

Méthodologie de mise en oeuvre d'un répertoire de toponymes et d'une ontologie forestière pour un accès aux ressources environnementales : application aux forêts du bassin congolais

Marius Massala, Martin Paegelow, Robert Laurini, Franck Favetta

► To cite this version:

Marius Massala, Martin Paegelow, Robert Laurini, Franck Favetta. Méthodologie de mise en oeuvre d'un répertoire de toponymes et d'une ontologie forestière pour un accès aux ressources environnementales : application aux forêts du bassin congolais. MASHS 2011, Jun 2011, Marseille, France. 9 p., 2011. <halshs-01065769>

HAL Id: halshs-01065769

<https://halshs.archives-ouvertes.fr/halshs-01065769>

Submitted on 18 Sep 2014

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.

METHODOLOGIE DE MISE EN OEUVRE D'UN REPERTOIRE DE TOPONYMES ET D'UNE ONTOLOGIE FORESTIERE POUR UN ACCES AUX RESSOURCES ENVIRONNEMENTALES: APPLICATION AUX FORETS DU BASSIN CONGOLAIS

Marius Massala* - Martin Paegelow* - Robert Laurini - Franck Favetta****

*Université de Toulouse II le Mirail

Maison de la Recherche,

5, Allées A. Machado, 31058 Toulouse, Cedex 9, France

mariumassala@gmail.com, paegelow@univ-tlse2.fr

**INSA de Lyon

Bat B. Pascal, av. A. Einstein F-69622 Villeurbanne Cedex, France

{prénom.nom}@liris.cnrs.fr

Résumé – *L'exploitation des données de télédétection a pour conséquence la production massive des données géographiques. Les acteurs impliqués dans leur gestion utilisent des catalogues d'archivage basés sur des normes géographiques complexes ne prenant pas en compte l'évolution spatiale, temporelle et des noms des lieux des objets géographiques. Dans cet article, nous présentons une solution basée sur l'exploitation des toponymes avec une application pour les pays d'Afrique centrale. La solution repose sur un répertoire de toponymes et une ontologie forestière intégrant des relations topologiques qualitatives. L'application finale est de fournir aux acteurs impliqués dans la gestion des forêts un outil qui facilite le suivi des objets géographiques et in fine faciliter l'accès aux ressources informationnelles.*

Mots clés – **Objet géographique, répertoire des toponymes, métadonnées, ontologie, forêt**

1 Introduction

Ces dix dernières années ont été fortement marquées par une prise de conscience planétaire sur les problèmes liés à l'environnement. Les sommets de Rio (1992), de Johannesburg (2000) ainsi que les initiatives régionales telles que le Partenariat des forêts du Bassin du Congo, témoignent de cette prise de conscience de la communauté internationale et des gouvernements sur la nécessité d'assurer un meilleur devenir aux écosystèmes forestiers très menacés. Les forêts d'Afrique Centrale représentent le deuxième plus grand massif boisé après celui de l'Amazonie, et constituent de ce fait un patrimoine de l'humanité toute entière [21,11]. Le développement des Etats financé à grands budgets, conjugué à l'urbanisation galopante des villes et à l'augmentation de la population, s'accompagnent des problèmes environnementaux qui se posent avec acuité. La dernière conférence sur le climat tenue à Copenhague (2009) a permis de mettre en lumière les enjeux liés à la préservation de l'environnement et des écosystèmes. C'est dans ce cadre que certains pays d'Afrique Centrale ont mutualisé leurs moyens pour mettre en œuvre des procédures de sauvegarde de leur écosystème forestier afin de mieux remplir leurs engagements liés au développement durable. A cet effet, ils se sont engagés à gérer leur patrimoine forestier en s'appuyant sur des données de télédétection et des bases de données communes. Une agence d'étude et de

