter 6. Share-PSI 2.0 Network contributed stories, use cases and evidences supporting them.

A subset of the best practices developed by the Network have not been formally agreed to be Best Practices by the majority of partners (according to the selection process is described in Chapter 3). They are included as **Additional Recommendations** in this deliverable's Annex.

4.1 Correspondence to the Directive

To capture the main aspects of the best practices, a list of elements of the PSI Directive was established (see Chapter 2 "Elements of the PSI Directive") that was used to extract aspects from the case studies that were relevant for the PSI Directive elements. Table 3 shows, for each of the elements of the PSI Directive described in Section 2:

- Which best practices have this element as their main focus (the *Primary Best Practices*) and
- Which best practices contribute to that element, although their main focus is on another element (the *Other Relevant Best Practices*).

The status of each best practice is indicated as PUB (PUBLISHED), W3C or REC (RECOMMENDATION), i.e. those from the Annex.

Table 3: Best Practice Correspondence to PSI Directive Elements

Element	Primary Best Practices	Other Relevant Best Practices
Policies and Legislation	• Develop and implement a cross-agency strategy (PUB) • Encourage crowd-sourcing around PSI (PUB) • Enable feedback channels for improving the quality of existing government data (PUB) • High level support (PUB) • Holistic metrics (PUB) • Develop an Open Data publication plan (PUB)	

Element	Primary Best Practices	Other Relevant Best Practices
	• Open up public transport data (PUB) • Support Open Data start-ups (PUB) • Maintain records of stakeholders' rights and interests (REC) • Respect legislation and stakeholders' rights (REC)	
Platforms	• Establish Open Government Portal for data sharing (PUB) • Provide complementary presentations (W3C) • Provide feedback to the original publisher (W3C)	• Encourage crowd-sourcing around PSI (PUB) • Enable feedback channels for improving the quality of existing government data (PUB) • Establish an Open Data ecosystem (PUB) • Standards for Geospatial data (PUB) • (Re)use federated tools (PUB) • Citizens participation to improve Open Data portal productivity and efficiency (REC)
Dataset criteria	• Dataset Criteria (PUB) • Cost-benefit analysis of the value of information (REC)	• Establish an Open Data Ecosystem (PUB)
Charging	• Provide PSI at zero charge (PUB)	• Establish an Open Data Ecosystem (PUB) • Holistic Metrics (PUB) • Cost-benefit analysis of the value of information (REC) • Open Data Business Models & Value Disciplines (PUB)
Techniques	• Publish statistical data in Linked data format (PUB) • Provide bulk download (W3C) • Use content negotiation for serving data available in multiple formats (W3C) • Make data available through an API (W3C) • Avoid Breaking Changes to Your API (W3C) • Enrich data by generating new	• (Re)use federated tools (PUB) • Standards for Geospatial data (PUB) • Establish Open Government Portal for data sharing (PUB) • Preserve identifiers (W3C) • Use persistent URIs as identifiers of datasets (W3C)

Element	Primary Best Practices	Other Relevant Best Practices
	metadata (W3C) • Make feedback available (W3C) • Choose the right formalization level (W3C) • Use Web Standards as the foundation of APIs (W3C) • Provide Subsets for Large Datasets (W3C) • Provide data up to date (W3C) • Provide version history (W3C) • Provide a version indicator (W3C) • Reuse vocabularies, preferably standardized ones (W3C)	• Use persistent URIs as identifiers within datasets (W3C) • Provide complete documentation for your API (W3C) • Provide Complementary Presentations (W3C) • Assign URIs to dataset versions and series (W3C) • Provide metadata (W3C) • Provide descriptive metadata (W3C) • Provide locale parameters metadata (W3C) • Provide structural metadata (W3C)
Organization	• Open Data Business Models & Value Disciplines (PUB) • Establish an Open Data Ecosystem (PUB)	• (Re)use federated tools (PUB) • High level support (PUB)
Formats	• Standards for Geospatial data (PUB) • Provide bulk download (W3C) • Provide data in multiple formats (W3C) • Use machine-readable standardized data formats (W3C) • Reuse vocabularies, preferably standardized ones (W3C)	• (Re)use federated tools (PUB) • Assign URIs to dataset versions and series (W3C) • Use persistent URIs as identifiers of datasets (W3C) • Use persistent URIs as identifiers within datasets (W3C)
Reuse	• (Re)use federated tools (PUB) • Study the Companies that Build on PSI at National Level (REC) • Cite the Original Publication (W3C) • Make data available through an API (W3C) • Provide data license information (W3C) • Follow Licensing Terms (W3C)	• Establish an Open Data Ecosystem (PUB) • Standards for Geospatial data (PUB) • Open Data Business Models & Value Disciplines (PUB) • Use persistent URIs as identifiers of datasets (W3C)

Element	Primary Best Practices	Other Relevant Best Practices
	• Provide locale parameters metadata (W3C) • Provide structural metadata (W3C) • Provide real-time access (W3C)	• Use persistent URIs as identifiers within datasets (W3C) • Provide Complementary Presentations (W3C) • Provide data provenance information (W3C) • Provide Subsets for Large Datasets (W3C) • Provide version history (W3C) • Provide a version indicator (W3C)
Persistence	• Assess dataset coverage (W3C) • Preserve identifiers (W3C) • Assign URIs to dataset versions and series (W3C) • Use persistent URIs as identifiers of datasets (W3C) • Use persistent URIs as identifiers within datasets (W3C) • Provide version history (W3C) • Provide a version indicator (W3C)	• (Re)use federated tools (PUB)
Quality	• Enable quality assessment of open data (PUB) • Provide data quality information (W3C) • Gather feedback from data consumers (W3C) • Make feedback available (W3C) • Enrich data by generating new data (W3C) • Provide feedback to the original publisher (W3C)	• Provide data provenance information (W3C) • Citizens participation to improve Open Data portal productivity and efficiency (REC)
Documentation	• Provide complete documentation for your API (W3C) • Provide Complementary Presentations (W3C)	• (Re)use federated tools (PUB) • Enrich data by generating new metadata (W3C) • Provide data license information (W3C) • Provide version history (W3C) • Assess dataset coverage (W3C) • Provide metadata (W3C) • Provide descriptive metadata (W3C)

Element	Primary Best Practices	Other Relevant Best Practices
		• Provide locale parameters metadata (W3C) • Provide structural metadata (W3C) • Provide data provenance information (W3C) • Provide data quality information (W3C) • Catalogues and Indexes for Reference (REC)
Selection	• Publish overview of managed data (PUB) • Categorise openness of data (PUB) • Identify what you already publish (PUB) • Select high value datasets for publication (REC) • Understand your internal needs and priorities (REC) • Understand demand for data (REC)	• Establish an Open Data Ecosystem (PUB) • Gather feedback from data consumers (W3C) • Provide feedback to the original publisher (W3C) • Provide real-time access (W3C)
Discoverability	• Provide metadata (W3C) • Provide descriptive metadata (W3C) • Provide locale parameters metadata (W3C) • Provide structural metadata (W3C) • Provide an explanation for data that is not available (W3C) • Catalogues and indexes for reference (REC) • Discover by location (REC) • Citizens participation to improve Open Data portal productivity and efficiency (REC)	• Establish an Open Data Ecosystem (PUB) • Standards for Geospatial Data (PUB) • Establish Open Government Portal for data sharing (PUB) • Assess dataset coverage (W3C) • Use persistent URIs as identifiers of datasets (W3C) • Use persistent URIs as (W3C) identifiers within datasets • Assign URIs to dataset versions and series (W3C)

Notes:

1. Where a technique best practice is primarily useful for one of the other elements, it has been categorised as a primary best practice for that element, rather than as a technique.
2. All the best practices related to documentation also aid discoverability, but this is not shown explicitly.

4.2 Share-PSI 2.0 Best Practices Template

In order to ensure consistency of the Share-PSI 2.0 best practices and also to ensure that contents of the best practices are easy to understand and use, a common template for the Share-PSI 2.0 best practices was developed. The structure for the best practices was defined in consultation with the network partners, based on comments from the EC reviewers in the first and second project reviews. Individual sections of this template are described in Table 4. Note that the W3C best practices follow a different template.

Table 4 Share-PSI 2.0 Best Practice Template

Template section	Description
Title	Title of the best practice, preferably in the imperative form, e.g. "*Develop and Implement a Cross Agency Strategy*".
Outline	Summary of the best practice, most importantly what the challenge is and how it can be solved.
Links to the Revised PSI Directive	Classification of the best practice with one or more PSI elements. This classification directly links every Share-PSI 2.0 best practice with one or more problem domains or issues related to the implementation of the PSI Directive. See section 2 for the PSI elements description.
Challenge	Description of the challenge that the best practice tries to address. In the context of the Share-PSI 2.0 best practices a challenge is usually a problem or barrier faced by some subject involved in PSI production, publication or reuse.
Solution	Description of the solution to the identified challenge. It should clearly state what should be done in order to overcome the problem or barrier.
Why is this a Best Practice?	Rationale for this being a best practice. This section summarizes the arguments and explains the expected impact of implementing the best practice.
How do I implement this Best Practice?	This section provides guidelines and recommendations for implementing the best practice. The aim of this section is to make the best practice actionable, i.e. it should provide substantial and sufficient information to allow the readers to use the best practice for local implementation purposes. However please note that the best practice needs to be applicable in different contexts of the member states. Therefore, the guidelines and recommendations should be generally applicable. Country specific guidelines should be provided in the localised guides - see the "*Where has this best practice been implemented?*" section of the best practice.
Where has this best practice been implemented?	This section provides a list of countries, their implementations of this best practice and the national contact points.

References	References to the relevant sources. Original workshop stories, papers or session notes should be referenced in order to link the best practice with evidence collected during the Share-PSI 2.0 workshops. Additional references such as books, papers, studies etc. could be provided as well.
Localised Guidance	A list of guides that offer advice consistent with the BP and/or that cite it directly. This allows a reader to follow up on the BP with further guidance in their own context and, often, in their own language.
Contact Info	Main contact point for the best practice. Name and email of the person should be provided together with an affiliation and country where appropriate.
Related Best Practices	Links to other best practices relevant to the topic discussed by the best practice.

4.3 Collecting Feedback

Implementation and use of the best practices might show that some of them need to be updated in one way or the other, even after project's completion. Therefore, collecting lessons learned and feedback is essential for ensuring that best practices address needs and challenges faced by PSI and open data practitioners and users, and that the guidelines provided by the best practices will always be up to date.

In order to achieve this objective, each BP is linked to an open issue on GitHub where any comments can be made in a forum familiar to many people. A copy of each of the BPs is also hosted within the Share-PSI GitHub repository[22] so that edits and updates can be made. It is then a relatively simple process to copy such changes in to the stable, non-editable environment of w3.org where the authoritative snapshots are published.

[22] https://github.com/w3c/share-psi/issues/

5 Published Share-PSI 2.0 Best Practices

5.1 Categorise openness of data

Outline

Establishing a simple system to categorise the openness of data makes it easier for public sector organisations to determine with whom data can be shared.

Links to the Revised PSI Directive

Selection

Challenge

How to ensure that data is shared with the authorised users even if some restrictions apply?

Public sector organisations often only consider

- what can be open for all and don't think about datasets which can be shared or
- get concerned about the difficult datasets which cannot be opened, rather than identifying the low hanging fruit which can be made easily available.

However, it may also be useful to publish data that can be shared under certain restrictions.

Solution

Descriptions on both open data and data with legal restrictions can be published with an indication of who it can be shared with. DCAT-APs "access right" property can be used for this purpose, combined with a colour code system for end users:

- :public (green) for data that can be made available with no restrictions;
- :restricted (yellow) for data that is not open for all and has some restrictions, and data which an organisation may be unsure of the status of;
- :non public (red) for data that is sensitive and can only be made available under strict conditions.

A colour code is not the only way to denote the category of data openness. Other schemes with a similar meaning could be used as well.

Why is this a Best Practice?

This best practice allows a twin-track approach: 'publish early even if imperfect' and a commitment to a 'high quality core.' This twin-track policy will maximise the benefit within practical constraints. It will reduce the excuses for poor or slow delivery; it says 'get it all out and then improve.'

The categorisation is simple. It is not designed to facilitate a deep and exhaustive mapping of data in an organisation, but rather it allows for an initial overview and categorisation to get the ball rolling.

By providing transparency on what information your agency holds, it is possible for reusers to provide feedback on your priorities and for other public sector bodies to discover datasets they potentially can reuse (under certain legal restrictions).

How do I implement this Best Practice?

The only major requirement here is that someone has responsibility to perform this for the organisation. Technical requirements are minimal. Some simple tools are necessary to assist in the identification and release of some data early.

Establish an overview of managed datasets in your organization and implement a basic legal examination of the content of each dataset. Use DCAT-APs access rights property to indicate level of openness.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Norway	Difi Traffic Light System	Heather Broomfield, Difi

References

- Samos Workshop Story: Traffic Light System For Data Sharing
- Open Data Institute guidance: The data spectrum
- Open Data Institute guidance: What makes data open?
- Open Data Institute guidance: What are the impacts of non-open licenses?

Localised Guidance

- (Austria) Framework for Open Government Data Platforms
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke o prilagodbi skupova podataka za javnu objavu i ponovno korištenje Open Data Guide, Croatia
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Finland) Avoimen Datan Opas Open Data Guide
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) Open Data Handbook, Solutions Bank
- (Ireland) Guide for publishers
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Malta) PSI Directive Implementation & Internal Data Sharing Platform (draft)
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information

- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (UK) Open Data Resource Pack

Contact Info

Heather Broomfield, Difi (Norway).

Related Best Practices

- Identify what you already publish
- Publish overview of managed data
- Provide data unavailability reference

5.2 Dataset criteria

Outline

This best practice sets out a number of criteria that can be used to prioritise the publication of some datasets ahead of others.

Links to the Revised PSI Directive

Dataset criteria

Challenge

To develop the criteria for 'high-value datasets' taking into consideration the likely reuse of open data and to help governments understand which datasets to prioritise for publication.

Solution

To follow this guidance on dataset criteria which has been developed through engaging with both users and reusers of the data. The characteristics of 'high-value datasets' are seen from three perspectives: re-usability, value for data owners, value for reusers.

Re-usability

- High-value data should reach at least 3-stars on Tim Berners-Lee's 5-star schema (making it available on the Web under an open license in a non-proprietary, structured format).

Value for data owner

A dataset may be considered of high-value when one or more of the following criteria are met:

- sharing it contributes to transparency;
- the publication is subject to a legal obligation;
- the data directly or indirectly relates to their public task;
- sharing it helps with cost reduction.

Value for reusers

The value of a dataset primarily depends on its use and reuse potential, which can lead to the generation of business activity. The potential of the dataset is defined by:

- the size and dynamics of the target audience;
- the number of systems or services that could use the dataset.

Datasets contributing to transparency have a strong social impact and reuser's interest in these datasets is high.

Engaging with Reusers

It is important to engage directly with reusers to understand the value of your dataset.

Recommendations:

- establish a communication channel, for example, with a mailing list or a community on Joinup or on the Open Data Portal that could be used to make announcements to reusers and to gather feedback;
- use collaborative tools. This encourages collaboration between a community or reusers and the cross-fertilisation of ideas and business opportunities.

Why is this a Best Practice?

It's important to have a shared understanding of what can be considered to be high-value datasets so that publication of these datasets can be prioritised.

Understanding which datasets should be published, under what criteria and priority, will help public authorities to see the benefits of publishing more high quality datasets.

How do I implement this Best Practice?

In order to implement this BP, you need:

- an understanding of high-value data,
- communication channels with data users and data reusers.

Where has this best practice been implemented?

