

Best Practice: (Re)use federated tools

25 July 2016

This version

<https://www.w3.org/2013/share-psi/bp/ft-20160725/>

Latest version

<http://www.w3.org/2013/share-psi/bp/ft/>

Previous version

<https://www.w3.org/2013/share-psi/bp/ft-20160627/>

This is one of [a set of Best Practices](#) for implementing the [\(Revised\) PSI Directive](#) developed by the [Share-PSI 2.0 Thematic Network](#).

Share-PSI Best Practice: Develop a federation tool for open data portals by [Share-PSI 2.0](#) is licensed under a [Creative Commons Attribution 4.0 International License](#).

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 re-users 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 (eg 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 (eg 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	Implementation	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](#) M^a 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

Local Guidance

This Best Practice is cited by, or is consistent with, the advice given within the following guides:

- (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) [Using Open Public Sector Information](#)
- (International) [DCAT application profile implementation 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
- (Latvia) [Atvērto datu vadlīnijas](#) Open Data Guidelines
- (Lithuania) [Viešojo Sektoriaus Informacijos platinimo gerosios praktikos](#) Best Practices for Sharing Public Sector Information
- (Luxembourg) [Recommandations pour l'ouverture des données publiques](#) Recommendations for opening data
- (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](#)

Issue Tracker

Any matters arising from this BP, including implementation experience, lessons learnt, places where it has been implemented or guides that cite this BP can be recorded and discussed on the project's [GitHub repository](#)