centralisation des informations géographiques a été ainsi mise en place : l'Agence pour le Développement de l'Information Environnementale (ADIE), qui a pour objectif principal la mutualisation de l'information environnementale récoltée par elle-même et par les Organisations Non Gouvernementales (ONG) spécialisées, les ministères en charge des forêts, les institutions et les projets. L'ADIE a été très vite confrontée au problème de la gestion des flux massifs de ces ressources. La gestion des données colossales dont les études peuvent s'étendre sur plusieurs thématiques, et aussi sur plusieurs localités et couvrant plusieurs types de végétations aux limites parfois floues est devenu problématique. Face à cette difficulté, la nécessité de disposer d'outils d'analyse de l'état des ressources naturelles face aux actions anthropiques est devenue une exigence; l'objectif visé étant d'améliorer la gestion et évaluer l'impact des actions de développement sur l'environnement. Le présent article présente le processus de modélisation d'une solution qui doit servir à l'ADIE pour accéder aux ressources informationnelles. Pour atteindre cet objectif, nous nous sommes appuyés sur les répertoires des toponymes et les ontologies. L'utilisation de ces deux concepts (ontologie et toponymes) semble être à nos yeux des pistes prometteuses dans le suivi de la dynamique spatiotemporelle des objets géographiques et aussi pour une recherche efficace des ressources informationnelles. L'Outil de Gestion des Forêts Equatoriales (Ogefe) résultant de ce travail et actuellement en cours de développement au laboratoire GEODE de l'Université Toulouse II Mirail, se veut d'abord un outil d'appui aux acteurs de la sous-région d'Afrique centrale. Il est destiné aux membres de la COMIFAC¹ pour leur servir d'outil d'aide à la décision pour une gestion efficace de tout le massif forestier du bassin du Congo.

2 Synthèse des travaux existants dans le domaine

La plupart des solutions qui portent sur la gestion des ressources géographiques se concentrent surtout sur la conception des catalogues des métadonnées. On en dénombre plusieurs dont les plus connus sont Geoseource, MDweb, Ecofor, Géoportail, Infoterre, Holloscope, etc. Comme le mentionne [1,13] de nombreux outils sont actuellement développés par les organismes producteurs d'informations géographiques, par des collectivités ou des services de l'Etat. Tous les travaux ainsi menés portent sur le catalogage et l'archivage des données géoréférencées, en s'appuyant sur les normes ISO 19115 et FGDC² [13]. Certains de ces outils ont intégré en plus du catalogage, la notion d'observatoire. C'est le cas par exemple du programme ROSELT (Réseau d'Observatoires de Surveillance Ecologique à Long Terme) mis en œuvre en Afrique par l'IRD³ qui utilise l'outil MDWeb [9] pour cataloguer les métadonnées. Il y a aussi l'Observatoire du Domaine Côtier (ODC), développé à l'Institut Universitaire Européen de la Mer dont les objectifs sont de recueillir des séries d'observations de la mer d'Iroise et de fournir aux scientifiques les données nécessaires à toutes interprétations se référant à l'évolution de cet écosystème [13]. Du côté de l'Afrique de l'Ouest précisément en Guinée Conakry, un observatoire pour le développement territorial adapté au contexte local a fait l'objet d'une analyse conceptuelle mais faute de moyens financiers, celui-ci n'a pas été finalisé [5]. Si les catalogues et les observatoires développés présentent des intérêts évidents en termes d'archivage, ils sont néanmoins limités pour la diffusion des connaissances environnementales. Cette insuffisance est en partie liée à l'absence d'éléments relatifs aux sources, à l'historique, à la généalogie : c'est-à-dire à l'« environnement » de l'information (ou traçabilité) [13]. De plus, la complexité et la lourdeur des normes sont parfois des facteurs décourageant pour les acteurs impliqués dans le processus de catalogage. Enfin, la majeure partie des outils n'intègrent pas toujours les évolutions spatio-temporelles et toponymiques