Country	Initiative	Contact Point

References

- Timisoara Workshop Talk: Good practices for identifying high value datasets and engaging with reusers: the case of public tendering data
- Timisoara Session: How good is good enough?
- Krems Workshop Session: Specific Critical Success Factors for Open Data Publication and Use
- Open Data Institute white paper: How to prioritise open data to drive global development
- Open Data Institute guidance: Engaging with reusers

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (International) Open Data Handbook, Solutions Bank
- (Ireland) Guide for publishers

- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Latvia) Atvērto datu vadlīnijas Open Data Guidelines
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Malta) PSI Directive Implementation & Internal Data Sharing Platform (draft)
- (Netherlands) Handreiking bij openen van data Guidance on Open Data
- (Romania) Ghid Pentru Publicarea Datelor Deschise Romanian Open Data Guide
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Spain) Government Data Openness and Re-use
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (UK) Open Data Resource Pack
- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Nicolas Loozen, PwC

Related Best Practices

- Develop an Open Data Publication Plan
- Open Up Public Transport Data
- Establish an Open Data Ecosystem

5.3 Develop an Open Data publication plan

Outline

Datasets that are fit for publication as Open Data need to be identified as well as the requirements of both the internal and external stakeholders. Open Data publication plan should be developed taking into account needs of the relevant stakeholders as well as the potential benefits, risks and costs of publication of the datasets.

Links to the Revised PSI Directive

Policies and Legislation

Challenge

How to manage publication of Open Data?

Public sector bodies hold a large number of various datasets. However, they also operate under budgetary and various other constraints that prevent them from making every potential dataset available for reuse at once.

Solution

Develop an Open Data publication plan that balances the requirements and the known constraints. Open Data publication plan should contain all the necessary tasks to implement the Open Data initiative and it should be supported by a roadmap for publication of the selected datasets.

Why is this a Best Practice?

This best practice aims at balancing the requirements and needs of the relevant stakeholders with the available resources by planning of the data release and other necessary steps that takes into account the demand for data, the potential benefits and risks associated with the publication of the identified datasets as well as the estimated costs of the publication.

How do I implement this Best Practice?

An analysis of the available datasets should be performed, needs and requirements of both the internal and external stakeholders should be understood and the Open Data publication plan should be developed. Publication of Open Data should be in line with the strategic aims of the organization, with the relevant policies such as the national Open Data strategy and it should meet the demand of the potential users of the data. Therefore, the requirements of both internal and external stakeholders should be identified and analysed. Open Data publication plan should contain all the necessary tasks to implement the Open Data initiative. With regard to the datasets planned to be published as Open Data a release roadmap should be developed. Release of the datasets should be prioritised taking into consideration the requirements of the stakeholders as well as the results of the analysis of the datasets, i.e. the identified benefits and risks and the cost/effort estimates. Progress should be tracked against the Open Data publication plan. This plan should be also regularly reviewed and updated.

Implementation steps could be summarised as follows:

- Obtain the top management support.
- Appoint a person responsible for development of the Open Data publication plan.
- Secure cooperation of the departments involved in the Open Data initiative.

- Identify requirements of both internal and external stakeholders. Voice of every group should be treated with due regards.
- Develop Open Data publication plan balancing the requirements and the known constraints.
- Develop Open Data release roadmap based on the dataset priorities.
- Set up a review process for the Open Data publication plan.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Austria	Open Government Data Implementation Model	Bernhard Krabina
Spain	Basque Country PSI reuse assessment	Martin Alvarez-Espinar

References

- Samos Workshop Story: Open Data Publication Plan
- Timisoara Workshop Report
- Open Data Institute guidance: Engaging with reusers
- Open Data Institute guidance: How to prioritise data to drive global development

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke o prilagodbi skupova podataka za javnu objavu i ponovno korištenje Open Data Guide, Croatia
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Finland) Avoimen Datan Opas Open Data Guide
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) ELI implementation methodology: Good practices and guidelines
- (International) Open Data Handbook, Solutions Bank
- (Ireland) Guide for publishers
- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Netherlands) Handreiking bij openen van data Guidance on Open Data
- (Romania) Ghid Pentru Publicarea Datelor Deschise Romanian Open Data Guide
- (Serbia) Open Data Handbook
- (Spain) Government Data Openness and Re-use

- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (Sweden) Vidareutnyttjande av information Om PSI och öppna data Reuse of PSI and open data
- (UK) Open Data Resource Pack
- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Jan Kučera, University of Economics, Prague (Czech Republic).

Related Best Practices

- Categorise openness of data
- Dataset Criteria
- Identify what you already publish
- Publish overview of managed data
- Establish an Open Data Ecosystem

5.4 Develop and implement a cross agency strategy

Outline

Developing and implementing a strategy on open data that coordinates the efforts of multiple agencies.

Links to the Revised PSI Directive

Policies and Legislation

Challenge

Different agencies in the public sector develop and implement their own strategy. These different strategies are unconnected so it is hard to apply central planning and evaluation.

Solution

There needs to be a strategy that coordinates the efforts of multiple agencies.

Why is this a Best Practice?

This best practice allows different agencies to understand what is required, plan accordingly and measure progress. Having a strategic plan is essential for decision makers at the highest level, i.e. ministers, to support an approach to implementation. Both the G8 Open Data Charter, published in 2013, and the Shakespeare Review of Public Sector Information, emphasised the need for a clear, visible, auditable plan for publishing data as quickly as possible, defined both by bottom-up market demand and by top-down strategic thinking, overcoming institutional and technical obstacles with a twin-track process which combines speed to market with improvement of quality:

- an ‘early even if imperfect’ track that is very broad and very aggressively driven, and
- a ‘National Core Reference Data’ high-quality track which begins immediately but narrowly;

and then moving things from Track 1 to Track 2 as quickly as we can do reliably and to a high standard. ‘Quickly’ should be set out by government through publicly committed target dates.

How do I implement this Best Practice?

- Give responsibility to an individual civil servant or department for developing the plan.
- Convene a meeting, or a series of meetings, between stakeholders - data producers, data users etc.
- Develop the plan through an iterative process before seeking high level endorsement.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Flanders, Belgium	The Flemish Innovation Projects	Noël van Herreweghe, Program Manager Open Data – Government of Flanders in Belgium

References

- Samos Workshop Talk: The Flemish Innovation Projects: promoting innovation through encouraging the use and reuse of government datasets
- Open Data Institute guidance: How to plan and budget an open data initiative
- Open Data Institute white paper: Open data in government, how to bring about change
- Open Data Institute method report: Supporting public sector open data leadership
- Open Data Institute technical paper: The Open Data Maturity Model and Pathway

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke o prilagodbi skupova podataka za javnu objavu i ponovno korištenje Open Data Guide, Croatia
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Estonia) Eesti avaliku teabe masinloetava avalikustamise roheline raamat Green Paper on machine-readable Estonian Public Information disclosure
- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) Open Data Handbook, Solutions Bank
- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Malta) PSI Directive Implementation & Internal Data Sharing Platform (draft)
- (Norway) Veileder i tilgjengeliggjøring av offentlige data Guide to making public data available
- (Romania) Ghid Pentru Publicarea Datelor Deschise Romanian Open Data Guide
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Spain) Reutilización de la Información de los Servicios Públicos Open University Development Guide
- (Spain) Guía metodológica para planes open data sectoriales Methodological Guide for Sectorial Open Data Plans
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use

- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Noël van Herreweghe, Program Manager Open Data – Government of Flanders in Belgium

Related Best Practices

- Develop an Open Data Publication Plan
- Establish an Open Data Ecosystem

5.5 Enable feedback channels for improving the quality of existing government data

Outline

The goal of this best practice is to improve the quality of government data by enabling feedback channels for users to report errors, inconsistencies, and incompleteness in already published data. It is aligned with the European Commission notice 2014/C 240/01 paragraph 3.2 that states: "To facilitate the use of data in the public sector while significantly increasing the value of datasets for subsequent reuse, it is recommended that datasets be: [...] subject to regular feedback from reusers (public consultations, comments box, blogs, automated reporting, etc.) to maintain quality over time and promote public involvement."

Links to the Revised PSI Directive

Policies and Legislation, Platforms

Challenge

Often reusers of governmental datasets make copies in order to curate their local copy of the data (for example, fixing errors or completing data). On the one side, this approach is not optimal as it leads to duplication of efforts and reduces the possibility of sharing and reuse, and on the other, the publisher would like to know about the user needs and the benefits of opening the data. Hence, the following questions can be raised by the publishing institution:

- Who is accessing and using my data? Does the data fulfil their need?
- What are their experiences?
- What is the true value of my data?
- Can we increase our revenue by providing better services to our customers?

Solution

Responsible public sector bodies should provide feedback mechanisms through which stakeholders can identify mistakes and correct them where possible. One possible practical approach could be to use a distributed versioning system for the published data, such as GitHub, in order to improve open data as is common for open source software. Alternatively, provide a simple feedback loop using comment boxes, forums etc. Public sector bodies should actively encourage stakeholders/reusers to use these mechanisms.

Why is this a Best Practice?

Anyone using data, whether they're part of the organisation that creates it or an external reuser, will want the data to be accurate. This is difficult to achieve, requiring time and effort - i.e. it can be expensive. By creating mechanisms through which datasets can be corrected by the community of users, the cost is distributed. Although the data itself may not be crowd-sourced, its curation can be, to the benefit of everyone, including the publisher.

How do I implement this Best Practice?

The publisher needs tools for collecting feedback. Innovative ways such as crowdsourcing can be used for collecting and improving the quality of existing government data sources.

Issue tracking and bug reporting platforms are commonplace in open source software projects and such tools can readily be used for collecting feedback about datasets.

Where has this best practice been implemented?

The approach is applicable to any Member State. For practical examples please, check the information below.

Country	Initiative	Contact Point
Poland	Implementation of the revised reuse Directive in Poland, Open Data Portal	Jacek Wolszczak, Ministry of Administration and Digitisation, Poland
Scotland	ALISS service	Peter Winstanley, Scottish Government, UK
EU	FP7 ENGAGE project: feedback mechanisms based on Web 2.0	Charalampos Alexopoulos, University of the Aegean, GR

References

- Heather Broomfield, Agency for Public Management and eGovernment, The Norwegian Public Roads Administration Case Study, The Norwegian Meteorological Institute Case Study, Norway
- Mateja Prešern, Ministry of Public Administration, Supervizor application, Reuse of PSI with a goal of strengthening the integrity and transparency, Slovenia
- Lorenzo Canova, Antonio Vetrò, Marco Torchiano, Raimondo Iemma & Federico Morando, Politecnico di Torino, OpenCoesione and Monithon - a Transparency Effort, Italy
- Benedikt Kotmel, Ministry of Finance, Experiences of identifying datasets for sharing, Czech Republic
- Jacek Wolszczak, Ministry of Administration and Digitization, Implementation of the revised reuse Directive in Poland, Poland
- Amanda Smith & Sumika Sakanishi, ODI, Publishing and improving the quality of open data with Open Data Certificates, United Kingdom
- Charalampos Alexopoulos, Euripidis Loukis, Yannis Charalabidis, 2014, A Platform for Closing the Open Data Feedback Loop based on Web2. 0 functionality, JeDEM 6 (1): 62

Localised Guidance

- (Austria) Framework for Open Government Data Platforms
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke o prilagodbi skupova podataka za javnu objavu i ponovno korištenje Open Data Guide, Croatia
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Estonia) Avaandmete loomise ja avaldamise juhend Open Data Guidelines

- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Latvia) Atvērto datu vadlīnijas Open Data Guidelines
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Malta) PSI Directive Implementation & Internal Data Sharing Platform (draft)
- (Norway) Veileder i tilgjengeliggjøring av offentlige data Guide to making public data available
- (Romania) Ghid Pentru Publicarea Datelor Deschise Romanian Open Data Guide
- (Serbia) Open Data Handbook
- (Spain) Reutilización de la Información de los Servicios Públicos Open University Development Guide
- (Spain) Government Data Openness and Re-use
- (Spain) Guía metodológica para planes open data sectoriales Methodological Guide for Sectorial Open Data Plans
- (Spain) Decálogo Open Data Open Data Decalogue
- (Spain) Guía para el desarrollo de la Universidad Abierta Open University Development Guide
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (Sweden) Guiding principles for working with digital cultural heritage
- (UK) Open Data Resource Pack
- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Valentina Janev, Institute Mihajlo Pupin

Related Best Practices

- Enable quality assessment of open data
- Gather feedback from data consumers

5.6 Enable quality assessment of open data

Outline

Data Quality DQ is primarily perceived to be a subjective term: What suffices, is “good enough” for one person, might be inferior to another. “Suffice” here means to be suitable to fulfil a certain need in a process. However beside the subjective aspect of DQ, there is an objective view on DQ which can be measured and help to establish provable and comprehensible metrics on DQ. The adherence to standards, enforced by tools which in turn are embedded in and used by processes, will help to raise DQ. In order to sustainably raise DQ, measures need to be in place all along the data pipeline and not only at the providing front end. DQ improvement has to be considered as a process rather than a one-time measure.

Links to the Revised PSI Directive

Quality

Challenge

The proliferation of open data as a mean to foster open innovation processes towards improved or new products and services, to increase transparency and to perform self-empowered impact measurement of policies also raises concerns about the quality of the provided resources. The early assumption that more data, even of uncertain origin and quality, will unconditionally result in better decisions as long as the right algorithms are used, gave again way to the insight that the principle of garbage-in, garbage-out still holds true. This fact combined with raising concerns regarding data platform usability, data literacy and trust put the quality aspect into the focus. Ironically government Data Quality became of an issue lately primarily due to the fact that government started to release data sets as Open Data which enables stakeholders to carry out citizens control rights. Bringing together data from diverse sources for the first time partially makes data issues like missing data obvious, but even more so deficiencies which arouse due to lacking or missing Master Data Management..

Solution

Traditional metrics to assess Data Quality like accuracy, applicability, and understandability remain relevant, and in the realm of Open Data, get extended by measures like openness, timeliness and primacy. Work carried out in the European Commission's Open Data Support project suggests seven aspects to consider:

- 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?

DQ improvement measures have to be in place all along the (open) data life cycle, otherwise quality measures will be perceived to be an additional burden, causing efforts and costing money. Also note, that the Open Data Life Cycle is - a cycle which suggest to set up data improvement measures as a process rather than a one time measure.

Why is this a Best Practice?

Lacking DQ will reduce data users trust and prevent the unfolding of an open data market. Investment into DQ will pay back internally to the administration, as the potential for interoperable data services will be risen as well as externally, as for data users it will become more easy to blend together data sets of diverse sources to create added value services.

How do I implement this Best Practice?

Implementation of this BP requires addressing the problem from a technical as well as organisational perspective.

Technically, DQ can be raised by adhering to conventions, norms and standards. However, the adoption of conventions, norms and standards requires governance at various levels. Set-up of governance structures is typically in the responsibility of the CIO or someone in charge with comparable powers and duties.

- It's within the CIO's responsibility to provide guidance on how to structure and implement ICT-systems, which use common and agreed conventions, norms and standards.
- The CIO should be responsible for identifying semantically equivalent data entities, describe standards according to which these data entities should be modelled and monitor the adherence to these standards.

Common data entities, where possible, should be modelled according to the core vocabularies.

CSV files could be annotated using W3C's CSV on the Web Recommendations, which also included a formalised model to describe the columns of CSV files.

Data descriptions should be made according to the DCAT-AP vocabulary.

During the data publishing stage, the W3C Data Quality Vocabulary (DQV) can be used. This provides a framework in which the quality of a dataset can be described either by the publisher or the wider audience.

Tools can automatically check a certain range of DQ domains, like adherence to claimed encodings (such as utf8) or the structural regularity of CSV files.

For assessing the quality of the dataset itself prior to publishing, e.g. for publishing statistical data in RDF format an RDF Data Cube validator (PDF) can be used.

To enrich the data with quality assessment information and track provenance in RDF integration process, e.g. the UnifiedViews tool can be used.

Organisation-wise

- The CIO should implement a data governance framework which comprises data architecture management, meta-data management, and master data management (MDM).

- The importance of data as a mission-critical asset can be risen by establishing the role of the Chief Data Officer (CDO).
- The principles of ISO 8000, like vocabulary usage, semantic encoding, provenance, accuracy and completeness can be taken into account.

The obligatory usage of minimum widespread technical standards like utf8 could be enforced by legal measures or order of the federal CIO.

To assess the publishing process, consider the steps described by ODI Certificates (or similar).

