



HAL
open science

Distributed and Collaborative Knowledge Management Using an Ontology-Based System

Weronika T. Adrian, Antoni Ligeza, Grzegorz J. Nalepa, Krzysztof Kaczor

► **To cite this version:**

Weronika T. Adrian, Antoni Ligeza, Grzegorz J. Nalepa, Krzysztof Kaczor. Distributed and Collaborative Knowledge Management Using an Ontology-Based System. 1st IFIP International Workshop on Artificial Intelligence for Knowledge Management (AI4KM), Aug 2012, Montpellier, France. pp.112-130, 10.1007/978-3-642-54897-0_7. hal-01256588

HAL Id: hal-01256588

<https://hal.science/hal-01256588v1>

Submitted on 15 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Distributed and Collaborative Knowledge Management Using an Ontology-based System

Weronika T. Adrian, Antoni Ligeza, Grzegorz J. Nalepa, Krzysztof Kaczor

AGH University of Science and Technology, Poland, email: {wta,ligeza,gjn,kk}@agh.edu.pl

Abstract. Semantic annotations and formally grounded ontologies constitute flexible yet powerful methods of knowledge representation. Using them in a system allows to perform automated reasoning and can enhance the knowledge management. In the paper, we present a system for collaborative knowledge management, in which an ontology and ontological reasoning is used. The main objective of the application is to provide information for citizens about threats in an urban environment. The system integrates a database and an ontology for storing and inferring desired information. While a terminology of the traffic danger domain is described by the ontology, the location details of traffic conditions are stored in the database. During run-time, the ontology is populated with instances stored in the database and used by a Description Logic reasoner to infer new facts.

1 Introduction

One of the important research fields of Artificial Intelligence (AI) is the area of Knowledge Representation and Reasoning (KR&R) [5]. The Semantic Web [3] initiative is sometimes perceived as the new incarnation of AI, tackling some of its problems and challenges. Although this worldwide project is not aimed at constructing intelligent machines, it has resulted in development of several effective KR&R methods. Representation of knowledge is done on a few levels of abstraction. For single objects, attributes and relations to other objects (resources) are defined, by use of semantic annotations. These attributes and relations are organized into semantic vocabularies for various domains. Classification of objects and classes definition using their interdependencies is done with use of ontologies [6] of different expressiveness and formality level. Stating logical axioms about classes enable automated reasoning and inferring conclusions about single objects. There is an ongoing research on integrating ontologies with higher-level representation of rules. Semantic applications can make use of this multilevel knowledge representation and exhibit semi-intelligent behavior.

Web-based information systems have been widely used to facilitate communication and distribution of information in a rapid and efficient way. Whether through official news portals or social systems like Facebook or Twitter, people inform each other about the events or dangers. Using GIS systems [14] that allow to store, represent and search geographic information, users can add location metadata to the information they provide or get useful data based on their localization (e.g. by the use of a GPS). One of the area which still needs a careful attention from the information systems is the local safety of citizens in the urban environment.

Projects such as Wikipedia has demonstrated that people are willing to cooperate if they find it worthwhile and the system is easy to use. Collaborative knowledge engineering and management can be enhanced by employing intelligent techniques, for example by using formal knowledge representation. However, a system interface must remain simple. An interesting and promising example of such a combination are semantic wikis [9,1], in which teams can collaboratively build and manage formally specified knowledge in an intuitive fashion and with little technical knowledge.

In this paper, we present a system for collaborative knowledge management enhanced with semantic knowledge representation and reasoning. The main objective of the system is to gather knowledge about threats of various sorts within a defined urban area. The system should serve the local community and the police. Our proposed solution combines social software features (commenting, ratings etc.) with a strong underlying logical representation of knowledge entered by users. The application employs AI methods, namely a domain ontology of traffic dangers and conditions, and a Description Logic (DL) [2] reasoner to infer knowledge from facts explicitly present in the system.

The rest of the paper is organized as follows: in Section 2 the motivation for our research is given with references to selected previous works. Section 3 gives an overview of selected existing solutions and related work. In Section 4, basic assumptions for the system are specified. Section 5 gives an overview of the system, including its functionality, architecture, a threat ontology, the integration of an ontology and a database in the system, the reasoning in the system and the user interface. The implementation is briefly discussed in Section 6. Evaluation of the approach is sketched in Section 7. The paper is summarized in Section 8 and future work is outlined in Section 9.

2 Motivation

Within the INDECT project ¹ several problems related to security and intelligent information systems are investigated. Task 4.6 of the project focuses on development of a *Web System for citizen provided information, automatic knowledge extraction, knowledge management and GIS integration* [8]. The main objective of our research is to develop a semantically enriched environment for collaborative knowledge management. Local communities should be able to quickly share information about current traffic dangers and threats, for instance closed roads, holes in the pavements and streets, dangerous districts or events that impede a regular traffic. The system proposed within the task should be a sort of a community portal that allows citizens to participate and cooperate in order to improve the security in the urban environment. Within the task several initial system prototypes have been developed [7,13] and the current work consists in integrating the best solutions into the final system.

The system should use some sort of intelligent processing to provide possibly most useful knowledge to the users. To this end, a Knowledge-Based System (KBS) should be proposed, with a formalized knowledge representation and reasoning mechanisms. Categorization of threats and possibility of inferring new facts based on the ones entered

¹ See <http://indect-project.eu>.

by users is a desired feature. To enhance the automated knowledge processing of the system, semantic technologies for GIS were analyzed and discussed in [10].

While a threat domain ontology can be the same for various locations, different system installations will vary depending on the locations they work in. Abstract knowledge can and should be shared across applications boundaries to facilitate change management. On the other hand, the system should be robust and easily adaptable to local conditions, so the access to the actual data should be optimized.

The system should encompass social features, such as possibility to comment on, discuss and rate information entered by other users. This way, the users can gain or lose credibility and the community can indirectly control spam information. The user interface (UI) should be intuitive and easy to use, potentially adaptable to various hardware platforms including desktop and mobiles. Encompassing these requirements should provide a useful intelligent system for improving urban safety.

3 Threat Information Systems – an overview of existing solutions

In this section, we give a brief overview of the existing approaches and solutions for threat information systems. Crime Mapping systems were originally a class of systems that map, visualize and analyze crime incident patterns using Geographic Information Systems (GIS). This name has been later extended to incorporate all applications that aid in improving the public safety. This includes natural disasters monitoring systems which are often designed for specific regions and the scope of their functionalities is usually limited to the specific types of disasters that are most common and most dangerous in those regions, systems monitoring threats on the roads and crime monitoring systems. A detailed survey of the existing crime mapping systems is given in [15].