¹ Conférence des Ministres en charge des Forêts d'Afrique Centrale

² Federal Geographic Data Commmity

³ Institut de Recherche et de Développement

des objets géographiques que les ressources informationnelles sont sensées décrire. Ce dernier paramètre d'évolution des objets est capital à nos yeux. En effet, comme le souligne [4] « *l'espace géographique est un espace évolutif au sein duquel prennent souvent place des dynamiques. Or les représentations actuelles de l'information géographique négligent souvent cette composante et fournissent une description statique de l'espace géographique* ». En définitive toute modélisation portant sur le territoire passe nécessairement par la mise en évidence de tous les objets impliqués, de la caractérisation de leur dynamique et de leurs interactions [14] in [5] ; c'est le but visé par nos travaux.

3 Les objets géographiques et le répertoire des toponymes dans la solution OGEFE

Notre proposition cherche à faire coexister les noms d'objets spatiaux aux limites géométriques exactes ou imprécises avec un ensemble de concepts reliés par des relations qualitatives. Nous prenons en compte deux types de relations : *toponymique et topologique*. La solution repose sur un modèle de représentation des dynamiques spatio-temporelles et toponymiques des objets. Il sert de socle à un SRIS⁴ dont les finalités sont le suivi spatio-temporel des objets et l'indexation sémantique des ressources informationnelles associées.

3.1 Les objets géographiques

Les experts représentent en général l'espace géographique sous forme d'objets [4]. C'est dans cette logique que nous avons représenté les objets géographiques du territoire cible de notre étude (sous région d'Afrique centrale). A cet effet, les objets géographiques ont été classifiés en deux catégories à savoir : les objets de type polygone et les objets linéaires. Tous les objets sont décrits par des géométries intégrant des limites minimales et maximales. Cette dernière limite (la maximale) est variable en fonction du temps. Ce mode de représentation des objets dit « *du jaune d'œuf* » permet de spécifier les incertitudes aux frontières pour les objets aux limites vagues [10] in [6]. C'est le cas par exemple des forêts ou de certaines régions urbaines. Toutes les géométries des objets se présenteront sous forme de couches vectorielles préalablement numérisées à l'aide des logiciels SIG classiques comme MapInfo ou ArcGIS. Chaque objet du territoire est référencé par un concept issu de l'ontologie et pointe vers un ensemble de noms provenant du dictionnaire des toponymes.

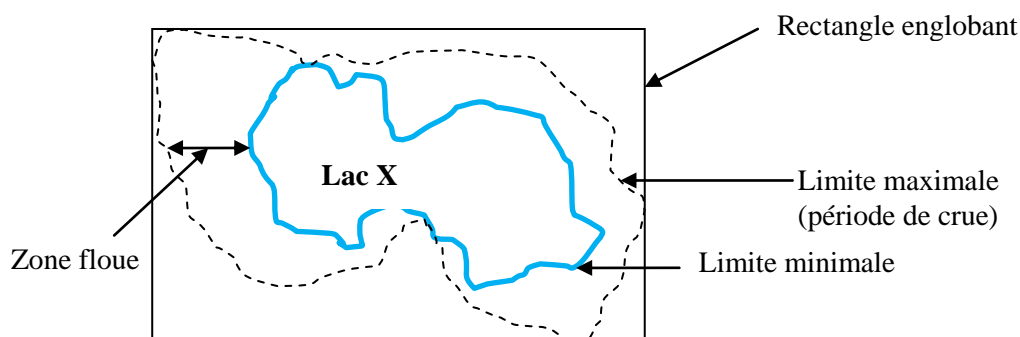


Figure 1. Exemple de représentation d'un objet géographique en «jaune d'œuf»

⁴ SRIS : système de recherche d'informations spatiales

3.2 Le répertoire des toponymes

Un toponyme est le nom officiel d'un lieu ou d'un site [17]. L'utilisation des toponymes permet d'identifier et/ou de localiser les objets géographiques [17]. Dans la modélisation que nous proposons, leur utilisation permet de prendre en compte non seulement les variantes linguistiques mais également les synonymes dialectaux et leurs évolutions historiques. Dans cette modélisation, l'entité *Appellation* permet à chaque entité géographique d'intégrer des variantes nominales sur le plan linguistique et aussi connaître des changements d'appellation tout au long de son histoire. En outre, le modèle permet d'avoir pour un objet plusieurs géométries et d'établir une relation d'interdépendance entre les changements géométriques et toponymiques.

