

## **Application of Semantic Services in ORCHESTRA**

Ulrich Bügel<sup>1</sup>, Désirée Hilbring<sup>1</sup>, Ralf Denzer<sup>2</sup>

<sup>1</sup> Fraunhofer IITB, Fraunhoferstr. 1, 76131 Karlsruhe, Germany

<sup>2</sup> Environmental Informatics Group, Goebenstrasse 40, 66117 Saarbrücken, Germany

### **ABSTRACT**

This paper is part of a series of ORCHESTRA contributions to ISESS 2007. For better understanding of this article, it is recommended to read the introduction paper (Usländer<sup>2</sup> (2007)) first.

The paper outlines the definition, implementation and application of semantic services in the context of a service-oriented architecture as developed by the ORCHESTRA project.

While up to date most standardisation work in the area of distributed geospatial processing has concentrated on syntaxes for data exchange and functions in terms of Web services, the ORCHESTRA architecture additionally covers interoperability at the semantic level. In ORCHESTRA, ontologies provide background knowledge which can be used to explicate the meaning of content.

The paper illustrates how semantic descriptions of information and services are covered by the ORCHESTRA architecture, describes the design of basic semantic services and outlines semantic scenarios and their implementation in ORCHESTRA pilots. It can be concluded that utilization of semantic technologies in open, multi-domain and cross-boarder IT environments is an ambitious task but can provide significant improvement for stakeholders.

### **1. Introduction**

The overall goal of ORCHESTRA (ORCHESTRA (2007), Usländer<sup>2</sup> (2007)) is to design and implement a service-oriented architecture (SOA) that will improve the interoperability among actors involved in multi-risk management. The ORCHESTRA architecture is open and based on standards.

By adapting the ISO Reference Model for “Open Distributed Processing” (RM-ODP) to the requirements of a SOA, a reference model for the ORCHESTRA (RM-OA) architecture has been specified (Usländer<sup>1</sup> (2007)) and published in the context of the OGC Working Group “Risk and Crisis Management”. To date most standardisation work in the area of distributed processing has focused on the interaction syntax (e.g. signature of interfaces) and on the structure of the transferred data. These two aspects provide for physical connectivity and exchange of data, and they allow a distributed architecture to be realized. However, similar information content in different nodes may refer to substantially different things and the systems consequently do not interoperate at the level of semantics.

In ORCHESTRA, ontologies provide background knowledge which can be used to explicate the meaning of content. Ontology-based semantic meta-information is attached to information items and services such that these can be related to each other on a semantic level. Thus, in ORCHESTRA work on semantic interoperability promoted by various standard organisations - e.g. the World Wide Web Consortium (W3C) - is of prime importance.

The following chapters illustrate how basic semantic services are combined with thematic ORCHESTRA services in various European scenarios implemented in the ORCHESTRA pilots, in which the ORCHESTRA architecture is being validated. The inclusion of semantic services leads to significant functional enhancements and is the foundation for semantic interoperability.

