



HAL
open science

Enrichissement et amélioration d'une base documentaire dédiée aux études environnementales de Bretagne

Benjamin Duhamel

► **To cite this version:**

Benjamin Duhamel. Enrichissement et amélioration d'une base documentaire dédiée aux études environnementales de Bretagne. domain_shs.info.docu. 2013. mem_00840171

HAL Id: mem_00840171

https://memsic.ccsd.cnrs.fr/mem_00840171

Submitted on 1 Jul 2013

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.

Benjamin DUHAMEL

MASTER1, MENTION ICD

(Option : Sciences de l'Information et du Document)

MEMOIRE DE STAGE

Mission effectuée du 22 avril au 31 Juillet 2013.

au

GroupeMENT d'intérêt public Bretagne environnement

6-A, rue du Bignon
35000 Rennes

**Enrichissement et amélioration d'une base documentaire
dédiée aux études environnementales de Bretagne**

Sous la direction de :

Mr B. JACQUEMIN (resp. universitaire)

Mme. C. LATHIERE (tuteur professionnel)

Soutenu le 24 Juin 2013 à l'UFR DECCID-SID

Université Charles de Gaulle, Lille 3 (Campus Pont de Bois)

BP 60 149, 59 653 Villeneuve d'Ascq Cedex

Année Universitaire 2012/2013

Mes sincères remerciements pour l'accueil et l'expérience vécue au sein du GIP BE.

*A ma tutrice, le responsable et toute l'équipe
ainsi qu'aux autres stagiaires ayant partagé le temps avec moi.*

*Mes remerciements également pour mon tuteur universitaire pour son aide
ainsi qu'aux enseignants de Lille 3 que j'ai sollicité et qui m'ont aidé également.*

Introduction.....	5
1 Le géocatalogue des études sur l’environnement en Bretagne : cadrage de l’outil dans son environnement directe et plus étendu	8
1.1 Le Groupement d’Intérêt Public Bretagne Environnement : le projet, les missions et les outils	8
1.2 Un environnement global politique, juridique, sociétal et technique	10
1.2.1 L’accès du public à l’information en matière d’environnement	10
1.2.2 L’ouverture en ligne des données publiques et la transparence de l’action publique (<i>Open source governance</i>) ainsi que la notion de donnée publique	11
1.2.3 Les données ouvertes (<i>Open data</i>) : retour terminologique	12
1.2.4 Le libre accès (<i>Open access</i>) à l’Information Scientifique et Technique	12
1.2.5 Les archives ouvertes (<i>Open archive</i>) et le protocole OAI-PMH	14
1.3 Le Géocatalogue des études sur l’environnement en Bretagne	14
1.3.1 Présentation	14
1.3.2 Constats : défaut d’accès aux données environnementales et déficit de gestion documentaire	15
1.3.3 Missions : co-construction d’une base documentaire pour un état des connaissances, référencement et propositions d’outils, normes et procédures documentaires	15
2 Le géocatalogue des études sur l’environnement en Bretagne : des services et un fonctionnement complexe	16
2.1 Le géocatalogue : fonctionnalités et services	16
2.2 Le géocatalogue dans le miroir du processus itératif visant à satisfaire un besoin d’information (besoins – sources – collecte – traitement – diffusion)	17
2.2.1 Définition des besoins : thématiques et typologie des documents	19
2.2.2 Sources d’informations	21
2.2.2.1 Les acteurs de l’environnement en Bretagne	21
2.2.2.2 D’autres sources exploitables ?	22
2.2.3 Collecte de l’information	26
2.2.3.1 Aspect collaboratif	26
2.2.3.2 Trois méthodes de collecte	27
2.2.3.3 La problématique de l’actualisation	29
2.2.4 Traitement de l’information	29
2.2.4.1 Jeu de métadonnées Dublin Core et matrice Excel	29
2.2.4.2 Indexation : thésaurus et matrices	30
2.2.4.3 Le géo-référencement	30

2.2.4.4 Méthodologie de traitement d'un fichier OAI volumétrique (plus de trois mille références)	31
2.2.5 Diffusion de l'information (et export)	33
3 Le géocatalogue des études sur l'environnement en Bretagne : gestion de grosse volumétrie, automatisation, rigueur et qualité de catalogage, l' « information durable » au cœur d'un équilibre difficile à trouver	34
3.1 Un équilibre à trouver entre les procédures automatiques et la rigueur ainsi que la qualité du catalogage	34
3.2 Croisement et ouverture avec le concept d'information durable	35
Conclusion	37
Bibliographie et Webographie	38
Sigles	39
Sommaire des annexes	40