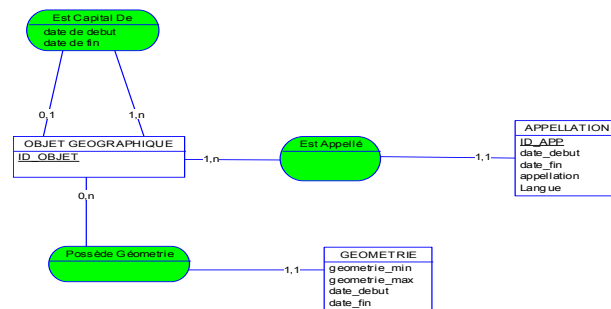


Figure 2. Modèle partiel des dynamiques spatio-temporelles et toponymiques des objets

4 Modélisation d'une ontologie appliquée au bassin du Congo

Une ontologie se définit comme un ensemble de concepts (classes) et les relations existant entre ces concepts [3]. Elle est définie aussi en tant qu'une spécification explicite d'une conceptualisation souvent considérée comme un modèle réutilisable et partageable [15]. En des termes simples, c'est le regroupement des connaissances et concepts relatifs à un domaine donné et les relations qui peuvent les entourer. Dans le domaine géographique, les ontologies et les répertoires des toponymes doivent intégrer des relations topologiques qualitatives et quantitatives [16, 17, 18]. Cela se matérialise par la définition des relations topologiques tournées vers les différents types de trait : par exemple une île doit être entourée d'eau, une gare doit être à côté d'une ligne de chemin de fer, une mangrove ne peut être qu'à côté de la mer etc. L'ensemble des relations permises peut aussi contenir des restrictions juridiques par exemple, une concession d'exploitation forestière et un parc ne peuvent pas se superposer entre eux, ou l'un ne peut contenir l'autre. La spécification de telles relations permet de résoudre les problèmes pour lesquels l'approche par catalogue est inefficace. Pour ce qui nous concerne, il s'agit d'inventorier et mettre en relation l'ensemble des concepts portant sur la forêt continentale du bassin du Congo.

Après avoir analysé les documents bibliographiques, nous avons tiré une taxonomie des concepts et des modèles génériques des relations topologiques qui peuvent exister entre eux. C'est ce que nous présentons dans la section ci-dessous.

4.1 Choix des concepts et mise en œuvre de la taxonomie

Modéliser les objets et les concepts du territoire dans le cadre de la conception des répertoires des toponymes est toujours une tâche ardue si les objectifs de départ ne sont pas suffisamment explicites. Comme le souligne [8], toute conception d'un modèle faisant référence au territoire doit nécessairement prendre en compte les différents points de vue que sont le milieu naturel, les

