



HAL
open science

Geospatial Web Services Semantic Discovery Approach Using Quality

Khaled Rezeg, Mohamed Tayeb Laskri, Sylvie Servigne

► **To cite this version:**

Khaled Rezeg, Mohamed Tayeb Laskri, Sylvie Servigne. Geospatial Web Services Semantic Discovery Approach Using Quality. *Journal of Convergence Information Technology*, 2010, 2, 5, pp.28-35. hal-01381617

HAL Id: hal-01381617

<https://hal.science/hal-01381617>

Submitted on 4 Dec 2020

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.

Geospatial Web services semantic discovery approach using quality

Rezeg Khaled ^{*1}, Laskri Mohamed Tayeb ^{*2}, Sylvie Servigne ^{*3}

^{*1} Computer science department, University of Biskra, 07000, Algeria

^{*2} Computer science department, University of Annaba, 23000, Algeria

^{*3} Research Center for images and information systems (LIRIS) Lyon, 69000, France

rezeg_khaled@yahoo.fr, laskri@univ-annaba.org, sylvie.servigne@insa-lyon.fr

Abstract

The use of geographic data through Web services shows the geospatial Web Services (GWS). A growing number of GWS designed to interoperate spatial information over the network have emerged. GWS are changing the way in which spatial information systems and applications are designed, developed, and deployed [1]. When GWS is increasing, the difficulty of the service discovery increases too. The service discovery represents one of the interactions between Web service components (supplier, customer and UDDI). It is unfortunately that it is based only on the syntax. The semantic aspect necessary to satisfy the user is absent. The use of semantic Web technology in the Web service discovery phase gives rise to the Web service semantic discovery. The latter is based on semantic reasoning. It enhances the accuracy of search results compared to traditional techniques of Web service discovery, the additional matching accuracy in terms of computing power.

In this paper we are interested to the GWS semantic discovery based on geographical metadata held in geo-catalogs. The metadata used in our case is designed according to the ISO 19119 standard reinforced by the quality criteria.

Keywords

Semantic discovery, Metadata, User profile, Service quality, Geospatial Web Service.

1. Introduction

The geographic information systems today provide access to a large number of heterogeneous and distributed sources. The multiplication of resources and increased data requires the use of Web services that appear as a solution to ensure interoperability between different resources. They are able to collaborate and share control of process data and information between

applications on different platforms. The use of geographic data through Web services gives the Geospatial Web Services (GWS). A growing number of GWS designed to interoperate spatial information over the network have emerged. GWS are changing the way in which spatial information systems and applications are designed, developed, and deployed [1]. When GWS is increasing, the difficulty of the service discovery increases.

The rest of our paper is organized as follows: section 2 introduces the GWS, section 3 shows quality concept (data quality, service quality and use quality). Section 3 presents our Geospatial Web services semantic discovery approach. Finally, this paper contains some conclusions and perspectives.

2. Geospatial Web Service (GWS)

The traditional uses of geographic information generate many problems as cited in [1]. The marriage of Web service with geospatial information leads to GWS [1].

Generally, a GWS can be viewed as a modular Web application that provides services on geospatial data, information, or knowledge. It refers to the use of Web services technologies to manage, analyze, and distribute spatial information. Furthermore, a GWS can be sorted and searched through its geospatial characteristics, such as location, area, neighborhood, and other spatial features. As a Web service, a GWS involves three actors: user entity (consumer), provider entity (provider), and register entity (broker). Basic operations during the life cycle of a GWS include publication, discovery, binding, invoking, and execution. The interfaces of some GWS have been standardized. The most important players for such standardization efforts are ISO/TC211 and Open Geospatial Consortium (OGC) [1].

