



Conception et élaboration d'un système d'information pour la gestion des ressources en géomatique

Florence Dujardin

► To cite this version:

Florence Dujardin. Conception et élaboration d'un système d'information pour la gestion des ressources en géomatique. Génie logiciel [cs.SE]. 2010. dumas-00538777

HAL Id: dumas-00538777

<https://dumas.ccsd.cnrs.fr/dumas-00538777>

Submitted on 23 Nov 2010

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.

CONSERVATOIRE NATIONAL DES ARTS ET METIERS
CENTRE REGIONAL ASSOCIE DE SAINT-ETIENNE

MEMOIRE

présenté en vue d'obtenir
le DIPLOME D'INGENIEUR CNAM

SPECIALITE : INFORMATIQUE
OPTION : SYSTEMES D'INFORMATION

par

Florence DUJARDIN

Conception et élaboration d'un système d'information pour la
gestion des ressources en géomatique

Soutenu le 23 septembre 2010

JURY

PRESIDENT :	M. Fouad BADRAN	<i>Professeur CNAM Paris</i>
MEMBRES :	M. Eric FAVIER	<i>Professeur C.R. CNAM de Saint-Etienne responsable filière informatique</i>
	M. Bernard BONCHE	<i>Professeur C.R. CNAM de Saint-Etienne</i>
	M. Gilles RUSAOUEN	<i>Professeur C.R. CNAM de Saint-Etienne</i>
	M. Didier GRAILLOT	<i>Directeur de recherche ENSMSE</i>
	M. Philippe BEAUNE	<i>Maître assistant ENSMSE</i>

Remerciements

Je tiens à exprimer toute ma reconnaissance ainsi que mes remerciements aux personnes qui, tant sur un plan professionnel que personnel, ont contribué à la réussite de ce travail.

Je remercie particulièrement Didier GRAILLOT, qui m'a permis de mettre un terme à un diplôme débuté il y a bien longtemps.

Je remercie Roland DECHOMETS grâce auquel ce travail a vu le jour.

Je remercie Philippe BEAUNE pour ses conseils avisés, ses idées et sa patience.

Je remercie très sincèrement Eric FAVIER pour sa disponibilité.

J'adresse mes plus vifs remerciements à Messieurs Fouad BADRAN, Bernard BONCHE et Gilles RUSAOUEN pour avoir accepté d'évaluer ce travail et d'être membres du jury.

Je tiens à exprimer ma reconnaissance et ma gratitude à tous ceux qui ont collaboré et rendu possible la réalisation de ce travail, particulièrement aux membres du centre SITE (Sciences, Information et Technologies pour l'Environnement) avec lesquels je souhaite continuer à collaborer.

Merci à l'Ecole Nationale Supérieure des Mines de Saint-Etienne dans laquelle j'évolue depuis 1991.

Merci au Conservatoire National des Arts et Métiers, et plus particulièrement au centre régional associé de Saint-Etienne et notamment à Corinne, Caroline et Sophie.

Merci à Adrien, Carmen, Catherine, Ambre, Laure, Christine, Christèle, Joël, Françoise, Mireille, Aurélien, Djamel, Frédéric, Jordan, Gaël, Florent, Christiane, Hervé, Éliane, Arlette, Ghislaine, Véronique, Agnès, Valérie, Pascale, Eric, Zahia, Valentine, Mariana et tant d'autres...

Enfin, merci à toi Eva, ma fille, pour avoir été aussi patiente et solidaire durant ces longues journées de vacances...

Florence DUJARDIN

Table des matières

Remerciements	2
Table des matières	3
Liste des figures.....	5
Liste des tableaux	6
Table des annexes.....	7
Liste des abréviations	8
Glossaire	12
Introduction et objectifs du travail.....	14
PARTIE 1 CONTEXTE DU PROJET	15
CHAPITRE 1 – SITE : UN CENTRE DE FORMATION ET DE RECHERCHE EN INGENIERIE DE L'ENVIRONNEMENT	17
1. L'Ecole Nationale Supérieure des Mines de Saint-Etienne	17
L'Ecole en quelques chiffres (source rapport d'activité 2009)	17
2. Les centres de recherche de l'Ecole.....	20
3. Le centre SITE	21
SITE en quelques chiffres (source rapport d'activité 2009)	22
4. Le département MIRI	25
5. Le département IDEE.....	26
Quelques partenaires d'IDEE.....	27
6. IDEE et les ressources en géomatique	28
7. L'acquisition de données	29
CHAPITRE 2 – ENVIRONNEMENT ORGANISATIONNEL ET MATERIEL	31
1. Organisation du centre SITE	31
Un ingénieur de recherche au centre SITE.....	31
2. Environnement informatique du centre SITE.....	31
Un technicien informatique au centre SITE	31
En support : le service informatique.....	32
3. Matériel informatique et logiciels utilisés par le centre SITE	34
Matériel informatique	34
Logiciels	36
CHAPITRE 3 – L'ORGANISATION DE L'INFORMATION GEOGRAPHIQUE NUMERIQUE D'IDEE	38
1. Historique et organisation actuelle	38
2. Gestion actuelle des données.....	40
PARTIE 2 PRESENTATION DU PROJET ET ANALYSE DES BESOINS	45
CHAPITRE 1 – ETAT DES LIEUX ET ANALYSE DES BESOINS	47
1. Constat	47
2. Les acteurs du projet	48
3. Utilisation des ressources géomatiques existantes	50
4. Un système d'information de métadonnées	53
5. Facteurs de risques et de réussites	55
CHAPITRE 2 – NATURE DE L'INFORMATION : DES DONNEES AUX METADONNEES	58
1. Description de l'information géographique utilisée.....	58
2. La norme internationale ISO 19115	60
La norme française NF EN ISO 19115	65

Autres normes ou directives	65
CHAPITRE 3 – LE MODELE CONCEPTUEL DE (META)DONNEES	69
1. Le cycle de vie de la métadonnée géographique	69
2. Modèle Conceptuel de (méta)Données proposé.....	69
3. Modèle Conceptuel de (méta)Données avec prise en compte des normes en vigueur ISO 19115	73
PARTIE 3 REALISATION ET MISE EN ŒUVRE DE L’OUTIL	79
CHAPITRE 1 – OUTILS EXISTANTS	81
1. MDweb.....	81
2. Geonetwork	82
3. Géosource.....	84
4. ArcCatalog	86
5. Ce qui existe dans le laboratoire : le géorépertoire de la ZABR	88
CHAPITRE 2 – SOLUTION CHOISIE	90
1. Les différentes alternatives.....	90
2. Les différents profils utilisateurs	92
3. Analyse financière.....	93
4. Les préférences des utilisateurs	93
5. Formation	94
CHAPITRE 3 – MISE EN ŒUVRE	96
1. Installation.....	96
2. Création de la métadonnée	98
3. Test sur un lot de données particulier	100
4. Gains attendus pour les usagers	100
Conclusion et perspectives	102
Bibliographie	104
Annexe 1 Formulaire de saisie des métadonnées au format Inspire via Excel.....	106
Annexe 2 Copies d’écran des interfaces graphiques d’édition des métadonnées de ArcCatalog.....	110
Annexe 3 Copies d’écran des interfaces graphiques des rubriques de saisie des métadonnées sous GeoNetwork, vue avancée	113

Liste des figures

Figure 1 : organigramme ENSMSE (source rapport d'activité 2009 ENSMSE)	19
Figure 2 : organigramme de l'UR SEPIT (source rapport d'activité 2009 Centre SITE)	24
Figure 3 : illustration des ressources informatiques du centre SITE.....	36
Figure 4 : illustration de l'arborescence de données géographiques sur SUN	40
Figure 5 : illustration de l'arborescence de données géographiques sur le serveur Windows	42
Figure 6 : schématisation des modalités de travail actuelles des enseignants-chercheurs	52
Figure 7 : Modèle Conceptuel de Données de la norme ISO 19115 au format Unified Modeling Language (UML) (source http://eden.ign.fr/std/iso_19115/uml19115/view)	64
Figure 8 : Modèle Conceptuel de (méta)Données proposé	70
Figure 9 : Modèle Conceptuel de (méta)Données avec prise en compte des normes en vigueur ISO 19115	74
Figure 10 : aperçu de l'interface graphique de MDweb (source http://demo.mdweb-project.org/mdweb/search/main.jsf consulté en juillet 2010).....	82
Figure 11 : aperçu de l'interface graphique de Geonetwork (source http://www.fao.org/geonetwork/srv/fr/main.home consulté en juillet 2010).....	83
Figure 12 : aperçu de l'interface graphique de Géosource (source http://geosourcedemo.brgm.fr/geosourcedemo/srv/fr/main.home consulté en juillet 2010) ...	85
Figure 13 : interface graphique de ArcCatalog copie d'écran.....	87
Figure 14 : interface graphique d'édition des métadonnées de ArcCatalog (copie d'écran)	87
Figure 15 : aperçu de l'interface graphique de métaZABR (source http://metazabr.lyon.cemagref.fr.fr/ consulté en juillet 2010)	89
Figure 16 : aperçu de l'interface graphique de saisie des métadonnées sous GeoNetwork, vue avancée (copie d'écran)	99

Liste des tableaux

Tableau 1 : arborescence du répertoire /sig256/dechomet	41
Tableau 2 : arborescence du répertoire du serveur Windows.....	42
Tableau 3 : les dix principaux éléments de métadonnées au sens de la norme ISO 19115.....	62
Tableau 4 : comparatif entre GeoNetwork et ArcCatalog.....	91

Table des annexes

Annexe 1 Formulaire de saisie des métadonnées au format Inspire via Excel.....	106
Annexe 2 Copies d'écran des interfaces graphiques d'édition des métadonnées de ArcCatalog ..	110
Annexe 3 Copies d'écran des interfaces graphiques des rubriques de saisie des métadonnées sous GeoNetwork, vue avancée	113

Liste des abréviations

ACV	Analyse du Cycle de Vie
ADELE	ADministration ELEctronique
ADEME	Agence de l'Environnement et de la Maîtrise de l'Energie
AERES	Agence d'Evaluation de la Recherche et l'Enseignement Supérieur
AE-RMC	Agence de l'Eau Rhône Méditerranée Corse
AFNOR	Association Française de NORmalisation
ANR	Agence Nationale de la Recherche
ArcSDE	Arc Spatial Database Engine
ARMINES	Association pour la Recherche et le développement des Méthodes et processus Industriels
BRGM	Bureau de Recherches Géologiques et Minières
CDD	Contrat à Durée Déterminée
CDI	Contrat à Durée Indéterminée
CEMAGREF	CENtre national du Machinisme Agricole, du Génie Rural, des Eaux et des Forêts
CIRIDD	Centre International de Ressources et d'Innovation pour le Développement Durable
CIS	Centre Ingénierie et Santé
CMPGP	Centre Micro-électronique de Provence Georges CHARPAK
CMS	Content Management System
CNAM	Conservatoire National des Arts et Métiers
CNES	Centre National d'Etudes Spatiales
CNIG	Conseil National de l'Information Géographique
CNRS	Centre National de la Recherche Scientifique
CR	Centre Régional

DGME	Direction Générale de la Modernisation de l'Etat
DGPS	Differential Global Positioning System
ENISE	Ecole Nationale d'Ingénieurs de Saint-Etienne
ENTPE	Ecole Nationale des Travaux Publics de l'Etat
ENSMSE	Ecole Nationale Supérieure des Mines de Saint-Etienne
EPA	Etablissement Public à caractère Administratif
ESRI	Environmental Systems Research Institute
FAO	Food and Agriculture Organization of the United Nations
FGDC	Federal Geographic Data Committee
FICCDC	Federal Interagency Coordinating Committee on Digital Cartography
G2I	Génie Industriel et Informatique
GAST	Geonetwork's Administrator Survival Tool
GEM	Groupe des Ecoles de Mines
Ghz	Giga oertz
GIS	Groupement d'Intérêt Scientifique
GPL	General Public License
Go	Giga octet
GPS	Global Positioning System
HDL	Hydroxyde Double Lamellaire
HDR	Habilitation à Diriger des Recherches
HP	Hewlett Packard
ICM	Ingénieurs Civil des Mines
IDEE	Information, Décision et Évaluation Environnementale
IGN	Institut Géographique National
IIT	Indian Institute of Technology

INSA	Institut Nationale des Sciences Appliquées
INSERM	Institut Nationale de la Santé Et de la Recherche Médicale
INSPIRE	Infrastructure for Spatial Information in the European Community
ISMIN	Ingénieurs Spécialisés en Microélectronique, Informatique et Nouvelles technologies
ISO	International Standardization Organization
ISTP	Institut Supérieur des Techniques de la Performance
ITII PACA	Institut des Techniques d'Ingénieurs de l'Industrie Provence Alpes Côte d'Azur
IRD	Institut de Recherche pour le Développement
LDAP	Lightweight Directory Access Protocol
LIRMM	Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier
MCD	Modèle Conceptuel de Données
MCF	Maître de ConFérence
MEF	Metadata Exchange Format
Mhz	Mega hertz
MIRI	Maîtrise des Impacts environnementaux et des Risques Industriels
MNT	Modèles Numérique de Terrain
MTD	Meilleures Technologies Disponibles
Mo	Méga octet
PACA	Provence Alpes Côte d'Azur
PNUE	Programme des Nations Unies pour l'Environnement
SEIU	Sciences de l'Environnement Industriel et Urbain
SEPIT	Sciences de l'Environnement pour les Processus Industriels et Territoriaux

SGBG	Système de Gestion de Base de Données
SI	Service Informatique
SIG	Système d'Information Géographique
SITE	Sciences, Information et Technologies pour l'Environnement
SMS	Sciences des Matériaux et des Structures
SPIN	Sciences des Processus Industriels et Naturels
SPOT	Satellite Pour l'Observation de la Terre
STDDAD	Systèmes Territoriaux, Développement Durable et Aide à la Décision
To	Tera octet
UML	Unified Modeling Language
UMR	Unité Mixte de Recherche
UNEP	United Nation Environmental Programme
UJM	Université Jean Monnet
UR	Unité de Recherche
W3C	World Wide Web Consortium
XML	Extensible Markup Language
ZABR	Zone Atelier Bassin du Rhône

Glossaire

Ce glossaire a été élaboré à partir des lexiques des sociétés ESRI, GeoNetwork Opensource et des sites web consultés.

Donnée : ensemble quelconque de faits ayant un rapport entre eux, organisés selon un format particulier. Souvent, il s'agit des éléments d'information de base produits, stockés ou traités par un ordinateur.

Données géographiques : informations relatives aux entités réelles, notamment leur forme, leur emplacement et leur description. Les données géographiques sont composées de données spatiales et de données attributaires.

eXtensible Markup Language (XML) : créé par le World Wide Web Consortium (W3C), le langage XML est destiné à la création de formats de texte conçus pour faciliter l'échange de données entre applications informatiques (par exemple, sur le Web). Le langage XML est un ensemble de règles permettant de créer des formats d'information normalisés à l'aide de balises personnalisées, et de partager des formats et des données entre différentes applications.

Géocatalogue : organisation d'un ensemble de métadonnées sous la forme d'un recueil (base de données ou autre) facilitant leur consultation et leur diffusion.

Géodatabase : ensemble de jeux de données géographiques de différents types stockés dans un dossier de système de fichiers communs, un fichier de base de données ou une base de données relationnelles multi-utilisateurs.

Géolocalisé : technologie permettant de positionner un objet sur une carte à l'aide de ses coordonnées géographiques.

Géomatique : application de l'informatique à la géographie, permettant la gestion des données à référence spatiale en utilisant les sciences et les technologies d'acquisition, de stockage, de traitement et de diffusion.

Géoréférencé : technologie permettant d'établir une relation entre des coordonnées spatiales sur une carte et les coordonnées réelles.

Géorépertoire : catalogue et moteur de recherche qui permet d'inventorier et de restituer des métadonnées.

Global Positioning System (GPS) : système de satellites et d'équipements de réception permettant de calculer la position de trois dimensions (x, y et z) sur le globe. Le système GPS est utilisé dans la navigation, la cartographie pour les SIG et l'arpentage.

Métadonnées : informations décrivant le contenu, la qualité, la condition, l'origine et d'autres caractéristiques d'un jeu de données ou d'informations. Les métadonnées sont constituées de propriétés et de documentation. Les propriétés proviennent de la source de données (par exemple, le système de coordonnées et la projection des données), tandis que la documentation est saisie manuellement (mots-clés utilisés pour décrire la donnée).

Open Geospatial Consortium (OGC) : organisation de l'industrie travaillant sur l'interopérabilité des communautés géospatiales et SIG, et assurant sa prise en charge.

Raster : est un format de données spatiales basé sur le pixel.

Système d'Information Géographique (SIG) : système informatique de matériels, logiciels, données et workflows permettant de collecter, stocker, analyser et diffuser des informations. Un SIG est un système avec lequel les utilisateurs interagissent pour intégrer, analyser et visualiser les données géographiques, identifier les relations, les schémas et les tendances et apporter des solutions aux problèmes. Chaque SIG représente généralement des informations sur les cartes sous la forme de couches de données utilisées pour l'analyse et la visualisation.

Vecteur : est un format de données spatiales basé sur le mode vectoriel.

Introduction et objectifs du travail

Les travaux relatifs à ce mémoire ont été effectués à l'Ecole Nationale Supérieure des Mines de Saint-Etienne au sein du centre d'enseignement et de recherche SITE sous la direction de Didier GRAILLOT.

Depuis le début des années 1990 le centre d'enseignement et de recherche SITE de l'Ecole Nationale Supérieure des Mines de Saint-Etienne effectuant une recherche concernant les enjeux territoriaux dispose d'un Système d'Information Géographique. Durant ces deux dernières décennies, les enseignants-chercheurs du centre ont produit et acquis de nombreuses ressources géographiques, données environnementales à fort caractère spatial. Elles représentent aujourd'hui une part importante de son patrimoine scientifique. Leur gestion dans un système organisé d'information est devenue une nécessité afin de pérenniser ces ressources.

Avant de pouvoir organiser les données et de maintenir la mémoire de ce patrimoine de données, il est primordial de bien savoir les décrire. C'est pourquoi, il est nécessaire de constituer un catalogue contenant des informations sur les ressources géographiques disponibles dans le centre, c'est-à-dire un catalogue informatisé de métadonnées.

Ce besoin s'inscrit dans une préoccupation rencontrée par de nombreux acteurs du domaine de l'information géographique. C'est pourquoi dès 2003, une norme internationale a été rédigée, elle organise par classes toutes les informations pertinentes susceptibles d'être disponibles pour décrire la donnée.

Après avoir situé le contexte du projet et analysé les besoins en matière de système d'informations environnementales, un Modèle Conceptuel de (méta)Données approprié et cohérent avec les normes en vigueur est étudié. L'analyse des conditions de sa mise en œuvre et son implantation clôtureront ce mémoire.

L'objectif principal du projet est de faciliter le travail des enseignants-chercheurs qui doivent acquérir, créer et traiter des données spatiales et temporelles dans le but d'élaborer une réponse à une problématique environnementale dans le cadre de leurs activités de recherche, mais aussi d'enseignement. La réalisation de cet objectif passe par une gestion plus adéquate et mieux organisée du patrimoine des ressources géographiques du centre. Cette gestion n'est possible que grâce à l'utilisation d'un géocatalogue.

Partie 1

Contexte du projet

Je travaille depuis 1991 à l'Ecole Nationale Supérieure des Mines de Saint-Etienne (ENSMSE). Depuis 2003, je suis en poste au sein du laboratoire Sciences, Information et Technologies pour l'Environnement (SITE). Le travail relatif à ce mémoire s'est déroulé dans le département Information, Décision et Évaluation Environnementale (IDEE).

Dans le premier chapitre de cette partie l'organisation professionnelle dans laquelle j'évolue sera décrite, puis dans le deuxième chapitre l'environnement informatique du centre SITE et plus précisément celui du département IDEE. Enfin le troisième chapitre présentera comment sont gérées les ressources en géomatique d'IDEE.

Chapitre 1 – SITE : un centre de formation et de recherche en ingénierie de l'environnement

1. L'Ecole Nationale Supérieure des Mines de Saint-Etienne

L'Ecole Nationale Supérieure des Mines de Saint-Etienne (ENSMSE) aura bientôt deux cents ans d'existence. Créée en 1816, c'est une institution académique dans la lignée des "Grandes Ecoles françaises". Dédiée à la formation d'ingénieurs et à la recherche scientifique, l'Ecole a pour mission d'accompagner le développement de ses élèves et de ses étudiants (Ingénieurs Civils des Mines, doctorants) et celui des entreprises.

L'ENSMSE est un Etablissement Public à caractère Administratif (EPA) depuis 1992, sous tutelle du ministère de l'Economie, de l'Industrie et de l'Emploi. Elle est l'une des trois écoles nationales supérieures du Groupe des sept Ecoles des Mines (GEM), dédiées à la formation d'ingénieurs, de chercheurs, ainsi qu'à la recherche finalisée et appliquée, c'est-à-dire orientée vers les besoins de l'industrie.

Forte de sa formation historique, Ingénieurs Civil des Mines (ICM), l'Ecole a enrichi son offre et se positionne sur toutes les filières de formation : elle accueille des Bac + 2 en formation d'ingénieur généraliste, ou ingénieur spécialisé, forme des doctorants, et apporte des solutions en terme de formation par apprentissage, ou alternance, tout au long de la vie professionnelle.

Son savoir-faire s'appuie sur un juste équilibre entre des formations académiques soutenues par une pédagogie innovante et des périodes de formation en prise directe avec les réalités de l'entreprise en France et à l'étranger.

L'Ecole en quelques chiffres (source rapport d'activité 2009)

Le budget est de **22,3 M€** hors masse salariale, dont 70% provenant des ressources propres de l'établissement et des ressources de l'activité.

L'Ecole comporte :

- **410 salariés** (CDI, CDD, tous statuts confondus, dont 20 personnes ARMINES, 41 mises à disposition) : scientifiques 62 %, techniques 18 % et administratifs 20%.
- **250 vacataires** issus du milieu économique

Pas moins de **1570 étudiants** sont inscrits à l'Ecole en formation tous cursus confondus (élèves-ingénieurs et 3^{èmes} cycles) et répartis sur les sites de Saint-Etienne (1161), de Gardanne (391) et de Florianópolis au Brésil (19) :

- 447 élèves ICM et 14 auditeurs libres
- 210 élèves-ingénieurs ISMIN (Spécialité Microélectronique, Informatique et Nouvelles Technologies)
- 42 étudiants en Mastères spécialisés
- 74 étudiants en Master Recherche
- 173 étudiants en formation doctorale (Ecole doctorale de Saint-Etienne : ENSMSE / Université Jean Monnet / ENISE)
- 625 en formation sous statut salarié (formation continue et par apprentissage) : en convention avec l'ISTP (469 à Saint-Etienne) et ITII PACA (156 à Gardanne)

Le montant des contrats de recherche (y compris subventions liées aux équipements scientifiques) est de **8,2 M€** pour l'année 2009.

La surface totale des deux sites de Saint-Etienne représente **33 000 m²**.

La surface totale du site de Gardanne est de **25 000 m²**.

Celle du futur site du Centre Ingénierie et Santé (CIS) est de **6 000 m²**.

La Société des Anciens Elèves compte plus de **3600 ingénieurs** en activité.