aménagements et les activités humaines, et enfin, les organisations sociales et/ou administratives. Une telle démarche permet de bien structurer l'espace et de mettre en cohérence les différents concepts qui le composent [5, 1, 9]. Dans notre cas, le choix des concepts de l'ontologie s'est fait sur la base du découpage territorial qui regroupe à la fois le découpage administratif et le découpage forestier. Ensuite, nous avons intégré le couvert végétal de la région et les différentes classes constituant le réseau hydrographique. Pour construire la taxonomie des concepts, le territoire a été découpé en quatre superclasses. La première est l'**espace administratif**, classe dans laquelle on retrouve le découpage administratif des pays. Nous avons appelé cette zone « *espaces administratifs* », et on y trouve les provinces, les départements, villes et villages. Ce découpage s'inspire de celui du Gabon. Ensuite on a l'**espace à réglementation**: ce sont des zones à infimes densité humaine. Nous les avons appelée « *espace à réglementation* ». On retrouvera particulièrement les concessions forestières (CFAD⁵, PFA⁶, RFP⁷), les parcs forestiers, les réserves et jardins d'ordres botanique, zoologique, etc. L'**espace en eau** : classe qui regroupe le réseau hydrographique. On y trouve les fleuves, rivières, torrents, lacs, les eaux maritimes et lagunaires. Enfin, l'**espace en terre ou d'occupation du sol**: ce sont les différents types de végétation que l'on rencontre dans chaque pays. A cet effet, nous avons repartis l'espace forestier gabonais en sous-concepts à savoir : La forêt, les savanes, les steppes, les mangroves, les plaines du littorales et les forêts inondées. La figure 3 ci-dessous présente quelques uns des concepts de la future ontologie.

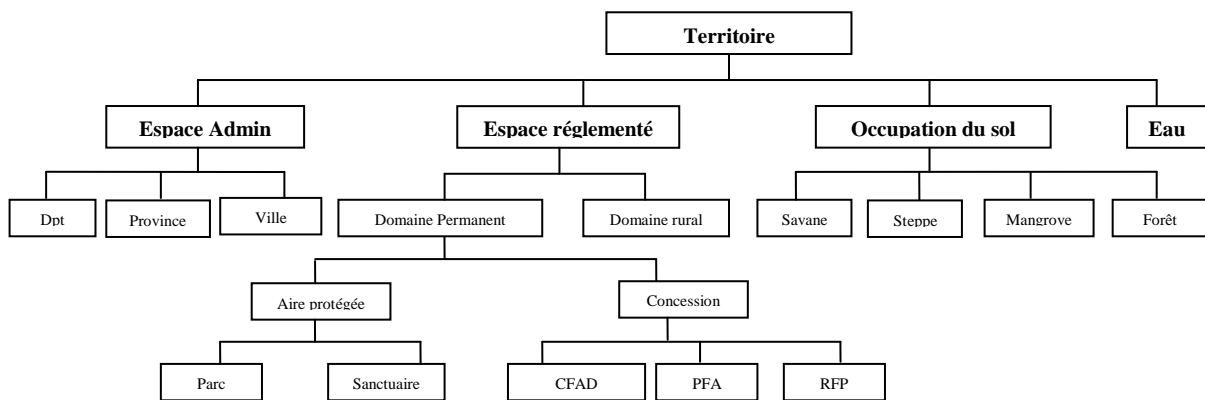


Figure 3. Taxonomie de quelques concepts dans le cadre d'un plan d'aménagement du territoire gabonais.

4.2 Limites floues et fixes dans les concepts

Les concepts peuvent être distingués en deux types à savoir : les concepts décrivant le paysage dont les limites sont difficilement identifiables, et les concepts dont on peut déterminer les limites. Cette distinction entre les concepts s'est faite sur la base des lois forestières mises à notre disposition, et aussi grâce aux orientations des acteurs de l'ADIE et de l'équipe de recherche participant aux présents travaux. Au nombre des concepts aux limites fixes, nous avons les sous-classes du concept *espace administratif*. Ensuite viennent les classes appartenant aux *zones à*

⁵ Concession Forestière sous Aménagement Durable (permis d'exploitation de grande taille)

⁶ Permis Forestier Associé (permis industriel de petite dimension)

⁷ Réserves Forestières de Production (forêt mis en attente avant attribution)

réglementation comme les parcs forestiers, et les concessions d'exploitations forestières. Toutes ces zones font l'objet d'une délimitation claire et précise par les autorités administratives. Par contre, un concept présente des limites floues lorsqu'il existe une zone de transition aux limites imprécises. Par exemple la limite géographique exacte entre une mangrove et une forêt inondée n'est pas clairement déterminable. Si sur le plan de la représentation cartographique cela semble claire, dans la réalité par contre, celle-ci n'est pas triviale. Aux nombres des concepts ayant des limites floues nous avons *les espaces en terres* ou *catégoriels* (forêt primaire et secondaire, la savane, mangrove, les steppes, etc.), et les *espaces en eaux*. Pour ce qui concerne les espaces en eau, nous avons pris en compte les zones de confluence entre les sous-classes du concept *espace en eau*.

