Proceedings of the 7th International Conference on HydroScience and Engineering

Philadelphia, USA September 10-13, 2006 (ICHE 2006)

ISBN: 0977447405

Drexel University

College of Engineering

GENERIC CONCEPTUAL MODELS FOR DATA AND KNOWLEDGE SHARING. APPLICATION TO THE ENVIRONMENTAL DOMAIN.

Barde Julien[1], Jean Christophe Desconnets[2], Thérèse Libourel[3] and Pierre Maurel[4]

ABSTRACT

Within the framework of the Syscolag and Roselt pluridisciplinary research programs, we propose generic conceptual models to set up a communication infrastructure to facilitate informational resources as well as knowledge sharing between the various stakeholders. We propose the use of a metadata service to share informational resources and the use of an ontology to share knowledge. We focus on the control of thematic and spatial descriptions. We propose a generic model to manage the semantic and spatial relationships between thematic and spatial concepts regardless of the domain. Finally we propose a generic ontology model to inventory and structure concepts and related knowledges in the particular case of the environmental domain.

1. INTRODUCTION

In this paper we propose to describe a common research work carried out in the framework of two research projects to enhance data and knowledge sharing. The first one, named Syscolag (COastal and LAGoonal SYStem), aims to support an Integrated Coastal Zone Management (ICZM) approach of the Languedoc-Roussillon coastal area, in the South of France. The second one, called Roselt (Long Term Ecological Monitoring Observatories Network) is a regional network designed to facilitate the implementation of common data collection and processing methods to study the desertification processes in the African Circum-saharian zone. This paper summarizes the main results related to a PhD thesis in computer sciences, which was part of the interdisciplinary Syscolag research program. During that period we studied in collaboration with our partners generic and specific concepts and tools to share data and knowledge in the environmental domain (applied to the ICZM domain in particular). The issue of data and knowledge sharing in this domain is challenging because of the variety of informational resources distributed among different stakeholders. Also challenging is the multidisciplinary framework which involves the setting up of a common vocabulary between the partners in order to share the meaning of terms used to describe and search data or knowledge. This is particularly true in the ICZM domain. In this paper we refer to the definitions of the terms *data*, *information* and *knowledge* given by Doody (2003) and Kay and

[1] Research engineer, CEMAGREF/UMR TETIS, Maison de la télédétection, 500 rue J. F. Breton, Montpellier, cedex 5, 34093, France (juldebar@gmail.com)
[2] Research engineer, IRD/US Désertification, Maison de la télédétection, 500 rue J. F. Breton, Montpellier, cedex 5, 34093, France (jcd@teledetection.fr)
[3] Computer Senior Research Scientist, LIRMM, 161 rue Ada 34392 Montpellier, cedex 5 – France (libourel@lirmm.fr),
[4] Research engineer, CEMAGREF/UMR TETIS, Maison de la télédétection, 500 rue J. F. Breton, Montpellier, cedex 5, 34093, France (pim@teledetection .com)

Christie (2001) in the ICZM context. We will use the generic term of *informational resource* or *resource* to designate any kind of data, information or knowledge.

Our work aims to set up generic methods to minimize syntactic and semantic interoperability problems in order to improve data and knowledge management in the environmental domain (applied to Syscolag and ROSELT contexts). An additional difficulty is due to the management of the spatial dimension of the environmental information. The proposed solution consists of an infrastructure of mutualization whose core is made of two main components: a metadata management service to describe and locate existing information coupled with a domain ontology (a semantic referential) to integrate and share expert knowledge (Barde, 2005 and Barde et al., 2005). In particular, one of our goals was to illustrate the close relationship between these components. Indeed we think that the metadata service and ontologies setting up should be done together since they depend on each other.

The structure of this paper follows the description of three main conceptual models (UML class diagrams cited above), which summarize our approach. In the first section, we propose a class diagram which explains our understanding of main steps to set up a metadata service from a standard metadata element set. In particular we will focus on the specific roles played by metadata elements related to spatial and thematic descriptions in the quality of data and knowledge sharing. In the second section, we propose a model to manage thematic and spatial descriptions in order to improve knowledge sharing and data sharing as well thanks to the setting up of a controlled vocabulary into our metadata service. We make explicit the relationships between thematic and spatial concepts and thereafter between a metadata service and a spatial ontology. This diagram summarizes the different kinds of concepts as well as relationships used to set up our ontology (regardless of the domain in question). In the third section, we present a model to represent basic knowledge in environmental domain in a manner that is compliant with the ecosystemic approach as well as semantic Web applications. It can be considered as a proposal of methodology to inventory and structure concepts of an ontology for the thematic field of environmental domain. In the last section, we present some possible applications as well as our technical realization. The whole solution can be accessed online through a Web portal. The details of the semantic referential can be accessed by the way of a terminological base, a semantic network as well as a digital atlas in the special case of spatial concepts (OGC WMS compliant). By integrating these components into our metadata service, we aim to facilitate and to improve metadata capture and retrieval tasks.

2. DATA SHARING WITH A METADATA SERVICE

More than data management, this section concerns in practice informational resources[5] management using a metadata service (to locate, acquire and treat informational resources) in the environmental domain (applied to coastal and marine domains). This step is an essential preliminary which stays necessary all along the integrated management process. Furthermore, this task is more complex in an interdisciplinary framework because of the lack of shared tools and methods as well as informational resources heterogeneity. This task will thus contribute in filling the lack of interdisciplinary representation of coastal and marine areas.

2.1 Syntactic interoperability

[5] The term “informational resource” refers to information elements such as data, documents, protocols that will be described and qualified at a higher level (metadata) using the metadata service. Depending on the organization and the individual, a resource might be a report, a map, an image, a video, a dataset, a database, a model, a value measured by a sensor, an image-segmentation software application, etc.

It is currently clearly accepted that metadata are essential (but not sufficient) to facilitate data sharing. Indeed, metadata helps humans and machines to locate as well as to access the physical data to which they are related.
To ensure the quality of the descriptions as well their syntactic interoperability around the world with similar metadata service using compatible standards, standards provide shared (or at least compatible) and relevant syntax which consists in defining sets of metadata elements and their relationships. Moreover, according to the kinds of metadata elements, standards provide sometimes enumeration or controlled lists to fulfil metadata elements in order to make their content as explicit as possible from a semantic point of view. It means that syntactic interoperability should, if possible, be complemented by semantic interoperability. However, it can not be the role of a metadata standard to furnish controlled vocabularies for all the kinds of metadata elements. Coastal and marine domains are characterized by the high heterogeneity of available informational resources. Among them, we think that geographic information (GI) is essential to share data and knowledge in that context because of the graphical representation which allows to partly shunt semantic problems and unite users during knowledge sharing process. These different kinds of informational resources can be described in a relevant way by using several standards. For our metadata service, in our specific case, the ISO 19115 standard (relevant for the description of spatial information) has been chosen to describe our data in a standard way.