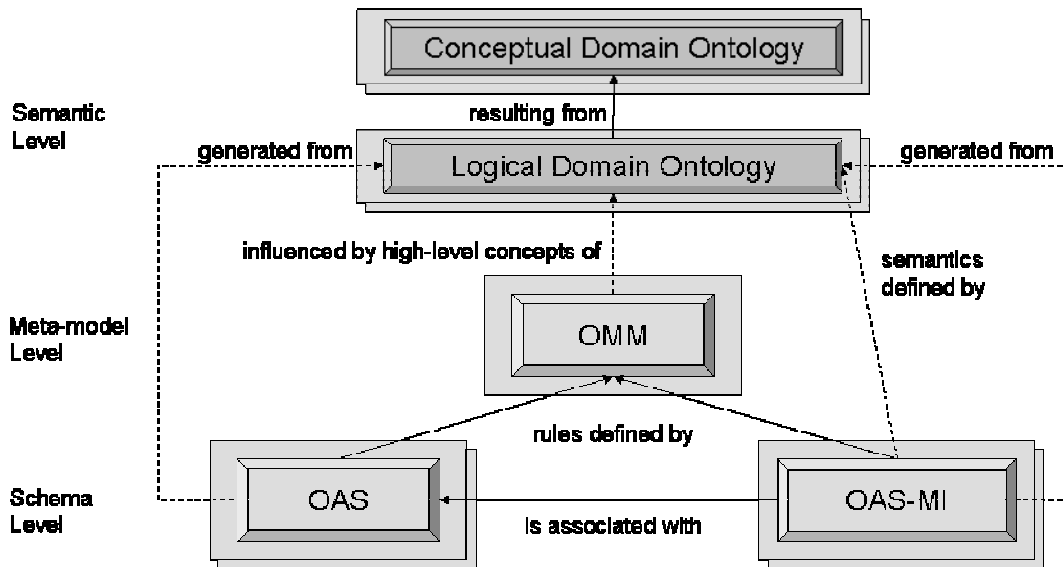
## **2. Semantic Information in the ORCHESTRA Architecture**

As outlined in (Usländer<sup>2</sup> (2007)), the RM-OA - Viewpoints „Information“ and „Service“ define a generic, platform neutral view of the architecture.

The information viewpoint is determined through the ORCHESTRA-Meta-Model (OMM), an extension of the ISO 19109 General Feature Model (GFM (2003)). The OMM defines rules for the building of concrete (ORCHESTRA) application schemas (OAS) and meta-information schemas (OAS-MI). The semantic level provides semantics to the information specified in the other levels, e.g. through explicit consideration of ontologies defined and shared in user communities (Figure 1).

An ontology is an explicit, formal specification of a shared conceptualisation (Studer (1998)). Ontologies may be thought of as a formal representation of the knowledge associated with a particular subject area (domain) or task. Their ultimate purpose is to enable machine understanding, which in turn provides the potential for data and service interoperability. In ORCHESTRA, ontologies are considered in two appearances according to the following two development stages of ontologies:

- A conceptual ontology is constructed by the domain expert. It is structured knowledge in an informal representation form which is understandable for the domain expert.
- A logical ontology is a transformation of a conceptual ontology into a machine-readable notation like e.g. OWL (2004). The transformation is done by an ontology expert



**Figure 1. Domain Ontologies in the Information Viewpoint**

Figure 1 illustrates some examples of relationships to the other levels of the specification framework, e.g.

- High level concepts defined in domain ontologies may influence the definition of new meta-classes in the OMM.
- OAS may be generated from logical ontologies if these have a sufficient level of detail, e.g. if they include typed slot definitions that may be mapped to feature properties types.
- The semantics of meta-information describing data or services (in an OAS-MI) can be defined by references to elements of an ontology. For description of service capabilities, the list of possible schemas is kept open and may include standard service meta-information schemas from ISO, OGC or any ontology-based standard suggested in the Semantic Web community (e.g. OWL for Services (OWL-S (2004) or the Web Service Modelling Ontology (WSMO (2006))).

### 3. Semantic Services

Similar to the information viewpoint, where the OMM provides a meta-model and a set of associated rules that control the specification of an OAS, in the service viewpoint the OMM provides corresponding rules for specification of service types (see Usländer<sup>2</sup> (2007)). The service viewpoint defines functional blocks of applications in form of consistent and redundancy-free service descriptions. There are two main service categories:

- Thematic services provide domain specific functionality such as execution of geo-statistic calculations or aggregations, simulation services etc.