The ability to represent geospatial semantics is a great importance when building geospatial Web

applications. The Semantic Web Services (SWS) provides technology solutions for intelligent service annotation, discovery, composition and invocation in distributed environments. Deployment of this technology in geospatial Web applications has the potential to improve the discovery and integration of geospatial information and its reuse in other contexts than the original one [2]. The GWS semantics is ensured through semantic classifications, such as that proposed by the ISO standard (ISO text for ISO 19119 Geographic Information - Services) in [3, 4, 5]. It classifies the different services even more detailed and distinguishes six classes of services. These classifications will present as follows:

1. Human interaction service: services for the control and management of user interfaces, graphics, multimedia, etc. (Catalog viewer, Geospatial viewer, Geospatial feature editor).
2. Model/Information management services: management development, processing and storage of metadata, concepts and datasets (Feature access service, Map access service, Coverage access service, catalog service).
3. Workflow/Task services: services for users through tasks and activities (Chain service definition, work flow enactment service).
4. Processing services: a Web processing service provides access to calculations and models that deal with spatial data.
5. Communication services: encoding services / decoding and transferring data within telecommunication networks.
6. Management services: control system components service, applications and networks (eg: management of user accounts).

3. Geospatial Web service quality

The quality is defined generally as: "all the properties and characteristics of a product or a service which gives it the ability to satisfy the needs expressed or implied (ISO 8402: 1994) [6].

The GWS quality is based on:

- **The spatial data quality:** this quality is ensured through criteria set by the spatial aspects for assessing the relevance of data from the geographical context [7]. They are described by ISO 19113 which provides a description of the quality of spatial data using seven criteria (three quantitative, three qualitative and one specific) [6, 8].
- **The WS quality:** the term has been used to

express non-functional requirements for different sectors such as the research community and network issues in real time [9]. The QoS properties show different characteristics and are expressed in different forms, as they may relate to the services themselves, particularly in terms (i) performance (reliability, execution time expected, etc..), (ii) execution context (latency, throughput, etc..) or (iii) the needs of the user (as the result of the economic cost calculation, etc.). In literature we find several classifications of QoS, including those cited in the work of A. Mani [10], R Shuping [9] and L. Médini [11].

- **The use quality:** in our context it represents the information requested by the user in terms of quality through its application. The "quality" dimension plays a very important role in personalization domain. This dimension data describe the quality expected by the user. These data will be confronted with the actual quality produced by the search information to narrow the search space [12].

A current trend is to exploit the quality aspect in the discovery of SWG. In this context there are several works like :

- The approach proposed in [13] uses semantic annotations for the automatic discovery and composition of GWS. The thesis first discusses how to combine semantics provided by metadata catalogues with semantics provided by ontologies to produce catalogue services capable of efficient searching for suitable Web services. Then, it addresses how to compose Web services from basic OGC Web services in an automatic or semi-automatic way. The thesis focuses on the composition of data services and portrayal services, and adopts existing SWS technology (WSMO/WSML/WSMX)
- The approach proposed by [14] which uses the descriptive metadata of the service attributes quality in the discovery process by matching between service quality and that presented in the use profile distinguishing the needs geospatial services for the customer.
- The work proposed by [15] deals with the SW can dynamically generate quality information for the results of spatial analysis taking into account the quality of input data.

- In the system proposed by [16], calculations of trust are established by the basic capabilities that match the parameters of QoS. Geospatial Services Ordering Metric project (GSOM) is used to assess quality of service and confidence building. It represents a complete architecture for a semantic awareness for security, discovery and orchestration of SWG.

4. Proposed Approach

Quality management plays a vital role to satisfy the user needs (explicit or implicit) [6, 7, 17]. Personalization aims to facilitate the expression of user needs and make information intelligible to the selected user and exploitable [12, 18]. The existence of certain standards provides taxonomies to facilitate

interoperability as the ISO 19119 standard which proposes taxonomy of GWS [5]. From the points mentioned above and the continuity of the work presented in [19] we propose a geospatial Web services semantic discovery approach reinforced by the criteria quality at different levels (data, service and use).

In this section we present the architectural and functional aspects of our approach. We specify the GWS model description, the proposed user query formalism, and we make a presentation of our implementation.