However none of them is generic enough to allow a fine description of all these informational resources without a prior profiling step to separate elements among their usefulness for the description of different kinds of informational resources. This work is essential to adapt any generic metadata standard to a specific context.

2.2 A generic model to set up a metadata service

We propose in Figure 1 a UML class diagram which summarizes the main steps of a metadata service setting up process regardless of the standard metadata elements set that is used.

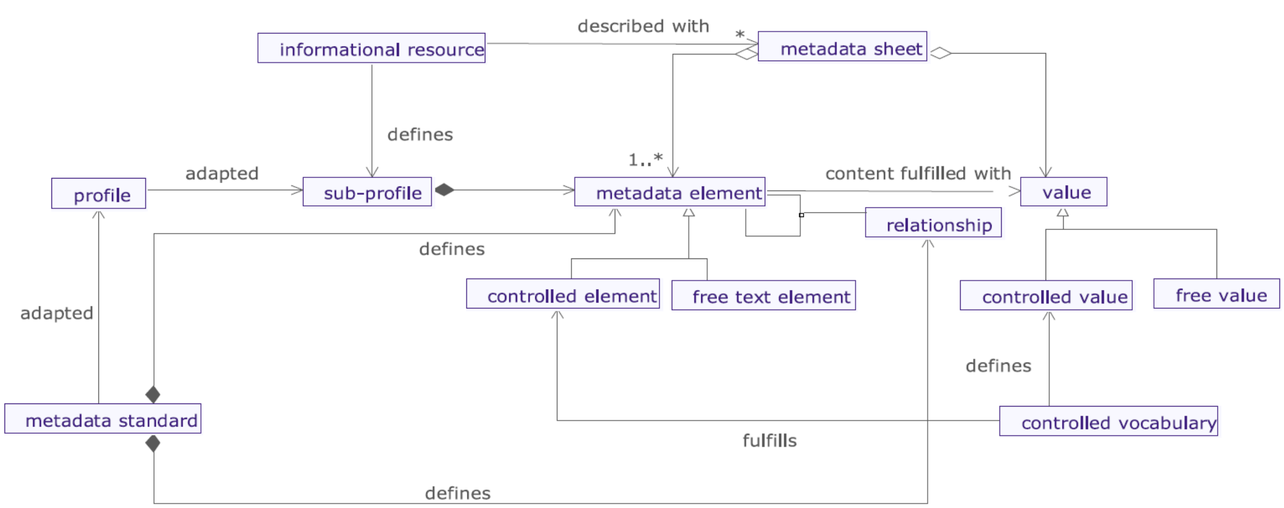

Figure 1 Manuscript layout Generic model to set up of a multi-standard metadata service: different elements to consider and different kinds of valuation related to them.

This diagram shows the different kinds of metadata elements to consider. First, we distinguish useful from useless MD elements (according to the contexts and the kinds of informational resources) by setting up of profiles:

- profile: according to the terminology of the OMG (Object Management Group), a profile, or adaptation, is an extension of a meta-model used to define the meta-model of a standard for a specific community (OMG). Thus, it specifies the options of the norm implementation in order to fulfil a particular need. A profile can not contradict the standard to which it refers. This is the reason why it always respects the core of this one. On the other hand, it can integrate elements which do not exist initially in the wide standard, or extended elements.
- sub-profile: a sub-profile is a new version of the profile (a profile of the profile) according to the kind of informational resource to describe.

Thereafter, we can distinguish different categories among the relevant elements according to their valuation modes:

- metadata elements which can be filled with free text:
  - some of them can't be automated because they are usually unique: *abstract*, *title*...
  - others because they are context dependent but can potentially be automated: *address* of *contacts*…
- metadata elements whose filling is controlled : *date*, ...

It seams that, the more users comes from a multi-disciplinary context, the less consensual metadata elements are several into the common profile set up. This is one of the reasons of the interest and the relevance of Dublin Core[6] metadata elements set. However the ISO 19115 is more interesting for an environmental application because of the possibility to describe information in a spatial way with standard geographical information (OGC recommendations compliant).

We implement this model to generate our physical data model in a RDBMS. For now, we just work with ISO 19115, but we can integrate in the corresponding tables any other standard (like the Dublin Core or Bibtex in particular).

The first part of that work was to study the ISO 19115 metadata standard to determinate which metadata elements present a particular interest for ICZM (in a multi-disciplinary context).

2.3 Set up of an ISO 19115 profile

In a very generic thematic field like ICZM, it appears after interviews with the different stakeholders of our program that it is difficult to have a consensus on metadata elements utility. For example, except the metadata element which answers the question "*Where* ? " common to a broad variety of informational resource type, it clearly appears after interviews with the partners involved in our project that other ISO 19115 metadata elements specific to geographical information only interest GIS users and not especially all kinds of ICZM stakeholders. A specific profile of the ISO 19115 has thus been implemented, and thereafter as many sub-profiles as kinds of informational resources used in our specific context (in accordance with Figure 1). However, our experience in ICZM shows that thematic and geographical description metadata elements are essential to describe and locate existing informational resources in that context. They should thus be included in every sub-profile and their filling should be assisted with a controlled vocabulary (not in free text as permitted by ISO 19115).

2.4 Interest of thematic and spatial descriptions

Whatever the chosen metadata standard, the thematic as well as spatial metadata elements as prior criteria are used in most of the informational resources description and research process because they answer the basic questions "what ? " and "where ? " (usually with "when ? " as well as "Type of information resource ? ") (Barde et al., 2004). Even if using a standard to control syntactic forms of metadata, data and knowledge sharing will not be efficient without controlling the content of those metadata elements. Nevertheless, it is less difficult to control the keywords related to temporal and data type descriptions than those related to spatial and thematic descriptions.

However thematic and especially spatial metadata elements are not available in all standards. According to ISO 19115 recommendations, thematic and spatial descriptions metadata elements belong to the main heading (*identification*). As is the case for several standards recommendations, they are not mandatory and can be filled in free text (see Figure 2). However ISO encourages the users to control them (by suggesting the use of *thesaurusName* metadata element, see Figure 2). As they are particularly important in environmental context we decided to make them mandatory (as part of the core of our ISO 19115 profile). In complement we propose to assist and to control the way people describe these elements to ensure the quality of metadata sheets.

