AUTOMATYKA • 2009 • Tom 13 • Zeszyt 2

Grzegorz J. Nalepa*, Weronika T. Furmańska*

Review of Semantic Web Technologies for GIS**

1. Introduction

Geographic Information Systems (GIS) [9–12] allow to store, represent and search geographic information. Today, they provide an effective foundation for digital maps, planning applications and localization services. GIS integrate number of technologies, including efficient storage solutions, relational databases that allow for searching geographic data, and advanced visualization components. These systems need to address number of specific problems, including: efficient representation of geographic data allowing for optimal storage, search and presentation; real time visualization; integration with other applications using GIS as data repository, e.g. mapping services.

An important improvement in the development of these systems was the introduction of Web-based GIS (or WebGIS). A flexible Web-based interface allows for an easy access for number of parallel clients. On the client side it requires only a simple web browser (optionally equipped with selected presentation plugins). The server side processing requires important amounts of storage and computing power. Such a Web-based client-server architecture proved to be an effecitve tool for making GIS both popular and useful for ordinary, non-technical users. Number of popular geo-localization services such as Google Maps (local.google.com), Via Michelin (www.viamichelin.com), or polish Zumi (www.zumi.com.pl) use the flexible Web GIS approach. They are based on different technical solutions, e.g. Java applets, Macromedia Flash or AJAX.

The next step in the evolution of GIS is the integration of semantic technologies developed within the W3C Semantic Web initiative [2]. Technologies such as formal metadata descriptions with RDF, and formal ontologies in RDFS and OWL allow for implementing enhanced search and classification features in GIS. Semantic Web technologies provide flexible knowledge representation and processing based on a machine-readable XML format, formally structured with XSD. Using this representation reasoning with rules is possible with the use of SWRL.

* Department of Automatics, AGH University of Science and Technology, Krakow
** This work has been partially performed in the framework of the EU ICT Project INDECT (FP7-218086).

In the paper selected aspects of using Semantic Web technologies for representing GIS metadata are considered. The technologies are to be used in the practical design of the GIS system concerning threats of various nature and aimed at improving safety of citizens in urban environments. The system for registering citizen-provided information is a part of the INDECT FP7 Project [8].

The rest of the paper is organized as follows: In Sect 2 the problem of representing metadata are discussed. In Sect. 3 the main issues of Semantic Web technology are presented. Based on this, in Sect. 4 the technology of Geospatial Semantic Web is discussed. Selected ontologies for GIS are briefly presented in the subsequent Sect. 5. Examples of semantically enriched GIS are descibed in Sect. 6. The paper ends with concluding remarks in the final section.

2. Metadata in GIS

GIS store mainly geographical data related to different aspects of map analysis and representation. In order to annotate it for the purpose of search and analysis, metadata is needed. Metadata plays an important role in building practical applications of GIS. Two main interpretations for metadata in GIS are as follows:

1) data about the GIS data quality, see ISO 19115 (TC 211), and
2) geospatial ontologies, knowledge about geo data semantics.

There are several bodies providing standards for GIS metadata on both local and international level.

The main international standard for low-level GIS metadata if the ISO 19115 standard (see the ISO TC 211, http://www.isotc211.org). It provides information about identification, extent, quality, spatial and temporal schema, spatial reference, and distribution of digital geographic data.

In the US the *Federal Geographic Data Committee* (FGDC, see http://www.fgdc.gov/metadata) is an interagency committee that promotes coordinated development, use, sharing, and dissemination of geospatial data on a national level. It provided the Content Standard for Digital Geospatial Metadata (CSDGM). Currently there is an ongoing effort to align this standard with the ISO regulation. The Clearinghouse Network is a community of distributed data providers who publish collections of metadata that describe their map and data resources, documenting data quality, characteristics, and accessibility.

The US goverment provides a large map and data repository at the GeoData.GOV portal (http://geodata.gov). It serves as a flexible platform for sharing and browsing publically available maps, including the metadata annotations.

