

A SCHEME OF BUILDING SEMANTIC DESCRIPTION OF GEOGRAPHIC PROCESSING SERVICES

Xia Zhang^{a,*}, Deren Li^b, Xinyan Zhu^b

^aSchool of Printing and Packaging, Wuhan University - zxx@whu.edu.cn

^bState Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University

Commission VI, WG VI/5

KEY WORDS: Internet GIS, Semantics, System Integration, Geographical Services, Interoperability, Intelligent GIS

ABSTRACT:

In this paper, a systematic scheme of building semantic description of Geographic Processing Services (GPServices) based on ontology is proposed. According to the principle of ontology, GPServices taxonomy with multi-semantic-granularity is first built by extending taxonomy in ISO 19119. Then common semantics of GPServices instances are studied and summarized. After that, semantic associations between GPServices concepts, instances and properties of concepts and instances are analyzed and redefined based on Ontology Web Language (OWL) and Semantics Markup for Web Services (OWL-S). Finally, a model for GPServices semantics description is proposed by extending OWL-S. GPServices classification ontology and geographic information ontology is introduced to describe semantics of GPServices concepts and geographic information data in services parameters. GPService Quality ontology and GPService Security ontology are newly defined to represent GPServices instances semantics. With the proposed model, description of GPServices semantics can be built and express the meaning of GPServices to clients.

1. INTRODUCTION

With the development of Semantic Web, semantic description of services has become a key problem. It is the basis for automatic service discovery, service execution and even service composition. Therefore, in geographic information domain, it is necessary to study the description of Geographic Processing Services (GPServices) semantics to keep up with the development of Semantic Web. Among theories for semantic representation, ontology is powerful and widely accepted for describing concepts and associations between concepts for certain domain. There has been a lot of study on semantic description of geographic information based on ontology. However, GPServices semantic description based on ontology is rarely studied. This paper focuses on studying how to describe semantics of GPServices with ontology. The purpose is to propose a model for GPServices semantic description.

To express semantics with ontology explicitly, Ontology Web Language (OWL) is recommended by W3C. In OWL suite, OWL-DL is based on Description Logic (DL), which provides powerful expressivity and decidability and guarantees that reasoning process can terminate and always return correct result. A DL knowledge base is analogously typically comprised by two components - a "TBox" and an "ABox". The TBox contains intensional knowledge in the form of a terminology ("taxonomy" could be used as well) and is built through declarations that describe general properties of concepts. TBox or taxonomy explicitly represents subsumption between concept expressions. The ABox contains extensional knowledge - also called assertional knowledge that is knowledge specific to the individuals of the domain of discourse.

According to OWL and DL, semantics of GPServices should be described with ontology in both abstract viewpoint and concrete

viewpoint which correspond with TBox and ABox. In this paper, GPServices taxonomy is firstly studied in abstract level. A method for classify diverse GPServices is proposed and subsumption between GPServices is established according with this method. In concrete level, common semantics of GPServices instances are analyzed and summarized. Then semantic associations between GPServices concepts are studied besides subsumption relationship.

Based on the work in both levels, a model for depicting GPServices semantics is proposed by extending Semantics Markup for Web Services (OWL-S) which is defined based on OWL for representing services semantics.

2. GEOGRAPHIC PROCESSING SERVICES TAXONOMY

2.1 Classification Method

Linear classification and surface classification are two basic classification methods. Both of them take use of attributes of the objects. However, each of them has its own advantages and disadvantages. In linear method, objects are classified according to some fixed attributes. With this method, objects can be classified into different hierarchies. Logical relationship between different classified items can be well represented. However, the position of some items cannot be decided just according to the fixed attributes in some cases. In surface classification, category surfaces are built according to the combination of all attributes of objects. This method has good adaptability and flexibility. Category surfaces can change with the requirement of application. But space redundancy can be produced in this method. There may be no items be classified into some surfaces. Therefore, in this paper, linear method and

* Tel: +86-27-68778144, Fax: +86-27-68778488, email: zxx@whu.edu.cn

surface method will be combined together to build GPSServices taxonomy.

