



Métadonnées et Information spatiale. Application à la gestion intégrée de la zone côtière.

Julien Barde, Thérèse Libourel Rouge, Pierre Maurel

► To cite this version:

Julien Barde, Thérèse Libourel Rouge, Pierre Maurel. Métadonnées et Information spatiale. Application à la gestion intégrée de la zone côtière.. SDN: Semaine du Document Numérique, Jun 2004, La Rochelle, France. sic_00001239

HAL Id: sic_00001239

https://archivesic.ccsd.cnrs.fr/sic_00001239

Submitted on 7 Dec 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Métadonnées et information spatiale

Application à la gestion intégrée de la zone côtière.

Julien Barde^{*,} — Thérèse Libourel^{**} — Pierre Maurel^{*}**

^{*}UMR-3S, Maison de la télédétection, 500 rue Jean François Breton, 34093 Montpellier Cedex 05

{julien, pim}@teledetection.fr

^{**}LIRMM, 161 rue Ada, 34392 Montpellier Cedex 05

libourel@lirmm.fr

RÉSUMÉ. Le travail que nous présentons a pour objectif la mise en place d'un outil d'échange d'information entre les acteurs impliqués dans la gestion du littoral du Languedoc Roussillon. Compte tenu des problèmes de sémantique sous jacents à la pluridisciplinarité dans ce domaine, nous proposons l'élaboration d'un service de métadonnées privilégiant l'échange d'information sur critères thématique et géographique. Nous avons, à ce titre, implémenté la norme ISO 19115 et souhaitons optimiser la description des éléments de métadonnées thématiques et spatiaux par le biais de vocabulaires contrôlés. L'utilisation complémentaire d'une interface cartographique permettra de visualiser et manipuler aisément les objets de référence du littoral en Languedoc Roussillon, représentations cartographiques des mots clés spatiaux. Ces derniers ont été inventoriés auprès des acteurs du domaine puis structurés de manière à modéliser et informatiser les relations qu'ils entretiennent.

ABSTRACT. This work aims to set up an information sharing tool between the stakeholders involved in the coastal management of the Languedoc Roussillon authority. Taking into account the semantic problems due to the pluridisciplinarity in this domain, we propose to set up a metadata service to share information by using thematic and geographic criteria. Indeed we consider that the use of spatial description is a good way to resolve the problem of vocabulary between the disciplines. We thus choose to implement the ISO 19115 standard. A cartographic interface associated to databases for geographic information will allow the management of the reference geographic objects for the coastal management of Languedoc Roussillon area. These objects have been collected nearby the stakeholders of this domain and have been structured with UML diagrams to modelize the interactions and relations between them.

MOTS-CLÉS : Métadonnées, interopérabilité, vocabulaire contrôlé, information géographique, gestion intégrée de la zone côtière.

KEYWORDS: Metadata, interoperability, controlled vocabulary, geographic information, integrated coastal management.

1. Introduction

Syscolag [SYS] est un programme de recherche pluridisciplinaire piloté et financé par la région Languedoc Roussillon qui recherche une amélioration de sa gestion actuelle du littoral et prône une démarche intégrée (Gestion Intégrée de la Zone Côtière, GIZC). L'objectif majeur du projet est, dans un premier temps, de construire une base de connaissances dont un des maillons essentiels est un service de métadonnées. Compte tenu de la variété des disciplines représentées dans ce projet (*géologie, sédimentologie, halieutique, écologie marine, sciences économiques, sciences politiques, géographie, informatique*), l'échange d'information s'avère difficile d'un point de vue sémantique et le support spatial est perçu comme un possible affranchissement des problèmes de vocabulaire interdisciplinaires. La proposition consiste donc à soutenir l'aspect intégré de l'échange d'information en s'appuyant en partie sur la description spatiale des ressources. Pour cela, le service de métadonnées s'appuie sur les recommandations de l'ISO 19115 [ISO] (norme pour l'échange de métadonnées d'information géographique). Nous en avons tiré un " profil " et nous complétons cette norme par un contrôle du vocabulaire qui remplit les éléments clés pour décrire les ressources ¹.

