Discovery of Geospatial Resources:

Methodologies, Technologies, and Emergent Applications

Laura Díaz
Universitat Jaume I de Castellón, Spain

Carlos Granell
Universitat Jaume I de Castellón, Spain

Joaquín Huerta
Universitat Jaume I de Castellón, Spain

Information Science
REFERENCE

Managing Director:	Lindsay Johnston
Senior Editorial Director:	Heather A. Probst
Book Production Manager:	Sean Woznicki
Development Manager:	Joel Gamon
Development Editor:	Mike Killian
Acquisitions Editor:	Erika Gallagher
Typesetter:	Nicole Sparano
Cover Design:	Nick Newcomer, Lisandro Gonzalez

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Copyright © 2012 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Discovery of geospatial resources: methodologies, technologies, and emergent applications / Laura Diaz, Carlos Granell and Joaquin Huerta, editors.
p. cm.
Includes bibliographical references and index.
Summary: “This book provides relevant theoretical frameworks and the latest empirical research finding to improve understanding of geospatial discovery methodologies and technologies, as well as techniques to design and deploy geospatial resources in Information Infrastructures”--Provided by publisher.
ISBN 978-1-4666-0945-7 (hardcover) -- ISBN 978-1-4666-0946-4 (ebook) -- ISBN 978-1-4666-0947-1 (print & perpetual access) 1. Geographic information systems--Research. 2. Geospatial data. 3. Spatial analysis (Statistics) I. Diaz, Laura, 1976- II. Granell, Carlos, 1976- III. Huerta, Joaquin, 1965-
G70.212.D64 2012
910.285--dc23
2011048796

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

Chapter 9
Methodologies for Augmented Discovery of Geospatial Resources

Mattia Santoro
National Research Council (CNR), Italy

Paolo Mazzetti
National Research Council (CNR), Italy

Stefano Nativi
National Research Council (CNR), Italy

Cristiano Fugazza
Institute of Environment and Sustainability, Italy

Carlos Granell
Universitat Jaume I de Castellón, Spain

Laura Díaz
Universitat Jaume I de Castellón, Spain

ABSTRACT

Presently, solutions for geo-information sharing are mainly based on Web technologies, implementing service-oriented frameworks, and applying open (international or community) standards and interoperability arrangements. Such frameworks take the name of Spatial Data Infrastructures (SDIs). The recent evolution of the World Wide Web (WWW), with the introduction of the Semantic Web and the Web 2.0 concepts and solutions, made available new applications, architectures, and technologies for sharing and linking resources published on the Web. Such new technologies can be conveniently applied to enhance capabilities of present SDIs—in particular, discovery functionality.

Different strategies can be adopted in order to enable new ways of searching geospatial resources, leveraging the Semantic Web and Web 2.0 technologies. The authors propose a Discovery Augmentation Methodology which is essentially driven by the idea of enriching the searchable information that is associated with geospatial resources. They describe and discuss three different high-level approaches for discovery augmentation: Provider-based, User-based, and Third-party based. From the analysis of these approaches, the authors suggest that, due to their flexibility and extensibility, the user-based and the third-party based approaches result more appropriate for heterogeneous and changing environments such as the SDI one. For the user-based approach, they describe a conceptual architecture and the main components centered on the integration of user-generated content in SDIs. For the third-party approach, the authors describe an architecture enabling semantics-based searches in SDIs.

DOI: 10.4018/978-1-4666-0945-7.ch009

Copyright © 2012, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

INTRODUCTION

In recent years, the World Wide Web (WWW) has undergone several important changes in terms of available applications, architecture, and related technologies. The need for a more effective resource sharing through the Web raised awareness on efforts aiming to enable machine-to-machine applications on top of the Web architecture by making semantics explicit. These efforts are currently coordinated in the W3C Semantic Web Activity which "provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries […]. It is based on the Resource Description Framework (RDF)" (W3C, 2011). At the same time, new use cases, new available applications and technologies have made possible the WWW revolution which is known as Web 2.0 (O'Reilly, 2005). This term actually refers to an entirely new paradigm in the use of the Web as a platform for applications characterized by features like: delivery of services instead of packaged software, with cost-effective scalability; control over unique, hard-to-recreate data sources that get richer as more people use them; trusting users as co-developers; harnessing collective intelligence; leveraging the long tail through customer self-service; design of software above the level of a single device; lightweight user interfaces, development models, and business models (O'Reilly, 2005; O'Reilly & Battelle, 2009).

These two main changes jointly make new resources available, and new technologies to discover them through semantic relationships. Unavoidably, these changes would and should affect the geo-information sharing domain that is mostly based on web paradigms and technologies. Recently many efforts aim to provide more powerful tools for the discovery of geospatial information that is made available through traditional or Web 2.0 services, basing on explicit or implicit semantics (Klien, et al., 2004; Smits & Friis-Christensen, 2007; Lemmens, et al., 2006; NASA-JPL, 2011).

Information technology and geo-science are worlds in continuous change. Semantics and Web 2.0 are the present challenges, but new ones will emerge in the future. This raises the conceptual issue of enhancing geospatial information discovery capabilities in order to accommodate present and, possibly, future needs. This chapter describes two approaches based on the methodology of augmenting semantically the discovery process to enhance the search and retrieval of geospatial resources.

Data Discovery in the Geospatial Information Domain

It is estimated that more than 80% of data that human beings have collected so far are geospatial data in a wide sense (Frankling & Hane, 1992; MacEachren & Kraak, 2001), i.e. data with an explicit or implicit spatial/temporal reference. Moreover, this spatial/temporal reference is relevant, and even fundamental, for many applications. Therefore, it is of major importance to be capable of discovering geospatial data according to their content and geospatial characteristics (i.e., spatial coverage and temporal extent), and to effectively describe them.

Presently, solutions for sharing geo-information implement service-oriented frameworks applying open (international or community) standards and interoperability arrangements (Nativi, 2010). Such frameworks take the name of Spatial Data Infrastructures (SDIs). Typically, in service-oriented frameworks such as SDIs, discovery functionalities are provided by catalog components. A formal definition of these components is given by the ISO/TC211-Geographic Information/Geomatics, stating that a catalog service is a "service that provides discovery and management services on a store of metadata about instances. The metadata may be for dataset instances, e.g., dataset catalogue, or may contain service metadata, e.g., service catalogue" (ISO, 2003b). At present, the Open Geospatial Consortium (OGC[1]) Catalog Service Specification (OGC, 2007) defines a common core specification for geospatial resources discovery providing a consolidated framework. Services implementing this standard can be queried using common geospatial constraints (i.e., basing on *what*, *where*, *when*, *who,* etc.).

In addition to standard solutions based on ISO metadata and data models and OGC Web Services (OWS), many other de-facto standards are extensively used, defined in the context of specific Communities-of-Practice (CoP), such as THREDDS Data Server and OPeNDAP in the Meteo-Ocean domain, Global Biodiversity Resource Discovery Service (GBRDS) in the biodiversity domain, and so on. The advent of Web 2.0 further widens the geo-information world with services providing geospatial resources (GeoNames[2], GeoCommons[3], OpenStreetMap[4], etc.).

Statement of the Problem

In state-of-the-art catalog implementations, while the information about geospatial reference is usually precise (through the spatial and temporal extent expressed with well defined Spatial and Temporal Reference Systems), the information about content is often inadequate for some relevant use-cases. Indeed, it is often expressed using free text in multiple metadata fields or using controlled vocabularies for some specific fields (e.g. keywords), as it happens in the INSPIRE profile (INSPIRE, 2010). This limits the effectiveness of geospatial data discovery because of the limitations of searching words in free text instead of matching concepts. Therefore, the fact that the same word may have different meanings in different domains, languages, or contexts makes the search error-prone.

Moreover, existing geospatial catalog solutions do not address the issue of searching and retrieving resources from the emerging world of Web 2.0 services. In such a context, new content and resources are continuously created and made publicly available by thousands of users. Such an amount of information cannot be ignored when designing and developing a modern SDI.

The conceptual challenge behind the above issues concerns the enablement of new ways of searching in present SDIs. To address this, we define a *Discovery Augmentation Methodology*. This methodology is essentially driven by the idea of enriching the searchable information that is associated with geospatial resources. In fact, whether we consider semantic information or user-generated content, the current capabilities of geospatial catalog solutions need to be enhanced for enabling searches on additional content.

Next section introduces three high-level approaches for discovery augmentation taking into account not only semantics and Web 2.0, but with a more general perspective in order to address possible future needs. In the following two sections, we analyze in depth the User-based Approach and the Third-Party Approach. For the former, we describe a conceptual architecture and the main components centered on the integration of user-generated content in SDIs. For the latter, we describe an architecture based on a third-party approach to enable semantics-based searches in SDIs. The chapter is finished with some concluding remarks.

DISCOVERY AUGMENTATION METHODOLOGY: HIGH-LEVEL APPROACHES

Enhancing geospatial resource discovery capabilities can be achieved by augmenting the searchable descriptions of resources. Examples of additional descriptions (that is, something not searchable with typical geospatial discovery services) are: semantic information and user-generated annotations. Searching for geospatial resources that are constrained by one or more of the previous descriptions implies to characterize such resources properly. Several approaches can be followed in order to cope with the required additional descriptions. Roughly, these can be classified in:

- Provider-based approach
- User-based approach
- Third-party approach

Each of the above approaches has advantages and drawbacks. To analyze them in the next sections it is important to underline which are the main issues to be addressed by a discovery augmentation solution, in particular: interoperability, accuracy, and extensibility.

Interoperability is critical for the geo-information domain, where large and distributed infrastructures must manage resources from different and heterogeneous scientific domains. A good solution for discovery augmentation must then be able to address the different interoperability issues related to interconnecting resources in such a heterogeneous environment. Secondly, Accuracy of search results is central to all discovery systems; it should be as high as possible also when discovery capabilities are augmented. An important aspect that impacts on accuracy is metadata quality because poorly documented resources (e.g. a wrong translation of a term indicating its semantics) may lead to lower precision-recall values. Finally, extensibility, which is more related to the augmentation approach. As stated in the introduction, a general approach should be able to address possible future needs and not only the present need raised by semantics interoperability and Web 2.0 resources. A general augmentation approach that is relatively easy to extend for satisfying new requirements is highly desirable.

Provider-Based Approach

This approach represents the straightforward solution. In this case, data providers enrich resource metadata by adding related semantic information based on controlled vocabularies or even on full ontologies. In fact, this approach consists of making explicit as much as possible additional metadata description for each resource. Clearly, this approach is provider-based since the characterization of resources with new searchable content is completely entrusted to the provider.

The main advantage of this approach consists of the high accuracy in the description of resources. In fact, they are directly supplied by the resource providers and then allow agents to execute queries against the additional description. Thus, an authoritative, quality description can be ensured.

On the other hand, it is not always possible to apply this approach. Resources, especially for global systems, are maintained by several providers at different stages. Two providers, one the role of creator and the other as a custodian, should generate additional descriptions in a consistent way. This implies extra synchronization tasks among diverse providers involved in the life cycle of resources, which is not always possible. This might be a quite complex and expensive operation, since the continuous changes in geo-information domain would require constant updates of large repositories. In conclusion, this approach does not result an extensible solution.

User-Based Approach

This approach moves the task of augmenting the searchable resources from data providers to data users. From this perspective, we address the discovery augmentation in two different ways: a) augmenting information sources by extending current systems in order to access Web 2.0 services, which are typically based on user-generated content; and b) by delegating to users the description of resources by adding the so-called *resource annotation* capability to resource sharing systems or providing mechanisms that allow users to annotate resources once they are found.

This approach distributes the task of enriching large repositories of metadata to a wide range of users, scaling and making potentially use of a much higher amount of knowledge. Allowing users to annotate resource with new meta-information highly increases the available knowledge; however, this raises the challenge of metadata quality control. This is still an open issue not only in the geospatial domain (Stvilia, et al., 2008), but it assumes great importance in this domain as scientific data must be described in a proper way to support scientific and decision-making applications (Craglia, et al., 2008).

Third-Party Approach

The main principle of the third-party approach is to *build on existing systems*, a widely applied concept in creating SDIs based on the System of Systems approach (Global Earth Observation System of Systems[5]). The idea is that existing systems continue to operate within their own mandates, because additional capabilities are provided by new components that interconnect with existing systems generating added value, for instance, additional meta-information to perform searches.

In our case, existing systems are: a) currently available discovery services and b) repositories of additional meta-information such as controlled vocabularies, ontologies, user-generated annotations, etc. Following this approach, a third-party component is in charge of classifying existing resources according to available meta-information.

Clearly, this approach is not as accurate as the provider-based one. Indeed metadata quality is still ensured by the data provider but the automatic classification required to elicit additional information (e.g. semantic information) may be inaccurate in some cases. However, this approach allows to characterize resources with proper additional meta-information (provided with a basic set of metadata) stored in existing archives and repositories, without any modification. Moreover, it does not prevent from adding explicit new meta-information on existing resources.

