DELIVERABLE

D4.2 Metadata specifications

Project Acronym:	PoliVisu	
Project title:	Policy Development based on Advanced Geospatial Data Analytics and Visualisation	
Grant Agreement No.	769608	
Website:	www.polivisu.eu	
Contact:	info@polivisu.eu	
Version:	1.0	
Date:	31 August 2018	
Responsible Partner:	HSRS	
Contributing Partners:	EVIV, GEOS, MACQ	
Reviewers:	Lieven Raes (AIV) Gert Vervaet (AIV) Michal Kepka (P4A) Dirk Frigne (GEOS)	External Reviewers: Andrew Stott Bart De Lathouwer
Dissemination Level:	Public	X
	Confidential – only consortium members and European Commission Services	

This project has received financial support from the European Union Horizon 2020 Programme under grant agreement no. 769608.

Revision History

Revision	Date	Author	Organization(s)	Description
0.1	29/05/2018	Tomáš Řezník	HSRS	Initial draft, added section 2
0.2	11/06/2018	Tomáš Řezník, Štěpán Kafka, Lieven Raes, Gert Vervaet, Matteo Satta	HSRS, AIV, GeoSparc, ISY	Added sections 3 and 4
0.3	21/06/2018	Tomáš Řezník, Štěpán Kafka, Karel Charvát	HSRS	Added sections 3 and 4
0.4	17/07/2018	10 stakeholders participating in the seminar at the JRC (lead by Tomáš Řezník, organised by Andrea Perego)	European Commission, DG Joint Research Centre	Revision of the all the metadata specifications by external experts
0.5	24/07/2018	Tomáš Řezník, Gert Vervaet	HSRS, GeoSparc	Text consolidations, added Annex A
0.6	24/08/2018	Tomáš Řezník		Version for internal and external review
0.7	28/08/2018	Tomáš Řezník Jiří Bouchal	HSRS ISP	Review remarks integrated
1.0	31/08/2018	Jiří Bouchal	ISP	Final Version

Table of Contents

List of Tables

List of Figures

Executive Summary

Information resources from several domains are required when proposing and/or revising a policy. It is difficult to share and re-use those information resources within diverse domains, like it is required within PoliVisu. Mechanisms for unified discovery of information resources need to be established as one of the first steps. This deliverable therefore enhances the PoliVisu D4.1 Technical specifications with high-detail technical information on metadata specifications and management. As such, it also represents a cornerstone for the D5.1 Technical Integration Plan.

We are nowadays witnesses of huge development of metadata specifications. The World Wide Web Consortium aims at tighter integration of data and metadata, the Open Geospatial Consortium revises its Catalogue Service for the Web, hundreds of data and service producers implement INSPIRE discovery services, geospatial and non-geospatial resources are being integrated through approaches like CKAN, Google tries to provide answers directly instead of metadata when feasible or in the latter case offers more user-friendly and added value rich cards/snippets etc.

So far, any concept or implementation does not provide sufficient support for consistent discovery of geospatial as well as non-geospatial resources. The PoliVisu approach re-uses the existing well-proven approaches and continues with their further enhancement as well as new developments.

The unified core metadata structure was defined when following the ISO 19115 Geographic Information – Metadata standard together with ISO 19100 family of standards, OGC implementation specifications and W3C recommendations. A guide to populating and managing the PoliVisu catalogue was written to support the common vision of metadata management.

The developed PoliVisu discovery concept comprises three layers. The first discovery layer is being used for backwards domain compatibility, based on the INSPIRE discovery services (backwards compatible to the OGC catalogue services) and CKAN domain-specific catalogues. The second discovery layer represents a gate to the world of Semantic Web through DCAT-AP, GeoDCAT-AP, RDF and Schema.org. The third discovery layer touches the users of mainstream full-text searching engines through the technologies like Google rich cards/snippets.

Moreover, such three layered discovery concept enables to handle geospatial and non-geospatial resources homogeneously. The benefits of such clear metadata management are following. The developed PoliVisu catalogue is capable of answering the questions like: “Show me which resources are related to those sensor measurements.” The results in such a case comprise links to e.g. a description of a traffic model that was used to process the measurements, an interpolated dataset originating from the measurements, map composition depicting the measurement as well as the interpolated dataset, legal act created upon the measurement, Web services publishing the sensor measurements and interpolated dataset, standardization document defining the framework for quality evaluation of such measurements, Website describing the purpose of the measurements and its outcomes to the public etc.

The PoliVisu metadata specifications have been discussed within and beyond the PoliVisu consortium. A seminar with metadata experts of the European Commission, DG Joint Research Centre as well as metadata experts within the Open Geospatial Consortium helped with improvement of metadata specifications. The relevant parts from the developed PoliVisu metadata specifications are being approved as the OGC GeoDCAT-AP Best practice report. Moreover, a white paper regarding the metadata management is planned during the further development of the PoliVisu project. Such white paper will aim at managing the (metadata of) visualizations, revised metadata formats etc. The intention of such a white paper is to support the PoliVisu development in the most up-to-date way, especially since the updates of the D4.2 Metadata specification document are not foreseen in the Description of Work.

1 Introduction

The PoliVisu project uniquely combines diverse information resources from several domains to evolve the traditional public policy making cycle. The underlying diverse information resources comprise, among others, big datasets, sensor data, social media, open data, metadata, geospatial data, etc. Those diverse information resources are commonly used in a wide range of human activities for different purposes. As such, they have a great potential to be re-used by different institutions and companies. On the other hand, the effective (re)use of all those information resources is paralysed by limited knowledge of where relevant data is stored and/or published. The term “data islands” could be used in this context. Discovering the localisation of information resources (data) storage is therefore the key to reusing information.

Re-use of information resources require reliable discovery mechanisms – namely, inventorying and searching capabilities. However, reliable metadata is not enough. It has to be published within reliable catalogue services to offer the value of metadata to as wide a public as possible. This deliverable therefore describes the PoliVisu metadata specifications that ensure unified and efficient handling with metadata, their linkages to data as well as inventorying and searching capabilities.

Regarding the state-of-the-art analysis, the major advantage of existing metadata specifications is that there are so many to choose from. Each domain has at least one metadata specification including a list of descriptive items that should be used to “label” the information resource. However, descriptions concerning titles, abstracts, publication dates, formats, and publishers, etc. are the same across all metadata specifications. It may be even stated that the basic set of metadata items is the same for geographical data/services as for any human product known in daily life, as depicted in Figure 1.

The metadata content has then to be encoded into an exchange format. XML (eXtensible Markup Language) is the most common, since it is used, for example, by the ISO 19139, Dublin Core (DC) Metadata Initiative, the Content Standard for Digital Geospatial Metadata (CSDGM), and semantics-oriented approaches. For further information on this topic, see e.g. Nogueras-Iso, Zarazaga-Soria and Muro-Medrano (2005) [22], Moellering et al. (2005) [20] or Reznik et al. (2016) [25]. The exchange format may be used directly for metadata transfer between organisations and/or used as an input for a catalogue service. In the end, relevant information resources may be discovered through the metadata of a catalogue, for example, a Web browser as a client application of a catalogue (service).

Catalogue services developed according to the implementation specifications of the Open Geospatial Consortium (OGC) represent today one of the most common means of metadata publication within the domain of geographic information. For more information on the details of OGC catalogue services, see Nebert, Whiteside and Vretanos (2007) [21].

The necessity of cataloguing in a spatial data infrastructure (SDI) seems to be universally accepted. This may also be supported by experience originating from Europe. In 2003, Directive 2003/98/EC (also known as PSI – Public Sector Information [8], revised version [10]) established a minimum set of rules governing both the re-use and the practical means of facilitating the re-use of existing documents held by public sector bodies in the European Union. In the end, Directive 2003/98/EC had only a partial impact in the field of geographical information. In 2007, Directive 2007/2/ES (also known as INSPIRE – INfrastructure for SPatial InfoRmation in Europe [9]) was established, chiefly to make it easier to discover available spatial data and services.

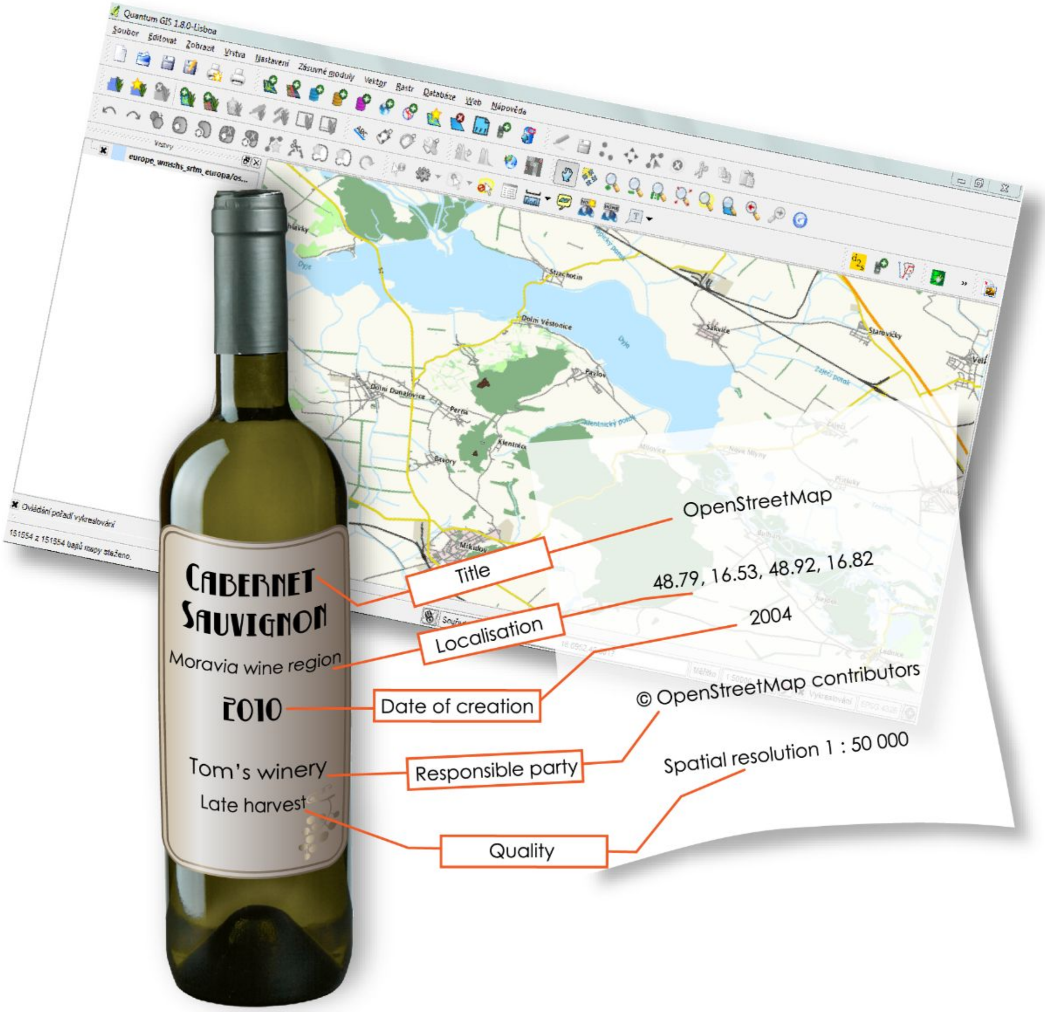


Figure 1: The basic set of metadata items required for the description of an information resource is the same for any human product known in daily life (adopted from Reznik et al., 2015 [25])

The trends of the last decade comprise shift to open (meta)data and cataloguing services. The primary motivation is to increase usage of information resources within as well as beyond the community/communities where those information resources were created. Integrating data portals and publishing metadata on the web in a linked open data format is a first step in making data more discoverable. The (semantic) linked approaches preserves methods and formats used within a community for managing and publishing metadata, while at the same being able to make them also discoverable in other environments.