Pour mieux cerner l'organisation de l'ENSMSE et ses différentes entités, il est utile de se référer à l'organigramme de l'établissement.

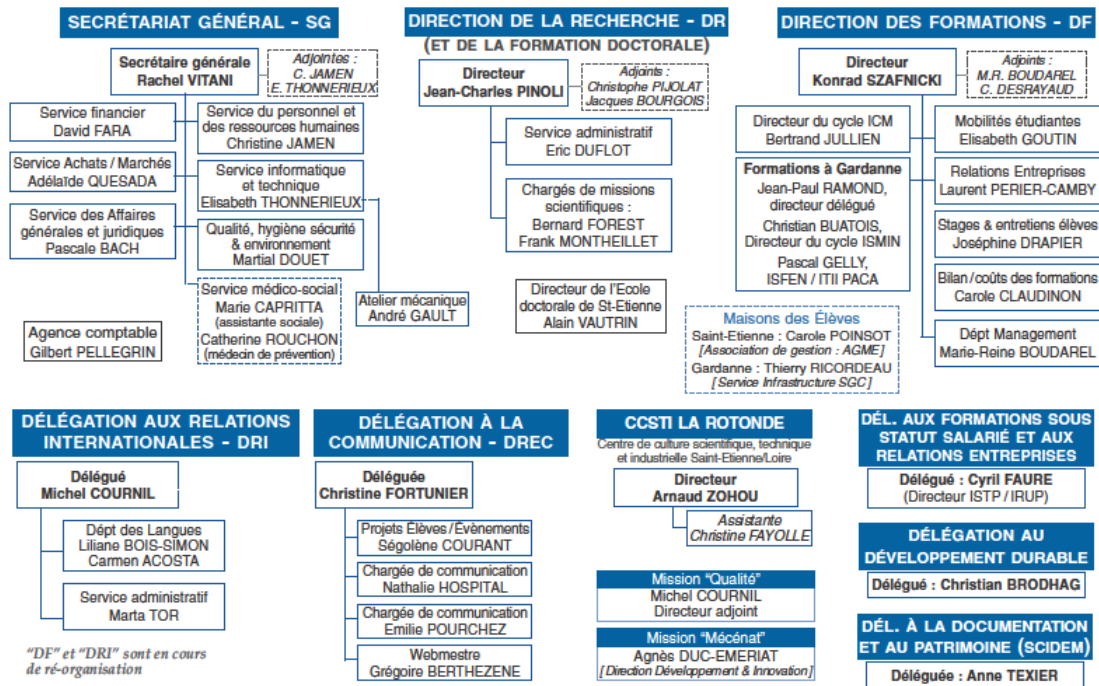


**Président du
Conseil d'administration**
François GERIN

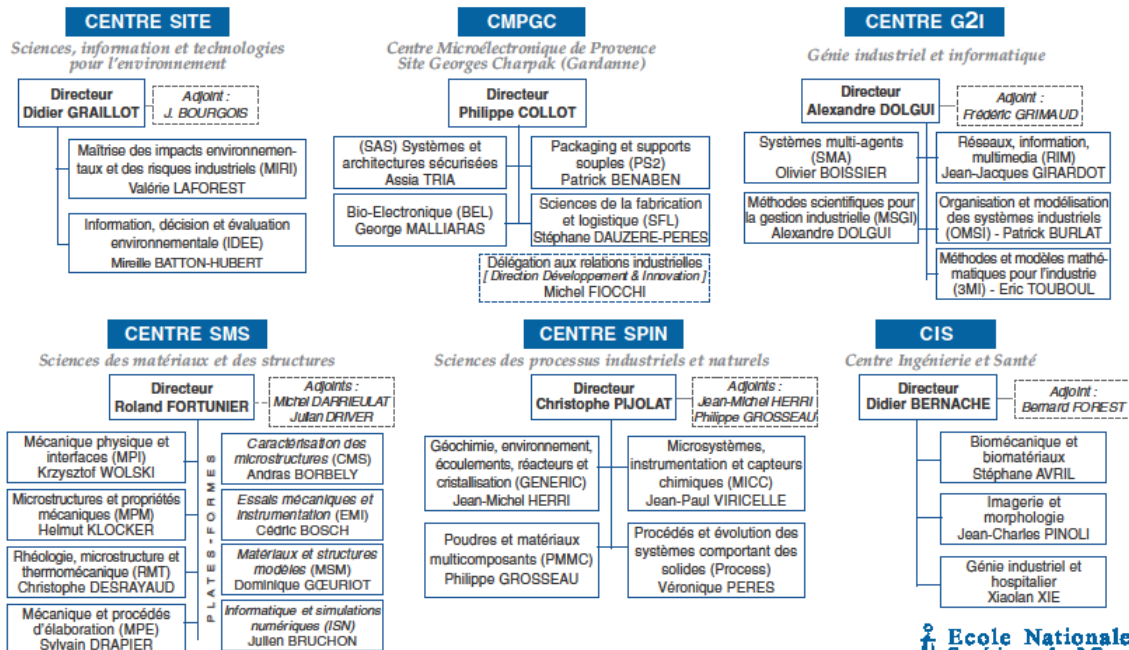
Directeur
Philippe JAMET

- Michel COUNIL, Directeur adjoint chargé des relations internationales et des réformes institutionnelles
- Jean-Charles PINOLI, Directeur adjoint chargé de la recherche
- Daniel OCHOA, Directeur adjoint chargé du développement et de l'innovation
- Rachel VITANI, Secrétaire générale de l'école
- Monique REYRE, Directrice déléguée du Site Georges Charpak [Service Infrastructure]

DIRECTIONS FONCTIONNELLES



CENTRES D'ENSEIGNEMENT ET DE RECHERCHE



PRÉSENTATION DE L'ECOLE - Avril 2010

**Ecole Nationale
Supérieure des Mines
SAINT-ETIENNE**

Figure 1 : organigramme ENSMSE (source rapport d'activité 2009 ENSMSE)

Après avoir présenté l'ENSMSE, nous allons décrire ses centres de recherches et plus précisément le centre SITE, son département IDEE pour lequel ce travail a été réalisé.

2. Les centres de recherche de l'Ecole

L'Ecole organise ses activités académiques en six centres de formation et de recherche impliqués sur des thématiques diversifiées (procédés, matériaux, environnement, informatique, santé, génie industriel). Ils sont membres actifs de plusieurs pôles de compétitivité et mènent de nombreux projets de recherche contractuelle dans le cadre de l'Agence Nationale pour la Recherche (ANR) ou directement auprès d'entreprises françaises et mondiales. Plusieurs d'entre eux sont labellisés par le CNRS ou l'INSERM. L'Ecole est membre de l'Institut Carnot.

Le fonctionnement de ces centres est assuré par 121 enseignants-chercheurs qui, avec l'aide de 152 doctorants, partagent leurs connaissances avec les élèves ingénieurs, ceci en lien avec de nombreuses collaborations industrielles.

Les six centres de formation et de recherche sont les suivants :

- Centre Sciences des Matériaux et des Structures (SMS) dont les thèmes de recherche sont :
 - le design et l'allègement des structures pour le transport
 - l'optimisation des matériaux et structures pour les énergies non-productrice de CO₂
 - les analyses multi-physiques locales pour les micro-nanotechnologies
- Centre Sciences des Processus Industriels et Naturels (SPIN) dont les thèmes de recherche sont :
 - l'interaction solides-gaz
 - les matériaux multi-composants
 - la cristallisation et transferts
- Centre Sciences, Informations et Technologies pour l'Environnement (SITE) dont les thèmes de recherche sont :

- l'aide à la décision dans les domaines des ressources en eaux, des déchets, des technologies propres, de la supervision de procédés industriels, de la qualité de l'air.
- Centre Génie Industriel et Informatique (G2I) dont les thèmes de recherche sont :
 - les outils informatiques supports aux activités d'ingénierie en conception et production
 - les systèmes distribués
 - les E-technologies
 - la Supply Chain Management
- Centre Microélectronique de Provence Georges Charpak (CMPGC) dont les thèmes de recherche sont :
 - la science de la fabrication et la logistique
 - les systèmes et l'architecture sécurisée
 - le micro-packaging et les supports souples
- Centre Ingénierie et Santé (CIS) dont les thèmes de recherche sont :
 - les biomatériaux et la mécanique
 - l'imagerie et les statistiques
 - le Génie Industriel Hospitalier
 - la dynamique des systèmes biologiques
 - le Bio-environnement Industriel

Le sujet de ce mémoire s'inscrit dans une problématique rencontrée au sein du centre SITE dont l'organisation va être développée ci-après.

3. Le centre SITE

Lancée en 1999, le centre SITE dispense des formations pour préparer aux métiers de l'environnement industriel et territorial.

L'option environnement industriel et territoire a pour objectif de préparer les élèves ingénieurs aux métiers de l'environnement industriel et territorial dans un contexte de développement durable, tout en conservant l'approche généraliste que doit posséder l'ingénieur.

Le centre SITE effectue des recherches autour de l'aide à la décision dans les domaines environnementaux, les technologies propres et la maîtrise des risques industriels et naturels. Ces actions s'inscrivent dans les grandes préoccupations liées au Développement Durable.

Le centre SITE est composé de deux laboratoires de recherche : MIRI (Maîtrise des Impacts environnementaux et des Risques Industriels) et IDEE (Information, Décision et Évaluation Environnementale).

Il comporte une unité de recherche Science de l'Environnement pour les Processus Industriels et Territoriaux (SEPIT).

Le domaine de recherche et les activités de SEPIT s'articulent autour de deux thématiques :

- en processus industriel : maîtrise des impacts environnementaux et des risques
- processus territoriaux : aide à la décision pour la maîtrise des ressources territoriales

SITE en quelques chiffres (source rapport d'activité 2009)

- effectif total : 43 y compris les C.D.D.
- enseignants-chercheurs : 11
- doctorants : 17
- nombre de thèses soutenues : 22 depuis 2005
- thèses en cours : 16
- budget : 560 k€ (hors masse salariale état)
- quatre axes d'enseignement en deuxième année d'ICM : Cindynique, Conduite Avancée des Processus Industriels, Modélisation dans et pour les Sciences de l'Ingénieur, Maîtrise de l'Environnement et de l'Energie

- deux Master recherche : Sciences de l'Environnement Industriel et Urbain (SEIU) et Systèmes Territoriaux, Développement Durable et Aide à la Décision (STDDAD).

Directeur adjoint : Valérie LAFOREST

Secrétariat : Zahia MAZER, Christiane RETRUS

Thématique 1 "Maîtrise des impacts environnementaux et des risques"		Thématique 2 "Maîtrise des ressources et aide à la décision en aménagement du territoire"	
Valérie LAFOREST (MCF-HDR)		Didier GRAILLOT (Prof1. HDR)	
BOURGOIS Jacques	Prof1. HDR	BREUIL Florent	IR
GONDRAN Natacha	MCF	BRODHAG Christian	Prof1. HDR
LONDICHE Henry	Prof2. HDR	DECHOMETS Roland	Prof1. HDR
PIATYSZEK Eric	MCF	MIMOUN Djamel	MCF
SZAFNICKI Konrad	Prof2. HDR	VAILLANT Hervé	IR
TARDY Alicja	IR		
CIKANKOWITZ Anne	IR	PARAN Frédéric	IR
NARCE DELOLME Cécile	IR		
Doctorants		Doctorants	
EL HAITE Hakima KARAGIANNIS Georges LEVEILLARD Florine MARREF Souâd MOREAU Valentine NGUYEN Viet Hai PERRIN Loïc PICHERIT Marie-Lou RAYMOND Gaëlle ROJO Gabriel VILLOT Jonathan (Rep. doctorants)		BOUBAKER Leila GAUR Shishir GUILLOUX Gaël MARSAUCHE Maud RE-BAHUAUD Jordan (Rep. Doctorants) TACNET Jean-Marc TOMBOZAFY Mamy	
Enseignant-chercheur invité		Enseignant-chercheur invité	
LADHARI Taoufik (2007-Tunisie) TAYIBI Hanan (2009-Espagne) KALKA Joanna (2007-Pologne)		ANIM ADDO Samuel (2006-Allemagne) CHAHAR Bhagu Ram (2009-Inde) WAAUB Jean-Philippe (2006-Québec CN) WOTTO Marguerite (fin 2006-2007-Québec CN)	
Responsabilités Formation		Responsabilités Valorisation-Publication	
1-Discipline doctorale "Sciences et Génie de l'Environnement" Jacques BOURGOIS 2-Master Recherche "SEIU" Valérie LAFOREST "STDDAD" Didier GRAILLOT 3-Formation "ICM" Eric PIATYSZEK		1-Commission publications, site web de l'UR Florent BREUIL 2-Valorisation des œuvres numériques Florent BREUIL, Konrad SZAFNICKI	
	Budget de l'UR		
	Hervé VAILLANT		
	Relations internationales		
	Alicja TARDY		
	Assistance technique informatique		
	Florence DUJARDIN – Roland DECHOMETS (ressources SIG)		

Figure 2 : organigramme de l'UR SEPIT (source rapport d'activité 2009 Centre SITE)

Après avoir présenté le centre SITE, une description des deux départements qui le composent sera faite en s'attardant plus sur IDEE, car le projet concerne en priorité cette unité.

4. Le département MIRI

Ce département développe des modèles permettant de mieux maîtriser les impacts environnementaux et des risques industriels. Il est composé de 16 personnes dont 8 doctorants.

Sa recherche s'organise autour des thèmes suivants :

- l'optimisation du fonctionnement des processus éco-industriels
- l'évaluation des performances des Meilleures Technologies Disponibles (MTD) en matière de gestion environnementale
- l'analyse des processus de danger/risque
- le contrôle, la supervision et l'aide à la décision en temps réel appliqués aux processus éco-industriels et vivants

Les principales problématiques abordées sont :

- le recyclage et la valorisation des déchets industriels
- la pollution/dépollution des sols
- l'analyse des risques industriels
- le management environnemental et le développement durable

Pour cela, MIRI développe principalement ses compétences dans les domaines suivants :

- les technologies et les productions plus propres et plus sûres
- le piégeage des métaux lourds dans les rejets industriels par HDL et cyclo-dextrines
- le contrôle/supervision et l'aide à la conduite des éco-procédés simulateurs pédagogiques de projets environnementaux

Exemples de réalisation :

- l'analyse multicritère des MTD

- les logiciels de supervision pour le traitement des effluents industriels
- l'analyse du Cycle de Vie (ACV) des voitures hybrides
- l'évaluation des MTD en industrie mécanique

5. Le département IDEE

Il développe des modèles pour l'aide à la décision et à la négociation, l'évaluation des impacts environnementaux (pollution eau/air/sol), évalue et analyse les phénomènes environnementaux, le transfert de connaissances et de compétences. Il est composé de 16 personnes dont 7 doctorants.

Sa recherche s'organise autour des thèmes suivants :

- l'optimisation, la modélisation, la simulation et le monitoring des impacts spatiaux des phénomènes continus environnementaux (air, eau, odeurs)
- les méthodologies d'évaluation en matière de développement durable
- l'élaboration de méthodes d'aide à la décision pour la prise en compte de l'environnement par les acteurs locaux

Les principales problématiques abordées sont :

- la gestion des ressources en eau
- l'analyse de dysfonctionnements en réseaux
- la pollution/dépollution des sols
- le management environnemental
- l'aide à la négociation territoriale

Pour cela, IDEE développe principalement ses compétences dans les domaines suivants :

- le développement durable
- l'hydrologie et l'hydrogéologie (interactions nappes-rivières)
- l'analyse des risques naturels (inondations)
- les systèmes d'information environnementaux (géographiques, territoriaux, d'entreprises)

Exemples de réalisation :

- l'évaluation de l'impact d'un centre de stockage des déchets sur la qualité de l'air
- les méthodes de discrétisation pour la modélisation par éléments analytiques en hydrogéologie
- la quantification des échanges nappes-rivières
- l'empreinte écologique

Ces thèmes font appel et mobilisent beaucoup d'informations spatialisées à l'échelle du processus industriel (MIRI) et à l'échelle des processus territoriaux (IDEE). Ces deux types de processus correspondent aux thématiques des deux départements et des deux thèmes de l'UR SEPIT du centre SITE.

Le projet réalisé s'ancre dans la nécessité de l'utilisation de ressources géomatiques du département IDEE.

Quelques partenaires d'IDEE

Des partenariats sont développés avec des institutions chargées de la maîtrise d'ouvrage de programmes de recherche, des organisations ou des collectivités territoriales.

IDEE est membre du Groupement d'Intérêt Scientifique (GIS) Zone Atelier Bassin du Rhône (ZABR), qui regroupe le CNRS, les Universités de Lyon 1 et de Saint-Etienne, l'INSA Lyon, La Maison du Rhône, l'Ecole vétérinaire de Lyon et l'ENTPE. La problématique de la ressource en eau tant sur les plans physiques, biologiques, sociaux et économiques constitue l'axe principal de recherche du groupement. Une des actions de cet accord scientifique consiste, par exemple, à mieux comprendre et décrire les phénomènes liés aux pollutions par l'intermédiaire des échanges entre les eaux souterraines et les eaux superficielles en particulier celles du bassin versant du Rhône. L'ensemble des actions qui impliquent les partenaires de cet accord s'inscrit dans un objectif global d'amélioration de la qualité des eaux du Rhône et de leur bon état écologique à l'horizon 2015.

IDEE participe au programme européen du Plan Loire Grandeur Nature. Dans ce cadre, ce programme favorise la contribution interdisciplinaire à la gestion équilibrée de la ressource en eau de l'hydro-système ligérien amont dans un contexte de changement climatique en collaboration avec l'UJM Saint-Etienne, le CEMAGREF Lyon, et d'autres organismes.

Dans le cadre du GEM, IDEE travaille sur un programme avec l'IIT de Delhi sur l'optimisation de la gestion en ressource en eau.

Sur le thème de la ressource en eau, il s'agit de traiter, de modéliser une masse de données spatiales importante (réseaux hydrographiques, localisation de nappes phréatiques, données physiques : débits, hauteurs d'eau, données de qualité).

Quelques autres principaux partenaires : ADEME, EDF, Agence de l'Eau Rhône Méditerranée Corse (AE-RMC), VEOLIA.

Après avoir présenté l'organisation et la structure dans lesquelles s'inscrit l'étude de ce mémoire, il s'agit d'aborder maintenant les ressources concernées par l'étude.

6. IDEE et les ressources en géomatique

Les travaux de recherche du centre SITE s'inscrivent dans une stratégie globale de gestion des ressources territoriales. Pour comprendre comment évoluent ces ressources, ces travaux portent sur deux thèmes principaux : l'analyse des mécanismes physiques mis en jeu (ressources en eau, air) et l'évaluation environnementale des scénarii d'aménagement dont ces ressources font l'objet.

L'analyse des mécanismes physiques génère des modèles phénoménologiques particuliers (écoulements d'eau, qualité de l'air) qui servent à simuler des scénarii pour une aide à la décision en aménagement territorial. La recherche des modalités d'utilisation de ces modèles dans des contextes de négociation territoriale conduit à étudier les différentes représentations possibles des résultats qu'ils produisent en assurant leur cohérence.

Les travaux des enseignants-chercheurs de SITE s'appuient sur une forte expérience en matière de modélisation environnementale dans le domaine de l'eau et de la qualité de l'air, en matière de systèmes d'information pour l'aide à la décision et à la négociation territoriale et sur une capitalisation des données collectées au cours des programmes de recherche développés jusqu'à présent.

Depuis 1992, la recherche en ingénierie de l'environnement est présente au sein de l'ENSME. Durant deux décennies de nombreuses données ont été acquises ou produites. Ces données correspondent à la matière première des chercheurs, qui en les rassemblant et les combinant à partir de diverses sources des informations territoriales analysent et

raisonnent sur une problématique, afin de comprendre les phénomènes et les événements ayant un impact sur le territoire.

7. L'acquisition de données

Le département IDEE utilise de nombreuses données spatiales : physiques, hydrogéologiques, géophysiques, mais aussi biologiques, physico-chimiques et sociales.

Elles sont souvent achetées à un producteur de données [ECOBICHON] ou mises à disposition par :

- l'Institut Géographique National (IGN), qui est un établissement public chargé de la géodésie, du nivellement et de la cartographie du territoire,
- le Bureau de Recherches Géologiques et Minières (BRGM), chargé de l'élaboration de la cartographie géologique ; il met à disposition des bases de données sur la géologie mais aussi sur l'hydrologie,
- le Centre National d'Etudes Spatiales (CNES), qui est un établissement public, qui conduit le programme des satellites d'observation de la terre et est le distributeur exclusif des images des satellites SPOT.

Les données géographiques peuvent être relatives aux contrats de recherche. Dans ce cas, les données proviennent soit des partenaires, soit d'autres acteurs : les concessionnaires de réseaux de distribution d'eau, les collectivités locales, les industriels, les cabinets spécifiques et autres organismes.

Les données géographiques peuvent être aussi produites directement par le département grâce à des relevés de terrain. Ceux-ci sont effectués avec un Global Positioning System (GPS) et un Differential GPS.

Un certain nombre de matériels d'analyse spatiale et environnementale est à la disposition des chercheurs :

- une station hydrométéorologique
- un stéréoscope : SOKKIA
- une table à digitalisation
- un Global Positioning System (GPS) : Trimble Pro XRS

- un Differential GPS : Trimble Pro XRS module de commande Recon usage DGPS de précision couplé au satellite différentiel Omnistart offrant une précision 0,50 m au lieu des 3 à 5 mètres
- un GPS léger : Dell Axim GPS
- un télémètre : IMPULSE IR 600
- un encodeur d'angle : Mapstar
- une sonde hydro : TRITECH PA200
- une sonde piézométrique
- un micro-moulinet pour les mesures hydrodynamiques

Les données géographiques d'IDEE sont sous divers formats numériques (raster, vecteurs, etc) ou non (papier). Elles sont composées de bases de données ortho-photographiques de l'IGN, de données satellites du CNES, de photographies, d'images scannées, etc. Elles peuvent avoir subi un traitement, leur utilisation nécessite alors un indice de fiabilité et une traçabilité, car leur introduction dans des modèles de simulation a une influence sur la qualité des résultats obtenus. Généralement, ces données sont géoréférencées et utilisées sous un Système d'Information Géographique (SIG).

Après avoir décrit l'organisation professionnelle du projet, l'environnement informatique sera décrit dans le chapitre suivant.

Chapitre 2 – Environnement organisationnel et matériel

L'organisation informatique de l'ENSMSE s'articule autour du Service Informatique (SI), avec une décentralisation de gestion dans les différents centres d'enseignement et de recherche.

Dans ce chapitre, seront d'abord décrits l'environnement organisationnel informatique du centre SITE, puis les matériels et logiciels utilisés.

1. Organisation du centre SITE

Un ingénieur de recherche au centre SITE

Le centre SITE emploie un ingénieur de recherche dont la mission principale est la réalisation de systèmes d'information pour le développement durable dans l'espace francophone.

Il met au point, pour Internet, des documents pédagogiques de sensibilisation dans le domaine de l'environnement et du traitement des déchets, de développement durable (Agenda 21 locaux).

Il conçoit des outils de valorisation de la recherche, met au point et maintien des systèmes d'information spécialisés, serveurs web, CMS (Content Management System), listes de diffusion, bases de données SQL, Access...