Ubiquity of the Internet allowed the Crime Mapping class systems to be available not only for specialized units, but for everyone with Internet access, thus allowing the improvement of public safety.

Crime Mapping systems can be divided into categories based on the nature of threats they focus on. The INDECT Crime Mapping class system developed at the AGH University does not specialize in any of them. Instead, it attempts to track threats in all of those categories.

3.1 Categorization of Crime Mapping systems

Crime Mapping is a very extensive class of systems. To facilitate analysis and comparison of its representatives, several factors can be separated that the systems can be categorized by. The following categorization and the list of systems are not exhaustive, but should give an overview of the state of the art in this area.

Types of threats This seems to be the most important criterion, since the system is chosen by user based on kind of threats they are interested in monitoring or learning about. Crime Mapping systems focus on:

- Natural disasters: Earthquakes, Fires, Floods, Dry weather, Storms (e.g. thunderstorms, hurricanes, etc.), Tsunami, Tornados, Volcanic activity.

- Threats caused by a man: 1) Crimes: numerous types of criminal activity, e.g. robberies, assaults, car or house break-ins, sex crimes, etc., 2) Industry level: larger scale disasters caused by technology, e.g. chemical spills, toxic fumes, radiation, etc., 3) Infrastructure-type (e.g. old buildings, slippery stairs)
- Transportation risks: 1) Hazards: various places or situations that may cause a risk to the driver, e.g. potholes, sharp curves, road bumps, etc., 2) Disruptions: not as much dangerous places, but more an inconvenience for people travelling, e.g. road works, traffic jams, narrowings or speed limits.

Geographical range Crime Mapping systems vary in area they cover. This also is an extremely important factor, as user will be looking for a system that covers the area of their interest. They are divided into following categories:

- Global: systems spanning the whole planet;
- Local: limited only to some area. Depending how the range boundaries are defined, they can be limited: geographically, for example, a system monitoring fires in only one forest or warning about tornados on some specific continent, or socially, depending by social structure, e.g. from single city quarter, through the whole city to an entire country or even several of them.

Availability Crime Mapping systems can be also divided depending on who the system is available to. The significance of this factor is that users need to have means to access the system if they are to use it. Categories are:

- Publicly available: many systems are open for browsing for everyone, without even the necessity of registering or logging in.
- With authorized access: users need to register with the system and log in to learn the information it offers. Registration itself could be open for everyone or somehow limited – for example, a registration fee could be collected.
- Restricted: only people satisfying some defined conditions could be accepted as users, e.g. inhabitants of monitored area, members of an organization or employees of a company.

Data Source Finally, Crime Mapping systems vary in where the data collected and presented comes from. The source of information matters, because it has impact on its quality and credibility. Data can come from:

- Agencies, cooperating with the system, e.g. law enforcement or university;
- Experts, analyzing data and providing the results to the system;
- Automated monitoring, collected by various measuring or observation apparatus (e.g. seismic sensors or cameras), often subjected also to processing or analysis, which itself also can be automated;
- Users – some systems allow users to provide data to the system, or even depend on that greatly.

3.2 Natural Disasters Information Systems

Natural disasters monitoring systems are often designed for specific regions and the scope of their functionalities is usually limited to the specific types of disasters that are most common and most dangerous in those regions. For example, the area of Great Britain has several different systems for warning about floods. Among others, there are: National Flood Hazard Mapping² in Wales, Managing Flood Risk and Flood Warnings³ in Scotland and Risk of Flooding from Rivers and Sea⁴ in London and its vicinity. Similarly, there is a California-Nevada Fault Map⁵ (see Fig. 1) , collecting information about earthquakes in the western states of USA, where those are a frequent occurrence due to the proximity to the San Andreas Fault.

Another system, developed as part of U.S. Geographical Survey's (USGS) Earthquake Hazards Program⁶ is available on the Internet to aid reducing earthquake hazard in the USA. However, their area of observations is not limited to the US only. The scientists behind the program study and keep track of earthquakes all around the world, providing not only local but also global information. This proves that aside from those regional systems for warning about national disasters there also are ones working globally. Other examples of such systems are Seismic Monitor⁷, also gathering information about earthquakes around the world, and Global Flood Map⁸.

3.3 Road Threats Information Systems

Another area where technological advancement is progressing constantly is the transportation. For common citizen especially important from the safety point of view is personal transport – car technology and road infrastructure.

GPS navigation devices become increasingly popular, and they not only serve as tools for finding optimal routes and guide the driver to their target, but often they also take into consideration and keep track of dangerous spots on the road, requiring the driver to be more cautious or avoid it completely. Places where, for example, accidents happen more often due to potholes or poor visibility, are sometimes included in maps the devices use, so that they inform the driver when they are approaching the danger.

Systems like that already exist. They collect information of such statistically risky places, but also keep track of other, temporary difficulties, like road works, narrowings, traffic jams or accidents. More often than in case of natural disasters monitoring systems these cover only some local areas – there is no actual global transport threats and difficulties monitor.

The examples of such systems are: Traffic Information Highways Agency (reporting bad weather, speed reductions, accidents and road works in England)⁹, Travel News