2.2 Taxonomy in ISO 19119

Owing to the diversity of GPSServices, it is difficult to build the classification of GPSServices. The main contribution in this field is from the work of OpenGIS Consortium (OGC) and ISO/TC211. In ISO 19119 document, OGC and ISO/TC211 proposed a geographic services category in Open System Environment (OSE). In the proposed category, the taxonomy of GPSServices is presented according to the property types for features given by the General Feature Model presented in ISO 19109, including spatial property, thematic property, temporal property and metadata property. The taxonomy is shown in Table 1.

Geographic processing services
- Geographic processing services – spatial
- Geographic processing services – thematic
- Geographic processing services – temporal
- Geographic processing services – metadata

Table 1. GPSServices Taxonomy in ISO 19119

In this taxonomy, GPSServices that modify the same type of attributes of features are classified into a same category. For example, coordinate conversion service and coordinate transformation service are classified into GPSServices - spatial. With this taxonomy, users can approximately decide the classification type of certain service according to the property types for features that the service deals with and vice versa.

However, it is very limited to get the classification semantics or subsumption between GPSServices with this taxonomy. There are some disadvantages in this taxonomy. First, the capability of this taxonomy is limited. The proposed four categories cannot cover all kinds of GPSServices. For services that do computation with two or more property types can not be classified into any category. For example, feature matching service is defined in ISO 19119 as service that determines which features and portions of features represent the same real world entity from multiple data sources. It is classified into GPSServices-spatial. But features cannot be decided whether they are matched just according to their spatial property. Temporal and thematic properties should also be taken into account.

Besides, subsumption between GPSServices cannot be fully represented with this taxonomy. There is only two-level classification. On the second level of each category, there is no further classification. For instance, spatial analysis service can be classified as a kind of GPSServices-spatial. Route determination service and proximity analysis service mentioned in ISO 19119 should be subclasses of spatial analysis service. But they are directly put into GPSServices-spatial.

Therefore, the semantics granularity of taxonomy in ISO 19119 is too simple to be used as TBox of GPSServices. In this paper, it is improved and extended with linear and surface classification method to build up a multi-semantic-granularity GPSServices taxonomy.

2.3 Multi-semantic-granularity GPSServices Taxonomy

According to classification methods summarized and disadvantages of ISO 19119 GPSServices taxonomy analyzed above, new service categories are firstly proposed to accommodate services that deal with features with two or more property types. With the combination of four feature property types, 11 new categories surface are produced with surface classification method, as listed in Table 2.

	GPSServices Category Surface
Two Property Types	Spatial-Temporal Spatial-Thematic Spatial-Metadata Temporal-Thematic Temporal-Metadata Thematic-Metadata
Three Property Types	Spatial-Temporal-Thematic Spatial-Temporal- Metadata Spatial-Thematic-Metadata Temporal-Thematic-Metadata
Four Property Types	Spatial-Temporal-Thematic-Metadata

Table 2. 11 New GPSServices Category Surfaces

Together with the original four categories, there are totally 15 GPSServices categories surfaces. Figure 1 depicts 15 GPSServices categories surfaces with UML packages.