Besides, another advantage of this approach consists of being able to accommodate future needs in a relatively easy way, i.e., it is extensible. In fact, the business logic necessary to classify resources is concentrated in a separate component (provided by a third party) that can be adapted to satisfy new requirements without affecting the other existing systems.

AN ARCHITECTURE TO INTEGRATE AND ANNOTATE USER RESOURCES

With the emergence of the Web 2.0, ordinary citizens have begun to produce and share geospatial information on the Internet. These Web 2.0-based geospatial activities show that users are willing to engage more actively in the production and provision of content. The aim of this section is to describe how web services offering volunteered geographic content can be considered to augment the number of data sources in Geospatial Information Infrastructures (GII) (Díaz, et al., 2011).

Data contained in GIIs are usually produced and maintained by the official providers, like scientific or government institutions that guarantee data quality and completeness. These providers also register metadata descriptions of data and resources in standard catalog services. In this way, common tasks of GII users (like data search, discovery, and evaluation for a particular purpose) are performed against these catalog services, since they contain metadata descriptions that should point to the resource itself or the data service serving it.

Web 2.0 services are meant to be easily accessible, via specific web sites or APIs, in order to integrate them in different applications. The contents of Web 2.0 services are mainly user-generated and are being continuously updated by non experts. There is then a lack of authoritative indicators regarding completeness, accuracy, or even veracity of data, but on the other hand, it can be easily rated, improved, and updated by users. Due to the easy deployment and publication mechanisms, the rate of participation is high and then resources are also quite up-to-date. In contrast, GII publication mechanisms are still complex and users do not have the knowledge to publish easily new content.

GII is built on services implementing OGC standard interfaces. Service interoperability is reached by applying the same standard interfaces to the different components deployed in the infrastructure to be used in as many use cases as possible, at the cost of certain abstraction level and format complexity. The use of OGC standards is then beneficial in terms of integration and interoperability, though, these can be rather complex when compared to Web 2.0 services, which are built upon simple application-level protocols (the so-called APIs) and lightweight data formats relying on simplicity, ease of implementation, and fast adaptation to user's needs. Open standards are used where they serve the keep-it-simple principle. Each service offers different functionalities, so each one provides its own public API. Some common operations (such as searching and accessing content) could benefit from simple standards as GeoRSS[6], GeoJSON[7], or KML (OGC, 2008), thus increasing interoperability to some extent.

An application of the user-based approach for discovery augmentation is the implementation of user annotation or tagging, which covers a descriptive perspective in terms of discovering resources through user-generated tags close to the content of resources. There are user-centered techniques that may improve the discovery experience from the user perspective. A recent study (Strohmaier, et al., 2010) differentiates between users who use tags for categorization and those who use tags for description purposes. The first group of users is motivated toward tagging because they want to construct and maintain a navigational aid for the resources being tagged. This implies a limited, stable set of tags that represent categories. As the tags are very close to the hierarchical, structured representations of certain models (e.g. forestry, environment, floods, etc.) they can act as suitable facilitators for navigation and browsing.

In the user-based approach, we propose a couple of techniques around the concept of tagging:

- Augmenting the discoverable sources: Users define tag-based search queries that expand over a great range of heterogeneous sources, augmenting the range of potential discoverable sources.
- Annotating the discovered resources: After discovery, when a user annotates and aggregates one or more search results into a collection to improve future searches.

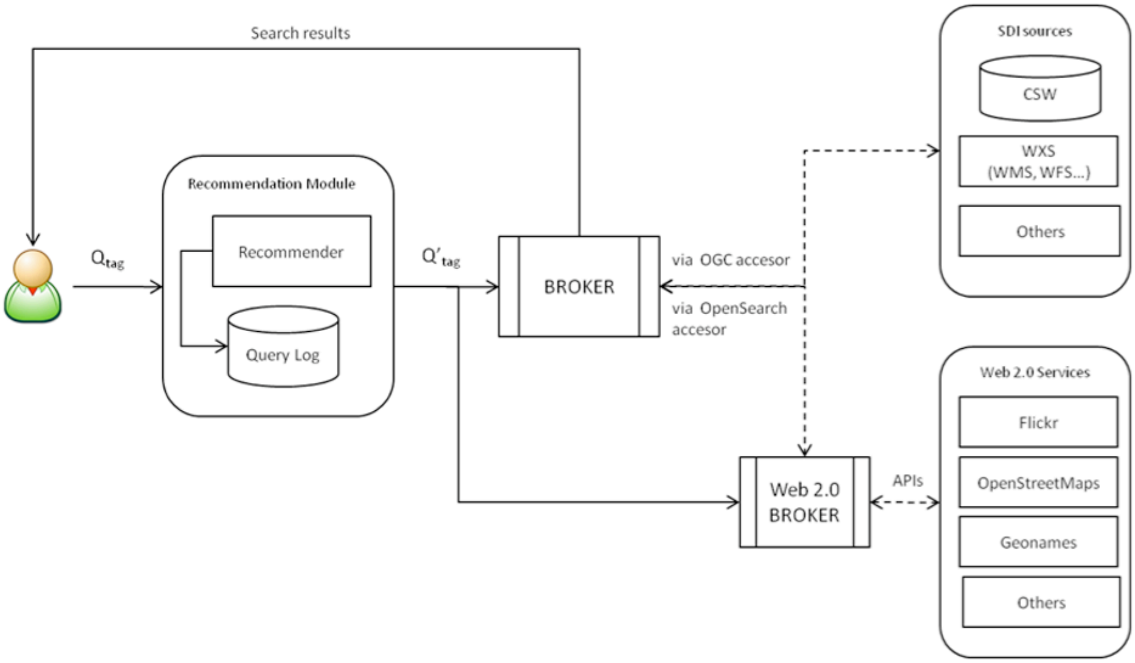
Augmenting the Discoverable Sources

The first architecture consists of exposing a simple common query interface and common response formats for several discoverable sources. These backend sources represent Web 2.0 services, which in principle contain distinct types of resources, both in nature and format. In order to provide interoperability across these services, increased data accessibility, and make client implementation much easier, we proposed the use of a Web 2.0 Broker (Nuñez, et al., 2011) that enables a simple common search interface to query the set of Web 2.0 services. In this context, tags play an essential role in this architecture by allowing users to discover heterogeneous resources from different sources.

Figure 1 illustrates the proposed architecture for tag-based discovery and the augmentation of discoverable sources via the Web 2.0 Broker integrated with the EuroGEOSS Discovery Broker (EuroGEOSS, 2010). This architecture is supported, at least, by the following components:

- Recommendation Module, which seeks recommended tags from previous user's queries.
- Web 2.0 Broker, which forwards the user's query over a large set of Web 2.0 services.

Figure 1. Tag-based discovery and sources augmentation via the web 2.0 broker



In a normal discovery scenario, the user starts typing the tags that she thinks best describe the information she is looking for. The Q_{tag} in Figure 1 denotes the initial user query in terms of typed tags in an unrestricted manner. In the following, we describe the main components involved in a tag-based query (Q_{tag}) augmented over discoverable resources.

Recommendation Module

The functionality of this module is to help users choosing the tags to improve search accuracy, on the basis of historical tags used in previous queries. The Recommender component takes a tag (Q_{tag} in Figure 1) and provides a list of tags (Q'_{tag}) related to the input tag. Its aim is to propose related tags in function of the actual tag by analyzing historical searches. The idea is to find correspondences between tags that, although syntactically different, could be semantically related. For instance, if a user types the tag "fire," this component should recommend other tags that were used in previous queries and have been related somehow with the actual tag, the strategy could consider synonyms that have been previously highly ranked, highly used, etc. In this sense, we benefit from the previous tags used by users to perform similar queries, that is, exploiting the social knowledge in terms of tag clouds built by the users. Recent works (Barragáns-Martínez, et al., 2010) are exploring with success the use of social tags to improve recommendations to users, other works like (Vilches-Blázquez & Corcho, 2009) disambiguate between syntactic terms to find semantic correspondences in known taxonomies. On the other hand, the Query Log component will store and keep track of the different tag-based queries performed by users. This allows us to improve the quality and accuracy of further tag-based queries.

The Recommendation Module will augment the original query Q_{tag} into an extended query Q'_{tag} ready to be consumed by the main Discovery Broker or directly by the Web 2.0 Broker. We also identify some challenges in implementing this module. One for example is to deal with multilingual tags. Another challenge is to identify the target fields in the ISO metadata records used to match the tags. Current metadata records are not annotated with user tags as other Web 2.0 services are.

Web 2.0 Broker Component: OpenSearch as Integration Protocol

As mentioned, Web 2.0 services offer different contents and functionalities and provide their own public API. In our approach, we look for a mechanism to access common functionalities (search interface, geographic content data type) by offering a common entry point and experimenting with simple open standards.

OpenSearch (Clinton, 2009) and its geo extension (Turner, 2010; Gonçalves, 2010) is proposed as the minimal query interface that can be used to access geospatial content, actually for both Web 2.0 services and SDI services. In this sense, OpenSearch becomes the "common query interface" to query the set of Web 2.0 services (Figure 1). This would allow for easy client implementations that could search and integrate all data sources regardless of their origin. In addition, the use of the OpenSearch as a common interface to access to all Web 2.0 services greatly alleviates the integration with the Discovery Broker. This approach implies that the Web 2.0 Broker is designed as a set of adapters that transform and propagates the original OpenSearch query to the different Web 2.0 APIs. The Web 2.0 broker encompasses all the adapters for the selected services for the integration scenario, which for an initial prototype would be: Twitter, Flickr, OpenStreetMap, and Geonames.

Annotating Discovered Resources

Complementary to the previous approach, where the Web 2.0 Broker component augmented the set of discoverable sources, the following approach is based on the use of annotation techniques around the concept of collections. This technique makes use of the OAI-ORE[8] (Open Archive Initiative—Object Reuse & Exchange) abstract model to compose and annotate collections of resources that are of interest to the users, along with the use of a RDF-based repository to persist such collections. In this context, tags play an essential role in this architecture in letting users annotate collections of heterogeneous resources from different sources.

Figure 2 illustrates the proposed architecture for the annotation of collections of search results (discovered resources) through the set of brokers. Note that this architecture may be an extension of the previous, to provide added-value functionalities over the previous tag-based discovery approach. For this reason, some components are shared by both approaches, like the Recommendation and Web 2.0 Broker components:

- Recommendation Module and Web 2.0 Broker.
- Annotation Module, which let users annotate collections of resources and then publish them in a repository.
- Repository Module, which persists resource collections as RDF triples.

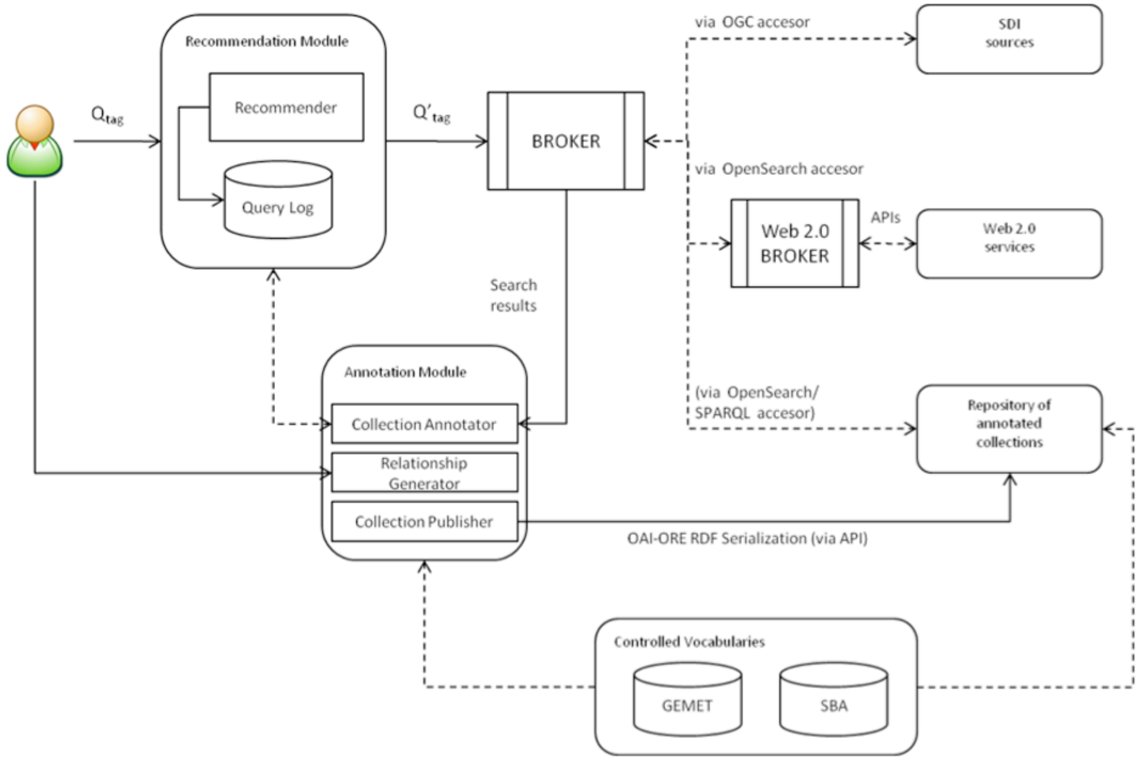
Before going into the description of the Annotation and Repository modules, we describe in the following the basic notions and entities behind the OAI-ORE specification.