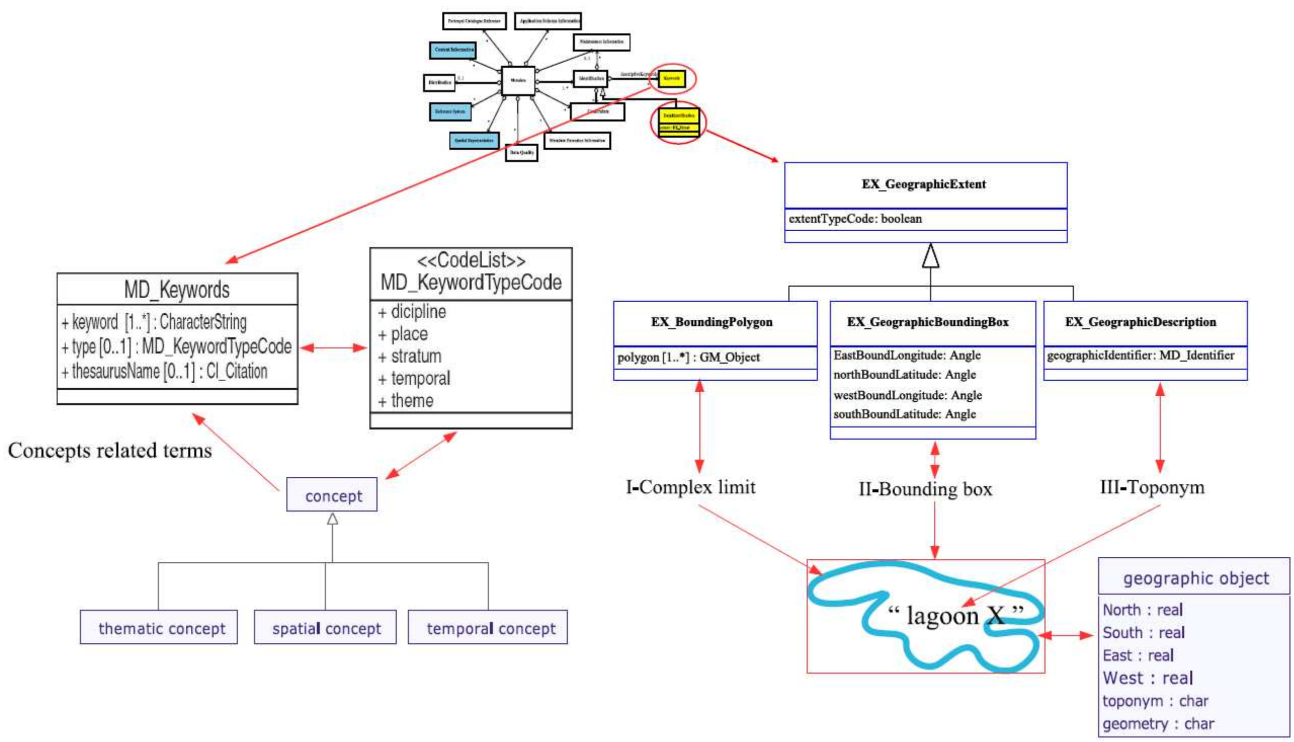

Figure 2 ISO 19115 metadata elements to describe thematic and geographic extent of an information resource.

We have chosen to use ISO 19115 to describe any kinds of informational resources (GI or not). Indeed, in environmental domain, except some particular information which can not be described by using a spatial extent (but have an universal interest). We want to describe, if possible, every type of resource by using these spatial metadata elements (e.g. *pictures*, *transects*, *tagging*, *measures*...).

To the minimal metadata sheet made of prior descriptions we thus add a geographical description which allows the management of metadata as any other type of information managed by the way of a GIS.

Our main goal is finally to propose a generic method to control or even to automate the filling of these fundamental but complex metadata elements. We detail in the following section why it is necessary to control the content of these elements and how this goal is strongly related to the process of knowledge sharing.

3. KNOWLEDGE SHARING WITH AN ONTOLOGY

Efficient knowledge sharing involves a good integration, representation as well as diffusion of the different stakeholders knowledge (scientific or not) by using methods adapted both to computer treatments and users needs.

The sharing of the meanings of the concepts, which is a prior necessary step to share knowledge, is permitted by the setting up of a controlled vocabulary. Moreover, the content of thematic and spatial metadata elements is made explicit and standard by using this kind of vocabulary. It is thus an essential complement for a metadata standard which standardizes the meaning of the metadata element itself (necessary as described in section 2.4). By working on knowledge sharing, we thus improve the quality of informational resources sharing.

As ontologies are the most sophisticated controlled vocabularies and allow a real knowledge sharing, we propose an ontology model compliant with current standards but adapted to our specific context, in particular to spatial descriptions.

3.1 Necessity of a semantic interoperability

As we previously said in section 2.3, thematic and spatial descriptions are core metadata elements to describe or search metadata sheets in our context. They have thus to be controlled in order to share the meaning of their contents as well. Such a control is made possible by completing our metadata service with controlled vocabularies dedicated to these kinds of elements content and built from a reference spatial database and from a semantic referential both adapted to our domain. Unfortunately, for now, we didn't find a relevant reference vocabulary related to coastal and marine domains to control the content of the related metadata elements.

We have therefore to inventory and structure the relevant *geographical objects* and *terms* (i.e. *concepts)* to describe the thematic content and the spatial extent of the different resources owned and distributed by the different stakeholders. This controlled vocabulary will thereafter serve as a basis to knowledge sharing.

We thus proposed to set up our own controlled vocabulary on coastal and marine domains for the specific case of Languedoc Roussillon area. For the same reasons as those of the metadata service, this controlled vocabulary aims to be as standard as possible to allow interoperability with other similar ones. Its aims are :

- from the users point of view : to answer the need of *common vocabularies* to share the meanings of concepts and the knowledges which are related to them. To assist the fulfilling of metadata elements during the description step or the search of metadata sheets into the metadata service. To represent knowledge with user friendly GUIs (e.g. with semantic networks).
- from the computer point of view : to allow the *matching between concepts* and thus the ability to *expand requests* by using the relationships introduced between concepts by the way of a thesaurus (or ontology) of ICZM domain. We hope this way to improve the relevance of metadata description and research.

In order to structure this inventory, we first propose to study the existing relationship between spatial and thematic concepts.

3.2 Link between spatial and thematic concepts

As suggested in ISO 19115 (see Figure 2), we have distinguished different categories among the collected concepts (i.e. thematic, spatial and temporal concepts). In particular, a part of that work focused on the existing link between spatial and thematic concepts in accordance with their specific importance (see section 2.4).