Introduction

Mon stage s'inscrit dans le développement du projet du géocatalogue des études expertes sur la connaissance environnementale en Bretagne. Ce projet mené par le Groupement d'Intérêt Public Bretagne Environnement se place dans la continuité des objectifs qu'il poursuit, c'est-à-dire d'accès à la connaissance et d'aide à la compréhension des problématiques régionales. Il correspond alors à la volonté de diffusion d'un centre de documentation sur internet. C'est à l'approche de son ouverture officielle au public que mon stage s'est dessiné, répondant aux besoins du projet d'enrichissement de la base de données en volume (catalogage, indexation et géo-référencement des documents) ainsi que d'amélioration de la cartographie des partenaires potentiels afin d'accroître le réseau des contributeurs.

A travers la réalisation de ces missions ainsi que de la découverte de l'intérieur du fonctionnement d'un tel projet, j'ai pu prendre conscience de la multitude de questions que pouvaient sous-entendre la création et le fonctionnement d'un catalogue relai comme le géocatalogue des études. Qui sont les différents partenaires ? Comment travaillent-ils au niveau de leur documentation ? Quels types de partenaires pourraient devenir contributeurs ? Comment créer des partenariats efficaces ? Sur quels fondements reposent les partenariats (convention d'Aarhus) ou les différentes initiatives dans le domaine du partage de l'information scientifique (ex : les archives ouvertes) ? Avec quels formats de description et quels protocoles documentaires va-t-on opérer, notamment par rapport à des partenaires aux pratiques documentaires hétérogènes ? Où sont les limites du corpus éligible au projet ? Comment gérer de grosses volumétries et ainsi cataloguer des milliers de notices issues d'un flux OAI ? Comment fonctionne le géo-référencement et qu'entend-on par un « état des connaissances » grâce à la cartographie dynamique ? Autant de questions qui m'ont amené à poser la problématique suivante : **Le géocatalogue des études sur l'environnement en Bretagne, un projet s'inscrivant dans un environnement documentaire, politique et technologique aussi riche que complexe.**

La méthodologie que j'ai employée pour répondre à cette problématique est la suivante. Dans un premier temps, j'ai réalisé un travail de cadrage sur la structure d'accueil, le GIP Bretagne environnement, sur sa raison d'être et les politiques sur lesquelles elle s'appuie, notamment la convention d'Aarhus et la directive 2003/4/CE

du 28 janvier 2003. Ce travail m'a conduit à m'interroger et à explorer au-delà du monde de l'environnement, les initiatives et politiques à l'œuvre aujourd'hui dans le domaine de l'information et de la connaissance, de sa diffusion, de son accès et de son partage. Ceci m'a permis de faire un point terminologique sur l'ouverture en ligne des données publiques, les données ouvertes et le libre accès. Autant de concepts contemporains mais ne relevant pas forcément des mêmes choses. Face à un paysage qui m'était alors confus, il m'est apparu intéressant de préciser ces notions afin de mieux en repérer les enjeux et placer le projet du géocatalogue des études dans un environnement plus global. Ainsi nous exposerons ces éléments pour que chacun puisse lui-même observer là où il pourrait y avoir des passerelles et là où il n'y en a pas. La fin du cadrage consiste en la présentation de la genèse du géocatalogue, celui-ci répondant au constat fait du défaut d'accès aux données environnementales ainsi que de l'absence d'un outil centralisateur sur ces thématiques en Bretagne.