The following text is dedicated to relevant aspects with higher technical emphasis. As such, section 2 discusses the structure of metadata for information resources related to the PoliVisu scope. Section 3 then deals with cataloguing capabilities including integration with contemporary development. Section 4 comprises implementation vision of a catalogue and meta-information system within the PoliVisu project. Finally, Annex A brings a 'cookbook' for populating and maintaining PoliVisu catalogue.

2 Metadata structure

The following text is related to the short descriptions on relevant metadata standards and related initiatives that aim at unifying the metadata structure. The basic information has been adopted from GeoDCAT-AP Best practice report [23] that is a document of the Open Geospatial Consortium mostly written and coordinated by the partners of the PoliVisu project as well as the NextGEOSS project (http://nextgeoss.eu/) as the linked activity.

2.1 Dublin Core

The Dublin Core Schema [11] is a small set of vocabulary terms that can be used to describe documents, Web resources (video, images, Web pages, etc.), as well as physical resources such as books or CDs, and objects like artworks. Dublin Core Metadata may be used for multiple purposes, from simple resource description to combining metadata vocabularies of different metadata standards, to providing interoperability for metadata vocabularies in Linked Data Cloud and Semantic Web implementations.

The Dublin Core standard has almost no strict regulations. Each of fifteen Dublin Core elements is optional and may be repeated. There is also no prescribed order for presenting or using the elements. Note that the Dublin Core also exists in its qualified version with 18 metadata elements and full version with 55 metadata elements. Nevertheless, the initial version with 15 metadata elements is the most used one as well as the only one version with ISO equivalent, ISO 15836:2009 Information and documentation -- The Dublin Core metadata element set.

2.2 DCAT

DCAT is an RDF [30] vocabulary designed to facilitate interoperability between data catalogues published on the Web.

By using DCAT to describe datasets in data catalogues, publishers increase discoverability and enable applications to easily consume metadata from multiple catalogues. It further enables decentralised publishing of catalogues and facilitates federated dataset search across sites. Aggregated DCAT metadata can serve as a manifest file to facilitate digital preservation. DCAT makes extensive use of terms from other vocabularies, in particular, Dublin Core.

DCAT uses a layered concept: Catalog, consisting of one or more datasets and one or more distribution for each dataset (see also Figure 2).

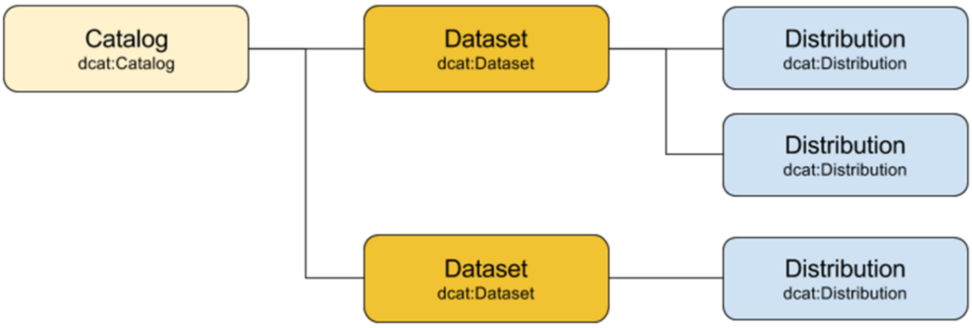


Figure 2: Meta-data structure according to DCAT

A dataset in DCAT is defined as a "collection of data, published or curated by a single agent, and available for access or download in one or more formats". A dataset does not have to be available as a downloadable file. For example, a dataset that is available via an API can be defined as an instance of dcat:Dataset and the API can be defined as an instance of dcat:Distribution. DCAT itself does not define properties specific to APIs descriptions. These are considered out of the scope of this version of the vocabulary. Nevertheless, this can be defined as a profile of the DCAT vocabulary.

2.3 DCAT-AP

The DCAT Application Profile (DCAT-AP) for data portals in Europe is a specification based on DCAT for describing public sector datasets in Europe. It is used to enable cross-data portal search for data sets and make public sector data better searchable across borders and sectors.

The DCAT-AP consists of a set of metadata elements, policies and guidelines defined for a particular application. An application profile is not complete without documentation that defines the used policies and best practices appropriate to the application.

The DCAT-AP is very common in many European data portals, including the EU data portal (see Figure 3). The DCAT-AP suggests the use of specific categorizations and describes which fields are mandatory, recommended or optional.

A good example of a DCAT-AP profile defining the policies is the Belgian DCAT-AP profile: http://dcat.be/ where the Federal level, the regions and individual cities are using the same DCAT-AP profile to describe their datasets. Other European countries like Italy, Netherlands, the United Kingdom, the Czech Republic, Germany, Sweden, and Norway have also implemented their own DCAT-AP profile.

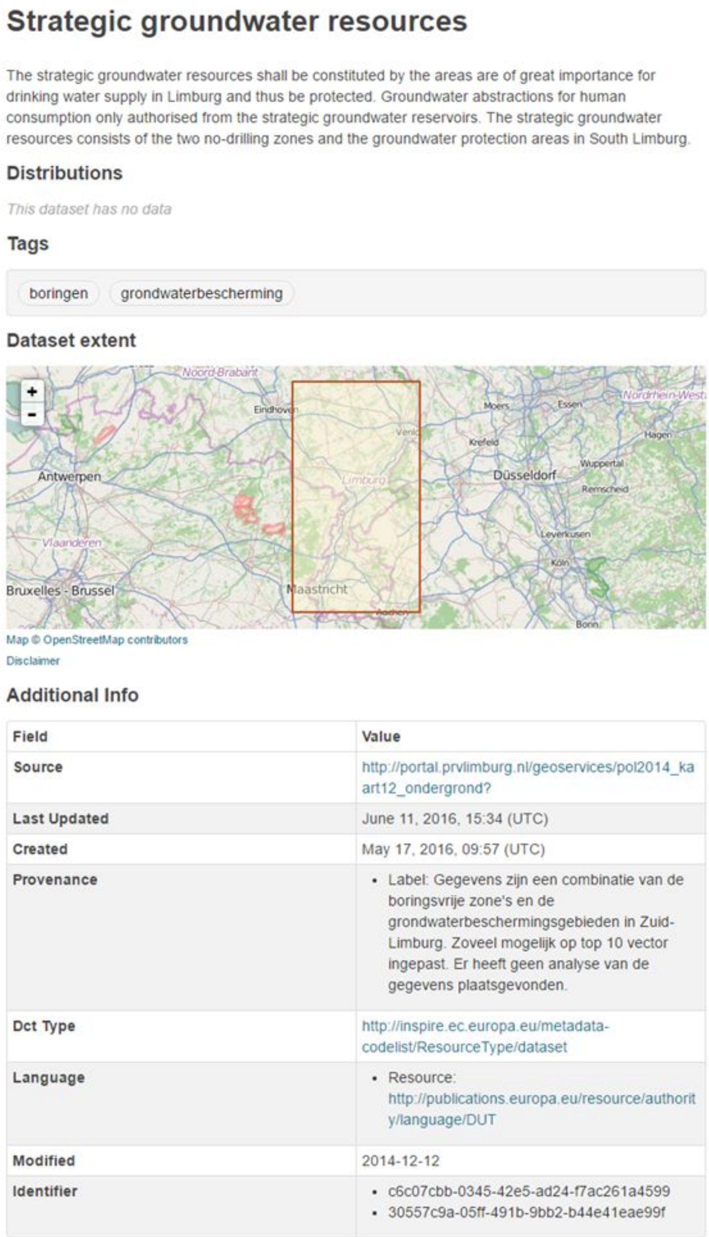


Figure 3: Example of DCAT-AP on the EU data portal (adopted)

2.4 Linked Open Data

Beside the open data metadata standards, other relevant non-thematic initiatives have a strong link with metadata standards and metadata publication. Linked Open Data (LOD) and Schema.org are related initiatives to make metadata more discoverable on the web.
The term Linked Data refers to a set of best practices for publishing structured data on the Web. These principles have been coined by Tim Berners-Lee in the design issue note Linked Data. The principles are:

1. use URIs as names for things,
2. use HTTP URIs so that people can look up those names,
3. when someone looks up a URI, provide useful information,
4. include links to other URIs so that they can discover more things.

The idea behind these principles is on the one hand side, to use standards for the representation and access to data on the Web. On the other hand, the principles propagate to set hyperlinks between data from different sources. These hyperlinks connect all Linked Data into a single global data graph, similar as the hyperlinks on the classic Web connect all HTML documents into a single global information space. Thus, Linked Data is to spreadsheets and databases what the Web of hypertext documents is to word processor files. Linked Open Data allows to open entire datasets on a metadata level and on the level of the data itself.

2.5 ISO standards

2.5.1 ISO 19115:2003 and ISO 19115:2014 - Geographic information metadata

ISO 19115 [15] is a standard of the International Organization for Standardization (ISO). The standard is part of the ISO geographic information Suite of Standards (19100 series). ISO 19115 and its parts define how to describe geographical information and associated services, including contents, spatial-temporal purchases, data quality, access and rights to use.
The objective of this International Standard is to provide a clear procedure for the description of digital geographic datasets and services so that users will be able to determine whether the data in a holding will be of use to them and how to access the data. By establishing a common set of metadata terminology, definitions and extension procedures, this standard promotes the proper use and effective retrieval of geographic data.
ISO 19115 was revised in 2013 to accommodate growing use of the internet for metadata management, as well as to add many new categories of metadata elements (referred to as code lists) and the ability to limit the extent of metadata use temporally or by the user.

2.5.2 ISO 19119:2005 and ISO 19119:2016 - Geographic information services

ISO 19119:2005 [16] identifies and defines the architecture patterns for service interfaces used for geographic information, defines its relationship to the Open Systems Environment model, presents a geographic services taxonomy and a list of example geographic services placed in the services taxonomy.
It also prescribes how to create a platform-neutral service specification, how to derive conformant platform-specific service specifications, and provides guidelines for the selection and specification of geographic services from both platform-neutral and platform-specific perspectives.
ISO 19119:2005 was revised and resulted in ISO 19119:2016. The 2016 version defines requirements for how platform neutral and platform specific specification of services shall be created, in order to allow for one service to be specified independently of one or more underlying distributed computing platforms. It also

defines requirements for a further mapping from platform neutral to platform specific service specifications, in order to enable conformant and interoperable service implementations.
ISO 19119 includes also the specific requirements and metadata model for metadata for services.
The INSPIRE Metadata specifications or Implementing Rules and the 'INSPIRE Metadata Technical Guidelines' are based on EN ISO 19115 and EN ISO 19119:2005 for describing metadata for data sets, data sets series and services [27].

2.5.3 ISO/TS 19139:2007 Geographic information - Metadata - XML Schema Implementation

Defines in [17] Geographic MetaData XML (gmd) encoding, an XML Schema implementation derived from ISO 19115. Geographic metadata is represented in ISO 19115 as a set of UML packages containing one or more UML classes. ISO 19115 provides a universal, encoding-independent view of geographic information metadata. ISO/TS 19139 provides a universal implementation of ISO 19115 through an XML schema encoding that conforms to the rules described in ISO 19118 (Encoding).