- Architectural services are used by thematic services. They provide generic, platform-neutral and domain-independent functionality. They can be distinguished into infrastructure services such as catalogue service, services for user management, authentication and authorization, monitoring services or services which provide access to features, documents or sensors, and architecture support services such as services for service chaining or semantic services.<sup>1</sup>

The semantic services currently defined in the ORCHESTRA architecture are categorised as being architecture services. Their interfaces are first specified on an abstract level in UML and then mapped to a selected service platform (here: W3C Web services) in an implementation specification. Their functionality is described in the following subchapters. These services are defined such that semantic scenarios can easily be composed, what is illustrated thereafter. The overall task is to provide a basic service framework which makes the ORCHESTRA architecture capable of providing interoperability on a semantic level. The main challenge was to smoothly incorporate ideas and technologies developed by the Semantic Web community into the ORCHESTRA architecture, however, respecting the guidelines and constraints of the corresponding ISO and OGC geospatial standards.

### *3.1 Annotation Service*

According to Berners Lee (2001), the Semantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. The basic idea is to annotate content (and also services, databases etc.) with formal meta-information which can be embedded into a commonly understood context – given by means of ontologies – and further processed by computer applications.

While today most tools for semantic annotation are user driven or semi-automatic, the ORCHESTRA Annotation Service can automatically generate meta-information from various sources and relate it to elements of an ontology (e.g. concepts, properties, instances). In the first implementation, the focus is on annotation of unstructured and semi-structured sources (e.g. texts in natural language, documents or web pages). It is based on a tool for automatic Information Extraction (IE) (see GATE (2005), which is able to identify and normalize named entities occurring in the text by means of Natural Language Processing (NLP). The assignment of named entities to ontological elements is based on pre-populated background information held in a repository, e.g a knowledge base (see chapter 3.3). For further implementations it is planned to extend the Annotation Service such that it cannot only identify entities known in advance, but also discover new knowledge and populate it to a knowledge base, from where it can be queried by means of query languages.

The current implementation of the Annotation Service supports applications such as highlighting of found entities in the text and relates them to ontological elements, such that the user gets support in text understanding.

---

<sup>1</sup> Abstract specifications of ORCHESTRA architectural services can be downloaded under:  
<http://www.eu-orchestra.org/publications.shtml#OAspects>

### *3.2 Ontology Access Service*

The Ontology Access Service supports the read access to the specification of a logical ontology and to export or import a complete specification of a logical ontology into an ontology store. It provides a high-level view to the content of the ontology, allowing the client to get information about the taxonomy (classes and properties) and to extract vocabulary statements for human/machine interpretation.

The Ontology Access Service is independent of any ontology technology, e.g. ontology languages. Some typical usages of this service are:

- Retrieve the list of ontologies this service provides access to.
- Store, update or delete ontology entries.
- Get a stored ontology, partially or fully.
- Get lists of concepts, properties or individuals, layouted according to inheritance relationships.

### *3.3 Knowledge Base Service*

The Knowledge Base Service provides access to a knowledge base in an ORCHESTRA service network. It conveys query requests received via the network to the knowledge base's local processing engine and returns the results to the requesting service. It abstracts from existing languages for knowledge representation and querying, but it assumes some concepts which are common to most of them:

- Knowledge is represented as a graph, i.e. a number of nodes and edges.
- The knowledge graph is divided into a number of sub-graphs, so called "models", which are units of knowledge with a unique identification.
- Models are described by a number of basic elements constituting the model graph; these elements describe the nodes and the edges. Updates of a model can be performed by adding/deleting basic elements.

RDF (2004) is an example for a standard which fulfils these assumptions. In RDF, for instance, basic elements are called "statements", while SPARQL (2006) could be used as query language for querying RDF models.