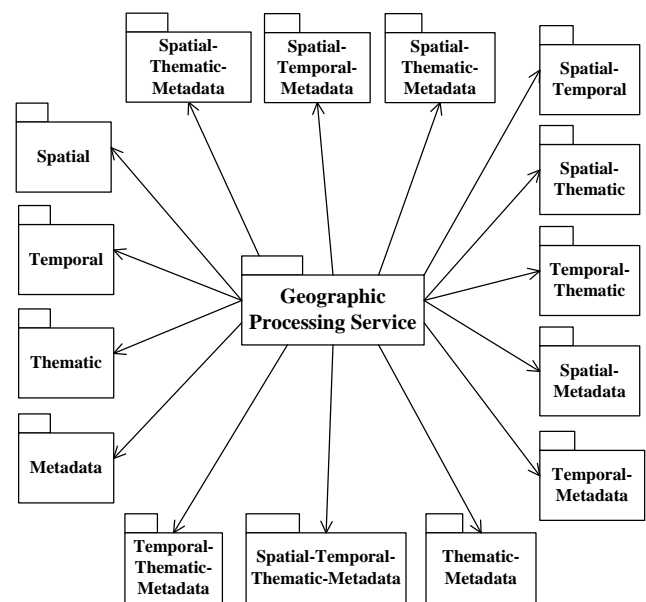


Figure 1. GPSServices Categories Surfaces

With 15 GPSServices categories surfaces, all kinds of GPSServices can be divided into certain category without confusion. For example, spatial analysis services that do computation only with features' spatial property will be classified into GPSServices-spatial category. But for spatial analysis services that deal with both spatial property and certain thematic property (such as population) will be classified into GPSServices-spatial-thematic category. These two kinds of spatial analysis services should be given different names to avoid conflict. For the former, spatial analysis service is

acceptable, for the latter, spatial- thematic analysis service can be used instead.

Under each surfaces, there will be many GPSServices. Further classification is needed to represent subsumption between them. Linear method is useful in the further classification. In this paper, subsumption between GPSServices in each surface are determined in terms of the semantic granularity of their processing. Processing semantic granularity of certain service can be determined according to semantics of its function or parameters. Services with same processing semantic granularity will live at the same layer. Services at different layer will have different semantic granularity. Coarser semantic granularity the service has, higher layer the service will be. Consequently, a multi-level GPSServices taxonomy will be produced. For example, as far as spatial analysis services under spatial surface are concerned, there are many kinds of spatial analysis services that do different computation, such as route determination services, buffer analysis services, overlap analysis services, and so on. They are subclass of spatial analysis services. For buffer analysis services, there will be different services that perform computation with features with different spatial type of features, such as point type, line type, surface type. For line-type or surface-type feature, they can be classified into simple and complicated line-type or surface-type features. Therefore, there can be two kinds of buffer analysis services to deal with these features with these two kinds of fine granularity spatial property. For complicated line-type or surface-type features, they can even be classified further. But we will not discuss more here. Part of the taxonomy tree for spatial analysis services under spatial category surface is represented with UML classes in Figure 2.

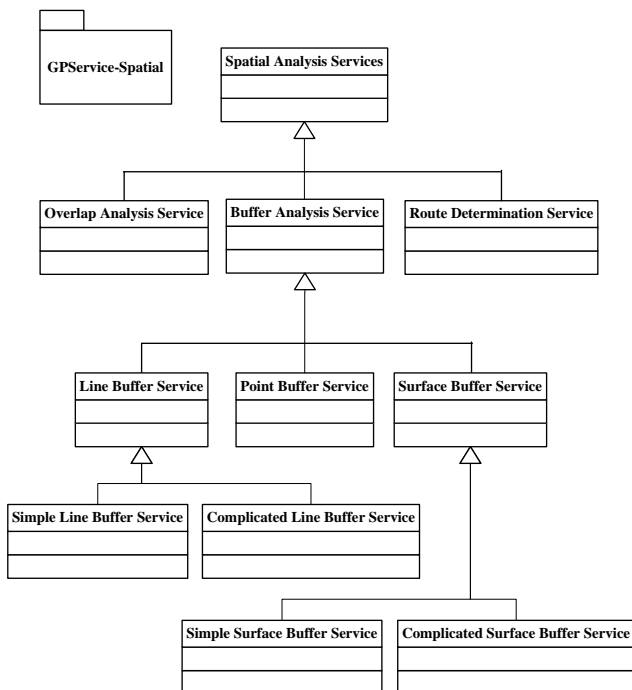


Figure 2. GPSService Taxonomy Example (Spatial Analysis Service)

Another example can be given about coordinate conversion services in spatial category surface. They can be reclassified

according to the coordinate systems the service manipulates. The classified services will have finer granularity semantics.