2.5.4 ISO 19150 Rules for developing ontologies in the Web Ontology Language (OWL)

ISO/TS 19150-1:2012 defines the framework for semantic interoperability of geographic information. This framework describes a high level model of the components required to handle semantics in the ISO geographic information standards with the use of ontologies.
ISO 19150-2:2015 [18] defines rules and guidelines for the development of ontologies to support better the interoperability of geographic information over the Semantic Web. The Web Ontology Language (OWL, [29]) is the language adopted for ontologies.
It defines the conversion of the UML static view modelling elements used in the ISO geographic information standards into OWL. It further defines conversion rules for describing application schemas based on the General Feature Model defined in ISO 19109 into OWL.
It does not define semantics operators, rules for service ontologies, and does not develop any ontology.

2.6 INSPIRE Metadata specification

The INSPIRE Directive aims to create a European Union's spatial data infrastructure for the purposes of environmental policies and policies or activities which may have an impact on the environment. This European Spatial Data Infrastructure enables the sharing of environmental spatial information among public sector organisations, facilitate public access to spatial information across Europe and assist in policy-making across boundaries.
INSPIRE is based on the infrastructures for spatial information established and operated by the Member States of the European Union. The Directive addresses 34 spatial data themes needed for environment-related applications. The Directive came into force on 15 May 2007 and is implemented in various stages, with full implementation required by 2021.
To make the thematic information sources coming from the 28 EU member states interoperable on the metadata level, INSPIRE established metadata implementing rules (i.e. legislation) and metadata technical guidelines.
The 'INSPIRE Metadata Implementing Rules' [5] is an EU Regulation which establishes the obligations under which public sector bodies in the EU should publish descriptive metadata on geographic data sets (series) and services. It describes a set of twenty-one metadata elements to be used to describe geographic datasets, dataset series and services. In order to further harmonize the implementation of these obligations, the

(non-binding) 'INSPIRE Metadata Technical Guidelines' [27] were defined. These technical guidelines show how the metadata elements in the Regulation match metadata elements in ISO 19115: 2003 and ISO 19119: 2007. The specification also contains examples how metadata records can be transformed into XML. Although the Technical Guidelines are not binding, it is strongly advised to follow them as otherwise it is very difficult to demonstrate that the metadata provided in a different model / format are in accordance with the INSPIRE Metadata Implementing Rules (which themselves are binding).
These technical Guidelines prescribe the use of ISO19115:2003. However, a metadata record that complies with the ISO19115: 2003 core elements is not fully compliant with the INSPIRE Metadata Implementing Rules and its derived Technical Guidelines. There are some additional metadata elements that INSPIRE requires on top of the ISO19115: 2003 core profile, e.g. the (degree of) ‘conformity’ of spatial datasets with the INSPIRE Data specifications.

2.7 Summarizing points

Closing the gap between open data standards and geospatial standards isn’t an easy process because of the organic growth of standards over the past decades (see also Figure 4). Because of a lack of knowledge of each other’s standards, there are several redundancies between metadata information, and there is a need for data mapping to allow re-using metadata and to make data and other resources discoverable over the Web, independently of the portal that hosts the metadata of the resource.
The figure below gives an overview of the most important standards and initiatives in the geospatial and open data fields. The standards section (metadata specifications) below, shows that these initiatives are clearly divided between specific geospatial oriented metadata standards (ISO 19115 & 19119) and non-geospatial standards as Dublin Core. The ISO 19115 based EU oriented INSPIRE data specifications are also part of the family of the geospatial metadata standards.
The machine to machine communication layer contains standards and initiatives to exchange information between data portals or business software with a data catalogue included. The most common formats are, XML, RDF[1] and Turtle[2]. Where ISO 19139 and DCAT are respectively geospatial and open-data oriented formats, the very general RDF triples and more particular GeoDCAT-AP format are multisector ones. (Meta)data publication as linked open data and GeoDCAT-AP can integrate open non-geospatial metadata and geospatial metadata in a way that it can be used for publication on the web. Both initiatives (GeoDCAT-AP and more general the publication of RDF triples) are intermediary formats/technologies.
The Web publication of geospatial and open datasets is nowadays done by using geoportals or open data portals. The latter have in some cases also a basic geo-visualisation function often displaying a bounding box. Another technique is using information feeds like Atom. The most commonly used feeds don’t have a geospatial element included. Schema.org is an interesting format to publish all the metadata directly on the web. Schema.org can also contain some basic geospatial data. The power of Schema.org lies in the fact that data can be found directly using a search engine.
GeoDCAT-AP as an extension on DCAT-AP is a very powerful intermediary format to translate ISO 19100 or INSPIRE stored metadata into an RDF compatible format that can be used to publish metadata on (geo)data portals and on the web. GeoDCAT-AP as an intermediary format today can exchange open data and geospatial data in a well-defined format.

[1] That is also an XML-based format.
[2] Terse RDF Triple Language

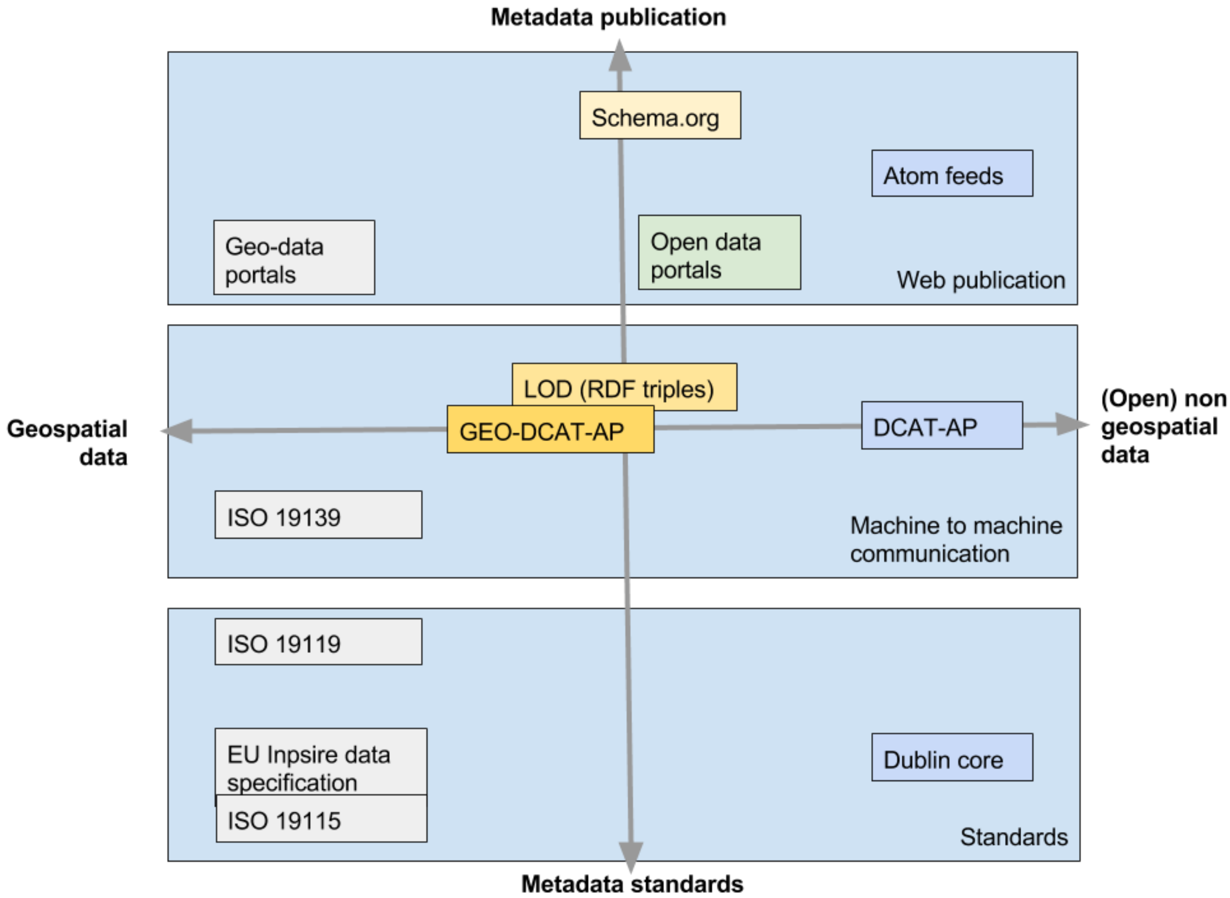


Figure 4: Metadata specifications quadrant overview

3 Metadata publication principles

3.1 OGC Catalogue Service for Web

As described by the OGC implementation specifications [21] and [26], catalogue services are the key technology for locating, managing and maintaining distributed geo-resources (i.e. geospatial data, applications and services). With catalogue services, client applications are capable of searching for (geo-)resources in a standardized way (i.e. through standardized interfaces and operations) and, ideally, they are based on a well-known information model, which includes spatial references and further descriptive (thematic) information that enables client applications to search for geo-resources in very efficient ways.

The following use cases may be used as the essential ones for any catalogue service:

- **Discover metadata:** A requestor discovers metadata entries in a catalogue service either by browsing the content of the catalogue or by placing certain query terms. If a service is discovered that fits his/her search terms, he/she can bind to this service in accordance with the information in the result sets of the catalogue service.
- **Harvest metadata:** A catalogue service may harvest metadata records from another catalogue (service) as long as they rely on the same standardization framework. For instance, about one thousand catalogue services have been developed under the INSPIRE Directive (2007/2/EC). Any of

them may be harvested for PoliVisu purposes as they rely on the same implementation specification as the PoliVisu catalogue.

- **Publish metadata:** A publisher describes (geo-)resources according to a standardization framework (could be one of those described in the sections from 2.1 to 2.6). In the Polivisu scope, a (geo-)resource might be a dataset, service, textual document, sensor measurement, video, tweet, model etc. As an alternative, the publisher might be a broker that does not own the geo-resource, but describes and publishes metadata descriptions to a catalogue service on behalf of a publisher. An example of a broker may be a consultant or subcontractor. Publisher or broker loads metadata into a catalogue. Technically speaking, described (geo-)resources are published by a catalogue service that is one kind of a Web service based on request/response pattern.

The Open Geospatial Consortium's implementation specification Catalogue Service for Web was issued in 2007. It is designed as a guide for a catalogue service development with wide support of underlying metadata frameworks. Metadata structure may therefore follow as default the structure according to the Dublin Core (see section 2.1), ISO standards (as described in sections 2.5.1, 2.5.2 and 2.5.3), INSPIRE (see section 2.6), Content Standard for Digital Geospatial metadata (used as the default one in the United States of America and Canada, see [13]), OASIS eBusiness Registry Information Model (ebRIM, see [19]) and others. Unfortunately, semantic approaches were not available at the time of the implementation specification creation.

The latest version of the OGC Catalogue Service for Web (CSW) is 3.0.0, however the most commonly used is the previous version 2.0.2. The huge support of the 2.0.2 version has been a consequence of its use in legislation, such as INSPIRE [6].

The minimal query language syntax is based on the SQL WHERE clause in the SQL SELECT statement. Implementations of query languages that are transformable to the Contextual[3] Query Language [7] are the OGC Filter Encoding Specification [28] and the CIP and GEO profiles of Z39.50 Type-1 queries [14]. The query language used within a catalogue service is extensible.

So-called queryables have been defined in order to ensure interoperability among catalogues that implement the same protocol binding and query compatibility among catalogues that implement different protocol bindings perhaps through the use of “bridges” or protocol adapters. Defining a set of core queryable properties also enables simple cross-profile discovery, where the same queries can be executed against any catalogue service without modification and without detailed knowledge of the catalogue's information model. This requires a set of general metadata properties that can be used to characterize any resource. Table 1 brings overview on the mandatory queryables that have to be supported in the PoliVisu catalogue for its backwards compliance to the OGC implementation specification.