Further reading

- Data Quality Vocabulary
- Introduction to ISO 8000
- Data Management Body of Knowledge
- Standards on eProcurement
- Standards on eInvoicing
- Open Data Certificates

Where has this best practice been implemented?

The approach is applicable to any Member State. For practical examples please, check the information bellow.

Country	Initiative	Contact Point
Austria	Mission Statement of the Sub-working Group Quality Assurance of Open Data Portals of the Cooperation Open Government Data Austria	Cooperation OGD Austria
Finland	Valmistele ja avaa - Prepare and open Section 3.6. Tiedon viimeistely ja laatu - Finishing the data and data quality	Prime Minister's Office Finland
Serbia	Validating RDF Data Cube Models	Valentina Janev, Mihailo Pupin Institute, University of Belgrade, Belgrade, Serbia
UK	ODI Certificate for the Westminster City Council	Westminster City Council
UK	Cross platform character encoding profile	

References

- David Corsar, Peter Edwards, Enhancing Open Data with Provenance, dot.rural Digital Economy Hub
- ProvenanceWeek 2014

- Giorgos Flouris, Yannis Roussakis, Marrıa Poveda-Villalon, Pablo N. Mendes, Irini Fundulaki, Using Provenance for Quality Assessment and Repair in Linked Open Data, 2nd Joint Workshop on Knowledge Evolution and Ontology Dynamics (EvoDyn-12) at the ISWC2012
- Makx Dekkers, AMI Consult, How good is good enough?
- Amanda Smith & Sumika Sakanishi, ODI, Publishing and improving the quality of open data with Open Data Certificates, United Kingdom
- Samos presentation: Examples from the Norwegian public sector
- Lisbon workshop session: Roadblocks in Commercial Open Data Usage
- Timisoara workshop session: How good is good enough? A common language for quality?
- Comparing the 5-star scheme with Open Data Certificates
- Lisbon workshop session: Roadblocks in Commercial Open Data Usage
- Samos Workshop Session: The Potential within the Government for Innovation and Efficiency from Open Data – Examples from the Norwegian public Sector

Localised Guidance

- (Austria) Framework for Open Government Data Platforms
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke o prilagodbi skupova podataka za javnu objavu i ponovno korištenje Open Data Guide, Croatia
- (Estonia) Avaandmete loomise ja avaldamise juhend Open Data Guidelines
- (Finland) Avoimen Datan Opas Open Data Guide
- (International) Open Data Handbook, Solutions Bank
- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Malta) PSI Directive Implementation & Internal Data Sharing Platform (draft)
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Spain) Government Data Openness and Re-use
- (UK) Open Data Resource Pack
- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Original Authors: Johann Höchtl, Valentina Janev

Contributors: Muriel Foulonneau, Lorenzo Canova

Editors: Valentina Janev, Johann Höchtl

Related Best Practices

- Enable feedback channels for improving the quality of existing government data
- Provide data provenance information
- Provide versioning information
- Reuse vocabularies

5.7 Encourage crowdsourcing around PSI

Outline

Preparing PSI for sharing can be time consuming, expensive and, sometimes difficult. Engaging the community in the task will increase the quality and quantity of available data as well as enthusing the potential users.

Links to the Revised PSI Directive

Policies and Legislation, Platforms

Challenge

To increase the quality and quantity of machine readable data within a constrained budget.

Solution

Crowdsourcing can be an efficient way to increase quality and availability of machine readable data, in particular for cultural heritage institutions. Innovative techniques, including gamification, can be used to harness the skill and enthusiasm of the community at large. On a practical level, datasets can be made available on platforms such as GitHub so that users can offer corrections (accepting such corrections remains under the control of the data owner). This is the approach undertaken by the City of Chicago. On a policy level, identifying community crowd sourcing projects outside government institutions can also be an indicator of valuable datasets that should be prioritised for open publication since the level of community involvement is generally proportional to the level of interest in that data.

Why is this a Best Practice?

Many institutions lack resources necessary to manually go through large collections of unstructured data that has been created over many years (e.g. in the cultural heritage sector). By engaging external communities to collaborate on this data it is possible to create more detailed machine readable data supporting a wider range of reuse cases.

More machine readable open data supports a wider range of use-cases in services and applications.

- Many institutions lack resources necessary to manually go through large collections of unstructured data
- By engaging external communities to collaborate on this data it is possible to create more detailed machine readable data supporting a wider range of reuses.
- Crowdsourcing engages the community that the end product serves.

How do I implement this Best Practice?

- Identify the exact need first and then seek groups able to support solving that need via crowdsourcing.
- Think of crowdsourcing as another tool to create/improve data sets and think about the phases of your data collection project and where crowdsourcing could best fit in.
- Involve stakeholders who could benefit from a free source of certain data sets and have them provide funding in order to sustain crowdsourcing efforts.

- Tasks need to be small to be able to be completed by volunteers with limited time.
- Utilize a gamification approach if possible, that is, by playing a game, users perform a useful task.
- It is possible to use crowdsourcing without the user's knowledge. The best known example of this is the use of CAPTCHAs to solve the micro task of reading words that optical recognition software cannot and by that method digitising hard to read texts.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Sweden	Guiding principles for digital cultural heritage	Digisam

References

- Dimitris Paraschakis, Crowdsourcing cultural heritage metadata through social media gaming, 2013, Malmö University
- Krems Workshop Session: Towards A Sustainable Austrian Data Market

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) Open Data Handbook, Solutions Bank
- (Latvia) Atvērto datu vadlīnijas Open Data Guidelines
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information

Contact Info

Peter Krantz peter@peterkrantz.se

Related Best Practices

- Enable feedback channels for improving the quality of existing government data

5.8 Establish an Open Data ecosystem

Outline

Organisations are taking giant steps towards freeing up public data with the aim of making it available for everyone to access this data and information, for personal as well as for business use. The uptake of this data and information has not been what was expected, however. Simply making data available to the public isn't enough to make that data useful. Citizens are not interested in data: they are interested in services being built with the available data and information. Therefore, more needs to be done: the establishment of an active open data network, an ecosystem to facilitate the uptake of data and information for reuse.

Links to the Revised PSI Directive

Organisation, Platforms, Dataset criteria, Charging, Reuse, Selection, Discoverability

Challenge

The open data landscape consists of different actors and users with different needs, challenges, problems, expectations, roles and envisaged opportunities. A strategy to coordinate all of those internal and external actors, stakeholders and interested parties, is needed. Only then can they be convinced to share data and information and make those available for reuse; only then will all involved know what can/must be available in order to realize innovation efforts and new products and services envisaged in the PSI directive.

Solution

An active network is needed, an "open data ecosystem," facilitating interaction and communication amongst everybody interested and/or involved in open data and the reuse of information and data, internal, as well as external to the organisation. Within such a regional, national and international ecosystem, information exchange such as success stories will increase visibility and therefore awareness with regards to the availability of data and information for reuse and the results achieved by doing so. Small infomediary companies could act as intermediaries between all involved parties and with other market segments. Examples can be found in Spain where over 80% of infomediary companies are older than 5 years old and generate approximately 4,500 to 5,500 jobs, mostly linked to ICT: analysis, processing and presentation of information [DH].

Why is this a Best Practice?

Most open data initiatives in most EU member states have focussed on making sure everything is in place on the supply side of opening up their data and information. Experience in many member states shows that it is not sufficient to just make the data open, i.e. simply making data available to the public isn't enough to make that data useful. All agree that focusing on bringing about the necessary strategic, legal and technical aspects with regards to the implementation of open data isn't enough either. Open data requires an approach based on the administration releasing it, the kind of data being released and, perhaps most importantly, it's targeted audience. Interaction, communication, contacts with journalists, raising awareness, training, collaboration, discussions etc. is needed to foster the development of new innovative products and services, thereby helping to realise the goals of the PSI Directive such as stimulating economic and business activity and ultimately providing value for society as a whole. Initiating an open and constructive dialogue could be the

beginning of an open data ecosystem, giving rise to sustainable business models of solutions making use of Public Sector Information, hopefully in combination with data and information from the private sector. Mostly forgotten, but not less important; this best practice will also achieve efficiency gains through sharing data inside and between public administration, fostering participation of citizens in political and social life and increasing transparency of government.

Implementing this best practice could help:

- To get the most out of scarce public resources that are available in our member states.
- To increase reuse, by the private and also by the public sector.
- To focus the government agencies' efforts on offering the relevant information.
- To improve available dataset formats, the information quality and its accessibility.
- To promote better national regulations on PSI.
- To enhance coordination between public and private sector in PSI reuse.
- To bring about real added value to the development of open data in the member states and help communities to enter the open data movement or improve on it.
- To use what already exists.
- To make use of best practices realised in other member states.
- To realise speed, scale of economy, cooperation.

How do I implement this Best Practice?

If member states want to adopt this Best Practice they will have to:

- make sure everything is in place to facilitate the implementation of open data (strategy, content, legal and technical);
- be willing to adopt a bottom-up approach to implement an open data program;
- bring about an open data ecosystem by listening to all stakeholders and interested parties;
- dedicate the necessary people and resources in realising this ecosystem;
- be willing to share data inside and between public administrations;
- establish the necessary follow-up procedures towards continuity.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Albania	Utilization Cases of Open Data Albania	Julia Hoxha – Albanian Institute of Science, Tirana, Albania. Aranita Brahaj – Albanian Institute of Science, Tirana, Albania. Benedikt Kämpgen – Institute AIFB, Karlsruhe Institute of Technology, Germany.

References

- Utilization Cases of Open Data Albania, Julia Hoxha and Aranita Brahaj, Institute AIFB, Karlsruhe Institute of Technology, Karlsruhe, Germany and Albanian Institute of

Science, Tirana, Albania. Presented at Share-PSI Samos Workshop Uses of Open Data Within Government for Innovation and Efficiency, July 2014
- Lisbon workshop presentation: Spanish Infomediary Sector Characteristics (PDF), M. D. Hernández, Ministerio de Hacienda y Administraciones Públicas. Madrid, Spain.
- Samos Workshop Talk: Open Government Data Austria - Organisation, Procedures and Uptake
- Samos Workshop Talk: The Flemish Innovation Projects: promoting innovation through encouraging the use and reuse of government datasets.
- Lisbon Workshop Session: Events, hackathons and challenge series - stimulating open data reuse
- Lisbon Workshop Session: Open Data Economy: from 'Wow' to 'How'
- Timisoara Workshop Talk: EU actions on Open Data – current policy and legal context
- Krems Workshop Talk: An Ongoing Open Dialog in an Open Data Ecosystem
- Berlin Workshop Talk: Government as a Developer - Challenges and Perils
- Open Data Institute guidance: Engaging with reusers
- Open Data Institute method report: Peer networks for open data leaders
- Open Data Institute white paper: How to improve agriculture, food and nutrition with open data

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (International) ELI implementation methodology: Good practices and guidelines
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (Latvia) Atvērto datu vadlīnijas Open Data Guidelines
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Norway) Veileder i tilgjengeliggjøring av offentlige data Guide to making public data available
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Spain) Government Data Openness and Re-use
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (UK) Open Data Resource Pack

- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Noël van Herreweghe, Program Manager Open Data – Government of Flanders in Belgium

Related Best Practices

- Develop and Implement a Cross Agency Strategy
- Support Open Data Start Ups
- Encourage crowdsourcing around PSI
- Establish Open Government Portal for data sharing

5.9 Establish Open Government Portal for data sharing

Outline

Data portals can facilitate the distribution of open data by providing easy-to-access, searchable hub for multiple datasets. They often also act as showcases for reuse of data and as a hub for the interested community.

Links to the Revised PSI Directive

Platforms, Techniques, Discoverability

Challenge

Public administrations consist of government bodies of differing sizes and often their information systems lack interoperability. Data sets are of different sizes sometimes relating to small localities only, which makes them less interesting to the broader community such as the infomediary sector. Furthermore, the datasets are in different formats which complicates their reuse.

Solution

A data portal is established by the government (national, regional or local). This may be managed in-house or by a contractor. Close collaboration with the R&D or education sector can help to meet fundamental goals at the beginning of the portal project. Researchers and computer engineering students are involved in direct reuse of data in development of innovative services (web and mobile applications).

Why is this a Best Practice?

A portal fulfils many functions:

- it acts as a platform through which datasets are made available, catalogued and made searchable;
- it promotes the provision of metadata and makes it easy for that metadata to be added at the time of publication;
- it acts as a showcase for applications that reuse the data;
- it can also act as a community hub.

How do I implement this Best Practice?

Data portal software can be developed from scratch, bought off the shelf or obtained as open source software. The best known example of an open source package is CKAN.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Spain	Canaries Open Data Portal	José Luis Roda-García, www.opendatacanarias.es

References

- Samos Workshop: Open Data to Improve Sharing and Publication of Information between Public Administrations PDF)
- Samos Workshop: A federation tool for opendata opendata portals (PDF)
- Samos Workshop Talk: Open Government Data - Fostering Innovation (PDF)
- Krems Workshop Session: Italian National Guidelines for the Valorization of the Public Sector Information (PDF)
- Berlin Workshop Session: European Interoperability: The ISA Core Vocabularies (PDF)
- Berlin Workshop Session: The EDP: A Technical View (PDF)
- Lisbon Workshop Talk: Model-Driven Engineering for Data Harvesters
- Timisoara BarCamp Talk: The Pan European Data Portal - Early Wireframes
- Berlin Workshop Talk: The European Data Portal - Opening up Europe's Public Data
- Berlin Workshop Talk: The EDP: A Technical View
- Berlin Workshop Talk: The Role of the Portal

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke_o_prilagodbi_skupova_podataka_za_javnu_objavu_i_ponovno korištenje Open Data Guide, Croatia
- (CzechRepublic) Standardy_publikace_a_katalogizace_otevřených_dat_veřejné_správy ČR Open Data Standards
- (Estonia) Eesti_avaliku_teabe_masinloetava_avalikustamise_roheline_raamat Green Paper on machine-readable Estonian Public Information disclosure
- (Germany) Open Government Data Deutschland
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (Ireland) Guide for publishers
- (Italy) Linee_Guida_Nazionali_per_la_Valorizzazione_del_Patrimonio_Informativo Pubblico National Development Guidelines for Public Sector Information
- (Latvia) Atvērto datu vadlīnijas Open Data Guidelines
- (Lithuania) Viešojo_Sektoriaus_Informacijos_platinimo_gerosios_praktikos The Best Practices for the PSI publication
- (Malta) PSI Directive Implementation & Internal Data Sharing Platform (draft)
- (Norway) Veileder i tilgjengeliggjøring av offentlige data Guide to making public data available
- (Romania) Ghid Pentru Publicarea Datelor Deschise Romanian Open Data Guide

- (Serbia) Open Data Handbook
- (Spain) Reutilización de la Información de los Servicios Públicos Open University Development Guide
- (Spain) Government Data Openness and Re-use
- (Spain) Decálogo Open Data Open Data Decalogue
- (Spain) Guía para el desarrollo de la Universidad Abierta Open University Development Guide
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (UK) Open Data Resource Pack

Contact Info

José Luis Roda-García, Director of the Canaries' Open Data Platform-University of La Laguna, Tenerife, Spain.

Related Best Practices

- Develop and Implement a Cross Agency Strategy
- Develop a federation tool for open data portals
- Provide metadata

5.10 High level support

Outline

Open data actions must be supported by senior officials.

Links to the Revised PSI Directive

Policies & Legislation, Organisation

Challenge

All open data actions need support at high level to be effective at overcoming those objections and, no less importantly, securing resources and refocusing priorities. When done correctly, the sponsor could also start advocating the actions him/herself.

Solution

Open data actions **MUST** be supported by senior officials who are empowered to provide top down authority where required.

Why is this a Best Practice?

Public sector bodies normally have their own procedures for securing high level support, such as ministerial support. However, such support is often most readily made public by the inclusion of an introduction to the document by the senior official. Written support in other formats (e.g., press releases, internal e-mail, social media, web page etc.) could also be beneficial for your open data actions.

How do I implement this Best Practice?