According to ISO 19115 recommendations, it is possible to use a term like "lagoon of Thau" to describe both key-word as well as geographic extent metadata elements. We thus proposed a model in order to clarify the use of that kind of term (see details in Barde et al., 2005). We propose to define a spatial concept as a kind of (specialization from object approach point of view) a thematic concept whose instances correspond to geographic objects (e.g. "lagoon" and "artificial reef" are both spatial concepts and "Thau lagoon" is a term which refer to a geographic object). Spatial concepts have thus to be distinguished from thematic concepts during the inventory.

Moreover we consider the spatial dimension as a potential property of any kind of concept.

We want by that way to allow map representation of existing data and knowledges as well as a thematic structuring of geographical information by following the thematic structure of our semantic referential. This approach helps to browse the content of our spatial database by using the same thematic structuring used in our semantic referential.

In the following sub-section we summarize the interest of different tools to set up a controlled vocabulary made of thematic and spatial concepts.

3.3 Different tools to control the vocabulary

Different kinds of tools exist to manage a controlled vocabulary. They differ from each other according to the diversity of relationships they allow to manage between concepts.

3.3.1 Different kinds of relationships between concepts

We decided to comply at least with existing consensual standards by using ISO 2788 and 5964 to set up our thesaurus/ontology (ISO, 1985 and ISO, 1986). With these standards it is possible to manage (in a multilingual context) the following classical relationships:

- *synonymy* and *polysemy* are managed by using *UF* (*Used For*) and *SN* (*Scope Note*) relationships which allow matching between terms and concepts (see Figure 4),
- *generalization* and *specialization* are managed by using *BT/NT* (*Broader Term/Narrower Term*) relationships and *RT* (*Related To*) relationship allow *requests expansions* towards other relevant concepts (Barde, 2005).

We notice that these kinds of relationships are independent of the thematic domain.

Indeed, with these relationships, it this impossible to express more specific kinds of knowledge like:

- "Swordfish eats squid". We thus proposed to manage a new kind of relationship which name is *RTS* (acronym for *Specialized RT*) which is related to a specific domain. *RTS* allows to specify the meanings of the *RT* relation: e.g. *protects*, *eats*, *pollutes*…In practice RTS can be related to an *URI* (*ie* to any concept of an ontology). We propose to use standardized concepts to name our RTS instances (see details in 3.3).
- "Thau lagoon is in south of Montpellier city". We thus proposed to manage Egenhofer spatial relationships too. This kind of relationships is currently standardized by the OGC.

In the following section, we present related tools to manage these kinds of concepts and relationships.

3.3.2 Link between the different existing tools

Figure 3 summarizes existing tools to control the vocabulary of a domain and propose a model to set up a spatial ontology. We list from the simplest to the most complex tool (Barde, 2005):

- a *catalogue* just consists of an inventory of concepts without describing their relationships,

- UF as well as SN additional relationships are managed by using at least a *glossary*,
- UF, BT (BTG/BTP), NT (NTG/NTP), RT additional relationships are usually managed by using a *thesaurus*,
- every additional kinds of relationships are managed by using an *ontology*. Among them, the previous standard relationships but as well RTS ones (see details in section 3.3). Ontology allows thus to integrate expert knowledge related to the concepts of a thesaurus in a specific domain.
- a *spatial ontology* : we consider that a spatial ontology is an ontology able to manage as well Egenhofer spatial relationships in the specific case of spatial concepts. A spatial ontology allows the geographic location of knowledges (see section 3.2 and Figure 4).