Table 1: Mandatory queryables according to the OGC Catalogue for Web 2.0.2 implementation specification

Name	Definition	Data type
Subject	The topic of the content of the resource.	CharacterString
Title	A name given to the resource.	CharacterString
Abstract	A summary of the content of the resource.	CharacterString
AnyText	A target for full-text search of character data types in a catalogue.	CharacterString
Format	The physical or digital manifestation of the resource.	CharacterString

[3] Formerly known as ‘Common’ as well ‘Common Catalogue’, sometimes also together with ‘OGC’ in its name.

Identifier		A unique reference to the record within the catalogue.	Identifier
Modified		Date on which the record was created or updated within the catalogue.	Date-8601
Type		The nature or genre of the content of the resource. Type can include general categories, genres or aggregation levels of content.	CodeList
BoundingBox		A bounding box for identifying a geographic area of interest.	Compound
	WestBoundLongitude	Western-most coordinate of the limit of the resource's extent, expressed in longitude in decimal degrees (positive east).	Numeric
	SouthBoundLatitude	Southern-most coordinate of the limit of the resource's extent, expressed in latitude in decimal degrees (positive north).	Numeric
	EastBoundLongitude	Eastern-most coordinate of the limit of the resource's extent, expressed in longitude in decimal degrees (positive east).	Numeric
	NorthBoundLatitude	Northern-most, coordinate of the limit of the resource's extent, expressed in latitude in decimal degrees (positive north).	Numeric
CRS		Geographic Coordinate Reference System (Authority and ID) for the BoundingBox.	Identifier
Association		Complete statement of a one-to-one relationship.	Association,
	Target	Referenced resource.	Identifier
	Source	Referencing resource.	Identifier
	Relation	The name of the description of the relationship.	CodeList or Identifier

Three classes of operations may be identified from the interface operations point of view (see also Figure 5):

- **service operations** which are operations a client may use to interrogate the service to determine its capabilities;
- **discovery operations** which a client may use to determine the information model of the catalogue and query catalogue records;
- **management operations** which are used to create or change records in the catalogue.

The following brief description of operations supports Fig. 5 as well as is a basis for the PoliVisu catalogue implementation:

- **GetCapabilities:** The mandatory operation allowing CSW clients to retrieve service metadata from a server. The response to a GetCapabilities request shall be an XML document containing service metadata about the server.

- **DescribeRecord:** The mandatory operation allowing a client to discover elements of the information model supported by the target catalogue service. The operation allows some or all of the information model to be described.
- **GetDomain:** The optional operation used to obtain runtime information about the range of values of a metadata record element or request parameter.
- **GetRecords:** The mandatory operation through which a client specifies his/her filters on data discovery, such as the title of a dataset I am searching for, its geospatial extent etc. The returned values are metadata, i.e. a description of a (geo-)resource.
- **GetRecordById:** The mandatory operation retrieving the default representation of catalogue records using their identifier. This operation presumes that a previous query has been performed in order to obtain the identifiers that may be used with this operation.
- **Record Locking:** Such optional mechanism does not define a locking interface, instead relying on the underlying repository to mediate concurrent access to catalogue records.
- **Transaction:** The optional operation defines an interface for creating, modifying and deleting catalogue records. The transaction operation may be used to "push" data into the catalogue.
- **Harvest:** An optional operation that may be used to create or update records in the catalogue. The primary motivation is "pull" data into the catalogue.

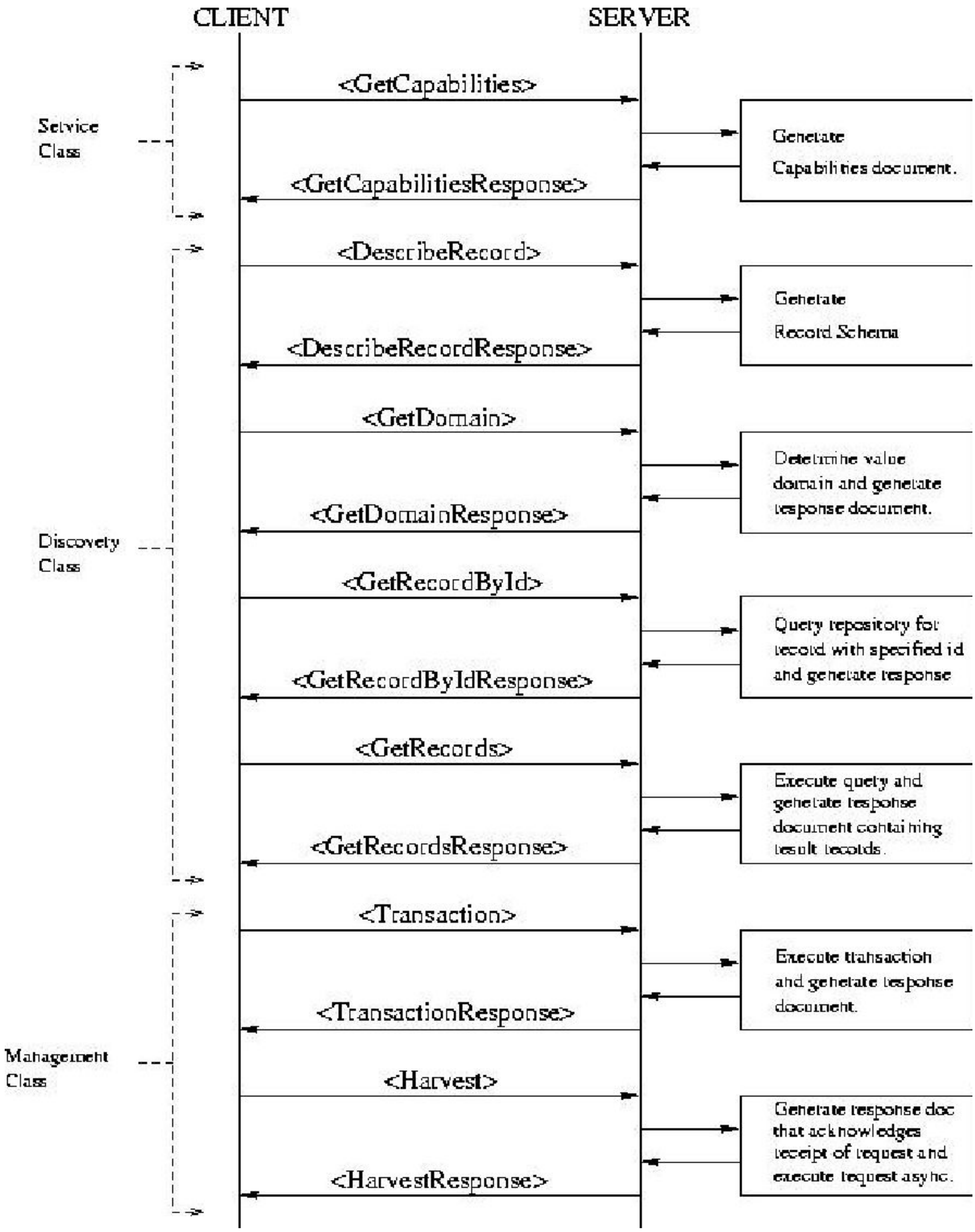


Figure 5: Operations and classes of operations including their purpose in a Catalogue Service for Web (adopted from the OGC implementation specification).

3.2 CKAN

Unlike the OGC Catalogue Service for Web, Comprehensive Knowledge Archive Network (CKAN, [4]) is an activity beyond the geospatial domain. The CKAN was developed as a Web-based open source management system for storage and distribution of open data. It is intended to power data hubs and data portals as it enables to publish, share and use data. The CKAN also aims at previewing discovered data through maps, graphs and tables.

From the technological point of view, the CKAN is built with Python on the backend and Javascript on the frontend, and uses The Pylons Web framework and SQLAlchemy as its object-relational mapping. Its database engine is PostgreSQL and its search is powered by SOLR (searching platform built on Apache Lucene [1]). It has a modular architecture that allows extensions to be developed to provide additional features such as harvesting or data upload.

The CKAN uses its internal model to store metadata about the different information resources, and presents it on a Web interface that allows users to browse and search those metadata. It also offers an API that allows third-party applications and services to be built around it.

All the dataset descriptive fields, as depicted in the Table 2, are searchable (see below for the metadata we bring out into the interface). Shortly said, data may be added and edited in the following ways:

- manually via the Web interface,
- using CKAN's rich JSON API,
- via custom spreadsheet importers.

Table 2: The CKAN descriptive fields of which all are searchable separately as well as commonly through full text searching (adopted from http://ckan.org)

Title	Definition
Title	Allows intuitive labelling of the dataset for search, sharing and linking
Unique identifier	Dataset has a unique URL which is customizable by the publisher.
Groups	Display which groups the dataset belongs to if applicable. Groups (such as science data) allow easier data linking, finding and sharing amongst interested publishers and users.
Data preview	Preview .csv data quickly and easily in browser to see if this is the dataset you want.
Revision history	CKAN allows you to display a revision history for datasets which are freely editable by users (as is thedatahub.org)
Extra fields	These hold any additional information, such as location data (see the description of geospatial features below) or types relevant to the publisher or dataset. How and where extra fields display is customizable.
Licence	Instant view of whether the data is available under an open licence or not. This makes it clear to users whether they have the rights to use, change and re-distribute the data.
Multiple formats (if provided)	See the different formats the data has been made available in quickly in a table, with any further information relating to specific files provided inline.
API key	Allows access every metadata field of the dataset and ability to change the data if you have the relevant permissions via API.

Automatic tagging is one of the advantages of the CKAN approach, as depicted in Fig. 6. The tags "geospatial" and "test" were derived automatically from data. Such tags are also discoverable through so-called faceted search which drills-down via facets like tags, format, licence or publisher. The faceted search is another discovery option to full-text search (à la Google) and fuzzy matching (an option to search for closely matching terms instead of exact matches). All the searching capabilities are also available over an API.