Il participe actuellement au développement du géorépertoire de métadonnées de la Zone Atelier Bassin du Rhône (ZABR).

Il collabore avec le Centre International de Ressources et d'Innovation pour le Développement Durable (CIRIDD), avec l'Organisation Internationale de la Francophonie et avec l'Institut de l'Energie et de l'Environnement de la Francophonie.

Il est membre du Comité Directeur et du Comité d'Orientation International du Système d'information mondial francophone pour le développement durable.

2. Environnement informatique du centre SITE

Un technicien informatique au centre SITE

Depuis 2003, le centre SITE dispose d'un technicien en informatique. Il a en charge la gestion du parc informatique : des serveurs aux postes clients, des différents systèmes et de l'ensemble des logiciels privés et libres, ainsi que leur mise en réseau.

Sa mission s'étend de l'analyse du besoin, au choix et à la mise en place de la solution, ainsi qu'à la gestion de son cycle de vie.

Il met en œuvre, de manière autonome, la politique informatique de l'Ecole au sein du centre SITE.

Ses fonctions sont :

- la définition annuelle des besoins informatiques et son suivi budgétaire
- la planification des évolutions
- l'installation, l'administration, la gestion, la maintenance et la mise à jour du parc informatique (environ 80 postes) et de l'ensemble des logiciels associés (recherche et enseignement)
- l'aide et l'assistance aux utilisateurs non-spécialistes dans l'utilisation et le fonctionnement des matériels informatiques et des logiciels de bureautique
- la gestion du parc d'enseignement : 15 postes
- la définition de la politique de sauvegarde ainsi que son installation, sa gestion centralisée et le pilotage de l'automate
- la gestion et suivi des contrats de maintenance matériel et logiciel, ainsi que les relations fournisseurs

Ses compétences s'étendent de la gestion des serveurs sous-systèmes d'exploitation Server 2003 et/ou Unix Solaris 10, à l'administration des postes clients Windows XP en environnement réseau. Tout le parc est administré à l'aide d'un gestionnaire d'administration informatique centralisé : Active Directory.

Le technicien est membre de la Commission Informatique de l'Ecole et est le correspondant informatique du service informatique.

En support : le service informatique

Le technicien informatique du centre SITE travaille en étroite collaboration avec le Service Informatique (SI) dont le rôle est le suivant :

- la mise en place du réseau informatique de l'Ecole

- le développement et le support des applications communes à l'ensemble de l'Ecole : annuaire LDAP (Lightweight Directory Access Protocol), messagerie, téléphonie, visio-conférence, etc.
- le développement et support de l'informatique des services administratifs
- la gestion des moyens audiovisuels
- l'organisation de la maintenance et la gestion des interventions
- le suivi du parc informatique de l'Ecole
- la gestion du marché informatique et le suivi des commandes
- la mise en place de la politique de sécurité informatique, ainsi que son administration et sa gestion

En liaison avec la commission informatique de l'Ecole, le SI définit le schéma informatique directeur, oriente les choix informatiques vers des produits standards et assure un support aux correspondants informatiques.

Le SI est composé de 16 personnes, dont deux affectées au centre G2I et quatre affectées au CMPGC. Ces deux centres ne disposent pas de technicien informatique.

Les compétences des membres du SI s'articulent autour de plusieurs axes :

- l'expertise de l'administration et de la sécurisation des systèmes d'exploitation des serveurs UNIX et Windows server, ainsi que la gestion de leurs sauvegardes
- l'administration du pare-feu
- l'administration réseau et la supervision des matériels actifs
- le suivi des applications spécifiques : ressources humaines, payes, comptabilité, etc.
- le développement d'applications internes : gestion des stocks, demandes d'intervention, ...
- l'administration, l'installation des postes clients et des périphériques et imprimantes, ainsi que des applications bureautiques et spécifiques du Secrétariat Général de l'Ecole
- assistance aux utilisateurs du Secrétariat Général de l'Ecole

- suivi du parc informatique et gestion des contrats de maintenance collectifs

3. Matériel informatique et logiciels utilisés par le centre SITE

Matériel informatique

Dans cette partie, seuls les matériels informatiques utilisés par le laboratoire IDEE seront présentés. Les matériels permettant l'acquisition de données ne sont pas traités puisqu'ils relèvent davantage de l'instrumentation sur les sites d'études.

Le département IDEE utilise les deux serveurs Windows Server 2003 du centre SITE comme tous les personnels du centre pour sauvegarder et partager leurs données. Une sauvegarde automatisée des deux serveurs est effectuée sur bandes.

Les serveurs sont des HP serveur Proliant, un ML350G4P cadencé à 3 Ghz avec 1 Go de mémoire vive et six disques durs de 300 Go pour l'un, et un ML350G5 bi-cœur cadencé à 2 Ghz avec 4 Go de mémoire vive et de six disques durs de 300 Go en RAID 5 pour l'autre. Ils ont été achetés respectivement en 2006 et 2007.

Sur le ML350G5, un automate de sauvegarde HP storageworks 1/8 ULT 448 autoloader d'une capacité de stockage de 1,6 To en natif et 3,2 To compressés, ainsi que son gestionnaire HP OpenView Storage Data Protector ont été installés.

Le département IDEE dispose de deux serveurs SUN dédiés au SIG : un Sun Enterprise 250 cadencé à 400 Mhz acheté en 2000 avec Solaris 2.7 en système d'exploitation et l'autre un Server SUN Fire V240 cadencé à 1,5 Ghz acheté en 2006.

La sauvegarde de ces serveurs s'effectue sur un troisième serveur du bâtiment un SUN Ultra60 acheté en 1998 cadencé à 300 Mhz sous Solaris 2.6 sur lequel est installé un automate piloté. Il s'agit d'un chargeur automatique SPARstorage DLT4700 avec un magasin de 7 cartouches géré par le logiciel Solstice Backup de Legato.

Chaque membre du département dispose d'au moins un poste informatique fixe ou portable selon ses besoins. Parfois la mobilité de l'utilisateur ainsi que les puissances de calcul et d'affichage des outils utilisés nécessite l'attribution de deux ordinateurs : un fixe et l'autre mobile.

Une dizaine d'ordinateurs fixes est dédiée aux personnels du département. Il s'agit d'ordinateur de type station de travail dont XP est le système d'exploitation :

- HP workstation Z400 doté d'un processeur Intel Xeon W3520 cadencé à 2,66 Ghz avec 4 Go de mémoire vive acheté en 2010,
- Dell Precision Workstation N-Series T5400 doté d'un processeur Intel Xeon E5400 cadencé à 2,5 Ghz avec 4 Go de mémoire vive acheté en 2007 ou 2008,
- HP workstation XW4400 doté d'un processeur Intel core 2 duo E6700 cadencé à 2,67 Ghz avec 2 Go de mémoire vive acheté en 2007.

De plus, deux autres ordinateurs de ce type sont des serveurs calcul et de logiciels de recherche.

Les stagiaires accueillis dans le département utilisent le même type de matériel mais plus ancien. Cinq ordinateurs leur sont réservés.

Une petite dizaine d'ordinateurs portables est dédiée aux personnels du département. Tous sont installés sous XP.

Il s'agit d'ordinateur de type :

- Dell Precision M4300, doté d'un processeur Intel core 2 duo T9300 cadencé à 2,5 Ghz avec 4 Go de mémoire vive, acheté en 2008,
- HP Compaq NC8430 doté d'un processeur Intel core 2 duo T5500 cadencé à 2 Ghz avec 2 Go de mémoire vive, acheté en 2007

De plus, quelques vieux ordinateurs portables (Intel PIV) sont réservés au terrain.

L'infrastructure réseau permet une connexion filaire à 100Mo/s ou Wifi.

Dans le schéma suivant, les ressources informatiques du centre SITE sont illustrées.

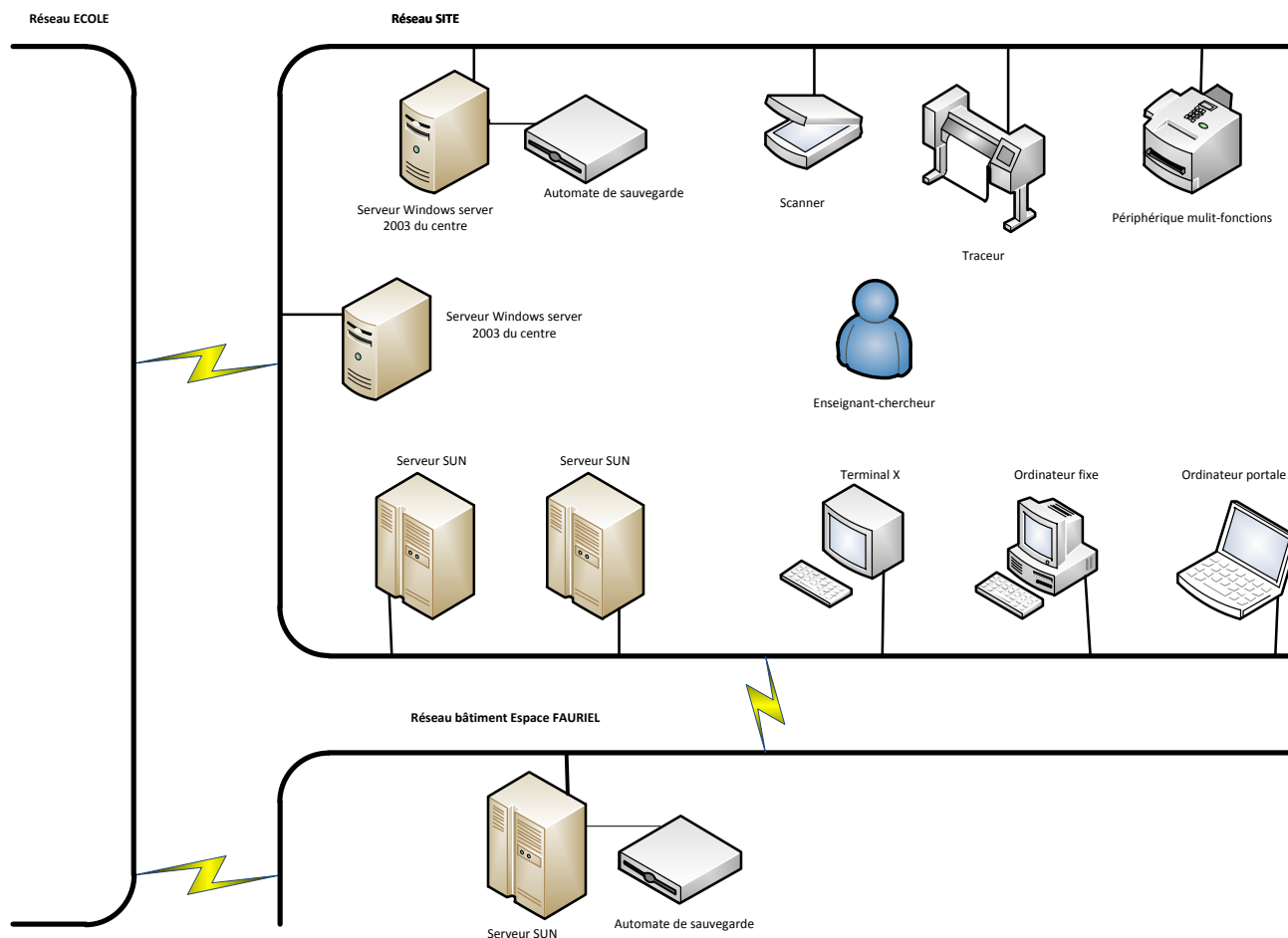


Figure 3 : illustration des ressources informatiques du centre SITE

Logiciels

Le centre dispose également de moyens de modélisation : ADMS pour la dispersion atmosphérique de polluants, FLUENT en mécanique des fluides, MODFLOW pour la simulation des écoulements souterrains et ARCGIS pour l'analyse spatiale.

Les logiciels de recherche dont dispose le laboratoire en plus de tous les logiciels de bureautiques communs à l'Ecole sont les suivants :

- un logiciel de Système d'Information Géographique : ArcGIS comprenant ArcInfo, ArcEditor, ArcView, les modules Spatial Analyst, 3D Analyst, Network Analyst, Publisher, GeoStatistical Analyst, Arc PAD, Schematics, ArcScan, Maplex ainsi que ArcGis Server Entreprise Basic
- un logiciel de cartographie par GPS : ArpentGIS GPS

- un logiciel de post traitement des données acquises par GPS Differential : GPS Pathfinder Office
- un logiciel pour les satellites géostationnaires : GPS Planning
- un logiciel de calcul par éléments finis de modélisation de l'écoulement transitoire des eaux souterraines : MicroFEM Hydrodynamique
- un logiciel de modélisation des écoulements souterrains : MODFLOW
- un logiciel de préparation des données (interpolation, maillage, raffinement): Visual 3D Builder
- un logiciel d'interpolation : SURFER
- un logiciel SIG en mode raster (image) : IDRISI
- un logiciel de modélisation hydrodynamique de surface : HEC-RAS
- un logiciel d'acquisition de données : NMEA
- un logiciel de dispersion atmosphérique : ADMS
- un logiciel de mécanique des fluides : ANSYS Fluent
- un logiciel d'analyse des données : SPAD
- un logiciel d'analyse de modélisation : ANSYS DesignXplorer
- un logiciel de modélisation de débit d'air : AIRPACK
- un logiciel de modélisation des systèmes dynamiques : VENSIM
- un logiciel de calcul à partir de données météorologiques : FLOWSTAR
- un logiciel de calcul scientifique : MATLAB

Après avoir décrit l'infrastructure informatique, il sera abordée l'organisation des ressources en géomatique dans le prochain chapitre.

Chapitre 3 – L'organisation de l'information géographique numérique d'IDEE

1. Historique et organisation actuelle

De par son origine, l'Ecole des Mines a toujours eu un attachement au territoire. L'enseignement et la recherche concernant la géologie ont tenu durant deux siècles une place importante au sein de l'institution. Aujourd'hui, cet attachement aux problématiques territoriales reste présent, en particulier pour les problématiques environnementales.

C'est dans les années 1990, qu'un enseignant-chercheur en ingénierie de l'environnement a introduit à l'ENSMSE le Système d'Information Géographique qui constituait à l'époque un nouvel outil.

Un système d'information géographique est un *"système informatique de matériels, de logiciels, et de processus conçus pour permettre la collecte, la gestion, la manipulation, l'analyse, la modélisation et l'affichage de données à référence spatiale afin de résoudre des problèmes complexes d'aménagement et de gestion"* selon le comité fédéral de coordination inter-agences pour la cartographie numérique (FICCDC, 1988).

Cet enseignant-chercheur a débuté sa carrière à l'ENSMSE au département de Géologie en 1972. Ses activités pluridisciplinaires (géologie, exploitation des mines, travaux publics, hydrogéologie, informatique scientifique, mathématiques appliquées, recherche opérationnelle, gestion de projets) l'ont conduit à mettre en place un ensemble d'outils graphiques d'aide à la décision, en particulier dans le domaine de l'information environnementale.

L'introduction de cet outil a nécessité une lourde prise en main. Tout d'abord, le SIG choisi, ArcGIS de la société ESRI, fonctionnait sur un serveur SUN dont le système d'exploitation est Unix. L'affichage était déporté sur un terminal X Tektronix. La maîtrise du logiciel a nécessité un apprentissage assez long. A cette époque, l'enseignant-chercheur a mis en place cet outil seul, ou simplement aidé par les fournisseurs. Il est devenu « l'expert » du SIG.

Il a encadré des chercheurs pour les applications géomatiques se rapportant à l'environnement avec un vif intérêt pour les mesures de précision par système DGPS (GPS Différentiel). Effectivement, beaucoup de données nécessaires aux recherches n'existaient

pas. C'est pourquoi, le département a acquis un certain nombre de matériels permettant la production de données.

Les données utilisées étaient stockées sur les disques durs du serveur SUN.

Sous chaque partition se trouvent des espaces utilisateurs dans lesquels sont stockées des données géographiques regroupées dans des répertoires correspondant à des projets de recherche. Les données sont copiées et recopiées selon les nécessités des projets. Les droits d'accès aux données sont restreints à l'utilisateur ou à un groupe d'utilisateurs.

A cette époque, le nombre d'utilisateurs était restreint : deux enseignants-chercheurs et leurs doctorants.

Une quinzaine d'années passaient. Les technologies évoluaient. Il n'est plus nécessaire aujourd'hui d'avoir des serveurs Unix pour utiliser les SIG et effectuer les calculs.

En 2006, l'enseignant-chercheur se rendant compte des limites de l'organisation des données et de leur accessibilité, décide de mettre en place un serveur de données spatiales qui permet de stocker, gérer et utiliser des données spatiales dans un SGBD. L'étude a débuté avec l'aide d'un étudiant en master spécialisé en Génie Logiciel, mais elle n'a jamais abouti.

Le projet visait l'implantation d'une géodatabase ARCSDE, produit ESRI. Cet outil aurait permis de stocker les informations dans une base de données relationnelles, telle que Oracle, Microsoft SQL Server, IBM DB2, PostgreSQL ou IBM Informix. Un des facteurs d'échec fut le nombre nécessaire de connaissances à maîtriser tant en système Unix qu'en SGBD Oracle.

En 2008, l'enseignant-chercheur effectue la migration d'une grande partie des données géographiques sous un serveur Windows du centre. Aucune restriction d'accès n'est mise en place.

Au fur et à mesure que les années passent, l'utilisation du SIG se généralise. Les projets se multiplient et les données sont multiples et acquises ou produites par près de dix enseignants-chercheurs.

Actuellement, les données sont achetées ou acquises par l'expert selon les besoins du département. C'est lui qui les commande ou les produit.

Aujourd'hui, ces données, qui sont de plus en plus propres à chaque utilisateur, sont en partie gérées individuellement et dispersées sur différents ordinateurs personnels et dans plusieurs répertoires, ainsi que sur deux serveurs de fichiers. Pour les plus récentes, elles ne sont pas partagées sur le réseau interne, leur échange s'effectue à la demande.

2. Gestion actuelle des données

Un certain nombre de données sont propres à chaque utilisateur. Il n'y a pas de mutualisation sauf pour deux ensembles de données existantes. La première datant de l'ère Sun et la seconde de l'ère Windows.

La première arborescence mise en place est la suivante :

- de /sig10 à /sig20
- et de /sig251 à /sig256

L'espace disque utilisé est de 138 Go au total, soit 17 partitions réparties sur 17 disques de durs respectivement de 2 Go à 18 Go.

Le schéma suivant illustre la structure de l'arborescence des données géographiques se trouvant sur l'un des serveurs SUN.

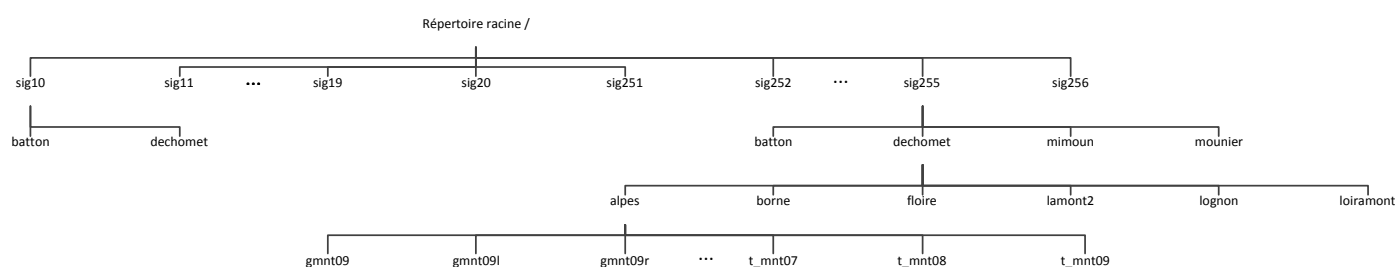


Figure 4 : illustration de l'arborescence de données géographiques sur SUN

Sous chaque partition, des espaces utilisateurs ont été créés et sous ces derniers des espaces projets existent, comme l'illustre le tableau 1 suivant pour la partition /sig256 répertoire dechomet.

Tableau 1 : arborescence du répertoire /sig256/dechomet

Répertoires	Fichiers de l'arborescence	Sous-répertoires de l'arborescence
/sig256/dechomet/alpes	g_mnt09.aux g_mnt09r.aux g_mnt09r.rrd g_slp09p.aux g_slp09p.rrd IGN93_517csp.jgw IGN93_517csp.jpg IGN93_517csp.jpg.aux IGN93_517csp.jpg.aux.xml IGN93_517csp.jpg.rrd log slope_g_mnt01 Slope_g_mnt01.aux Slope_g_mnt01.rrd	g_mnt09 g_mnt09l g_mnt09r g_slp09lp g_slp09p info Slope.mxd t_mnt07 t_mnt08 t_mnt09
/sig256/dechomet/borne	Les arborescences de ces répertoires ne sont pas développées.	
/sig256/dechomet/floire		
/sig256/dechomet/lamont2		
/sig256/dechomet/lognon		
/sig256/dechomet/loiramont		

Cette organisation de données était pertinente lors des balbutiements du SIG. Dès que le nombre d'utilisateurs et de projets a cru cette structure a atteint ses limites.

Cette arborescence existe toujours actuellement, bien que les serveurs aient été changés. Elle comporte des données qui n'ont pas été « migrées » du système SUN au système Windows

La seconde arborescence, mise en place en 2008, sur le serveur Windows commun au centre SITE, utilise un espace disque de 224 Go et est organisée en répertoires portant :

- soit le nom du producteur de données nationales telles que l'IGN
- soit celui du projet auquel elles se rapportent

Le schéma suivant illustre la structure de l'arborescence des données géographiques se trouvant sur le serveur Windows commun au centre.

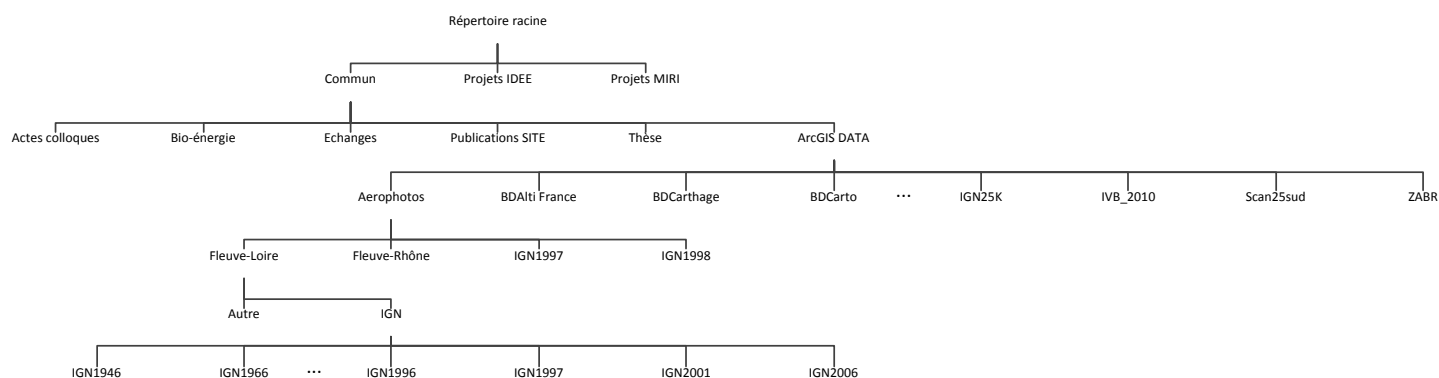


Figure 5 : illustration de l'arborescence de données géographiques sur le serveur Windows

La structuration de cette seconde arborescence est présentée dans le tableau suivant.