OAI-ORE Specification

The Open Archive Initiative - Object Reuse and Exchange (OAI-ORE) protocol (OAI, 2008) defines an abstract data model (OAI, 2008b) for describing, reusing, and exchanging collections of Web resources. The OAI-ORE protocol is initially conceived of in the context of digital libraries and e-print resources, in order to expose rich content (text, images, data, video, etc.) into aggregations to be then reused by client applications.

Conceptually, the OAI-ORE's abstract data model builds strongly on the notion of "address-

Figure 2. Tag-based discovery and annotation of collections with the broker

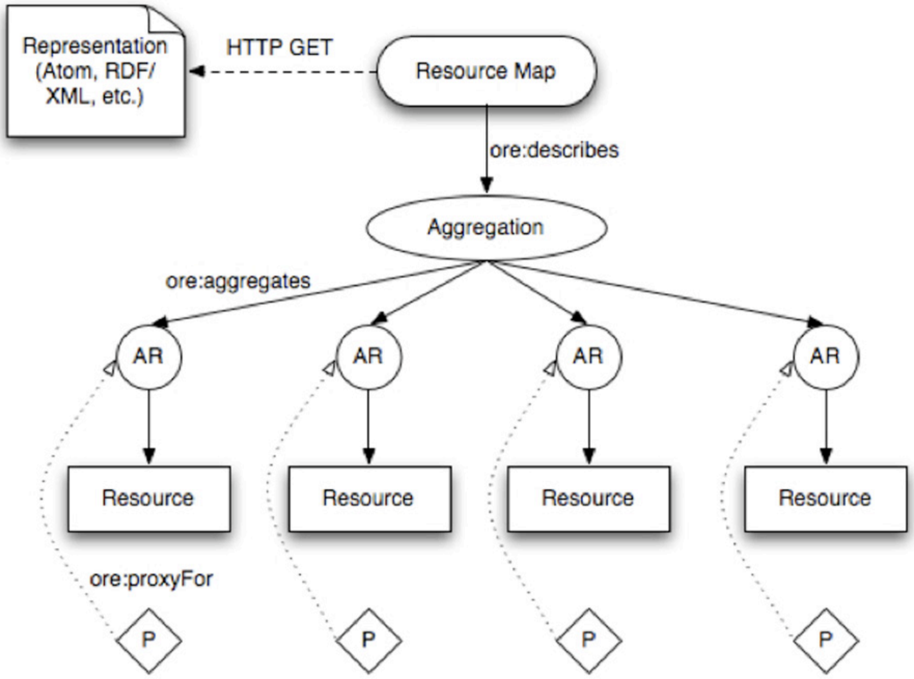


able resources" to indicate that any resource (file, image, text document, metadata, process, etc.) is identified using HTTP URI (Uniform Resource Identifiers[9]). The simplified diagram in Figure 3 shows the main entities that form part of the OAI-ORE's abstract model. The entity *Aggregation* plays a central role as it represents a collection of addressable resources that in turn are called *Aggregated Resources* (AR). The *ore:aggregates* relation denotes the inclusion of related resources in the same collection.. In addition, both *Aggregations* and *AR* entities are addressable resources in the sense that both use HTTP URI as referencing method, i.e., which just means looking up a URI on the Web in order to get either the resource itself or its representation.

Aggregation and *AR* are abstract entities that must still refer to concrete resources, which can be discovered resources of any type and even a chain of geospatial services (Abargues, et al., 2010). The OAI-ORE specification makes use of the *Resource Map* entity to provide a concrete representation for the whole aggregation, mostly derived from RDF. Some suggested formats in the specification are the Atom syndication format[10], RDF/XML[11], and RDFa[12] (a microformat for extending XHTML to support RDF).

Finally, OAI-ORE also defines a useful abstract entity called *Proxy* (P) by which it is possible to express the role an aggregated resource has explicitly in the context of an aggregation. The intended meaning of the role is expressed via *ore:proxyFor* relationships. For example, two resources may have a temporal relationship that connects to each other and this is only meaningful within the aggregation context in which they are defined.

Figure 3. Simplified OAI-ORE's abstract data model



Annotation Module

Annotating search results is a means to refine future searches. The use of collections (aggregations) of resources combined with user knowledge can improve the discovery of these concepts and others related to them. Currently, there exist different techniques to model collections or aggregations of heterogeneous resources. After some analysis, we selected the OAI-ORE paradigm because it defines a straightforward abstract data model for the description and exchange of aggregations/ collections of Web resources. The resources that compose an aggregation are identified by their URIs. Although OAI-ORE specification does not specify a concrete serialization format, RDF, and Atom are the preferred ones.

The aims of the Annotation Module is to allow the user to group different search results (resources) into OAI-ORE-based collections, annotate them, and finally publish all this information in a repository. Several components are required to carry out these tasks. The Collection Annotator (Figure 3) component allows users to create a collection from a set of aggregated resources. Such resources are the result of a search made by the Discovery Broker. An interesting point here is that, by using the OAI-ORE abstract model, the collection is no further a plain list but a graph or map of resources with typed links and established relationships among them and even with entities outside of the collection (i.e., Geonames[13] or DBpedia[14] entries). Users should interact with the Annotation Module to discard and select only those resources that better fits their requirements to form a collection. At this point, the user can add extra metadata (tags, terms) to the aggregated resources but also and most important to the aggregation itself. Tags at collection-level may describe common features of the aggregated resources and meaningful information that makes only sense as a whole.

Building a collection amounts to grouping a list of resources with regard to some links or relationships. In a first moment, maybe only a basic set of relationships can be set, such as those specifying the internal structure of the aggregation, and some others regarding the resources' metadata obtained in the search. These types of relationships could be incremented by using the Relationship Generator component (Figure 2) that generates new ones based on the geometric and temporal topology that the different resources exhibit among them. These new relationships could be generated automatically by gathering the resource's information collected by the broker, and by calculating the different relationships using for instance ontologies or vocabularies that could formally specify them, similarly to the ones used currently by the Ordnance Survey[15]. Terms from controlled vocabularies and taxonomies might be also used at this stage to identify aggregated resources and typed links.

As the OAI-ORE defines an abstract model, the newly-created collections need to be serialized in a concrete format. The Collection Publisher component serializes an abstract collection into a RDF or Atom representation to be added in suitable repositories.

Repository Module

The Repository Module, in right side of Figure2, will make available all the information, structured and annotated as collections, to the user. This module will be composed by a specialized RDF storage system commonly known as "triple store"[16]. These systems usually offer a SPARQL endpoint as a way to access and query the underlying repository.

As the Discovery Broker supports OpenSearch-interfaced sources (accessors), the OS2SPARQL component will address the translation of OpenSearch-based queries into SPARQL syntax. The OpenSearch interface ensures the communication between the broker and the RDF-aware repository. As the collections of resources can appear in the search results, the contained resources may point to other related resources, enriching then the search experience. Nevertheless, a dedicated SPARQL-interfaced accessor would be preferred to grant access to a great amount of public RDF triples available elsewhere.

SEMANTIC AUGMENTATION

Grounding Semantics-Aware Discovery

Organizing the concepts that web contents (and, among them, geospatial resources) are referring to is essentially what the Semantic Web (SW) is all about. Despite the immense work done in this research field in the last decade, most of the design principles driving the development of the SW can be found in the enlightening book, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor* (Berners-Lee & Fischetti, 1999). Since contents on the Web are constituted by (or at least annotated by) text, the only means for indexing them has traditionally been text-matching criteria. Of course, hyperlinking of web resources has been playing a predominant role in the ranking of search results (Brin & Page, 1998); however, hyperlinking functionalities are not available in all formats currently delivered by the Web and, particularly, not in the spatial data sets and services we are addressing, nor in the metadata annotating them. Moreover, text-matching techniques are characterised by a number of shortcomings that make them inefficient; among these, the more relevant are related to homonymy, synonymy (biasing, respectively, precision, and recall in the discovery process) and of course the difficulty of carrying out searches in a multilingual context.

In a nutshell, the SW consists of *statements* that relate resources to each other or, alternatively, relate resources to literal values (e.g., the string "foo"); statements are defined by specifying a *subject*, a *predicate* (or *property*) and an *object* according to the RDF data model (W3C, 2004), a *triple* in RDF jargon. Resources are singled out by using URIs, which provide a straightforward means to provide unique identifiers to complex entities. As an example, consider the sentence "chapter 'Methodologies for Augmented Discovery of Geospatial Resources' is included in the book *Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications*." Apparently, the sentence is defining a single statement relating a chapter to a book. Instead, if we want to unambiguously define the resources involved, we need to specify the URIs representing them, increasing the number of necessary statements. A set of sentences for expressing this may be the following:

- chapter <http://example.com/chapter123> is entitled 'Methodologies for Augmented Discovery of Geospatial Resources
- book <http://example.com/book456> is entitled *Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications*
- resource <http://example.com/chapter123> is included in resource <http://example.com/book456>

As customary in RDF triple representation, URIs are distinguished by enclosing them in angle brackets. Even in the extended set of sentences, some of the semantics conveyed by the original one is still not formalised. Specifically, the information that the two resources are, respectively, a book chapter and a book is missing; also, properties "is entitled" and "is included in" do not explicitly refer to entities that an automated agent may understand. The former can be expressed by using the type predicate of RDF for referencing elements from some widely acknowledged data schema for structuring publications (here, we consider the chapter and book elements of the DocBook format[17]). The titles associated with the two resources can be rendered by specifying the title Dublin Core metadata term[18]; instead, for expressing compositions of chapters in a book we can leverage on predicate *aggregates* from the OAI-ORE standard (OAI-ORE, 2008), that has been introduced earlier in this chapter. The resulting triples are shown in Figure 4.

Lists of triples are quite verbose and not apt to human consumption; a more convenient representation format is that of *directed labelled graphs*: In this formalism, ellipses represent resources (note that, in the SW, everything that is not a literal value is a resource) whose identifiers are often shortened by substituting a long *namespace* (e.g., "http://example.com/") with prefixes (e.g., "ex:"). Rectangles represent literal values, such as the string 'Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications.' Finally, arcs represent predicates. Figure 5 is showing the graph representation of the triples in Figure 4. Despite its simplicity, the RDF model is capable of representing very complex data structures, such as n-ary relations and statements about statements (that is, considering a whole triple as subject or object of a statement) (see Figure 5).

The graph representation of RDF makes apparent one of the main strengths of the data model, that is, its unstructured nature that makes it possible to seamlessly aggregate information (statements) from heterogeneous sources. However, RDF is just the medium for expressing information on the SW; for structuring it as specific data models, more standards were layered on top of RDF, with different degrees of expressive power and computational capabilities. These will be briefly introduced in the following section.