– *Une première partie* introduira brièvement l'importance du contrôle de ces éléments clés de description à travers la définition de mots-clés de référence. L'objectif est de maîtriser la terminologie des concepts employés et rejoint en ce sens l'objectif d'élaboration de thésaurus ou ontologie de domaine.

– *Une seconde partie* abordera spécifiquement le contrôle de la description spatiale. Nous avons en effet convenu dans le cadre de ce service de métadonnées, qu'à toute ressource, indépendamment de sa nature (information géographique ou non), serait attaché (durant la phase de saisie de la métadonnée) un élément contrôlé de description spatiale, de sorte qu'ultérieurement, l'interrogation de la métadonnée par critère spatial (phase de consultation) soit toujours possible. Afin de faciliter cette description spatiale des ressources, nous avons proposé de privilégier, au niveau de l'interaction avec les utilisateurs, deux moyens de saisie (et de consultation) : une interface cartographique et des listes toponymiques (de mots-clés spatiaux). Ce sont en fait deux modes d'exploitation différents d'un même travail d'inventaire des objets géographiques de référence dans le domaine de la GIZC. Nous verrons comment, sur la base des spécifications de l'ISO 19115, nous permettons facilement la mise en oeuvre de ces méthodes de description.

– *Une dernière partie* sera consacrée à la méthodologie employée pour informatiser la gestion des éléments clés de description spatiale inventoriés. En particulier, nous décrirons une base d'objets géographiques qui va permettre la réalisation des objectifs précédents. Finalement, côté utilisateur, l'interface d'interrogation permettra de coupler, selon la nature des demandes, les listes toponymiques et l'utilisation de cartes interactives.

1. Sous le vocable ressource nous regroupons des données (images, documents, statistiques, etc.), des traitements (outils d'analyse, interprétations..).

2. Contrôle des éléments de métadonnées clés pour la description de ressources

L'étude de la norme ISO 19115 montre qu'elle encourage une adaptation de ses recommandations d'ordre général selon le contexte de son implémentation [LIB 03]. En particulier, la formulation de la description ou de la recherche de ressources est, généralement, principalement exprimée par l'intermédiaire de mots clés synthétisant la pensée de l'utilisateur. Or si la norme ISO 19115 définit les descripteurs à utiliser, le choix des descriptions (mots clés en particuliers) est souvent laissée libre. Elle suggère néanmoins l'utilisation d'un thésaurus de référence pour contrôler la terminologie des différents types de mots-clés (thématiques, spatiaux, temporels...). En effet, comme le souligne [MAT 02], la description à l'aide de mots clés est une tâche subjective, puisqu'il est possible au sein d'une organisation ou d'une discipline, de trouver des descriptions différentes d'une même ressource. Il convient donc de les contrôler. Parmi ces éléments de descriptions clés : la description thématique et la description spatiale constitue la charnière essentielle de la description minimale d'une ressource (quoi, où, mais aussi quand, qui...) les deux autres éléments (temporelles et personnes concernées) étant beaucoup moins polysémiques puisque beaucoup moins subjectifs. A ce titre, nous proposons de contrôler prioritairement les vocabulaires des descriptions thématique et spatiale.

L'exercice qui revient à maîtriser la sémantique du vocabulaire employé pour remplir les éléments clés de métadonnées est donc une partie aussi importante que le choix des éléments de métadonnées (puisque complémentaire) [KAS 01]. Ce travail relève donc de la constitution d'un vocabulaire contrôlé (de manière plus ou moins approfondie selon qu'on parle de glossaire, de thésaurus voire d'ontologie) [EGE 99].

Par ces arguments nous justifions la nécessité de mettre en place des ontologies thématique et spatiale. Toutefois, sans détailler d'avantage l'élaboration de l'ontologie thématique, nous allons plutôt développer la manière dont le travail s'effectue pour la description spatiale, mais nous retiendrons qu'il existe des liens forts entre descriptions thématique et spatiale.