Tableau 2 : arborescence du répertoire du serveur Windows

Répertoires	Sous-répertoires	Sous sous répertoires
Aerophotos	Fleuve-Loire	Autre IGN <i>contenant les répertoires</i> → 2829 → 2830 → IGN0000 → IGN1946 → IGN 1966 → IGN 1974 → IGN 1979 → IGN 1980 → IGN 1985 → IGN 1990 → IGN 1991 → IGN 1996 → IGN 1997 → IGN 2001 → IGN 2006
	Fleuve-Rhône IGN1997 IGN1998	
Autres	Les arborescences de ces répertoires ne sont pas développées.	
BDAlti France		
BDCarthage		
BDCarto		
BDortho2001		
BDortho2006		
CalculIBV		
Cartes géol		

CorinneLC	
DDE Roanne	
Dept42	
Dore	
Fleuve_Loire	
Fleuve_Rhône	
Galets_Mamy	
Geodesie-nivellement	
Geol 1M BRGM	
Geol SE BRGM	
IGN25K	
Images Sat	
IVB_2010	
Modflow	
PLGN_3	
Python	
Scan25_RRA	
Scan25sud	
Scan_divers	
ZABR	

Les « aérophotos » sont des photographies aériennes dont l’emprise géographique concerne plusieurs sites d’étude dans le département de la Loire, le long du fleuve Rhône et ailleurs. Ces photographies permettent des analyses diachroniques par comparaison à différentes périodes.

Les données contenues dans les répertoires IGN « année » sont des données de type raster, composées de quatre fichiers. Un certain nombre d’informations décrivant la ressource est contenu dans ces fichiers, il s’agit de l’étendue géographique, de la référence spatiale, des informations sur le raster et sur la source de donnée.

Les ressources du répertoire « BDAlti France » sont des données spatiales altimétriques hydrologiques (limites de bassin versant), géologiques, couverture et occupation du sol, etc.

Les données sont disponibles pour tous les utilisateurs du centre SITE.

Cette seconde organisation des données s’affranchit du propriétaire. Cette structure par projet ou fournisseur est plus pertinente et intuitive. Par contre, certaines données sont dupliquées dans plusieurs répertoires, selon qu’elles ont été utilisées pour différents projets. De plus, les enseignants-chercheurs peuvent travailler sur différents projets concernant la même zone géographique en simultanément et effectuer des traitements sur les

données en parallèle aux différentes études afin d'obtenir des résultats. Cette manière de travailler génère des problèmes pour connaître la « version » de données.

Cette organisation permet un accès aux données, mais ne fournit aucune information sur les données.

De façon historique, les données géographiques du département, ont été organisées par rapport à leur support physique. Cette organisation assez erratique doit évoluer vers une organisation plus performante des données. L'installation et l'utilisation d'un serveur de données géographiques doit être envisagé.

Mais avant d'envisager une réorganisation, il est important pour une utilisation pertinente des données géographiques de connaître un certain nombre d'informations qui permettent de les caractériser sans avoir à les examiner en détail.

Une gestion plus adéquate et mieux organisée s'avère donc indispensable pour capitaliser et organiser l'information environnementale générée par les activités du département. C'est l'objectif général du présent travail.

Dans cette première partie, l'organisation professionnelle, l'environnement informatique du projet ont été exposés. Puis la gestion actuelle des ressources en géomatique du département IDEE a été présentée. Il s'agit maintenant de proposer des solutions possibles pour mieux gérer et organiser l'ensemble de ces informations.

Partie 2

Présentation du projet et analyse des besoins

Dans cette partie il s'agit de formaliser la problématique, d'exposer la méthode de travail et d'aborder les normes en vigueur pour aboutir à un Modèle Conceptuel de (méta)Données approprié.

Chapitre 1 – Etat des lieux et analyse des besoins

1. Constat

Comme cela a été précisé au chapitre 3 de la partie précédente, voici presque vingt ans que l'utilisation de SIG a débuté au centre SITE.

Il est indéniable que la quantité importante de données accumulées durant plusieurs années constitue le patrimoine du laboratoire.

Le départ en retraite imminent de l'enseignant-chercheur responsable du SIG a mis en exergue l'intérêt de la patrimonialisation des données du département. Ceci a favorisé le démarrage du projet fin 2009.

Effectivement, il est pratiquement le seul actuellement à détenir l'expertise des données géographiques. Il a produit ou participé à la production de toutes les ressources en géomatique. Il détient l'historique des données achetées ou mises à disposition et des contrats réalisés.

Personne ne maîtrise aussi bien que lui les mesures de précision permettant la création de données géographiques. Cette perte de compétence qu'il est prévu de renouveler pourra être réduite par le développement de l'utilisation (DGPS) et la maîtrise des outils d'acquisition de données par les enseignants-chercheurs.

De plus, l'assistance technique qu'il fournissait aux enseignants-chercheurs pourra être remplacée par des formations permettant de maîtriser les outils informatiques utilisées.

Par contre, la perte de l'expertise sur les données en géomatique ne pourra pas être comblée par une compétence humaine. En effet, la multiplicité des projets et des données ne peuvent plus être gérées uniquement au niveau individuel. Une grande quantité de documents hétérogènes (cartes, plans, photo aériennes, images satellites, documents textuels) existe actuellement et intervient dans le développement des projets de recherche.

Sans l'élaboration et la mise en place d'un système d'information organisé comportant une couche de métadonnées pertinentes les problèmes suivants peuvent devenir irrémédiables :

- oubli de l'existence de la donnée
- difficulté de connaître le contenu et l'étendue des données
- ignorance des documents disponibles

- difficulté de localiser physiquement les documents
- incapacité de connaître quels traitements la donnée a subi
- impossibilité de connaître l'origine de la donnée

En effet, personne d'autre que l'enseignant-chercheur retraité ne peut répondre aussi rapidement que lui aux questions suivantes :

- Existe-t-il des données sur telle zone géographique ?
- Quelles données ont été utilisées pour tel projet ?
- Il y a quinze ans un Modèle Numérique de Terrain (MNT) a été élaboré pour telle zone, aujourd'hui un nouveau projet est à l'étude sur la même zone. Est-il possible de réutiliser ce modèle ?

Face à ce constat, ce projet s'est avéré. Il s'agit d'un véritable projet de département qui a pour finalité de faciliter les activités académiques en structurant et organisant l'information afin de permettre la capitalisation et l'exploitation des données géographiques de manière structurée et durable.

Le département IDEE a besoin de mettre en valeur et de partager en interne ces données. Il doit donc disposer d'un système d'information permettant leur gestion et leur exploitation. Dans le cadre de ce travail, il s'agit d'étudier la faisabilité et la mise en service d'un système de distribution d'informations sur les données selon les besoins des utilisateurs.

Plus précisément, l'objectif de ce projet est la définition et la mise en place d'un système d'information permettant de référencer, d'indexer, de gérer, d'administrer et d'archiver les informations concernant les ressources géographiques disponibles du laboratoire de recherche IDEE, tout en permettant leur traçabilité et leur accessibilité, avec une notion de qualité des données à l'image de ce que l'on trouve sur le WEB (recherche thématique et géographique, géo portail, ...). Le système d'information est destiné à diffuser les données uniquement en interne et doit garantir leur caractère confidentiel.

2. Les acteurs du projet

Le commanditaire est le directeur du centre SITE, qui a une vision précise du besoin et une connaissance fine du domaine. J'ai eu trois entretiens avec lui, qui ont permis d'affiner ma perception du métier et du besoin.

Les acteurs du projet sont au nombre de douze. Il s'agit des sept enseignants-chercheurs du département. Quatre sont directement concernés et ont participé activement aux réflexions sur ce projet. Les cinq doctorants du département ont pris une part importante dans les discussions et dans les propositions relatives à ce projet.

A partir du besoin exprimé « capitaliser et organiser l'information environnementale générée par les activités du département » je les ai accompagnés dans leur démarche pour aider à définir ce qu'ils attendaient d'un système d'information pour la gestion des ressources en géomatique.

J'ai formalisé l'expression des besoins, identifié les enjeux et contraintes du projet et défini les objectifs. Pour cela, j'ai analysé l'existant et les méthodes de travail utilisées.

Bien sûr, chaque étape du projet a été validée par les acteurs.

Devant l'imminence du départ en retraite de l'enseignant-chercheur, j'ai privilégié des longs entretiens avec lui afin de recueillir sa vision du projet. L'enseignant-chercheur expert des données part en retraite début avril 2010. Je l'ai ainsi interviewé quatre fois.

Durant ces entretiens, j'ai pu acquérir une vision globale des données dont disposait le département :

- des cartes de géologie à différentes échelles : 1/25000^{ième}, 1/500000^{ième} et 1/80000^{ième} sous format papier
- des cartes IGN au 1/25000^{ième} sous format papier
- des scans 25 de l'IGN sous format numérique
- des BD Ortho de 2001 pour les régions Rhône-Alpes et PACA et de 2006 pour Rhône-Alpes sous format numérique
- des photographies de l'IGN ou non sous format papier
- une BD ortho créée par l'enseignant-chercheur sur la base de photos aériennes de différentes années sous format numérique
- des études sur le fleuve de la Loire datant de 1850 sous format papier
- des chroniques de mesures sous format papier ou numérique
- des modèles sous format numérique
- des images satellites sous format numérique

- des données brutes sous format papier ou numérique
- des données produites sous format numérique

Afin d'affiner ma compréhension du métier et d'analyser les besoins, j'ai organisé et animé quatre réunions largement ouvertes. Les doctorants étaient toujours présents à ces réunions et leur participation active traduisait un vif intérêt porté au projet et un réel besoin de connaître les ressources en géomatique présentes dans le département. Les enseignants-chercheurs, étant présents au sein d'IDEE depuis plusieurs années, étaient moins avides de connaissances sur les ressources. Leur présence aux réunions était moins systématique.

J'explique cette différence de participation des doctorants et des enseignants-chercheurs par le fait que les plus jeunes sont plus impactés par la nécessité de connaître l'existence de la donnée et son information.

Effectivement, les enseignants-chercheurs bénéficient de leur expérience. Ils ont acquis une connaissance au fil du temps sur les données disponibles au sein du département, et ont constitué un certain nombre de réseaux leur permettant d'obtenir l'information plus facilement.

Le but de ces réunions était en outre :

- de préciser la demande et les attentes des utilisateurs
- de connaître la nature de l'information, son origine et son utilisation
- de clarifier les besoins, d'identifier les enjeux
- de modéliser le travail actuel

Grâce à cette méthode j'ai pu comprendre et formaliser à quel moment de leur recherche intervenait le besoin de connaître les ressources disponibles.

3. Utilisation des ressources géomatiques existantes

Lors de ces réunions, les informations principales relatives à l'utilisation des ressources géomatiques dans le département ont été mises à plat.

Lorsqu'une problématique concernant une zone géographique délimitée est soumise à un enseignant-chercheur (étape 1 du schéma suivant), ce dernier évalue s'il

existe une problématique similaire ayant été déjà traitée par le département et débute la prospection permettant de connaître les données disponibles sur le territoire concerné.

La prospection s'effectue :

- en interne, auprès de la personne référente de l'historique des données et des projets (étape 2 du schéma)
- en externe, auprès de producteurs de données, mais aussi des concessionnaires de réseaux de distribution d'eau, des collectivités locales, des industriels, des cabinets spécifiques et d'autres organismes ou des partenaires (étape 3 du schéma)

Le résultat de sa prospection lui permet de connaître l'ensemble des données disponibles.

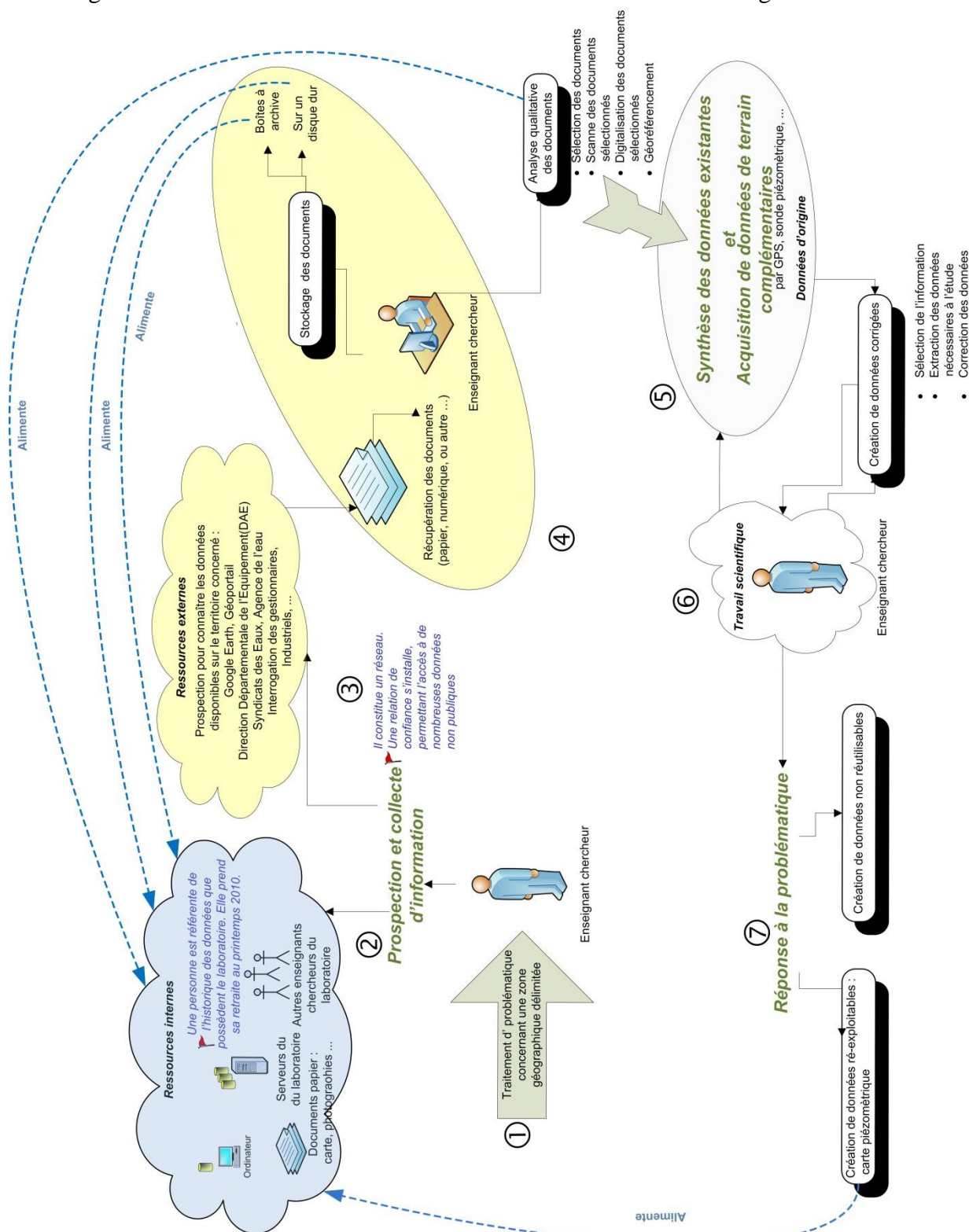
Après avoir étudié la problématique et défini quelles données lui seraient nécessaires, il établit un bilan des données manquantes. Si elles n'existent pas, l'enseignant-chercheur doit pratiquer des relevés de terrain pour les créer. Si elles existent, il doit les collecter (étape 4 du schéma).

La collecte des données se fait soit en interne, soit en externe en achetant la donnée ou en la récupérant à titre gracieux auprès d'acteurs, parfois en établissant des conventions d'utilisation. Dans ce dernier cas, une confiance s'établit entre l'enseignant-chercheur et le partenaire externe. Cette confiance est primordiale et elle doit être préservée.

Puis l'enseignant-chercheur établit une synthèse des données existantes et acquises sur le terrain (étape 5 du schéma).

Après avoir collecté la donnée, l'enseignant-chercheur effectue un certain nombre d'opérations en vue de la corriger et de la traiter (étape 6 du schéma), afin qu'elle permette de répondre à la problématique posée (étape 7 du schéma).

Figure 6 : schématisation des modalités de travail actuelles des enseignants-chercheurs



La prospection pour connaître de façon la plus exhaustive possible toute l'information disponible sur une zone géographique précise est fortement consommatrice de temps.

L'enseignant-chercheur essaie d'obtenir la matière première utile à sa recherche (au sens de son métier) et de connaître en interne l'information disponible.

Il est indéniable que la personne référente de l'historique de la donnée détient un rôle primordial.

Avoir la mémoire de l'existence de la donnée géographique et des traitements qui lui ont été appliqués est une véritable fonction dans le département qui a pour finalité de faciliter les activités académiques en structurant et organisant de l'information afin de permettre la capitalisation et la valorisation des données géographiques.

Il devient donc imminent de mettre en place un système d'information afin de faciliter :

- le réemploi des données
- la mémoire de l'existence des données, de leur création et des traitements qu'elles ont subi, ainsi que de leurs conditions d'utilisation
- le partage et l'accès de l'information géographique
- la localisation de la donnée
- la description des données

Sans solution, un temps précieux sera perdu par les enseignants-chercheurs lors de leurs travaux. En effet, la prospection des données disponibles sur un territoire délimité, ainsi que l'acquisition de la donnée sont des actions fortes consommatrices de temps. Lorsqu'il y a achat de la donnée, le coût financier est important.

4. Un système d'information de métadonnées

Suite à cette analyse, je propose la mise en place d'un outil de communication, de mutualisation et de diffusion autour de l'information géographique, un système d'information permettant de localiser les données et de fournir un certain nombre d'informations sur la donnée, d'acquérir l'information qualitative au sujet des documents disponibles : un système de métadonnées adapté aux besoins des utilisateurs.

Effectivement, une métadonnée est une donnée qui a pour but de décrire un ensemble ou un lot de donnée.

Selon le vocabulaire de la géomatique, publié par l'Office de la langue française en 1993, une métadonnée est "*une donnée qui renseigne sur la nature de certaines données et qui permet ainsi leur utilisation pertinente*".

Les métadonnées sont aux données ce que l'étiquette est à un produit : il faut informer les usagers de leur existence, de leurs caractéristiques et des précautions d'emploi à respecter, ainsi que d'habituer les créateurs à décrire leurs données.

On s'oriente donc vers un système d'information contenant des métadonnées, qui doit permettre de :

- trouver la donnée
- connaître la donnée
- savoir qui détient la donnée
- connaître sa valeur et sa fiabilité

Ce système doit permettre de créer une mémoire officielle et exhaustive des données spatiales pour le département.

L'analyse des besoins effectuée a permis d'identifier deux axes principaux :

- la capitalisation de la connaissance sur le patrimoine de données du département
- l'organisation des données géographiques

Avant de pouvoir organiser les données, il faut les connaître. C'est pourquoi dans ce projet, il s'agit de mettre en place un système de métadonnées précédant un inventaire exhaustif des données. Cette phase ultérieure est non prise en charge dans ce travail. En effet, cette étape ne peut être réalisée qu'après avoir mis en place un système de métadonnées approprié et cohérent avec les normes en vigueur.

5. Facteurs de risques et de réussites

La principale difficulté à l'issue de ce projet est le travail considérable à réaliser pour établir un recensement et un inventaire exhaustif du patrimoine en ressources géomatiques. Aujourd'hui, il n'existe aucun recensement des données géographiques du département IDEE. Ce dernier possède des données pouvant dater de 1850, et non seulement en France, en Europe, mais aussi au Maroc et en Amérique. Le mode d'acquisition des données n'est pas aisé à identifier.

Le recueil des informations liées aux données existantes est un travail qu'il faudra effectuer après la validation de l'outil et de sa maquette, afin de constituer un géocatalogue. Il sera difficile de générer des métadonnées a posteriori.

Les fonctions de l'enseignant-chercheur retraité n'ont pas été reventilées entre les membres du département. Aucune certitude n'existe quant au maintien de son poste budgétaire dans la structure.

Cette charge de travail, importante, risque de justifier le report des renseignements des métadonnées. C'est un point à ne pas négliger.

Pour que le système de métadonnées puisse vivre plus facilement, il faudra que la saisie de la métadonnée intervienne dès l'acquisition ou la production d'un lot de données. Dans le cas de l'acquisition de données auprès d'un autre organisme (IGN, partenaire, ...), il faudra veiller à ce que les métadonnées soient livrées en même temps que les données.

Comme dans beaucoup de laboratoire de recherche, les enseignants-chercheurs du département IDEE travaillent en flux tendu, ils ont peu de temps à consacrer pour le renseignement des métadonnées. Lors des rencontres avec les utilisateurs, j'ai constaté que le manque de temps et donc d'implication était un frein à la réussite du projet, malgré la motivation de la plupart des chercheurs.

Actuellement, les données nécessaires aux projets en cours sont préparées par l'enseignant-chercheur retraité pour les membres d'Idee, ce qui ne permet pas encore de se rendre compte de la réelle perte de compétence. Ce manque sera ressenti pleinement dans les mois prochains lorsque de nouveaux projets débiteront.

Tant que l'enseignant-chercheur est en poste, les membres d'Idee ne ressentent pas la nécessité de dissocier de la personne la connaissance sur les données. Aucune

sensibilisation et prise de conscience sur la nécessité de mettre en place des métadonnées n'a été faite.

Durant les réunions que j'ai animées, nous nous sommes rendus compte que l'utilisation des outils de la société ESRI n'est pas optimale (problème de formation sur les nouvelles versions, évolution des fonctionnalités d'ArcGIS, interface de la couche métadonnées difficilement accessible). C'est pourquoi, trois rencontres thématiques avec les représentants de cette société et les membres du département IDEE, ont été organisées. Ceci dans le but de préparer le groupe à l'évolution des méthodes de travail, de démontrer l'intérêt du partage d'information en interne et de la nécessité d'y passer du temps. Il s'agit de faciliter par là-même, l'appropriation du futur système d'information par les acteurs, mais aussi d'amorcer les réflexions sur la création de métadonnées.

Ces rencontres ont connu un véritable succès, tous les membres du département étaient présents.

Le facteur humain peut être aussi bien un facteur de réussite que d'échec pour ce projet. C'est pourquoi, il était important d'associer dès le démarrage du projet les utilisateurs, en prenant en compte leur problématique. Il faut toujours garder à l'esprit qu'il est plus facile de moderniser un outillage qu'un état d'esprit. Effectivement, l'outil désiré ne peut exister et vivre que si les utilisateurs se l'approprient et que la charge de travail liée à cette appropriation est raisonnable.

Au niveau technique, le projet bénéficie des réflexions sur la problématique engagée dans le milieu des années 1990. Le recul et les échecs rencontrés durant cette quinzaine d'années devraient en être un gage de réussite.

En effet, le développement d'un géorépertoire est possible, mais sa maintenance et son évolution sont scabreuses. L'exemple de MétaZABR que j'aborderai en partie 3 illustre cette difficulté.