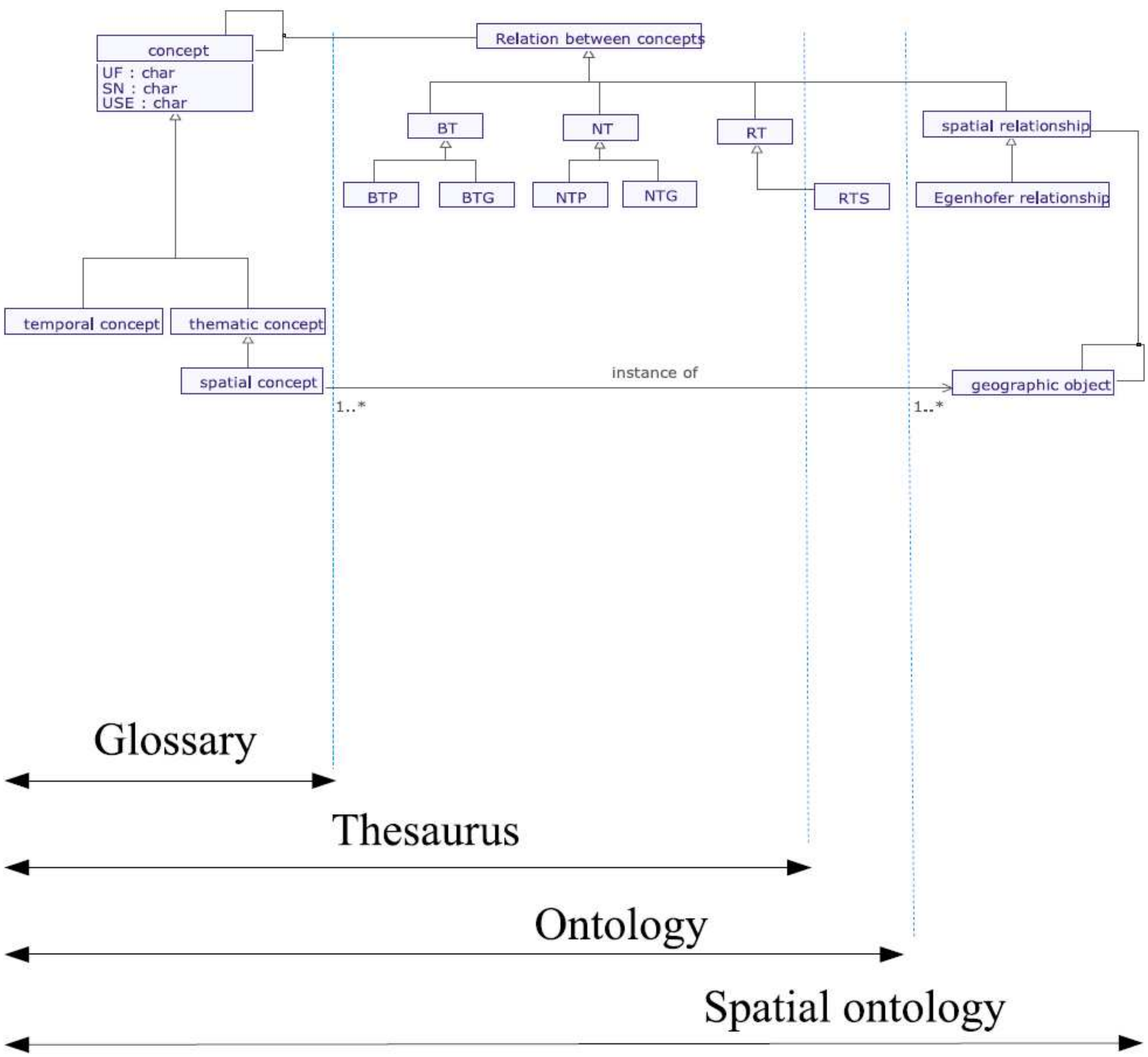

Figure 3: Different tools to control the vocabulary

We propose to use this simple spatial ontology model into our metadata service to generate, complete and use controlled vocabularies (compliant with standards from a semantic point of view) to facilitate thematic descriptions (by the way of the *key-words* metadata element) as well as spatial descriptions (by the way of the *geographical extent* metadata element) into our metadata service.

From this diagram we propose a synthesis with the Figure 1 which illustrates the link between a metadata service and a controlled vocabulary.

3.4 A generic model for the semantic control / Link between MD and ontology

If we consider the data sharing as well as the knowledge sharing conclusions, focusing on spatial and thematic description (see section 2.4), we can summarize the previous generic diagrams (Figure 1 and 3) as following in Figure 4.

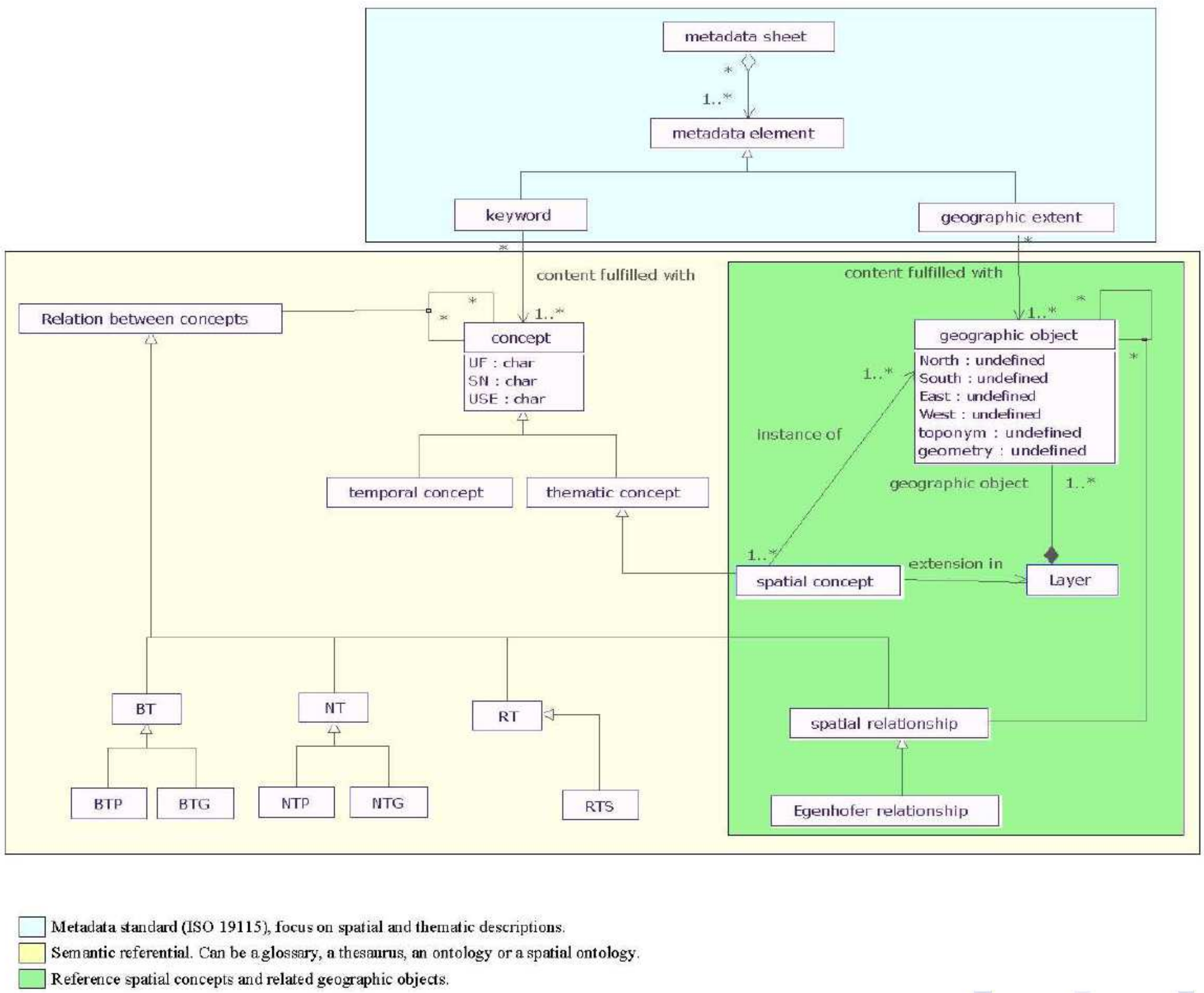

Figure 4 Links between keyword and geographic extent metadata elements into a metadata service and the concepts to fulfil them coming from a semantic referential (itself made of thematic and spatial referential).

3.4.1 Ontology into the metadata service

In our case we mainly use the controlled vocabulary into our metadata service:

- to assist users and control the fulfilling of metadata elements during description or searching steps :

  - by using a *completion component* (proposals comes from all the terms (preferred and alternate literal labels related to the inventoried concepts) of the available controlled vocabularies.
  - we nevertheless propose another component which is a *popup to browse the content of the different thesauri* for the case where people don't know a relevant term but want to browse the content to find a proposition.
- to allow the system to improve the quality of queries results (during the searching process) by expanding requests thanks to the relationships between concepts.