3. Contrôle de la description spatiale

Malgré la définition que nous avons donné au terme ressource, nous avons souhaité, dans le service de métadonnées, que chaque ressource soit décrite d'un point de vue spatiale indépendamment de sa nature.

L'ISO 19115 par ses spécifications, rend cet objectif concrétisable, et nous ne limitons donc pas la description spatiale des ressources à l'information géographique au sens strict (SIG...) mais plutôt à l'information spatiale telle que définie dans [TUR 02]. En effet, si chaque ressource n'est pas manipulable par un SIG toutes peuvent généralement profiter d'une partie au moins des descripteurs spatiaux proposés par l'ISO 19115. Il s'agit concrètement d'ajouter à la description générale d'une ressource au moins un descripteur spatial essentiel : nommer la zone concernée par la ressource, par exemple, un article de revue scientifique qui traite de l'étang de Thau. Le mot clé *étang de Thau* faisant office ici de description spatiale de la ressource de type *article*

de revue scientifique.

On note toutefois, que les recommandations de l'ISO 19115 seront pleinement exploitées pour une ressource de type "information géographique vraie". En effet, on profite alors d'un ensemble détaillé de descripteurs spatiaux proposés par la norme qui permettent de décrire en profondeur les caractéristiques propres à ce type de ressource.

3.1. Les descripteurs spatiaux

Nous allons détailler dans cette partie quels sont les principaux types de descripteurs spatiaux envisagés et de quelles manières nous comptons les manipuler.

3.1.1. Les différents types de descripteurs de l'étendue spatiale retenus par l'ISO 19115

Si on détaille les éléments de métadonnées relatifs à la description de l'étendue géographique d'une ressource dans l'ISO 19115 (voir diagramme des classes UML Figure 1), trois types essentiels sont distingués :

- *un mot (mot clé spatial ou identifiant alpha-numérique)* qui permet d'identifier de manière unique une étendue spatiale : par exemple "*étang de Thau*".
- *les limites extrêmes (Nord, Sud, Est, Ouest)* du rectangle limitant l'emprise au sol de la zone concernée par la ressource (*Bounding Box*).
- *l'objet géographique* lui-même, sous forme de la liste des coordonnées des points qui représentent les *limites exactes* de l'étendue spatiale (le premier et le dernier point de la liste étant identiques).

REMARQUE. — la *Bounding Box* n'est autre qu'une limite rectangulaire décrivant grossièrement les limites de l'étendue spatiale d'un objet, et la connaissance de l'*objet géographique* engendre donc celle de la *Bounding Box*.

3.1.2. L'inventaire des descripteurs pertinents dans notre domaine d'application

Le bon fonctionnement de l'utilisation des différents descripteurs proposés ci-dessus passera donc par un travail préalable d'inventaire des objets géographiques de référence pour le domaine de la gestion du littoral en général et plus particulièrement pour la région Languedoc Roussillon.

Sur la base de travaux de référence à l'échelle nationale [ALL 00] et par l'intermédiaire d'un stage d'analyse proposé à des DESS en informatique, nous avons entamé, à l'échelle régionale, un travail d'inventaire et de structuration des concepts auprès de chaque discipline impliquée dans ce projet (collecte des mots clés spatiaux et, autant que possible, de l'information géographique associée).

Nous comptons en extraire les principales catégories thématiques d'objets d'intérêt pour la gestion du littoral en Languedoc Roussillon, selon les différentes disciplines impliquées.