To reach high level support for your open data action, you should:

- Understand, whose support do you mostly need? How broad is the action? Who are your main targets? E.g., national government strategies will typically need ministerial support, ideally including the Head of Government him/herself. When you need to reach IT-specialist, a government/ministerial CIO is more suitable.
- Be very simple and clear in the message that you need from him/her. It's best that you have a fine draft ready for reviewing.
- Think of ways to reach him/her. It could be just as simple as a e-mail or could include several meetings with lower officials to convince them of this need
- Ask for permission to use his/her name in different media. Most probably you have to coordinate it with his/her communication experts.
- Have the name of the senior official supporting the open data action in question readily discoverable. You can also include a quote from him/her also in a press release, internal e-mail, social media, web page etc.

Where has this best practice been implemented?

This best practice has the following implementation examples:

- In Flanders, since 2011 Open Data and Shared Services have been at the center of the eGovernment strategy. The Open Data Programme was set up that realising that in a well-functioning, transparent and democratic society, citizens and business must be able to access government data and information and to share and reuse that information, freely and with minimal restrictions. The government decided on a top-down approach to implement the open data strategy. A concept note was drawn up and signed by all ministers.
- In Germany, successful open government needs broad political commitment at all levels of government due to its cross-level importance and the necessary cultural transformation of policy-making and administration. It will be especially important to find a powerful sponsor and further high-level supporters for the project who identify with the development of open government.

Country	Initiative	Contact Point
Belgium	The Flemish Open Data Program	noel.vanherreweghe@bz.vlaanderen.be
Germany	Open Government Data Germany (EN, short version)	jens.klessmann@fokus.fraunhofer.de

References

- The Flemish Open Data Program, Noël Van Herreweghe

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke o prilagodbi skupova podataka za javnu objavu i ponovno korištenje Open Data Guide, Croatia
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Estonia) Eesti avaliku teabe masinloetava avalikustamise roheline raamat Green Paper on machine-readable Estonian Public Information disclosure
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Malta) PSI Directive Implementation & Internal Data Sharing Platform (draft)
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Sweden) Ramverket för öppna data - SKL The Framework for Open Data
- (Sweden) Vidareutnyttjande av information Om PSI och öppna data Reuse of PSI and open data
- (UK) Open Data Resource Pack

Contact Info

Noël van Herreweghe, Program Manager Open Data – Government of Flanders in Belgium.

Related Best Practices

- Cross Agency Strategy

5.11 Holistic metrics

Outline

The costs of public sector information in respect to added value has to be assessed taking into account large-scale detour effects and not merely at the level of the publishing organisation. This BP clarifies why this is a best practice, what can be the obstacles and approaches to actually implement the BP.

Links to the Revised PSI Directive

Policies and Legislation, Charging

Challenge

The publication of data and information according to regulations, principles, best practices or recommendations generally has a positive effect. However, in cases where the publisher creates a substantial revenue by monetizing the concerned data, the loss of income represents a hard to deny fact which may contribute to the decisions against publishing data or information. Therefore more sophisticated assessments are required which are suitable to justify loss of income on one level in exchange to gained overall benefits.

Solution

A range of metrics taken at a higher level will often show significant benefits to the organisation as a whole, such as greater efficiency, improved fulfilment of the public task and increased transparency.

Why is this a best practice?

Generally, assessing the benefits of actions exclusively at local scale will lead to micro-optimisations and missed opportunities at the larger context. This is not only disadvantageous in the case of deciding on opening up data and information, but is an administrative leftover from times where holistic measure was mostly impossible due to non-existent integrated Information Systems which can provide a quick and comprehensive overview on policy making.

How to implement this Best Practice?

The actual implementation of public sector value assessment is very much influenced by the administrative organisational setup. If and open data strategy is implemented at the federal level, chances are high, that impact assessment will also take place at that level.

To facilitate the transition from assessing efforts and value from the local level to a higher level, techniques and methodologies of management by objectives, rigorous data analytics by using dashboards, data mining and predictive modeling as means of evidence based government could be used.

The actual implementation will further depend on the level where the PSI directive got implemented. While all EU member states are required to implement EU directives into national law, some member states forward this obligation to provinces (mostly on NUTS-2 level), which adds additional difficulty when erecting holistic measures of costs and effects of data and information publication. Some individual departments, or sub-departments, are

likely to see increased costs with no direct benefit to that department but at a higher level, the benefits should be evident and measurable.

Further Reading

- Impact Monitoring Framework für Open Government Data Ein Impact Monitoring Framework für Open Government Data am Beispiel von »OGD Schweiz« (DE) PDF
- From Evidence-Based Policy Making to Policy Analytics (PDF)
- Big Data and Analytics, [1 PDF]
- The impact of open data, Mininstry of Finance, Finland. Heli Koski, Elinkeinoelämän tutkimuslaitos

Where has this best practice been implemented?

Country	Initiative	Contact Point
Austria	Wirkungsorientierte Steuerun	ABTEILUNG III/9: WIRKUNGSCONTROLLINGSTELLE DES BUNDES, VERWALTUNGSINNOVATION
United Kingdom	Performance UK	
Finland	Government´s analysis, assessment and research activities	Prime Minister's Office Finland
Italy	Impact of open data measured by indicators (in progress)	Giorgia Lodi

References

- Samos Workshop Session: The Potential within the Government for Innovation and Efficiency from Open Data – Examples from the Norwegian public Sector
- Berlin Workshop Session: The Impacts of Open Data: Towards Ex Post Assessment
- Open Data Institute method report: Assessment tools for open data initiatives
- Open Data Institute technical report: Benchmarking open data automatically

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Estonia) Avaandmete loomise ja avaldamise juhend Open Data Guidelines

- (Finland) Avoimen Datan Opas Open Data Guide
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) ELI implementation methodology: Good practices and guidelines
- (International) Open Data Handbook, Solutions Bank
- (Ireland) Guide for publishers
- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Spain) Government Data Openness and Re-use
- (Spain) Guía metodológica para planes open data sectoriales Methodological Guide for Sectorial Open Data Plans
- (Spain) Decálogo Open Data Open Data Decalogue
- (Spain) Guía para el desarrollo de la Universidad Abierta Open University Development Guide

Contact Info

Editor: Johann Höchtl, Danube University, Krems
Contributors: Anne Kauhanen-Simanainen, Ministry of Finance, Finland, Giorgia Lodi, AgID, Italy

Related Best Practices

- Establish an Open Data Ecosystem
- Develop an Open Data Publication Plan

5.12 Identify what you already publish

Outline

Organisations might find deciding what information resources should be made available for reuse in machine-readable formats challenging. Information already published by organisations represent a good candidate for datasets to be published as open data. Therefore, organisations should create and maintain inventory of already published information. However, the amount of such information is often too large to be catalogued manually. Therefore, automated scraping techniques should be applied to create inventories of already published information.

Links to the Revised PSI Directive

Selection

Challenge

Where to start when deciding what information resources should be made available for reuse in machine-readable formats?

Identifying what information should be made available in machine-readable formats for reuse might be challenging due to the lack of knowledge what information is already published and the amount of information might be too large to be catalogued manually.

Solution

An inventory or catalogue of already published data and information assets should be developed and maintained. This may be achieved manually or by using automated scraping techniques to gather details of information assets that are already published on the Web site.

Why is this a Best Practice?

Information is frequently published under a distributed process using a content management system. Inventory of already published information assets might be missing and it might be difficult to create it manually. Therefore, organisations might find it challenging to know where to start when deciding what information resources should be made available for reuse in machine-readable formats.

An inventory of already published information helps organisations to understand what information they provide and what assets they can make more re-usable. Understanding of what datasets an organisation can possibly publish as open data is essential for selecting datasets for publication. Techniques such as site scraping allow organisations to periodically audit their Web site in order to assess what information assets they publish and in what form (open, closed, etc.).

How do I implement this Best Practice?

A simple spreadsheet might serve as an inventory of the data/information assets, but depending on the volume of information and the requirements of the organisation, cataloguing solutions such as CKAN might be deployed.

Scraping software/libraries are needed, such as Scrapy. Metadata gathered using the scraping software can be used as facets for sorting and grouping the links. Faceted browsing

features are provided by application such as Exhibit. If security is the concern, the scraper should be run on an isolated machine and only the headers should be processed.

Whether created manually or by automated means, the inventory should contain at least basic metadata about the data/information assets like the title, location, current format and terms of use. Additional metadata like the responsible person/unit, target data format or update frequency might help to manage the future publication process and it helps to make more precise estimates of the effort and costs needed to publish and maintain the open datasets.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Scotland	The Scottish Government	Dr Peter Winstanley, The Scottish Government

References

- Timisoara Workshop Session: Identifying what you already publish
- Krems Workshop Session: Extracting Structured Data from Unstructured Open Data

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke o prilagodbi skupova podataka za javnu objavu i ponovno korištenje Open Data Guide, Croatia
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (International) Open Data Handbook, Solutions Bank
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Malta) PSI Directive Implementation & Internal Data Sharing Platform (draft)
- (Netherlands) Handreiking bij openen van data Guidance on Open Data
- (Serbia) Open Data Handbook
- (Spain) Guía metodológica para planes open data sectoriales Methodological Guide for Sectorial Open Data Plans
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (Sweden) Vidareutnyttjande av information Om PSI och öppna data Reuse of PSI and open data
- (UK) Open Data Resource Pack

- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Dr Peter Winstanley, The Scottish Government

Related Best Practices

- Publish overview of managed data
- Categorise openness of data
- Dataset Criteria
- Develop an Open Data Publication Plan

5.13 Open Data business models and value disciplines

Outline

Open data holds considerable economic and social value beyond the walls of the governments and institutions that share their data. In response to the economic opportunities presented by the increasing availability of open data, a business model needs to be developed by any open data-driven organisation (at all levels). This will describe how value is created and captured through the decisions made and the resulting consequences.

Links to the Revised PSI Directive

Organisation, Reuse, Charging

Challenge

There are still many open data-driven organisations, especially at midstream and downstream levels, that are struggling to comprehend how to generate revenue and survive by adapting to the changes brought on by the ubiquitous growth of open data and 'Big Data.' In addition, open data-driven organisations have difficulty distinguishing different business models and understanding which one suits their organisational goal.

Solution

To exploit the value of open data, to maximise the benefits, and to enable the creation of innovative products and services, data-driven organisations should develop and implement a business model before starting their business. This is required to ensure that the products and services generate necessary value proposition and meet the needs of the customers/users and eventually generate substantial revenue. The 6-Values Open Data Business Model Framework ensures that managers are taking into an account all aspects of an effective and efficient business model and understand the effect of the different aspects on each other.

Why is this a Best Practice?

On the one hand, many open data-driven organisations, specifically at upstream level, have successfully designed, developed and implemented their business model but they need to understand the demand side (from mid and downstream organisations) in order to serve them better. On the other hand, midstream and downstream organisations are facing difficulties in developing a business model that allows them to better understand that demand. This best practice helps all types of organisations to overcome the challenges of developing a business model that is effective in identifying what is going on in the open data industry and what more needs to be done to support and feed both demand and supply side. Moreover, developing and implementing an effective and efficient business model can lead to customers/users' satisfaction, emerging innovative products and services, revenue generation and survivability of the organisation and eventually can lead to maximising the economic value of open data.

How do I implement this Best Practice?

In order to be able to start designing and developing this best practice, an organisation needs the following:

- Team up and have required expertise (preferably people with both open data and business development knowledge).
- Patience and courage in searching and sensing the market, existing open data products and services, potential collaborators, and existing competitors.
- Define market niche.
- Open data value discipline/s must be identified before developing the business model.

Open data value disciplines help organisations to focus on delivering superior customer value. Products or services must meet one or multiple value disciplines. **There are four Open Data Value Disciplines:**

1. **Usefulness** - tailors value proposition to directly support the needs of consumers in one way or another.
2. **Process Improvement** - tailors value proposition to match to the needs of the customer specifically for improving processes.
3. **Performance** - tailors value proposition for a better performance.
4. **Customer Loyalty** - tailors value proposition to target customer loyalty.

Further reading

Emerging Business Models for the Open Data Industry: Characterization and Analysis (PDF); Fatemeh Ahmadi Zeleti, Adegboyega Ojo, Edward Curry; 2014, INSIGHT Centre for Data Analytics

Where has this best practice been implemented?

Country	Initiative	Contact Point
Ireland	The Marine Institute	
United States	Open Data Impact Map	

References

- Lisbon Workshop Session: Open Data Business Model Generation (PDF) Fatemeh Ahmadi Zeleti, Insight Centre for Data Analytics
- Lisbon Workshop Paper: Realising an Open Data Marketplace in Greece (PDF) Charalampos Alexopoulos, Yannis Charalabidis, University of the Aegean
- Krems Workshop Talk: Business models for Linked Open Government Data: what lies beneath? (PDF) Nicolas Hazard; PwC
- Krems Workshop Paper: Linked Data Business Cube – Modelling Semantic Web business models (PDF) Tassilo Pellegrini; FH St. Pölten, Christian Dirschl & Katja Eck; Wolters Kluwer
- Fatemeh Ahmadi Zeleti, Adegboyega Ojo, Edward Curry (2014): Emerging Business Models for the Open Data Industry: Characterization and Analysis (PDF)
- Open Data Institute guidance: How to make a business case for open data
- Open Data Institute research: Open data means business
- Open Data Institute white paper: Open enterprise: how three big businesses create value with open innovation

Localised Guidance

- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Finland) Avoimen Datan Opas Open Data Guide
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (Ireland) Guide for publishers
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (UK) Open Data Resource Pack

Contact Info

Fatemeh Ahmadi-Zeleti, Insight Centre for Data Analytics, NUI Galway, Ireland

Related Best Practices

- Establish an Open Data Ecosystem
- Support Open Data Start-Ups

5.14 Open up public transport data

Outline

One of the EU's top priorities is promoting **smart, green and integrated transport** for the benefit of all citizens, the economy and society. In fact of that, under the Horizon 2020 research and innovation scheme, the European Commission allocated €756 million for transport-related activities within the 2016-2017 Work Programme. This work programme includes activities aiming at increasing the take up and scale-up of innovative solutions to achieve sustainable mobility in urban areas, increasing the attractiveness of public transport, creating new coordination and service concepts. One of the expected outputs is the elaboration of new business models for public transport through technological (such as IT and app-oriented services) and social innovations, taking into account possible social and demographic barriers.

In this sense, public transport information – i.e. timetables, service disruptions, stops, accessibility, etc. – is considered as high-value data, essential piece to enable this innovation, guaranteeing services more efficient, reliable and attractive both for operators and customers. Usually this information is shown at stops (on noticeboards or electronic displays) so the information is considered non-sensitive and public, so the evolution to open the information through the Web should be only a technical issue, no legal or strategic constrains must be applied. Thus, the high value of this dataset for the whole society in contrast with the minimum effort to open the already existing information makes public transport data a top priority for open data and PSI re-use initiatives.

Although transport services may be run by private companies, governments should guarantee the openness of this information. This can be done through advocacy or legal actions like issuing specific policies or adding clauses in public procurement.

All stakeholders will benefit: a better user experience for users; greener cities by using collective transport; more efficient company without noteworthy additional costs.

.

Links to the Revised PSI Directive

Policies and Legislation

Challenge

Although most of this information is non-sensitive, and considered as of high potential for reuse, is not always publicly exposed. Most of the transport companies already manage the operating information (timetables, status of the service, etc.) by electronic means. So, the cost of releasing the information openly should not be too high. Although the potential benefit is for all the society (including the government and the transport company), many of these companies are reluctant to open the data up —even having direct requests from the local/regional/national governments.
Cities tend to foster the use of smarter and greener public transport in order to reduce traffic, the subsequent air pollution, making it efficient and easy to use. More information will enable better user applications and services that will enhance citizens' experiences (e.g., journey plans, real-time waiting times, disruption alerts, etc.). This is not always possible due to the lack of open data, data already held by the companies running the service.

Solution

Local governments should make an effort opening the public transport information in machine-readable and easy to use formats –mainly addressed to companies and developers that will be able to create new services or products on top of it. If the government does not have either full access or control of the information, it should get it published, convincing the transport company by all possible means (e.g., issuing a local mandate, adding specific open data clauses in public contracts, etc.).