Je préconise que soit constitué un groupe de trois personnes comprenant un expert en géomatique, l'ingénieur de recherche travaillant sur le projet ZABR et un informaticien. Ce groupe prendra en charge l'inventaire exhaustif des données disponibles et formalisera

la meilleure façon de créer les métadonnées sur le patrimoine existant. Dans la troisième partie, j'aborderai ce point plus précisément.

Après avoir établi l'état des lieux du projet, analysé les différents besoins, identifié les acteurs et les risques encourus, nous allons maintenant décrire l'information géographique utilisée et les normes associées.

Chapitre 2 – Nature de l'information : des données aux métadonnées

1. Description de l'information géographique utilisée

Il devient impératif de constituer un jeu de métadonnées permettant de décrire les données pour les connaître, les faire connaître et les utiliser à bon escient, dans le respect des droits et des responsabilités qui y sont attachés. Une description précise de leurs caractéristiques techniques intrinsèques, des conditions, des précautions d'emploi et des restrictions techniques et juridiques éventuelles doit être faite.

Les métadonnées sont des sources d'informations importantes et des outils de travail indispensables pour l'usage optimal des données.

Les métadonnées se composent d'éléments relatifs à l'identification, la représentation spatiale, la qualité, le contenu, les modalités d'accès et de diffusion des données, etc. Elles permettent de les classer, de les décrire, elles offrent une traçabilité de la qualité de la donnée, sans avoir à la consulter.

Lors des réunions des groupes de travail, un certain nombre d'informations nécessaires (métadonnées) à la connaissance de la donnée ont été listées, afin de bien identifier les documents et d'en comprendre le contenu exact :

- nom officiel
- échelle
- producteur
- distributeur
- modes d'acquisition
- mise à jour
- nature
- source
- formats disponibles
- support
- norme utilisée
- localisation physique

- obstacle légaux à l'accès
- coût financier
- personne à contacter
- description du système de référence
- couverture spatiale
- couverture temporelle
- structure de la donnée
- systèmes d'exploitation
- utilisation
- valeur

Cette liste non-exhaustive a pour vocation d'acquérir de l'information qualitative au sujet des documents disponibles, de créer une mémoire officielle pour le département.

Ceci en vue de favoriser la consultation de données et de permettre une réponse rapide aux questions suivantes :

- Quels sont les documents disponibles sur un territoire donné ?
- Comment puis-je les obtenir ?
- Que contient-il ?
- Sous quels formats sont-ils disponibles ?
- Sur quoi porte la donnée ?
- Qui produit la donnée ?
- A quelle occasion la donnée a-t-elle été produite ?
- Comment la donnée a-t-elle été produite ?
- Quelle est la qualité de la donnée ?
- Comment peut-on accéder à la donnée ?
- Que décrit la donnée ?
- Qui a produit les métadonnées renseignant cette donnée ?

Cette liste de questions provient des réflexions menées en commun avec les utilisateurs mais aussi de retours d'expériences comme par exemple pour le MétaZABR abordé en partie 3.

La métadonnée permet de répondre aux questions : qui (à créer la donnée), quoi (ce qu'elle contient), quand (sa date de création, sa date de validité), où (zone géographique concernée), pourquoi (domaine d'usage) et comment (méthode de traitement).

Après avoir fait le constat de la nécessité de ne pas perdre la mémoire des ressources géomatiques et après avoir mis en évidence les facteurs de réussite et d'échec de ce projet, il s'agit d'aborder maintenant l'environnement extérieur se rapportant à la problématique, à savoir les systèmes normalisés de métadonnées pour des applications dans le domaine de l'environnement.

Le Modèle Conceptuel de métaDonnées (MCD) développé dans le cadre de ce travail sera présenté dans le chapitre 3 de cette partie.

2. La norme internationale ISO 19115

La prise de conscience de l'importance de la géographie et de la manière dont les éléments sont associés spatialement, combinée avec le progrès de la technologie électronique, a abouti, à l'échelle internationale, à un développement rapide de l'utilisation de l'information géographique numérique et de systèmes d'information géographique.

A mesure que les producteurs et les utilisateurs de données géographiques traitent de plus en plus de données, il est nécessaire d'avoir une documentation appropriée leur assurant une connaissance plus approfondie de leurs fonds de bibliothèque et leur permet de mieux gérer la production, le stockage, la mise à jour et la réutilisation des données.

Les métadonnées doivent être de qualité et apporter des informations comprises par tous, c'est pourquoi une normalisation s'est avérée nécessaire. Une action internationale a été entreprise par le Comité technique 211 de l'ISO. L'ISO est l'Organisation Internationale de Normalisation.

C'est dans ce cadre que fut développée la norme ISO 19115 [AFNOR 2005], qui a le statut de norme internationale depuis 2003. Elle a pour objectif de fournir une structure

permettant de décrire les données géographiques numériques. Elle définit les éléments de métadonnées, fournit un schéma et établit un ensemble de terminologies, de définitions.

Cette norme abstraite de contenu définit en les organisant par classes toutes les informations que l'on peut mettre à disposition pour décrire la donnée.

Cette norme a pour but de :

- fournir aux producteurs de données des informations appropriées permettant de caractériser correctement leurs données géographiques, de sorte que les utilisateurs puissent comprendre les contenus, les limitations et évaluer l'applicabilité de l'ensemble de données à leur utilisation prévue
- faciliter l'organisation et la gestion des métadonnées pour les données géographiques
- permettre aux utilisateurs d'appliquer les données géographiques de la façon la plus efficace grâce à leur connaissance des caractéristiques de base des données ; une documentation appropriée permettra à ceux qui sont peu familiers avec les données de mieux les comprendre et de les utiliser correctement
- faciliter la découverte, la recherche et la réutilisation des données ; les utilisateurs pourront mieux localiser, évaluer, acheter, utiliser les données géographiques et mieux y accéder ; une documentation appropriée assure une connaissance plus approfondie de leurs fonds de bibliothèque et leur permet de mieux gérer la production, le stockage, la mise à jour et la réutilisation des données
- permettre aux utilisateurs de déterminer si les données géographiques présentes dans un fond de bibliothèque leur seront utiles

Le document de 140 pages définit les sections des métadonnées, des entités de métadonnées et des éléments obligatoires et conditionnels de métadonnées, l'ensemble minimum de métadonnées requis pour servir la gamme complète des applications de métadonnées.

Cette norme fournit des informations sur l'identification, l'étendue, la qualité, le schéma spatial et temporel, la référence spatiale, et la répartition des données géographiques numériques.

Elle est affirmée comme une référence pour l'information géographique dans le domaine des métadonnées.

Ses atouts principaux résident dans son caractère modulaire et extensible qui la rend aisément adaptable.

Au sens de la norme ISO 19115, les dix principaux éléments de métadonnées sont présentés dans le tableau suivant.

Tableau 3 : les dix principaux éléments de métadonnées au sens de la norme ISO 19115

Éléments de la métadonnée	Attributs de l'élément
Informations d'identification	<ul style="list-style-type: none"> • intitulé, description, dates de référence, version, résumé, intervenants, ... • étendue des données • aperçus sur les données • informations sur les emplois possibles contraintes légales et de sécurité, ...
Description du contenu	
Système de coordonnées	
Informations de géolocalisation et d'organisation des données	
Informations de qualité (ou de qualification)	
Mesures de qualité	<ul style="list-style-type: none"> • précision géométrique, temporelle et sémantique, • exhaustivité, • cohérence logique
Informations de généalogie	<ul style="list-style-type: none"> • description des sources • description des processus appliqués aux sources
Modalités d'affichage (légendes)	
Modalités de diffusion	
Modalités de maintenance	

Dans la norme ISO 19115 les ressources géomatiques sont classées en 20 thèmes :

- Agriculture
- Flore et faune
- Limites politiques et administratives
- Climatologie et météorologie
- Economie
- Topographie
- Ressources et gestion de l'environnement
- Géosciences
- Santé
- Carte de référence de la couverture terrestre
- Infrastructure militaire
- Hydrographie
- Localisation des réseaux
- Océans
- Planification et gestion du territoire
- Société
- Aménagements urbains
- Infrastructure et transport
- Réseaux de télécommunication et d'énergie
- Inconnu

La figure suivante illustre le Modèle Conceptuel de la norme ISO 19115.

Ce modèle conceptuel de métadonnées a été développé pour répondre aux besoins des utilisateurs.

Ce modèle s'appuie sur des types de données spécifiés par les normes de base suivantes :

- ISO 19103 pour les types de données de base (chaînes de caractère, nombre,...)
- ISO 19107 pour les primitives géométriques servant notamment à la description de l'étendue de la ressource
- ISO 19108 pour les primitives temporelles servant notamment à la description des informations de maintenance liées à la ressource

La norme ISO 19139 [AFNOR 2006] permet l'encodage la norme 19115 au format XML pour permettre une standardisation et l'interopérabilité des systèmes en vue d'échanger et de mutualiser les métadonnées. C'est une norme complémentaire de la norme ISO 19115.

La norme française NF EN ISO 19115

La norme européenne EN ISO 19115 a le statut d'une norme française. Elle reproduit intégralement la norme internationale ISO 19115 qui n'existe actuellement qu'en version anglaise.

Autres normes ou directives

INSPIRE : Infrastructure for Spatial Information in the European Community

La directive INSPIRE est une directive européenne datant de 2007. Elle établit une infrastructure d'information spatiale en Europe pour soutenir les politiques environnementales communautaires et des politiques ou activités qui peuvent avoir un impact sur l'environnement. C'est une spécialisation de la norme ISO 19115.

Le projet européen INSPIRE, dont la vocation est d'insister sur l'importance de disposer de métadonnées harmonisées, libres d'accès, afin de faciliter l'usage des données géographiques, vise à mettre en place une Infrastructure de Données Géographiques en amenant les acteurs publics de l'information géographique vers le géocatalogage, l'harmonisation des pratiques, la mise en réseaux et l'élaboration d'accords de partage et d'accès aux données.

Avec la directive INSPIRE, la mise en œuvre des métadonnées devrait, en 2010, se généraliser dans tous les établissements publics producteurs de données géographiques.

Elle crée plusieurs obligations telle que la fourniture des données selon des règles de mise en œuvre communes, la constitution de catalogues de données (métadonnées), l'application de règles d'interopérabilité, l'accès gratuit aux métadonnées, l'accès aux données pour les acteurs réalisant une mission rentrant dans le cadre d'INSPIRE, les services pour permettre ces accès, l'existence d'une organisation adaptée pour s'assurer de la bonne mise en œuvre de la directive.

Dublin core

Le Dublin Core Metadata Initiative, est une organisation qui s'occupe du développement des normes de métadonnées interopérables qui supportent un large éventail d'applications.

Le Dublin Core naît de travaux de bibliothécaires qui de tout temps ont rédigé des notices, ou des index, pour décrire les documents disponibles. Effectivement, une partie de leur travail est de créer des données sur les ouvrages afin d'en faciliter la recherche.

Cette démarche s'applique particulièrement bien aux pages Web, car il est intéressant de coder certains renseignements essentiels des pages : le titre, l'auteur, la date, les mots-clefs, etc.

Mais pour que cette pratique puisse se développer, il est nécessaire de parler de la même chose. C'est pourquoi il a été établi une définition et une nomenclature communes des métadonnées.

En 1995, un atelier à Dublin (Ohio) a réuni un certain nombre d'équipes impliquées dans la sémantique sur le Web. De cet atelier est né le Dublin Core Metadata Initiative.

C'est la définition d'une liste précise de quelques métadonnées liées aux pages Web. Une norme a été établie, il s'agit de la norme ISO 15836-2003, RFC 5013.

L'utilisation de cette description de métadonnées peut être appliquée à différents type de métadonnées.

Le Dublin Core est un format de métadonnées indépendant du standard conceptuel ISO 19115.

Federal Geographic Data Committee (FGDC)

Né en 1990, le Federal Geographic Data Committee est un comité inter-organismes faisant la promotion de l'utilisation, de la diffusion et du partage coordonnés de données géospatiales aux États-Unis.

Ce comité a décrit une norme structurant en une hiérarchie de champs de données et de champs composés qui définissent le contenu informationnel des métadonnées, permettant de documenter des données géospatiales numériques.

Cette norme permet aux utilisateurs de données de déterminer :

- quelles sont les données disponibles
- si les données répondent à leurs besoins particuliers
- où trouver les données
- comment accéder aux données

Les normes sont des référentiels, des guides de méthodologie, des règles. Elles sont le résultat d'un consensus raisonnablement établi entre les instances représentatives dans le domaine au niveau international. Elles sont le recueil des bonnes pratiques, de l'état de l'art sur un sujet déterminé et délimité à un instant donné. Elles posent un cadre pour mieux travailler dans le domaine. Elles sont essentiellement descriptives. Elles sont élaborées en vue d'une diffusion large à vocation pédagogique.

Leur utilisation est un gage de cohérence avec ce qui se fait actuellement dans le domaine. Elles ont été rédigées pour être utiles à la communauté. Réaliser un projet en se basant sur elles, permet d'employer et de modéliser les projets de manière à ce qu'ils soient compris par le plus grand nombre.

Les utiliser permet de bénéficier de l'expérience de ses rédacteurs, malgré les contraintes et obligations qu'elles engendrent. En outre, elles sont les garantes d'une cohérence, d'une compatibilité et de l'interopérabilité des données.

Les normes ne sont pas directement utilisables et appliquées totalement car elles sont :

- trop complètes et détaillées
- trop compliquées

- trop longues à renseigner
- pas forcément adaptées aux thématiques de travail

La norme ISO 19115 comporte plus de 600 champs.

Si nous devons développer un outil spécifique, il doit prendre en compte le cadre de référence que représentent ces normes.

Après avoir formalisé la problématique, exposé ma méthode de travail et abordé les normes en vigueur, je vais maintenant présenter le modèle conceptuel de données que j'ai déterminé.

Chapitre 3 – Le Modèle Conceptuel de (méta)Données

1. Le cycle de vie de la métadonnée géographique

La définition du cycle de vie des métadonnées permettant de décrire les données mobilisables pour traiter les problématiques environnementales, constitue un préalable à l'élaboration d'un modèle conceptuel de métadonnées.

Le cycle de vie des métadonnées comprend trois phases principales :

- une phase d'élaboration et de saisie des métadonnées : les données acquises aujourd'hui doivent l'être obligatoirement avec leurs métadonnées. La création a posteriori des métadonnées du patrimoine géomatique existant demeure un travail considérable.
- une phase de gestion des métadonnées comprenant la mise à jour ; les métadonnées et les données qu'elles renseignent forment un couple indissociable et ont un cycle de vie similaire. Toute modification de la donnée entraîne la mise à jour de sa métadonnée.
- une phase d'archivage permettant de conserver la donnée et ses métadonnées.

Les données sont dans notre cas a priori des informations pérennes, car elles peuvent servir pour comprendre l'évolution des jeux de données eux-mêmes et des problématiques environnementales dans lesquelles elles interviennent.

Compte-tenu de l'état des lieux, de l'analyse des besoins, des normes en vigueur et du cycle de vie des métadonnées, il est proposé le modèle conceptuel suivant.

2. Modèle Conceptuel de (méta)Données proposé

Ce modèle s'articule autour de l'entité « Métadonnées » qui comporte des informations sur la donnée géographique.

Autour de cette entité principale gravitent d'autres entités pouvant être regroupées en trois groupes : la définition de la donnée géographique, l'origine de sa création et les personnes physiques ou morales ayant une incidence sur la donnée (créateur, contact, utilisateur).

Ce modèle est constitué de 16 entités et de 18 associations.

Les cardinalités allant de « Métadonnées » aux autres entités sont des relations 1,1 sauf pour « Information qualité » qui est 0,1 et pour « Domaine d'usage » qui est de 1,n.

Effectivement, l'information sur la qualité de la donnée n'est pas toujours disponible, une donnée peut se rapporter à plusieurs domaines d'usage.

L'entité au cœur du modèle « Métadonnées » a pour attributs :

- identifiant de la métadonnée
- titre de la métadonnée
- description de la donnée
- résumé de la donnée
- étendue de la donnée
- information sur les emplois possibles de la donnée
- contraintes légales et de sécurité
- échelle de la donnée
- origine de la donnée
- précision de la donnée
- unité de la donnée
- date de référence de la donnée
- date de modification de la donnée
- rédacteur de la métadonnée
- localisation de la donnée

Les entités suivantes permettent la compréhension de la donnée géographique :

- « Emprise spatiale » définissant l'étendue du territoire.
- « Système de coordonnées » précisant la projection utilisée.
- « Description du contenu » décrivant le type de données (modèle, carte, document), le format (papier, numérique), l'extension de fichier et la nature (biologie, écologie, hydrologie, météorologie, sociologie, ...). Les attributs

de cette entité sont : l'identifiant de la ressource, son type, son format initial, son extension fichier et sa nature.

- « Information qualité » portant sur la qualité de la donnée, l'indice de qualité est son attribut. Il est prévu pour l'instant d'évaluer la qualité de la donnée à l'aide de trois critères qualitatifs (indice de qualité) : bon, moyen, mauvais ; cette échelle peut être amenée à évoluer en fonction de l'introduction progressive des données et des métadonnées.
- « Mesure de qualité » détaillant la précision géométrique qui dépend des outils d'acquisition des données (GPS, DGPS). La cohérence logique de la donnée permet d'identifier les interruptions éventuelles dans des séries de mesure par exemple. Les attributs de cette entité sont : la précision géométrique et la cohérence logique de la ressource.
- « Mots-clés » regroupant les termes qui décrivent et caractérisent la ressource. L'utilisation de mots-clés requiert toute fois un consensus des utilisateurs à travers l'élaboration d'un thésaurus commun. Il existe des normes de mots-clés pour l'environnement.
- « Logiciel » précisant le logiciel utilisé pour consulter la donnée. Il peut s'agir d'outils de statistique, d'outils d'analyse spatiale ou de modèles de simulation. Les attributs sont : le nom du logiciel, sa version, son éditeur, son distributeur, son possesseur.

Les entités suivantes permettent de comprendre l'origine de la création de la ressource :

- « Généalogie des sources » caractérisant l'auteur de la donnée, les sources et les processus appliqués aux ressources ; une entité « projet » lui est rattachée afin de lier la source à un projet (étude de site ou étude d'impact par exemple). Les attributs de « Généalogie des sources » sont : l'auteur et la description des sources.
- « Modalité de diffusion » indiquant l'accessibilité à la ressource et la restriction d'utilisation ; cette entité est reliée au groupe par l'association avec l'entité « Contact ». Ses attributs sont : l'accessibilité à la ressource et ses restrictions d'autorisation.

- « Domaine d'usage » précisant les domaines pour lesquels la ressource peut être utilisée.

Les entités suivantes permettent d'identifier les personnes physiques ou morales liées à la création, à l'acquisition et la maintenance des données :

- « Contact » indiquant la personne référente à contacter pour la ressource. Ses attributs sont : prénom et nom.
- « Partenaire » précisant les liens de partenariat, par exemple, en cas de création conjointe de lot de données.
- « Organisation » identifie l'organisme. Ses attributs sont : organisme, nom, adresse, tutelle, site web et coordonnées.
- « Laboratoire » précisant le laboratoire à laquelle appartient la personne référente. Ses attributs sont : nom et adresse.

Ce premier modèle a été élaboré suite aux réflexions menées avec les utilisateurs pour prendre en compte, principalement, leurs préoccupations « métier ». Ce modèle conceptuel constitue un ensemble organisé d'informations, qui peuvent être renseignées par les utilisateurs et dont ils peuvent avoir besoin pour mobiliser des ressources géomatiques utiles, indispensables au développement de leurs projets de recherche ou à leurs activités d'enseignement.

Il s'agit d'une première ébauche qui ne prend pas en compte jusque-là les normes en vigueur. L'analyse de ces normes au chapitre précédent va nous permettre d'aboutir à un modèle plus évolué en bénéficiant de l'expérience de la communauté internationale travaillant sur la métadonnée de l'information géographique.

3. Modèle Conceptuel de (méta)Données avec prise en compte des normes en vigueur ISO 19115

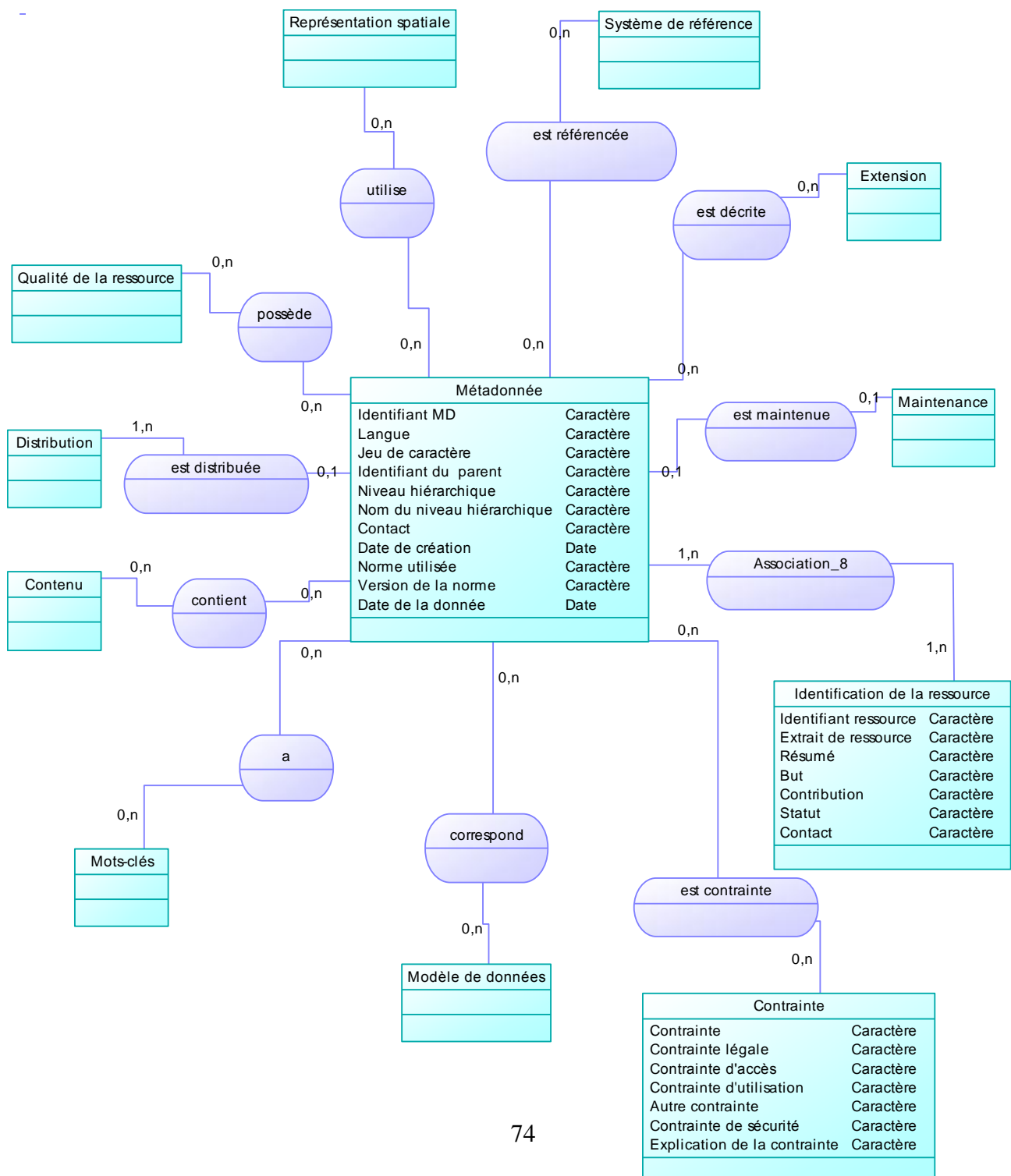
Comme nous l'avons vu précédemment, les normes de métadonnées sont complexes, très complètes et leur utilisation peut s'avérer difficile et problématique pour les usagers. Il s'agit donc de trouver un compromis entre les réponses à apporter en terme de métadonnées aux utilisateurs et les contraintes de normalisation existantes.