4.3 Relations topologiques entre concepts

Nous avons ici défini un modèle générique en le spécifiant toutefois au cas du bassin Congo et particulièrement du Gabon. A cet effet, nous avons défini des opérateurs binaires pour modéliser l'espace. Ce sont les opérateurs *Meet*, *Touch*, *overlap*, *include*, *CoveredBy*. Ces opérateurs qui résultent des différents modèles topologiques existant tels qu'ils ont été défini par [12], serviront à définir des associations qualitatives spatiales entre les différentes régions de l'ontologie. Il faut retenir qu'une relation est dite qualitative lorsqu'elle fait intervenir des concepts spatiaux sans toutefois disposer de leur géométrie [19, 2]. Cette définition cadre parfaitement avec notre modèle puisque dans notre cas, il n'est pas nécessaire de disposer implicitement d'une connaissance sur la géométrie des régions, mais plutôt de définir les relations d'interdépendance entre les concepts. Le choix des cinq opérateurs topologiques ci-dessus est lié au caractère des concepts que nous utilisons. Ainsi, l'opérateur *Meet* par exemple, met en présence deux concepts aux limites fixes, pouvant être contigües et qui ne se chevauchent pas. C'est le cas des *parcs animaliers* et des *concessions d'exploitations forestières*. L'opérateur *Touch* joue le même rôle que le *Meet*, mais est dédié aux concepts dont les limites sont vagues et à la classe *espace en eau*. *Touch* sera par exemple utilisé pour définir les relations *lac – fleuve* ou *mangrove – eaux marines*. Il faut mentionner que l'utilisation de cet opérateur (*Touch*) met en exergue des contraintes qu'il est important de respecter notamment les relations de proximités entre les concepts. La figure ci-dessous présente quelques relations que l'on peut avoir avec l'opérateur *Meet*.

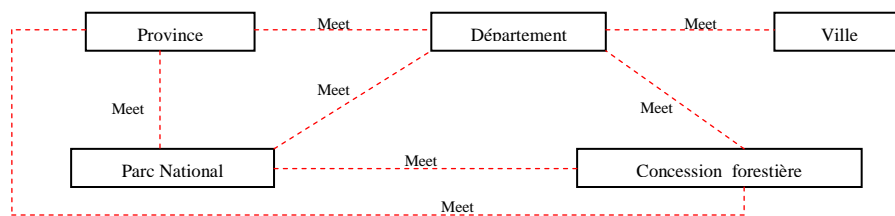


Figure 4. Exemple de relations de proximité possible avec *Meet*

Une fois les relations clairement établies, elles peuvent être modélisées comme indiqué dans la figure 5. Le modèle s'appuie sur les entités *Concept* et *Relation Topologique*. Toutefois, nous avons intégré l'entité *Objet géographique* que nous avons présenté en figure 2. L'entité *Concept* est l'élément central de cette partie du modèle. Il comprend deux relations réflexives qui prennent respectivement en compte les supers et les sous-classes d'un concept et aussi les relations qu'il (le concept) peut avoir avec d'autres concepts. Aussi, la modélisation permet également d'attribuer un type à chaque objet géographique.