This would benefit many stakeholders:

- municipalities and governments following 'Smart City' strategies are interested in reducing traffic in the city, enhancing public transport and encouraging its use;
- current public transport users would experience a better service; and newcomers will be attracted only if they see a real value in the service;
- private companies that would be able to produce new products and services for travellers.

Why is this a Best Practice?

Public transport information already exists within transport companies. The service information usually has no legal restrictions, so its publication should not be rejected based on these kind of issues. Also, the release cost should not be too high due to the existing management information systems of the transport companies. Encouraging this openness, all the parties will experience benefits.

How do I implement this Best Practice?

The first step is having government's **political commitment**. This engagement must be included into official digital agendas, stating clear roadmaps and plans for implementation. Specific clauses that guarantee the public access and publication of the transport data must be included explicitly in the agreements between public administrations and public transport companies.

In order to motivate openness and re-use of the information, the municipality must plan and perform actions to raise the awareness among stakeholders (publishers, private companies and citizens).

Where has this best practice been implemented?

Country	Initiative	Contact Point
Spain	Local open data initiatives: • Barcelona • Bilbao • Gijón • Cáceres • Madrid • Gipuzkoa • Granollers • La Palma (Island) • Las Palmas de Gran	Martin Alvarez-Espinar, CTIC

	Canaria • Málaga • Sabadell • Santander • Terrasa • Tenerife • Valencia • Zaragoza	

References

- Samos Workshop Talk: Open Traffic Information Standard & Experimentation for Enhanced Services (PDF)
- Samos Workshop Talk: Public Transport Data in the City of Gijon (PDF)
- Krems Workshop Session: OpenMove: How Trentino opened public transportation data and benefitted of a mobile ticketing solution for free (PDF)

Localised Guidance

- (Austria) Framework for Open Government Data Platforms
- (Belgium) Open Data Handleiding Open Data Handbook
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Finland) Helsinki Region Infoshare
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Portugal) Guia Dados Abertos - AMA | Dados.gov
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (Spain) Government Data Openness and Re-use
- (UK) Open Data Resource Pack
- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Martin Alvarez-Espinar, CTIC

Related Best Practices

- Develop an Open Data Publication Plan
- Establish an Open Data Ecosystem
- Establish Open Government Portal for data sharing
- Use machine-readable standardized data formats
- Provide real-time access
- Identify what you already publish
- Standards for Geospatial Data
- Provide metadata
- Dataset Criteria

5.15 Open up research data

Outline

Open science is a huge and complex area wherein research performance has the public sector as one of its important consumer. In many cases the public sector cannot get up-to-date and precise information about ongoing research due to commercial interests or bad routine. The benefits of a national research evaluation platform makes scientific achievements discoverable and measurable, and thus can affect innovation, economics and education in the country.

Links to the Revised PSI Directive

Scientific research data is not within the scope of the revised PSI Directive, however, it is a closely related topic that raises many of the same issues and so is included here.

Challenge

Open science is a huge and complex area gaining more and more momentum. It has some overlaps with PSI as scientific results and statistics are consumed by the public sector for strategic decisions, evaluations, education and research management. In many cases the public sector cannot get up-to-date and precise information about ongoing research due to commercial interests or bad routine. Currently, the collection of research results is fragmented, some domains have international networks, while other domains may remain without any national aggregator (especially in the area of humanities). This lack of information may lead to a situation when there is no correct view on either micro or macro level on the national research outcomes.

Solution

With the “Amsterdam Call for Action on Open Science” the issues are revealed and significant effort will be put into finding and evaluating new solutions. Currently, many countries establish open access mandates to ensure the visibility of research activities on national level. This enables and encourages researchers to disclose more information about their work. As a next step, national research monitoring platforms can be built to collect and organize data about research. This can be done on several levels:

- Sharing registry data about research results and publications,
- Sharing full texts of publications
- Providing a platform to easily archive and share research experiments.

Why is this a Best Practice?

Collecting information about research requires large-scale cooperation and infrastructure. As a key consumer of collected information, the public sector is suitable to make this effort. The benefits are also affecting the whole country as innovation and research are important for the economic growth. As a result, the scientific achievement becomes discoverable and measurable, which helps researchers in cooperation and improvement as well as funders to

get statistics and overview. Validated scientometric data can also be used to assess individuals in case of assigning grants and degrees.

How do I implement this Best Practice?

A public sector body or a dedicated organization should take the role for guiding the process. A legal background is needed with several elements: declaration of mission, open access mandate, help to resolve copyright issues with publishers, and regulations for data provision. An IT platform has to be established with central and distributed components. The IT requirements may vary in a large range depending on the tasks undertaken. The persistent archival of certain research experiments requires huge storage space and computing power. Furthermore, researchers need to be educated to understand why and how they need to open up their research.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Sweden	Swepub	Peter Krantz
Hungary	MTMT	András Micsik, SZTAKI
Finland	JUULI, Etsin	
The Netherlands	Narcis	

References

- Amsterdam Call for Action on Open Science
- Samos Workshop: MTMT: The Hungarian Scientific Bibliography.
- Timişoara Workshop: Role of Open Data in Research Institutions with International Significance (notes)
- Timişoara Workshop bar camp session: Making research data repositories discoverable
- Krems Workshop talk: re3data.org - Making research data visible and discoverable
- Krems Workshop bar camp session: Open Science & Technology

Localised Guidance

- (Austria) Framework for Open Government Data Platforms
- (Belgium) Open Data Handleiding Open Data Handbook
- (Estonia) Eesti avaliku teabe masinloetava avalikustamise roheline raamat Green Paper on machine-readable Estonian Public Information disclosure
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information

- (Spain) Government Data Openness and Re-use
- (Spain) Guía para el desarrollo de la Universidad Abierta Open University Development Guide

Contact Info

András Micsik, MTA SZTAKI

Related Best Practices

- Establish an Open Data Ecosystem
- Develop and Implement a Cross Agency Strategy

5.16 Provide PSI at zero charge

Outline

Publishing PSI for reuse at no charge unlocks maximum commercial and non-commercial potential.

Links to the Revised PSI Directive

Charging

Challenge

It has been noted that the Public Sector has difficulties determining the commercial and non-commercial potential of PSI. It is also conceivable that by charging for PSI, the potential for reuse will be reduced.

Solution

As much as possible, Public Sector organisations should publish PSI at zero charge to lower the barriers for reuse.

Why is this a Best Practice?

For commercial re-use, zero charge for PSI means that developing products and services will have lower cost, increasing their viability; for non-commercial re-use, zero charges facilitate re-use that otherwise would not be possible.

How do I implement this Best Practice?

Develop and implement a policy that encourages Public Sector organisations to lower marginal cost and to provide access at zero charges. Where necessary, additional funding for agencies that provide PSI at zero charges may be necessary.

Where has this best practice been implemented?

This best practice has been implemented in an overwhelming number of data portals. The following list gives a number of examples but there are many more.

Country	Initiative	Contact Point
EU	European Data Portal	
EU	European Union Open Data Portal	
Slovenia	Slovenia Meteo office	
Austria	Data Portal Austria	
Germany	Data Portal Germany	
The Netherlands	Data Portal Netherlands	

Spain	Open Data Portal Spain	
Spain	Open Data CNMC	
Spain	Open Data Euskadi (Basque Country)	
Spain	Municipality of Madrid Open Data	
Spain	Open Data Gipuzkoa	
Spain	Municipality of Las Palmas de Gran Canaria	
Belgium	Open Data Flanders	
Finland	Open Data Portal of Finland	
Finland	Helsinki Region Infoshare	
Finland	The Finnish Meteorological Institute	
Finland	National Land Survey of Finland	

References

- Samos Workshop paper: A Federation Tool for Open Data Portals (PDF) Mª Dolores Hernández Maroto
- Samos Workshop: Many presentations and discussions at the Samos Share-PSI workshop mentioned how difficult it is for public sector bodies to assess the commercial potential of the information that they publish. Ease of publication is an identified requirement of the Open Group's Open Public Sector Data Business Scenario.
- Lisbon Workshop paper: A Federation Tool for Open Data Portals (PDF) Mª Dolores Hernández Maroto
- Timişoara Workshop paper: OpenMove: How Trentino opened public transportation data and benefitted of a mobile ticketing solution for free (PDF), Lorenzo Modena, CEO OpenMo

Localised Guidance

- (Austria) Framework for Open Government Data Platforms
- (Belgium) Open Data Handleiding Open Data Handbook
- (Croatia) Preporuke o prilagodbi skupova podataka za javnu objavu i ponovno korištenje Open Data Guide, Croatia
- (Finland) Helsinki Region Infoshare
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information

- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Latvia) Atvērto datu vadlīnijas Open Data Guidelines
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Norway) Veileder i tilgjengeliggjøring av offentlige data Guide to making public data available
- (Portugal) Guia Dados Abertos - AMA | Dados.gov
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Sweden) Vidareutnyttjande av information Om PSI och öppna data Reuse of PSI and open data
- (UK) Open Data Resource Pack

Contact Info

Makx Dekkers, AMI Consult.

Related Best Practices

- Open Data Business Models & Value Disciplines
- Establish an Open Data Ecosystem
- Holistic Metrics

5.17 Publish overview of managed data

Outline

When the user community is asked about what datasets they would prefer for release they often cannot respond because they are not aware of what datasets are available. Therefore, an overview of datasets managed by an organisation should be published.

Links to the Revised PSI Directive

Selection

Challenge

How to make sure that community is aware of datasets that could be possibly made available as open data?

There is often a 'catch 22' situation when identifying data for release. The public sector asks the user community what data they would like and they will prioritise this for release. However, the user community are often not aware of what exists and therefore cannot respond meaningfully.

Solution

Publish an overview of datasets managed by an organisation.

Why is this a Best Practice?

The user community is not always fully aware of datasets that an organisation can possibly make available as open data, at least not to a full extent. This might prevent the community from providing useful feedback about what datasets it would have preferred for release. Following this practice should improve efficiency of gathering feedback about datasets requested for release.

The overview when made public gives enough information for users (both public and private sector) to prioritise the most interesting data for release.

Transparency on what datasets public sector is managing combined with a simple classification of legal right to access the data is adding value to both public sector and reusers because:

- The knowledge that this data exists has value in itself
- It makes it possible to crowd-source your priorities
- You are providing transparency on what information your agency is managing.

How do I implement this Best Practice?

The technical requirements are minimal. A simple published spreadsheet and a route through which feedback can be sent should be sufficient. The overview can also be provided within an existing open data catalogue.

The impact of the practice can be improved when the best practice Categorise openness of data is being followed. This can be solved within an existing open data catalogue by using the "access rights" property in DCAT-AP.

This allows potential external users to examine the data categorised as yellow (datasets that could be possibly shared within the public sector but not with the external users) and bring forward arguments about why some of it could be re-categorised as 'green' data if they see fit.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Finland	Helsinki Region Infoshare: Helsinki Opens A Window to its Information Systems (press release)	Pekka Koponen, Forum Virium Helsinki.

References

- Samos Workshop Talk: Traffic Light System For Data Sharing
- Samos Workshop Talk: Public Transport Data in the City of Gijon (PDF)

Localised Guidance

- (Austria) Open-Government-Vorgehensmodell Open Government Process Model
- (Belgium) Open Data Handleiding Open Data Handbook
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Estonia) Eesti avaliku teabe masinloetava avalikustamise roheline raamat Green Paper on machine-readable Estonian Public Information disclosure
- (Finland) Helsinki Region Infoshare
- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Norway) Veileder i tilgjengeliggjøring av offentlige data Guide to making public data available
- (Portugal) Guia Dados Abertos - AMA | Dados.gov
- (Romania) Ghid Pentru Publicarea Datelor Deschise Romanian Open Data Guide
- (Serbia) Open Data Handbook
- (Sweden) Vidareutnyttjande av information Om PSI och öppna data Reuse of PSI and open data

Contact Info

Heather Broomfield, Difi (Norway).

Related Best Practices

- Categorise openness of data
- Dataset Criteria
- Identify what you already publish
- Provide metadata
- Establish an Open Data Ecosystem

5.18 Publish statistical data in Linked Data format

Outline

Publishing statistical data as Linked Data on the basis of W3C's RDF Data Cube vocabulary which specifies an approach for the expression of the data in a standardised machine-readable way as well as identifying a recommended set of metadata terms to describe the datasets.

Links to the Revised PSI Directive

Techniques

Challenge

Statistical data is currently published in a range of formats and standards that do not allow linking across datasets. It is used as the foundations for policy prediction, planning and adjustments, and therefore has a significant impact on the society (from citizens to businesses to governments). The process of collecting and monitoring socio-economic indicators can be considerably improved if the data produced by government organizations such as Statistical Offices, National Banks, Employment services, etc. are published in Linked Data format.

Solution

Linked Data paradigm has opened new possibilities and perspectives for government organizations to open data and interchange information. Data is open if it is technically open (available in a machine-readable standard format, which means it can be retrieved and meaningfully processed by a computer application) and legally open (explicitly licensed in a way that permits commercial and non-commercial use and reuse without restrictions), see the World Bank Open Data Essentials.

The Linked Data approach enables datasets to be linked together through references to common concepts. A dataset is represented in the form of a graph, using the Resource Description Framework (RDF) as a general-purpose language. Linked Data publication process refers to a set of activities related to extraction, transformation, validation, exploration and publication of RDF datasets originating from different sources (e.g., databases) on the Web. The ready for use RDF datasets can be either stored locally or registered at a metadata catalogue, e.g. build with CKAN open-source tool.

In 2014, the RDF Data Cube Vocabulary was published by the W3C Government Linked Data Working Group as a recommendation for publishing multi-dimensional data on the Web.

Why is this a Best Practice?

The approach contributes to the standardization of the process of publishing and reuse of multi-dimensional data on the Web. The approach is based on RDF Data Cube vocabulary that is mature enough to be used for publishing statistical data as it improves interoperability and allows comparison of data from different statistical sources. The vocabulary underlies SDMX (Statistical Data and Metadata eXchange), an ISO standard for exchanging and sharing statistical data and metadata among organizations and provides a layer on top of data to describe domain semantics, dataset's metadata, and other crucial information needed in the process of statistical data exchange. The RDF Data Cube Vocabulary has shown to be applicable in different use cases (Use Cases and Lessons for the Data Cube Vocabulary)

and to be a good trade-off between simplicity of use and formalization (Choose the right formalization level).

Cost implication: Costs of publication should be minimised unless there are clear benefits. Public sector body should analyse the current status of data availability, the demand for data and thus avoid unnecessary costs of transformation of data in Linked Data format. Public sector bodies publishing information SHOULD either:

- Publish it in the manner that involves lowest cost, consistent with making it available effectively and openly, or
- Carry out cost-benefit analyses of the possible measures to assess potential use and stimulate take-up, methods of publication, and formats for publication, and select measures, methods and formats in the light of those analyses.

The risk of deciding what publication form will best deliver value (commercial or other value of public information), and the work of converting it to that form, could be left to commercial product and service providers, and other consumers. If due to cost implications it is not possible to publish statistical data in that format, it is important to ensure possible transformations by third parties from the provided format to the RDF Data Cube Vocabulary. The multidimensional data model (with n-dimensional data cubes as datasets with observations, dimensions, measures) used by the RDF Data Cube Vocabulary is sufficiently generic to not restrict publishers.

A possible transformation has been shown for other common data formats for statistical data such as SDMX, XBRL, and the Dataset Publishing Language. If sufficient metadata is provided, transformation scripts are also possible from CSV and spreadsheet (e.g., Microsoft Excel) data.

How do I implement this Best Practice?

This best practice is based on a set of tools for automating the data extraction and publication process. However, the EU research community delivered many open-source tools for publishing the statistical data in Linked Data format, see for example the LOD2 Statistical Workbench, or the OpenCube toolkit.

Where has this best practice been implemented?

Many EU States (especially the Statistical Offices) already publish their data in Linked Data format e.g. http://statisticsbeta.com/ from Scotland. Most often these services are available on national Web portals, while the metadata is harvested on European level e.g. by the Publicdata.eu. Additionally, the European Commission maintains the Open Data Portal as a metadata catalogue available as Linked Data, see open-data.europa.eu Linked Data.