In the construction of GPSService taxonomy, settlement of services' name is also very important, which influence the semantic meaning of services. The names of services are the basic set of GPSServices terminology. A lot of work is needed to give each service a formal name because there are many homonyms, synonyms and colloquialism names about services name. This is what we will focus on in the future.

With the combination of linear and surface classification method above, services with different processing semantics can be classified into corresponding category. GPSServices taxonomy with multi-semantic-granularity levels can be established. The proposed taxonomy can overcome the disadvantages of taxonomy in ISO 19119 in some extent.

3. COMMON SEMANTICS OF GEOGRAPHIC PROCESSING SERVICES INSTANCES

According to Description Logic, intensional knowledge (TBox) is usually thought not to change. But extensional knowledge (ABox) is usually thought to be contingent, or dependent on a single set of circumstances, and therefore subject to occasional or even constant change. This means that it is not practical to depict semantic of GPSServices instances with uniform standards. They may change with their circumstances. However, items for describing common semantics of GPSServices instances are less influenced by physical condition. Therefore, it is possible to propose a uniform model for GPSServices instances semantic description.

3.1 Work of OWL-S and OGC

In the work for service semantic description, OWL-S is widely accepted even though it has some disadvantages. It is OWL-S is built based on OWL and recommended by W3C. It is an ontology language used for depicting service semantics. In OWL-S, three ontologies are defined to represent three aspects of service semantics, including ServiceProfile, ServiceModel and ServiceGrounding. Among them, ServiceProfile tells what the service does. It plays part of the role of GPSServices taxonomy, but not all. ServiceModel tells a client how to use the service and ServiceGrounding tells a user how to access the service. It can be seen that semantics represented with ServiceModel and ServiceGrounding are about services instances. They describe the inputs/outputs and precondition/effect (IOPE) of one service. They also represent processing steps of one service and the way for accessing services.

In geographic Information domain, OGC released a set of abstract specifications for Geographic Information Services (GIServices). In abstract specification for GPSServices (in OGC document called Web Processing Services), general information about GPSServices instances are expressed in a formatted document called "Capabilities". In "Capabilities" document, "OperationMetadata" and "ProcessOffering" elements are defined to represent services operations and the way for accessing services. Specific information about certain operation of GPSServices instances are depicted in document for the operation. In schema of operation description, "ProcessDescription" element is defined to represent

inputs/outputs and execution status of specific process of the service.

3.2 Common Semantics of GPService Instances

From the work of OWL-S and OGC, it can be summarized that two items of semantic description for GPService instances are necessary. One is functional semantics which express how the service works. It covers semantics of service operations, service IOPEs, operations IOPEs and processing steps of one process. Another is execution semantics which tells how to access the service. It includes semantics of provider, location, access ways (such as network protocol), and execution status of the service. It also involves the way of data mapping and message binding for accessing service.

Besides these two semantic description elements, description items for service quality and service security are also necessary. Both of them are not defined in OWL-S or OGC specifications. By analyzing related work for web services, the elements for describing GPService quality and GPService security are generally summarized. Due to the complexity of service quality and service security, summary in this paper just cover part of what service clients are interested.

The semantic contents of service quality may include maximum data volume afforded, stability, response time, reliability, cost and grade of the service. Stability shows difference of response times resulted from different network condition. Response time refers to time between request and response. Reliability represents maintenance of the service. Service cost means the fee paid for the service. Grade is the evaluation from service customer.

For service security, security mechanism should be described. It may include signature, encryption, protocol, and security notation. Protocol may involve data transfer protocol and key protocol. Security notation refers to authentication, authorization, access control, confidentiality, privacy, exposure control, anonymity, negotiation, key distribution and etc.