4.1. Architectural and functional aspect

The general architecture of the proposed model is as follows:

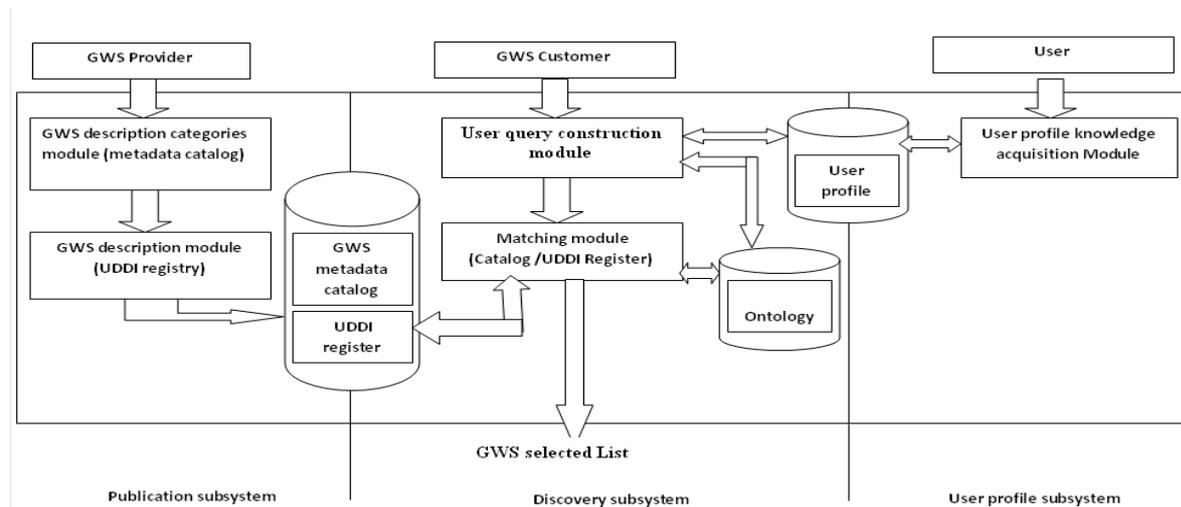


Figure 1: Proposed model architecture

4.1.1. GWS publication subsystem

In our context the GWS description is ensured through the following phases:

- The GWS annotation by specifying a geographic category and its qualification criteria for quality;
- The GWS standard description using WSDL.

Once this description is completed, the publication stage shall be as follows:

- A registration of GWS classes instances and their annotations is performed in the metadata catalog;
- Published in the UDDI registry.

4.1.2. GWS discovery subsystem

The user query is formulated by referring to user profile which allows us to enrich it. Once the request is made, the GWS discovery phases are:

- Find the mapped between quality criteria of GWS and those of the user query. The correspondence between the catalog entries metadata categories of GWS and the application is ensured through the use of the matching algorithm.
- Find the mapped the GWS functionalities and the query once: the results of the previous phase will be exploited in the discovery phase in the UDDI registry to restrict the search space in the class involved in applying the

matching algorithm.

4.1.3. User profile Subsystem

This subsystem is to acquire the user profile knowledge using the generic model of the user profile proposed by Mr. Bouzeghoub and D. Kostadinov in their work [12,18], by an instantiation of the model in the GWS. We are particularly interested to the personal dimension, the interest area dimension presented in our case by ontology and the quality dimension by adding the user preference for each quality concept. This preference is represented by a value between 0 and 1.

4.2. The GWS proposed model

The goal of our model is to facilitate the GWS semantic discovery based on enriching the GWS description, for suppliers to submit their GWS clearly and visibly, for customers to restrict the discovery space through user requests enriched by the user profile knowledge. The GWS description is ensured through: (i) a description GWS categories using ISO 19119 classification and quality criteria and (ii) the GWS description using the WSDL.

Our model: QGWSR-Model (Geospatial Web Service Quality Model Representation) is composed of 4 UML packages as shown in the following figure:

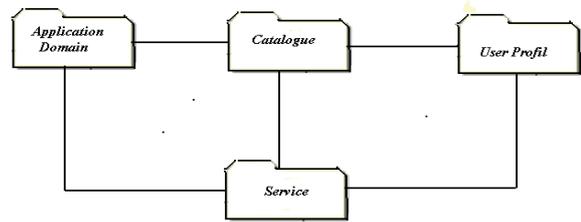


Figure 2: QGWSR-Model Overview

Catalogue: this package represents the core of the package model; it is the GWS categories description based on criteria of quality and ISO 19119 standards.