Country	Initiative	Contact Point
Greece	A Methodology for Publishing Linked Open Statistical Data	George Papastefanatos IMIS / RC Athena, Greece
Italy	LinkedStat	SpazioDati and Istat
UK (Scotland)	Scottish Government Statistics	Scottish Government
Serbia	Publishing and Consuming	Valentina Janev, Institute Mihajlo Pupin

Country	Initiative	Contact Point
	Linked Open Data with the LOD Statistical Workbench	

References

- Samos Workshop presentation: A Methodology for Publishing Linked Open Statistical Data (PDF), George Papastefanatos IMIS / RC Athena, Greece
- Samos Workshop presentation: Publishing and Consuming Linked Open Data with the LOD Statistical Workbench, Valentina Janev, Institute Mihajlo Pupin

Localised Guidance

- (Belgium) Open Data Handleiding Open Data Handbook
- (Estonia) Avaandmete loomise ja avaldamise juhend Open Data Guidelines
- (Finland) Helsinki Region Infoshare
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Serbia) Open Data Handbook
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (UK) Open Data Resource Pack

Contact Info

Valentina Janev, Institute Mihajlo Pupin;

Contributor: Benedikt Kämpgen, FZI Research Center for Information Technology

Related Best Practices

- Reuse vocabularies
- Use machine-readable standardized data formats

5.19 (Re)use federated tools

Outline

Different countries have developed federated/distributed tools for open data collection that enable automatic publication of the (meta)data corresponding to the data sets published on the websites of each public entity. A global index of reusable public information can be thus created and accessed by users to locate reusable data. Aggregation enables it without the need to know and find the website of the public entity holding the data in which reusers are interested in.

Links to the Revised PSI Directive

Platforms, Techniques, Organisation, Formats, Reuse, Persistence, Documentation

Challenge

The number of data owners is large and from various levels of government, so a country-wide central system, where everyone is authorized to log in, would be hard to implement. Also, the data owners are quite autonomous of each other, so no unified technical solution could be implemented into their internal processes. In addition, the end-users are from different domains, so there is a need for common understanding.

Solution

To overcome the challenges a federated/distributed solution, that is scalable and can integrate numerous counterparts, is to be implemented. Counterparts are integrated via common output data format that poses little requirements on internal processes. Also, a common vocabulary (in the format of data schema) is to be put in practice.

Why is this a Best Practice?

- An aggregated view of the collections from many agencies can be offered to the users, allowing them to explore those collections at a single point of contact.
- A lot more public servants can take part in (meta)data creation and verification.
- The creation of (meta)data is close to the source of the data itself.
- The public agencies hold (and govern themselves) the data in their own environment.
- No need for centrally managed individual access management or lengthy (meta)data input forms.

How do I implement this Best Practice?

To implement this best practice, you will need some elements, among them:

- Use (or set up) a data portal as a single point where to federate all published datasets of the different public entities. Agree on the organisation of persistent URIs, where from the origin of the data can be accessed. URIs assigned should persistently identify the same thing over time and the thing identified should be also persistently available.
- Select (or agree on) an intermediate data structure/format. As more and more systems are interconnected with each other, standardised solutions are preferable

(e.g., W3C recommendation in a DCAT/RDF or ATOM format feed). Nevertheless, in some exceptional domain/local cases it could be as simple as CSV with agreed column names/descriptions.

- Select an existing (or develop a new) set of tools to push the data to a publication portal (e.g., a harvester and a national portal, European Open Data Portal). It is best that these work well with the process and tools the public servants use for their daily life (so get to know the work of public servants). Configure the tools to collect the distributed data to be published to the portal.
- Appoint an agency responsible for maintaining the data structure/format and support the use of tools. If needed, a more sophisticated coordination structure between the different administrative levels (state, regional, local) should be established.
- Complement (or establish) a legal and technical framework ensuring that each public entity will federate their datasets at the national data portal in a standard manner.
- Document and train, how the public servant can create the data themselves.
- Monitor and support the use of selected tools (by the appointed agency).

Where has this best practice been implemented?

This best practice has the following implementation examples:

- An extension of Spanish National Catalogue datos.gob.es enables aggregation and automatic publication of the metadata corresponding to the data sets published in the own catalogues on the websites of each public entity and also at the in a consistent way. A global index of reusable public information is thus created and can be accessed. The tool ensures maximum coherence between the information being made available by the public entities in their own catalogues and the National Catalogue itself. This solution enables the existence of a global reuse scenario that provides greater visibility for the public data made available by the three levels of government (central, regional, local and universities), as well as a general overview of how public sector information is being reused in Spain.
- Since the first publication of the DCAT-AP in 2013, many member states have implemented national application profiles based on the European profile. A revision of the DCAT-AP was developed based on contributions from various Member States, the European Commission and independent experts. It has been implemented on national/regional level, with code lists recommended by the DCAT-AP. The reuse of a common structure has enabled to aggregate regional level and domain specific data catalogues to national level, and now on European level.
- Estonia's wealth of services has clearly indicated a need for a more structured and methodical approach to national-level service portfolio management. Data suppliers provide information on the public services they provide in the data format agreed either using the data extractions tools developed or manual data entry. That information is collected and stored centrally in a searchable format.
- To increase interoperability in the exchange of data between public agencies, Germany has developed a set of different free to use tools (under the name of XÖV meaning "XML for public administrations"). These tools aim to support the standardisation of data structures and codes lists. The tools can be used to create and manage code lists (Genericoder), to browse in public agency data standards (InteropBrowser) or data structures (XRepository).

Country	Initiative	Contact Point
Spain	The Spanish National Catalogue	soporte@datos.gob.es
European Union	European Data Portal	help@europeandataportal.eu
Estonia	Estonian Public Service Catalogue	riigiteenused@mkm.ee
Germany	Coordination Office for IT standards	joerg.hofmann@finanzen.bremen.de

References

- Samos workshop: A Federation Tool for Open Data Portals, Mª Dolores Hernández Maroto
- Berlin Workshop: Implementing The DCAT Application Profile For Data Portals In Europe, Nikolaos Loutas, Makx Dekkers
- Berlin Workshop: Estonian Metadata Reference Architecture, Hannes Kiivet
- Berlin Workshop: German XML for public administration "XÖV" tool chain in action, Sebastian Sklarß]init[AG, Lutz Rabe

Localised Guidance

- (Belgium) Open Data Handleiding Open Data Handbook
- (CzechRepublic) Standardy publikace a katalogizace otevřených dat veřejné správy ČR Open Data Standards
- (Estonia) Avaandmete loomise ja avaldamise juhend Open Data Guidelines
- (Germany) Open Government Data Deutschland
- (Hungary) Nyílt Adatok kézikönyv Open Data Handbook
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (International) DCAT application profile implementation guidelines
- (Ireland) Guide for publishers
- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Latvia) Atvērto datu vadlīnijas Open Data Guidelines
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Norway) Veileder i tilgjengeliggjøring av offentlige data Guide to making public data available
- (Serbia) Open Data Handbook
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use

Contact Info

soporte@datos.gob.es or via Dirección de Tecnologías de la Información y las Comunicaciones.

Related Best Practices

- Develop and Implement a Cross Agency Strategy
- Develop an Open Data Publication Plan
- Establish Open Government Portal for data sharing

5.20 Standards for Geospatial Data

Outline

Public administration bodies need to work together with architecture, engineering, and construction firms as well as building owners, brokers, component vendors, operators, insurers, inspectors, tenants, finance companies, fire departments, health and social services and more. For almost all PSI, location is critical. Therefore, it is essential that location/geospatial data is shared in a way most likely to be re-usable by partner organisations - and that means adhering to standards. Most standards relevant to geospatial data are developed by the Open Geospatial Consortium (OGC).

Links to the Revised PSI Directive

Platforms, Techniques, Formats, Reuse, Discoverability

Challenge

The overall challenge is to ensure that information from different sources that relate to the same location can be used together. This is particularly challenging because 'location' can be defined by many different means: name (in different languages and/or with different abbreviations), coordinates, boundaries, administrative district names, NUTS code, floor plans, points, centroids, polygons, rasters, subway maps, bus stops, time series, left/right directions etc. The goal is to make all this data available as open data, following open standards and open data models. Precision, uncertainty, provenance, rights and access control are often factors. Given all of this complexity, developing software that involves location data can be difficult. The requirement to make location data and location services sharable, re-usable and interoperable makes the task even more challenging.

Another important challenge is that Web technologies are evolving, so OGC standards and location information architectures also need to evolve.

Solution

By applying standards, particularly those developed by the OGC, public sector geo-/ and location information can be provided in an efficient and interoperable way to many other data sets and processing or visualisation components. OGC and ISO standards such as Web Feature Service WFS, WMS, GML, IndoorGML, CityGML and SOS ensure standardised access to all public sector information with spatial characteristics. Some (IndoorGML, in particular) are new, but most of these standards have been used for years by public sector organizations around the world. They are essential components for PSI architectures.

The Spatial Data on the Web Working Group is a joint undertaking by both OGC and W3C to make spatial data interoperate readily with more general data available on the Web. It operates in collaboration with a parallel group in W3C of the same name, with overlapping membership. Further information on the membership arrangements for the groups can be found in the charter. On 19 January 2016 the W3C and OGC Spatial Data on the Web Working Group published the First Public Working Draft of its Best Practices document for Spatial Data on the Web. This is a concerted attempt to bring together techniques used by the geospatial industry and Web technologists, especially those making use of Linked Data techniques. Typical use cases include environmental and cartographic data, transport and administrative data. Although clearly a lot remains to be done, the editors seek to illustrate the full scope of the best practices.

Why is this a Best Practice?

By using OGC standards to publish public sector information with spatial characteristics, it becomes much easier to integrate this information with other data sets that are served at similar interfaces. Data becomes discoverable using standardised catalogues and can be used as part of initiatives such as INSPIRE, the European Spatial Data infrastructure. This best practice (and the many best practices that have been developed and adopted by the OGC Technical Committee) are best practices that describe the best ways to implement OGC standards.

Most of valuable public sector information has spatial components to it. In order to make maximal use of this data, it should be made available through standardised interfaces following standardised formats. Using open standards from OGC, W3C and others ensures a very high level of interoperability, paving the way to new businesses and further commercialisation.

How do I implement this Best Practice?

This best practice is based on OGC technologies and also W3C technologies, so the way to begin, is by learning about those technologies.

Where has this best practice been implemented?

OGC standards are widely implemented throughout the world and form the basis of many critical industries and government activities.

Country	Initiative	Contact Point
Finland	Helsinki Region Infoshare	Pekka Koponen, Forum Virium Helsinki
Czech Republic	Czech implementation of the INSPIRE Directive	

References

- Timişoara Workshop Session: Free Our Maps (PDF)
- Berlin Workshop: Location Track

Localised Guidance

- (Belgium) Open Data Handleiding Open Data Handbook
- (Finland) Helsinki Region Infoshare
- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (International) Open Data Handbook, Solutions Bank
- (International) Using Open Public Sector Information
- (Ireland) Guide for publishers

- (Italy) Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico National Development Guidelines for Public Sector Information
- (Lithuania) Viešojo Sektoriaus Informacijos platinimo gerosios praktikos The Best Practices for the PSI publication
- (Portugal) Guia Dados Abertos - AMA | Dados.gov
- (Serbia) Open Data Handbook
- (Slovenia) Priročnik za odpiranje podatkov javnega sektorja Manual for the opening of public sector information
- (Spain) Guía de aplicación de la Norma Técnica de Interoperabilidad de reutilización de recursos de información Application Guide for Technical Interoperability Standard on PSI re-use
- (UK) Open Data Resource Pack
- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Dr. Ingo Simonis, Director Interoperability Programs & Science, Open Geospatial Consortium (OGC)

Related Best Practices

- Use machine-readable standardized data formats
- Use a trusted serialisation format for preserved data dumps
- Provide data in multiple formats
- Reuse vocabularies

5.21 Support Open Data start ups

Outline

An academic business accelerator is an organisational unit within a university that aims to mobilise and support people to build their own enterprise. The unit is responsible for transferring the innovation and entrepreneurship mentality to the next generation of skilled entrepreneurs. It also maintains collaboration with private and public structures for funding and mentoring. Open data can provide a very useful basis for entrepreneurship, allowing for development of added value services by citizens and small enterprises. The open data sub-unit enhances the collaboration between universities (potential entrepreneurs) with private and public funding organisations (chambers of commerce, municipalities, start-up investors) and experts (coaches and mentors) from the private sector in order to foster innovative open data start-ups to go live.

Links to the Revised PSI Directive

Policies and Legislation

Challenge

Getting a PSI / open data inspired based business up and running requires a multitude of fields of knowledge and mentoring.

Solution

Universities are well inter-connected, they have the required knowledge necessary for open data start-ups like (statistics, visualisation, programming), and they usually have the necessary relationships to practitioners, and other academics to complete missing skills and knowledge.

Furthermore, they can provide:

- Resources: They often own facilities which are not used up to their full capacity.
- Mentoring: University nodes can act as a trusted third party intermediary, who can establish the first contacts, or bring the first 5 customers.
- Networking to start-ups: Professors, Professionals.
- Potential Entrepreneurs like students, Alumni, SMEs, Individuals.

Why is this a Best Practice?

It contributes to sustainable growth and entrepreneurship based on Open Data, in Europe.

How do I implement this Best Practice?

A new structure in a University or other similar educational institute is needed. Detailed steps towards the academic start-up incubator:

- make open data via web services available to students: provide some initial tools, training, technology groups;
- expand on successful startup examples, generalize and apply to other areas of business;

- make long lasting competitions, as one month is not enough in order to distill ideas;
- connect with the outer world: journalists;
- provide legal advice for the establishment, IPR and privacy;
- mentors should also come from successful startups;
- The loudest bird survives: Blog on every activity, involve students into that process;
- team up with another academic school (unite media and technics);
- provide intercultural communication education, as startups are likely to provide services on an international scale;
- first identify friendly customers, then make them enthusiasts.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Greece	The University of the Aegean Startup Incubator	Yannis Charalabidis, University of the Aegean
Greece	Gov4All platform: University of the Aegean/Microsoft Greece open data incubator	Yannis Charalabidis, University of the Aegean, Greece
Australia	DataStart: A public-private partnership	Tim Neal, Data Policy, Department of the Prime Minister and Cabinet, Australia
Spain	Gijón OpenDataLab: local start-up incubator	Martin Alvarez-Espinar, CTIC

References

- Lisbon Workshop Session: Open Data Startups: Catalyzing open data demand for commercial usage
- Krems Workshop Session: University Business Accelerators on Open Data: Activities, Challenges and Best Practices
- Berlin Workshop Talk: An Intelligent Fire Risk Monitor Based on Linked Open Data (PDF)
- Presentation slides
- Open Data Incubator Europe

Localised Guidance

- (Belgium) Open Data Handleiding Open Data Handbook
- (Finland) Avoimen Datan Opas Open Data Guide
- (Germany) Open Government Data Deutschland
- (International) Using Open Public Sector Information
- (Latvia) Atvērto datu vadlīnijas Open Data Guidelines
- (Lithuania) Viešojo_Sektoriaus_Informacijos_platinimo_gerosios_praktikos The Best Practices for the PSI publication

- (Serbia) Open Data Handbook
- (Spain) Government Data Openness and Re-use
- (Spain) Guía metodológica para planes open data sectoriales Methodological Guide for Sectorial Open Data Plans
- (UK) Open Data Resource Pack
- (UK) Birmingham and West Midlands Localised Guide for Open Data

Contact Info

Yannis Charalabidis, University of the Aegean

Related Best Practices

- Establish an Open Data Ecosystem
- Establish Open Government Portal for data sharing
- Use machine-readable standardized data formats
- Provide real-time access

6 W3C Best Practices

The Best Practices developed by the Share-PSI 2.0 Thematic Network and the Best Practices created by the W3C Data on the Web Best Practices Working Group complement each other. This group focuses on the technical aspects of sharing data on the Web, meaning that the kind of policy issues addressed in previous sections of this document are out of scope. From the point of view of implementing the (Revised) PSI Directive, the 'multiple sets' of Best Practices should be seen as one. Table 3 on page 29 indicates how the W3C Data on the Web Best Practices relate to the different elements of the PSI Directive. The W3C Data on the Web Best Practices have been grouped in thirteen categories, each devoted a section in this chapter.

