Semantic metadata for spatial planning documents

Iwona Kaczmarek
Wroclaw University of Environmental and Life Sciences, Department of Spatial Economy
C.K. Norwida 25
Wrocław, Poland
iwona.kaczmarek@up.wroc.pl

Jaromar Łukowicz
Wroclaw University of Environmental and Life Sciences, Institute of Geodesy and Geoinformatics
C.K. Norwida 25
Wrocław, Poland
jlukowicz@gmail.com

Adam Iwaniak
Wroclaw University of Environmental and Life Sciences, Institute of Geodesy and Geoinformatics
C.K. Norwida 25
Wrocław, Poland
adam.iwaniak@up.wroc.pl

Marek Strzelecki
Wroclaw University of Environmental and Life Sciences, Institute of Geodesy and Geoinformatics
C.K. Norwida 25
Wrocław, Poland
marek.strzelecki@gmail.com

Abstract

The basic regulatory instrument for land management are spatial plans.. The wide range of issues covered and the complex structure of these documents creates a barrier for creating a universal data model for use in spatial planning. It is difficult to reconcile traditional notation used in legal regulations with modern standards for data storage, which allow searching and analysis of the data. The authors' goal is to use Semantic Web technologies to represent spatial planning issues. The aim is to meet the legal requirements for legal act amendment, while circumventing current restrictions associated with the lack of standards for spatial planning systems. The authors analyse a case which enriches spatial planning documents, provided in the form of a simple metadata containing XHTML resource. They define the type of content and allow to identify regulations specifying land use and parameters for development. The solutions proposed by the authors give more opportunities for the sharing of spatial planning documents than is done at present. This is important not only for officials, but also for investors as well as ordinary citizens.

Keywords: spatial planning, semantic web, metadata, RDFa

1 Introduction and motivation

The basic regulatory instrument for creating spatial order are spatial plans. Each plan includes two integrated parts: a graphical component in the form of a drawing and a textual part in the form of a municipal council resolution. There exist some relations between spatial features placed on drawings and findings described in the text, which need to be articulated in a clear way.

The nature of the normative act imposes historically developed principles for the creation of resolutions. It dictates a particular structure of the textual part of the plan such as the document hierarchy or the separation of general and specific decisions.

The wide range of issues covered and the complex structure of these documents presents another barrier for creating a universal data model for use in spatial planning. It is difficult to reconcile the traditional notation used for legal regulations with modern standards of data storage, which allow searching and analysis of the data. Another problem is in creating data-sharing policies that are compliant with regulations and will meet the needs of various user groups that will use this data for different purposes and with different tools [1,6].

In Poland, the XML-based digital format for exchanging legal documents is focused on the presentation and structure of legislative documents. However, it does not include the possibility of using metadata to describe content or to relate to external resources such as geometric data of spatial features specified by local plan legislation. If a legal act contains a graphical attachment (in practice usually TIFF or PDF files), it is included as a reference to a file.

The authors' goal is to use Semantic Web technologies (RDF graph model, ontologies based on description logic expressed in the OWL language, linked data) to represent spatial planning issues. The aim is to meet requirements of legal act amendment, while circumventing current restrictions associated with the lack of standards for spatial planning system.

There are different approaches to tackle this problem which do not necessarily rule each other out. However, formalizing legal act notation creates new possibilities for spatial planning resource utilization by various users in accordance with their specific needs. In addition Semantic Web technologies can enrich the usability of spatial planning resources with the possibility of searching as well as performing analysis and inference. Three possibilities for the use of Semantic Web technologies for sharing spatial planning resources can be considered:

1. The first case assumes that documents available in their current form, do not use any standard. They can be indexed in RDF repositories, together with metadata describing their content. Ontologies are used to provide relations between those documents [1,5].
2. In the second case, the textual part of spatial planning documents is available as XML, and the graphical part is available as a GML file. They can be integrated with the use of RDF graphs containing relations with parts of a legal act and spatial features available on the basic layer of the planned land use. Logical structure and semantic description of spatial planning features can be expressed according to the available ontologies. They can be created by various institutions and users, such as investors, and can reflect the legal status and their individual needs [6].

3. In the third case, ontologies can become tools that support the creation and publishing of spatial planning documents. This phase would involve the enrichment of available electronic documents (text documents specifically provided in HTML / XHTML, XML, ODF, PDF or SVG graphics, KML, GML, and others) for metadata defined in vocabularies in RDF. The aim is to enable new information access methods and to extend the existing one. The semantic annotations are used, which in general, are about assigning the entities in the text links to their semantic descriptions [3]. It is possible to use GRDDL, Microformats or RDFa standard techniques. Tags related to parts of the text and phrases can provide information to which planning categories the specified term, setting, size, or parameter belongs. They will therefore constitute the metadata describing the resource on the level of individual objects and attributes. This solution does not interfere directly with the structure of the document, because the metadata creates a new semantic layer [4]. It can provide RDF graphs that can be extracted from the source document by means of XSLT (the GRDDL standard). This will provide additional, machine readable information, describing spatial resources for discovery and aggregation of information on the Internet [7].

2 Description of experiments

In the article the authors carry out an analysis of the third case, in which spatial planning documents are enriched with the additional metadata, provided in the form of a simple XHTML resource. They define the type of content and allow to identify regulations specifying land use and parameters for development.