Service: this package represents the service description through the WSDL.

Application Domain: the Application Domain package describes the domain associated with the services represented.

User Profile: this package includes the user profile in our model to construct the user query according to his profile.

For the catalog package model description, we use the UML class diagram shown in Figure 3. This diagram is composed of two parts P1 and P2. The part P1 is based on the classification of GWS by ISO 19119 Standard presented in [5] and the part P2 represents our proposal to classify the types of GWS by quality criteria.

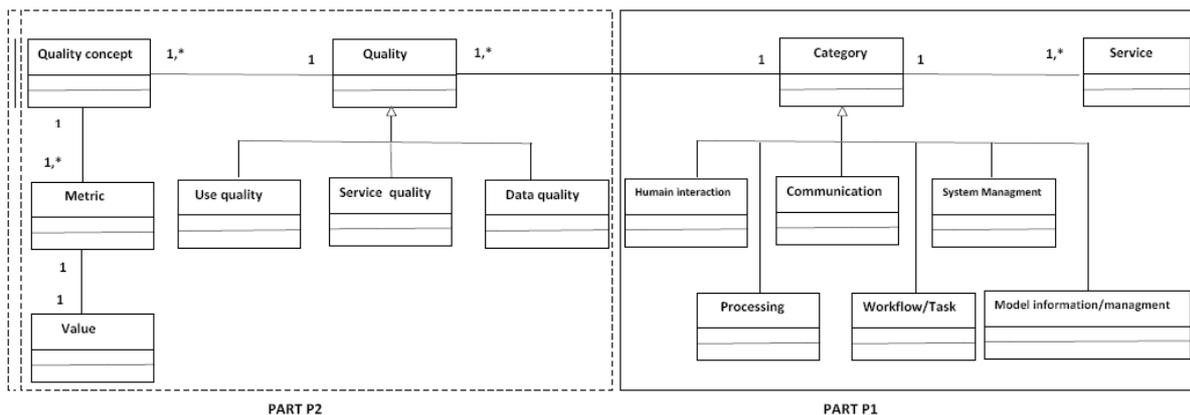


Figure 3: Diagram of GWS category model description

The diagram class is shown in Table 1.

Table 1. Class model describing GWS categories

Classe	Description
GWS category	abstract class designating GWS, it can be described in two facets: the identification and function
Human interaction	Services for monitoring and managing user interfaces, graphics, multimedia, etc.
Model/Information management	Management services for the development, processing and storage of metadata, concepts and datasets.
Workflow/Task	Support services to users through tasks and activities .
Processing	Treatment services can be configured to provide any type of GIS functionality to clients connected to a network, which includes access to calculations and / or computational models that apply to preprogrammed data reference space.
Communication	Services encoding / decoding and transferring data within telecommunication networks.
System management	The service control components of application systems and networks.
Service	The service class represents the GWS that includes the identifier of the UDDI registry to which the service belongs.
Quality	Abstract class designating quality classes.
Data quality	The criteria of spatial data quality (ISO 19115).
Service quality	The GWS quality criteria.
Use quality	Quality criteria of user profile
Quality Concept	Represents a quality concept.
Metric	Represents the indicators of quality evaluations
Values	Values corresponding to metric

In our work we propose three criteria for a quality class studied namely:

- **Fresh as data quality:** data freshness has been identified as one of the most important data quality attributes in information systems, it introduces the idea of data age. The metric for this quality are shown in the following table2.

Table 2. Summary Freshness factors and metrics [18]

Factor	Metric	Definition
Currency	Currency	The time elapsed since data was extracted from the source (The difference between query time and extraction time)
	Obsolescence	The number of updates transactions/operations to a since the data extraction time
	Freshness rate	The percentage of tuples in the view that are up-to-date (have not been updated since extraction time)
Timeless	Timeless	The time elapsed from the last update to a source (The difference between query time and last update time)

- **The response time as quality of service:** The web services response time can be divided into several components. We make a division as shown in Equation 1.

$$rws = tconn + treq + preq_resp + tresp \quad (1)$$

The total response time rws is in this model composed of the TCP connection setup (tconn), the request transmission delay (treq), the server-side processing delay for request parsing and response generation (preq_resp), and the response transmission delay (tresp). In addition to these delays, there can also be delays at the client side to compose the request message and to parse the response message, respectively [20].