6.1 Metadata

Data will not be discoverable or reusable by anyone other than the publisher if insufficient metadata is provided. Metadata provides additional information that helps data consumers better understand the meaning of data, its structure, and to clarify other issues, such as rights and license terms, the organization that generated the data, data quality, data access methods and the update schedule of datasets.

- Best Practice 1: Provide metadata: Metadata must be provided for both human users and computer applications
- Best Practice 2: Provide descriptive metadata: The overall features of datasets and distributions must be described by metadata
- Best Practice 3: Provide locale parameters metadata: Information about locale parameters (date, time, and number formats, language) should be described by metadata.
- Best Practice 4: Provide structural metadata: Information about the schema and internal structure of a distribution must be described by metadata.

6.2 Data licenses

A license is a very useful piece of information to be attached to data on the Web. According to the type of license adopted by the publisher, there might be more or fewer restrictions on sharing and reusing data. In the context of data on the Web, the license of a dataset can be specified within the metadata, or outside of it, in a separate document to which it is linked.

- Best Practice 5: Provide data license information: Data license information should be available.

6.3 Data provenance

The challenge in publishing data on the Web is providing an appropriate level of detail about its origin. The data producer may not necessarily be the data provider and so collecting and conveying this corresponding metadata is particularly important. Without provenance, consumers have no inherent way to trust the integrity and credibility of the data being shared.

Data publishers in turn need to be aware of the needs of prospective consumer communities to know how much provenance detail is appropriate.

- Best Practice 6: Provide data provenance information: Data provenance information should be available.

6.4 Data quality

The quality of a dataset can have a big impact on the quality of applications that use it. As a consequence, the inclusion of data quality information in data publishing and consumption pipelines is of primary importance. Usually, the assessment of quality involves different kinds of quality dimensions, each representing groups of characteristics that are relevant to publishers and consumers. The Data Quality Vocabulary defines concepts such as measures and metrics to assess the quality for each quality dimension [*VOCAB-DQV*]. There are heuristics designed to fit specific assessment situations that rely on quality indicators, namely, pieces of data content, pieces of data meta-information, and human ratings that give indications about the suitability of data for some intended use.

- Best Practice 7: Provide data quality information: Data Quality information should be available.

6.5 Data versioning

Datasets published on the Web may change over time. Some datasets are updated on a scheduled basis, and other datasets are changed as improvements in collecting the data make updates worthwhile. In order to deal with these changes, new versions of a dataset may be created.

- Best Practice 8: Provide version indicator: Information about dataset versioning should be available.
- Best Practice 9: Provide version history: A version history about the dataset should be available.

6.6 Data identifiers

Data discovery, usage and citation on the Web depends fundamentally on the use of HTTP (or HTTPS) URIs: globally unique identifiers that can be looked up by dereferencing them over the Internet [*RFC3986*].

- Best Practice 10: Use persistent URIs as identifiers of datasets: Datasets must be identified by a persistent URI
- Best Practice 11: Use persistent URIs as identifiers within datasets: Datasets should use and reuse other people's URIs as identifiers where possible
- Best Practice 12: Assign URIs to dataset versions and series: URIs should be assigned to individual versions of datasets as well as the overall series.

6.7 Data formats

The format in which data is made available to consumers is a key aspect of making that data usable. The best, most flexible access mechanism in the world is pointless unless it serves data in formats that enable use and reuse. Below we detail Best Practices in selecting formats for your data, both at the level of files and that of individual fields. W3C encourages use of formats that can be used by the widest possible audience and processed most readily by computing systems.

- Best Practice 13: Use machine-readable standardized data formats: Data must be available in a machine-readable standardized data format that is adequate for its intended or potential use.
- Best Practice 14: Provide data in multiple formats: Data should be available in multiple data formats.

6.8 Data vocabularies

Vocabularies define the concepts and relationships (also referred to as “terms” or “attributes”) used to describe and represent an area of interest. They are used to classify the terms that can be used in a particular application, characterize possible relationships, and define possible constraints on using those terms. Several near-synonyms for 'vocabulary' have been coined, for example, ontology, controlled vocabulary, thesaurus, taxonomy, code list, semantic network. There are different means to create ontologies, such the RDF Schema [*RDF-SCHEMA*] language, or the Web Ontology Language [*OWL2-OVERVIEW*]. Simpler models have thus been proposed to represent and exchange them, such as the ISO 25964 data model [*ISO-25964*] or W3C's Simple Knowledge Organization System [*SKOS-PRIMER*].

- Best Practice 15: Reuse vocabularies, preferably standardized ones: Shared vocabularies should be used to provide metadata.
- Best Practice 16: Choose the right formalization level: When reusing a vocabulary, a data publisher should opt for a level of formal semantics that fit data and applications.

6.9 Data access

Providing easy access to data on the Web enables both humans and machines to take advantage of the benefits of sharing data using the Web infrastructure. By default, the Web offers access using Hypertext Transfer Protocol (HTTP) methods. This provides access to data at an atomic transaction level. When data is distributed across multiple files or requires more sophisticated retrieval methods approaches like bulk download and APIs can be adopted.

- Best Practice 17: Provide bulk download: Data should be available for bulk download.
- Best Practice 18: Provide subsets for large dataset: If your dataset is large, enable users and applications to readily work with useful subsets of your data.
- Best Practice 19: Use content negotiation for serving data available in multiple formats: Use content negotiation in addition to file extensions for serving data available in multiple formats.
- Best Practice 20: Provide real-time access: When data is produced in real-time, it should be available on the Web in real-time.

- Best Practice 21: Provide data up to date: Data must be available in an up-to-date manner and the update frequency made explicit.
- Best Practice 22: Provide an explanation for data that is not available: For data that is not available, provide an explanation about how the data can be accessed and who can access it.
- Best Practice 23: Make data available through an API: Offer an API to serve data if you have the resources to do so.
- Best Practice 24: Use Web standards as the foundation of APIs: When designing APIs, use an architectural style that is founded on the technologies of the Web itself.
- Best Practice 25: Provide complete documentation for your API: Provide complete information on the Web about your API. Update documentation as you add features or make changes.
- Best Practice 26: Avoid breaking changes to your API: Avoid changes to your API that break client code, and communicate any changes in your API to your developers when evolution happens.

6.10 Data preservation

For a wide variety of reasons, data publishers are likely to want or need to remove data from the live Web. Simply deleting a resource from the Web is bad practice. In that circumstance, dereferencing the URI would lead to an HTTP Response code of 404 that tells the user nothing other than that the resource was not found. The following Best Practices offer more productive approaches.

- Best Practice 27: Preserve identifiers: When removing data from the Web, preserve the identifier and provide information about the archived resource.
- Best Practice 28: Assess dataset coverage: The coverage of a dataset should be assessed prior to its preservation.

6.11 Feedback

Data publishers want to ensure that the data published is meeting the data consumer needs and for this purpose, user feedback is crucial. Feedback has benefits for both publishers and consumers, helping data publishers to improve the integrity of their published data, as well as encouraging the publication of new data. Feedback allows data consumers to have a voice describing usage experiences (e.g. applications using data), preferences and needs. When possible, feedback should also be publicly available for other data consumers to examine. Making feedback publicly available allows users to become aware of other data consumers, supports a collaborative environment, and allows user community experiences, concerns or questions are currently being addressed.

- Best Practice 29: Gather feedback from data consumers: Provide a readily discoverable means for consumers to offer feedback.
- Best Practice 30: Make feedback available: Make consumer feedback about datasets and distributions publicly available.

6.12 Data enrichment

Data enrichment refers to a set of processes that can be used to enhance, refine or otherwise improve raw or previously processed data. This idea and other similar concepts contribute to making data a valuable asset for almost any modern business or enterprise. It is worth noting that some of these techniques should be approached with caution, as ethical concerns may arise. In scientific research, care must be taken to avoid enrichment that distorts results or statistical outcomes. For data about individuals, privacy issues may arise when combining datasets. That is, enriching one dataset with another, when neither contains sufficient information about any individual to identify them, may yield a combined dataset that compromises privacy.

- Best Practice 31: Enrich data by generating new data: Enrich your data by generating new data from the raw data when doing so will enhance its value.
- Best Practice 32: Provide complementary presentations; Enrich data by presenting it in complementary, immediately informative ways, such as visualizations, tables, Web applications, or summaries.

6.13 Republication

Reusing data is another way of publishing data; it's simply republishing. It can take the form of combining existing data with other datasets, creating Web applications or visualizations, or repackaging the data in a new form, such as a translation. Data republishers have some responsibilities that are unique to that form of publishing on the Web.

- Best Practice 33: Provide feedback to the original publisher: Let the original publisher know when you are reusing their data. If you find an error or have suggestions or compliments, let them know.
- Best Practice 34: Follow licensing terms: Find and follow the licensing requirements from the original publisher of the dataset.
- Best Practice 35: Cite the original publication: Acknowledge the source of your data in metadata. If you provide a user interface, include the citation visibly in the interface.

7 Conclusions

This document is the collaborative effort of all Share-PSI 2.0 partners coming from different organisations (public institutions, research organisations, non-government organisations and standardisation bodies) that have been actively involved in Open Data initiatives in European countries in the last few years. As documented in this deliverable, the Share-PSI 2.0 workshops were occasions where consortium partners presented stories that address particular practical issues from the PSI Directive implementation. Organised at different locations in Europe and co-located with other EU events (e.g. 5th annual Samos Summit on ICT-enabled Governance, or the CeDEM conference), the Share-PSI 2.0 workshops were events that gathered participants with different backgrounds ranging from politicians, public servants, EU representatives, via IT professionals and businessmen to researchers, students and other representatives of society.

Consortium partners from the government sector (e.g. MAREG-Greece, SCOT-United Kingdom, CORVE-Belgium, AMA-Portugal, DIFI-Norway, MNZ-Slovenia) contributed stories based on their responsibility for and the ownership of some or all PSI policies. As a result, we have the Best Practices Establish an Open Data Ecosystem, Develop and Implement a Cross Agency Strategy, Categorise openness of data, Develop an Open Data Publication Plan and others.

On the other hand, research organisations, non-government organisations and standardisation bodies based on their experience and technological background contributed to the technology oriented Best Practices such as Establish Open Government Portal for data sharing, Publish overview of managed data, Develop a federation tool for open data portals, Identify what you already publish, Enable quality assessment of open data and others.

For the stable list of best practices that have been published, the consortium partners expressed agreement that these methods were already applied, or cited in public documents. The consensus was reached in a voting process, where the main criterion was the opinion of consortium members. Additionally, a subset of recommendations have not been accepted by the majority of partners, but being a valuable pieces of knowledge gathered throughout of more than two years by the hard work of all partners, have been retained and included in the Annex of this deliverable.

Having created, updated or simply read many guides and policy documents on open data, public sector information sharing and PSI Directive implementation from across Europe, it is possible to assess which Best Practices are most widely agreed upon. This is covered in detail in deliverable D7.3 “Localised implementations guides for the best practice” and accessible online[23] as well. From D7.3, it can be seen that each of the BPs is covered in *at least* 6 such guides and that some are covered in almost all those looked at. Having an open data publication plan and a portal that supports feedback is clearly seen as an appropriate method of approach, whilst publishing statistics in Linked Data is seen as less of a priority.

It needs to be noted that we designed Share-PSI 2.0 Best Practices on a sufficiently high level to enable their implementation in actual national or local environments. As the starting point and open data maturity are not uniform across countries, the local guides are designed to suit the peculiarities of each country. The Share PSI 2.0 Best Practices will be a source of inspiration for every consortium participant in order to improve the open data situation in their country and to advance in implementation of the PSI Directive. As such, the Best Practices are the tools to be used by the network partners in creating or enhancing their local guides.

[23] https://www.w3.org/2013/share-psi/lg/

Annex. Additional recommendations stemmed from Share-PSI workshops

A. 1. Catalogues and indexes for reference

Outline

Information published on the web should be referenced by catalogues and indexes. (A catalogue is a list of publications. An index is a list of terms appearing in publications, with links to the publications that contain them.)

Links to the Revised PSI Directive

Discoverability, Documentation

Challenge

Data is published by the responsible administrative departments, usually acting independently. Descriptions of what is published are not always produced. Where they are produced, they are not co-ordinated, so that similar data published by different administrations may be described in quite different terms. There generally are no catalogs or indexes showing the totality of what is available.

This makes it hard to establish whether the data required for a particular purpose is available, and to obtain it where it is available.

Solution

Publication on the web of catalogues and indexes of public sector information can make a significant contribution to discoverability, both by search engines and by people browsing the web.

The catalogues and indexes should include metadata for the documents that they reference. This will facilitate discovery of those documents even when they themselves do not include metadata.

Why is this a Best Practice?

Adoption of this best practice will significantly improve discoverability of public sector information, by people accessing public sector websites directly, and by people using search engines.

How do I implement this Best Practice?

1. Define the catalogues and indexes to be used. This will require consultation with information producing and publication departments, and with other stakeholders
2. Define the formats of the catalogues and indexes. These should be web-based, with hyperlinks to the publications that they reference.
3. Add to the information publication process a step that ensures that the published material is catalogued and indexed appropriately.

References

- The Open Public Sector Business Scenario
- Share-PSI 2.0 Samos Workshop: Uses of Open Data Within Government for Innovation and Efficiency:
- Samos Workshop Talk: Open Data to Improve Sharing and Publication of Information between Public Administrations
- Krems Workshop Talk: Data banks - Data as an asset under the control of owner/custodian

Related Best Practices

- Publish overview of managed data
- Identify what you already publish

A. 2. Citizens participation to improve Open Data portal productivity and efficiency

Outline

The lack of transparency on how public money is spent and the lack of control over the status of project are two of the main reasons for the slow pace in implementing public projects, frequently causing inefficiencies of all kinds (e.g., time and cost inefficiencies). Open data about projects' financings in conjunction with a platform that enables projects' on-site monitoring (and sharing of the results) can help solve these problems.

Links to the Revised PSI Directive

Platforms, Quality, Discoverability

Challenge

The lack of transparency on how public money is spent and the lack of control over the projects status (both from citizens and government at the central level) are two of the main reasons for the slow pace in implementing public projects, frequently causing inefficiencies of all kinds (e.g., time and cost inefficiencies). Moreover, citizens are frequently not involved and are not aware about projects that are taking place in the area where they live.

Solution

Open data about projects' financings in conjunction with a platform that enables projects' on-site monitoring (and sharing of the results) can help solve the problems mentioned above.

Why is this a Best Practice?

- It fosters participation of the citizens and efficiency of the public sector bodies.
- Making the citizens and the policy makers aware about a problem is the first step for solving it and improving the efficiency of public expenditures.
- Exposing open data about public projects financing and offering to the citizens a collaborative platform for controlling those projects, helps to eliminate some inefficiency in public spending and helps citizen to be actively involved in public projects.
- The platform for monitoring public projects is a method and a model whereby citizen monitoring may be initiated and a tool for civic partners to: press forward, report on malpractice, but also collaborate in making all these projects work, in accelerating their completion and understanding whether they actually respond to local demand.
- This approach fosters a civic use of open data, so that citizens can feel a closer connection with the ways in which public money is being employed and ultimately with public policies and decisions.

How do I implement this Best Practice?

Prerequisites for implementing this best practice are:

- Good quality open data (that in the case explained in the paper4 are published on OpenCoesione), they must be at least:
    - Understandable (through use of good metadata)

    - Machine processable
    - Complete
    - Up to Date
- A portal, like Monithon (that is an independently developed initiative), that permits active involvement of communities and furnishes a shared methodology for checking the actual state of the projects described by the open data on the portal described above.

Where has this best practice been implemented?