Seule la norme ISO 19115 est prise en compte, car elle est LA référence en matière de métadonnées géographiques.

Le but est de proposer un ajustement du modèle précédent permettant à la fois de répondre aux besoins des utilisateurs tout en restant cohérent avec les standards de métadonnées.

Le Modèle Conceptuel de (méta)Données ajusté est présenté dans la figure suivante. Il est composé de 12 entités et de 11 associations, il est donc plus simple que le modèle présenté précédemment.

Figure 9 : Modèle Conceptuel de (méta)Données avec prise en compte des normes en vigueur ISO 19115



Ce nouveau modèle s'articule autour de l'entité « Métadonnées » qui comporte des informations sur la métadonnée de la ressource géographique.

L'entité « Métadonnée » constituant le corpus central comporte 11 attributs qui sont les suivants :

- Identifiant de la métadonnée : code identifiant de manière unique la fiche de métadonnée concernée, c'est une chaîne de caractère.
- Langue : indique la langue utilisée pour décrire la métadonnée. Cet attribut est particulièrement utile lorsque des chercheurs étrangers travaillant dans le laboratoire vont insérer des jeux de données avec leurs métadonnées.
- Jeu de caractère : indique la codification utilisée par le logiciel de gestion des métadonnées.
- Identifiant du parent : décrit la relation entre une collection de données générée par thématiques et un lot de données (mesure, carte). Les collections de données peuvent être associées à des thématiques environnementales particulières (gestion de l'eau, qualité de l'air, gestion des déchets). Ils peuvent être aussi liés à des sites d'étude particuliers et importants (Rhône, Loire, ...). Ces collections de données peuvent comporter plusieurs lots de données de nature différente.
- Niveau de hiérarchie : spécifie si la métadonnée porte sur une collection, un lot ou une série chronologique de mesure, par exemple.
- Les six autres attributs : nom du niveau hiérarchique, contact (personne responsable de la métadonnée), date de création de la métadonnée, norme utilisée, version de la norme, date de la donnée rejoignent les définitions classiques.

Autour de cette entité principale gravitent onze entités qui caractérisent :

- « Identification de la ressource » qui intègre la nature de la donnée (image satellite photographie aérienne, série chronologique de mesure, autres documents). Ses attributs sont : identifiant de la ressource, extrait de la ressource, résumé, but, contribution, statut et contact. Ils seront repris plus tard.

- « Contraintes » sur la ressource, les contraintes peuvent être légales (données protégées par un organisme accrédité (IGN, INSEE) qui peut s'engager sur la validité des données) ou sécuritaires (accès à la ressource soumise à l'autorisation du fournisseur/producteur de données ou accès temporaire à la donnée). Les contraintes peuvent aussi être des contraintes d'accès tel que l'accès à la ressource via un site web dédié. Les attributs de cette entité sont : contrainte légale, contrainte d'accès, contrainte d'utilisation, autre contrainte, contrainte de sécurité, explication de la contrainte (texte libre expliquant la contrainte).
- « Qualité de la ressource » comporte aussi la généalogie de la ressource ainsi que les méthodes d'acquisition ou de production.
- « Maintenance » de la ressource, indique quelle maintenance lui est appliquée.
- « Représentation spatiale » indique quel mode de représentation spatiale est utilisé pour représenter la ressource (vecteur, raster).
- « Système de référence » décrit les références spatiales (système de coordonnées, comme par exemple le Lambert I ou le Lambert II) et temporelles (pas de temps d'acquisition des mesures : horaire, journalier, mensuel, annuel) utilisées.
- « Contenu » de la ressource décrit la couverture spatiale de la donnée. Il peut s'agir de couverture administrative (commune, département, région), il peut s'agir aussi de sites d'étude (bassin versant).
- « Modèle de données » permet de visualiser et d'afficher la ressource.
- « Distribution » de la ressource définit auprès de qui la ressource est disponible : collectivités territoriales, institutions, ADEME, Agence de l'eau, etc. Il peut s'agir aussi d'une personne ressource particulièrement impliquée dans la constitution, la maintenance du lot de données voir sa validation.
- « Extension » de la métadonnée permet l'ajout d'attribut de métadonnée, tel que le moissonnage (récolte de métadonnées).

- « Mots-clés » définissent le type et les références des sources, ils comprennent les mots couramment utilisés pour décrire le sujet, le type de données. Ils peuvent être contenus dans un thésaurus.

Par rapport au modèle précédent, la métadonnée est reliée spécifiquement à une entité « Identification de la ressource » vue précédemment, qui comporte sept attributs. Ce sont les attributs d'identification de la ressource :

- Identifiant de la ressource : permet l'identification unique de la ressource
- Extrait de la ressource : aperçu de la donnée
- Résumé : bref résumé du contenu de la ressource. Ce résumé est conçu pour disposer d'un aperçu rapide sur la nature de la donnée, son type et son contenu.
- But : but dans lequel a été créée la ressource. Cet attribut représente entre autre le domaine d'usage des données : campagne de mesure, étude d'impact, modélisation.
- Contribution : personne ayant contribué à la création de la ressource. Cet attribut permet d'identifier la ou les personnes référentes liées à son acquisition, sa création, sa production et/ou sa maintenance.
- Statut : statut de la métadonnée
- Contact : cet attribut permet d'identifier la ou les personnes qu'il faut contacter pour avoir des informations sur la métadonnée.

Ce modèle conceptuel de données identifie avec distinction les informations concernant la ressource de celles de la métadonnée, qui se trouve en son cœur. Le premier modèle conceptuel de métadonnées avait pour principal défaut de représenter une vision plus métier de la ressource. Il est moins orienté métadonnée.

En terme de bilan, si nous comparons les avantages de ce nouveau modèle qui tient compte des normes en vigueur avec le précédent, nous pouvons résumer les modifications apportées de la manière suivantes.

De nouvelles entités sont apparues :

- « Représentation spatiale » remplace « Emprise spatiale ».
- « Système de références » remplace « Système de coordonnées ».
- « Maintenance » est totalement nouvelle, cette information n'avait pas été évoquée lors de l'élaboration du premier modèle.
- « Extension » est totalement nouvelle, en effet, il est important de prévoir des extensions permettant l'ouverture de l'outil.
- « Qualité de la ressource » regroupe les entités « Information qualité », « Mesure de qualité », « Généalogie des sources » et « Projet ».
- « Modèle de données » contient l'entité « Logiciel ».
- « Contrainte » contient l'entité « Modalité de diffusion ».
- « Distribution » contient les entités « Organisation », « Laboratoire », « Partenaire ».
- « Contenu » contient l'entité « Domaine d'usage ».
- « Identification de la ressource » contient « Description du contenu », ainsi que des attributs de l'entité « Métadonnées » du premier modèle.

L'entité « Mots-clés » a été conservée.

L'entité « Métadonnées » présente dans les deux modèles ne modélise pas du tout la même information. En effet, au départ cette entité caractérisait la ressource et sa métadonnée, maintenant la métadonnée a une existence propre et des attributs qui la décrivent. La clé de voûte de l'évolution du modèle se situe là.

Au final, suite à ce travail d'analyse nous aboutissons à un système de métadonnées optimisé mais perfectible qui répond aux besoins exprimés des utilisateurs et qui prend en compte les systèmes de norme en vigueur.

Après avoir présenté la structure générale du système d'information de métadonnées souhaité, il s'agit maintenant d'identifier un outil de géocatalogage approprié permettant d'implanter ce modèle conceptuel de (méta)données. C'est ce qui sera exposé dans la 3^{ème} partie.

Partie 3

Réalisation et mise en œuvre de l'outil

Dans cette partie, le tour d'horizon des outils de géocatalogage existants sur le marché et de ceux dont le département IDEE dispose sera effectué. Puis, le croisement entre les possibilités existantes et les besoins sera réalisé afin de proposer un outil à mettre en œuvre.

Il s'agit de trouver un outil permettant d'implémenter ou d'implanter le Modèle Conceptuel de (méta)Données développé dans le chapitre précédent. L'examen de plusieurs outils de géocatalogage (MDweb, GeoNetwork, GéoSource, ArcCatalog et MétaZABR) est effectué dans le premier chapitre de cette partie. Le second chapitre présentera la solution retenue. Une étude de coût sera jointe. Enfin, la mise en œuvre de l'outil sera présentée.

Chapitre 1 – Outils existants

Dans ce chapitre, on examinera les possibilités offertes par quatre outils parmi les plus connus du marché : MDweb, GeoNetwork, GéoSource et ArcCatalog.

1. MDweb

MDweb est un outil libre de catalogage, de localisation de ressources et d'indexation de l'information géospatiale. Il s'appuie sur les standards de métadonnées en vigueur (ISO, Dublin Core) et de communication de l'information géographique.

Depuis 2007, MDweb a été développé en partenariat entre l'Institut de Recherche pour le Développement (IRD) et les laboratoires informatiques et thématiques régionaux et nationaux : le LIRMM (Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier) et d'autres Unités Mixtes de Recherche (UMR) de l'Université de Montpellier II, ainsi que le CEMAGREF.

Il intègre un serveur de cartographie MapServer et utilise PostgreSQL.

C'est un logiciel libre Open Source.

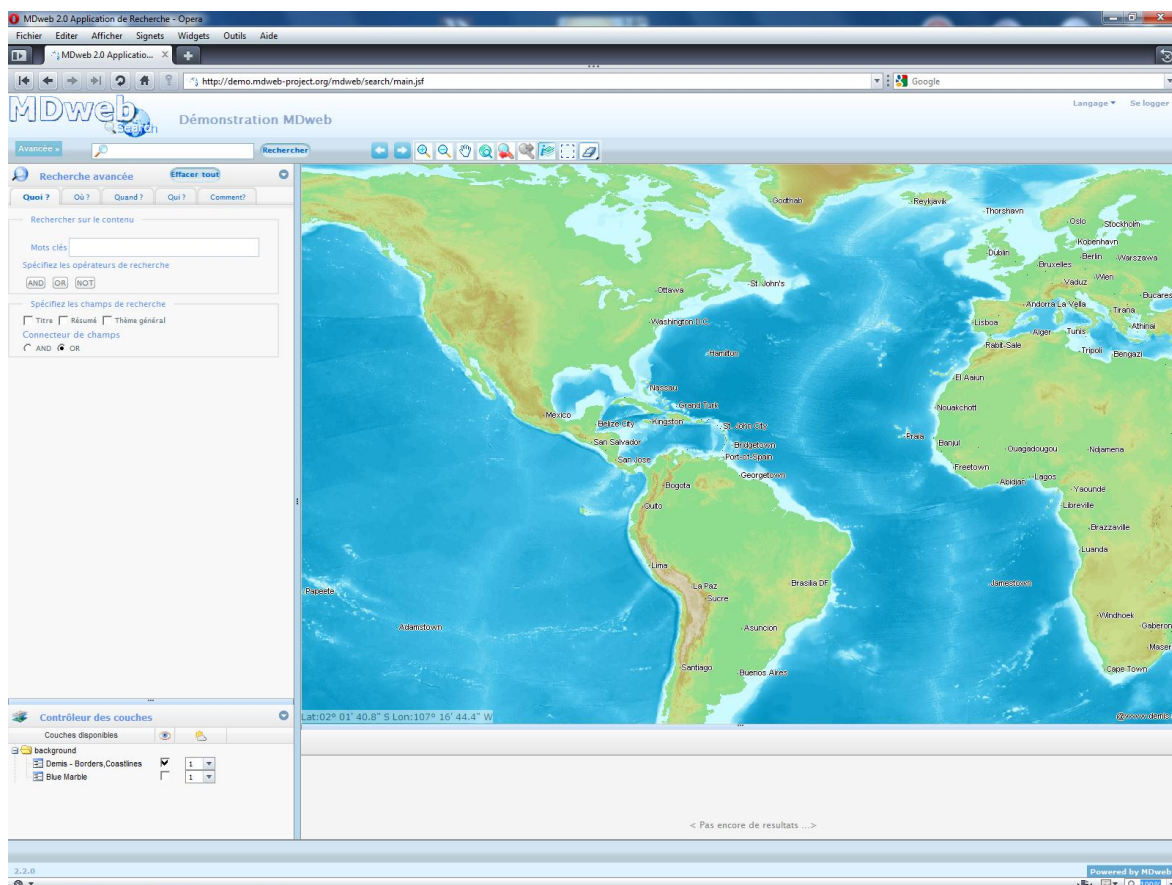
La dernière version est disponible depuis juillet 2010 et est très bien documentée. Il est à noter que durant deux ans aucune version n'a été distribuée. La version 1.6 est sortie en juillet 2007.

Il est mis à disposition sur la plate-forme : <http://www.mdweb-project.org>

Aucun recul sur la nouvelle version n'est possible.

La figure suivante est une copie d'écran de l'interface graphique de MDweb.

Figure 10 : aperçu de l'interface graphique de MDweb (source <http://demo.mdweb-project.org/mdweb/search/main.jsf> consulté en juillet 2010)



2. Geonetwork

Geonetwork est un gestionnaire de système d'information géographique. Il est développé conjointement par le Programme des Nations Unies pour l'Environnement (PNUE) et l'Organisation des Nations Unies l'alimentation et l'agriculture (la Food and Agriculture Organization of the United Nations - FAO), la version 2 a vu le jour en 2004.

A partir de la version 2.4 parue en juillet 2009 (en août 2010 est sortie la version 2.6), les utilisateurs peuvent travailler avec de multiples standards de métadonnées (ISO 19115, Federal Geographic Data Committee (FGDC) et Dublin Core). Il est aussi possible de partager les métadonnées entre les catalogues grâce au moissonnage.

Geonetwork est diffusé sous licence gratuite General Public License (GPL). Il intègre un large support des différentes normes internationales, et dispose de larges possibilités de personnalisation. Il supporte un grand nombre de langues.

Cet outil offre, de façon autonome, les fonctions classiques d'un dispositif simple de gestion d'informations descriptives : recherche, consultation et publication de fiches de métadonnées.

C'est un logiciel robuste, stable et respectant des normes officielles qui bénéficie d'une grande communauté de développeurs. L'accès aux informations se pratique via un navigateur internet.

Il est utilisé par de nombreux organismes.

C'est un outil datant de 2003, il est donc difficile d'avoir un véritable recul sur son utilisation.

GeoNetwork est mis à disposition sur la plate-forme : <http://sourceforge.net/projects/geonetwork/>

La figure suivante est une copie d'écran de l'interface graphique de GeoNetwork.

Figure 11 : aperçu de l'interface graphique de Geonetwork (source <http://www.fao.org/geonetwork/srv/fr/main.home> consulté en juillet 2010)



3. Géosource

Le Ministère de l'écologie et du développement durable a impulsé une démarche donnant naissance à un outil de géocatalogage (Géosource) en vue d'alimenter le géocatalogage français.

Le logiciel GéoSource basé sur GeoNetwork, est libre et est distribué sous licence GPL. Il a été réalisé dans le cadre du projet ADELE, sous maîtrise d'ouvrage de la Direction Générale de la Modernisation de l'Etat (DGME) et maîtrise d'œuvre du Bureau de Recherche Géologique et Minière (BRGM).

Certaines fonctionnalités ont été spécialisées pour la France, notamment la charte graphique et le profil français. Dans la mesure du possible, GeoNetwork et GéoSource seront diffusés simultanément lors des prochaines versions.

Géosource s'adresse essentiellement aux organismes qui ne disposent pas déjà de leur propre dispositif de catalogage conforme aux normes ISO relatives aux métadonnées.

Il a également été conçu afin d'alimenter facilement le géocatalogue, catalogage fédératif national et moteur de recherche du Géoportail, portail de l'information géographique publique.

Géosource a vocation à gérer les fiches de métadonnées géographiques selon la norme internationale 19115 et sa déclinaison française (profil français) et selon les règles de mise en œuvre de la Directive INSPIRE. Ces règles de mises en œuvre concernent :

- la saisie et la validation de fiches descriptives de métadonnées, ainsi que les documents complémentaires sur la description des données (documents associés, fichiers d'aperçu)
- la publication des fiches et la production des fiches dans un format interopérable qui permet de les mettre à disposition du plus grand nombre (norme XML 19139)
- la recherche et la consultation des fiches descriptives définies par un partenaire

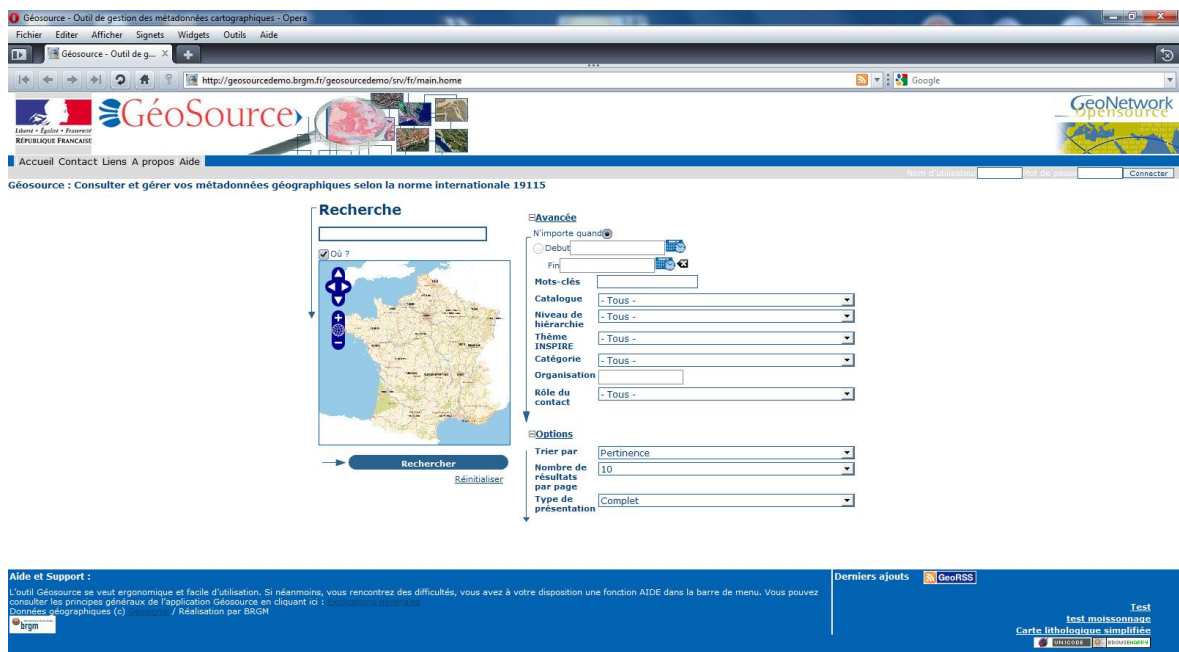
Géosource ne gère pas les données décrites par les métadonnées. En conséquence il n'est pas possible d'accéder aux données géographiques directement.

Géosource est mis à disposition sur la plate-forme : <http://www.geosource.fr/>

Géosource est utilisé par le BRGM, qui met à disposition un certain nombre d'outil, notamment un formulaire de saisie des métadonnées au format Inspire via Excel qui se trouve en annexe 1.

La figure suivante est une copie d'écran de l'interface graphique de Geosource.

Figure 12 : aperçu de l'interface graphique de Géosource (source <http://geosourcedemo.brgm.fr/geosourcedemo/srv/fr/main.home> consulté en juillet 2010)



4. ArcCatalog

Le logiciel ArcGIS Desktop développé par la société ESRI dispose d'un module ArcCatalog. Ce module est utilisé pour organiser et gérer toutes les données et métadonnées SIG. Il est distribué sous une licence commerciale.

ArcCatalog permet aux utilisateurs de créer, gérer et modifier les métadonnées stockées dans une représentation XML du Federal Geographic Data Committee (FGDC) ou de la norme ISO 19115. Les métadonnées comprennent les propriétés et la documentation associée. Les propriétés sont dérivées à partir de la source de données, tandis que la documentation est fournie par le producteur de données.

La fonction Métadonnées affiche les informations descriptives sur l'élément sélectionné dans l'arborescence du catalogue. Les métadonnées sont stockées en tant que langage de balisage extensible (XML). La métadonnée fait partie de la couche elle-même, elle est copiée et déplacée avec les données et est mise à jour automatiquement par ArcCatalog. Il est impossible d'accéder aux informations des métadonnées via un navigateur internet sans passer par le logiciel ArcServer.

Les métadonnées ne sont accessibles que par l'interface ArcCatalog, leur consultation ne peut se faire sans l'utilisation de logiciel de la société ESRI, ce qui implique un coût financier. De plus, il est compliqué à renseigner.

Les figures suivantes sont les copies d'écran de l'interface graphique de ArcCatalog et celle de son module d'édition de métadonnées.