Ce cadrage fait, la deuxième partie du travail est consacrée à un raisonnement sur les modalités concrètes de fonctionnement du projet. Est mis en avant dans un premier temps l'aspect service ainsi que les fonctionnalités délivrées par le projet au-delà même du simple catalogue mettant à disposition des références et documents. Cette dimension de service étant essentielle aujourd'hui, il était important de commencer par là. Ensuite j'ai fait le choix méthodologique de présenter le fonctionnement d'activité du géocatalogue sous l'angle du processus itératif répondant à un besoin d'information, processus souvent repris notamment pour expliquer le cycle de la veille. Je précise bien ici qu'il n'est pas question de parler de veille mais simplement de s'appuyer sur un processus simple permettant de séquencer les différentes étapes d'un projet de référencement documentaire. Ainsi les étapes de définition des besoins, de *sourcing*, de collecte, de traitement et de diffusion de l'information seront étudiées au prisme du projet du géocatalogue des études. Chaque étape permettra d'aborder un point de la complexité de l'organisation du projet. Les réalisations de mes missions seront alors clairement explicitées à l'intérieur même du processus. Naturellement, elles trouveront leur place dans les étapes source, collecte et surtout traitement de l'information.

La troisième et dernière partie sera l'occasion d'étudier plus profondément deux enjeux majeurs apparus durant l'étude : d'une part la tension et le difficile équilibre à trouver entre la gestion de grosses volumétries ainsi que de contributeurs aux pratiques documentaires variées, la nécessité d'automatiser les procédures et la rigueur et l'exigence de qualité catalographique nécessaires à avoir pour un géocatalogue pertinent. D'autre part, je reviendrai sur le concept « d'information durable » qui sous-entend une gestion responsabilisée, raisonnée et collective de l'information. Ce concept est à mon sens tout à fait en adéquation avec le projet.

1 Le géocatalogue des études sur l'environnement en Bretagne : cadrage de l'outil dans son environnement direct et plus étendu

1.1 Le Groupement d'Intérêt Public Bretagne Environnement : le projet, les missions et les outils

Le Groupement d'Intérêt Public Bretagne Environnement (GIP BE) a été créé par l'Etat et le Conseil Régional de Bretagne en 2007¹. Financé par la préfecture, la région de Bretagne, cofinancé par l'Union Européenne, et assurant une mission de service public, son but est de permettre à chacun de trouver les renseignements qu'il recherche sur l'environnement en Bretagne, afin de développer ses connaissances et d'être aidé dans ses prises de décision.

Ainsi le GIP BE rassemble, valorise et assure la diffusion des données et connaissances sur la thématique environnementale en Bretagne en s'appuyant sur des partenariats avec les producteurs de données ainsi qu'avec les organismes experts. Outre la facilitation de l'accès aux données et le développement de la culture environnementale en Bretagne, le GIP BE mène des actions d'observations régionales, conduisant au développement d'outils d'aide à la décision en cohérence avec les politiques publiques régionales ou locales. L'action du GIP BE se positionne donc aussi bien à destination du grand public que des publics experts et décideurs.

Les missions confiées au GIP BE afin de répondre aux objectifs identifiés sont les suivantes : diffusion d'un centre de documentation sur Internet, diffusion d'informations sur les acteurs et leurs actions, diffusion de synthèses sur les connaissances, diffusion d'informations sur les territoires, vulgarisation des connaissances, gestion de la connaissance, développement d'éléments de stratégie et de prospective, amélioration de l'organisation régionale des flux de données, amélioration de l'échange d'informations.

L'activité du GIP BE s'inscrit dans un cadre juridique précis, portant sur les obligations concernant l'accès du public à l'information en matière d'environnement, défini par la Convention d'Aarhus et la directive 2003/4/CE du 28 janvier 2003.

¹Arrêté de création du GIP pour la période 2007-2012 : <http://www.bretagne-environnement.org/Media/Fichiers/Arrete-portant-approbation-de-la-convention-constitutive-du-groupement-d-interet-public-Bretagne-environnement>

Nous pouvons mentionner que l'actualité a vu le GIP BE renouveler son activité pour les huit prochaines années. En effet, en janvier 2013, le groupement d'intérêt public a été prorogé jusqu'à 2020².

Le réseau est également un élément central du projet du GIP BE qui se veut collaboratif. Ainsi, la structure entretient un réseau de partenaires et se place au cœur des échanges avec les services de l'Etat, les associations, les établissements publics et les collectivités. Elle assure ainsi la cohésion du régional vers le local ou vers le national.