Country	Initiative	Contact Point
Italy	Open Coesione together with Monithon	Luigi Reggi, Head of the Monithon initiative Lorenzo Canova, Researcher at POLITO

References

- Samos Workshop Talk: Open Spending in Albania
- Samos Workshop Talk: OpenCoesione and Monithon - a Transparency Effort
- Timisoara Workshop Session: The Electronic Public Procurement System, open data and storytelling in Romania

Contact Info

Luigi Reggi, Lorenzo Canova

Related Best Practices

- Establish Open Government Portal for data sharing
- Select high value datasets for publication

A. 3. Cost-benefit analysis of the value of information

Outline

Costs of publication should be minimised unless there are clear benefits.

Links to the Revised PSI Directive

Dataset criteria, Charging

Challenge

It is difficult for public sector bodies to estimate the commercial or other value of their information. The risk of deciding what publication form will best deliver that value, and the work of converting it to that form, should be left to commercial product and service providers, and other consumers.

Solution

Costs of publication should be minimised unless there are clear benefits. Public sector body should analyse the current status of data availability, the demand for data and thus avoid unnecessary costs of data publication.

Why is this a Best Practice?

Public sector bodies should always ensure that their funding, which generally comes directly or indirectly from the citizens that they serve, is used wisely.

How do I implement this Best Practice?

When considering publication of a set of information, either

1. Publish it in the manner that involves lowest cost, consistent with making it available effectively and openly, or
2. Carry out cost-benefit analyses of the possible measures to assess potential use and stimulate take-up, methods of publication, and formats for publication, and select measures, methods and formats in the light of those analyses.

References

- The Open Public Sector Business Scenario (This business scenario was largely based on the Samos Share-PSI 2.0 workshop.)
- Share-PSI 2.0 Samos Workshop: Uses of Open Data Within Government for Innovation and Efficiency
- Krems Workshop Talk: RDB Rechtsdatenbank - Legal database for free research
- Krems Workshop Session: Towards A Sustainable Austrian Data Market

Related Best Practices

- Select high value datasets for publication

- Understand your internal needs and priorities
- Holistic metrics
- Provide PSI at zero charge

A. 4. Discover by location

Outline

Spatial information is helpful as a way of finding information.

Links to the Revised PSI Directive

Discoverability

Challenge

To make information discoverable by relating it to location.

Solution

Publishers should enrich information relating to real-world features (cities, rivers, mountains, etc.) or geo-political and geo-statistical entities (administrative areas, census and survey areas, etc.) with identifiers and links to their respective geospatial entities using standard approaches such as ISO 19115.

Why is this a Best Practice?

Information needs to be discovered in order to be used. It is also a core requirement of the PSI Directive. Improvements to the discoverability of information improve the chances that it will be reused. Incorporating geospatial identifiers helps people to develop applications that facilitate the discovering information relevant to specific locations e.g. through GPS-enabled devices such as mobile phones.

How do I implement this Best Practice?

In order to implement this best practice, you need

- an ability to publish information in open formats;
- to link the information elements to an authoritative geospatial reference entity.

References

- Lisbon Workshop Session: The Central Role of Location
- Berlin Workshop Track: Location Track

Related Best Practices

- Standards for Geospatial Data

A. 5. Maintain records of stakeholders' rights and interests

Outline

A number of stakeholders have legitimate interests in what is done with information and they might also have legal rights to confidentiality or privacy. Bodies collecting or holding public sector information must respect these rights and should conform to the interests and wishes of the relevant stakeholders. Therefore, bodies collecting or holding public sector information must maintain records of concerned stakeholders and their rights in the information and their wishes regarding it.

Links to the Revised PSI Directive

Policies and Legislation

Challenge

What should I do in order to ensure that privacy and other rights and interests are respected?

Operation within the law requires due respect for privacy and intellectual property rights. A number of stakeholders have also legitimate interests in what is done with information that bodies collecting or holding public sector information should respect.

Solution

Bodies collecting or holding public sector information should maintain records of concerned stakeholders and their rights in the information and their wishes regarding it. Bodies collecting information must inform concerned stakeholders of the reasons for collecting it, of how it will be used, and of the visibility that it will have. Bodies collecting information should also collect the wishes of stakeholders regarding its treatment and, in particular, its visibility.

Why is this a Best Practice?

Operation within the law requires due respect for privacy and intellectual property rights. Giving due consideration to the stakeholders' interests gives citizens and companies more confidence in government, and makes them more willing to supply information. Also, it involves greater engagement with them, which in itself contributes to the quality of government and the health of democratic society

Following this best practice:

- helps to ensure that the stakeholders' interests are considered and their rights are respected;
- helps to categorise data according to its openness;
- provides an input into the assessment of risks related to the data release.

How do I implement this Best Practice?

Identify rights and interests of the relevant stakeholders such as:

- People and companies that the information is about,
- People and companies whose property the information is about, and
- People and companies that have created or added value to the information.

Bodies collecting or holding public sector information include public-sector bodies and also commercial bodies involved in the collection, processing, and publication of the information.

References

- Story: Stakeholders' Interests and Rights
- The Open Group Business Scenario: Open Public Sector Data
- Samos Workshop Talk: Open Government Data Austria - Organisation, Procedures and Uptake
- Lisbon Workshop Session: Steps to a suitable redress mechanism
- Krems Workshop Session: Making and implementing a governmental open data policy

Related Best Practices

- Categorise openness of data
- Develop an Open Data Publication Plan

A. 6. Respect legislation and stakeholders' rights

Outline

Legislation might place restrictions on what information can be made available for reuse. Stakeholders such as people or organizations that the information is about might also have legitimate interest in what is done with the information. Organizations selecting datasets to be made available for free reuse should only select datasets that can be published so without violating the relevant legislation or the stakeholders' interests.

Links to the Revised PSI Directive

Policies and Legislation

Challenge

Not every dataset or piece of information held or collected by an organisation can be made available for free re-use due to the legislative restrictions or due to the wish of a stakeholder.

Solution

When selecting datasets to be made available for free reuse, check what legislation applies to the candidate datasets and whether it allows the datasets to be published. Check also that sharing the datasets is not against the interests of the relevant stakeholders. If required check that permission was given by all the required stakeholders.

Why is this a Best Practice?

Impact of the best practice is that only datasets that can be made available for reuse without violating the legislation or the relevant stakeholders' interests are selected for publication.

How do I implement this Best Practice?

Relevant legislation and stakeholders need to be identified for every datasets that is considered to be made available for free reuse. Classification of data/information sources according to their confidentiality or openness level helps to select datasets can be made available for reuse without restrictions.

Contact Info

Jan Kučera, University of Economics, Prague (Czech Republic).

Related Best Practices

- Maintain records of stakeholders' rights and interests
- Categorise openness of data

A. 7. Select high value datasets for publication

Outline

There are certain categories of data that are in demand and whose reuse has already proven to bring economic or social benefits. High value data such as geodata, public transport data or public spending data should be made available as open data.

Links to the Revised PSI Directive

Selection

Challenge

Which categories of data have already proven that are worth publishing?

Value of datasets might not be always evident. This fact complicates selection of datasets for publication.

Solution

In general, value of datasets for both reusers and the data owner should be assessed.

However, there are certain categories of data that are in demand and whose reuse has already proven to bring economic or social benefits. Geodata, public transport data or public spending data are examples of datasets with possible high value. High value datasets should be made available for reuse.

Why is this a Best Practice?

Publishers operate under resource constraints. At the same time, value of datasets for reusers or the data owner might not be always evident. Knowing what datasets represent high value datasets might help the publishers to focus their effort on datasets that are in demand and where the value can be demonstrated by successful case studies.

Implementing this best practice should lead to increased availability of high value datasets in machine-readable formats and under conditions permitting its reuse. Consequently, reuse of these datasets might lead to economic and social benefits.

How do I implement this Best Practice?

The Best Practice Dataset criteria provides generic guidelines for identification of high value datasets. Share-PSI 2.0 workshops shown that datasets from the following domains can be considered as high value datasets:

- Public transport information
- Public spending
- Geodata

You can also see datasets or data domains used for benchmarking in studies such as Open Knowledge Global Open Data Index or Open Data Barometer.

The European Commission also provides Guidelines on recommended standard licences, datasets and charging for the reuse of documents that contains a list of categories of data that should be made available for reuse.

However, it is necessary to keep in mind that even if some dataset does not fall into the 'high value' category today, it might become a high value dataset in the future. Therefore, this best practice should not be interpreted as a recommendation to publish only the high value datasets. It just implies that high value datasets are a good starting point of an Open Data initiative. Analysing reuse of the high value datasets might also help to understand, how value is created through data reuse which can subsequently help to assess possible value of other datasets.

References

- Samos Workshop Story: Opening Up Public Transport Information to Save Costs
- Timisoara Workshop Story: Free our maps
- Samos Workshop Story: Supervizor - An Indispensable Open Government Application (Transparency Of Public Spending)
- Timisoara Workshop Talk: Good practices for identifying high value datasets and engaging with reusers: the case of public tendering data
- Timisoara Workshop Session: How benchmarking tools can stimulate government departments to open up their data

Contact Info

Jan Kučera, University of Economics, Prague (Czech Republic).

Related Best Practices

- Dataset Criteria
- Open Up Public Transport Data
- Publish overview of managed data
- Develop an Open Data Publication Plan

A. 8. Study the companies that build on PSI at national level

Outline

The Spanish companies that create applications, products or value added services for third parties, from public sector information are studied periodically to better know their necessities in order to adapt the public sector information offer. Hereinafter we will call this set of companies: the infomediary sector.

Links to the Revised PSI Directive

(Re)use

Challenge

Discovering the main features of infomediary companies and their offered products and services which constitute the key agent in the reutilization process. Delineating the main characteristics of the primary information supply, which constitute the raw material from which the products and services are generated. Learning about the market and the demand for applications, products and services. Obtaining some proposals for improvement.

Solution

The study provides and analyses the key elements that define the infomediary sector in Spain, a sector made up of companies that produce products and / or services for sale to third parties from the public sector information. Results mainly focus on the following sectors:

- Basic characteristics of the infomediary industry
- Products, services and applications offered from reused information
- Marketing models and revenue

Why is this a Best Practice?

It allows fulfilling the objectives of the directive - reuse - and deepening the progress of the information society

- Fosters the development of new digital products and services, thereby stimulating economic and business activity and ultimately providing value for society as a whole.
- Helps to accomplish PSI Directive to each entity of its duty to publish public data and make that data available for reuse, adapting the offer to the private sector demand.
- Enables the existence of a global scenario that fosters the extraction of general conclusions and a general overview of the PSI situation in Spain, facilitating the use of this information to extract meaningful and actionable knowledge regarding the open data landscape.

Implementing this best practice could help:

- To foster job creation in the infomediary sector
- To get the most out of scarce public resources that are available in our country
- To increase reuse in Spain, by the private and also by the public sectors
- To focus the government agencies efforts on offering the relevant information to the infomediary sector

- To improve available datasets formats, the information quality and its accessibility.
- To promote better national regulations on PSI
- To enhance coordination between public and private sector in PSI reuse

How do I implement this Best Practice?

In order to implement this best practice, you need to:

- obtain a high-level mandate to endorse the study;
- identify companies that make up the infomediary sector;
- design and conduct a survey to find the necessary aspects;
- collaborate with reuse PSI industry business associations;
- apply the lessons learned from the study results;
- periodically repeat the survey to measure the progress made.

References

- Lisbon Workshop Talk: Spanish Infomediary Sector Characteristics
- Study June 2011 edition (in Spanish), http://www.ontsi.red.es/ontsi/sites/default/files/1308555551216.pdf
- Study July 2012 edition, Executive summary (in English), http://www.ontsi.red.es/ontsi/sites/default/files/121001_red_007_final_report_2012_edition__vf_en_1.pdf
- Study 2014 edition, Executive summary (in English), http://www.ontsi.red.es/ontsi/sites/default/files/executive_summary_public_infomediary_sector_2014.pdf (this edition has been published in march 2015)

Contact Info

soporte@datos.gob.es
http://administracionelectronica.gob.es/general/verContacto.htm

Related Best Practices

- Establish an Open Data Ecosystem
- Develop Open Data Business Models & Value Disciplines

A. 9. Understand demand for data

Outline

Not every dataset that an organisation can possibly publish as Open Data is equally relevant or interesting to the reusers. Understanding the demand for data might help organizations to focus on datasets that are relevant to the reusers which in turn might foster reuse of Open Data.

Links to the Revised PSI Directive

Selection

Challenge

What datasets are worth making available for reuse?

Selecting datasets to be published as Open Data might be challenging because it might not be always clear what datasets are worth making available for reuse.

Solution

Understanding who the potential reusers are and what datasets they demand helps publishers to focus on datasets that are relevant and interesting to the reusers.

Why is this a Best Practice?

Misalignment between the supply and demand sides of the Open Data ecosystem might negatively affect the reuse of data. Following this practice should lead to better alignment between the supply and demand for Open Data. Providing data that is in demand might help to attract reusers and thus it might help to develop the PSI reuse market.

Following this practice should also help the data publishers to better understand who the reusers are and how they reuse the data. Understanding the user base is a prerequisite for an effective user engagement and it might also help to facilitate development of an Open Data ecosystem.

How do I implement this Best Practice?

When selecting datasets for publication (potential) reusers should be identified for every candidate dataset. Identifying community crowd sourcing projects outside government institutions can also be an indicator of valuable datasets that should be made available as Open Data.

Engaging the reusers in the planning phase of an Open Data initiative might result in valuable feedback what datasets they need. Data journalists represent a user group that can help promote the released data.

References

- Samos Workshop Story: Getting Journalists Involved In the Process Of Opening Up Data
- Samos Workshop Talk: Open Spending in Albania

- Lisbon Workshop Bar Camp: Open Data Life Cycle and Infrastructure (notes, slides, diagram)
- Timisoara Workshop Session: Crowd sourcing alternatives to government data – how should governments respond? (paper, notes)
- Methodology for publishing datasets as open data (COMSODE)
- Open Data Handbook
- 5 stars of Open Data engagement model

Contact Info

Jan Kučera, University of Economics, Prague (Czech Republic).

Related Best Practices

- Encourage crowdsourcing around PSI
- Develop an Open Data Publication Plan
- Publish overview of managed data
- Dataset Criteria
- Establish an Open Data Ecosystem
- Understand your internal needs and priorities

A. 10. Understand your internal needs and priorities

Outline

If publication of Open Data is driven only by the external pressures, public administrations might miss the opportunity to seize benefits for themselves. Therefore, it is important for public sector bodies to understand their internal needs and priorities regarding their Open Data initiatives and take these needs and priorities into account when selecting datasets for publication.

Links to the Revised PSI Directive

Selection

Challenge

How to internalize possible benefits from making data available for reuse?

Focusing too much on the legal obligations and external pressures to release data for reuse leads to a situation when public sector bodies as data providers often cannot internalize possible benefits from data liberation.

Solution

Governmental bodies thus realise that, besides responding to external pressures, the actual justification for opening data assets is that public bodies could directly reap tangible benefits, resulting in efficiency, effectiveness and hive of public sector innovation. Therefore, it is important for public sector bodies to understand their internal needs and priorities regarding their Open Data initiatives and take these needs and priorities into account when selecting datasets for publication.

Why is this a Best Practice?

Needs and priorities of the data owners are no less important than the demand for data. Following this practice should help public sector bodies to reap benefits resulting from making data available for reuse, e.g. reduction of data requests and elimination of unnecessary duplication made possible by the fact that information assets are no more trapped into data silos, economies of scope fuelled by collaborative coding among public agencies and software reuse on top of an accessible common data layer, recourse to collective intelligence for the purpose of solving tough governmental problems through data science.

Following this practice would contribute to make a further step towards the obtainment of an outcome-based government whose actions demonstrate a clear link with their results generated (i.e., outcomes) in terms of value that, in turn, could be internalised by the governments (e.g., efficiency, effectiveness) without overlooking the quest for the creation of value for society at large ("public value").

How do I implement this Best Practice?

Prerequisites for implementing this best practice:

- An outcome-based mind-set

- Cost accounting routines able to quantify advantages that could be achieved thanks to Open Data
- Repertoire of case studies serving as source of inspiration

References

- Lisbon Workshop Story: Open Data 2.0 - Changing Perspectives

Related Best Practices

- Develop an Open Data Publication Plan
- Dataset Criteria
- Understand demand for data