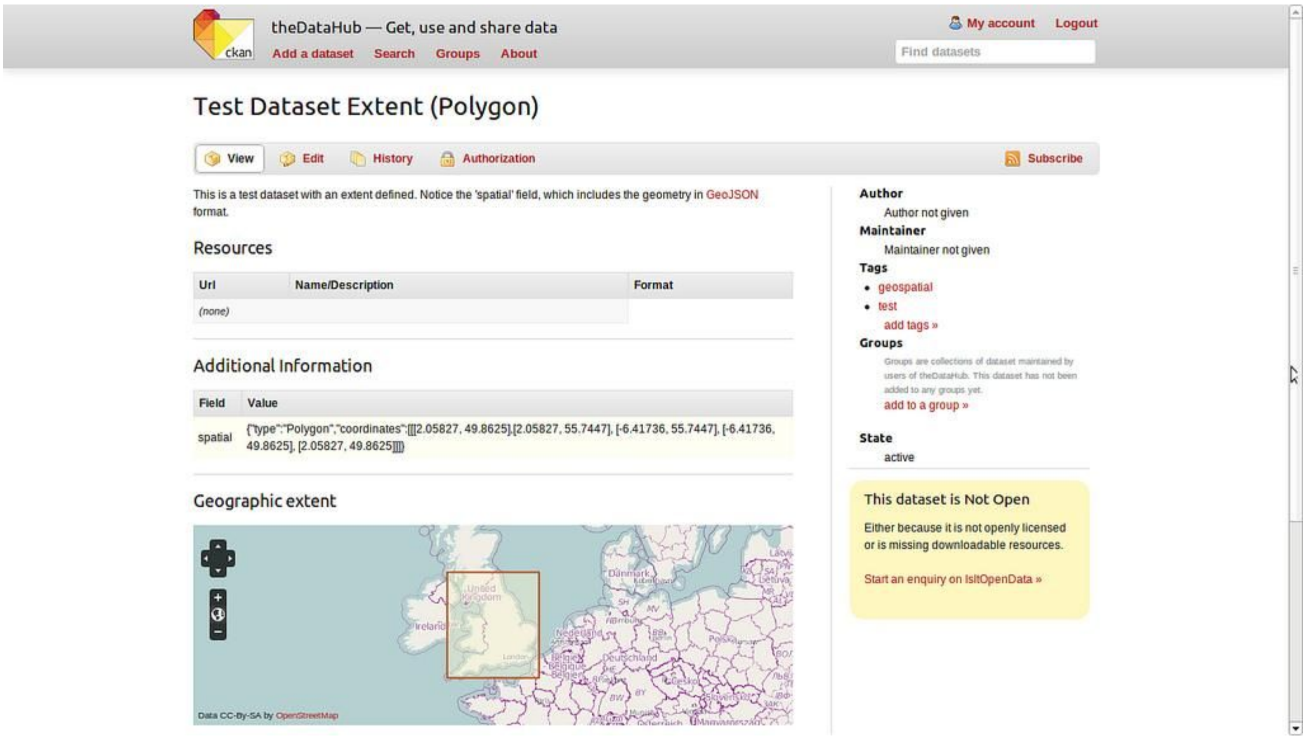


Figure 6: An example of the CKAN discovery GUI with geospatial and tagging properties (adopted from https://github.com/ckan).

CKAN is the de facto standard software for publishing open data with more than 200 community extensions. In a very short time, it has been widely adopted by high profile public projects, such as http://data.gov.uk (UK), http://data.gov (USA) and http://data.gov.au (Australia). It is a mature, open project with a sound model of development and a growing community of developers.

Moreover, the CKAN can be used to create a federated network of data portals which share data between each other. The CKAN supports the DCAT standard [2] for data catalogue metadata, so data can also be federated from other non-CKAN catalogues. Harvesting capabilities comprise the following:

- geospatial CSW Servers,
- existing Web catalogues,
- simple HTML index pages or Web Accessible Folders,
- ArcGIS, Geoportal Servers and Z39.50 databases,
- other CKAN instances.

From the geospatial perspective, the CKAN can display data on an interactive map when structured data with location information are loaded into CKAN's DataStore. The spatial extension [3] enables a CKAN instance to understand a location associated with a dataset, and use this to offer geospatial search capabilities via the Web interface and API. A user searching for datasets can filter the results by geographical location,

specifying a bounding box to limit the area of interest. The CKAN understands different coordinate geometries and parses location information accordingly.
The CKAN includes tools to import geo-coded metadata in a number of formats and make it queryable ('discoverable') according to the INSPIRE directive. It can import major metadata schemas such as ISO 19139, GEMINI 2.1 and FGDC (also known as CSDGM) can handle records hosted in a variety of ways, including the geospatial OGC CSW standard, WAFs, ArcGIS portals, Geoportal Servers and Z39.50 databases. CKAN can also serve geospatial packages via its own CSW interface. The architecture is extensible, making it easy to support other standards and distribution services.

3.3 Google rich cards/snippets

The business-oriented terms Google rich cards and Google rich snippets are synonyms for the same concept enhancing the 'traditional' full-text searching capabilities. Both stand for a metadata approach that comprises results of a full-text searching together with additional details (metadata) about the discovered results. We may identify plenty of examples, for instance displaying photos of a certain place, reviews for restaurants, ingredients for a receipt, movies or social media accounts for individuals (see also Fig. 7). However, geospatial-based examples, beyond photos related to some location, are scarce. Google rich cards/snippets are provided since 2016 in selected countries, globally since 2017.
So far, up to June 2018, the following set of extensible properties are provided through special interactive features within the Google rich cards/snippets:

- Calculator,
- time zone, currency, and unit conversions,
- word translations,
- flight status,
- local film showings,
- weather forecasts,
- population and unemployment rates,
- package tracking,
- word definitions,
- "do a barrel roll" (search page spins),
- "askew" (results show up sideways).

Shortly said, the shift in paradigm is from 'bringing links to information resources with answers' to 'bringing the answers'. For instance, I will receive thanks to the in-built calculator interactive feature the answer '2' when searching for '1+1' instead of pointing to resources where the answer for the question '1 + 1' is. Such shift brings the advantage of unambiguous simple answer. The weakness is on the contrary a trustfulness of the discovery algorithm. It is very easy in such an approach to promote e.g. one dataset as the best one, for instance when answering the question 'base map for the Czech Republic'.

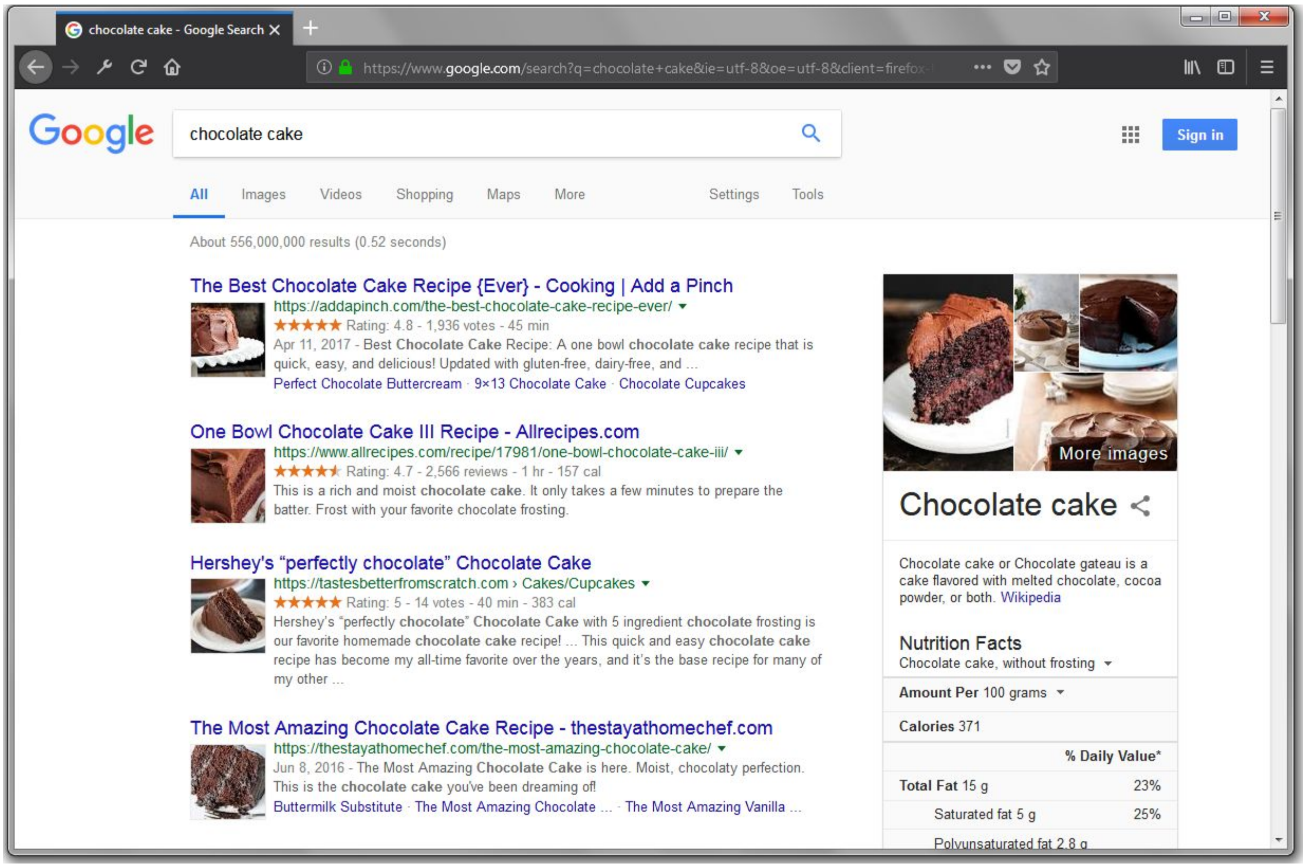


Figure. 7: Example of the Google rich cards/snippets when searching for a "chocolate cake" (adopted from http://google.com). Additional metadata like review rating, cook duration, number of reviews, nutritional value etc. are displayed together with the result of the discovery process.

3.4 Summarizing points

Similarly to section 2, we may identify two cataloguing bubbles: geospatial and non-geospatial. The concept of metadata publication and discovery has been in the geospatial domain settled down for last 15 years thanks to the OGC implementation specifications on catalogue services. Such implementations have been commonly used in spatial data infrastructures and have significantly contributed to discovery and re-use of geospatial data, Web services and other resources.

Nevertheless, the OGC-based concept for metadata publication and discovery is a specific one as it only partly overlaps with the latest mainstream IT publication and discovery mechanisms. The open-source initiatives like CKAN, as well as official activities like the GeoDCAT-AP [23] are aimed to bridge geospatial and non-geospatial domains. The benefits of such metadata publication and discovery are clear, it enables to link all the related resources like "show me which resources are related to those sensor measurements". The results in such a case comprise links to e.g. a description of a traffic model that was used to process the measurements, an interpolated dataset originating from the measurements, map composition depicting the measurement as well as the interpolated dataset, legal act created upon the measurement, Web services publishing the sensor measurements and interpolated dataset, standardization document defining the framework for quality evaluation of such measurements, Website describing the purpose of the measurements and its outcomes to the public etc.

4 Implementation Vision

The implementation of publication and discovery capabilities within the PoliVisu project aims at (1) connecting the existing concepts described above as well as (2) development of new extensions to existing standardization activities. The PoliVisu metadata development is therefore based on the BDVA architecture, as described in the PoliVisu's D4.1 Technical specifications. The whole metadata concept is understood and as a PoliVisu metadata pipeline, depicted in Figure 8. The PoliVisu approach re-uses all the available metadata existing developments within and beyond the pilot cities. Namely, it:

- harvests metadata from existing OGC Catalogue Service for Web servers (no matter whether being developed within legally defined framework such as INSPIRE Discovery Service and/or during any other provider's activity as a 'pure' OGC CSW),
- harvests metadata from endpoints of OGC-compliant Web services,
- adopts metadata from file-based formats, such as XML in a well-described structure,
- incorporates metadata from semantically-rich repositories, such as CKAN and/or SPARQL endpoints.