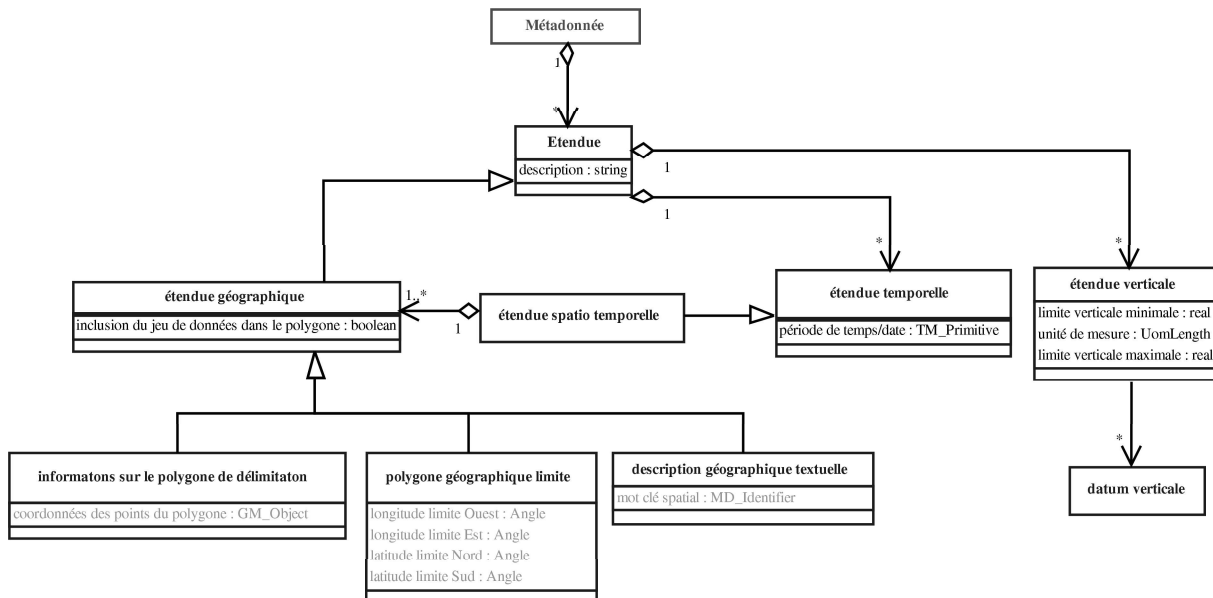


Figure 1. Les spécifications de l'ISO sur le type étendue spatio-temporelle

3.1.3. Les relations entre objets : structuration et liens transversaux

La Figure 2 illustre, à la suite d'un entretien avec des géologues impliqués dans le projet, l'inventaire et la structuration des objets géologiques jugés d'intérêts dans le domaine de la gestion du littoral du Languedoc Roussillon. Nous utilisons le langage UML afin de rendre ces diagrammes implémentables par le système.

La synthèse de l'ensemble de ces approches disciplinaires devrait nous permettre de proposer une modélisation unifiée et synthétique des concepts géographiques du domaine valorisable auprès des gestionnaires du littoral [LEH 02].

3.2. Gestion informatique des descripteurs spatiaux

Nous allons évoquer ici les choix informatiques qui vont permettre, du point de vue du système, la réalisation des objectifs, côté utilisateur, évoqués dans la partie 1.