From a data sharing point of view, the tools to control vocabulary present various interests into a metadata service:

- a *catalogue* is sufficient to control the fulfilling with a completion component (and thus all others more sophisticated tools) but can't be used to expand requests,
- a *glossary* with the UF and SN relationships can expand requests on synonyms into a metadata service,
- a *thesaurus* facilitates the fulfilling with a thesauri content browser but as well searching thanks to more sophisticated requests expansions,
- the management of any kind of relationships is allowed with an *ontology*. Expert knowledge related to the concepts can thus be used to drive the system.
- a *spatial ontology.* As we explained previously, we proposed the following link between thematic and spatial concepts: a spatial concept is a kind of thematic concept whose instances are geographical objects (see Figure 4). By using this link, a spatial ontology will be useful to map informational resources and knowledge with the related keywords.

We thus aim the setting up of a spatial ontology to manage spatial as well as semantic relationships in a most efficient way. As spatial concepts will be related to geographical objects (managed into layers), queries expansions are then possible by using both semantic and spatial relationships.

3.4.2 Metadata into ontology

Contrariwise, we can use this model to generate into a concept description links to related metadata. Users will thus have the opportunity to consult some related metadata while browsing the semantic referential.

In order to share knowledge and to facilitate data sharing, we have used standards which concern controlled vocabularies (in particular Web semantic languages allow both management of metadata and ontologies) as well as knowledge representation using *semantic networks* to apply them, in a second time, to the modelling of coastal and marine systems. Ontologies will allow, in the long term, the setting up of a knowledge base (sensus stricto) in information systems.

Finally, with this diagram, we thus propose an infrastructure useful to manage a metadata service, a controlled vocabulary as well as a spatial database. Thereafter, as the these two conceptual models are not specific of ICZM at all but generic enough to a broad variety of applications, we then propose in the following section a *top level* model adapted to the thematic structuring of ICZM concepts in our domain. We propose to use an *ecosystemic approach* which is relevant to analyze the characteristics of complex coastal and marine systems and compliant with ontology's goals (inventory and structuring of concepts).

4. PROPOSAL OF A MODEL TO INVENTORY AND STRUCTURE THE CONCEPTS OF THE ICZM DOMAIN

The expert or local knowledge on coastal and marine ecosystems can be managed by using techniques which comes both from the study of knowledge sharing and representation in informatics as well as in ecology. The sharing of an efficient representation of the interactions between the elements of the coastal and marine systems (natural and/or anthropic elements) is an important aim to improve communication and sharing between communities involved in different disciplines but as well between machines.

As described in section 2, the setting up of a semantic referential for a specific domain involves the inventory of relevant concepts as well as the description of some of their relationships (more or less according to the kind of set up controlled vocabulary see Figure 3). In this section, we focus on a methodology proposal to realize a relevant (from a thematic point of view) concepts inventory which will be compliant with reference thesauri or ontologies as well as with the integration of basic knowledges related to ICZM.

We want to underline in this part the existing feedback from data sharing towards knowledge sharing. Indeed, the metadata service plays a main role by directly contributing to inventory terms related to constituent concepts of a controlled vocabulary.

4.1 Thematic analysis of the domain

By studying sustainable development as well as ICZM references (Hénocque and Denis, 2001, UNESCO, 1997 and 2003) promoting the ecosystemic approach, we consider that this approach facilitates the set up of an ontology, as described in Coccossis and Van Der Weide (1999):

“Systems theory is based upon the idea that the real world can be described by a set of elements and the interactions between these elements. In general, the system only includes a small part of the universe […]. The first step in any analysis of a system is the selection of its boundaries and the decomposition of the area within these boundaries into a set of elements and interactions.”

The different kinds of elements (natural or anthropics) and their basic interactions in the coastal and marine areas are represented by using this approach.

In order to propose a relevant ontology, we thus propose to apply the ecosystemic approach to inventory and structure elements and interactions (ie the concepts which are related to) in our domain. First, we propose to divide the world into two categories / sub-systems (see Figure 5):

- the first category concerns constituent *elements* of the ecosystem grouped into sub-systems:
    - the static environmental elements (natural or artificial), e.g. *fish*, *lighthouse*, *harbour*, *trawler…*
    - the humans (actors/stakeholders) who can be considered as a specific kind of natural element, e.g. *biologist*, *fisherman*, *windsurfer…*
- the second category concerns the *interactions* between those elements, which can be separated in different classes among the involved elements, e.g. *erosion*, *pollution*, *catalysis*, *use conflict…*

This way we propose a pattern to inventory the *static elements* we want to consider in a given system by designating concepts in accordance with the ecosystemic approach which is compliant with ontologies setting up. Then dynamic of the elements in the described system is modelled by introducing a *relationship* between these concepts.

Furthermore this approach is compatible with the representation used in computer sciences languages to represent knowledge (in particular the object approach).

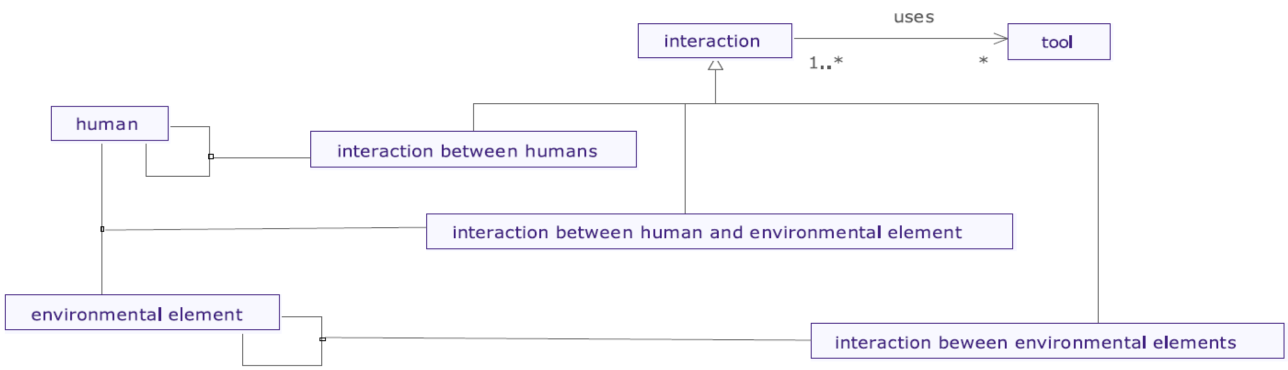

Figure 5 Model III : ontology model proposal to structure thematic concepts of ICZM.

4.2 Conformity with Web semantic languages

The model presented in Figure 5 is very close to the basic representation of knowledge used in Web semantic languages like RDF. Indeed, a triple is made of {Subject, Predicate, Object} where, in our case :

- Subject=actor | resource | interaction,
- Predicate=interaction,
- Object=actor | resource | interaction.