An open community resource is the GeoNetwork (http://geonetwork-opensource.org). It is a standards based, Free and Open Source catalog application to manage spatially referenced resources through the web. It may serve as a foundation to build an open GIS environment presenting both geo data and metadata with a web-based interface.

The development of the Semantic Web technologies provided a flexible, Web-based technology for data structurization, interpretation, and semantic annotation. Currently it is the source of the most important technologies used to represent and process metadata.

3. Semantic Web technologies

Semantic Web is a project inspired by a vision presented in [2]. Its main goal is to create a Web that will be "an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation'' [2]. In such an environment intelligent software agents could solve complex searching, reasoning and planning tasks. The idea behind it is to represent the Web content in a form, which is understandable and processable by machines and thus enable for automated reasoning and data integration, as well as for exchanging data among applications, acting on users behalf.

The way to achieve these goals is the development of standards, by means of which one can turn unstructured data into meaningful facts. These facts can be then accumulated into various knowledge bases. The next step is to provide formalisms of knowledge representation and provision of mechanisms, such as knowledge integration and reasoning, to operate on them. The Semantic Web standards are developed in a layered architecture observed in Figure 1.

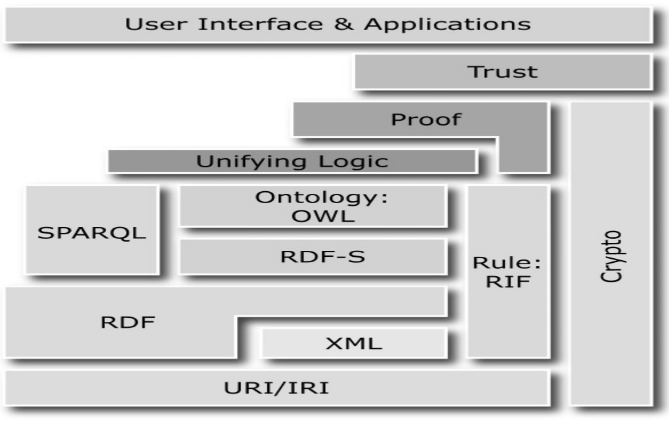

Fig. 1. The Semantic Web Layer Cake [3]

At the bottom of the so-called semantic stack (or layer cake) the identification and encoding of resources is done with use of existing standards such as Unicode and URI. The structure of the Semantic Web documents is provided by XML. Semantic annotations and relations between Web resources are modelled in Resource Description Framework (RDF) and formalized with use of ontologies. Ontologies enable for building taxonomies of the modelled domains and provide reasoning support. Moreover, they allow for knowledge integration, sharing and exchange across application boundaries. Web Ontology Language OWL is a syntactic variant of an expressive Description Logic [1]. Hence, its semantics is well defined and automated reasoning is provided.

Ongoing research on reasoning on the Web includes various integration proposals between ontology languages and rules. Several solutions have been proposed, including RuleML Language, Description Logic Programs (DLP) [6], and Semantic Web Rule Language (SWRL) [7]. Apart from that, a Rule Interchange Format is being developed at W3C. It aims at providing a universal standard for rule export and interchange.

While Semantic Web is mainly concerned with standards, Semantic Technologies include much more, e.g. software agents, systems using metadata annotation, or complex logical procedures etc. The number of available semantic tools has risen steadily for the last decade. There has been a noticeable increase of interest in semantic technologies from big companies over the past few years [herman2008semweb-state]. For the use with GIS, the proposal of the Geospatial Semantic Web technologies has been formulated.

4. Geospatial Semantic Web

Geospatial Semantic Web denotes application of semantic technologies to geospatial data, which include geospatial databases, texts, images, maps, tables etc. A great amount of geospatial data on the Web, their growing popularity and variety of use scenarios made geographic systems a promising area of application of the semantic technologies. A major enabling technology of the Geospatial Semantic Web is the service-oriented architecture. Main semantic technologies include metadata annotation and geospatial ontologies supporting the GIS systems.