3.2.1. La base de données associée

Conformément aux recommandations de la norme (Figure 1) nous proposons donc

6 Modèles et Applications pour le Document Cartographique Numérique (SDN 04).

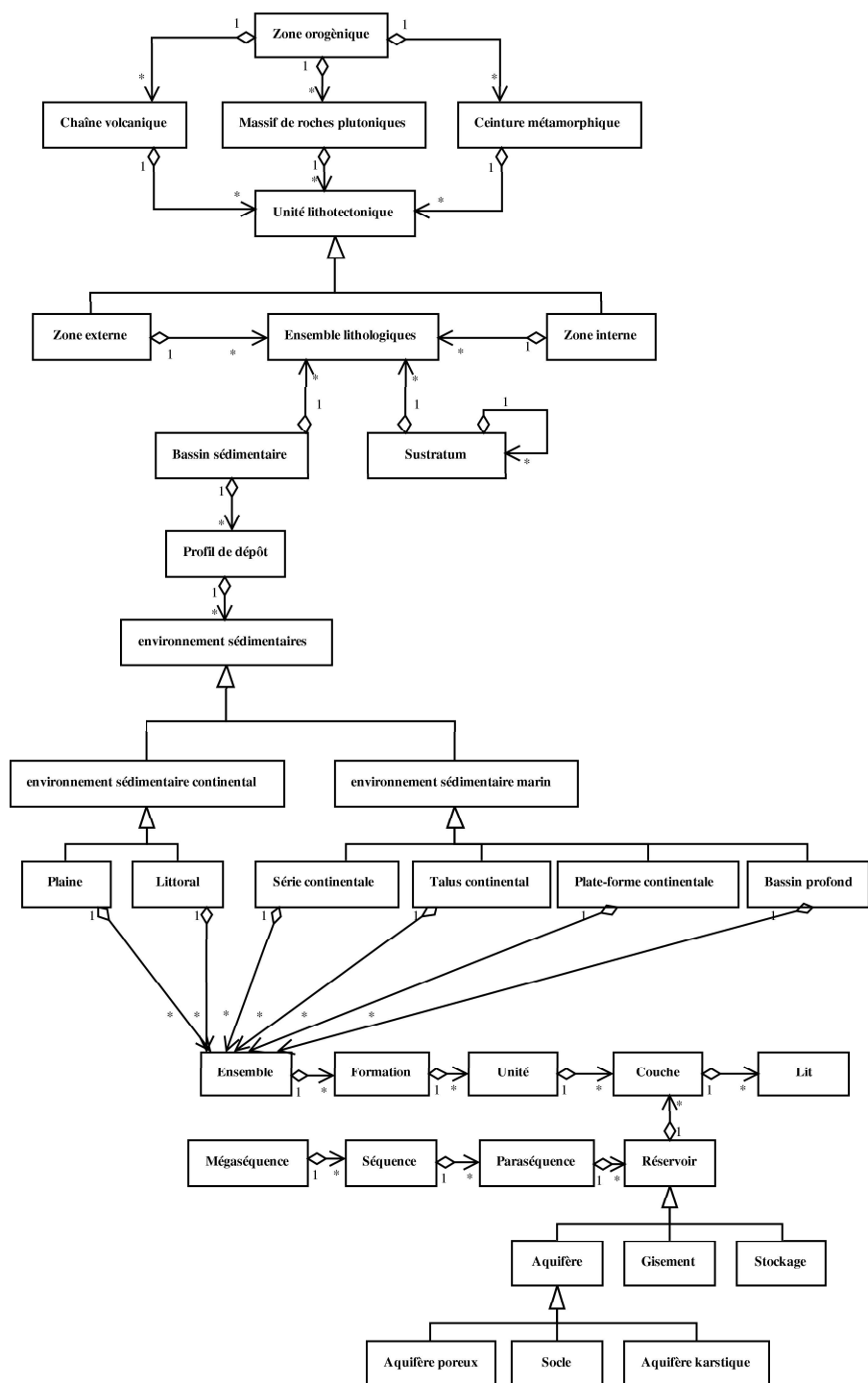


Figure 2. Modélisation des liens entre objets géographiques jugés d'intérêt en géologie pour la gestion du littoral

Table Objets géographiques
mot clé spatial/identifiant : MD_Identifier
Objet [1..*] : GM_Object
Limite Nord : Angle
Limite Sud : Angle
Limite Est : Angle
Limite Ouest : Angle
Système de référence spatial : string
Altitude minimale : real
Altitude maximale : real
Unité de mesure de l'altitude : string
Datum vertical : string
Période/Date : TM_Primitive

Figure 3. Base d'objets géographiques issue des recommandations de l'ISO (voir Figure 2)

de faire des trois types de descripteurs recommandés des attributs d'une même base d'objets géographiques tel que cela est illustré dans la Figure 2. En effet, à chaque objet géographique on peut idéalement attacher ces trois types de descriptions (*identifiant*, *Bounding Box*, *limite exacte*).