- **The popularity like use quality:** the popularity refers to the degree of actual use of information by users [21]. The metric for this

test are shown in the following table3.

Table 3. Summary popularity factors and metrics

Factor	Metric	Definition
Popularity	Number of consultation	The number of GWS using
	Rating	The evaluation or recommendation given by a user group
	Number of citations	The number of citations referencing the GWS

4.3. Use query formalism proposed

The request reflects an approximate expression of user need. It is enriched by the profile that reduces the size of the results produced [12]. To assist the user during the preparation of its application, we propose a component query construction based on the user profile ontology and the scope through a friendly Web interface. This component provides the user with a prior request, according to its profile and gives the possibility to determine the degree of correspondence between the desired needs and offers potential GWS and refines its application using a pseudo language. This request must contain elements representing the semantic and syntactic aspect of GWS to discover. It must have a structure to the matching in the catalog metadata based on quality criteria, and matching at UDDI registry based on elements of the functional aspect of the service to be discovered.

The user query must be constructed by using the following grammar:

```

<User_query> ::= <Query_functionality>
               <query_function_param> (Where
               <Condition> ) ?
<Condition> ::= <QCC> ('AND' <condition>)*
<QCC> ::= <Metric_Q> <Operator> <Valeur_q>
          (<Op_Connectivity> <QC>)*
<Operator> ::= '=' | '<' | '<=' | '>' | '>=' | '<>' |
             <OP_ONT>
<Op_Connectivity> ::= 'AND' | 'OR'

```

The user query components are described in the following table 4.

Table 4. User query components

Query component	Description

Query_functionality	Specific category of GWS, in our case it is one of the following features: display, analysis, abstraction, abstraction or archiving. In our work we use two features: the analysis and display
Query_function_param	Represents the list of parameters specifying the functional aspect of GWS.
Condition	Represents simple or composed conditions of the request.
QCC	Category quality
Metric_Q	Represents the quality indicator in our case the metric components are basic.
Valeur_Q	Represents an atomic value of quality
OP_ONT	Operator related to the domain ontology of application
?	The element is optional
	This symbol defines an alternative
*	The element is optional or repetitive.
::=	The item on the left of the symbol is defined by the elements of straight

This grammar is a kernel that will be completed according to the application domain function such as the list of values and operators (the operators used are the comparison operators. We can use geospatial comparison operators captured from the geographical domain ontology).

4.4. Implementation

For the implementation of our approach, we propose the tourism field as application domain; we use a prototype ontology which includes only the part of the terminology field of tourism. We develop a Web application, an adaption of a traditional Web architecture; it is composed of the following:

- The provider introduce the description of the GWS (class-level and service level), the client submits requests and transmits data to the Web server.
- The Web server receives requests from the user. He reformulates and transmits the application server. This sends the XML response. The Web server presents the response HTML to the user.
- The application server performs all operations

on the metadata other than formatting: compliance audit of the basic metadata model.

- The database servers store the meta-databases.

Our implementation is based on the use of metadata and the algorithm of "matching". To make these persistent metadata, we chose the pair XML / XML Schema, based on the principle of markup languages SGML family. An example of XML schema is presented in Figure 04.

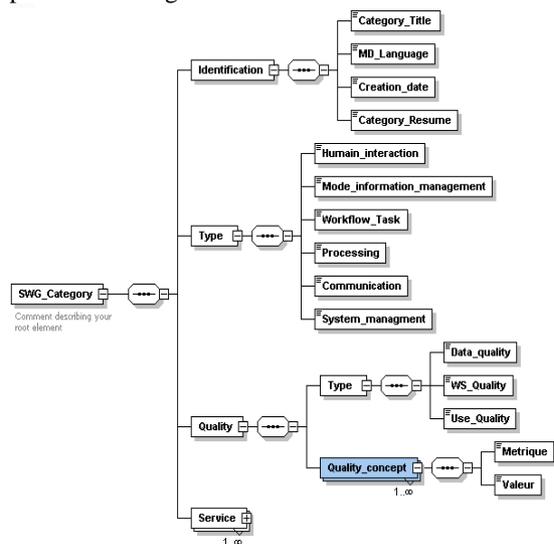


Figure 4. Extract File XML Schema of GWS class category (XML Spy 2005)