² See: <http://www.floodmaps.ie>

³ See: <http://go.mappoint.net/sepa>

⁴ See: <http://www.environment-agency.gov.uk/homeandleisure/floods>

⁵ See: <http://earthquake.usgs.gov/earthquakes/recenteqscanv>

⁶ See: <http://earthquake.usgs.gov/earthquakes>

⁷ See: <http://www.iris.edu/seismon>

⁸ See: <http://globalfloodmap.org>

⁹ See: <http://www.highways.gov.uk/traffic/traffic.aspx>

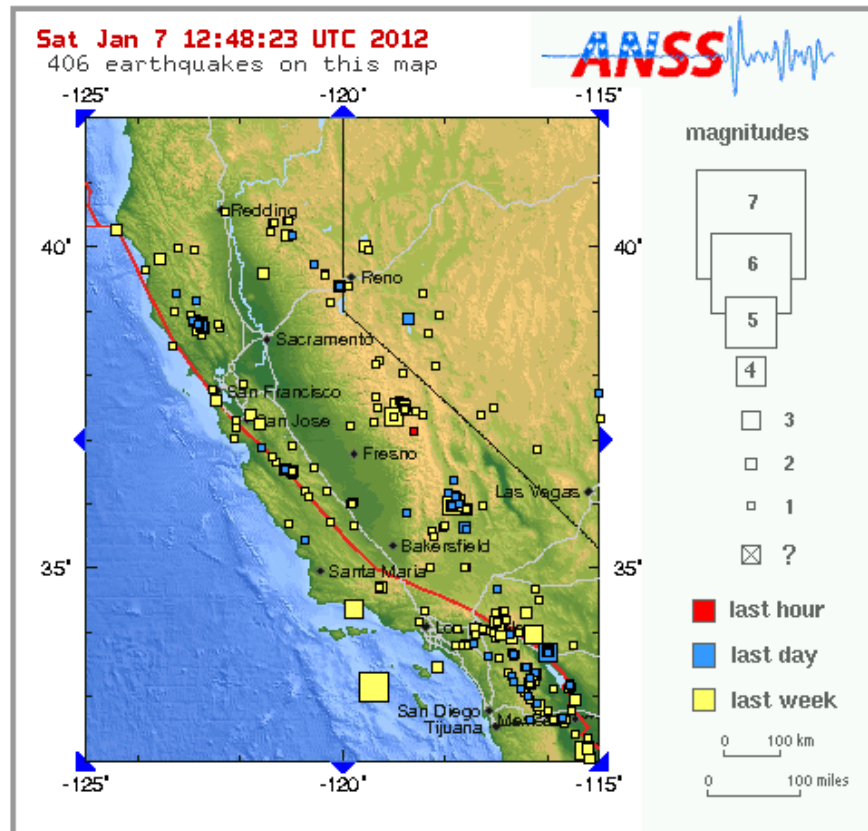


Fig.1. California-Nevada Fault Map. Source: <http://earthquake.usgs.gov/earthquakes/recenteqscanv>.

Scotland (accidents, road works, other difficulties)¹⁰, TravelMidwest.com¹¹ (road capacity, road works, accidents, speed cameras, etc. in Chicago and several other cities and areas around Midwestern states). NAVTEQ¹² is one of the few of such services that attempt to provide traffic information from all around the world, yet it only provides information about capacity and delays, not about threats or accidents.

Similar Polish service with traffic information, "Serwis dla kierowców" (Service for Drivers)¹³, available on the website of General Directorate of National Roads and Motorways (GDDKiA), bears comparison with Traffic England. Although the interactive map is a little more difficult to use (mouse roller does not work as zoom in and out,

¹⁰ See: <http://www.bbc.co.uk/travelnews/scotland>

¹¹ See: <http://www.travelmidwest.com>

¹² See: <http://www.traffic.com>

¹³ See: <http://www.gddkia.gov.pl/pl/10/serwis-dla-kierowcow>

drag and drop map moving sometimes fails), the road network this tool covers is much more dense and is not limited only to motorways. Moreover, user can switch between clear road map and information about road capacity (see Fig. 2).

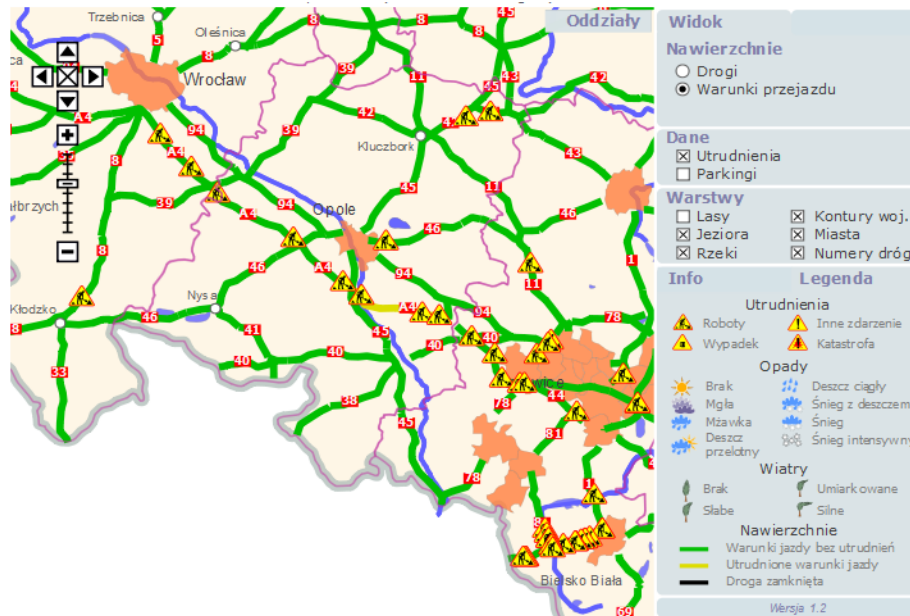


Fig. 2. Service for Drivers. Source: <http://www.gddkia.gov.pl/pl/10/serwis-dla-kierowcow>.

3.4 Crime Mapping systems

The name of Crime Mapping system class comes from the original idea for publicly available systems providing information about crimes and threats in urban area. The possibilities for what such systems could incorporate and the vast amount of functionalities they could provide are growing along with the development of the Internet. Crime Mapping systems gain much in credibility and accuracy of presented data when they cooperate with law agencies and police departments. Since this way the agencies aid in increasing public safety, profits from such cooperation are mutual.

Because law differs between parts of the world, such systems are mostly local, varying in range from one police unit (e.g. Crime Tracker¹⁴ for Jefferson Parish, Louisiana, USA), through a province or a county (Cornwall Crime Explorer¹⁵ or Arizona Crime

¹⁴ See: <http://crimestats.jpso.com/crimetracker/externalmanager/index.html>

¹⁵ See: https://www.amethyst.gov.uk/crime_atlas/atlas.html

Search¹⁶) to a whole country (Crime Reports¹⁷ or Crime Mapping¹⁸, both of which span over USA and Canada). Sadly, there are no global Crime Mapping systems.