Spatial planning documents are accessible in XML (the standard for describing legal acts). They are provided by the WFS service, in the form of XHTML documents (the textual part) and SVG, KML and GML files (the graphical part).

The data encoded in XHTML, which is human-readable, is described by metadata for automatic computer processing using the RDFa standard. The RDFa attributes included in the XHTML code allow input of information about:

- the identification of the resource (object) which is the subject of regulations (the 'about' attribute)
- the properties of the object specified in the text (the 'property' attribute which contains the IDs of the properties compatible with RDF Schema vocabulary, sometimes associated with the attribute 'content', similar to the DataProperty element in OWL)
- the property defined as a reference to other resources, such as types of planned land use defined in the dictionary (the 'rel' attribute, containing the identifier of properties compatible with RDF Schema vocabulary, associated with the 'resource' attribute, which is equivalent to the OWL ObjectProperty).

The values of the attributes refer to the metadata defined in the appropriate dictionary (in the form of an RDF Schema). For example, the 'property' or 'rel' attributes are defined in the dictionary as a rdf: Property, the 'resource' attribute indicates a particular resource which indicates the property assigned to the attribute, the 'rel' and 'content' attributes contain a literal, which is the value of the property assigned to the 'property' attribute.

In order not to lose information, RDFa attributes retain an editorial description of the elements (paragraphs, sections, indents). The use of this technology allows users to introduce information about regulations, for instance regarding future land use or development, into a spatial plan.

An HTML block (e.g.) containing regulations for the area has an attribute 'about', defining the featured resource: a specific planning area. It performs a similar role as an individual in OWL ontologies. All the properties appearing within this block refer to the resource. Particular properties are defined in inline elements, such as , in property = "master plan: zoneNumber" content = "01" or rel = "landUseDesignation" resource = http://www.semgis.st7.eu/ rdf / mpzp_voc.rdf # landUseDesignation.

All the features of the objects presented in the document can thus be described in a formal way. For this purpose, domain ontology for spatial planning is created. It provides a generic set of concepts and relations in the field of spatial planning. There are many methodologies for building ontologies. Inorder to achieve semantic interoperability, the authors suggest that one common ontology for annotating spatial plans should be used. In the case of multiple ontologies, ontology matching techniques should be used to solve semantic heterogeneity [8].

Using the mechanisms of generating an XML document and adding information about geometry, for example from GML or a relational database with geometry, one can enrich the document by:

1. representing geometry and topological relations (e.g. between areas in a spatial plan) using geo-ontologies such as WGS84,NeoGeo or GeoSPARQL
2. georeferencing to other resources, which are collected inSDI (obj_id in GML / WFS) [2]:
   a. spatial objects of the presented spatial development plan, but accessible from another web service,
   b. spatial objects from separate resources (e.g. forms of natural protection, flood risk areas)

Additionally an application which creates XHTML/RDFa can add links to resources that describe the location, e.g. coming from Geonames ontology or the NTS system. They can also reference facts from DBpedia.

Having vocabulary in the form of a suitably labeled RDF Schema and XHTML document, one can extract an RDF graph of this document, which can be explored using SPARQL.

3 Conclusions

The concerned work has practical implications. The solutions proposed by the authors allow more opportunities for the sharing of spatial planning documents than is done at present and allow users to access the knowledge in multiple ways. By adding semantic annotations, it is possible to provide new services, such as ontology-based searching of large databases of spatial plans. It is an important issue, because the access to information about future land use becomes more and more critical, not only for officials but also investors, as well as ordinary citizens.

4 References

[1] Carla Geovana N Macário and Claudia Bauzer Medeiros. Specification of a framework for semantic annotation of geospatial data on the web. SIGSPATIAL Special. Volume: 1, Issue: 1, Publisher: ACM, Pages: 27-32. ISSN: 19467729, 2009

[2] Kaoru Hiramatsu and Femke Reitsma. GeoReferencing the Semantic Web: ontology based markup of geographically referenced information. Joint EuroSDR/EuroGeographics workshop on Ontologies and Schema Translation Services, 2004

[3] Atanas Kiryakov , Borislav Popov , Damyan Ognyanoff , Dimitar Manov and Kirilov Miroslav Goranov. Semantic annotation, indexing, and retrieval. Journal of Web Semantics, 2004

[4] Patrick Maué, Henry Michels and Marcell Roth. Injecting semantic annotations into (geospatial) Web service descriptions. Semantic Web Journal SWJ, 2010

[5] Patrick Maué. An extensible semantic catalogue for geospatial web services.. Arbeit Volume 3, 2008

[6] Andrei Mihai, Berre Arne, Costa Luis, Duchesne Philippe, Fitzner Daniel, Schade Sven, Steinmetz Nathalie, Tertre Francois and Vasiliu Laurentiu. SWING: A Geospatial Semantic Web Service Environment. Proceedings of AGILE 2008 Workshop Semantic Web meets Geospatial Applications

[7] Robert E. McGrath. Semantic Extensions to Defuddle: Inserting GRDDL into XML. July 28, 2008

[8] Lorenzino Vaccari, Pavel Shvaiko, Juan Pane, Paolo Besana and Maurizio Marchese. An evaluation of ontology matching in geo-service applications. GeoInformatica 16 (1), pp. 31–66, 2012

This research work has been supported through the projects granted by the National Science Centre on the basis of DEC-2012/05/B/H/HS4/04197 and DEC-2011/03/N/HS4/03819 decisions.