Figure 13 : interface graphique de ArcCatalog copie d'écran

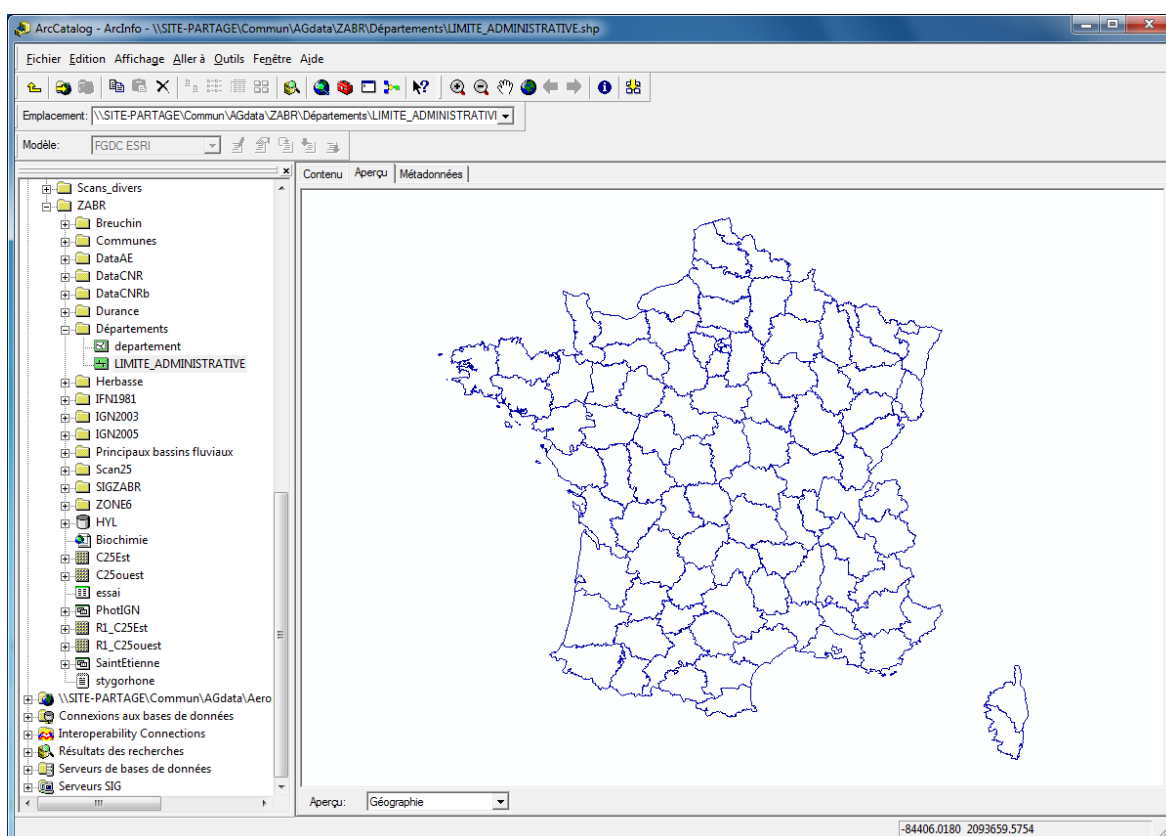
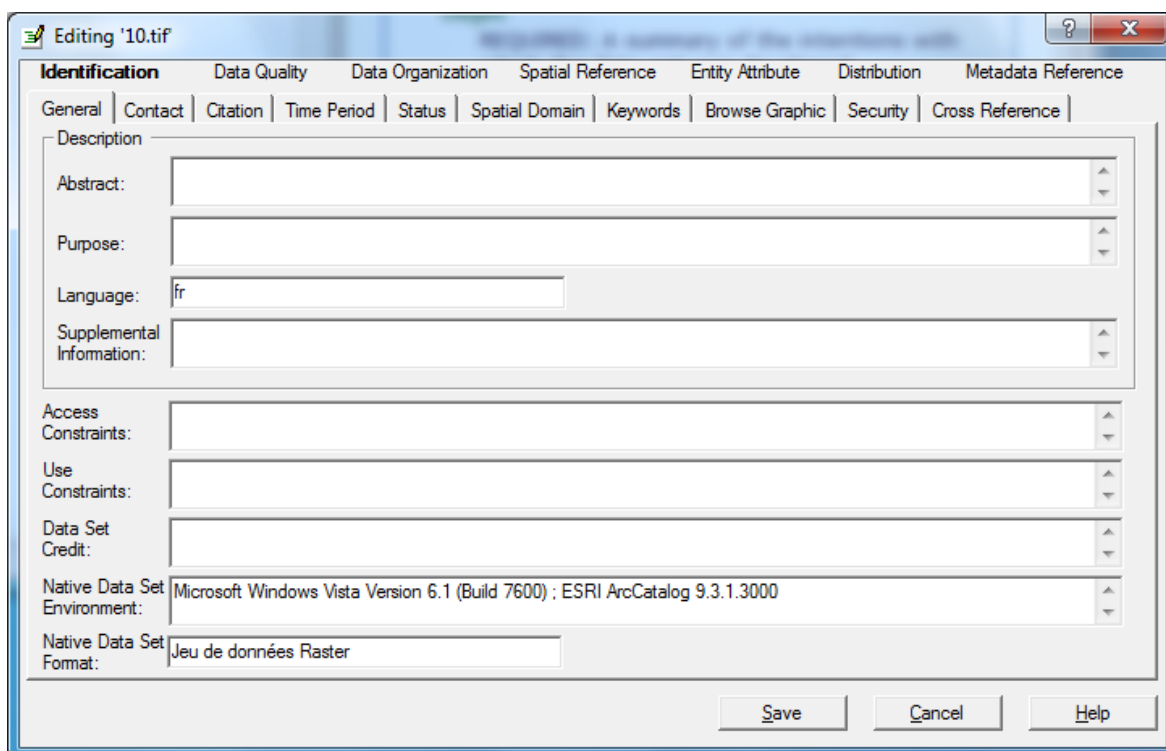


Figure 14 : interface graphique d'édition des métadonnées de ArcCatalog (copie d'écran)



En annexe 2 se trouvent les copies d'écran des autres onglets de l'interface graphique d'édition des métadonnées de ArcCatalog.

5. Ce qui existe dans le laboratoire : le géorépertoire de la ZABR

La Zone d'Atelier du Bassin du Rhône (ZABR) rassemble plusieurs établissements de recherche compétents dans de nombreuses disciplines : climatologie, écologie des hydrosystèmes fluviaux, économie, ethnologie, épidémiologie, géomorphologie fluviale, histoire des hydrosystèmes et des peuplements humains, hydraulique, hydrobiologie, hydrologie, physico-chimie des eaux, sociologie, etc.

Les équipes impliquées dans la ZABR, sont amenées à collecter, gérer et utiliser des informations de type divers (des photographies aériennes, des chroniques de mesure, etc) et de nature très variée (hydraulique, biologique, social, etc). Afin d'aider les chercheurs à accéder à ce patrimoine de connaissances dispersé parmi différents producteurs, laboratoires et institutions, la ZABR a décidé en 2002 de réaliser un géorépertoire : catalogues électroniques, accessibles notamment par Internet, permettant la recherche de données géoréférencées.

Cet outil « MetaZABR » [ENSMSE CEMAGREF] a vu le jour en 2004, sous l'action du centre SITE et l'équipe Informatique de l'unité Hydrologie-Hydraulique du CEMAGREF de Lyon.

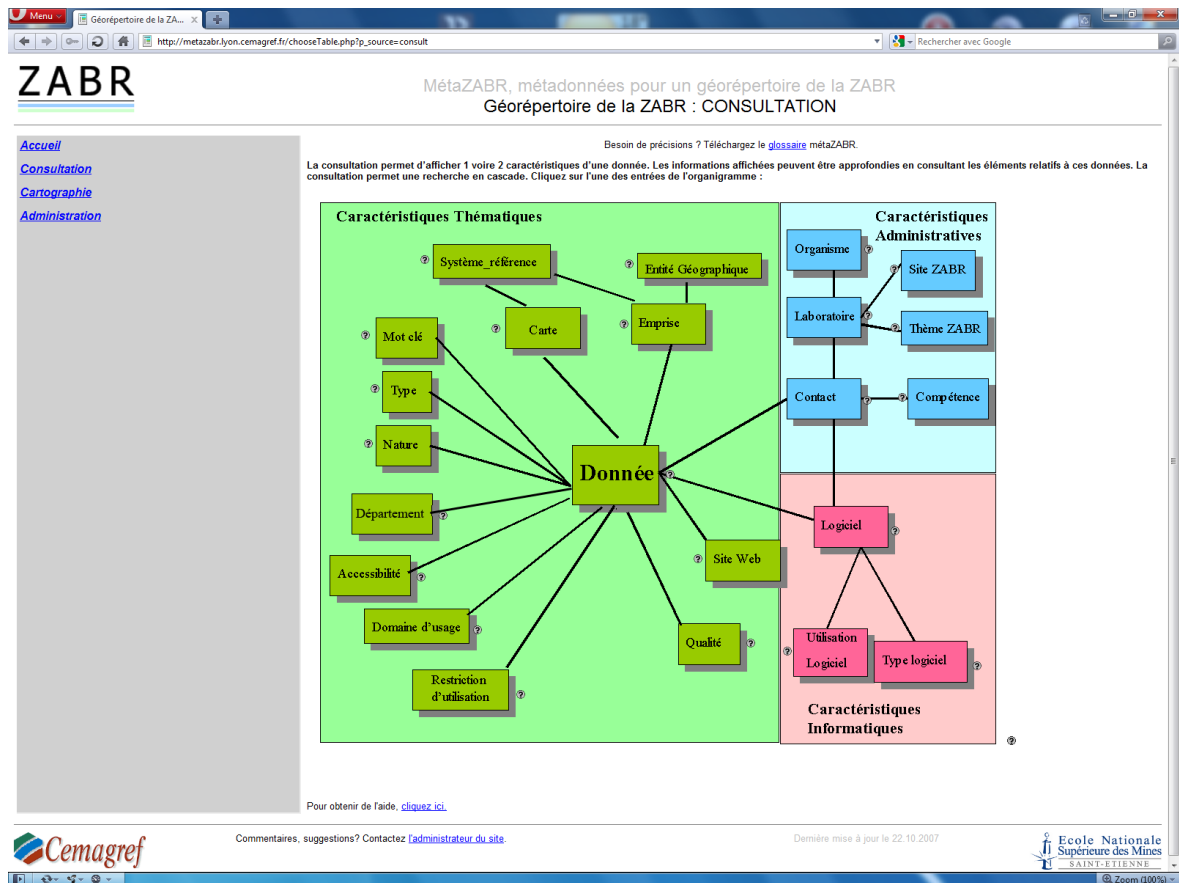
Il a été produit sous PostgreSQL avec un environnement php accessible via internet. Il a été développé en interne et toute évolution génère une quantité non négligeable de développements. Sa maintenance est trop lourde pour permettre son évolution rapide. C'est pourquoi, depuis sa mise en ligne en 2004, il a subi peu de modifications.

En 2009, devant ce constat, la ZABR décide de missionner le centre SITE, l'UMR 5600 du CNRS et le CEMAGREF pour étudier les nouvelles orientations que cet outil pouvait avoir.

Cette étude a abouti au choix par la ZABR de migrer les métadonnées de « MetaZABR » sous GeoNetwork. Cette migration est programmée.

Les figures suivantes sont les copies d'écran de l'interface graphique de métaZABR et celle de son module d'édition de métadonnées.

Figure 15 : aperçu de l'interface graphique de métaZABR (source <http://metazabr.lyon.cemagref.fr.fr/> consulté en juillet 2010)



Chapitre 2 – Solution choisie

1. Les différentes alternatives

Dans le précédent chapitre les quatre principaux outils de géocatalogage disponibles sur le marché ont brièvement été décrits : MDweb, GeoNetwork, GéoSource et ArcCatalog.

Le projet MétaZABR, qui doit très prochainement migrer sous GeoNetwork a été également abordé. Les membres du département participant à la ZABR doivent mettre en ligne les métadonnées des données géographiques dont ils disposent.

En prenant en compte que le facteur d'échec le plus important est le temps nécessaire à la réalisation du recensement et de l'inventaire exhaustif du patrimoine en ressources géomatique du département IDEE, MDweb ne sera pas choisi car cet outil n'est pas aujourd'hui présent dans l'environnement de travail direct des enseignants-chercheurs et nécessiterait du temps supplémentaire pour la prise en main.

ArcCatalog intégré à la suite ArcGIS d'ESRI est disponible dans le département IDEE depuis plus de 15 ans. Il respecte et supporte les normes en vigueur. Le support technique est relativement performant.

Durant ce projet, nous avons constaté que ce logiciel était sous utilisé. C'est pourquoi, il a été programmé un certain nombre de réunions avec ESRI France afin de favoriser le partenariat avec cette société dans le but de faire évoluer les méthodes de travail.

Ces réunions où tous les membres utilisateurs de ArcGis étaient présents ont permis tout d'abord de former un collectif de travail, d'acquérir une meilleure maîtrise du logiciel, et d'aborder les notions de géocatalogage et de géodatabase.

GéoNetwork choisit par la ZABR va être utilisé par tous les membres du département travaillant sur ce projet. L'ingénieur de recherche du département a participé à la migration des métadonnées de MétaZABR à GéoNetwork.

C'est un logiciel libre développé par la FAO. Il bénéficie d'une large communauté de développeurs et d'utilisateurs.

La GéoSource est éliminée car certaines données dont disposent le département portent sur des zones géographiques hors de nos frontières : Suisse, Inde, Pologne.

Le choix va donc porter entre GeoNetwork et ArcCatalog que nous allons comparer de façon plus détaillée.

Les principaux critères de choix sont : le support des normes en vigueur, la facilité d'installation et de maintenance, le tarif, l'échange de données.

Le tableau suivant résume cette comparaison.

Tableau 4 : comparatif entre GeoNetwork et ArcCatalog

Outils	ArcCatalog	GeoNetwork
Version	9.3.2	2.6
Editeur	ESRI	FAO
Normes et standards	FGDC ISO 19115 ISO 19139 Dubli Core	FGDC ISO 19115 ISO 19139 Dubli Core
Langues	Anglais, allemand, français, espagnol, néerlandais, chinois	Anglais, allemand, français, espagnol, néerlandais, chinois, italien, arabe
Licence	Commerciale	Libre et Open Source
Systèmes d'exploitation	Red Hat Suse Linux Sun Solaris 9, 10 Windows Windows Server 2003	Windows Linux Mac OS X
Serveur Web	Apache Tomcat	Apache Tomcat ou Jetty
Base de données	Oracle 9i Microsoft SQL Server PostgreSQL	MySQL PostgreSQL Microsoft SQL Server
Autres logiciels nécessaires	ArcGIS Server Standars Enterprise (incluant ArcSDE) et ArcGIS Desktop	Aucun
Interface au service annuaire centralisé	Microsoft ActiveDirectory LDAP	LDAP
Création de métadonnées et publication	1 – la création de métadonnée s'effectue en utilisant ArcCatalog ou n'importe quel éditeur XML 2 – pour les publier il faut les rendre disponible sur le serveur ArcGIS Server	1 – la création de métadonnées s'effectue en utilisant l'éditeur en ligne 2 – pour les publier il suffit de mettre en publication
Profil utilisateur	Administrateur, éditeur	Gestion des groupes et des utilisateurs Possibilité de définir des privilèges, des profils et des groupes Administrateur, relecteur, rédacteur
Import-Export de données	Oui format XML	Oui format XML

Il est à noter que l'importation et l'exportation des métadonnées est possible dans les deux logiciels sous un format XML.

Rappelons que l'évolution des ressources humaines du département n'est pas actuellement garantie. Il faut donc garder à l'esprit que le responsable du système d'information sera probablement une personne interne au département qui n'est pas informaticien. La maintenance et l'administration du système doivent être simples.

2. Les différents profils utilisateurs

Afin de garantir la fiabilité des métadonnées, il est primordial d'identifier les compétences nécessaires pour créer l'information.

Il s'agit de connaissances de base en géomatique :

- projections cartographiques
- systèmes de coordonnées
- système de référence géodésique
- prises de vue aérienne
- image satellite

Afin de garantir la pérennité des informations, il convient de décrire quatre profils d'utilisateurs:

- administrateur : il s'occupe d'administrer l'outil
- valideur : il coordonne et anime l'utilisation du géocatalogue, il s'occupe de publier et de valider les fiches de métadonnées saisies par les catalogueurs, il est garant de sa bonne utilisation et de la cohérence des informations
- catalogueur : il saisit et met à jour les métadonnées relatives aux données
- lecteur : il accède au géocatalogue pour de la consultation

Il faut identifier officiellement les personnes rattachées à ces différents profils. Ces affectations seront validées ultérieurement après le test de l'outil mis en place.

3. Analyse financière

L'analyse financière présentée est succincte, car les logiciels comparés n'induisent pas de frais supplémentaires directs : ArcGIS est déjà utilisé et GeoNetwork est libre de droit.

L'ENSMSE dispose d'une licence site pour l'utilisation de MS Microsoft SQL Server. Donc l'implémentation de l'outil sur ce SGBD n'induit pas de charge financière supplémentaire.

Le montant de la licence ORACLE Standard Edition s'élève à 3 113,20 euros T.T.C. par an.

Les logiciels ESRI ont un coût global en maintenance annuelle de 3 824,20 euros T.T.C.

Cette maintenance comprend :

- 6 licences ArcInfo Desktop, ArcEditor utilisation concurrente, ArcView utilisation concurrente, ArcGIS Spatial Analyst, ArcGIS 3D Analyst, ArcGIS Network Analyst, ArcGIS Geostatistical Analyst, ArcScan pour ArcGIS, ArcGIS Schematics, ArcGIS Publisher, et Maplex pour ArcGIS extensions pour licences concurrentes
- ArcPad Application Builder (incluant ArcPad)
- ArcGIS Server Enterprise Basic for 9.2 only

Si dans le futur, le choix est fait de mettre en place ArcGIS server, le coût de la version standard d'ArcGIS Server Enterprise est de 2 392 euros T.T.C. et le coût de la version advanced est de 4 784 euros T.T.C. par an.

Ces coûts sont associés aux choix des outils logiciels, cependant d'autres outils d'accès libres ne génèrent aucun coût.

4. Les préférences des utilisateurs

Les membres du département IDEE débutent leur réflexion sur la création de métadonnées. Ils travaillent en flux tendu. Ils sont conscients de l'importance des informations concernant les données et leur gestion. Par contre, aucune métadonnée n'est disponible dans le département, sauf dans MétaZABR.

Les enseignants-chercheurs souhaitent que la solution mise en œuvre :

- leur permette d'aborder concrètement la création des métadonnées
- soit accessible à partir de leur poste informatique en mode création, modification et maintenance
- n'engendre pas de coût matériel
- soit conforme aux normes en vigueur
- soit un outil libre qu'ils pourront proposer à leurs partenaires
- permette d'appliquer aux données des contraintes légales. En effet le statut légal de la donnée préoccupe les enseignants-chercheurs, c'est un point qu'il faudra particulièrement analyser lors de la création de la métadonnée
- s'installe, dans sa version standard, sans connaissance informatique nécessaire
- vive sans trop passer de temps à sa maintenance
- permette le moissonnage des métadonnées par et chez les partenaires du département IDEE
- permette l'import et l'export de métadonnées

Dans ce contexte, le choix se porte finalement sur l'outil GeoNetwork. L'outil est libre accès, et ne génère pas de charge financière supplémentaire.

Il sera installé sur un ordinateur déjà présent au département, donc le projet pourra voir le jour sans coût matériel ni logiciel supplémentaires.

De plus, cet outil n'a pas besoin d'être supporté par un autre logiciel contrairement à ArcCatalog, qui doit être supporté par ArcGIS Server. Il bénéficie de mises à jour périodiques (août 2009 version 2.4, août 2010 version 2.6) et ne nécessite pas de connaissance en développement informatique dans une utilisation standard.

5. Formation

Aucune formation en tant que telle n'est prévue. Par contre, un manuel d'aide au remplissage des métadonnées sera édité afin d'expliquer les normes utilisées et les gabarits définis pour la création des métadonnées, selon le Modèle Conceptuel de (méta)Données finalement retenu dans la 2^{ème} partie.

Ce manuel sera élaboré à partir des réflexions menées par le groupe (abordé en partie 2 chapitre 1 Facteurs de risques et de réussites) devant réfléchir à la meilleure façon de créer les métadonnées sur le patrimoine existant.

Pour que le projet puisse s'inscrire dans le temps, nous continuerons à accompagner les utilisateurs du système en organisant des réunions dont le but sera de gérer la création des métadonnées et de poursuivre le travail de réflexion sur le patrimoine du département.

Chapitre 3 – Mise en œuvre

Après l'étude de faisabilité technique et organisationnelle du système d'information, la mise en place de l'outil choisi GeoNetwork est présenté dans ce chapitre.

GeoNetwork a diverses fonctionnalités, il permet :

- l'accès direct à la recherche dans les catalogues géographiques locaux et distribués
- le téléchargement vers et depuis le serveur de données
- d'être informé des entrées mises à jour grâce à un fil RSS et GeoRSS
- l'édition en ligne des métadonnées
- la gestion des métadonnées aux formats ISO 19115 / ISO 19119 / ISO 19139, FGDC et Dublin Core
- la planification du moissonnage et la synchronisation des métadonnées entre les catalogues distribués
- le contrôle des accès
- la gestion des groupes et des utilisateurs
- l'interfaçage multilingue

Toutes ces fonctionnalités ne serviront sans doute pas dans un premier temps, mais leur utilisation reste possible.

1. Installation

Exigence système

GeoNetwork fonctionne sous MS Windows, Linux ou Mac OS X. Il nécessite un processeur d'au moins 1 Ghz et 512 Mo de mémoire vive. Un espace disque de 250 Mo est nécessaire. Cet espace sera augmenté selon les besoins de l'utilisation de l'outil.

Les logiciels suivants peuvent être utilisés à la place de Jetty (serveur web) et McKoiDB (SGBD) : Java Runtime, Apache Tomcat et un gestionnaire de base de données (MySQL, Postgresql, Oracle).

GeoNetwork supporte les navigateurs Firefox, Internet Explorer et Safari.

Installation de GeoNetwork

Le logiciel GeoNetwork [GeoNetwork opensource] est installé sur un ordinateur de type station de travail HP Compaq xw4400 doté de 4 Go de mémoire vive et d'un processeur Intel Core duo E6700 cadencé à 2,67 Ghz acheté en 2007.

L'installation est faite en utilisant le serveur web Jetty inclus dans l'installation de base de GeoNetwork et le gestionnaire de base de données McKoiDB car à terme nous ne savons pas qui maintiendra le système de métadonnées. Il est probable que ce soit un non-informaticien. De plus, dans la configuration actuelle uniquement 12 personnes sont susceptibles d'accéder en simultané au système.

Après avoir téléchargé le logiciel GeoNetwork version 2.4.3 faisant 169 Mo sur le site : <http://sourceforge.net/projects/geonetwork/files/>, son installation a été effectuée.

Le déroulement est le suivant :

- double cliquer sur geonetwork-install-2.4.3-0 pour démarrer l'installation
- suivre les instructions sur l'écran

Pour utiliser GeoNetwork, il faut tout d'abord démarrer le serveur :

`$INSTALL_PATH/bin/start-geonetwork` (pour l'arrêter il faut utiliser la commande `stop-geonetwork`)

Au premier lancement le système demande quelques minutes pour complètement démarrer.

Il faut se connecter au serveur geonetwork en utilisant un navigateur web à l'adresse suivante : <http://localhost:8080/geonetwork>

Paramétrage

La configuration et l'administration de GeoNetwork peuvent s'effectuer de deux manières :

- par l'application web, cette interface permet de paramétrer la majorité des options
- par l'outil GAST : Geonetwork's Administrator Survival Tool, cette application fournit des tâches spécifiques à l'administrateur du catalogue

L'administrateur de GeoNetwork prendra en charge la gestion et le paramétrage des groupes et des utilisateurs.

J'ai configuré l'outil pour permettre l'authentification des utilisateurs par le serveur LDAP de l'ENSMSE. Ainsi les utilisateurs n'auront pas de mot de passe spécifique à retenir.

L'attribution des profils aux utilisateurs sera finalisée durant la phase de conception des métadonnées.

2. Création de la métadonnée

GeoNetwork dispose de deux outils permettant la saisie des métadonnées selon le Modèle Conceptuel de Métadonnées définie en deuxième partie :

- l'éditeur en ligne disposant d'une interface graphique permettant l'affichage de trois vues de métadonnées, une par défaut, une avancée et une XML
- l'outil d'insertion basé sur les documents XML

Le géocatalogue est fourni avec les modèles de métadonnées pour créer de nouvelles métadonnées en utilisant l'éditeur en ligne. Ces modèles permettent de décrire divers types de ressource (données vecteur ou raster, etc).

Un format MEF (Metadata Exchange Format) a été spécialement conçu pour l'échange de métadonnées et de données.