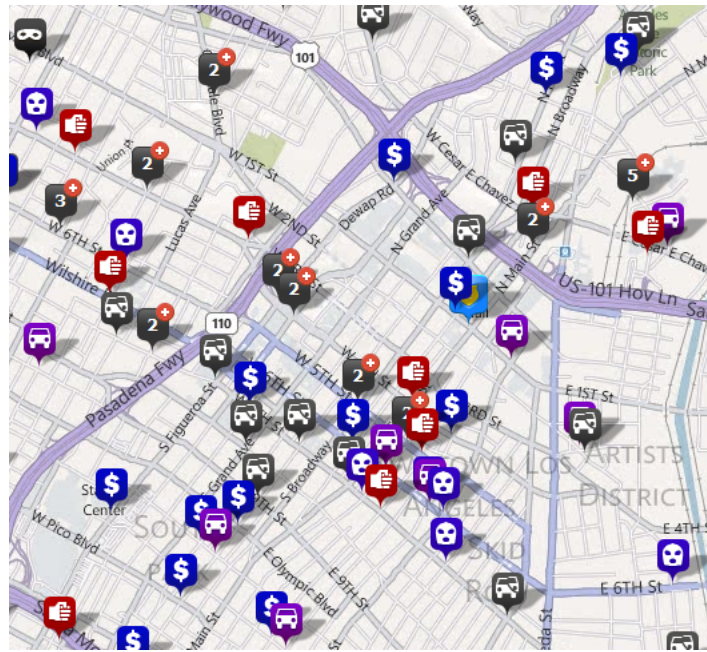


Fig. 3. Interactive map with crime markers. Source: <http://www.crimemapping.com>.

The systems present the threats in a visual form on a map and provide various output channels, e.g. e-mail notifications, text messages or even Twitter alerts. They operate on mobile devices and make use of their GPS systems. The apparent lack, however, is that the information presented to the users is strictly that which was entered. The original contribution of our approach is to supply the system with intelligent processing techniques based on ontological reasoning. Moreover, our approach aims at encompassing various kinds of threats by using a threat ontology.

4 A Conceptual Model for the System

The main focus of our research was on tools to process the information provided by citizens via a specialized website. In fact, a Web System software for citizen provided information, automatic knowledge extraction, knowledge management and GIS integration is to be developed. This task is intended to complement other work oriented

¹⁶ See: <http://www.azcentral.com/CrimeMaps>

¹⁷ See: <https://www.crimereports.com>

¹⁸ See: <http://www.crimemapping.com>

towards automated information extraction from existing web resources by building an Internet-based, distributed information acquisition and automated knowledge management system.

Such a system should combine a CMS (Content Management System) system with a KMS (Knowledge Management System) incorporating intelligent information processing tools based on knowledge engineering. It should allow storing and retrieving of partially analyzed investigations. For spatial information processing and information presentation a GIS (Geographical Information Systems) technology should be used and logic-based systems technology for automated inference should be incorporated.

The main goal of the system is to serve as a distributed knowledge acquisition system for data, information and knowledge provided by citizens, as well as to enable limited automated knowledge management. In principle, the working scenario for the systems is as follows. The system offers a web interface (based on thin-client technologies; in practice a standard web viewer). The interface offers various functionalities for different types of users. The main functionality refers to enabling definition of a new threat. The provided knowledge is then checked for syntactic correctness and processed in an automatic way. The ultimate result is stored in the internal knowledge base. The knowledge base should enable automatic knowledge extraction and processing in order to maximize the effectiveness of the system.

5 System Overview

The proposed system is an ontology-driven application. It integrates a database and an ontology for storing and inferring knowledge about traffic dangers in a given area. While the abstract of traffic danger domain is described by the ontology, the location details of traffic conditions and geographical information (e.g. relations among concrete streets, districts and postal codes) are stored in a relational database. During run-time, the information from the database is integrated (synchronized) with the core ontology (the terminology is populated with instance data stored in a database). The synchronization is done automatically at the application start and at any time on a user demand. The synchronized ontology is then used by a DL reasoning engine to infer facts about chosen area. The deduction is based on definitions of threats which depends on specific traffic conditions present in specific locations.

5.1 Functionality and User Roles

The main objective of the proposed system is to provide citizens with a real-time data about dangers occurring in a chosen area. Part of this information is entered into the system by so-called *trusted users*. The rest is automatically inferred based on the axioms of the threat ontology and instance data of current conditions (facts).

Three kinds of users are distinguished within the system. *Regular users* can browse the system knowledge and ask questions about specific locations and dangers. They address the system with dynamic questions and get results of inferred traffic dangers. *Trusted users* can modify the information stored in the database, e.g. they can update the locations of traffic conditions occurrences. The information is validated and stored

in the database. Updated knowledge can be used for dangers deduction process, after *synchronization* with the ontology. Finally, the *experts* can modify the core ontology.

5.2 Architecture and Data Flow

The system is divided into three functionally different layers:

1. a web dashboard layer dedicated to the interaction with users (through browser clients),
2. a platform layer which is the core of system responsible for processing knowledge, and
3. a storage layer, where all the data is stored, in a database and an ontology (see Figure 4).

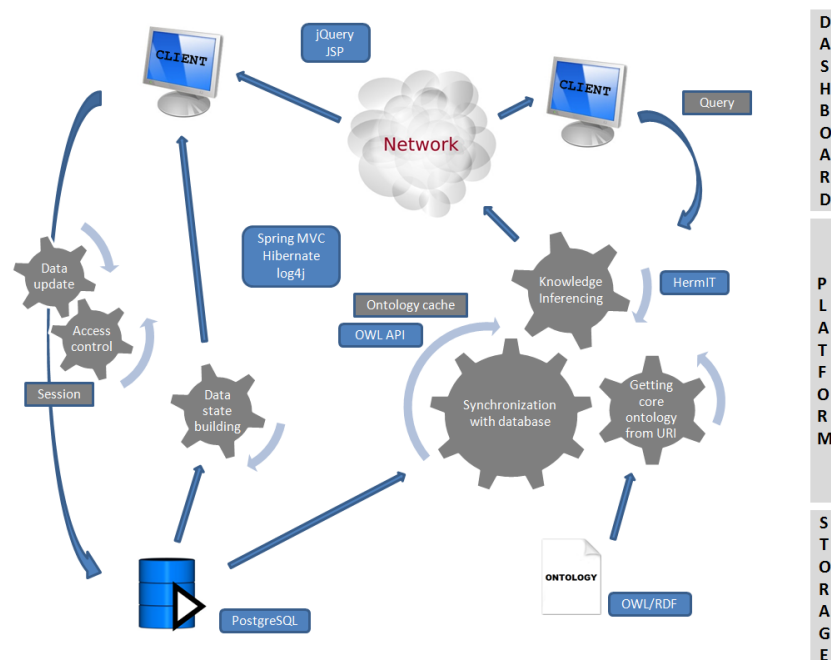


Fig. 4. Data flow in the system

All users can interact with the web-based dashboard for querying system, to get desired information. The main logic of the system (presented in Fig. 4 as three cogged wheels) consists in: downloading the core ontology (*"Getting core ontology from URI"*), synchronizing the core ontology with the current data uploaded by trusted users (*"Synchronization with database"*), and inferring the ontology dependencies (*"Knowledge inferencing"*).