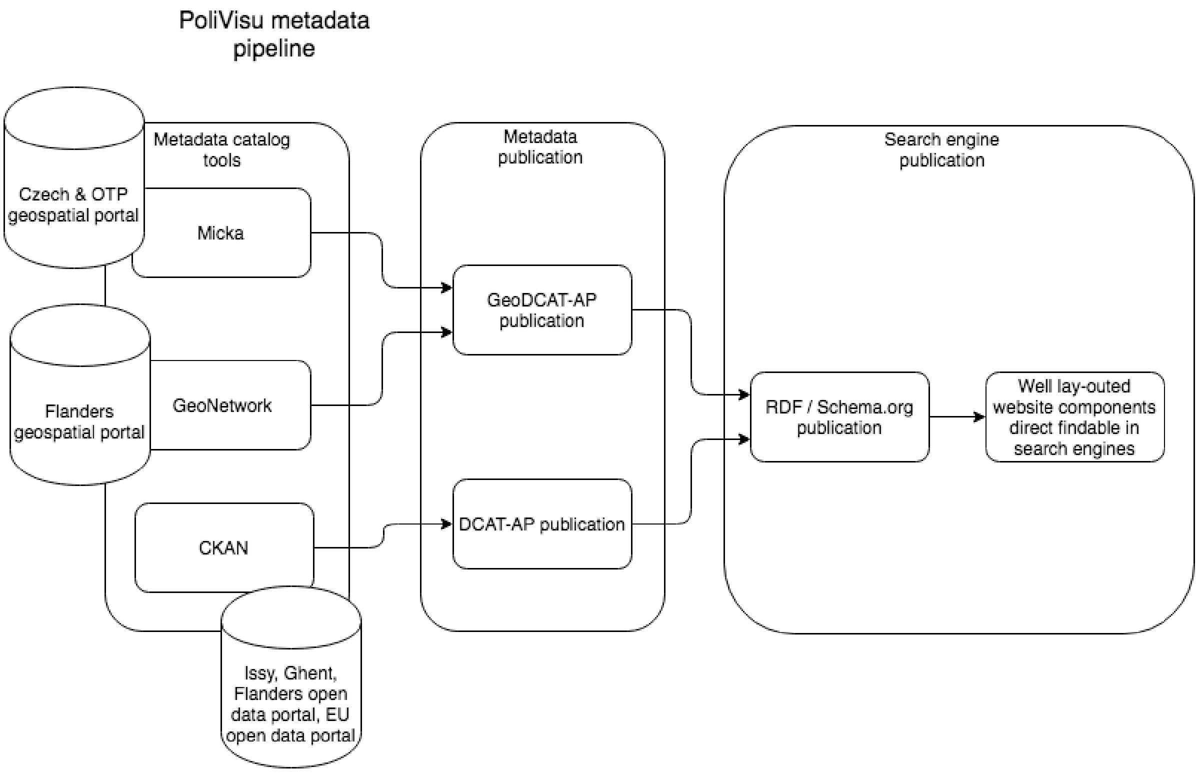


Figure 8: The basic schema of the PoliVisu metadata pipeline.

Figure 9 then shows the isolated approaches of various de jure and de facto metadata and discovery services standardization efforts. The depicted "worlds" also demonstrates the PoliVisu backwards compatibility to the underlying standardization efforts.

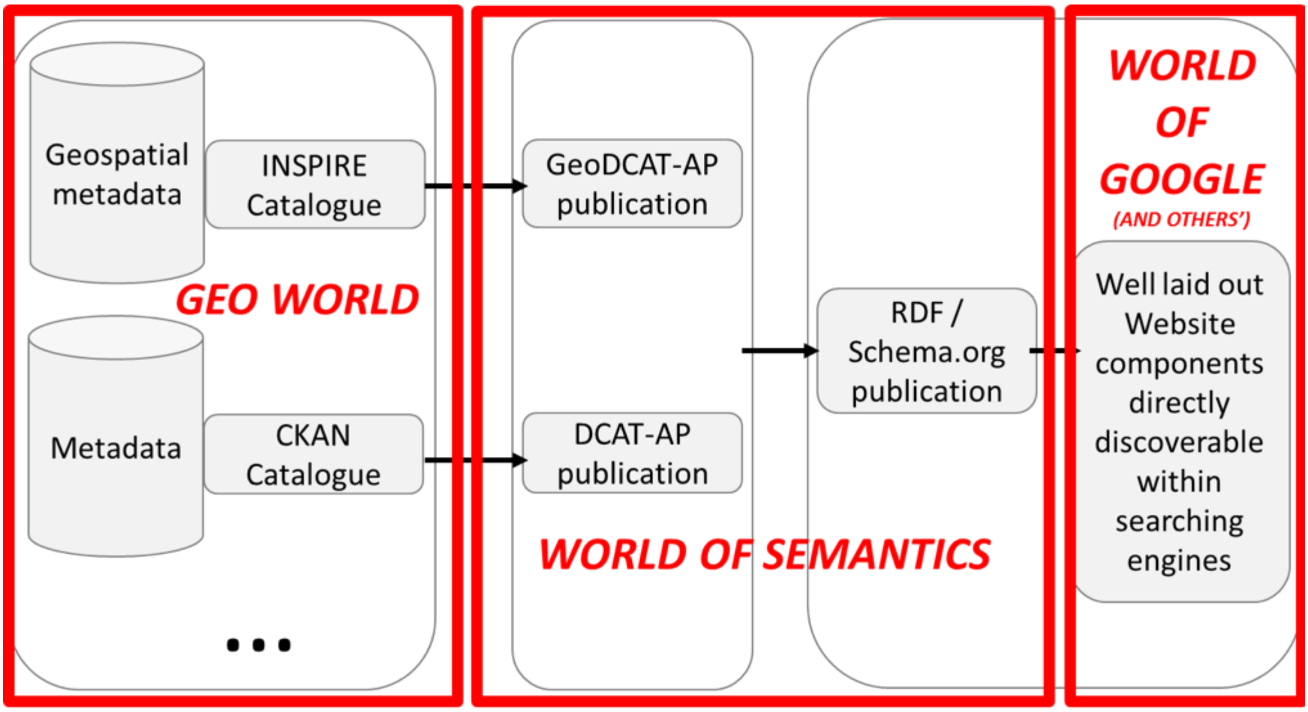


Figure 9: Backwards compatibility of the PoliVisu metadata pipeline to existing standardization efforts.

From the integration perspective, the (Open)Micka tool represents the main module for metadata management (creation, harvesting, updates, edits, publication, linking and discovery) as described in the D5.1 Technical Integration Plan. All relevant technical integration details are written in that deliverable including deployment UML diagrams, input and output methods in the API, installation and licensing details etc.

The PoliVisu (Open)Micka testing environment has been set up under the URL https://dev.bnhelp.cz/micka-demo/en/.

The Annex A of this deliverable brings a cookbook for populating the PoliVisu catalogue and further metadata management.

4.1 The core metadata structure

As agreed within the PoliVisu consortium, the EN ISO 19115: 2003 Geographic Information – Metadata including Cor. 1:2006 is the essential standard for the development and integration. The ISO 19115 core metadata elements were selected to provide the basis for interoperability on the one hand and not to add a new burden in form of metadata management on the other hand. The list of ISO 19115 core metadata elements is depicted in Table 3. **PoliVisu metadata are customizable, scalable and extensible. The core metadata elements presented in Table represent the basis that needs to be followed by all the PoliVisu stakeholders. On the contrary, the core metadata elements may be customized and/or further metadata elements may be added as far as still following the principles defined in the ISO 19115 standard.**

Table 3: The core metadata elements including details that are being used to describe PoliVisu geospatial and non-geospatial resources

Title	Required as	Definition	XPath in ISO 19139
Title	mandatory	Name by which the cited resource is known.	MD_Metadata/MD_DataIdentification/citation/CI_Citation/title
Reference date	mandatory	Reference date for the cited resource.	MD_Metadata/MD_DataIdentification/citation/CI_Citation/date
Responsible party	optional	Identification of, and means of communication with, person(s) and organization(s) associated with the resource(s).	MD_Metadata/MD_DataIdentification/pointOfContact/CI_ResponsibleParty
Geographic location	conditional (if available)	Geographic position of the dataset.	MD_Metadata/MD_DataIdentification/extent/EX_Extent/EX_GeographicExtent/EX_GeographicBoundingBox or EX_GeographicDescription
Language	mandatory	Language(s) used within the resource.	MD_Metadata/MD_DataIdentification/language
Character set	conditional (mandatory if differs from UTF-8)	Full name of the character encoding standard used for the dataset.	MD_Metadata/MD_DataIdentification/characterSet
Topic category	mandatory	Main theme(s) of the dataset.	MD_Metadata/MD_DataIdentification/topicCategory
Spatial resolution	optional	Factor which provides a general understanding of the density of spatial data in the dataset.	MD_Metadata/MD_DataIdentification/spatialResolution/MD_Resolution/equivalentScale or MD_Resolution/distance
Abstract	mandatory	Brief narrative summary of the content of the resource(s).	MD_Metadata/MD_DataIdentification/abstract
Distribution format	optional	Provides a description of the format of the data to be distributed.	MD_Metadata/MD_Distribution/MD_Format/name and MD_Format/version
Additional extent	optional	Time and/or vertical extent covered by the content of the resource.	MD_Metadata/MD_DataIdentification/extent/EX_Extent/ EX_TemporalExtent or EX_VerticalExtent
Spatial representation type	optional	Method used to spatially represent geographic information.	MD_Metadata/MD_DataIdentification/spatialRepresentationType

Reference system	optional	Description of the spatial and temporal reference systems used in the dataset.	MD_Metadata/MD_ReferenceSystem
Lineage	optional	Information about the events or source data used in constructing the data specified by the scope or lack of knowledge about lineage.	MD_Metadata/DQ_DataQuality/lineage/LI_Lineage
On-line resource	optional	Location (address) for on-line access using a Uniform Resource Locator address or similar addressing scheme.	MD_Metadata/MD_Distribution/MD_DigitalTransferOption/onLine/CI_OnlineResource
Metadata file identifier	optional	Unique identifier for this metadata file.	MD_Metadata/fileIdentifier
Metadata standard name	optional	Name of the metadata standard (including profile name) used.	MD_Metadata/metadataStandardName
Metadata standard version	optional	Version (profile) of the metadata standard used.	MD_Metadata/metadataStandardVersion
Metadata language	conditional	Language used for documenting metadata.	MD_Metadata/language
Metadata character set	conditional	Full name of the character coding standard used for the metadata set.	MD_Metadata/characterSet
Metadata point of contact	mandatory	Party responsible for the metadata information.	MD_Metadata.contact/CI_ResponsibleParty
Metadata date stamp	mandatory	Date that the metadata was created or revised.	MD_Metadata/dateStamp

4.2 Enhancing the metadata structure

As written above, one of the goals of the PoliVisu project is to adopt the latest development while remaining backwards compatible to the existing approaches, such as ISO and INSPIRE. A customization of the ISO 19115 metadata elements seems as a feasible way of achieving both goals. The ISO 19115 core metadata elements are in PoliVisu used as the cornerstone for description of all the resources that are not falling under INSPIRE and/or those not having already their own scheme. The mandatory ISO 19115 core metadata elements consist of:

- title,
- reference date (creation/revision/publication),
- abstract,
- language,
- metadata point of contact,
- metadata date stamp,
- topic category.

The ISO 19115 does not support linking capabilities between the core metadata records. Such support is only partial between the conditional and optional metadata elements:

- a relationship from subset to superset (only one directional relationship assumed in the ISO 19115) may be discovered through the *ParentIdentifier* metadata element,
- a relationship from a Web service to a dataset is realised through the *OperatesOn* metadata element,
- a relationship from a dataset to a Web service is unambiguously realised through the *transferOptions.onLine* metadata element (more relationships feasible under the same metadata element),
- a relationship between a dataset derived from another dataset could be realised through the *lineage.source.description* metadata element represented as *gmx:Anchor* or *xlink:href* XML concepts; in any case, a register of the relations would be beneficial.

The PoliVisu approach therefore aims at customizing the ISO 19115 metadata to incorporate mechanisms for linking metadata and related resources. So far, the following relationships have been identified from the PoliVisu pilot resources as described in the deliverable D6.1 Pilot Scenarios:

- A is **published** by B (a dataset being published by a Web service),
- A is **derived from** B (an interpolated dataset derived from a sensor measurement),
- A is **developed in line with** B (a dataset developed in line with legislation/model/…),
- A is **a part of** B (a sensor measurement being a part of an initiative/campaign/…),
- A is **described** by B (a dataset being described by a document/Web site/…),
- A **has a successor** B (a ceased dataset is being replaced by a new dataset),
- A **has visualisation** B (a policy goal is visualized through a map composition).