Therefore, functional semantics, execution semantics, quality semantics and security semantics are elements of common semantic of GPService instances. With these four elements, the model for GPService instances semantic description can be established. But specific semantics of each element should be decided according to specific circumstance. Further research need to be done in this field.

4. SEMANTIC ASSOCIATIONS BETWEEN GEOGRAPHIC PROCESSING SERVICES

According to the definition of ontology, semantic associations between concepts, instances and properties are important parts of GPService semantics. They are the foundation for semantic reasoning.

4.1 Axioms and Constructs in OWL

In OWL, there are some axioms or constructs for defining semantic associations among concepts, instances and properties of concepts or instances. Semantic associations between concepts include "subClassOf", "equivalentClass" and "disjointWith". "subClassOf" is used for stating subsumption relationship between two concepts (classes), "equivalentClass"

for equal relationship and "disjointWith" for disjoint relationship. Set association constructs "intersectionOf" and "unionOf" shows that one concept is the intersection or union of another two concepts. And "complementOf" indicates that one concept is the complement of another. Associations between concepts and instances can be defined with "instance of". "sameAs" and "differentFrom" constructs explain that two instances are same or different. Constructs for properties involve "equivalentProperty", "TransitiveProperty", "SymmetricProperty", "FunctionalProperty", "InverseOf" and "InverseFunctionalProperty". They are very useful for reasoning according to properties of concepts or instances.

4.2 Constructs for GPService Composition

Axioms and constructs in OWL can be used to establish the basic relationship between GPService concepts and instances. According to geographic domain knowledge, there is explicit and fixed process sequence between some geographic processing. There is also some potential or optional composition between services. For example, some spatial-thematic analysis service may be formed with one spatial analysis service and one thematic statistic analysis. Therefore, there is a kind of composition relationship between these two services. The relationship of service composition is a kind of knowledge for intelligent service composition which can contribute to full use of services resources. It should be expressed in service semantic description.

To represent composition associations between GPService, it is necessary to define some axioms or constructs which are not in OWL constructs set. In OWL-S, there are some structural constructs for inner process control of one service's processing. They are "Sequence", "AnyOrder", "ConditionOf", "Split + Join", "Iterate" and "Choice", which cover the basic control of processes in a service operation. They can be used together to produce complex constructs. Since certain GPService carry out certain processing work, these constructs can also be used as service composition constructs. With a triple (service1, composition association, service2), composition relationship between service1 and services 2 can be expressed with "composition association".

If a service is a composite service that aggregates several services, these constructs can be used to together represent the sequence among the services aggregated. Restricted cardinality of the association can indicate the relationship between two services is required or optional.

5. GEOGRAPHIC PROCESSING SERVICES DESCRIPTION MODEL

Based on the research on GPService taxonomy, common semantics of GPService instances and semantic associations among GPService, a description model can be proposed. Since OWL-S is the recommended standards for service ontology representation, the proposed model is presented by extending OWL-S.

OWL-S has defined ServiceProfile, ServiceModel and ServiceGrounding to express what the service does, how the service work and how to access the service. However, it dose not provide a way to represent GPService's classification semantics, to express geographic information data semantics in IOPEs, and to deliver semantics of service quality and service

security. These problems are considered and modelled in this paper. The whole model is shown in Figure 3~7.