For working with most recent data, provided by trusted users, the *synchronization* mechanism integrates core ontology, describing the abstract of traffic dangers, with specific real time data. The process is executed for the first time on application start, i.e. the first request to the server while accessing the main page of the system. This functionality is also available on demand. After synchronization, the populated ontology is cached in memory and used for inferencing.

A single installation of the system (for instance for a single city) has its own database, in which the information about streets, districts and actual conditions are stored. The *core ontology* on the other hand can be shared by several installations of the system. It is accessible by an URI and can be stored on local or remote server.

5.3 Traffic Danger Ontology

As noted in [11], "In recent years the development of ontologies has been moving from the realm of Artificial-Intelligence laboratories to the desktops of domain experts". Sharing common understanding of the knowledge domain is one of the most critical reason for developing ontologies. Explicit domain assumptions provide a clear description of the domain knowledge and simplify the knowledge extensibility. In our case, the domain ontology consists of concepts of geographical locations (streets, districts, postal codes), traffic dangers (e.g. *LowFrictionDanger*, *RoadConstructionDanger*), traffic conditions (including among others a hierarchy of *WeatherConditions* and *RoadConstructionConditions*) and describes multiple relations among them. An excerpt of the ontology is shown in Figure 5. The ontology enable the system to reason upon stored facts and answer the questions of the following types:

- What traffic dangers can be encountered within a specific area?
- Is there any danger within area of specific postal code or specific district?
- What kind of dangers are connected with specific atmospheric conditions?
- Are there any dangers connected with specific condition (e.g. low friction) in a specific area?
- What are the sub-areas of a specific location?
- Are there any traffic conditions provided for a specific location?

In order to answer these questions either a semantic reasoner is used (which performs classification of concepts within the ontology) or appropriate DL queries are constructed. Based on the definitions of *TrafficDangers* in the ontology and information about actual condition occurrences, implicit knowledge may be deduced (what kind of danger results from given conditions in a selected area).

5.4 Integration of the Ontology and the Database

While the abstract domain knowledge is expressed by the ontology axioms, the operational knowledge of the system is stored in a relational database. The database schema can be observed in Figure 6. The knowledge stored in the database consists of the locations structure and the actual traffic conditions in these locations. Specifically, the locations of the traffic conditions occurrences are defined by postal codes. The postal

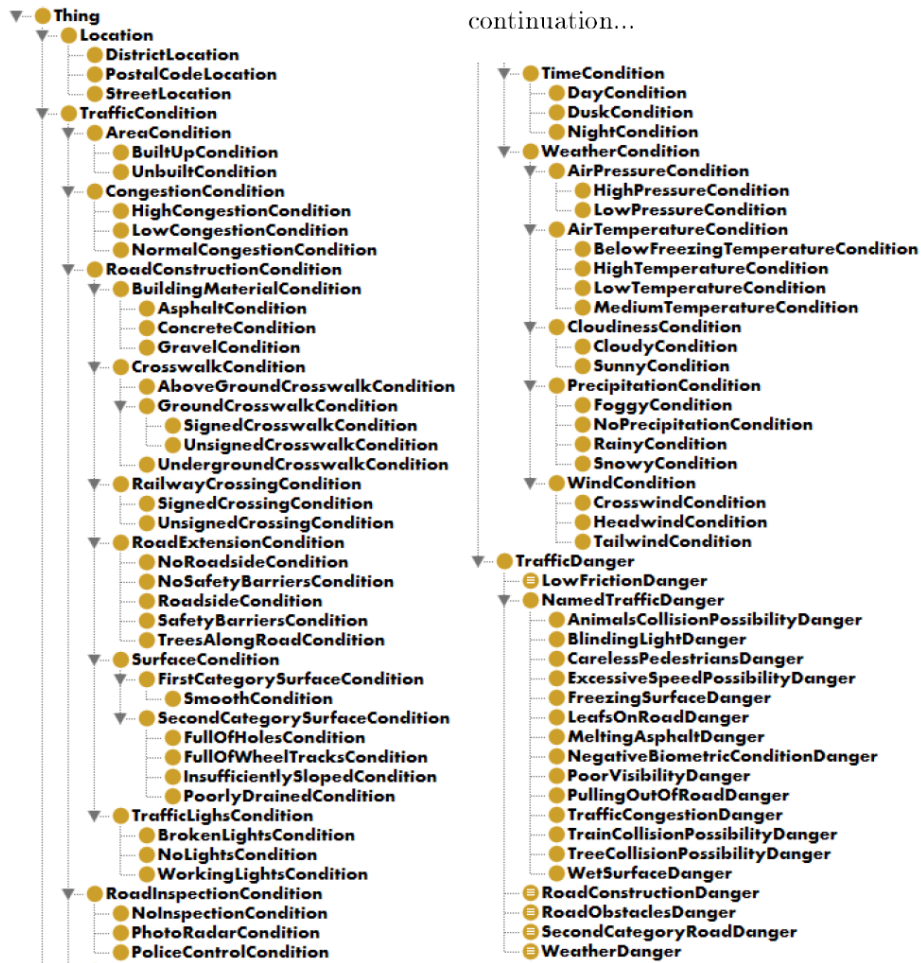


Fig. 5. Traffic Danger Ontology: Asserted class hierarchy

codes are connected to streets, which in turn are connected to districts. For instance, one can add an information that a particular street is under construction (a *RoadConstructionCondition* or one of its subclasses occurs) or that there is a specific weather condition in a specific district.