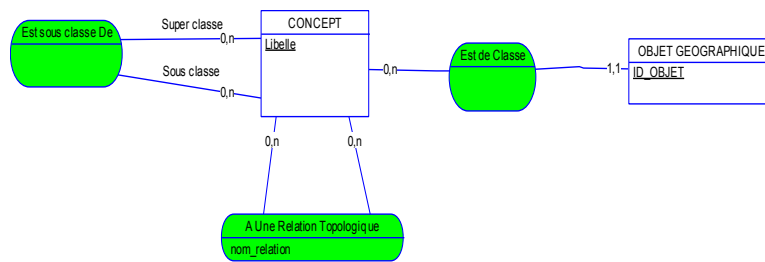
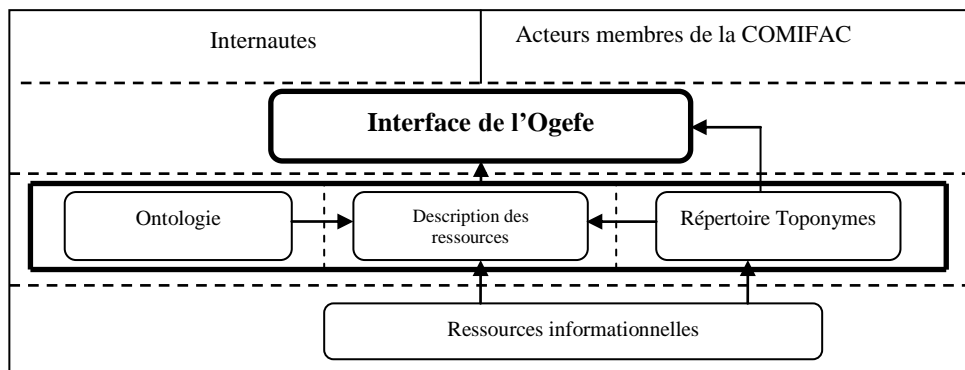


Figure 5. Portion de modèle formalisant les relations topologiques entre les concepts de l'ontologie

5 Architecture de l'outil

Pour arriver à mettre en place les objectifs que nous avons présentés, nous envisageons de développer une solution globale qui prenne en compte toutes les caractéristiques que nous avons eues à énoncé ci-dessous. Le système que nous mettrons en place est un *service de toponymes* qui se présente sous la forme d'un *portail web*. Il comportera une partie utilisateur au sein de laquelle tous les internautes auront accès, et aussi une partie fédérative et descriptive où l'on retrouvera à la fois le répertoire des toponymes, l'ontologie et la description des ressources. Enfin la dernière et troisième partie est un serveur de fichiers qui sert de rattachement physique des ressources à la base de données. La **figure 6** ci-dessous présente le schéma architectural de l'outil Ogefe.



6 Conclusions et perspectives

La démarche méthodologique que nous venons de décrire dans cet article, présente les contours de la solution que nous sommes entrain d'implémenter. L'outil *OGEFE* a l'avantage d'intégrer non seulement les dimensions lexicographiques et les évolutions toponymiques des noms des lieux, mais aussi les évolutions spatio-temporelles des objets géographiques. Notre solution permet de typer les toponymes au moyen d'une ontologie. Aussi, l'usage des relations topologiques dans l'ontologie, est un atout majeur qui facilitera l'accès aux ressources informationnelles dans la future plate forme.

Références

- [1] Barde Julien (2005), *Mutualisation des données et des connaissances pour la gestion intégrées des zones côtières. Application au projet SYSCOLAG*, Thèse de doctorat, Université Montpellier II.
- [2] Bedel Olivier (2009), *GEOLIS: Un Système d'information logique pour l'organisation et la recherche de données géolocalisées*, Thèse de doctorat, Université de Renne I.