Au niveau des outils mis en place par le GIP BE, nous pouvons mentionner en premier lieu le portail³, qui est également la porte d'entrée pour tous les autres outils (sites web des observatoires, ma commune, cartographie dynamique et géocatalogue des études). Il relaie quotidiennement une information pluri thématique dédiée à l'environnement en Bretagne issue d'acteurs très variés allant du secteur public au monde associatif en passant par le domaine universitaire. Ces informations sont structurées selon une logique d'accès soit par typologie de documents : articles, données, brèves, atlas, documentation, agenda ; soit par thématiques : eau, air et climat, énergie, déchets, risques, sols, sous-sol, mer-littoral, patrimoine naturel, territoire-activités. Le portail tient également à jour un annuaire des acteurs, une lettre d'information, des rubriques illustrations. Dans son ensemble la variété et la richesse de ses informations en font un véritable écosystème informationnel complet et ouvert. L'esprit de vulgarisation et de démocratisation de l'accès aux données, informations, événements et connaissances est fort présent sur le portail dans un souci de développement de la culture environnementale.

Les autres outils sont davantage tournés vers des publics initiés, experts ou décideurs, permettant à travers des outils interactifs⁴, d'accéder à des données à l'échelle d'un territoire spécifique (ex : commune, département, EPCI, SAGE, bassin versant, etc...). Les observatoires, au nombre de quatre : eau, énergie et gaz à effet de serre, biodiversité et patrimoine naturel, déchets, ont pour vocation d'être des

² Arrêté de renouvellement du GIP pour la période 2013-2020 : <http://www.bretagne-environnement.org/Media/Fichiers/Arrete-du-8-janvier-2013-portant-approbation-du-renouvellement-du-GIP-Bretagne-Environnement>

³ Portail de l'information environnementale en Bretagne : <http://www.bretagne-environnement.org/>

⁴ Ma commune, mon environnement : L'accès aux données locales de Bretagne (<http://communes.bretagne-environnement.org/>) et la Cartographie dynamique des données environnementales en Bretagne (<http://cartographie.bretagne-environnement.org/>)

plateformes de connaissances sur des thématiques précises. Ils diffusent et valorisent les données et informations, à travers notamment des indicateurs, des connaissances et des bilans facilitant la prise de décision.

Dernier outil majeur sur lequel nous mettrons un coup de projecteur dans ce travail, et qui du reste constituera la trame essentielle de celui-ci, le géocatalogue des études⁵. C'est sur l'alimentation et le développement de cet outil que j'ai réalisé mon stage a pris source.

1.2 Un environnement global politique, juridique, sociétal et technique, riche

1.2.1 L'accès du public à l'information en matière d'environnement

Les missions du GIP Bretagne Environnement et par suite celles du géocatalogue des études, découlent directement des obligations liées à la Convention d'Aarhus et à la Directive 2003/4/CE du 28 Janvier concernant l'accès du public à l'information en matière d'environnement⁶.

Le portail de l'Etat sur l'information publique environnementale décrit la convention d'Aarhus comme un pilier de la démocratie environnementale et la synthétise en ces mots : « *Toute personne a le droit d'être informée, de s'impliquer dans les décisions et d'exercer des recours en matière d'environnement. (...) Ce texte essentiel contribue à créer la confiance du citoyen envers ses institutions et, plus largement, leur fonctionnement démocratique. En offrant au citoyen une place dans les débats environnementaux, elle rencontre les exigences de transparence et de proximité, synonymes de bonne gouvernance publique.* »⁷. La convention a pour objet l'accès à l'information, la participation du public au processus décisionnel ainsi que l'accès à la justice en matière d'environnement⁸.

Nous allons inscrire dans les points suivant le projet du géocatalogue des études conduit par le GIP BE, au-delà de la dimension d'accès du public à l'information en matière d'environnement, dans le panorama actuel vaste caractérisé

⁵ Géocatalogue des études : <http://etudes.bretagne-environnement.org/>

⁶ Journal Officiel de l'Union européenne : <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:041:0026:0026:FR:PDF>

⁷ Portail *Tout sur l'environnement* : <http://www.toutsurlenvironnement.fr/aarhus/la-convention-daarhus-pilier-de-la-democratie-environnementale>

⁸ Convention d'Aarhus : <http://www.unece.org/fileadmin/DAM/env/pp/documents/cep43f.pdf>

par le mouvement de l'ouverture et l'accès aux données. Voyons donc précisément ce qu'il en est et les liens qui peuvent être faits avec le géocatalogue des études.