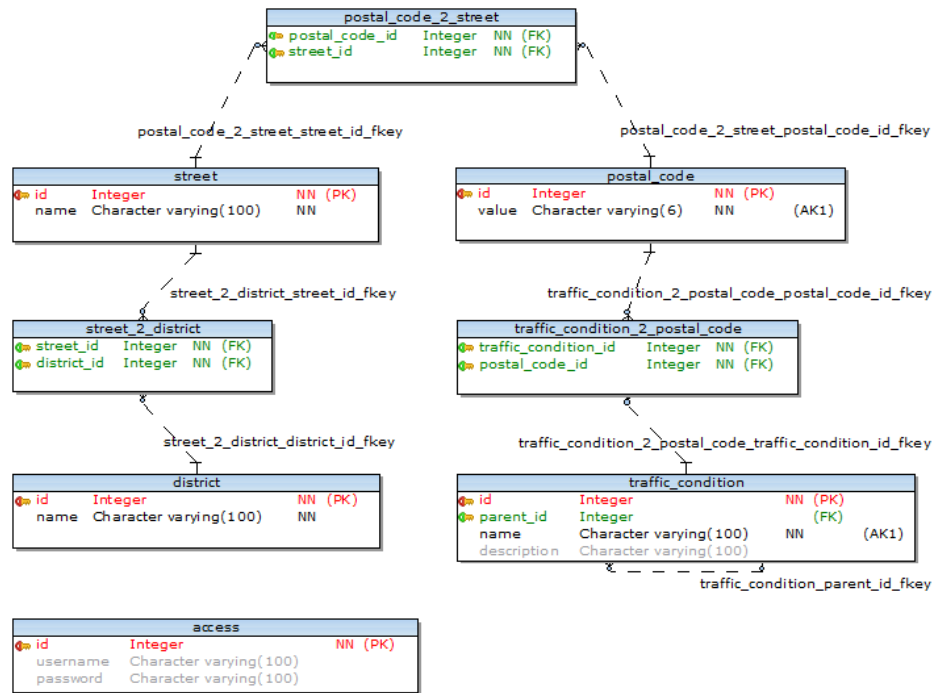


Fig. 6. ER diagram of the traffic database

One of the most important aspects of the system is the possibility of an integration of data from the database and the ontology. Upon the *synchronization* process, the core ontology is cached and populated with the data from the database becoming a *synchronized ontology*. While the core ontology describes a terminology of traffic danger, the synchronized one is related to a specific environment and used for reasoning. Consequently, synchronized ontology can differ between the various environments where the system is deployed. For example, traffic conditions information for Cracow can vary significantly from those in Montpellier. Although it is possible to have a single installation of the system and synchronizing the ontology at once with all global data, it can result in system overloading and decreasing performance while inferring dependencies.

5.5 Reasoning in the System

Reasoning in the system is provided by invoking a DL reasoner on a synchronized ontology. The sequence diagram (see Figure 7) the required steps for the reasoning

process. Once the trusted users have provided traffic conditions facts, a regular user can

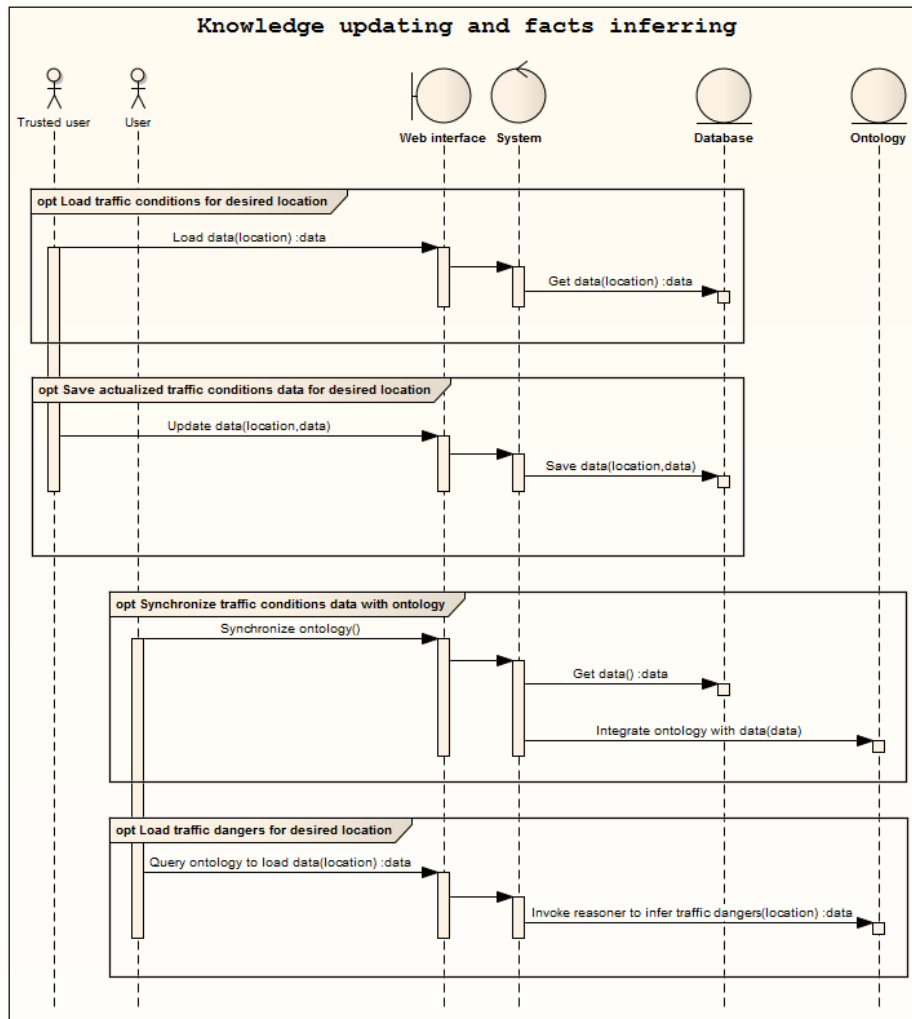


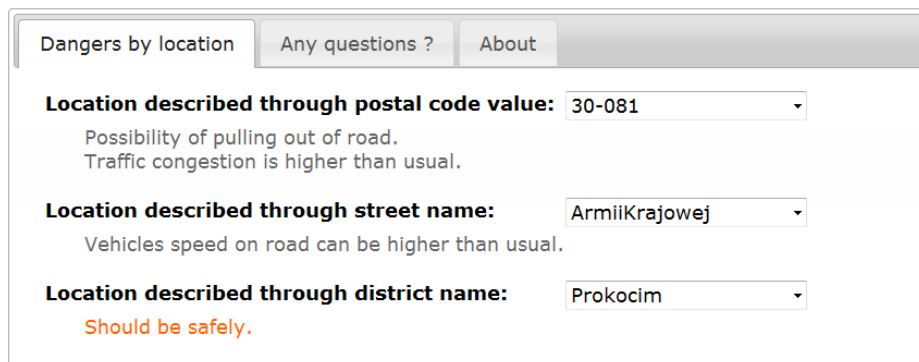
Fig. 7. Sequence diagram for updating and inferring data

check what threat they may expect in a specific area. Responding to the user request, the system imports the up-to-date facts into a locally stored ontology (synchronizes the ontology), and then query the ontology by posing appropriate DL queries. From a user perspective, a query is constructed by selecting a desired location through a web-based interface. Once the query is created, the DL reasoner is invoked to process it on the cached ontology. The inferred set of information is provided to the user. The reasoning

takes relatively little time given the ontology having less than 100 classes. However, no benchmark testing has been done yet and it this is a subject for further work.