In order to understand the Geospatial Semantic Web specifics, certain concepts must be introduced (based on [4]). The basis is *geospatial data*, such as maps, geographic data, digital imaginery of the Earth which is usually more complex than conventional data with respect to both syntax and semantics. Applications that process it are called *geospatial services*. They use *geospatial features* which are features associated with a location, classified into types (e.g. bays, rivers, etc.). Geospatial fields are functions that map points of a domain region into values from a given set which results in various kinds of maps (e.g. a vegetation map). *CRS* (coordinate reference system) defines how geospatial data relates to real locations on Earth. *Gazetteers* define vocabularies consisting of identifiers, location descriptions, and attributes for a set of geospatial features. They allow for locating places on the map easily and building efficient search functionality into applications.

A number of challenges for Geospatial Semantic Web have been identified. Among them recognition and interoperability of geospatial information resources is a prime one. There exist various information resources which differ in terms of organization of data – from loosely structured webpages to well-organized spatial databases and web services. Another issue is the development and management of geo-ontologies. This includes translation of existing taxonomies into Semantic Web standards, and agreement between ontologies from various domains. In order to match existing concepts in web pages to geo-ontologies, research on innovative methods of building ontologies from maps, images and sketches available on the web is carried on. Finally, due to different categorisations of the relationships between geographical concepts, an integration of multidisciplinary ontologies is required for them to be able to operate efficiently.

5. Selected Ontologies for GIS

Geospatial Ontologies provide formalized means to express geospatial data. They serve as predefined and possibly standarized vocabularies. Early works of W3C aimed at developing a comprehensive *W3C Geospatial Vocabulary*

(http://www.w3.org/2005/Incubator/geo/XGR-geo/).

It should be a basis for *W3C Geospatial Ontologies*

(http://www.w3.org/2005/Incubator/geo/XGR-geo-ont).

An example case study [5] demonstrates how semantic technologies can improve GIS metadata. Selected ontologies from this project are discussed below.

Ordnance Survey is a Great Britain's national mapping agency. Within the agency, the GeoSemantics team has been working on development of a Topographic ontology (see http://www.ordnancesurvey.co.uk/oswebsite/ontology). The process is staged, and the ontology is being built from ontologies of the sub-domains organized as modules. These ontologies include various Domain Ontologies, such as Hydrology, Administrative Geography, or Buildings and Places. Ontology Modules developed so far are Topography, Mereological Relations, Network Relations and Spatial Relations.The Semantic Web for Earth and Environmental Terminology (SWEET) project (http://sweet.jpl.nasa.gov/ontology) provides an upper-level ontology for various Earth sciences. It aims at developing „a common semantic framework for various Earth science initiatives". The SWEET ontologies include several thousand terms and relations between them. Besides providing hierarchy of concepts and geospatial features, they use OWL class restrictions to capture the semantics of the geospatial concepts [4].

Geospatial Semantic Web Research and Trends are traced on a Geospatial Meaning website (http://www.geospatialmeaning.eu/). Among others an european framework programme EU FP6 Cinespace, provides an ontology presented in Figure 2.