1.2.2 L'ouverture en ligne des données publiques et la transparence de l'action publique (*Open source governance*) ainsi que la notion de donnée publique

L'accès du public à l'information en matière d'environnement peut se voir également dans la perspective de la politique nationale d'ouverture des données publiques. Ce projet, piloté par la mission Etalab⁹, s'inscrit dans le cadre de la modernisation de l'action publique et a pour objectif la mise à disposition gratuite des données publiques, conformément au principe général de réutilisation libre, facile et gratuite. Ceci est conduit dans l'optique de la transparence de l'action publique comme stipulé dans la Circulaire du Premier ministre du 26 mai 2011 relative à l'*Open Data*¹⁰.

Bien que ce terme d'Open data soit fréquemment utilisé par les initiatives du gouvernement français, nous parlerons dans ce travail plutôt d'*Open Source Gouvernance*, qui à l'image de le célèbre *Open Government Initiative* de l'administration Américaine « *vise à créer un niveau sans précédent de transparence et d'ouverture du gouvernement [et] se situe dans une tendance émergente (...) de l'Open source governance, qui prône l'application en politique et dans la gouvernance des démocraties de philosophies telles que l'open source et les contenus ouverts (Open data), pour permettre à tout citoyen intéressé de contribuer à créer les contenus de la Politique, et pour permettre aux gouvernements de mieux bénéficier des savoirs et savoir-faire locaux.* »¹¹. Nous reviendrons dans le point suivant sur la notion d'*Open data*.

Au sujet de la notion de donnée publique, notions que celle-ci couvre l'ensemble des données qui sont ou devraient être publiées ou tenues à disposition du public, et qui sont produites ou collectées par un État, une collectivité territoriale, un organe parapublic, dans le cadre de leurs activités de service public¹². En France, c'est la Commission d'Accès aux Documents Administratifs (CADA) qui est l'autorité

⁹ Site Etalab : <http://www.etalab.gouv.fr/>

¹⁰ Etalab. *La feuille de route du Gouvernement en matière d'ouverture et de partage des données publiques* : <http://www.etalab.gouv.fr/article-la-feuille-de-route-du-gouvernement-en-matiere-d-ouverture-et-de-partage-des-donnees-publiques-115767801.html>

¹¹ Wikipédia. *Open Government Initiative* : http://fr.wikipedia.org/wiki/Open_Government_Initiative

¹² Wikipédia. *Donnée publique* : http://fr.wikipedia.org/wiki/Donn%C3%A9e_publique

administrative indépendante et consultative chargée de veiller à la liberté d'accès aux documents administratifs.

Une plateforme unique interministérielle d'ouverture des données public¹³ est administrée par *Etalab*, afin de rassembler et de mettre à disposition librement l'ensemble des informations publiques de l'Etat, de ses établissements publics et, si elles le souhaitent, des collectivités territoriales et des personnes de droit public ou de droit privé chargées d'une mission de service public.

1.2.3 Les données ouvertes (*Open data*) : retour terminologique

Si l'*open data* (donnée ouverte) est bien souvent assimilée à l'ouverture des données publiques, il est néanmoins juste de préciser que cette dernière n'est qu'une forme de l'*open data*. En effet, comme le rappelle l'association *liberTIC*, l'*Open data* est à la base une démarche plus globale et neutre qui vise à la « *publication de données numériques en ligne selon une série de dix critères garantissant leur libre accès et réutilisation par tous* »¹⁴. Les dix critères font d'une donnée brute, une donnée dite « ouverte » si sa publication est complète, primaire, opportune, accessible, exploitable, non-discriminatoire, non-proprétaire, libre de droits, permanente et gratuite.

Il me semble donc opportun de bien préciser qu'une donnée ouverte peut être publique ou privée et que son caractère ouvert vient de la manière par laquelle elle est publiée. Cette dernière, structurée, reprend « *une méthodologie qui garantit son libre accès et sa réutilisation par tous, sans restriction technique, juridique ou financière* »¹⁵.