## **4. Scenarios and Combined Semantic Services**

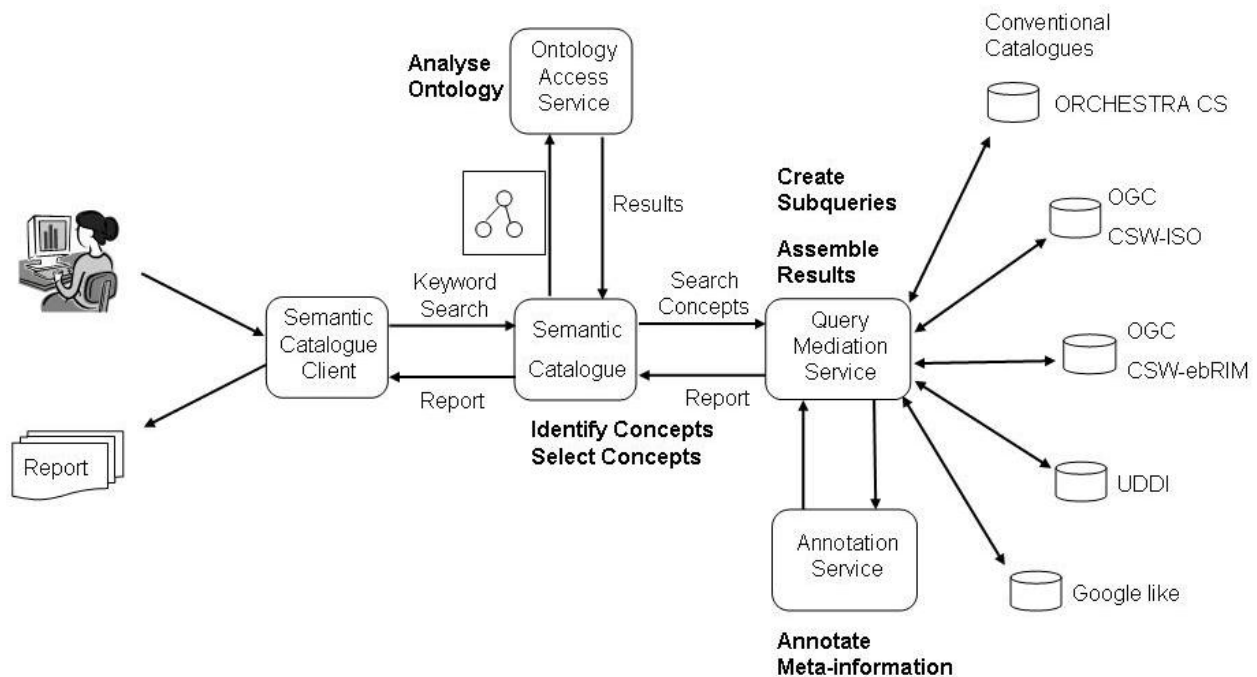
The presented semantic services can be flexibly combined to build use case scenarios supporting semantic interoperability.

### *4.1 Semantic Catalogue*

The Semantic Catalogue integrates different kinds of catalogue standards (e.g. ORCHESTRA Catalogue Services (CS), OGC Catalogue Services (CSW), UDDI registries or Google like meta-

information). It creates a consistent report containing meta-information gathered from several conventional catalogues. The conventional catalogues contain meta-information data and/or services for example dedicated to risk management. The resulting report is structured according to the concepts defined in a logical ontology, which provides a unique view on different vocabularies used in the underlying conventional catalogues.

Figure 2 illustrates the architecture. The user formulates a keyword-based query, which is sent to the Semantic Catalogue via its client. The Semantic Catalogue provides the interface of a conventional ORCHESTRA Catalogue Service. This ensures that clients of the Semantic Catalogue do not need to handle complex ontology based queries (in the sense of a knowledge base). The Semantic Catalogue Service identifies concepts of the ontology which can be related to the search keywords and expands the search query. It is assisted by the Ontology Access Service which is used to analyse the ontology. The Semantic Catalogue Service then searches for matching meta-information entries in the underlying conventional catalogues.



**Figure 2: Semantic Catalogue**

The access to the conventional catalogues is managed by the ORCHESTRA Query Mediation Service (QMS) which creates an appropriate query for each identified concept and each underlying conventional catalogue. The results of the conventional catalogues are delivered in their original meta-information schemas and are assembled by the QMS using the structure of the logical ontology into a report representing a coherent view. This meta-information report is delivered as result of the query to the user.