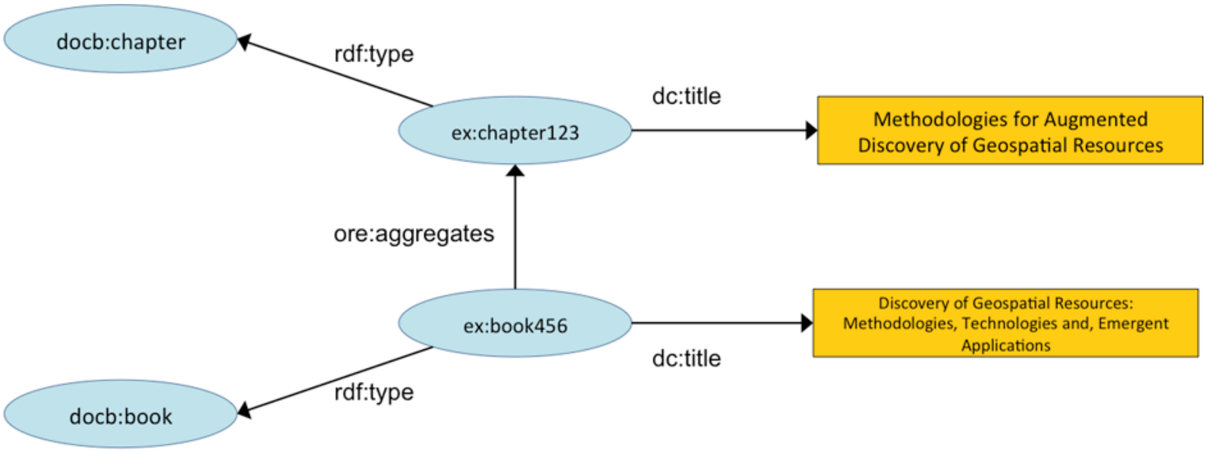
Representing Domain Knowledge

An aspect of data modelling that the basic RDF data model cannot handle is characterizing the entities that are referred to by triples. As an example, we

Figure 4. Example of RDF triples

```
<http://example.com/chapter123>

    <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
    <http://www.oasis-open.org/docbook/xml/4.2/docbookx.dtd#chapter>

<http://example.com/book456>
    <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
    <http://www.oasis-open.org/docbook/xml/4.2/docbookx.dtd#book>

<http://example.com/chapter123>
    <http://purl.org/dc/terms/title>
    ‘Methodologies for Augmented Discovery of Geospatial Resources’

<http://example.com/book456>
    <http://purl.org/dc/terms/title>
    ‘Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent
Applications’

<http://example.com/book456>
    <http://www.openarchives.org/ore/terms/aggregates>
    <http://example.com/chapter123>
```

Figure 5. RDF graph induced by the statements in Figure 4


docb:chapter
rdf:type
dc:title
ex:chapter123
Methodologies for Augmented Discovery of Geospatial Resources
ore:aggregates
ex:book456
Discovery of Geospatial Resources: Methodologies, Technologies and, Emergent Applications
dc:title
rdf:type
docb:book

may need to characterise *book456* as a scientific book (that is, ascribe the resource to a specific *class* of resources) and not, for instance, as a novel, as a cooking manual, etc. Still we may want to retain some of the characteristics that all books are sharing, such as the editor, the number of pages, the selling price, etc. In a word, we would like to structure data in a more object-oriented fashion. The schema languages that were defined for the SW, such as the RDF Schema (RDFS) (W3C, 2004b) and the Web Ontology Language (OWL) (W3C, 2004c), are addressing this requirement; schemata defined through the language primitives defined by RDFS and OWL are typically referred to as *ontologies*. Another desirable feature is the capability of structuring predicates and defining relations among them. As an example, OAI-ORE defines predicate *isAggregatedBy* for expressing part-whole relations (not the whole-part relations expressed by predicate *aggregates*) but the relation between these (one being the inverse of the other) is not made explicit. The RDF/XML serialisation of an OAI-ORE resource map containing the last triple in Figure 4 may, or may not, feature also the following triple:

- <http://example.com/chapter123>
- <http://www.openarchives.org/ore/terms/isAggregatedBy>
- <http://example.com/book456>

In fact, one of the main advantages of SW data formats is that an RDF graph may convey more information than that explicitly stated by the graph itself. This is due to the semantics underlying the schema languages, which defines the *entailments* (that is, the logical implications) that shall hold. These may be given at different formalization levels; for example, RDFS provides a set of rules indicating exactly which inferences shall be supported by RDFS *reasoners* (the automated agents deriving implications). Instead, most OWL sub-languages (e.g., OWL-Lite, OWL-DL) reflect the expressive power of some (description) logics (Baader, et al., 2003) and then do not define entailments as a finite set of production rules. RDFS and the different flavors of OWL have a well-defined expressiveness. As an example, RDFS is expressive enough to arrange classes and predicates in a hierarchal way (e.g., for defining class *ScientificBook* as a specialization of class *Book*) but OWL expressiveness is required to express that predicates *aggregates* and *isAggregatedBy* are one the inverse of the other. Because of the associated inference capabilities, data sources that are defined by SW schema languages are typically referred to as *Knowledge Bases* (KB) (Russell & Norvig, 2005).

A prominent example of OWL ontology in the geospatial domain is the schema grounding the GeoNames geographical database[19]. Queries to the GeoNames web service can specify the parameter *type=rdf* for obtaining results encoded as RDF according to the schema defined by the service. As an example, Figure 6 is showing a fragment taken from the RDF response retrieved from the service for the search pattern "france." On the basis of the associated schema, a reasoner can derive that France (represented by the URI http://sws.geonames.org/3017382) is a *parentFeature* of Île-de-France because predicate *parentCountry* is a sub-property of the former. Another interesting feature of GeoNames is that it also provides data as LinkedData: this means that, by accessing any of the URIs in a query response (such as the one corresponding to France) a user agent may retrieve the RDF data fragment corresponding to the resource. This feature allows clients to selectively navigate the huge RDF graph represented by GeoNames data without requiring the download of the whole resource.

Other than creating data schemas that can be instantiated by individuals, such as *ex:chapter123*, SW schema languages can also be used for the sole purpose of structuring domain knowledge (that is, defining classes and predicates) with only few or no *individuals* (members of a class) instantiating them. This approach allows for a high

Figure 6. Fragment of the RDF output of the GeoNames web service

```
<gn:Feature rdf:about="http://sws.geonames.org/3012874/">
        <rdfs:isDefinedBy>http://sws.geonames.org/3012874/about.rdf</rdfs:isDefinedBy>
        <gn:name>Île-de-France</gn:name>
        <gn:alternateName xml:lang="fr">Île-de-France</gn:alternateName>
        <gn:alternateName xml:lang="en">Île-de-France</gn:alternateName>
        <gn:alternateName xml:lang="es">Isla de Francia</gn:alternateName>
        <gn:alternateName xml:lang="de">Île-de-France</gn:alternateName>
        <gn:featureClass rdf:resource="http://www.geonames.org/ontology#A"/>
        <gn:featureCode rdf:resource="http://www.geonames.org/ontology#A.ADM1"/>
        <gn:countryCode>FR</gn:countryCode>
        <gn:population>11598866</gn:population>
        <wgs84_pos:lat>48.5</wgs84_pos:lat>
        <wgs84_pos:long>2.5</wgs84_pos:long>
        <gn:parentCountry rdf:resource="http://sws.geonames.org/3017382/"/>
        <gn:childrenFeatures rdf:resource="http://sws.geonames.org/3012874/contains.rdf"/>
        <gn:locationMap rdf:resource="http://www.geonames.org/3012874/region-ile-de-
france.html"/>
</gn:Feature>
```

expressiveness when defining *controlled vocabularies* (i.e. *thesauri*) for modelling a given domain. In the context of SDI-related thesauri, this approach is exemplified by the Semantic Web for Earth and Environmental Terminology (SWEET) (NASA-JPL, 2011), a set of 201 highly modularised OWL ontologies whose terms are expressed as Class instances and arranged hierarchically by using predicate *subClassOf* from RDFS. The lexical representation of terms is constituted by the fragment identifiers of URIs (e.g., "#ClimateChange"). The exploitation of constructs from the meta-language (OWL) allows for defining custom predicates that class instances may have and constrain ontology constructs according to them. As an example, Figure 7 is portraying the definition of class #Mesoclimate whose members are defined as either instances of super-class #Climate and also instances whose property #hasSpatialScale takes values from class #Mesoscale.

Developing Thesauri for SDIs

The discovery of geospatial data and services on the Internet is a topic that, despite the hype nourished by the novel notion of SDI (Masser, 2005), still lacks efficient technologies and techniques for enabling an effective retrieval of resources. The non-textual nature of this category of resources, the scarce set of metadata that is often annotating them, the multilingual gap (and in general the linguistic issues hindering the indexing of all categories of resources) motivates the recourse to specialized applications, *geoportals* (Maguire & Longley, 2005), for retrieval of spatial data and services. Even when an exhaustive set of metadata is provided, the distance between the terminology adopted by the domain expert during annotation and that of the end user (which can be herself a thematic user, but coming from a different domain) makes it difficult to reconcile the metadata descriptions on the one hand and the search patterns on the other. Also, specifying metadata in

Figure 7. Definition of class #Mesoclimate in the SWEET ontology

```
<owl:Class rdf:about="#Mesoclimate">
      <rdfs:subClassOf rdf:resource="#Climate"/>
      <rdfs:subClassOf>
            <owl:Restriction>
                  <owl:onProperty rdf:resource="&scale;#hasSpatialScale"/>
                  <owl:allValuesFrom rdf:resource="&scale;#Mesoscale"/>
            </owl:Restriction>
      </rdfs:subClassOf>
      <rdfs:comment xml:lang="en">The climate of a natural region of small extent, for
example, valley, forest, plantation, and park. Because of subtle differences in elevation
and exposure, the climate may not be representative of the general climate of the
region.</rdfs:comment>
</owl:Class>
```

a language that is different from that of the end user may utterly prevent retrieval of a resource. Albeit easing access to geospatial information, geoportals often fall short of providing advanced search functionalities that allow the end user for bridging this gap.

Acknowledging these shortcomings, the SDI community is increasingly considering recourse to thesauri as one of the most promising means to annotate resources in a consistent way and to ease retrieval of this category of resources in a multilingual and cross-thematic context. In particular, thesauri based on SW formats provide language-neutral identifiers (URIs) that can be related with distinct syntactic forms in order to be selected for annotation through multiple, language-dependant textual representations. Also, the internal structure of thesauri (relating terms with each other according to specificity and relatedness) allows for advanced functionalities such as query refinement on the basis of semantic properties among terms. Finally, by aligning independent thesauri produced in distinct thematic areas, it is possible to easily bridge between different domains. Using SW schema languages for articulating thesauri, such as for the SWEET ontologies, may be awkward to domain experts because of the inherent technicalities. Moreover, applying inference to large schema definitions may be particularly expensive on the computational side.

Instead, a different approach that is gaining wide acceptance is constituted by relying on a fixed OWL schema and articulate thesauri by instantiating the classes and predicates defined by the former. The clear advantage with respect to the previously described approach is the capability of bounding the complexity of data structures on the basis of the underlying schema. As for the expressiveness, it is depending on the specific data schema that is taken into consideration. The research on *Knowledge Organisation Systems* (KOS) (Bechhofer & Goble, 2001) is focusing on the development of thesauri organizing terms in a coherent way. The intuition this research elaborated on is that it is possible to take advantage of the Semantic Web infrastructure in a lightweight, easily implementable way. The Simple Knowledge Organisation System (SKOS[20]) constitutes the main output of this research thread: it is a low complexity ontology that allows to easily structure terms (also providing translations into multiple languages), to group them into collections of terms and to relate terms from independent thesauri to one another.

The Simple Knowledge Organisation System

SKOS represents a breakthrough in knowledge organisation inasmuch it provides an exhaustive set of constructors for creating thesauri without involving the technicalities of more expressive ontology schema languages. Making a comparison with XML technologies, it is like exploiting an XML custom vocabulary for expressing a given category of data structures versus dealing with XML Schema constructs for expressing them. Figure 8 is showing a general view of the entities defined by SKOS, as rendered in the Protégé ontology editor[21] (the names in bold building on the entities of OWL and RDFS). Class *ConceptScheme* is used for creating a whole thesaurus; *Collection* and *OrderedCollection* allow to group terms; finally, class *Concept* represent individual terms in the thesaurus (see Figure 8).

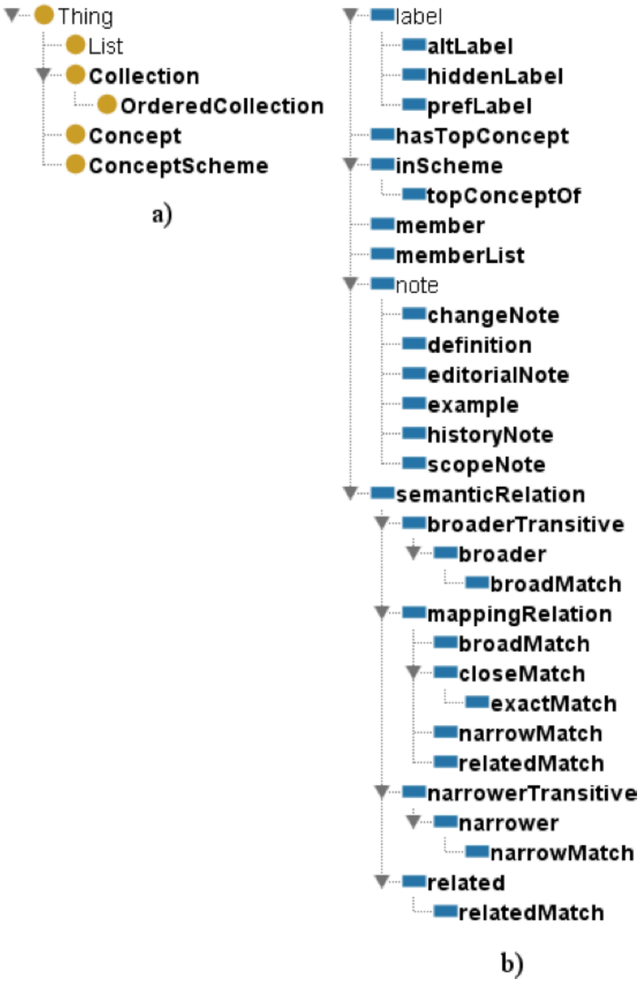
More complex is the organisation of properties. A set of properties can be used to associate *labels* (that is, human-readable text representations) to terms that are, by themselves, represented by URIs: *prefLabel*, *altLabel,* and *hiddenLabel* represent, respectively, preferred labels, alternative labels and labels that are possible misspellings of term labels. A second set of properties allow the grouping of entities in a thesaurus: properties *inScheme* and *topConceptOf* allow to associate terms (*Concepts*) with the thesaurus (*ConceptScheme*) they belong to and allow to specify, among terms, which are the more general. Finally, a set of properties can be used to categorise terms according to specificity (*broader*, *narrower*) and relatedness (*related*); their variants terminating with "Match," together with properties *exactMatch* and *closeMatch*, are used for relating terms from distinct thesauri while those terminating with "Transitive" are introduced to support the transitive closure of sub-properties but are typically not introduced explicitly because it is up to the underlying reasoning facilities to deduce them. As we will see, all categorisations of SDI-related terms considered in this Chapter are based on SKOS or on an extension to it.