1.2.4 Le libre accès (*Open access*) à l'Information Scientifique et Technique

Le mouvement du libre accès vient du constat et de la réaction face à la difficulté d'accès aux résultats scientifiques. De la *Lettre ouverte de la Public Library of Science* (PLoS, 2001)¹⁶ point de départ du mouvement, puis *l'Initiative de Budapest pour l'Accès Ouvert* (BOAI, 2002)¹⁷ qui recommande les voies verte et dorée ainsi que l'auto-archivage, la *déclaration de Bethesda pour l'édition en libre*

¹³Plateforme française d'ouverture des données publiques françaises : <http://www.data.gouv.fr/>

¹⁴ liberTIC, « l'open data ». In *Actions Open data* : <http://actionsopendata.org/l-open-data/definition/>

¹⁵ Wikipédia. *Open data* : http://fr.wikipedia.org/wiki/Open_data

¹⁶Traduction (INIST) de la *Lettre ouverte de Public Library of Science* <http://openaccess.inist.fr/?La-lettre-ouverte-de-Public>

¹⁷ Traduction (INIST) de *l'Initiative de Budapest pour l'Accès Ouvert* : <http://openaccess.inist.fr/?Initiative-de-Budapest-pour-l>

accès (2003)¹⁸, et enfin la *déclaration de Berlin* (2003)¹⁹ qui élargit la notion du Libre Accès à tous les biens culturels et à toutes les données issues d'un travail de recherche.

Pour synthétiser le mouvement du libre accès, l'enjeu principal est l'ouverture au plus grand nombre de l'accès aux résultats de la recherche scientifique, ce dans la logique même du mécanisme scientifique ainsi que dans sa démocratisation, qui tous deux nécessitent le partage. Cela conduit alors à la volonté de se libérer des freins à son accès que peuvent représenter le coût important de l'accès à l'Information Scientifique et Technique (IST) à travers notamment le système traditionnel de la revue et de l'édition scientifique.

En effet, la recherche étant déjà financée la plupart du temps par des fonds publics, pourquoi les œuvres de l'esprit ainsi créées sont-elles alors cédées dans leurs droits patrimoniaux à des éditeurs qui monnayent à coût élevé leur accès ? Francis André, directeur adjoint de la Direction de l'Information Scientifique et Technique (DIST) du CNRS, expose que c'est le cheminement de cette réflexion qui a conduit les chercheurs à vouloir se réapproprier leur production scientifique dans la perspective de l'intérêt général et de la stimulation de la création scientifique.²⁰

Ce même auteur nous expose également comment la revue scientifique a permis, outre l'instauration de la validation scientifique par les pairs à travers les comités de lecture, d'instaurer l' « *économie de la réputation* » que ce soit au niveau des revues elles-mêmes, qu'au niveau des chercheurs. Aussi, les revues scientifiques vont-elles jouer un rôle essentiel dans « *la constitution d'une science officielle, validée, identifiée, véritable capital scientifique de référence* ». Ainsi la validation par les pairs, mais aussi les métriques, la notoriété et les calculs de citations vont constituer le socle de fonctionnement du système de l'édition scientifique.²¹

Le *libre accès* va-t-il faire perdre en termes de validation ce qu'il a fait gagner en liberté et en partage ? Va-t-il remettre en cause une économie de la réputation

¹⁸ Traduction (INIST) de la *déclaration de Bethesda pour l'édition en libre accès* :

<http://openaccess.inist.fr/?Declaration-de-Bethesda-pour-l>

¹⁹ Traduction (INIST) de la *déclaration de Berlin sur le Libre Accès à la Connaissance en sciences exactes, sciences de la vie, sciences humaines et sociales* : <http://openaccess.inist.fr/?Declaration-de-Berlin-sur-le-Libre>

²⁰ André F. *Libre accès aux savoirs*.

²¹ Idem

instaurée chez les scientifiques au profit d'une science désintéressée ? L'édition scientifique est-elle en péril ou se dirige-t-on vers une vision de l'information de type « durable », prenant ainsi en compte plus de partage, d'équité, d'innovation et de responsabilité²² ? De nouvelles formes d'intermédiation sont-elles en gestation ? Autant de questions à débattre et sur lesquelles nous pouvons également appréhender le projet du géocatalogue du GIP BE.

1.2.5 Les archives ouvertes (*Open archive*) et le protocole OAI-PMH

Une archive ouverte est un site, aussi appelé entrepôt, où les documents déposés par les chercheurs sont en libre accès. On y trouve les références bibliographiques des articles et lorsqu'ils y ont été associés, leurs textes intégraux. Les archives ouvertes suivent le protocole d'échange des données OAI-PMH, permettant le moissonnage sur le réseau Internet au travers d'un client via une requête.