To combine the advantages of a DBMS and XML, we opt for the solution presented in [22] which is to split the metadata XML into several files identified by their id, and stored in BLOBs (Binary Large object) of a relational database.

In the metadata level, the matching algorithm "LARKS" presented in [23] is adjusted by applying the first three filters (context, profile, similarity) on the quality criteria contained in the request and description of classes available in GWS metadata catalog. This solution enables us to have a result like "relaxed match". In the UDDI registry level, the application of this algorithm [24] is based on the interrogation mechanism of the UDDI registry using the different pages (White pages, yellow pages, green pages). A list of GWS classified by order of correspondence (quality aspect, functional aspect) is the result of the discovery.

A set of interface is available to the user (Provider / Customer / User), the interface is made using PHP that provides ease of interfacing with databases [25]. These interfaces are:

- Interface for acquiring knowledge of the user profile;
- GWS publication interface;
- GWS discovery interface

5. Discussion

Paolucci [24] proposes an approach based on a description of Web service at one level by adding a semantic layer, based on the services capabilities discovery by using DAML-S as capabilities description language and adapting with the UDDI registry to allow services to be searched by keyword and also by semantic inference. Correspondence developed between a service and a request described in DAML-S is achieved by applying an algorithm that recognizes various degrees of equivalence determined by the minimum distance between concepts in a global ontology. Our approach is based on two levels of service description, the first concerning the categories of services based on the ISO19119 standard and quality criteria through the use of metadata catalog. The second one concerning the service description by using WSDL. The semantic layer in our case is represented by the metadata catalog. The discovery process proposed in our approach is also made at two levels using a global ontology. The first is the metadata catalog where we use the matching algorithm "LARKS", applying the first three filters (context, profile, similarity) on the quality criteria contained in the request and the descriptions of GWS classes in the catalog metadata. The second concerns the UDDI registry using the matching algorithm based on the interrogation mechanism of the UDDI registry using the different pages (White pages, yellow pages, green pages). The connection between the metadata catalog and the UDDI Registry is provided through a index table. So we can conclude that our approach provides a richer description than that proposed in the Paolucci approach and it provides a restriction of the search space at the UDDI registry. But it requires more knowledge to complete services description and it needs a component that provides the association between the catalog metadata and the UDDI registry.

6. Conclusion and prospects

Our discovery approach based on semantic metadata organized in catalogs, proposed in the publication phase of GWS two levels of description. The first level concerns the GWS classes catalog based on the ISO19119 standards and quality criteria; the second level concerns the UDDI registry through the WSDL. In the discovery phase, this solution allows to

restrict the search space and increase the number of relevant services through the application of the algorithm of "matching". This ensures the correspondence between semantic description of the GWS and the elements of the user query to the catalog level for quality criteria and the UDDI standard for the functional aspect using the application domain ontology.

As perspectives of this work, we want prospects, on one hand the enrichment of metadata to support other quality criteria of GWS to enhance the semantic discovery through experimentation of our approach in different areas. Secondly we will interest to the influence of quality criteria on the GWS composition.