SKOS Thesauri in the SDI-Domain

The development of thesauri is a key factor for geographic information retrieval and led to the inclusion in the GEO Work Plan 2009-2011 (GEO, 2010) of a sub-task of the GEOSS Common Infrastructure devoted to ontology and taxonomy development. Early adopters of SKOS as either the native or an alternative format for representing thesauri include the United Nations Food and Agriculture Organization which is maintaining AGROVOC (FAO, 2011), a multilingual resource comprising over 28.000 terms covering all subject fields in agriculture, forestry, fisheries, food and related domains. The United States National Agricultural Library is maintaining a bilingual agricultural thesaurus made up by more than 80.000 terms (USDA). In Europe, the GEMET Thesaurus (EIONET, 2011) is providing a narrower set of terms (around 6000 terms) but is covering all languages spoken in the European Union.

Also in Europe, the EU Publications Office is providing the SKOS version of EuroVoc (2011), a multilingual and multidisciplinary thesaurus covering all areas of activity of the EU. Recently, the categorization of Societal Benefit Areas (SBAs) underlying GEOSS has been encoded as SKOS and translated into four more languages in the context of the EuroGEOSS FP7 project[22]. Even if this resource can barely be considered a thesaurus, since it is featuring only 68 terms, it constitutes a good example of how resources not natively encoded as SKOS can be easily translated and extended in order to support multiple languages. Also, a large amount of data sets and services are expected to be categorized according to SBAs and, consequently, the availability of this small thesaurus, when properly interconnected with other reference thesauri, may prove very useful for resource discovery.

Figure 8. The SKOS ontology: a) class hierarchy and b) property hierarchy



Finally, several other institutions and projects are similarly digitising SDI-related domain knowledge by using SKOS as the encoding format. Starting from the CUAHSI Hydrologic Information System[23], a thematic thesaurus for water has been developed in GEOSS AIP-3 and is currently being translated into SKOS for inclusion in the knowledge base considered in this Chapter. The Australian agency CSIRO is developing vocabulary services[24] for thematic thesauri that are made accessible through an API akin to the one implemented by GEMET; among these, the extensive water thesaurus derived from the WDTF schema of the Australian Bureau of Meteorology. In general, since knowledge management has a direct application in the internal functioning of companies and institutions, more and more domain-specific thesauri are expected to be developed in the future.

With regard to the thesauri that were exploited as reference thesauri during development of the semantics-aware component that is described in this Section, the source that was taken into consideration has been the RDF repository developed in the context of the GENESIS FP7 project[25]. The repository is an instance of the Sesame framework[26], an open source Java web application for storing and querying RDF data, that is currently hosting the following thesauri in the SKOS format:

- The General Multilingual Environmental Thesaurus (GEMET[27]): 28 of the 31 languages currently provided by the EIONET portal.
- The INSPIRE Feature Concept Dictionary and Glossary: 21 of the 23 EU official languages for INSPIRE Themes, monolingual the other terms.
- The ISO 19119 categorisation of spatial data services: 21 of the 23 EU official languages.
- The GEOSS Societal Benefit Areas: 5 languages.

However, as explained in the following Section, in principle any SPARQL-compliant repository can be seamlessly integrated and also sources based on a different protocol can be easily integrated. Another important precondition for applying the semantic augmentation paradigm to the retrieval of geospatial resources was the harmonization of the distinct thesauri by aligning corresponding terms among the resources listed above.

An Architecture for Transparent Semantic Augmentation

One possible concrete application of the third-party discovery augmentation approach is to enable semantic discovery of geospatial resources. The objective is to develop a flexible framework that aims to provide users with semantics-aware query capabilities to discover geospatial resources from one (or more) traditional/standard discovery services. The semantic augmentation mechanism is transparent to resource providers because no additional meta-information needs to be explicitly added to existing metadata repositories.

According to the third party approach, the framework we designed makes use of a specific Discovery Augmentation Component (DAC) that implements the business logic required for semantic augmentation. The DAC implements a query expansion strategy according to which multiple, traditional geospatial queries are derived from a single semantic query. Existing semantic services are used to expand queries and combine the related results in a meaningful way.

In fact, the DAC is able to accept a query, expand it as multiple, semantically-related queries by accessing a customizable set of external semantic services (thesauri, ontologies, gazetteers, etc.), and finally issue them to a geospatial discovery service. The DAC combines query results in a meaningful way and sends them back to the client.

This framework realizes the Information Expert design pattern (Larman, 2005) assigning specific tasks (responsibilities) to components that have the information needed to carry out the task, making the architecture flexible and scalable. Moreover, it does not affect existing geospatial services interfaces by implementing a loosely-coupled solution. However, in order to enable this architecture, specific solutions are required in order to define the exposed interface, to address resources heterogeneity and performances issues.

This framework was implemented and tested in the context of the FP7 EuroGEOSS (A European approach to GEOSS) project[28] to provide semantics-aware capabilities to the EuroGEOSS discovery broker (EuroGEOSS, 2010; Nativi & Bigagli, 2009).

System Architecture Overview

The transparent semantic augmentation framework was designed applying the following general principles:

- Layered Architecture;
- Separation of Concerns;
- System of Systems (SoS) Approach.

The layered architecture is a well-known approach for designing complex systems (ISO, 1994), where functionalities are grouped and layered according to their abstraction level. Figure 9 shows the three layers of the proposed architecture, following the SOA taxonomy:

- **Service Consumer Layer:** This layer implements presentation functionalities, in fact it is comprised of components which implement Graphical User Interfaces (GUIs);
- **Semantic Augmentation Service Layer:** This layer is composed of service components which implement the business logic necessary to integrate semantic and geospatial services;
- **Query Service Provider Layer:** This layer contains services providing both semantic and geospatial query functionalities.

The DAC clearly falls into the semantic augmentation service layer, which makes use of the query service provider layer.

By separating the concerns, we obtain a set of "dedicated" components: each of them implementing a precise and well-defined set of functionalities. The *catalog service* component is dedicated to execute "traditional" geospatial queries. The *semantic service* provides access to semantic repositories (e.g. thesauri, ontologies, etc.). The *DAC* is in charge of executing the query expansion (accessing the semantic services) and distributes the resulting set of "traditional" queries to the federated catalog service. Finally, the *user agent* provides results visualization.

The System of Systems approach implies the capability of connecting with existing systems without any modification to their interfaces and/or functionalities. This is achieved by providing interoperability arrangements to mediate from a geospatial semantic query to a set of semantic queries and traditional geospatial queries.

The choice of the service interfaces was mainly driven by the need of being compliant with widely adopted catalog service specifications and to ensure interoperability with existing systems. For the interaction between the DAC and the catalog service, the OGC CSW/ISO AP (OGC, 2007b) (Application Profile) interface is used. Among the present OGC CSW APs, ISO AP is presently one of the most widely implemented. Besides, it is the INSPIRE-compliant catalog service interface.

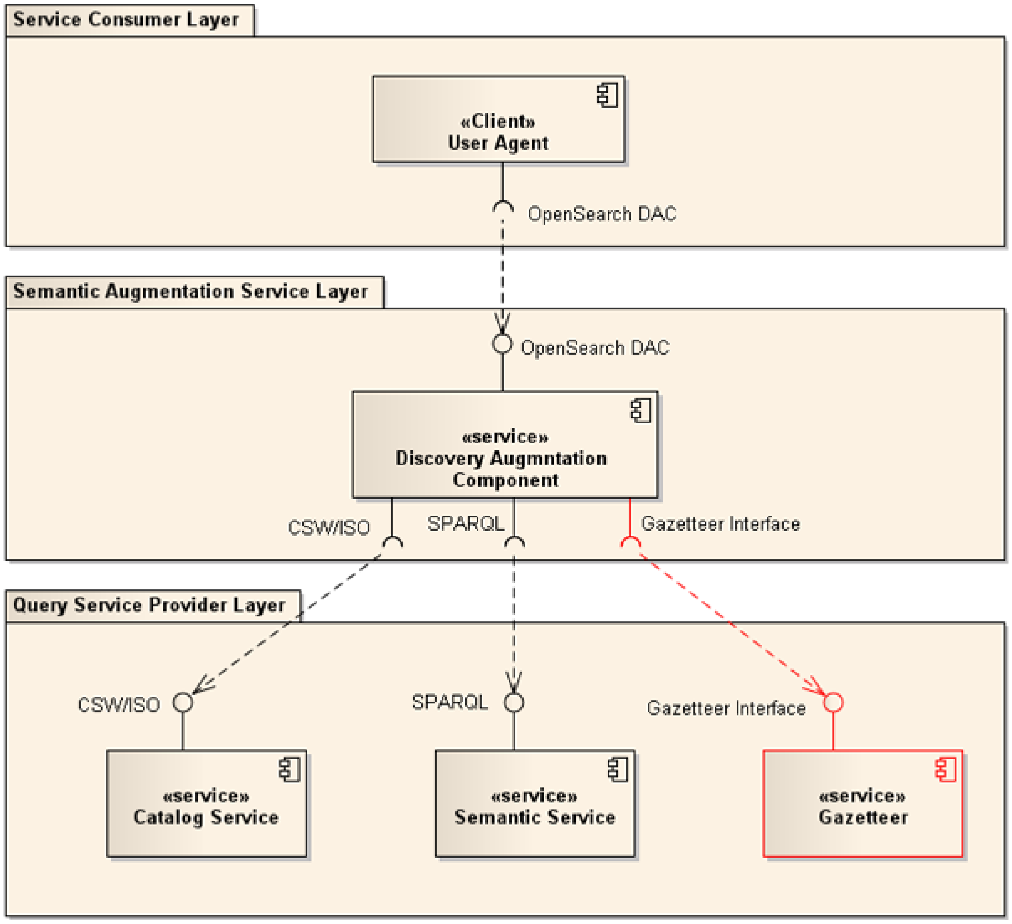
The query syntax to access the *semantic service* follows the W3C standard SPARQL (W3C, 2008). In keeping with the SoS approach, the DAC was conceived to be flexible enough to interoperate with other query languages published by different types of semantic services.

The *DAC* publishes an interface for *user agent* bindings. This interface allows clients to submit queries with a possible combination of semantic, geospatial, and free text constraints. At the time being, there is no standard interface or syntax allowing such combined queries. Hence, we decided to introduce an extension of the OpenSearch interface (Clinton, 2009). The OpenSearch is a lightweight interface that allows agents to query catalogs with a simple free text search. There exist several extensions of the basic OpenSearch syntax; two widely used extensions for geospatial queries are:

- The Geo extension (Turner, 2010; Gonçalves, 2010): this allows agents to specify a spatial extent/location as constraint in a query;
- The Time extension (Gonçalves, 2010): this allows agents to build queries based on time and time span constraints.

We introduced a new extension called: "Concept-driven extension." This allows the discovery of well-defined concepts and their relations from semantic services. The DAC query interface implements all these three extensions.

Figure 9. Architecture of the transparent semantic augmentation framework



OpenSearch Concept-Driven Extension and Query Expansion

The introduced extension is based on the concept of RDF triple (see see paragraph 'Grounding semantics-aware discovery') (W3C, 2004). Considering the basic OpenSearch interface, three additional semantic parameters can be specified in the query:

A. Subject
B. Predicate (semantic relation)
C. Object

All parameters are optional and can be combined with the free text (*searchTerm*) parameter in many different ways. The three parameters are used to represent an RDF statement that the selected concepts must satisfy. When the *searchTerm* (free text) parameter is also specified, the selected concepts are filtered to match the given *searchTerm*—that is, concepts in whose descriptions the free text (*searchTerm*) appears.