L'archive ouverte pluridisciplinaire HAL²³, destinée au dépôt et à la diffusion d'articles scientifiques de niveau recherche émanant des établissements d'enseignement et de recherche français ou étrangers ainsi que des laboratoires publics ou privés, est une référence aujourd'hui en terme d'archive institutionnelle. Elle a d'ailleurs, nous le reverrons plus tard, fait l'objet d'une moisson par le GIP BE dans la perspective de l'alimentation du géocatalogue des études.

1.3 Le Géocatalogue des études sur l'environnement en Bretagne

1.3.1 Présentation

Le géocatalogue des études est un outil de référencement et de géolocalisation des études environnementales menées sur le territoire breton. Il s'agit d'une base documentaire experte complémentaire de la documentation « grand public » disponible sur le portail. Si la géolocalisation permet d'offrir un état des connaissances par type de territoire (*exemple : le nombre d'études réalisées par communes*, voir **annexe I**), l'interface offre l'autre avantage de centraliser et de valoriser l'ensemble des études produites sur la thématique.

²² Bourret Christian *et al.*, « Capital immatériel et information professionnelle. L'émergence d'un concept nouveau : l'information durable », *Documentaliste-Sciences de l'Information*

²³ Entrepôt HAL : <http://hal.archives-ouvertes.fr/>

Le géocatalogue a été ouvert et mis en ligne officiellement au public le 10 Juin 2013 à l'occasion d'une journée de communication du GIP BE à Dinan.

1.3.2 Constats : défaut d'accès aux données environnementales et déficit de gestion documentaire

Le géocatalogue est né du constat qu'il y avait un défaut d'accès aux données environnementales. En effet de nombreuses études sont réalisées chaque année sur le territoire breton et faute d'outil de centralisateur, ces études sont dispersées et éparpillées sur le Web (pages de blogs, sites web, bases de données). Ces contenus experts, de type littérature grise, proviennent de plus d'une centaine de structures identifiées par le GIP BE. Ces structures, si elles ont pris le train en marche de la diffusion sur le Web, ne disposent pas par ailleurs toutes ou à des degrés de développement disparates, de gestion documentaire optimisée de leurs contenus. L'identification, le partage et la structuration de leurs données (au sens documentaire) apparaît alors comme un enjeu essentiel aujourd'hui.

1.3.3 Missions : co-construction d'une base documentaire pour un état des connaissances, référencement et propositions d'outils, normes et procédures documentaires

Le géocatalogue du GIP BE se positionne donc comme un acteur relai des connaissances et administre le référencement de cette documentation de manière responsabilisée en privilégiant la co-construction de la base de données avec la participation active des contributeurs. Ainsi il est proposé aux partenaires, ce que l'on pourrait appeler un « écosystème documentaire », leur offrant une « fenêtre » au sein d'un catalogue et d'une collection originale mais aussi une normalisation de description des représentations bibliographiques : un format *Dublin core* adapté à certaines spécificités, et enfin des procédures d'intégration des notices adaptées aux différents interlocuteurs : saisie manuelle, matrices Excel à compléter, dépôt de documents sur le portail suivant des procédures de catalogage, récupération des exports de bases de données, protocole OAI-PMH.

L'objectif étant de représenter un état des connaissances expertes sur l'environnement en Bretagne, le projet s'étend également à toute littérature blanche de niveau expert, pertinente et cohérente au projet.

2 Le géocatalogue des études sur l'environnement en Bretagne : des services et un fonctionnement complexe

2.1 Le géocatalogue : fonctionnalités et services

Le géocatalogue fonctionne avec le CMS libre *EZpublish*. Son développement et sa charte graphique ont été sous-traités à la société *SMILE*. Le GIP BE doit donc faire appel à cette société pour les développements informatiques. Par contre, il accède directement à l'interface administrateur qui lui permet d'éditer le site ainsi que d'alimenter la base de données du géocatalogue. Pour cela, le logiciel est paramétré pour importer des fichiers au format CSV²⁴, ces fichiers devant par ailleurs respecter une matrice sur laquelle nous reviendrons plus en détail plus loin.

L'interface d'accueil du géocatalogue (voir **annexe II**) offre les fonctionnalités suivantes : le visionnement des dernières nouveautés triées par date d'édition la plus récente, un accès par recherche cartographique ainsi qu'un accès par recherche textuelle et enfin, des statistiques dynamiques sur l'état du fonds du catalogue (nombre de documents, nombre d'organismes représentés, quantités par type de document).