## *4.2 Combination of Semantic Catalogue and Annotation Service*

The Semantic Catalogue Scenario can be extended through the usage of the Annotation Service, which annotates the meta-information results given back by each underlying catalogue. It then classifies these results, eliminates results which do not adhere to the given ontology concepts and orders result sets according to their relevance w.r.t. the ontology. This leads to reduced and more meaningful result sets of the Semantic Catalogue.

## **5. Semantics in ORCHESTRA Pilots**

The definitions of the ORCHESTRA pilot applications have been carefully adjusted in liaison with the intended pilot stakeholders. While on one side it had to be assured that functionality to be realised is really desired by the stakeholders, each pilot had the task to focus at certain aspects by which the ORCHESTRA architecture can be validated. Therefore, each pilot had to document requirements by means of UML use cases in a pilot assessment template.

For each pilot application scenarios in a cross-boarder and/or cross-domain context have been defined. Web Services according to the W3C specifications have been selected as the implementation platform. While the development of services and applications in the first phase has concentrated on obligatory services needed in most service networks – e.g. Catalogue Service, Feature Access Service – in the ongoing second phase one focus is on the definition and implementation of semantic services and scenarios.

Semantic scenarios in the pilots are developed along some common goals, e.g.

- Users should be able to formulate required resources by means of queries to an ontology.
- The ORCHESTRA service network should be able find and combine services and data to provide the required results, e.g. in form of generated maps.
- When exploring the results, the user should be assisted by semantic services such that he/she can better understand and interpret the results.

While in other pilots the focus is on automatic service composition and chaining, the following examples concentrate on semantic enhancements for catalogue search and assistance for exploring obtained results.

### *5.1 The ORCHESTRA Meta-pilot*

The goal of the ORCHESTRA Meta-pilot is to create a portal providing catalogue search and access to all ORCHESTRA pilots. To achieve this, the Semantic Catalogue Scenario as outlined in chapter 4 is applied. The logical ontology applied here contains concepts which cover the risk management vocabulary as used throughout the ORCHESTRA pilots, while data and services accessible at the pilot sites are described by means of meta-information in the native catalogues.

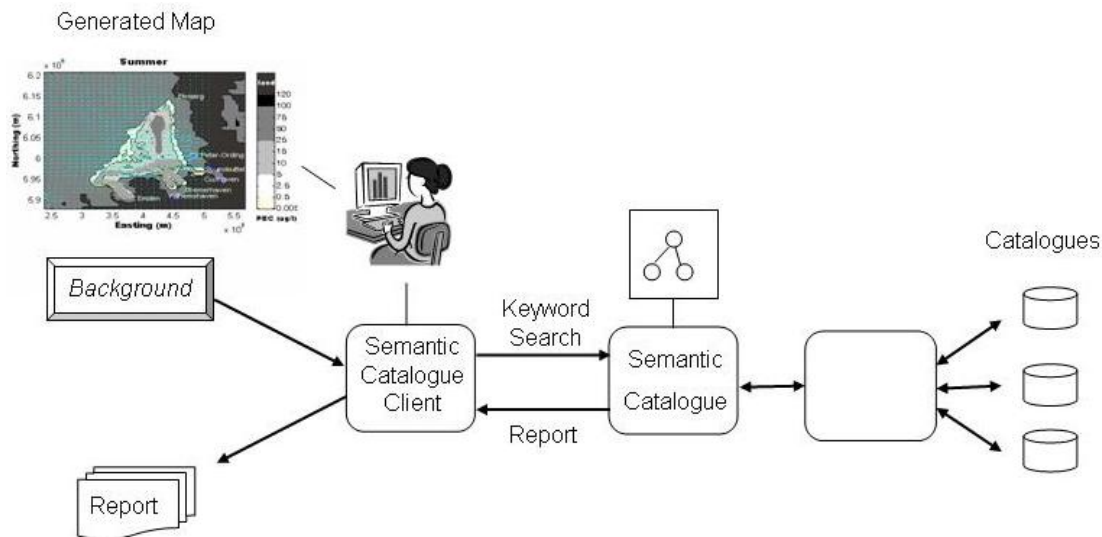
## 5.2 The ORCHESTRA Pilot “German Bight”