The three semantic parameters represent semantic concepts; thus, they can be represented in different ways by semantic services. When the *user agent* and the *semantic service* make use of different representations, the DAC is able to implement the necessary mediation functionality.

Examples of possible query parameters combinations are given in Tables 1, 2, 3, 4. They use the URI "http://www.eionet.europa.eu/gemet/concept/9221" from the GEMET Thesaurus for representing the semantic concept "water resource."

The implementations of the "Concept-driven extension" must support at least the GeoRSS and X-Suggestion[29] return types. GeoRSS is a recent extension of the RSS[30] feed format; it enriches traditional feeds with geographic and temporal information. When the query asks for a GeoRSS return type, the DAC builds a set of traditional geospatial queries using the multilingual text representation of the selected concepts in order to retrieve metadata records form the catalog service.

In order to create requests such as those in Table 1, 2, 3 the client software must be able to retrieve the URIs (or any other kind of representation of semantic concepts) from the semantic service. To achieve this, the X-Suggestion return type is used. When this return type is requested, the DAC selects the concepts without submitting queries to the *catalog service*. The selected concepts and their textual description are returned to the components of the service consumer layer (*user agents*) in order to be viewed by the user and, eventually, used in subsequent queries to the DAC. This feature enables the use of the DAC for browsing the content of the semantic service.

More return types can be defined and implemented for addressing specific needs. Examples of suitable return types are: XML documents for RDF graphs, ISO 19115 (ISO, 2003), metadata encodings for geospatial resources.

Using the concept-driven extension, it was possible to implement two different types of query expansion:

1. Automatic query expansion
2. User-assisted query expansion

For the first type, the main discovery steps are depicted in Figure 10.

A. The query keywords (the *what* constraint) are "expanded" with the concepts retrieved from the *semantic services* accessed by the DAC;

Table 1. Semantic query parameters, example 1

searchTerm	Subject	Predicate	Object
drought	http://www.eionet.europa.eu/gemet/concept/9221	skos:narrower	-
This query selects all concepts more specific than the concept "water resource" and whose description contain the keyword "drought"			

Table 2. Semantic query parameters, example 2

searchTerm	Subject	Predicate	Object
-	http://www.eionet.europa.eu/gemet/concept/9221	skos:narrower	-
This query selects all concepts more specific than the concept "water resource"			

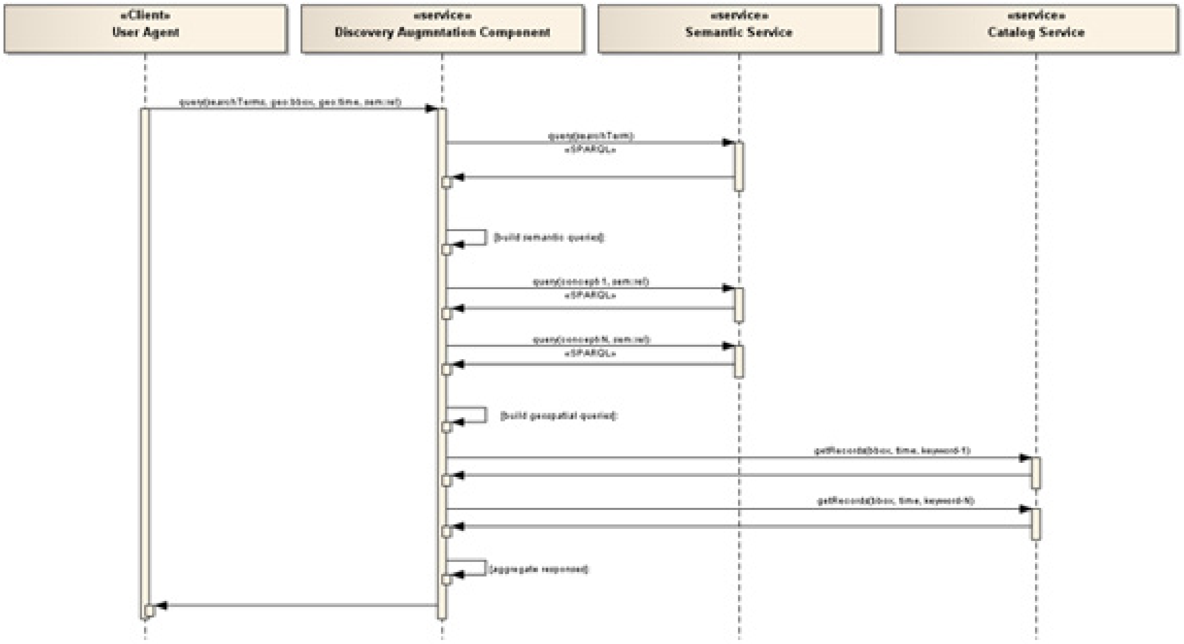
Table 3. Semantic query parameters, example 3

searchTerm	Subject	Predicate	Object
drought	http://www.eionet.europa.eu/gemet/concept/9221	-	-
This query selects all concepts which are related to the concept "water resource" according to the DAC default relation (customizable) and whose description contain the keyword "drought"			

Table 4. Semantic query parameters, example 4

searchTerm	Subject	Predicate	Object
drought	-	-	-
This query selects all concepts whose description contain the keyword "drought"			

Figure 10. Sequence diagram of the automatic query expansion


«Client»
User Agent
«service»
Discovery Augmntation Component
«service»
Semantic Service
«service»
Catalog Service
query(searchTerms, geo:bbox, geo:time, sem:rel)
query(searchTerm)
«SPARQL»
[build semantic queries]
query(concept-1, sem:rel)
«SPARQL»
query(conceptN, sem:rel)
«SPARQL»
[build geospatial queries]
getRecords(bbox, time, keyword-1)
getRecords(bbox, time, keyword-N)
[aggregate responses]

B. The selected concepts are used to build a set of "traditional" geospatial queries to be submitted to the *catalog service* accessed by the DAC;

C. The DAC performs a "smart" aggregation of the queries results and provides them back to the client (*user agent*).

The second type differs from the first one only for the first step. In fact, in this case, the selection of the concepts of interest is not automatic; the user can freely navigate the content of the semantic services and select which concepts will be used to query the catalog service. Figure 11 depicts the steps of this second type of query expansion

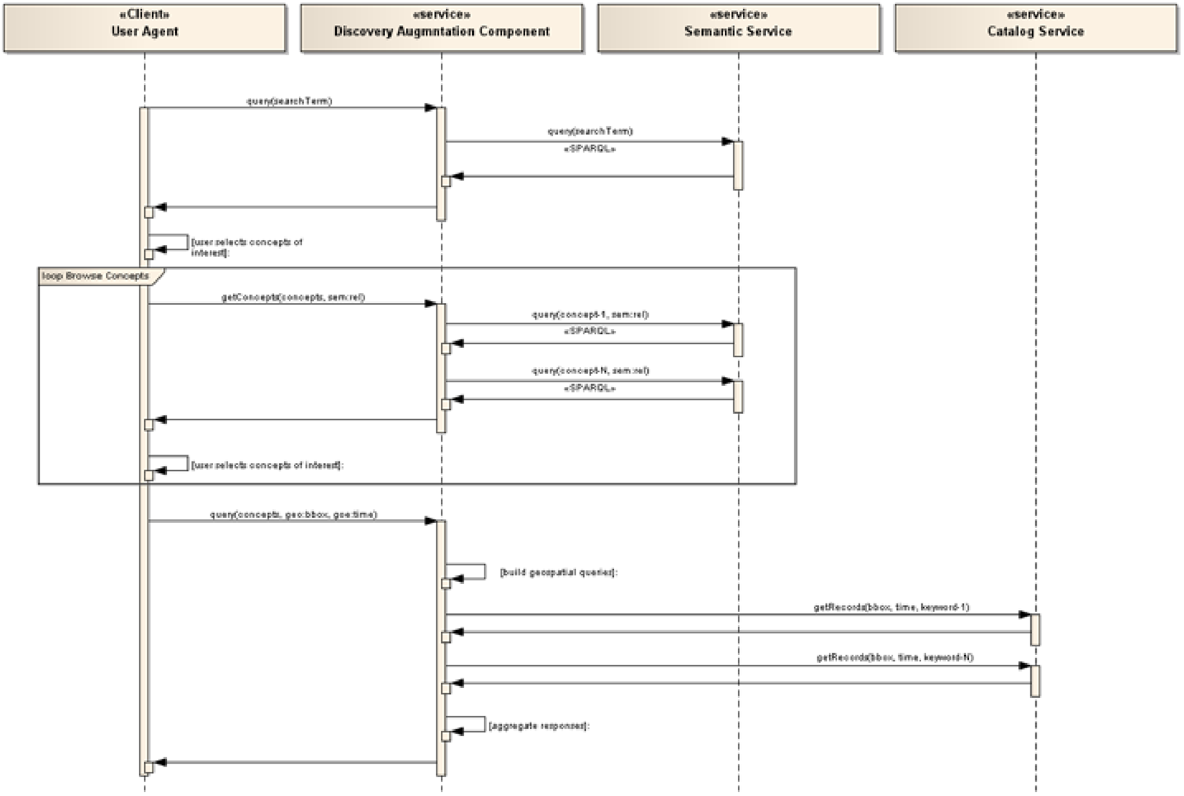
Discovery Augmentation Component

In order to implement the architecture described above, the DAC design is crucial. As for the functional requirements, the main functionality of the DAC is the implementation of the two query expansion strategies; the DAC must be able to:

- Interpret a semantic geospatial query;
- Expand the semantic geospatial query into one or more semantic queries;
- Build a set of geospatial queries and submit them to the catalog service;
- Combine the results of a set of geospatial queries in a meaningful way;
- Transform the results to the requested return type.

The design of the DAC internal modules is based on the *Mediation* pattern: a central module (the mediator) orchestrates a set of modules that operate independently from each other. This way, it is possible to concentrate the business logic that is required to expand the query into one central module: the *Orchestrator*. The other modules provide specific functionalities, such as generating and executing semantic queries, generating geospatial queries, executing geospatial queries, and transforming the results to the desired return type. In order to achieve the desired flexibility, communication between the Orchestrator and

Figure 11. Sequence diagram of the user-assisted query expansion


«Client»
User Agent
«service»
Discovery Augmntation Component
«service»
Semantic Service
«service»
Catalog Service
query(searchTerm)
query(searchTerm)
«SPARQL»
[user:selects concepts of interest]:
loop Browse Concepts
getConcepts(concepts, sem:rel)
query(concept-1, sem:rel)
«SPARQL»
query(concept-N, sem:rel)
«SPARQL»
[user:selects concepts of interest]:
query(concepts, geo:bbox, goe:time)
[build geospatial queries]:
getRecords(bbox, time, keyword-1)
getRecords(bbox, time, keyword-N)
[aggregate responses]:

the other modules is completely decoupled. The Orchestrator module interprets the semantic geospatial query and, accordingly, calls the necessary modules in the right order to process the request. Figure 12 depicts the main DAC internal modules.

The *OpenSearch Profiler* reads incoming requests, checks for validity, and forwards them to the Orchestrator in a more structured form. These functionalities must not be implemented by the Orchestrator to guarantee more flexibility; in fact, the Orchestrator must be agnostic to the communication protocols published by the DAC. Thus, future useful protocols can be added just by implementing new profilers.