5.6 User Interface

The web-based interface of the system allows its users to create dynamic questions, and get the results about inferred traffic dangers. The prototype implementation [12] uses simple forms by use of which the users can construct questions for the system, e.g. a user can choose a desired location from a drop-down list and ask what threats may be encountered in this particular area (see Fig. 8).



The screenshot shows a web-based user interface with three tabs: "Dangers by location", "Any questions?", and "About". The "Dangers by location" tab is active. Below the tabs, there are three location selection options, each with a dropdown menu and associated text:

- Location described through postal code value:** 30-081 (dropdown).
Possibility of pulling out of road.
Traffic congestion is higher than usual.
- Location described through street name:** ArmiiKrajowej (dropdown).
Vehicles speed on road can be higher than usual.
- Location described through district name:** Prokocim (dropdown).
Should be safely.

Fig. 8. An excerpt of web-based user interface of the system

Full development of a Graphical User Interface (GUI) with a map component is in progress. There is a process going on to integrate the logical layer of the system with an interface that uses maps and provides social features for the community of users. A fully-fledged GUI with an interactive map on which the users can navigate and filter threats by location, date, severity etc. has been developed (see Fig. 9), but is not yet fully integrated with the logical layer. In this version, a spatially-enabled database will be used which allows to store geographical data in an efficient way. The usability of the system is expected to increase, due to the possibility of visually choosing an area of interest (see Fig. 10) and the social features like rating threats and discussing them.

6 Implementation and Deployment

The ontology has been developed in a top-down process with the Protégé¹⁹ editor integrated with the HermiT DL Reasoner²⁰. The ontology is provided in different formats

¹⁹ See <http://protege.stanford.edu/>

²⁰ See <http://owlapi.sourceforge.net/>

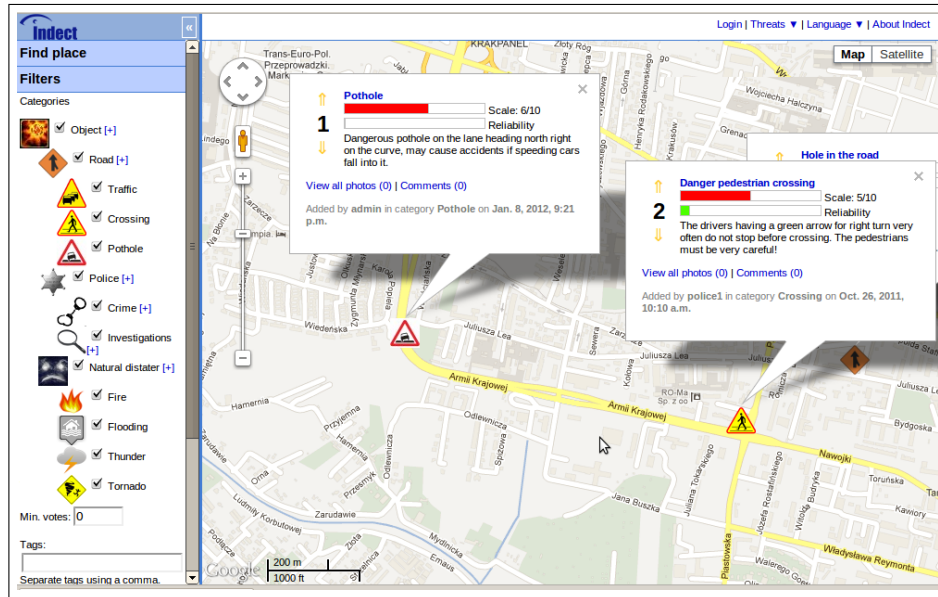


Fig. 9. GUI for the system with an interactive map and search filters.

(OWL2 XML ²¹, RDF/XML ²² or OWL2 Manchester Syntax) ²³). The synchronization is based on the OWL API library ²⁴ and provides up-to-date information (cache in memory) for the HermiT DL Reasoner.

The ontology can be stored on local or remote server and is accessed by an URI. The cooperation with a database is provided through the Hibernate ORM ²⁵ technology. The simple form-based user interface has been built with the JavaServer Pages (JSP) ²⁶ and jQuery JavaScript Library ²⁷, while requests from users and appropriate responses, are controlled by Spring MVC ²⁸. For logging the results of particular operations, log4j Java-based logging utility ²⁹ is used. PostgreSQL ³⁰ is chosen as SQL database. The application has been written in Java using the Eclipse Java IDE ³¹. Dependencies man-

²¹ See <http://www.w3.org/TR/owl2-xml-serialization/>

²² See <http://www.w3.org/TR/rdf-syntax-grammar/>

²³ See <http://www.w3.org/TR/owl2-manchester-syntax/>

²⁴ See <http://owlapi.sourceforge.net/>.

²⁵ See <http://www.hibernate.org/>

²⁶ See <http://www.oracle.com/technetwork/java/javase/jsp/index.html>

²⁷ See <http://jquery.com/>

²⁸ See <http://static.springsource.org/spring/docs/3.0.x/>

²⁹ See <http://logging.apache.org/log4j/>

³⁰ See <http://www.postgresql.org/>

³¹ See <http://www.eclipse.org/>

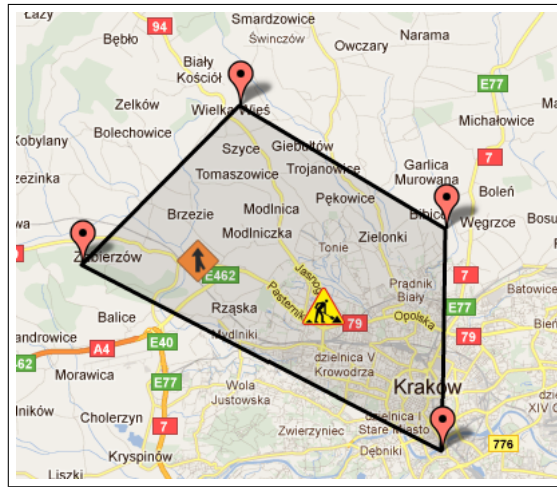


Fig. 10. GUI for the system: a map with selected area.

agement and versioning is the task of Apache Maven tool ³². All these technologies are free software or open source.

7 Evaluation

With respect to the criteria defined in Section 3, the system can be categorized as:

- General-purpose: it is designed to provide information about a wide variety of threats,
- Local: it is specifically aimed at being used in small communities (neighborhood, city, etc.)
- Publicly available (with limited functionality) and with authorized access (for more advanced features),
- Depending on users as main data providers.