- [3] Belhadef Hacene., Mohamed-Khireddine K (2007), *Conception d'une nouvelle ontologie pour l'interopérabilité des systèmes d'informations géographique*, 4th International Conference Sciences of Electronic, Technologies of Information and Telecommunications, Tunisia, March 25-29.
- [4] Bucher Bénédicte (2002), *L'aide à l'accès à l'information géographique: un environnement de conception coopérative d'utilisations de données géographiques*, Thèse de doctorat, Université Paris 6.
- [5] Camara A., Michel Passouant., JP Cheylan (2010), Conception d'un observatoire pour le développement territorial, *Revue internationale de géomatique*, **volume 20**, N° 3, p. 331-361
- [6] Cohn A.G and Gotts N.M (1996), the 'egg-yolk' representation of regions with indeterminate boundaries, Burrough and Frank.
- [7] COMIFAC (2007), *Rapport de la mission d'experts pour la réflexion sur la conception d'un système de suivi-évaluation du plan de convergence de la COMIFAC*, Yaoundé-Cameroun, 21 Mars.
- [8] De Sede M-H., Moine A (2001), Systémique et bases de données territoriale, des concepts et des outils pour une gestion raisonnée des territoires, SIG et aménagement, *Revue internationale de géomatique*, **volume 11**, N° 3-4, p. 333-358
- [9] Desconnets JC, Libourel Thérèse, S. Clerc (2007), *Cataloguer pour diffuser les ressources environnementales*, Actes de la conférence INFORSID 2007, Perros-Guirec, p. 344-361.
- [10] Dumoncel Franck (2006), *Interaction pour exprimer des requêtes spatiales guidées par des adjacences conceptuelles*, Thèse de doctorat, Université de Cean.
- [11] Dupuy B., Gérard C., Maître H.-F., Marti A., Nasi R (1998), *Gestion des écosystèmes forestiers denses d'Afrique tropicale humide*, Collection les bibliographies du Cirad, volume 9, p.207
- [12] Egenhofer M (1993), A model for detailed binary topological relationships, *Geomatica*, **volume 47**, p. 261-273.
- [13] Faucher C., Gourmelon F, Lafaye J-Yves., Rouan M (2009), Mise en œuvre d'une mémoire environnementale adaptée aux besoins d'un observatoire du domaine côtier : MENr, *Revue internationale de géomatique*, **volume 19**, N° 1, p. 7-26
- [14] Gate O., Libourel T., Cheylan J-P., Lardon S (1997), *Conception des systèmes d'information sur l'environnement*, Hermes, Paris 1997.
- [15] Gruber Thomas Robert (1993), A translation approach to portable ontology specifications, *Knowledge Acquisition*, **volume 5**, N°2, p. 199–220.
- [16] Jaillard Pierre (2008), *Entre guerres et conflits : la planète sous tension*, Festival international de géographie, Saint Dié-des-Vosges, SIG la lettre N° 10.
- [17] Janowicz K., Kebler C (2008), The role of ontology in improving gazetteer interaction, *Int. Journal of Geographical Information Science*, **volume 22**, N° 10, p. 1129–1157.
- [18] Janowicz K., Schwarz M., Wilkes M. (2009), *Implementation and Evaluation of a Semantics-based User Interface for Web Gazetteers*. At the Visual Interfaces to the Social and the Semantic Web (VISSW 2009) Workshop in conjunction with the International Conference on Intelligent User Interfaces (IUI 2009). *CEUR-WS Proceedings Volume 443*, Sanibel Island, Florida.
- [19] Miron Alina Dia., Jérôme Gensel., Marlène Villanova-Oliver., Hervé Martin (2007), *Relations spatiales qualitatives dans les ontologies géographiques avec ONTOAST*. In Int. Conf. on Spatial Analysis and GEomatics - SAGEO, p. 110
- [20] Proulx Marie José., Yvan Bédard., François Létourneau (2003), Information géospatiales dans internet : Application pour un contexte de renseignement militaire, *European Journal of GIS and Spatial Analysis - Revue internationale de géomatique*, **volume 13**, N° 3, p.323-338.
- [21] WRI (2000), *A first Look at Logging in Gabon*. World Resources Institute, Washington, p.50.