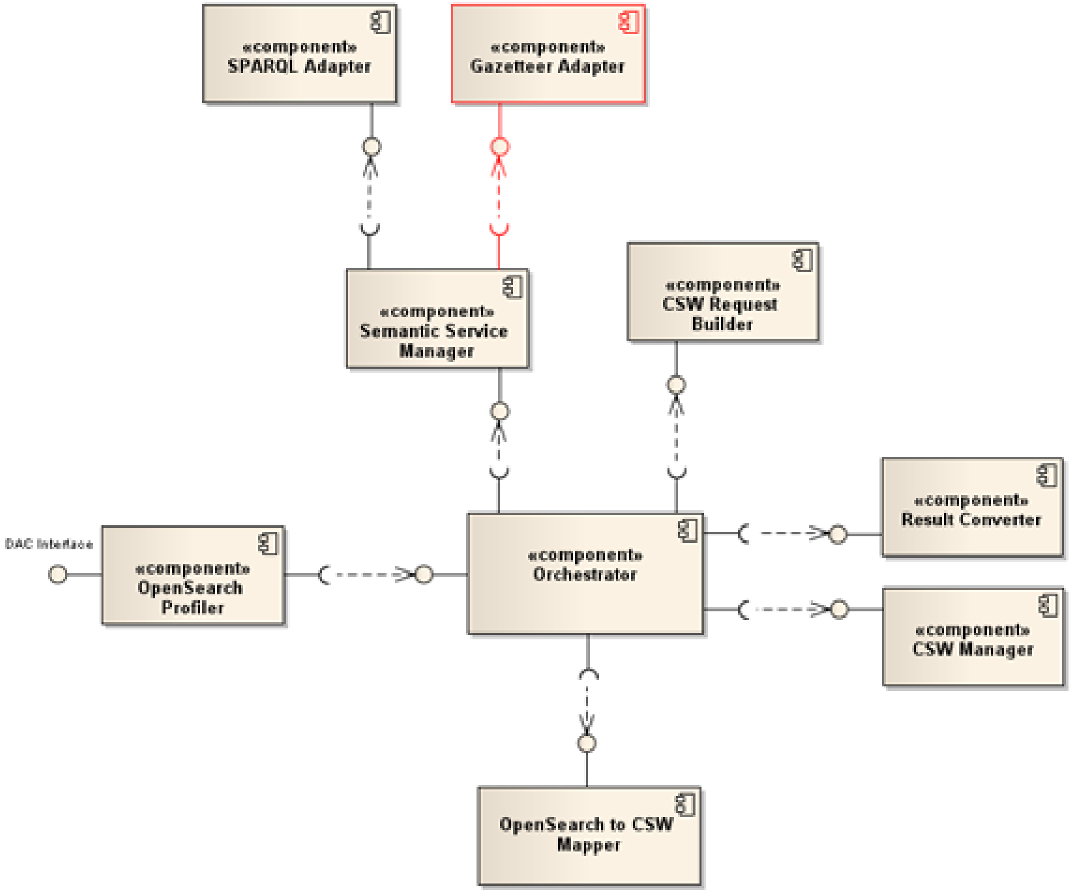
The *Semantic Service Manager* is in charge of generating and submitting requests to the *semantic service(s)*. This module makes of use of a set of *Adapters* to manage semantic services heterogeneity. These adapting modules mediate between the internal *Semantic Service Manager* interface and the specific semantic service query language (e.g. SPARQL). Thus, in order to add a new semantic service type (e.g. a gazetteer service), it is sufficient to implement the corresponding *Adapter*.

The *OpenSearch2GetRecords Mapper* module is able to read the non-semantic parameters of the request (i.e. spatial and temporal constraints) and create an ISO GetRecords request—which can be submitted to a standard OGC CSW/ISO catalog.

Given an ISO GetRecords request and a set of concepts (represented as terms), the *CSW Request Builder* module expands the original request generating a set of related ISO GetRecords requests.

The *CSW Manager* module communicates with the catalog service. It should be noted that this component does not make use of any adaptation functionalities. In fact, solutions for mediating different types of geospatial discovery services

Figure 12. DAC internal modules


«component» SPARQL Adapter
«component» Gazetteer Adapter
«component» Semantic Service Manager
«component» CSW Request Builder
«component» Result Converter
DAC Interface
«component» OpenSearch Profiler
«component» Orchestrator
«component» CSW Manager
OpenSearch to CSW Mapper

already exist (EuroGEOSS, 2010). Thus, where necessary, they should be coupled with the DAC.

Finally, the *Result Converter* module encodes the results (which can be both metadata records and semantic concepts) into the desired return type.

The DAC data model considers two main information elements:

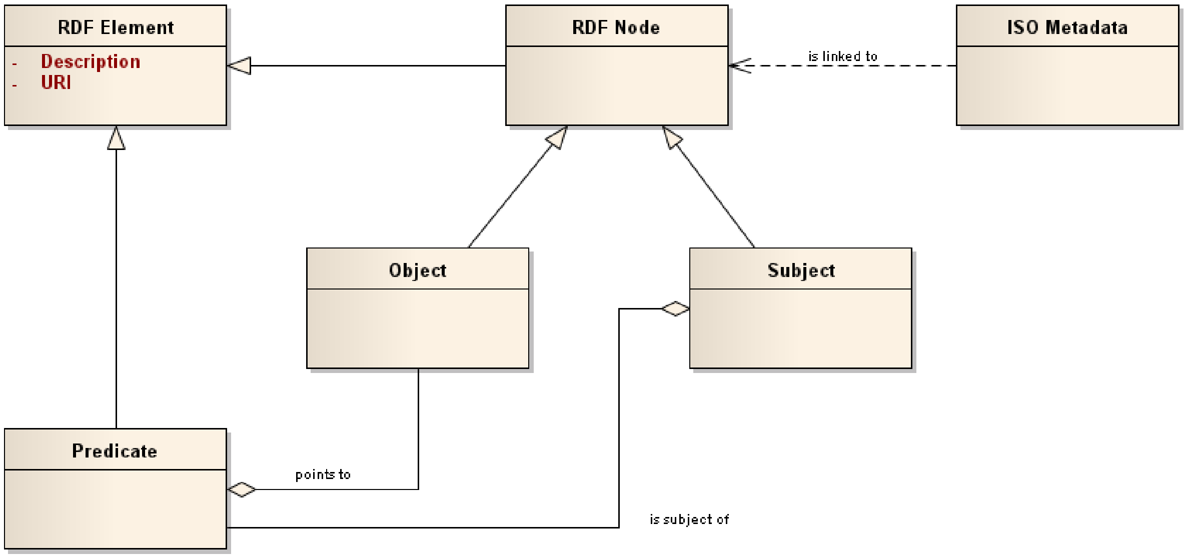
- Metadata Records
- RDF elements (concepts and semantic relations).

Figure 13 depicts a simplified schema of the DAC data model. This model connects ISO metadata records and concepts represented as RDF nodes, allowing building an RDF graph between metadata records (see Figure 14).

Referring to Figure 14, relations among metadata records are more "relaxed" than those obtained by linking the metadata elements directly: Figure 15a) depicts the case of direct links; while Figure 15b) shows the case of metadata elements linked through concepts.

In the design of the data model, an important choice is the strategy adopted to link metadata records to concepts. In fact, these two kinds of elements can be linked in several ways, determining the reliability of the *Relaxed Relation*. Basically, this can be decided case by case depending on the overall system needs.

Figure 13. DAC data model



Metadata records are composed of several elements (e.g. title, abstract, keywords, etc). These elements are used to link the metadata records to one or more matching semantic concepts—see Figure 15

One strategy is to match the "keyword" element content to one or more semantic concepts. This strategy generates a very reliable *Relaxed Relation*, resulting quite restrictive however.

A more general but less reliable strategy consists of taking into account also other metadata elements such as "title," "abstract," etc. We adopted this strategy, defining the link between metadata records and semantic concepts as it follows: *a metadata record is linked to a semantic concept if this appears in one of metadata record "keyword" elements OR in any other textual element of the metadata record*.

CONCLUSION

In this chapter, we discussed the enhancement of geospatial resources discovery services to support semantic queries. After explaining the general vision, we exposed specific issues related to the semantic discovery of existing geospatial resources provided through standards services and Web2.0 services. Methodologies, architectures, and current experimentations are also discussed. A possible methodology to enhance discovery capabilities of SDIs is to augment the searchable content that is associated with geospatial resources. We have defined three high-level approaches for providing the required additional meta-information: provider-based, user-based, and third party. We have selected the more extensible and flexible ones: the user-based approach, and the third party approach.

Two architectures have been presented, each implementing one of the selected approaches. In the first one, regarding the user-based approach, the architecture makes use of tags in two different but complementary scenarios. The first scenario considers tagging at discovery time, when users launch a query based on a set of tags. These tags are either introduced by the users or suggested by the system given the pool of previous tags used by others. In this case, the previous queries give clues to non-expert users to successfully discovery

Figure 14. Graph connecting concepts and metadata records

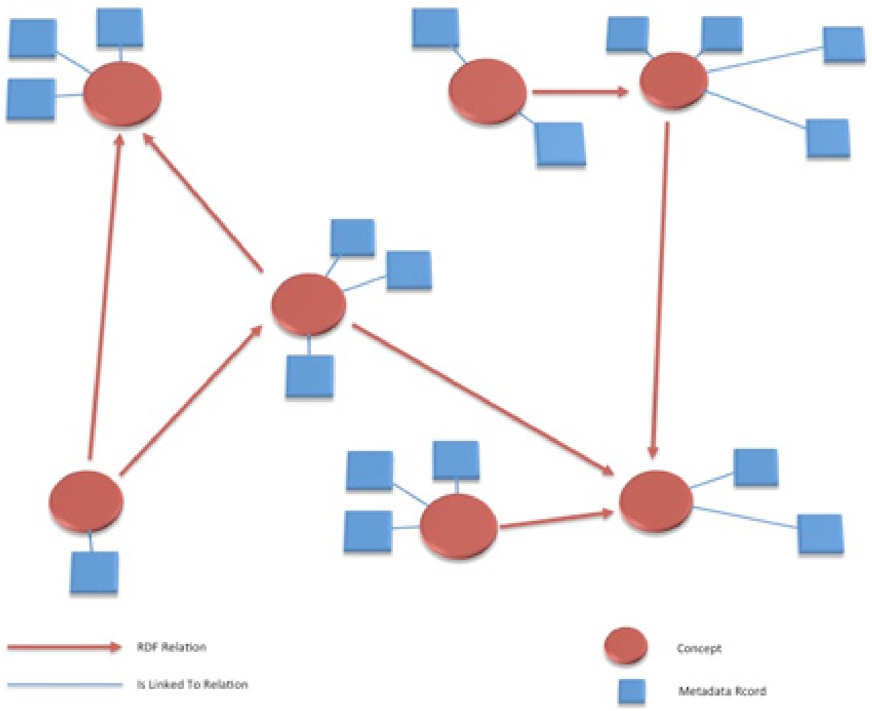
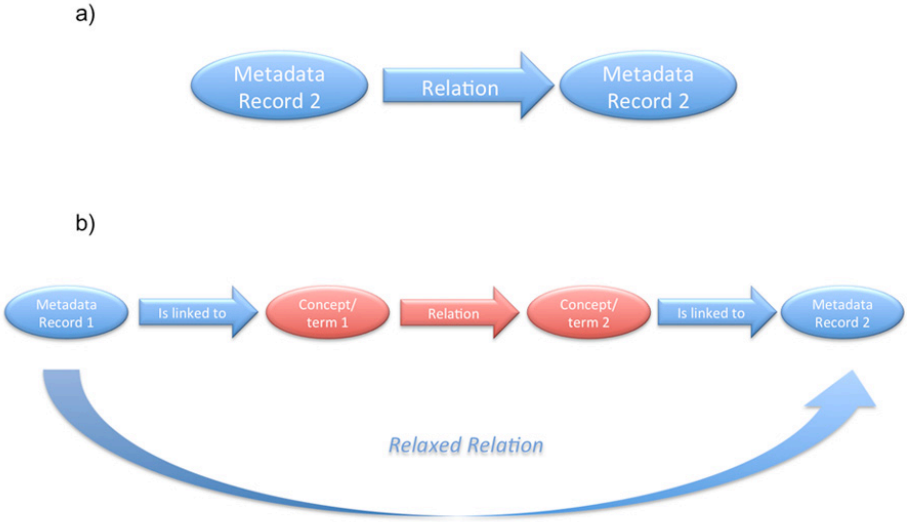


Figure15. Relations between metadata records in the DAC data model



geospatial resources. The second scenario involves the grouping of the resources that are discovered into collections. Users are allowed to annotate with tags the collection of discovered resources on the basis of certain relationships. In this case, not only individual resources are annotated but whole collections that contain resources that are related to the same topics.

The second solution we presented implements the third party approach in order to extend traditional discovery functionalities with semantic capabilities. Leveraging the Semantic Web approach, we described a flexible framework that transparently augments user queries with concepts related to the free-text constraint in the original query. This architecture is centered on a third party component, the Discovery Augmentation Component (DAC), which implements the business logic needed to integrate semantic and geospatial capacities.

The two architectures have been prototyped in the context of the EC-funded EuroGEOSS project and are part of the EuroGEOSS multidisciplinary interoperability infrastructure. This was successfully tested in several use scenarios of the GEOSS AIP-3 (Architecture Implementation Pilot—Phase 3) in collaboration with the GENESIS project (Nativi, et al., 2011; Fugazza, et al., 2011; Pozzi, et al., 2011). Demonstration videos are available at http://www.ogcnetwork.net/pub/ogcnetwork/GEOSS/AIP3/pages/Demo.html.

With regard to the user-centered approach, future research threads will be the inclusion of new Web 2.0 services, exchange formats and protocols, and the potential integration with official SDI content, the recommender module is still an open issue and existing and innovative techniques must be tested.