That way, it becomes possible to describe a new kind of basic expert knowledge like "Swordfish eats squid", which is not possible with standard thesauri relationships (see section 3.3). This new kind of relationship corresponds to RTS (see Figure 4 in section 3.4) which is an ontology relationship.

4.3 Inventory of domain related concepts

For a given (eco)system, by using the *ecosystemic approach*, we can thus process to the inventory of concepts by specializing these three top-level classes (element, actor and interaction, see Figure 5) in more specific ones (*top-down approach*) (by using *BT* and *NT* relationships). The bottom-up approach is fore sure possible too.
Moreover, in Syscolag ICZM research program case, in order to feed the lack of our concepts inventory as well as to improve its semantic interoperability, we want the users to specialize these generic concepts in our own domain by designating corresponding concepts from terms (both prefered or alternate labels) inventoried in reference thesauri (such as AGROVOC, GEMET, ASFA…). It is nevertheless possible to introduce a context specific concept (or simply a term) which could lack in the proposed reference thesauri.
In practice, this concepts inventory step is made by designating related terms either through the complementary use of our metadata service, mainly by describing (and even by searching) a resource with the key-word metadata element or thanks to dedicated GUIs to manage (and thereafter structure) the concepts inventory of our specific thesaurus (see section 3.4.1). But thanks to the metadata service, concepts can be designated unbeknownst to the users.
That way, we have integrated a first level of knowledge: knowledge of known elements in the system.
Once this inventory made or in course, we have a basis to build/create our specific thesaurus and ontologies by structuring these standard (from a semantic point of view) concepts with standard relationships (from a syntactic point of view) and specific knowledge related to these elements (as shown in Figure 3).

4.4 Structuring of inventoried concepts

As explained in section 3.3, this concepts inventory is then structured by introducing both kinds of relationships (thanks to dedicated GUIs):

- by complying with thesauri setting up standards (ISO 2788 and 5964) and their domain independent relationships (see details in section 3.3.1). It is thus necessary in this step to detail which terms are related to which concepts (if they are preferred or alternate labels), related definitions, broader and narrower terms…. This set up thesaurus will become thereafter a basis to decline ontologies.
- from the previous ontology model (see Figure 5) which helps to structure the basic knowledge related to these concepts (which is the most interesting relationship from a user point of view) by using domain dependent (RTS) relationships. It is so possible to introduce some relationships between these elements: interactions between static elements which will model the dynamic functioning of the concerned ecosystem.

Thanks to these relationships it is now possible to integrate more complicated knowledges: interactions between elements which render a first level of modelling of the system.
Finally, we want to highlight that the conception of this model is still made in a generic way so that it is possible apply it to manage knowledge in different thematic fields (at least in the environmental domain). This model has been validated by computer scientists as well as thematic experts in ICZM. Currently, prototypes are implemented to represent the knowledge linked to different basic scenarios for the coastal and marine resources.
Outside the use of an ontology to complement a metadata service (to improve the quality of fulfilling and thus of data sharing) and to set up a knowledge base (for knowledge sharing), we now propose in the following section some other possible interests of such a tool

5. APPLICATIONS

From an user point of view it allows the diffusion of *knowledge bases* in particular by the way of *terminological bases* as well as friendly *semantic networks* viewers (e.g. http://www.ontopia.net/omnigator/models/index.jspOmnigator). We currently develop a similar application to display the content of our ontology as a semantic network.
It is then possible to use these knowledge modelling to be interpreted by computers and drive a broad kind of applications integrated in information systems dedicated to coastal and marine domains :

- the *sharing of the terminology* (the meaning of the concepts and the representation of their interactions : semantic properties). This is a basic step to set up information systems and common language between stakeholders. For example, information systems dedicated to fisheries can use controlled vocabularies to improve information and knowledge treatment (e.g. *FIGIS*, *FOS*, *ASFA* controlled vocabularies proposed by the *FAO*). That way it is possible to propose *requests expansion*. For example by indicating to the system that a *fish species* (*swordfish*…) has different names (*Xiphias gladius*…) which represent the same *natural element* (synonyms management) and that this species in the particular context of coastal and marine areas is linked to other elements like *fisheries elements* (*longliners*…), other species (*squids*…) or *marine geomorphological elements* (*seamount*), *marine current*, *SST*…. That way the system can drive the user by the proposal of new relevant hyperlinks or queries.
- the *thematic structuring of geographic information* in a GIS or in a spatial database for the *Web Mapping Services*. It allows in particular the management of the layers display for a

marine or coastal online atlas by using thematic tree of layers (dynamically generated by using the thematic knowledge modelling). Indeed thanks to the relationship between thematic and spatial concept (see Figure 3) we answer the need of *thematic structuring of geographical information*. That way in our web mapping tool, we use the thematic relationships inherited by the spatial concepts to generate dynamically the *layers tree* with which the users choose the layers they want to display and the geographical objects they are looking for (in order to fulfil metadata or simply improve their knowledge of the system). That way the users use a similar approach that the one with which they choose a thematic concept in our controlled vocabulary. That way we want to unify thematic classifications of the coastal and marine elements to avoid the uses of different models which complicate knowledge sharing.

- *object oriented image treatment* which uses the semantic relationships between geographical objects to improve the automatic classification process of remote sensing images (e.g. *Ecognition* software).
- the setting up of one or several spatial databases to manage reference geographic information of different coastal and marine domains contribute thus to locate existing data. Moreover, it allows the physical sharing of reference layers between different stakeholders as well as a state of the art on existing and needed data. The quality of acquired, produced or diffused information is then improved.

6. TECHNICAL REALIZATION

6.1 Metadata service: technical realization

The tools used to develop our application are Opensource and respect the reference specifications in the domain of geographic information (in particular *OGC* as well as ISO). Moreover, as these solutions are free, efficient and allow their interoperability, they are especially interesting in the interdisciplinary context of coastal and marine domains. These normalized methods ensure more durability to the proposed systems.
The setting up of spatial databases uses Opensource spatial *RDBMS* like *Postgres* and its spatial cartridge *Postgis* or *MySQL Spatial* (similar to *Oracle Spatial*).
The content of such spatial *RDBMS* is more easily treated by using Opensource and standardized Web Mapping Services (*Mapserver*, *Geoserver*, *Geotools*...) which allow the online diffusion and treatment of information (coastal and marine atlas). These tools implement *OGC* specifications (*WMS*, *WFS*, *WCS*…) and ensure the interoperability of the proposed system with other similar ones around the world.
Furthermore, the setting up of Opensource GIS clients (as UDIG, Jump, Quantum GIS...) allows remote connections to the content of these spatial DBMS and propose to the users spatial analysis functions on the geographic information collected without software licences troubles.
The relevance of these proposals is currently confirmed by the conclusions of several international projects.