La figure suivante représente l'interface graphique de saisie des métadonnées sous GeoNetwork.

Figure 16 : aperçu de l'interface graphique de saisie des métadonnées sous GeoNetwork, vue avancée (copie d'écran)

The screenshot shows the GeoNetwork metadata entry interface. On the left, a sidebar contains a tree view of metadata sections: Vue par défaut, Par groupe, ISO Minimum, ISO Core, ISO Complet, Par section, Métadonnées, Identification, Maintenance, Contraintes, Information Spat., Système de Réf., Distribution, Qualité des données, Schéma d'application, Catalogue, Content Info, Ext. Info, and XML. The main content area is titled 'Métadonnées' and contains several input fields: Identifiant du fichier (8534900-f6bc-48c6-8239-09e97f5946b5), Langue (Anglais), Jeu de caractères (UTF8: 8-bit variable size UCS Transfer Format, based on ISO/IEC 10646), Identifiant du parent, Niveau hiérarchique, Nom du niveau de hiérarchie, Date de création (2010-08-25T13:32:45), Nom du standard de métadonnées (ISO 19115:2003/19139), and Version du standard de métadonnées (1.0). Below these fields are sections for 'Contact' (Nom de la personne, Organisation, Position, Rôle) and 'Informations de contact' (Téléphone, Adresse, Ressources on-line). At the bottom, there are buttons for 'Réinitialiser', 'Sauver', 'Sauver et fermer', 'Sauver et valider', 'Vignettes', and 'Annuler'. A 'Type Métadonnées' dropdown is also visible.

En annexe 3 se trouvent les copies d'écran d'une partie des autres rubriques.

A gauche de la fenêtre se trouvent les rubriques :

- Métadonnées
- Identification
- Maintenance
- Contraintes
- Information Spatiale
- Système de Référence
- Distribution
- Qualité des données
- Schéma d'application

- Catalogue
- Content Information
- Extension Information

Les modèles de création de données comportent des champs obligatoires et des champs fortement recommandés. Ils sont repérés dans l'interface grâce à un astérisque rouge.

Une information (info-bulle) sur chaque champ est accessible.

3. Test sur un lot de données particulier

Aucune métadonnée n'étant actuellement disponible, le test a été effectué à partir d'informations communales extraites de BD topo de l'IGN avec des métadonnées renseignées et fournies par la société ESRI.

4. Gains attendus pour les usagers

Avec cet outil, les enseignants-chercheurs du département disposent d'un outil leur permettant la création des métadonnées sur des données environnementales et spatiales. Cet outil facilitera l'inventaire du patrimoine de données du département.

Durant la période estivale 2010, GeoNetwork a été installé et paramétré. Mais sa présentation est prévue au cours du dernier trimestre 2010. Il est effectivement important d'effectuer un retour aux enseignants-chercheurs sur l'intérêt des métadonnées en leur indiquant les champs correspondants à renseigner. Une première rencontre sera programmée courant septembre selon les agendas de chacun.

Certains doctorants ont déjà commencé à élaborer des métadonnées sur leurs données produites. Pour l'instant elles se présentent sous un format texte. Il faudra donc les migrer.

L'étude effectuée durant ce projet permet ainsi aux enseignants-chercheurs d'avoir une vision globale sur les métadonnées et d'utiliser instantanément GeoNetwork.

Pour rendre possible la valorisation du patrimoine de ressources géomatiques du département, il faut maintenant mettre en place et faire vivre une structure organisationnelle dont la fonction sera de mettre en place une organisation pour créer les métadonnées a posteriori.

En conséquence, nous proposons :

- que cette action soit labellisée comme projet prioritaire par le département et le centre,
- qu'une organisation composée de l'ingénieur de recherche travaillant sur la ZABR, d'un enseignant-chercheur expert en géomatique et moi-même mette en place une logique de création de métadonnées sur les données qui ont été produites en interne,
- que des contacts soient pris avec les fournisseurs de données acquises pour savoir s'il existe une possibilité d'importer les métadonnées correspondantes.

Conclusion et perspectives

L'objectif principal du projet était de faciliter le travail des enseignants-chercheurs pour la constitution de collections ou de lots de données spatiales et temporelles permettant de traiter des thématiques scientifiques dans le domaine de l'environnement. Pour atteindre cet objectif une gestion plus organisée du patrimoine des ressources géographiques du centre s'avère indispensable.

La première partie du mémoire qui était consacrée à l'analyse du contexte du projet a établi et souligné plusieurs aspects :

- la spécificité des activités d'acquisition, de création et de maintenance des données dans le cadre du travail d'enseignant-chercheur en sciences de l'environnement. Les informations mobilisées sont hétérogènes, multi-sources, multi-échelles, variables dans le temps et dans l'espace.
- l'organisation professionnelle est un facteur important pour établir un environnement informatique compatible avec les pratiques des enseignants-chercheurs.
- la nécessité de recourir à un outil de géocatalogage des données, alimenté par un modèle conceptuel approprié.

La seconde partie initiée par l'analyse des besoins des enseignants-chercheurs a permis de mettre en évidence les facteurs de risques et d'échecs du projet et de montrer qu'à partir d'un premier cadre de formalisation de ces besoins, un modèle conceptuel de métadonnées était possible à condition de l'ajuster en prenant en compte les systèmes normatifs existants. Ce type de projet correspond en effet à une préoccupation de plus en plus fréquente dans les laboratoires de recherche en environnement ou en géographie et il est indispensable de se référer à des standards en la matière.

La troisième partie s'est attachée à montrer que ce modèle conceptuel de métadonnées était implantable à travers un outil de géocatalogage convenablement choisi après une analyse critique de ceux existants sur le marché. L'ensemble constitue ainsi une solution possible pour mieux gérer et organiser les informations qui représentent ainsi le patrimoine des chercheurs et celles qui devront être acquises dans le futur. Nous avons montré que GeoNetwork était l'outil de géocatalogage propice pour implanter le modèle conceptuel de métadonnées.

Un test préliminaire a été réalisé à partir d'un jeu de métadonnées fragmentaire. Cette étape de mise en œuvre a montré que les enseignants-chercheurs doivent s'approprier l'outil GeoNetwork. Il est nécessaire de déployer une infrastructure de données spatiales complète comprenant notamment l'installation d'un serveur de données géographique permettant l'archivage et la sécurisation des ressources. Il s'agit de perspectives intéressantes en priorité les activités du centre de recherche en tant que projet à part entière.

Ces perspectives demeurent réalistes si nous considérons plusieurs aspects. Une analyse financière a montré que le coût était supportable par le centre de recherche. Les modifications organisationnelles et structurelles liées à l'implantation de ce système d'information sont possibles.

Dans ces perspectives un groupe référent capable d'assurer une gestion durable des métadonnées devra être mis en place.

Les perspectives à court terme de développement du projet concernent une phase d'appropriation du système plus approfondie : formation, évaluation du système de métadonnées, retour d'expérience, ajustement modification.

Quant à l'ouverture du système à un autre public que les enseignants-chercheurs du centre (partenaires académiques, partenaires institutionnels, étudiants), cette possibilité ne sera envisageable qu'après avoir clairement analysé le statut légal de ces nouvelles ressources.

Bibliographie

Ouvrages imprimés

AFNOR 2005 Information géographique : Métadonnées Norme européenne Norme Française NF EN ISO 19115 Août 2005 Indice de classement : Z 52-020 140 pages

AFNOR 2006 Information géographique : Services Norme européenne Norme Française NF EN ISO 19119 Décembre 2006 Indice de classement : Z 52-025 67 pages

Agence pour le développement de l'administration électronique – Ministère du budget et de la réforme de l'Etat – Information géographique : recommandation relative aux métadonnées – version 1.0 du 5/10/1006 – 57 pages

ECOBICHON Claude L'information géographique 1^{ère} édition Paris : Hermès, 1994, 122 pages, collection Perspectives

GeoNetwork opensource : The complete manual By the developpers version 2.4 Copyright 2007-2009 The Open Source Geospatial Foundation – 146 pages

Travaux universitaires

ENSMSE CEMAGREF Rapport de phase 1 de l'étude « Structuration des données de la Zabr » Mars 2003 – 105 pages

ENSMSE CEMAGREF Rapport de phase 1 et 2 de l'étude « Structuration des données de la Zabr » Mai 2004 – 81 pages

ENSMSE CEMAGREF Document de synthèse « Géorépertoire de la ZABR » : évaluation de la transférabilité des métadonnées du Géorépertoire vers un logiciel professionnel Septembre 2009 – 18 pages

ENSMSE Présentation de la politique générale, de la recherche et de la valorisation à l'Ecole Nationale Supérieure des Mines de Saint-Etienne - Campagne d'évaluation AERES 2010 Unités de Recherche : LCG, LPMG, LSTI, MOC, SEPIT - pages

ENSMSE Rapport d'activités 2009 - 17 pages

ENSMSE Rapport d'activités 2008 – 16 pages

Articles de périodiques imprimés

Fiche du CNIG n°106 – 2007 – 8 pages

Eric PIATYSZEK Denis BARBET Jean-François PERRIN Un géorépertoire : Système de métadonnées hydro-socio-économique pour la Zone du Bassin du Rhône et ses partenaires
5^{ème} édition des journées STIC & Environnement, Lyon 2007 – 9 pages

Sites web

Conseil National de l'Information Géographique <http://www.cnig.gouv.fr/> dernière consultation juillet 2010

ESRI France - Les solutions SIG et cartographiques <http://www.esrifrance.fr/> dernière consultation juillet 2010

ENSMSE : Ecole Nationale Supérieure des Mines de Saint-Etienne <http://www.emse.fr>
dernière consultation août 2010

GeoRezo : le portail francophone de la géomatique <http://georezo.net/> dernière consultation juillet 2010

Annexe 1
Formulaire de saisie des métadonnées au format Inspire via
Excel

Saisie des métadonnées d'une donnée géographique

Description

*Titre de la donnée ^{obligatoire}	<input type="text"/>
*Description de la donnée ^{obligatoire}	<input type="text"/>
*Identifiant de la donnée ^{obligatoire}	<input type="text"/>
*Thèmes ISO concernés ^{obligatoire}	(1) : <input type="text"/> (2) : <input type="text"/> (3) : <input type="text"/>
*Thème INSPIRE ^{obligatoire}	(1) : <input type="text"/>
*Extension géographique ^{obligatoire} : Rectangle de l'emprise des données en degrés décimaux (par défaut, France métropolitaine)	Lat N / S : <input type="text"/> <input type="text"/> Long O / E : <input type="text"/> <input type="text"/>
*Référence temporelle ^{obligatoire}	Date de la donnée (création) : <input type="text"/> j/mm/aaaa Date de la donnée (publication) : <input type="text"/> j/mm/aaaa Date de la donnée (révision) : <input type="text"/> j/mm/aaaa
*Généralité de la donnée ^{obligatoire} (données source, traitement,...)	<input type="text"/>

Organisme Contact(s)

*Contact sur les métadonnées ^{obligatoire}	Organisme : <input type="text"/> Adresse : <input type="text"/> Ville : <input type="text"/> Code postal : <input type="text"/> e-mail : <input type="text"/>
Responsable de la ressource ^{obligatoire}	Organisme : <input type="text"/> Adresse : <input type="text"/> Ville : <input type="text"/> Code postal : <input type="text"/> e-mail : <input type="text"/>

Contraintes d'accès

	*sur les métadonnées	*sur les données
--	----------------------	------------------

Mots-clés décrivant les données

<input type="text"/>

Accès aux données

	Adresse URL	Nom
Lien Internet ^{obligatoire}	<input type="text"/>	<input type="text"/>
Lien vers services OGC	<input type="text"/>	<input type="text"/>
Lien vers le téléchargement	<input type="text"/>	<input type="text"/>



* la saisie des informations marquées d'un astérisque est obligatoire

Définition des termes et modalités de saisie



Le formulaire permet de remplir les métadonnées décrivant une donnée géographique selon la norme internationale ISO 19115 (Métadonnées).
Ce formulaire indique uniquement les métadonnées "minimales" nécessaires pour inclure la description de ces données dans le Géocatalogue (www.geo catalogue.fr). Il est conseillé de compléter par des informations proposées par la norme permettant de constituer une vue réaliste et détaillée des données cartographiques que vous disposez.
L'asterisque rouge indique les informations à remplir obligatoirement.

Définition des termes	
Titre de la donnée :	Nom de la donnée, plus généralement de la ressource, décrite par les métadonnées. Ce titre doit être le plus explicite possible, notamment sur le contenu de la donnée, sa version... Ce titre doit être court (moins de 100 caractères est conseillé).
Description de la donnée :	Résumé explicatif du contenu de la donnée.
Thèmes concernés :	Thème principal du jeu de données selon une typologie définie au niveau international. Agriculture, Faune et flore, Frontières, Limites administratives, Climatologie, Météorologie, Atmosphère, Economie, Altitude, Environnement, Information géo-scientifique, Santé, Cartes de base, Installations militaires, Eaux continentales, Localisation, Culture, utilisation des sols, Sociétés, Constructions, Transport, Communication
Extension géographique :	L'empreinte géographique globale que couvre les données. Cette information est indiquée par les quatre coordonnées du rectangle qui contient vos données : Point Nord, Point sud, Point Ouest et Point Est. Chaque donnée est indiquée en degrés décimaux selon le système de projection WGS84 (Longitude, latitude).
Date de la donnée (création) :	La date de création de la donnée en JJ/MM/AAAA. Ne pas confondre avec la date de création des métadonnées.
Généalogie de la donnée (données sources, traitement,...) :	La généalogie de la donnée permet d'indiquer, en quelques lignes, l'origine et les modalités de production de la donnée.
Contact :	Le contact permet d'indiquer le ou les organismes assurant l'une des deux fonctions suivantes : - Point de contact et d'informations sur les métadonnées ; - Point de contact et d'information sur la donnée. Le contact est généralement le même mais il est possible d'indiquer deux organismes différents si nécessaire.
Contraintes d'accès :	Cette information permet d'indiquer le niveau de confidentialité des métadonnées (c'est-à-dire de la fiche elle-même) et des données. Les contraintes s'appuient sur une typologie internationale définie par l'ISO.
Mots-clés :	Ensemble de mots-clés décrivant votre donnée. Chaque mot-clé est séparé par un ";". On privilégie des mots-clés relatifs au contenu des données (thématique, thème,...).
Accès aux données :	Cette rubrique permet d'indiquer d'éventuels liens Internet en lien avec les données décrites. Il s'agit notamment : - d'un site Internet permettant d'accéder aux données ou à leur visualisation ; - d'une adresse permettant un accès direct vers la visualisation de vos données au travers des services web de l'OGC - d'une adresse permettant de télécharger directement la ressource (le fichier shape de la couche cartographique,...)

 Métadonnées complémentaires relatives à la donnée 	
Informations sur la métadonnée	
*Identifiant de la métadonnée :	<input type="text"/>
Langue décrivent les métadonnées ^{INSPIRE} :	<input type="text"/>
Type de ressource ^{INSPIRE} :	<input type="text"/>
*Limite d'utilisation des métadonnées :	<input type="text"/>
Informations complémentaires sur la ressource	
Langue décrivent les données ^{INSPIRE} :	<input type="text"/>
Jeu de caractères des données :	<input type="text"/>
Type de représentation spatiale :	<input type="text"/>
Principal système de projection :	<input type="text"/>
Limites d'utilisation des données ^{INSPIRE} :	<input type="text"/>
Résolution spatiale ^{INSPIRE}	
Résolution indiquée en échelle : 1/ <input type="text"/>	
OU Résolution indiquée en taille de pixels (mètres) : <input type="text"/>	
Autres organismes	
Producteur des données	
Organisme :	<input type="text"/>
Adresse :	<input type="text"/>
Ville :	<input type="text"/>
Code postal :	<input type="text"/>
e-mail :	<input type="text"/>
Conformité (INSPIRE)	
Titre de la spécification ^{INSPIRE} :	<input type="text"/>
Date de publication ^{INSPIRE} :	<input type="text"/> /mm/aaaa
La ressource est-elle conforme à la spécification? ^{INSPIRE} :	<input type="text"/>

Annexe 2

Copies d'écran des interfaces graphiques d'édition des métadonnées de ArcCatalog

The screenshot shows the 'Editing '10.tif' dialog box with the 'Data Quality' tab selected. The 'General' sub-tab is active, displaying three sections: 'Logical Consistency Report' and 'Completeness Report' (both with empty text boxes and vertical scrollbars), and 'Cloud Cover' (with an empty text box). The bottom of the dialog features 'Save', 'Cancel', and 'Help' buttons.

The screenshot shows the 'Editing '10.tif' dialog box with the 'Data Organization' tab selected. The 'General' sub-tab is active, displaying spatial reference and raster properties. The 'Direct Spatial Reference Method' is set to 'Raster'. The 'Indirect Spatial Reference' is empty. Below, a grid of fields includes: 'Raster Object Type' (Pixel), 'Raster Display Type' (pixel codes), 'Row Count' (3100), 'Column Count' (4622), 'Vertical Count' (1), 'Raster Origin' (Upper Left), 'Number of Bands' (3), 'Cell Size X Direction' (0.003333), 'Cell Size X Units' (empty), 'Cell Size Y Direction' (0.003333), 'Cell Size Y Units' (empty), 'Bits Per Pixel' (8), 'Background Nodata Value' (empty), 'Pyramid Layers' (TRUE), 'Image Format' (TIFF), 'Image Type' (empty), 'Image Colormap' (FALSE), and 'Compression Type' (None). The bottom of the dialog features 'Save', 'Cancel', and 'Help' buttons.

Editing '10.tif'

Identification Data Quality Data Organization **Spatial Reference** Entity Attribute Distribution Metadata Reference

General Horizontal Coordinate System Vertical Coordinate System

Geographic Coordinate System Name:

Projected Coordinate System Name:

Horizontal Datum Name:

Ellipsoid Name:

Semi-major Axis:

Denominator of Flattening Ratio:

Save Cancel Help

Editing '10.tif'

Identification Data Quality Data Organization Spatial Reference **Entity Attribute** Distribution Metadata Reference

Detailed Description Overview Description

Entity Type Attribute

Label:

Type:

Count:

Definition:

Definition Source:

+ x y z Detailed Description (+) of 0

Save Cancel Help

Editing '10.tif'

Identification Data Quality Data Organization Spatial Reference Entity Attribute **Distribution** Metadata Reference

General | Distributor | Standard Order Process | Available Time Period

Resource Description: Downloadable Data

Distribution Liability:

Custom Order Process:

Technical Prerequisites:

+ X < > Distribution Information 1 of 1

Save Cancel Help

Editing '10.tif'

Identification Data Quality Data Organization Spatial Reference Entity Attribute Distribution **Metadata Reference**

General | Extensions

Metadata Date: 20100816 Metadata Standard Name: FGDC Content Standards for Digital Geospatial I

Metadata Review Date: Metadata Standard Version: FGDC-STD-001-1998

Metadata Future Review Date: Metadata Time Convention: local time

Language of Metadata: fr

Contact: REQUIRED: The organization responsible for Details...

Metadata Access Constraints:

Metadata Use Constraints:

Metadata Security Information

Classification System:

Classification:

Handling Information:

Save Cancel Help

Annexe 3

Copies d'écran des interfaces graphiques des rubriques de saisie des métadonnées sous GeoNetwork, vue avancée

Rubrique système de référence

The screenshot shows the 'Système de référence' (Reference System) metadata entry form in GeoNetwork. The form is divided into several sections:

- Information sur le système de référence**: This section contains fields for 'Nom du système de référence' (Reference system name), 'Autorité' (Authority), 'Date de la dernière actualisation' (Last update date), 'Date d'édition' (Edition date), 'Identificateur de provenance' (Provenance identifier), 'Forme de la présentation' (Presentation form), 'Séries' (Series), 'Autres informations de référence' (Other reference information), 'Code' (Code), 'Nom de l'identifiant' (Identifier name), and 'Version' (Version).
- Options de transfert digital**: This section contains fields for 'Unité de distribution' (Distribution unit), 'Taille transfert' (Transfer size), and 'Séries on-line' (Online series).
- Sources on-line**: This section contains three sub-sections for 'Ressource en ligne' (Online resource), each with fields for 'Adresse Internet' (Internet address), 'Protocole' (Protocol), 'Profil d'application' (Application profile), 'Description' (Description), and 'Fonction' (Function).

The form is displayed in a web browser window with the URL 'http://localhost:8080/geonetwork/srv/en/metadata.update'. The browser's address bar shows the URL, and the page title is 'GeoNetwork - Portail p...'. The browser's status bar shows the zoom level as 100%.

Rubrique distribution

The screenshot shows the 'Distribution' metadata entry form in GeoNetwork. The form is divided into several sections:

- Informations sur la distribution**: This section contains fields for 'Format de distribution' (Distribution format), 'Distributeur' (Distributor), 'Options de transfert digital' (Digital transfer options), 'Unité de distribution' (Distribution unit), 'Taille transfert' (Transfer size), and 'Séries on-line' (Online series).
- Sources on-line**: This section contains three sub-sections for 'Ressource en ligne' (Online resource), each with fields for 'Adresse Internet' (Internet address), 'Protocole' (Protocol), 'Profil d'application' (Application profile), 'Description' (Description), and 'Fonction' (Function).
- Sources off-line**: This section contains fields for 'Adresse Internet' (Internet address), 'Protocole' (Protocol), 'Profil d'application' (Application profile), 'Description' (Description), and 'Fonction' (Function).

The form is displayed in a web browser window with the URL 'http://localhost:8080/geonetwork/srv/en/metadata.update'. The browser's address bar shows the URL, and the page title is 'GeoNetwork - Portail p...'. The browser's status bar shows the zoom level as 100%.

Conception et élaboration d'un système d'information pour la gestion des ressources en géomatique

Mémoire d'Ingénieur C.N.A.M., Saint-Etienne

RESUME

La gestion du patrimoine de données environnementales du centre d'enseignement et de recherche SITE de l'ENSMSE est une préoccupation actuelle. Le partage et la pérennisation de cette information essentiellement géographique et spatiale ne peuvent pas être mis en place sans disposer de l'information sur les données.

Dans ce travail il s'agit de réaliser un modèle conceptuel de métadonnées qui prenne en compte les thématiques de recherche environnementales, les besoins des utilisateurs et qui puisse être durablement exploité. La structure de ce modèle est présentée, elle prend en compte les normes en vigueur existantes dans le domaine. L'outil d'implantation finalement choisi est de type géocatalogue : GeoNetwork.

Les contraintes et les limites d'utilisation de l'outil sont examinées dans le but d'identifier les facteurs de réussite du projet.

Mots clés : géomatique, métadonnées, géocatalogue, données géographiques, Système d'Information Géographique.

SUMMARY

The management of environmental data in the education and research center SITE of the ENSM-SE is an ongoing concern. The sharing and perpetuation of this information, which is essentially geographical and spatial in nature, cannot be implemented without disposing information on the data itself.

In this work, a conceptual metadata model was realized. Environmental research issues and user needs are taken into account in this model, which can be sustainably operated. The structure of this model, integrating existing standards in the field, is presented in this document. A geocatalog tool (GeoNetwork) was finally chosen for the implementation of this task.

The constraints and limitations in the use of this tool are analysed in order to identify success factors for this project.

Key words : geomatics, metadata, geocatalog, geographic data, geographic information system.