With regard to the DAC, future work will extend the set of semantic service interfaces that are currently supported. Besides, new controlled vocabularies and ontologies will be integrated. In our experimentation, the alignment of different thesauri and ontologies was carried out manually by domain experts; an automatic alignment approach for the DAC is another challenge that will be considered in the future.

ACKNOWLEDGMENT

The work was partially funded by the European Commission's Seventh Framework Programme under contracts 226487 and 223996.

REFERENCES

Abargues, C., Granell, C., Díaz, L., & Huerta, J. (2010). Aggregating geoprocessing services using the OAI-ORE data model. *International Journal on Advances in Intelligent Systems*, *3*(3-4).

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., & Patel-Schneider, P. (2003). *The description logic handbook: Theory, implementation and applications*. Cambridge, UK: Cambridge University Press.

Barragáns-Martínez, A. B., Rey-López, M., Costa-Montenegro, E., Mikic-Fonte, F. A., Burguillo, J. C., & Peleteiro, A. (2010). Exploitation of social tagging for outperforming traditional recommender system techniques. *IEEE Internet Computing*, *14*(6), 23–30.

Berners-Lee, T., & Fischetti, M. (1999). *Weaving the web: The original design and ultimate destiny of the world wide web by its inventor*. San Francisco, CA: Harper.

Brin, S., & Page, L. (1998). *The anatomy of a large-scale hypertextual web search engine*. Paper presented at the Seventh International World-Wide Web Conference. New York, NY.

Clinton, D. (2009). *OpenSearch 1.1 specification (draft 4)*. Retrieved June 8, 2011, from http://www.opensearch.org /Specifications/OpenSearch/1.1/Draft_4.

Craglia, M., Goodchhild, M. F., Annoni, A., Camara, G., Gould, M., & Kuhn, W. (2008). Next-generation digital earth: A position paper from the Vespucci initiative for the advancement of geographic information science. *International Journal of Spatial Data Infrastructures Research*, *3*, 146–167.

Díaz, L., Granell, C., Gould, M., & Huerta, J. (2011). Managing user generated information in geospatial cyberinfrastructures. *Future Generation Computer Systems*, *27*(3), 304–314. doi:10.1016/j.future.2010.09.002

EIONET. (2011). *The general multilingual environmental thesaurus (GEMET)*. Retrieved May 20, 2011, from http://www.eionet.europa.eu /gemet.

EuroGEOSS WP2. (2010). *Deliverable 2.2.2 – Specification of the EuroGEOSS initial operating capacity.* Retrieved May 20, 2011, from http://www.eurogeoss.eu /Documents/ EuroGEOSS_D_2_2_2.pdf.

EuroVoc. (2011). *EuroVoc – The EU's multilingual thesaurus.* Retrieved May 20, 2011, from http:// eurovoc.europa.eu/drupal/.

FAO. (2011). *FAO – The AGROVOC thesaurus*. Retrieved May 20, 2011, from http://aims.fao.org /website/ AGROVOC-Thesaurus/sub.

Franklin, C., & Hane, P. (1992). An introduction to GIS: Linking maps to databases. *Database*, *15*(2), 17–22.

Fugazza, C., Nagai, M., Nativi, S., Santoro, M., & Pozzi, W. (2011). *GEOSS AIP-3 engineering report – Vocabularies and semantics scenario.* Retrieved May 20, 2011, from http://www.ogc-network.net /pub/ogcnetwork/GEOSS/AIP3/ documents/AIPscenario_Semantics_1.5.pdf.

Gonçalves, P. (2010). *OpenGIS® OpenSearch geospatial extensions draft implementation standard, ver 0.0.2*. OGC 09-084r3. Retrieved from http://www.ogcnetwork.net.

Group on Earth Observation. (2010). *GEOSS – 2009-2011 work plan.* Retrieved May 20, 2011, from http://www.earthobservations. org /documents/work%20plan/ geo_wp0911_ rev3_101208.pdf.

INSPIRE. (2010). *INSPIRE metadata implementing rules: Technical guidelines based on EN ISO 19115 and EN ISO 19119, version 1.2*. Retrieved May 20, 2011, from http://inspire.jrc.ec.europa.eu /documents/Metadata/ INSPIRE_MD_IR_and_ ISO_v1_2_20100616.pdf.

International Organization for Standardization. (1994). *Information technology -- Open systems interconnection -- Basic reference model: The basic model. ISO/IEC 7498-1*. Geneva, Switzerland: ISO.

International Organization for Standardization. (2003a). *Geographic information – Metadata*. *ISO*, *19115:* 2003. Geneva, Switzerland: ISO.

International Organization for Standardization. (2003b). *Geographic information – Services. ISO/IS, 19119:2003*. Geneva, Switzerland: ISO.

Klien, E., Einspanier, U., Lutz, M., & Hubner, S. (2004). An architecture for ontology-based discovery and retrieval of geographic information. In F. Toppen & P. Prastacos (Eds.), *Proceedings of the Seventh Conference on Geographic Information Science (AGILE 2004),* (pp. 179-188). Heraklion, Greece: Crete University Press.

Larman, C. (2005). *Applying UML and patterns - An introduction to object-oriented analysis and design and iterative development* (3rd ed.). Upper Saddle River, NJ: Prentice Hall PTR.

Lemmens, R., Wytzisk, A., de By, R., Granell, C., Gould, M., & van Oosterom, P. (2006). Integrating semantic and syntactic descriptions to chain geographic services. *IEEE Internet Computing*, *10*(5), 42–52. doi:10.1109/MIC.2006.106

MacEachren, A. M., & Kraak, M. J. (2001). Research challenges in geovisualization. *Cartography and Geographic Information Science*, *28*, 3–12. doi:10.1559/152304001782173970

Maguire, D. J., & Longley, P. A. (2005). The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems*, *29*, 13–14.

Masser, I. (2005). *GIS worlds: Creating spatial data infrastructures*. New York, NY: ESRI Press.

NASA-JPL. (2011). *Semantic web for earth and environmental terminology (SWEET).* Retrieved May 20, 2011, from http://sweet.jpl.nasa.gov/.

Nativi, S. (2010). The implementation of international of geospatial standards for earth and space sciences. *International Journal of Digital Earth*, *3*(1), 2–13. doi:10.1080/17538941003764412

Nativi, S., & Bigagli, L. (2009). Discovery, mediation, and access services for earth observation data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *2*(4), 233–240. doi:10.1109/JSTARS.2009.2028584

Nativi, S., Santoro, M., Dubois, G., Skoien, J. O., De Jesus, J., De Longueville, B., & Fugazza, C. (2011). *GEOSS AIP-3 engineering report – eHabitat.* Retrieved May 20, 2011, from http://www.ogcnetwork.net /pub/ogcnetwork/GEOSS/AIP3/documents/ CCBio-eHabitat-ER-v2.0-FINAL.pdf.

Nuñez, M., Díaz, L., Granell, C., & Huerta, J. (2011). Web 2.0 broker: A tool for massive collection of user information. In *Proceedings of the European Geosciences Union (EGU) General Assembly 2011 (*EGU 2011). Retrieved 8 June, 2011, from http://meetings.copernicus.org /egu2011/.

O'Reilly, T. (2005). *What Is web 2.0 - Design patterns and business models for the next generation of software*. Retrieved May 20, 2011, from http://oreilly.com /pub/a/web2/archive/what-is-web-20.html.

O'Reilly, T., & Battelle, J. (2009). *Web squared: Web 2.0 five years on.* Retrieved May 2, 2011, from http://www.web2summit.com /web2009/public/ schedule/detail/10194.

Open Archives Initiative. (2008a). *Object reuse and exchange specification.* Retrieved May 20, 2011, from http://www.openarchives.org /ore/.

Open Archives Initiative. (2008b). *ORE user guide – Abstract data model.* Retrieved May 20, 2011, from http://www.openarchives.org /ore/1.0/datamodel.

Open Geospatial Consortium. (2007a). *OpenGIS® catalogue services specification, ver. 2.0.2*. OGC 07-006r1. Retrieved from http://www.ogcnetwork.net.

Open Geospatial Consortium. (2007b). *OpenGIS® catalogue services specification - ISO metadata application profile, ver. 1.0.0*. OGC 07-045. Retrieved from http://www.ogcnetwork.net.

Open Geospatial Consortium. (2008). *OGC KML, ver 2.2.0*. OGC 07-147r2. Retrieved from http://www.ogcnetwork.net.

Pozzi, W., Fugazza, C., Brewer, M. J., Santoro, M., Nativi, S., Lee, B., & Enenkel, M. (2011). *GEOSS AIP-3 engineering report – Global drought monitoring service through the GEOSS architecture.* Retrieved May 20, 2011, from http://www.ogcnetwork.net /pub/ogcnetwork/GEOSS/AIP3/documents/ GEOSS_AIP3_DroughtWater_ER.pdf.

Richards, D. (2007). A social software/Web 2.0 approach to collaborative knowledge engineering. *Information Sciences*, *179*(15), 2515–2523. doi:10.1016/j.ins.2009.01.031

Russell, S., & Norvig, P. (2005). *Artificial intelligence: A modern approach*. Upper Saddle River, NJ: Prentice Hall.

Smits, P., & Friis-Christensen, A. (2007). Resource discovery in a European spatial data infrastructure. *IEEE Transactions on Knowledge and Data Engineering*, *19*(1), 85–95. doi:10.1109/TKDE.2007.250587

Strohmaier, M., Koerner, C., & Kern, R. (2010). Why do users tag? Detecting users' motivation for tagging in social tagging systems. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (ICWSM 2010).* AAAI.

Stvilia, B., Twidale, M. B., Smith, L. C., & Gasser, L. (2008). Information quality work organization in Wikipedia. *Journal of the American Society for Information Science and Technology*, *59*, 983–1001. doi:10.1002/asi.20813

Turner, A. (2010). *OpenSearch Geo extension 1.0 (draft 2)*. Retrieved June 8, 2011, from http://www.opensearch.org /Specifications/OpenSearch/ Extensions/Geo/1.0/Draft_2.

USDA. (2011). *National agricultural library thesaurus and glossary.* Retrieved May 20, 2011, from http://agclass.nal.usda.gov /agt.shtml.

Vilches-Blazquez, L. M., & Corcho, O. (2009). A heuristic approach to generate good quality linked data about hydrography. In *Proceedings of the Database and Expert Systems Applications, International Workshop,* (pp. 99-103). IEEE Press.

World Wide Web Consortium. (2004a). *Resource description framework (RDF): Concepts and abstract syntax.* Retrieved May 20, 2011, from http://www.w3.org /TR/rdf-concepts/.

World Wide Web Consortium. (2004b). *RDF vocabulary description language 1.0: RDF schema.* Retrieved May 20, 2011, from http://www.w3.org /TR/2004/REC-rdf-schema-20040210/.

World Wide Web Consortium. (2004c). *OWL web ontology language reference.* Retrieved May 20, 2011, from http://www.w3.org /TR/2004/REC-owl-ref-20040210/.

World Wide Web Consortium. (2008). *SPARQL query language for RDF.* Retrieved May 20, 2011, from http://www.w3.org /TR/rdf-sparql-query/

World Wide Web Consortium. (2011). *Semantic web activity.* Retrieved May 20, 2011, http://www.w3.org /2001/sw/.

ENDNOTES

[1] http://www.opengeospatial.org/

[2] http://www.geonames.org/about.html

[3] http://geocommons.com/help/About

[4] http://wiki.openstreetmap.org/wiki/Main_Page

[5] http://www.earthobservations.org/geoss.shtml

[6] http://www.georss.org/Main_Page

[7] http://geojson.org/

[8] http://www.openarchives.org/ore

[9] http://tools.ietf.org/html/rfc3986

[10] http://www.ietf.org/rfc/rfc4287.txt

[11] http://www.w3.org/TR/rdf-syntax-grammar/

[12] http://www.w3.org/TR/rdfa-syntax/

[13] http://www.geonames.org

[14] http://dbpedia.org

[15] http://www.ordnancesurvey.co.uk/ontology

[16] http://en.wikipedia.org/wiki/Triplestore

[17] http://www.docbook.org/specs/cs-docbook-docbook-4.2.html

[18] http://www.dublincore.org/documents/dcmi-terms/

[19] http://www.geonames.org/ontology/ontology_v2.2.1.rdf

[20] http://www.w3.org/2004/02/skos/

[21] http://protege.stanford.edu/

[22] http://www.eurogeoss.eu/

[23] http://his.cuahsi.org/

[24] http://auscope-services.arrc.csiro.au/vocab-service/client/

[25] http://www.genesis-fp7.eu/

[26] http://www.openrdf.org

[27] http://www.eionet.europa.eu/gemet

[28] http://www.eurogeoss.eu/

[29] http://www.opensearch.org/Specifications/OpenSearch/Extensions/Suggestions/1.1

[30] http://www.rssboard.org