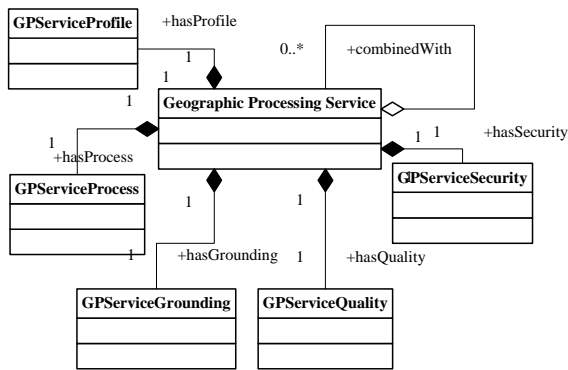


Figure 3. GPServices Semantic Description Model

Figure 3 shows five top ontologies that GPService ontology resembles with defined properties. Among them, GPServiceProfile, GPServiceProcess and GPServiceGrounding represent similar information to ServiceProfile, ServiceModel and ServiceGrounding in OWL-S. GPServiceQuality and GPServiceSecurity depict service quality and security semantics. One GPService can be combined with one or more GPServices with composition property “combinedWith” which is defined with basic composition constructs.

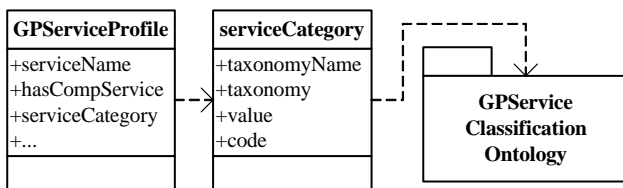


Figure 4. GPServiceProfile Ontology

Properties of GPService Profile ontology are shown in Figure 4. Property “serviceCategory” relates GPServiceProfile with GPService Classification Ontology which is built based on GPServices taxonomy. In class “serviceCategory”, property “taxonomyName” can be used to represent name of GPServices category surface. “taxonomy” refers to the hierarchical classification of GPServices under specific category surface. “value” and “code” is assigned to each GPService when building GPServices classification taxonomy.

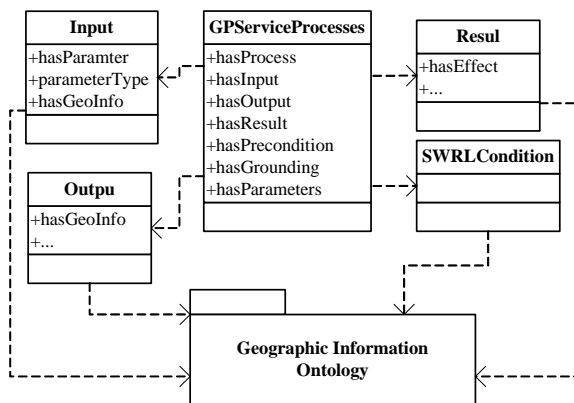


Figure 5. GPServiceProcess Ontology

As shown in Figure 5, GPServices Process ontology describes the IOPE of the service with classes Input, Output, Result and SWRLCondition. In the parameters of IOPEs, there always are geographic information data. Therefore, Geographic Information Ontology is introduced represent geographic information semantics. For Geographic Information Ontology, a lot of work needs be done. It is out of this paper.

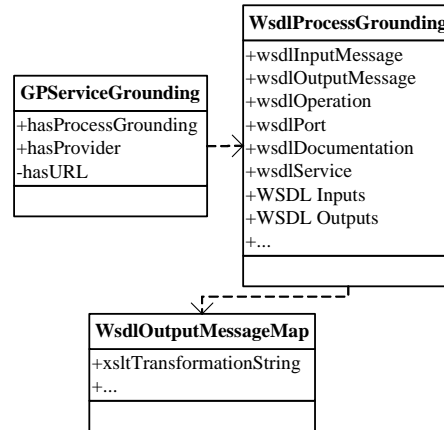


Figure 6. GPServiceGrounding Ontology

Same as that in OWL-S, GPServiceGrounding in Figure 6 can use WSDLProcessingGrounding to support service access.