The above presented list of relationships was discussed within and beyond the PoliVisu consortium. Among others, the above mentioned relationships were discussed within the seminar entitled "Developing a Bright Discovery Service Combining Geo- and Non-Geo Resources More Efficiently Than Google (at least so far)" that was organized within the European Commission, Directorate General Joint Research Centre (JRC) on 17 July 2018. The outcomes of this seminar underlined the necessity of such relationships. On the contrary, such relationships could not be published through the INSPIRE Registry (http://inspire.ec.europa.eu/registry) as originally planned. The items published in the INSPIRE Registry have to be first approved by the representatives of the European Union Member States. Since such procedure is longer than the time left till the end of the PoliVisu project, the consortium will develop their own register of relationships for linking metadata and related resources. Moreover, the ongoing work aims at mapping the relationships to GeoDCAT to enhance the vague *xlink* concept. A PoliVisu proof-of-concept implementation of an approach identifying parent and child datasets for the Open Land Use dataset[4] is depicted in Figure 9.

Note that the above mentioned list of relationships is an extensible one.

[4] See https://sdi4apps.eu/open_land_use/ for more information.

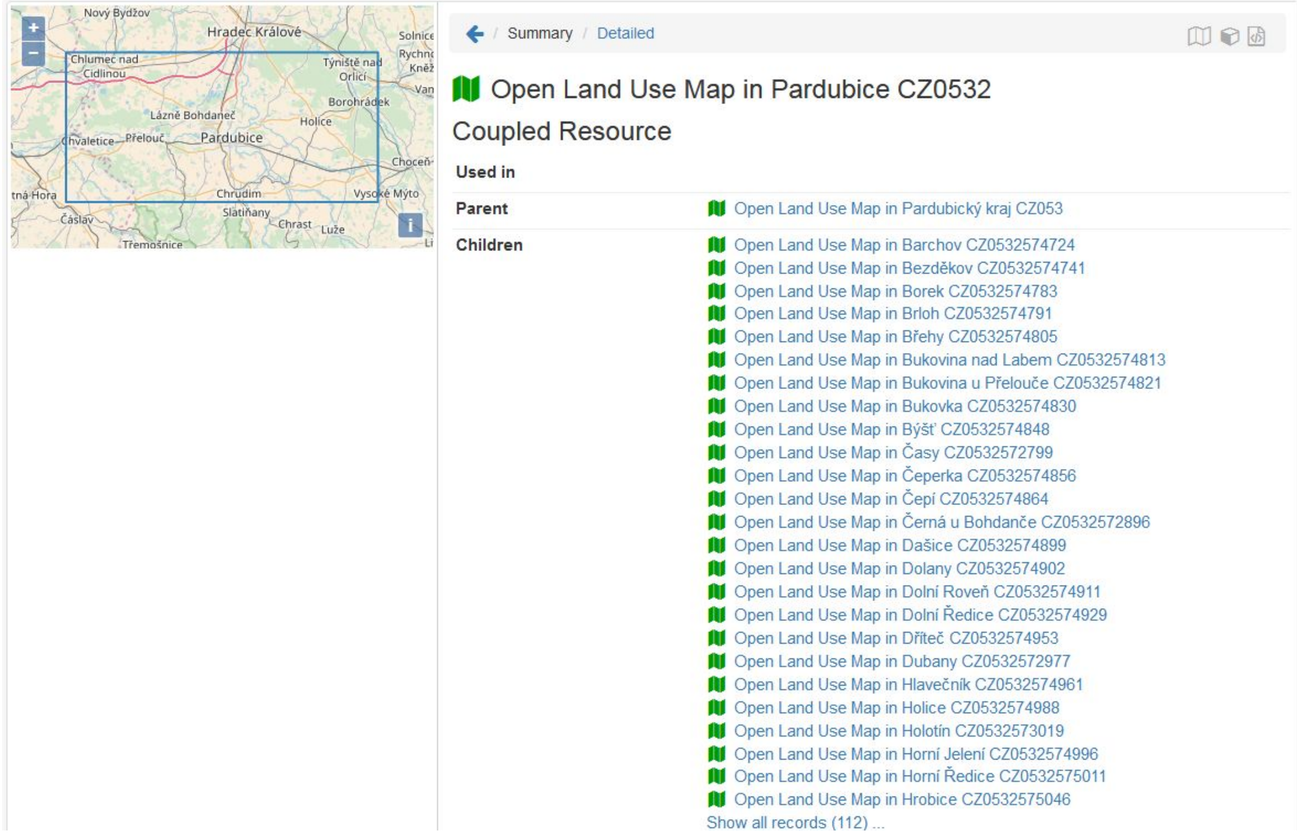


Figure 9: The PoliVisu proof-of-concept implementation for (meta)data linking: the Open Land Use dataset

Metadata descriptions in PoliVisu are considered as scalable. It means that any resource provider may customize, modify and/or extend metadata according to his/her purposes unless he/she does not conflict with existing standardization documents and/or internal PoliVisu metadata specifications. Annex A is presented at the end of this document to minimize such risks.

5 Conclusions

Equal handling and linking of geo- and non-geo- resources is a basic prerequisite of the PoliVisu project. Therefore, a structure and meaning of ISO 19115 core metadata elements are being re-used within PoliVisu to consistently describe geo- and non-geo- resources. Such descriptions are adopted from existing solutions and automatically transformed into the required structure whenever feasible. On contrary, manual metadata population seems necessary for resources that, so far, have not been described by metadata.

Metadata structure and storage is only the beginning of the PoliVisu metadata specifications. Three layers of discovery services have been designed to properly integrate geo- and non-geo- resources as well as to provide functionality to various kinds of users:

- The first discovery layer targets the geospatial and domain experts.
    - The INSPIRE compliant discovery services are supported for the geospatial experts as well as for the interoperability to the catalogues existing within the geospatial domain, including the spatial data infrastructures.
    - The CKAN catalogues are used for domain experts that are beyond the geospatial domain. Such catalogues support resources like Web pages, models, documents etc.
- The second discovery layer targets users of semantics-based applications. As such, this layer offers metadata in form following the relevant de jure and de facto standards: DCAT-AP, GeoDCAT-AP, RDF and Schema.org.
- The third discovery layer targets the masses of mainstream search engines' users. Well layouted Website components are published to become directly discoverable within search engines. Google rich cards/snippets are the best examples of the PoliVisu third discovery layers outcomes.

A PoliVisu catalogue was established to unify metadata management within the PoliVisu project. A set of metadata management use cases was described in the form of a guide (see Annex A).

The future work on metadata specifications will focus on further integration with ongoing standardization efforts of the World Wide Web Consortium and Open Geospatial Consortium as well as on further integration between data and metadata.

6 References

[1] *Apache SOLR* [online]. 2017. Cited 19 June 2018. Available at: <http://lucene.apache.org/solr/>.

[2] *GitHub - ckan/ckanext-dcat: CKAN ♥DCAT* [online]. Cited 19 June 2018. Available at: <https://github.com/ckan/ckanext-dcat>.

[3] GitHub - ckan/ckanext-spatial: Geospatial extension for CKAN [online]. Cited 19 June 2018. Available at: <https://github.com/ckan/ckanext-spatial>.

[4] CKAN - The open source data portal software [online]. Cited 19 June 2018. Available at: <https://ckan.org>.

[5] *Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata* [online]. Official Journal of the European Union L326/12. Published 4 December 2008. 19 pages. Cited 19 June 2018. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32008R1205&from=EN>.

[6] *Commission Regulation (EC) No 976/2009 of 19 October 2009 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards the Network Services* [online]. Official Journal of the European Union L274/9. Published 20 October 2009. 10 pages. Cited 19 June 2018. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32009R0976&from=EN>.

[7] *Contextual Query Language* [online]. The Library of Congress. Published 30 August 2013. Cited 19 June 2018. Available at: <https://www.loc.gov/standards/sru/cql/>.

[8] *Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information* [online]. Official Journal L 345. Published 31 December 2003, 7 pages. Cited 29 March 2018. Available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32003L0098&from=en>.

[9] *Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)* [online]. Official Journal L 108. Published 25th April 2007, p. 0001 – 0014. Cited 29 May 2018. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32007L0002&from=EN>.

[10] *Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information* [online]. Official Journal L175/1. Published 27 June 2013, 8 pages. Cited 29 March 2018. Available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013L0037>.

[11] DCMI. 2018. *Dublin Core Metadata Initiative* [online]. Cited 12 June 2018. Available at: <http://dublincore.org/>.

[12] European Commission. 2017*. DCAT Application Profile for data portals in Europe* [online]. Cited 13 June 2018. Available at: <https://ec.europa.eu/isa2/solutions/dcat-application-profile-data-portals-europe_en>.

[13] Federal Geographic Data Committee. 1998. *Content Standard for Digital Geospatial Metadata* [online]. Cited 19 June 2018. Available at: <https://www.fgdc.gov/standards/projects/metadata/base-metadata/v2_0698.pdf>.

[14] *ANSI/NISO Z39.50-2003 Information Retrieval: Application Service Definition and Protocol Specification* [online]. National Information Standards Organization. Cited 19 June 2018. Available at: <https://www.loc.gov/z3950/agency/Z39-50-2003.pdf>.

[15] *ISO 19115 Geographic Information – Metadata*. International Organization for Standardization: Geneva. 140 pages.

[16] *ISO 19119 Geographic Information – Services*. International Organization for Standardization: Geneva. 67 pages.

[17] *ISO 19139 Geographic Information – Metadata – XML Schema Implementation*. International Organization for Standardization: Geneva. 111 pages.

[18] *ISO 19150-2:2015: Geographic information - Ontology - Part 2*. http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=57466, July 2015.

[19] Martell, R. (ed). 2009. *CSW-ebRIM Registry Service - Part 1: ebRIM profile of CSW* [online]. Open Geospatial Consortium. Cited 19 June 2018. Available at: <http://portal.opengeospatial.org/files/?artifact_id=31137>.

[20] Moellering, H. et al. *World spatial metadata standards*. International Cartographic association. Elsevier Ltd., London 2005. 247 s. ISBN: 0-08-043949-7.

[21] Nebert, D., Whiteside, A., Vretanos, P. (ed). 2007. OpenGIS Catalogue Service Specification [online]. Open Geospatial Consortium. Cited 29 May 2018. Available at: <http://portal.opengeospatial.org/files/?artifact_id=20555>.

[22] Nogueras-Iso, J., Zarazaga-Soria, F. J., Muro-Medrano, P. R. 2005. Geographic Information Metadata for Spatial Data Infrastructures. Berlin Heidelberg: Springer. 263 pages. ISBN 3-540-24464-6.

[23] Perego, A., Cetl, V., Friis-Christensen, A., Lutz, M. 2017. GeoDCAT-AP: Representing geographic metadata by using the "DCAT application profile for data portals in Europe" [online]. Cited 19 June 2018. Available at: <https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2017/mtg3/2017-UNECE-topic-i-EC-GeoDCAT-ap-paper.pdf>.

[24] Raes, L., Vandenbroucke, D. 2017. *GeoDCAT-AP Best practice report* [so far internal OGC document] 94 pages.

[25] Reznik, T., Chudy, R., Micietova, E. Normalized evaluation of the performance, capacity and availability of catalogue services: a pilot study based on INfrastruture for SPatial InfoRmation in Europe. International Journal of Digital Earth 9, 325-341 (2016). doi: 10.1080/17538947.2015.1019581.