Une même table d'objets géographiques (voir Figure 3) réunira l'ensemble des objets retenus lors du travail d'inventaire et de structuration (évoqué dans la partie 3.1.2), chacun de ses enregistrements ayant donc pour attributs :

- *un identifiant*, soit alphanumérique sans signification soit un mot clé spatial comme identifiant unique (type *MD-Identifier*),
- *l'objet polygone* : ensemble de points (x,y) (type *GM-Object*),
- *les coordonnées extrêmes N, S, E, O* en degrés décimaux (type *Angle*),
- *mots clés thématiques associés* ce qui permettra de gérer des élargissements de recherche sur critère spatiale en exploitant un lien thématique en plus du seul lien géographique (par exemple une recherche sur l'*étang de Thau* sans réponse mais une proposition de réponse sur la thématique *étangs* bien qu'hors zone géographique concernée).

La liste de ces attributs est en grande partie tirée des spécifications de l'ISO sur le type *étendue* : *EX-Extent* (voir Figure 1).

3.2.2. Les logiciels utilisés

Postgres étant le logiciel utilisé pour gérer les éléments de notre base de métadonnées, et par souci de rester dans un environnement de développement libre nous

préconisons l'utilisation de *PostGis*, extension spatiale de *Postgres*, conçue spécifiquement pour la gestion d'information géographique.

3.2.3. *Exploitation du contenu de la base*

L'interrogation (ou le remplissage du contenu de la base) se fera en utilisant le langage *PHP* par le biais de listes toponymiques ou d'une interface cartographique pour désigner les objets géographiques d'intérêt à l'écran.

Là encore, nous avons opté pour l'utilisation d'une application *SVG* pour sa gratuité, la qualité du graphisme (visualisation et manipulation) et la rapidité de réponse de ces applications lorsqu'elles sont interfacées avec *PostGis*.

3.3. *Implémentation dans le service de métadonnées*

Nous allons décrire ici, du point de vue de l'utilisateur, la manière dont nous allons, en pratique, utiliser les infrastructures décrites précédemment.

3.3.1. *Décrire l'étendue spatiale lors de la saisie d'une fiche de métadonnées*

A la base de la démarche, durant la saisie d'une fiche de métadonnées, nous proposerons à l'utilisateur de décrire l'étendue géographique de la ressource :

- la première proposition revient à décrire l'étendue spatiale d'un document par le biais d'un toponyme. Nous souhaitons donc proposer des listes toponymiques (sous forme de menus déroulants) correspondant aux objets géographiques de référence pour le littoral Languedoc Roussillon (équivalent d'un vocabulaire contrôlé). Cependant l'utilisation exclusive de listes toponymiques nécessite une connaissance experte du littoral de la région.

- afin de rendre plus accessible et convivial aux personnes non expertes du domaine la description de l'étendue spatiale qui les intéresse, nous souhaitons donc proposer, comme alternative à la désignation textuelle, la visualisation des objets géographiques géoréférencés correspondants à ces toponymes par le biais d'une interface cartographique interactive où il sera possible :

- de désigner à la souris le ou les objets à retenir pour décrire une ressource.
- de matérialiser à l'écran l'emprise au sol qui décrit sa ressource. Ce dernier point est important comme possibilité de représentation propre à chacun (approche plus cognitive) lorsque l'utilisateur ne trouve pas (ni dans la liste toponymique, ni dans les objets géographiques associés dans l'interface cartographique) l'objet qui correspond à sa zone d'intérêt.

3.3.2. *Rechercher une information par requête sur critère spatiale*

De même, pour la recherche d'information, la zone d'intérêt pour lancer la recherche (élaboration d'une requête sur critère spatiale) sera effectué d'une des trois manières citées ci-dessus.

3.4. Liens entre descripteurs thématiques et géographiques

Contrairement aux mots-clés thématiques pour lesquels toutes les relations entre concepts sont à spécifier, les mots clés spatiaux disposent d'un avantage majeur puisqu'étant des objets concrets qui entretiennent nécessairement des relations d'ordre topologiques en plus des relations exprimées par les diagrammes UML tels qu'en Figure 2. Or les systèmes actuels (*SIG* par exemple) ou base de données spatiales gèrent ce type de relations [DES 01]. Ceci va permettre en pratique lorsqu'une requête spatiale sur un objet sera sans résultat de proposer une approximation de la réponse sur un critère de distance, inclusion, exclusion avec d'autres objets. Nous proposons ici un lien supplémentaire de nature thématique, géré par le biais d'un attribut thématique propre dans la base d'objets géographiques (voir partie 3.2.1). Ainsi nous permettons de rendre compte de l'intérêt de lier critères thématiques et spatiaux dans le service de métadonnées.