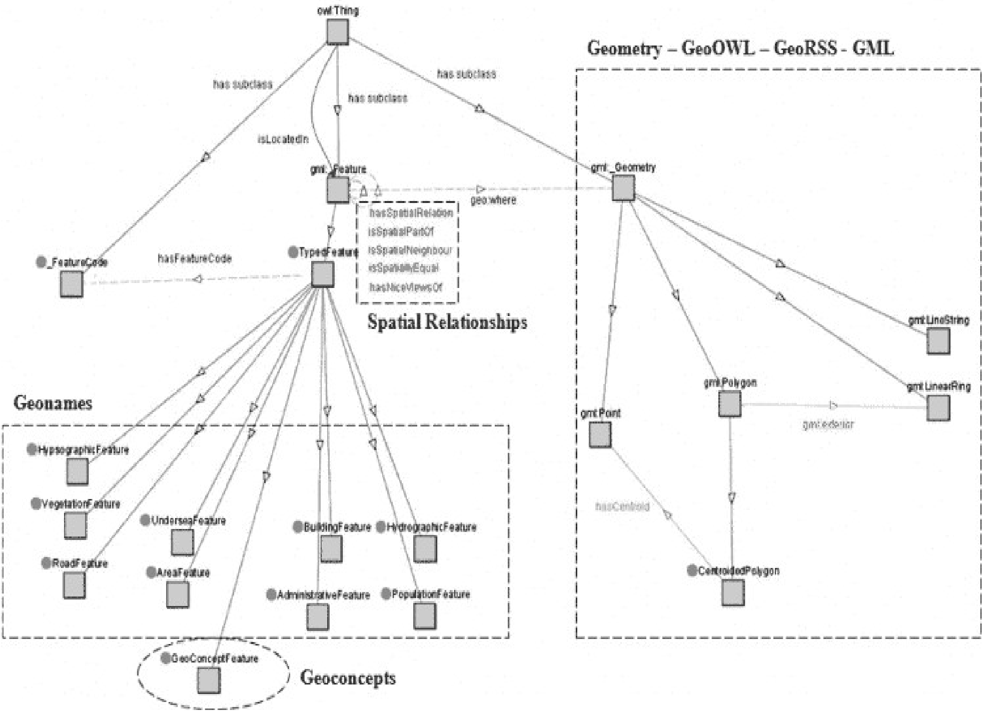

Fig. 2. A skeletonof an ontology developed within the 6th Framework Programme CINeSPACE project

6. Semantically Enriched GIS

The growing popularity and role of GIS systems, including open tools and interfaces posed some new challanges for GIS-based applications. The need to semantically annotate GIS resources becomes apparent. Flexible semantic layer allowed for combining typical map-presenting and route planning systems with features related to facilities localization (hotels), extended descriptions on external websites, etc. In order to gather and organize this data, GIS applications made some collaborative interfaces available. Using them, users can extend the content.

From the perspective of the INDECT project, systems allowing citizens to work with authorities to build an up-to-date threat database, thus allowing for an improvement of security, is essential. Such a system should allow the authorities to present to citizens the information they have in a transparent, accessible fashion. At the same time, citizens should be able to provide feedback, by extending, modyfying the information contained in the system. This kind of a distributed, collaborative environment, possibly linked to other data sources, is not trivial to be built.

Attempts to present such an security-related information have already been undertaken. Several systems providing some kind of GIS-based information with semantically enhanced interaface are available, mainly in the US. Examples of such systems include: http://www.crimemapping.com, http://www.crimereports.com, http://www.spotcrime.com. However, in most of the cases they have a unidirectional interface, aimed at information presentation and semantic search, with no feedback available. An example of a session with such a system is given in Figure 3.