The ORCHESTRA pilot “German Bight” enables end users and providers of environmental information systems access to and creation of multi-risk maps for risks caused by ship traffic. The generation of the maps is based on integrated use of information about risk factors (e.g. chemical pollution), shipping routes, environmental databases (e.g. about distribution of species), numeric models (e.g. currents, dispersion) and cartographic data. This information is exploited in simulation models. In “what if” scenarios (e.g. introduction of new shipping routes) the generated risk maps can visualize potential risks for certain areas. This functionality is of high use for fishery, fish farms and tourism.

When a user has executed a simulation run and then examines the created risk map, various questions may arise concerned with the correct interpretation of the map, e.g.

- What is the impact of tributyltin (TBT) and/or copper on dedicated species?
- Which limits of pollution must not be exceeded?
- Which are the relevant regulations therefore?
- What are molluscs, gastropods and what causes imposex on these species?

In order to get answers to such questions, the user needs background information which is usually not directly shown on the generated maps. Therefore, the Semantic Catalogue Scenario is applied in the German Bight pilot (see Figure 3).



**Figure 3: Usage of Semantic Catalogue in “German Bight”-Pilot**

Existing conventional catalogues providing background information on German Bight issues can be attached to the Semantic Catalogue. Queries to the Semantic Catalogue can be directly issued when looking at a generated map. The ontology used here addresses concepts dedicated to describe potential risks and their impact on species in the German Bight.

The Annotation Service can assist the Semantic Catalogue in multiple ways:

- The generated risk maps may have associated legends or documents. Annotations relate text passages to elements of the ontology which can be used as a reading assistant (see Figure 4).
- In the same way, semantic annotation can be applied to documents received in query results of the Semantic Catalogue.
- Annotation can help to optimize query result sets (as described in chapter 4).
- Annotations can be populated as new knowledge to a knowledge base, from where it can be retrieved by means of queries to a Knowledge Base Service (as described in chapter 3).

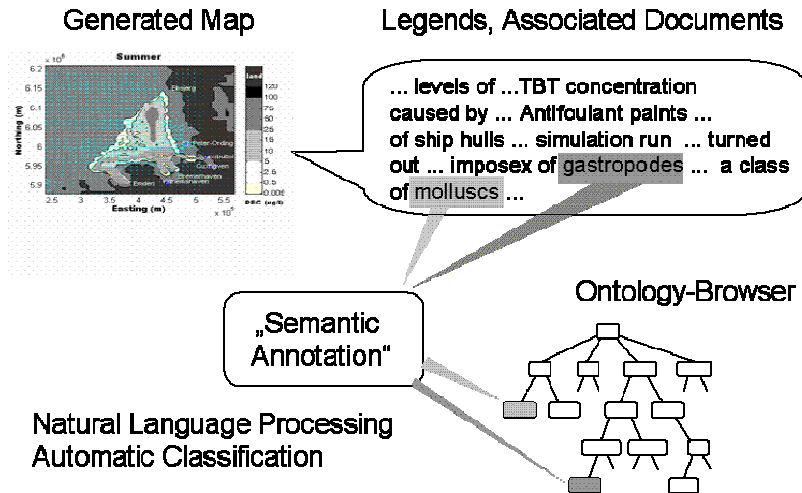


Figure 4: Semantic Annotation of Risk Maps

## 6. Conclusions

The development of geospatial environmental information systems up to date is mainly guided by utilization of standards promoted by OGC community. These standards are well reflecting the need of stakeholders and they have led to interoperability of systems on the syntactic level. Although there is also work in OGC standardization groups on semantic issues, the main driving force for semantic interoperability currently comes from the Semantic Web initiative and standardisation within the W3C community. However, Semantic Web technologies still have an academic flavour, there is comparatively little impact coming from the stakeholders site.