[26] Senkler, K., Voges, U. (ed). 2007. *OpenGIS Catalogue Service Specification – ISO Metadata Application Profile* [online]. Open Geospatial Consortium. Cited 18 July 2018. Available at: <http://portal.opengeospatial.org/files/?artifact_id=20555>.

[27] Technical Guidance for the implementation of INSPIRE dataset and service metadata based on ISO/TS 19139:2007 [online]. 2017. Cited 19 June 2018. Available at: <http://inspire.ec.europa.eu/file/1705/download?token=iSTwpRWd>.

[28] Vretanos, P. (ed). 2010. *OpenGIS Filter Encoding 2.0 Encoding Standard* [online]. Open Geospatial Consortium. Cited 18 July 2018. Available at: <http://portal.opengeospatial.org/files/?artifact_id=39968>.

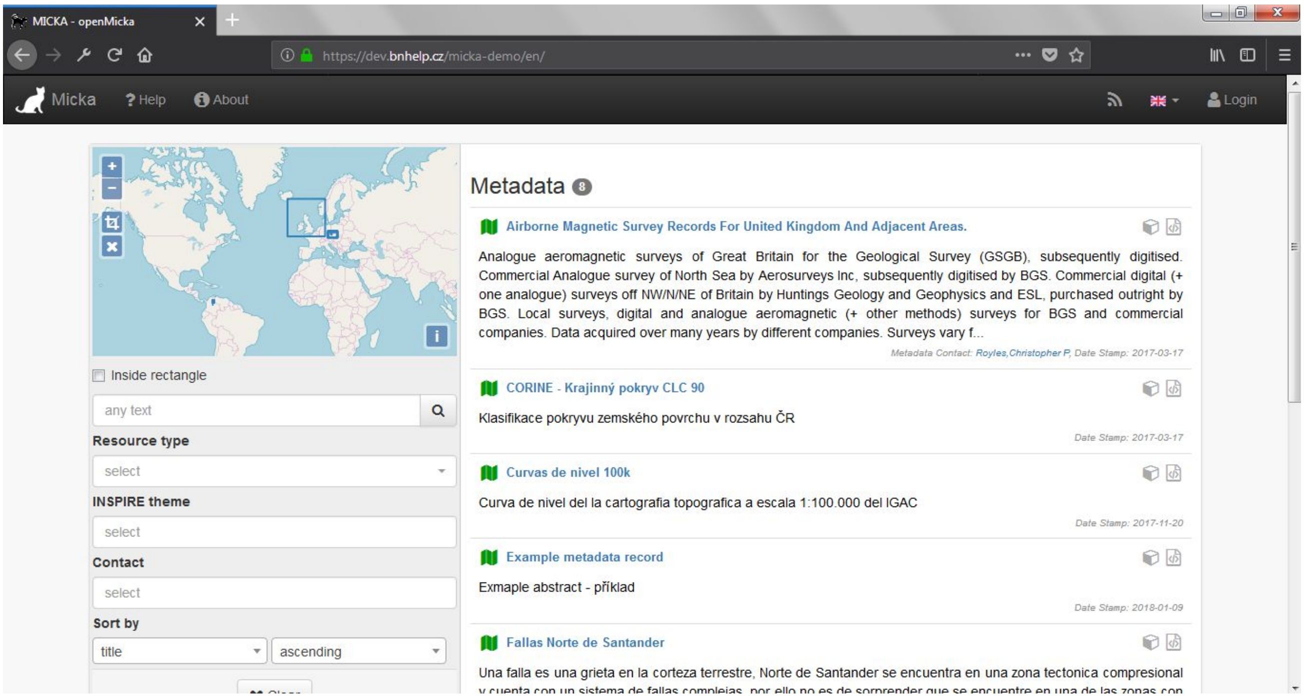
[29] W3C. 2012. *OWL - Semantic Web Standards* [online]. Cited 12 June 2018. Available at: <https://www.w3.org/OWL/>.

[30] W3C. 2014. *Resource Description Framework (RDF)* [online]. Cited 12 June 2018. Available at: <https://www.w3.org/RDF/>.

Annex A – Guide for populating PoliVisu Catalogue

1. Check, whether your organisation/department/colleague/… already has a catalogue. In such a case contact PoliVisu metadata administrator, Tomáš Řezník. Hopefully all the existing metadata could be migrated from your metainformation system into the PoliVisu catalogue.
   In case that you are not operating your own metainformation system, please proceed to point 2.

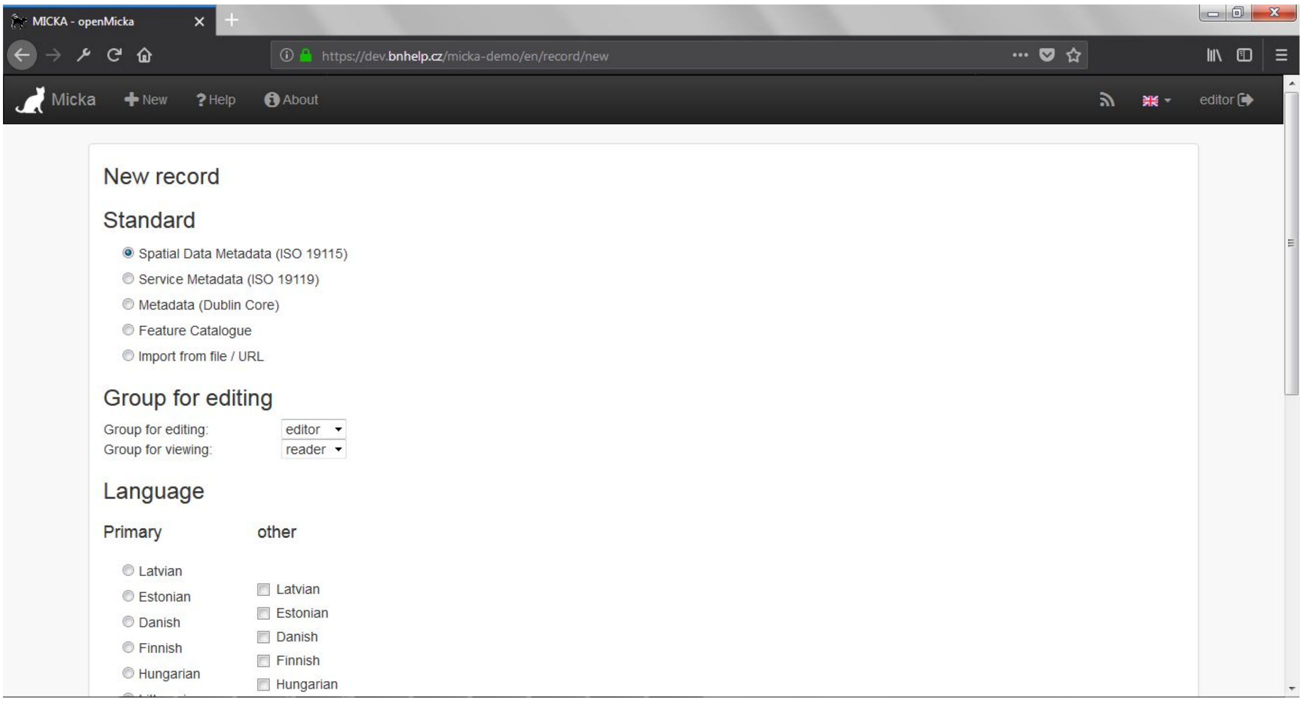
2. Insert into a Web browser an address for the PoliVisu Catalogue https://dev.bnhelp.cz/micka-demo/en/. Your screen should now contain the initial screen of the (Open)Micka metainformation application as in the figure below.



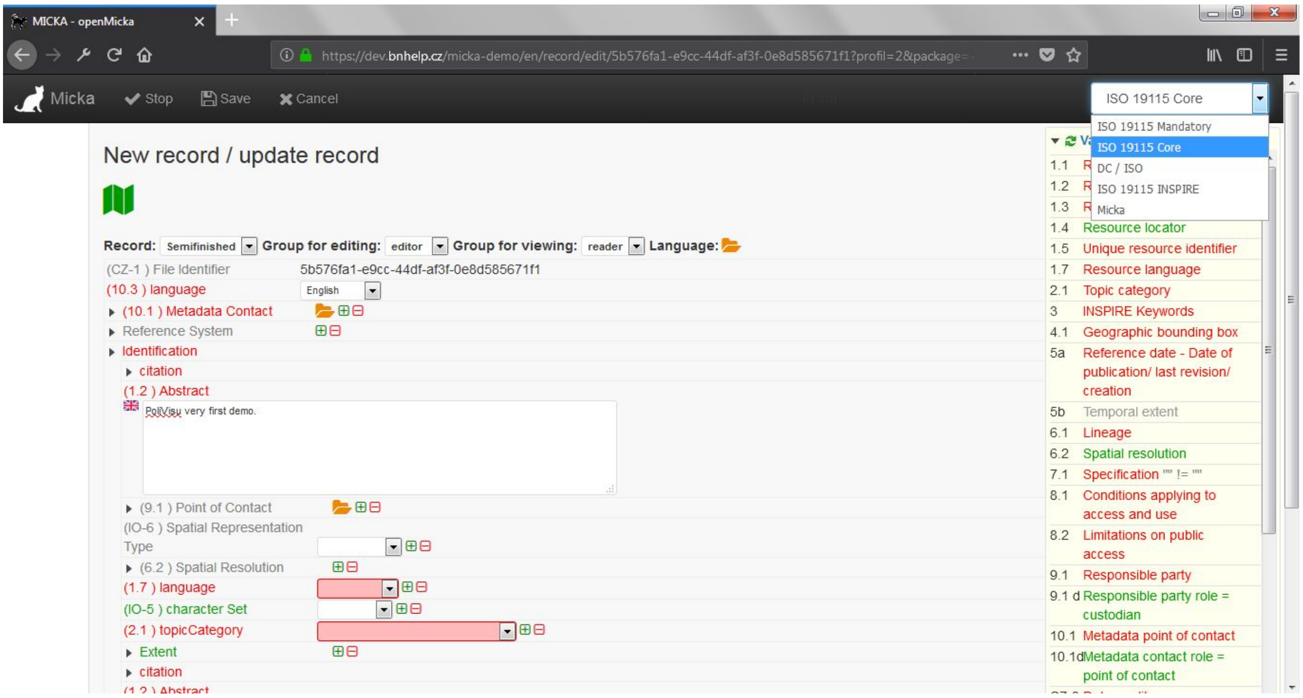
3. Log into the application by clicking on the  icon on top right with the credentials provided for each interested PoliVisu project partner.

4. Click on the  icon that appeared on top left menu after logging into the application. This icon will start inserting a new metadata record(s).

5. Check that the default option "Spatial Data Metadata (ISO 19115)" has been selected. This enables you to follow the metadata structure as agreed within the PoliVisu consortium. You can select the user(s) that may edit/view your metadata. Also languages for metadata management could be modified. Please select one out of 24 official EU languages and between zero and 23 additional languages in case that your metadata are intended to be multilingual. You may just choose your default language without adding any additional languages. After finishing your settings, press the Save icon at the very bottom of the Web page.



6. Select ISO 19115 Core in the top right selection as depicted in the figure below. You can then start populating the metadata elements by your information. Metadata elements highlighted with red colour are the ones that are mandatory. Metadata elements highlighted with green colour are the ones that are conditional. Metadata elements highlighted with grey colour are the ones that are optional. The list on the right side shows you the remaining/inconsistently filled in metadata elements.



7. After adding all the relevant (meta)information, click on  the icon. Your metadata record is available in the PoliVisu catalogue. Congratulations!