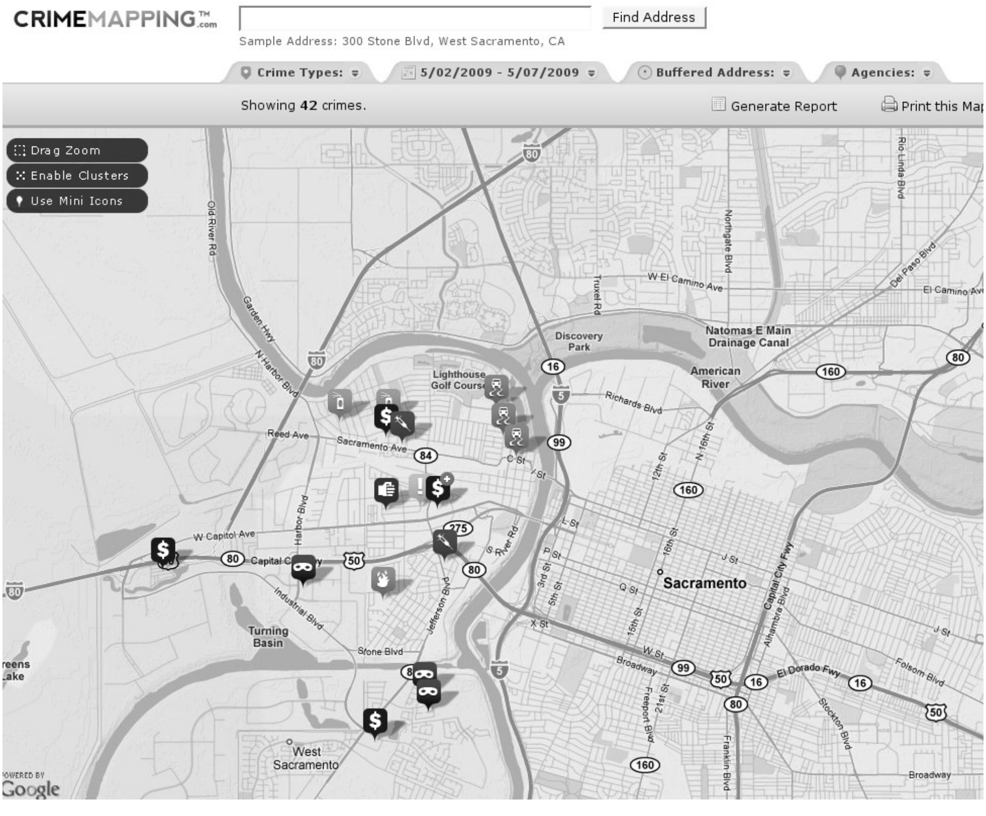

Fig. 3. An example of threaths information from http://www.crimemapping.com

7. Concluding Remarks

In the paper selected issues related to extending GIS systems with semantical information are discussed. Semantic Web technologies play an important role in flexible metadata representation. The Geospatial Semantic Web technology is an important step forward in the construction of semantically enhanced GIS.

Semantic annotations and specialized ontologies have been developed and applied in number of projects and solutions. Some of them are shortly introduced in this paper. Geospatial information and reasoning has a great potential and may introduce a significant improvements in numerous applications, developed for various purposes and different end-users.

The motivation for this research is the development of a semantically enriched environment with a GIS component for citizens and the police. The system proposed within the FP7 INDECT project will be a community portal that allows citizens to participate and cooperate in order to improve the security in the urban environment. Using semantic solutions the system will provide a coherent technology for threats taxonomies integration, as well as knowledge management, integration and automated reasoning.

References

[1] Baader F., Calvanese D., McGuinness D.L., Nardi D., Patel-Schneider P.F., *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, 2003.
[2] Berners-Lee T., Hendler J., Lassila O., *The Semantic Web.* Scientific American, May 2001.
[3] Bratt S., *Semantic Web, and Other Technologies to Watch*. World Wide Web Consortium, January 2007.
[4] Breitman K.K., Casanova M.A., Truszkowski W., *Semantic Web: Concepts, Technologies and Applications*. Springer-Verlag, 2007.
[5] Dolbear C., *Case Study: Semantic Web Technology at Ordnance Survey.* Ordonance Survey, http: /www.w3.org/2001/sw/sweo/public/UseCases/OrdSurvey, 2007.
[6] Grosof B.N., Horrocks I., Volz R., Decker S., *Description logic programs: combining logic programs with description logic*. Proc. of the Twelfth International World Wide Web Conference, WWW2003.
[7] Horrocks I., Patel-Schneider P.F., Boley H., Tabet S., Grosof B., Dean M., *SWRL: A Semantic Web Rule Language Combining OWL and RuleML*. W3C Member Submission 21 May 2004.
[8] Intelligent information system supporting observation, searching and detection for security of citizens in urban environment. FP7, 2009–2013, http://www.indect-project.eu/.
[9] Litwin L., Myrda G., *Systemy Informacji Geograficznej. Zarządzanie danymi przestrzennymi w GIS.* Helion 2005.
[10] Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W., *Geographic Information Systems and Science.* John Wiley and Sons, Ltd. 2005.
[11] Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W., *Geographical Information Systems. Principles, Techniques, Applications, and Management*. John Wiley and Sons, Inc. 1999.
[12] Wilson J.P., Fotheringham A.S., *The Handbook of Geographic Information Science*. Blackwell Publishing Ltd. 2008.