7. References

- [1] P. Zhao, G. Yu, L. Di, Geospatial.Web.Services , Idea Group Inc., 2007.
- [2] R. Dumitri, E. Klien., SWING – A Semantic Framework for Geospatial Services , The Geospatial Web, Springer London ,2007.
- [3] Open GIS Consortium Inc., OpenGIS Web Services Architecture, 2003.
- [4] G. Percivall, ISO 19119 and OGC Geographic Information Service Architecture ,NASA Geographic Interoperability Office Global Science and Technology, Inc, 2002.
- [5] Yuqi Bai ,Liping Di , Yaxing Wei, A taxonomy of geospatial services for globalservice discovery and interoperability, Elsevier Computers & Geosciences , 2008.
- [6] G. Troispoux, « La qualité des données géographiques au service des utilisateurs », CERTU (Networks center studies : transport, urbain and public works, ecology, durable development and management ministry), France, 2007.
- [7] C.Gutiérrez, S. Servigne, R. Laurini, Towards real-time metadata for sensor-based network and geographic databases, 5th International Symposium Spatial Data Quality, ITC, Enschede, The Netherlands, 2007
- [8] S. Servigne, N. Lesage, T. Libourel, Spatial data quality: an introduction. Fundamentals of Spatial data Quality. Isted : London. ISBN 1905209568, 2006, pp. 179-208.
- [9] R Shuping , A Model for Web Services Discovery With QoS , ACM, 2003.
- [10] A Mani ,A Nagarajan, Understanding quality of service for Web services, <http://www.ibm.com/developerworks/library/ws-quality.html> , 2002.
- [11] L. Médini, «Systèmes d'Information Avancés (et répartis), Introduction aux services Web», CCI (Competence Complementary in Computer science) Master 2 , Network option, Lyon 1 University, 2005.
- [12] D Kostadinov, «Personnalisation de l'information : une approche de gestion de profils et de reformulation de requêtes»; Ph.D. Thesis, Versailles Saint-Quentin-En-Yvelines University, 2007.
- [13] E. R. Sacramento, Automatic Discovery and Composition of OGC Web Services, Phd Thesis, Cesra Federal University, Brasil, 2009.
- [14] M.S. Aktas, G. Aydin, G.C. Fox, H.Gadgil, M.Pierce, A.Sayar, Information Services for Grid/Web Service Oriented Architecture (SOA) Based Geospatial Applications, Scientific commons, 2005.
- [15] A. Donaubaue, T.Kutzner, F. Straub, Towards a Quality Aware Web Processing Service, GI-Days, 2008.
- [16] A. Alam, G. Subbiah, B. Thuraisingam, L.Khan, Reasoning with Semantics-aware Access Control Policies for Geospatial Web Services, Proceedings of the 3rd ACM workshop on Secure web, 2006.
- [17] L.Berti-Équille, Data quality assurance: a permanent challenge for information system, databases and warehouses, journal, British Library Direct, Number 74, 2005, pp. 13-21.
- [18] M. Bouzeghoub, V. Peralta, A Framework for Analysis of Data Freshness, 1st Int. Workshop on Information Quality in Information Systems (IQIS), Paris, France, 2004.
- [19] K. Rezeg, M.T. Laskri, O. Kazar, S. Servigne, Geospatial Web Services Discovery Approach Using Metadata and Multi-Agents System, International Computers and Software (IRECOS) ISSN 1828-6003, Vol 4 , 2009, pp 26-33.
- [20] J. Garcia, P. Hurtig, A. Brunstrom, The effect of packet loss on the response times of Web services , In proceeding of the IASTED International, 2006.
- [21] M. Bouzeghoub, S. Calabretto, N. Denos, R. Harrathi, D. Kostadinov, A. Nguyen, V. Peralta, «Accès personnalisé aux informations : approche dirigée par la qualité», In INFORSID 2007, Perros-guirec, 2007, pp. 105-120.
- [22] A. Yann, "Conception et exploitation d'une base de métadonnées de traitements informatiques, représentation opérationnelle des connaissances d'expert, Application au domaine géographique", Phd. Thesis, IGN France, CEAN University, 2006.
- [23] K. Sycara, S. Widoff, M. Klusch, M., J. Lu, Larks : Dynamic matchmaking among heterogeneous software agent in cyberspace. Autonomous Agents and Multi-Agent Systems , 2002, pp 173–203.
- [24] M. Paolucci, T. Kawamura, T.R. Payne, Katia Sycara1, Semantic Matching of Web Services Capabilities, The Semantic Web — ISWC 2002, Springer Berlin / Heidelberg, 2002.
- [25] PHP langage, <http://cerig.efpg.inpg.fr/Note/2002/php.htm>, 2002