6.2 Knowledge Base : Technical realization

Our conceptual formalization of knowledge uses the *UML* language which is a user graphical friendly object oriented language. *UML* facilitates the participation of the different stakeholders and can be interpreted by machines to generate different kind of codes. From a development point of view, *UML* models facilitate the implementation of the proposed models in different ways, according to the goals: into a *spatial RDBMS* (which is an interesting solution to manage together

thematic and spatial concepts with spatial functions) or by using ontologies languages of the *Semantic Web* (which are essentially XML syntaxes, e.g. *RDF* and *OWL*). In particular, there is a possibility to unify the management of metadata and ontologies by using only RDF or OWL.
SKOS language constitutes an interesting way to implement ISO 2788 and 5964 standards on the Web. Furthermore as SKOS is a RDF vocabulary, there is a close relationship with our metadata service aim. However SKOS isn't sufficient to integrate more complicated knowledge related to the concepts which constitute a thesaurus (see section 3.3.1). It means that, in order to create an ontology we need to complement SKOS language (which is possible with OWL language, for example). We are currently studying a mixed approach with a spatial RDBMS as well as SKOS files by using JENA API.
Moreover, Web Solutions are still relevant to graphically represent and diffuse the knowledge towards a broad range of people. Nevertheless, we currently still use a DBMS for the specific interest of using a spatial cartridge.

7. CONCLUSION

There is a strong complementarity between data and knowledge sharing through cooperation between a metadata service and controlled vocabularies. In particular, knowledge is basically expressed by the way of thematic and spatial as well as temporal elements. As a proof, when users search data in order to improve their knowledge, the corresponding metadata elements play a main role to summarize the content of the searched knowledge.
However, concepts related to thematic and spatial descriptions are the most difficult to control, because they differ according to the thematic domain and the geographical context of implementation. Furthermore, it is for now still impossible to find a relevant controlled vocabulary which satisfies the ICZM domain needs to share data and knowledge.
We thus propose generic models which summarize the relationships between metadata and ontologies taking care to clarify the existing relationships between thematic and spatial concepts. Moreover we think that data sharing by the way of a metadata service is an essential component to inventory the relevant concepts in a specific domain and thus to assist knowledge sharing by feeding the lack of controlled vocabularies (in addition of dedicated GUIs to manage controlled vocabularies). Conversely controlled vocabularies, especially ontologies, ensure the quality of fulfilling into a metadata service and thus the quality of data sharing. There is thus a feedback from metadata to ontologies and vice versa. We finally think that these tools (metadata service and spatial ontology) should be set up together into a single application thanks to adapted conceptual models.
Knowledge sharing in environmental domain leans on the inventory of basics interactions between elements (components) of the considered ecosystems. By translating the ecosystemic approach into a simple UML diagram it becomes possible to integrate this kind of knowledge in a compliant way both with users needs and machines interoperability (with Web semantic languages).
In practice, to share knowledge and to facilitate data sharing, we have used standards which concern controlled vocabularies (in particular Web semantic languages allow both management of metadata and ontologies) as well as knowledge representation using *semantic networks* to apply them, in a second time, to the modelling of coastal and marine systems. Indeed, ontologies will allow, in the long term, the setting up of knowledge bases in our information system.
We are currently working on the integration of the temporal dimension to improve knowledge management. That way, we will introduce a chronology into inventoried knowledge (probably as a property of a triple). This is particularly interesting and necessary to introduce links between different kinds of interactions (causes and effects) and directly related to State-Pressure-Impact-Response diagrams.

REFERENCES

Barde, J. (2005). "*Mutualisation de données et de connaissances pour la Gestion Intégrée des Zones Côtières*". *Application au projet SYSCOLAG.* PhD thesis, universite Montpellier II.

Barde, J., Libourel, T. and Maurel, P. (2005). "A metadata service for integrated management of knowledges related to coastal areas". *Multimedia Tools and APplications*, 25(3):419–429.

Barde, J., Libourel, T., Maurel, P., Desconnets, J.C., Mazouni, N. and Loubersac, L. (2005). "A metadata service for managing spatial resources of coastal areas". In *Coastgis 2005*, Data Models, Databases and Metadata, Aberdeen, Scotland. University of Aberdeen.

Barde, J., Libourel, T. and Maurel, P. (2004). "Ontologies et métadonnées pour le partage d'information géographique". *Rvue Internationale de Géomatique*, 14(2):199–217.

Coccossis, H., Burt, T. and Van Der Weide, J. (1999). "*Conceptual Framework and Planning Guidelines for Integrated Coastal Area and River Basin Management*". Split, Croatia.

Doody, J.P. (2003). "Information required for integrated coastal zone management: Conclusions from the european demonstration programme". *Coastal Management*, (3):163 173.

Gruber, T. (1993), "A translation approach to portable ontology specifications. Knowledge Acquisition", *An International Journal of Knowledge Acquisition for Knowledge-Based Systems*, vol. 5, no. 2, pp. 199-220.

Hénocque, Y. and Denis (2001). "*Des outils et des hommes pour une Gestion Intégrée des Zones Côtières, guide méthodologique*", volume IOC Manuals and Guides No. 42 UNESCO.

ISO (2003). Geographic Information – Metadata, ISO 19115, Geneva, Switzerland.

ISO 2788:1986 (1986). "*Guidelines for the establishment and development of monolingual thesauri, ISO 2788*". International Organization for Standardization (ISO), Geneva, Switzerland.

ISO 5964:1985 (1985). "*Guidelines for the establishment and development of multilingual thesauri, ISO 5964*". International Organization for Standardization (ISO), Geneva, Switzerland.

Kay, R. and Christie, P. (2001). "An analysis of the impact of the internet on coastal management". *Coastal Management*, (29):157 181.

Pullar, D. and Egenhofer, M. (1988), "Toward formal definitions of topological relations among spatial objects ", ed. Marble, D., *Proc. 3rd International Symposium on Spatial Data Handling, SDH'88, Sydney, Australia*, pp. 225-241.

Object Management Group (OMG). *http://www.omg.orghttp://www.omg.org*.

UNESCO (2003). "*A Reference Guide on the Use of Indicators for Integrated Coastal Management-ICAM Dossier 1*". IOC Manuals and Guides No. 45 UNESCO, 2003.

UNESCO (1997). "*Guide méthodologique d'aide à la gestion intégrée de la zone côtière*", volume IOC Manuals and Guides No. 36 UNESCO.