The work on semantic interoperability in the ORCHESTRA consortium attempts to bring the two worlds together on a practical and user-driven level. This has turned out to be an ambitious task, but ORCHESTRA has contributed to a proof of concept by developing an architecture which is capable to integrate semantic services whilst remaining compatible to OGC standards. Positive feedback from stakeholders leads to the conclusion that the application of semantic services can provide significant functional enhancements and opens the field of cross-domain and cross-boarder inter-working.

## 7. Acknowledgements

This work has been co-funded by the European Commission's IST programme, under Unit G5 "ICT for the Environment", through the Integrated Project FP6-IST 511678 ORCHESTRA (see <http://www.eu-orchestra.org>). Special thanks to the members of the sub-project 3 "Open Architecture" and the ORCHESTRA pilot team "Environmental Risks Caused by Ship Traffic in the German Bight in sub-project 4 for their good co-operation.

## 8. References

- Berners Lee (2001): Berners Lee, T., Hendler, J., Lassila, O.: *The Semantic Web*. Scientific American, Mai 2001
- Clement, L., Hatley, A., von Riegen, C., Rogers, T. (2004): UDDI Version 3.0.2, UDDI Spec Technical Committee Draft, OASIS, [http://uddi.org/pubs/uddi\\_v3.htm](http://uddi.org/pubs/uddi_v3.htm), 2007/02/21
- GFM (2003): ISO FDIS 19109 Geographic Information – Rules for Application Schema.
- Hilbring, D., Coraboeuf, D. (2007): WP3.4 – OA Service Specifications, Abstract Specification of the Catalogue Service, [http://www.eu-orchestra.org/docs/OA-Specs/Catalogue\\_Service\\_Specification\\_v1.1-BRGM-IITB.pdf](http://www.eu-orchestra.org/docs/OA-Specs/Catalogue_Service_Specification_v1.1-BRGM-IITB.pdf), visited in February 2007
- Nebert, D. Whiteside, A. (2005): OpenGIS Catalogue Services Specification, OpenGIS Implementation Specification, OGC 04-021r3
- ORCHESTRA (2007): Integrated Project ORCHESTRA (Open Architecture and Spatial Data Infrastructure for Risk Management), IST FP6-511679, 9/2004-8/2007, <http://www.eu-orchestra.org>, 2007/02/21
- OWL (2004): World Wide Web Consortium (W3C<sup>®</sup>): OWL Web Ontology Language. <http://www.w3.org/TR/owl-features>, 2007/02/21
- OWL-S (2004): OWL-S: Web Ontology Language for Services <http://www.w3.org/Submission/2004/07/>, 2007/02/21
- RDF (2004): World Wide Web Consortium (W3C<sup>®</sup>): Resource Description Framework (RDF). <http://www.w3.org/RDF/>, 2007/02/21
- RM-ODP (1998): ISO/IEC 10746 RM-ODP Reference Model for Open Distributed Processing
- SPARQL (2006): World Wide Web Consortium (W3C<sup>®</sup>): SPARQL Query Language for RDF. W3C Working Draft 4 Oct. 2006. <http://www.w3.org/TR/rdf-sparql-query/>, 2007/02/21
- Studer (1998): Studer, R.; Benjamins, V. R.; Fensel, D.: Knowledge engineering: Principles and methods. *Data and Knowledge Engineering (DKE)*, 25(1-2):161-197, 1998.
- Usländer<sup>1</sup>, T. (Ed.) (2007): RM-OA - Reference Model for the ORCHESTRA Architecture. Deliverable D3.2.3 of the ORCHESTRA Consortium, Version 2.0, January 2007, <http://www.eu-orchestra.org>, 2007/02/21
- Usländer<sup>2</sup>, T.; Denzer, R; Güttler, R (2007): Open Service-oriented Architecture for Environmental Risk Management Applications. International Symposium on Environmental Software Systems (ISESS), Prague, May 22-25, 2007
- WSMO (2006): WSMO: Web Services Modelling Ontology <http://www.wsmo.org/>, 2007/02/21