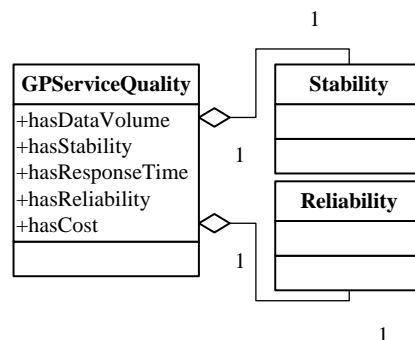


Figure 7. GPServiceQuality Ontology

In Figure 7, GPServiceQuality is defined with five parameters as its properties. Object property “hasStability” relates GPServiceQuality with Stability class and “hasReliability” with Reliability class to describe service stability and reliability.

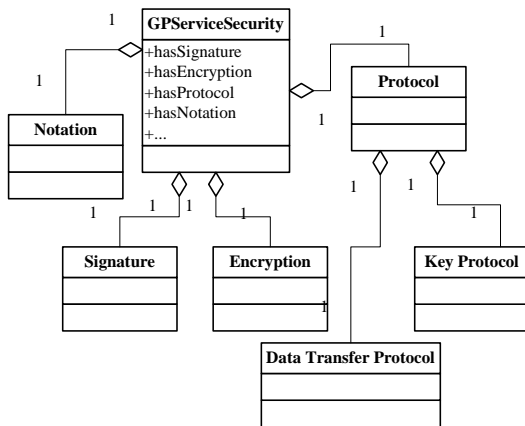


Figure 8. GPSecurityService Ontology

It is shown in Figure 8 that GPSecurityService ontology aggregates Signature, Encryption, Protocol and Notation through four properties. Class Protocol is composed of Data Transfer Protocol and Key Protocol whose semantics needs further study.

With these five top ontologies, GPServices semantics can be represented in both abstract viewpoint and specific viewpoint.

6. CONCLUSION AND FUTURE WORK

It is expected that the semantic description of GPServices built with the proposed scheme can help the implementation of semantics-based discovery, integration or composition of GPServices on Internet/Web, which may contribute to the development of GPServices toward intelligent services.

However, the research in this paper just provides a framework for GPServices semantic description. To reach the aim for intelligent GPServices application, a lot of work need be done, such as problems of homonyms, synonyms words in building GPServices terms, application semantics of GPServices in specific condition and more other problems. These are our future interested work.

ACKNOWLEDGEMENTS

This paper is supported by project of Open Research Fund of State Key Laboratory of Information Engineering in Surveying,

Mapping and Remote Sensing: No. WKL (06) 0301, National Natural Science Foundation of China (No.40701156) and National Key Basic Research and Development Program of China (No.2004CB318206). Thanks to the work of members in our groups.

REFERENCES

- Daniel A Menasc., 2002, QoS Issues in Web Services. IEEE Internet Computing, 6(6), pp.72-75.
- Daniele Nardi, Ronald J. Brachman., 2003. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, pp.16-20.
- David Martin, et al., 2006, OWL-S: Semantic Markup for Web Services. <http://www.ai.sri.com/daml/services/owl-s/1.2/overview/> (accessed 10 Sep. 2007)
- Denker G,Kagal L,Finin T., 2005, Security in the Semantic Web Using OWL. Information Security Technical Report, (10), pp.51-58.
- George Percivall, 2002, OpenGIS Service Architecture (same as ISO 19119). OpenGIS Abstract Specification Topic 12, pp.27-30.
- Michael K. Smith, et al., 2004, OWL Web Ontology Language. Guide. <http://www.w3.org/TR/2004/REC-owl-guide-20040210/> (accessed 10 Jun. 2007)
- Mike Dean, et al., 2004, OWL Web Ontology Language Reference. <http://www.w3.org/TR/owl-ref/> (accessed 25 Jun. 2007).
- Peter Schut, 2007, OpenGIS® Web Processing Service, OGC Standard 05-007r7.
- R. Stevens et al., 2007, Using OWL to model biological knowledge. *International Journal of Human-Computer Studies*, 65(2007), pp.583-594.
- U. Kuter et al., 2005, Information gathering during planning for Web Service composition. *Journal of Web Semantics*, (2005), pp.183-205