The system has several advantages which make it suitable for being used in the aforementioned local environment. The architecture of the system, with loose coupling of the database and ontology, enables using the same core ontology in various installations. Synchronized ontologies (populated with real time data) can differ between various environments. This decentralized way of cooperation, when each client have cached in memory its own synchronized-on-demand ontology instance, is chosen for performance optimization reasons. Comparisons to situation in which single instance of ontology is the centralized part, accessible for synchronization to all clients, provide obvious performance drawbacks. This is why it was considered as an anti-pattern and depreciated by design.

³² See <http://maven.apache.org/>

The web-based, collaborative nature of the system makes it possible for citizens to communicate with each other, discuss and rate threats, and let them – to some extent – control the quality of the information entered into the system. The effectiveness of using the system depends largely on the engagement of the community. The so-called *network effect* is an important factor, because the more people honestly engage in the knowledge management process, the better the quality of information given by the system.

The architecture and design principles, especially the use of an ontology, have also their drawbacks. Because ontologies should be designed to be reusable, it is required to spent relatively lots of effort to provide a good design and tests. Domain experts should be involved in developing these tasks in parallel with programmers, to provide short iterations, as results of quick responses for every inconsistencies in ontology (irrelevant hidden relationships, domain descriptions mistakes or usability problems).

The data in the system is currently acquired from users. However, the system would benefit from incorporating some knowledge from external data sources, e.g., weather forecast services etc.

8 Summary

AI techniques may be successfully used in various applications for Knowledge Management. Using an ontology in a KM system allows to store abstract data, share it across several installations and manage changes in a centralized way. A loose coupling of the ontology with a relational database allows to store concrete data about conceived area in a database and populate the ontology with instance data during application run-time. Embedding a Description Logics reasoner enable the system to reason upon explicit knowledge entered by users and give back a useful response. A graphical user interface with a map component and social software features make the system user friendly and has a gradual learning curve.

To the best of our knowledge, there does not exist a crime mapping system that uses ontologies and DL reasoning to provide rich information based on knowledge gathered in the system. Although there exist numerous solutions for various danger information systems, none of them describe the threats in a formalized ontological way, relate weather or road conditions to the possible dangers and reasons about these dependencies. We believe that this is our original contribution compared to existing work.

9 Future work

The system has been tested with several Web browsers and can be used on any device that support Web browsing. However, for mobile devices, some adaptations are needed. The current prototype implementation has a limited user interface. The intended integration with a GUI providing interactive map and social features is not yet finalized. A possible direction for further development could be focused on extensions for heterogeneous application-to-application communication. The RESTful Web Services³³ can be

³³ See http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.

considered. External systems would be perceived as software agents. Their tasks could be focused on periodic connections to the system, getting some information set, and creating statistics about the traffic dangers. The statistics could visualize frequencies of particular dangers on a specific area or classify the safety of the selected district.

Another important research thread is the *context-awareness* of the system. Using the system should be seamlessly incorporated into the daily routines of its users. Therefore, certain adaptations are considered [4] that will allow installations of the system to work on mobile devices in an adaptive way.

Acknowledgement

The research presented in this paper is carried out within the EU FP7 INDECT Project: "Intelligent information system supporting observation, searching and detection for security of citizens in urban environment" (<http://indect-project.eu>).

References

1. Adrian, W.T., Bobek, S., Nalepa, G.J., Kaczor, K., Kluza, K.: How to reason by HearT in a semantic knowledge-based wiki. In: Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2011. pp. 438–441. Boca Raton, Florida, USA (November 2011), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6103361&tag=1
2. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.): The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press (2003)
3. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American (May 2001), <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&pageNumber=1&catID=2>
4. Bobek, S., Nalepa, G.J., Adrian, W.T.: Mobile context-based framework for monitoring threats in urban environment. In: Multimedia Communications, Services and Security: 6th International Conference, MCSS 2013: Kraków, Poland. June 6-7, 2013. Proceedings (2013), accepted
5. Brachman, R., Levesque, H.: Knowledge Representation and Reasoning. Morgan Kaufmann, 1st edn. (2004)
6. Guarino, N.: Formal ontology and information systems. In: Proceedings of the First International Conference on Formal Ontologies in Information Systems. pp. 3–15 (1998)
7. Ligęza, A., Adrian, W.T., Ernst, S., Nalepa, G.J., Szpyrka, M., Czapko, M., Grzesiak, P., Krzych, M.: Prototypes of a web system for citizen provided information, automatic knowledge extraction, knowledge management and gis integration. In: Dziech, A., Czyżewski, A. (eds.) Multimedia Communications, Services and Security, Communications in Computer and Information Science, vol. 149, pp. 268–276. Springer Berlin Heidelberg (2011)
8. Ligęza, A., Ernst, S., Nowaczyk, S., Nalepa, G.J., Szpyrka, M., Furmańska, W.T., Czapko, M., Grzesiak, P., Kałuża, M., Krzych, M.: Towards enregistration of threats in urban environments: practical consideration for a GIS-enabled web knowledge acquisition system. In: Dańda, J., Derkacz, J., Głowacz, A. (eds.) MCSS 2010: Multimedia Communications, Services and Security: IEEE International Conference: Kraków, 6-7 May 2010. pp. 152–158. Poland (2010)

9. Nalepa, G.J.: Collective knowledge engineering with semantic wikis. *Journal of Universal Computer Science* 16(7), 1006–1023 (2010), http://www.jucs.org/jucs_16_7/collective_knowledge_engineering_with
10. Nalepa, G.J., Furmańska, W.T.: Review of semantic web technologies for GIS. *Automatyka: półrocznik Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie* 13(2), 485–492 (2009)
11. Noy, N.F., McGuinness, D.L.: *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford University, Stanford, CA, 94305
12. Waliszko, J.: *Knowledge Representation and Processing Methods in Semantic Web*. Master's thesis, AGH University of Science and Technology (2010)
13. Waliszko, J., Adrian, W.T., Ligeza, A.: Traffic danger ontology for citizen safety web system. In: Dziech, A., Czyżewski, A. (eds.) *Multimedia Communications, Services and Security, Communications in Computer and Information Science*, vol. 149, pp. 165–173. Springer Berlin Heidelberg (2011)
14. Wilson, J., Fotheringham, A. (eds.): *The Handbook of Geographic Information Science*. Blackwell Publishin Ltd (2008)
15. Żywioł, M.: *Analysis and evaluation of crime mapping systems* (2012)