4. Conclusion

Par souci d'échanger de l'information dans un contexte pluridisciplinaire sur un sujet de gestion de l'environnement, nous proposons la mise en place d'un service de métadonnées permettant une description spatiale détaillée pour chaque type de ressource. L'ISO 19115 permet ces descriptions et pour l'adapter à notre domaine d'application (celui de la gestion intégrée de la zone côtière en Languedoc Roussillon), nous mettons en place des vocabulaires contrôlés destinés à maîtriser la terminologie des éléments clés de description que sont la thématique et l'étendue géographique concernée par la ressource. Cette démarche garantit la qualité du service de métadonnées. Un travail d'inventaire et de structuration des concepts géographiques manipulés par les différentes disciplines impliquées dans la gestion de la zone côtière permettra d'unifier les représentations en établissant des références partagées par les utilisateurs et en contrôlant le sens des termes de description et de recherches de métadonnées du point de vue de la machine. Nous espérons ainsi faciliter les processus de description et les problèmes de sémantique rencontrés dans ce type de projet tout en optimisant les modalités de stockage et de recherche d'information. En particulier nous souhaitons faire de la description géographique une démarche conviviale et efficace.

5. Bibliographie

- [ALL 00] ALLAIN S., GUILLAUMONT B., LE VISAGE C., LOUBERSAC L., POPULUS J., « Données géographiques de référence en domaine littoral marin », vol. SHOM/IFREMER, Version du 26/04/2000, 2000.
- [DES 01] DESCONNETS J., LIBOUREL T., P. M., A. M., M. P., « Proposition de structuration des métadonnées en géosciences », *Journées CASSINI'01, Montpellier, France*, vol. Géomatique et Espace Rural, 2001, p. pp. 69-82.

- [EGE 99] EGENHOFER M., FONSECA F., « Ontology-Driven Geographic Information Systems », (ED.) C. B. M., Ed., *7th ACM Symposium on Advances in Geographic Information Systems*, Kansas City, MO, 1999.
- [ISO] ISO, « International Organisation for Standardization (ISO), ISO 19115 Geographic Information - Metadata, Technical Committee 211. », [http ://www.isotc211.org/publications.htm](http://www.isotc211.org/publications.htm).
- [KAS 01] KASHYAP V., « Information Modeling on the Web : The Role of Metadata, Semantics and Ontologies », *Practical Handbook of Internet Computing CRC Press M. Singh (editor)*, , 2001.
- [LEH 02] LEHFELDT R., PIASECKI M., « Metadata in Coastal Information Systems », *5th International Conference on Hydrosience and Engineering*, Warsaw, 2002.
- [LIB 03] LIBOUREL T., DESCONNETS J., MAUREL P., MOYROUD N., PASSOUANT M., « Les métadonnées : pourquoi faire ? », *Géoévénement 2003*, vol. in CD-ROM, 2003.
- [MAT 02] MATA E., BAÑARES J., GUTIERREZ J., MURO-MEDRANO P., RUBIO J., « Semantic Disambiguation of Thesaurus as a Mechanism to Facilitate Multilingual and Thematic Interoperability of Geographical Information Catalogues », *5th AGILE Conference on Geographic Information Science, Palma (Balearic Islands, Spain) April 25th-27th 2002*, , 2002.
- [SYS] SYSCOLAG, « Systèmes côtiers et lagunaires », [http ://www.syscolag.org](http://www.syscolag.org).
- [TUR 02] TURBOUT C., GAIO M., MADELAINE J., « De la gestion à la réutilisation d'information géographique hétérogène multi-sources sur Internet », vol. Journées Cassini 2002, 2002.