Du point de vue de la recherche, les deux approches textuelles et cartographiques offrent l'utilisation de filtres que l'on peut assimiler aux facettes de plus en plus utilisées que ce soit depuis quelque temps par les sites Web d'e-commerce et puis plus récemment par les nouvelles solutions documentaires comme les Web OPAC nouvelles générations ou le projet *Flamenco*²⁵. Ce type d'approche pour la recherche d'information renvoie conceptuellement à la classification à facettes apparue en 1936 par l'indien bibliothécaire et mathématicien Ranganathan. Aujourd'hui, le concept a évolué avec les usages liés à l'avènement d'Internet et on parlera davantage dans les SIC de « *navigation à facettes* »²⁶ (*faceted navigation*) concernant la recherche d'information. La recherche cartographique propose d'utiliser les filtres par organisme, par type de territoire (ex : communes) et par thème, alors que la recherche textuelle offre le filtre par type de document ainsi que

²⁴ Comma-separated values

²⁵ Exemple de l'interface facetée pour les vainqueurs du Prix Nobel : <http://orange.sims.berkeley.edu/cgi-bin/flamenco.cgi/nobel/Flamenco>

²⁶ Maisonneuve Marc & Touitou Cécile. *Une nouvelle famille d'Opac*. BBF 2007, t. 52, n. 6 (dossier : nouveaux outils pour la recherche) : <http://bbf.enssib.fr/consulter/bbf-2007-06-0012-002>

par territoire (ex : Brest) en plus. Aussi, le filtre thématique sur la recherche textuelle offre un niveau de plus de précision dans le plan de classement que sur la recherche cartographique. En effet, le thésaurus interne est constitué par cinq niveaux d'indexation allant du plus générique au plus spécifique. Les deux premiers niveaux sont donc utilisés sur le portail à la manière d'un plan de classement pour l'accès à l'information.

La recherche textuelle correspond à ce que nous connaissons davantage traditionnellement dans le cadre de la recherche bibliographique. La recherche libre permet d'interroger sur un seul champ de requête tous les champs de la notice bibliographique. La recherche avancée permet d'interroger différents champs par critères de titre, auteur, collection, éditeur et année d'édition.

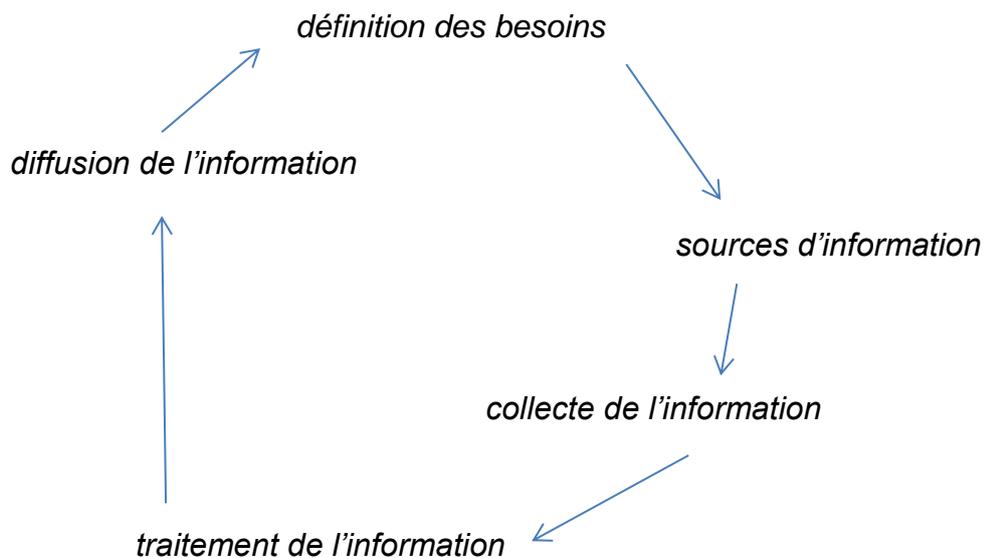
En plus des fonctionnalités précédemment évoquées, le site propose des services supplémentaires pour le géocatalogue des études :

- Un flux RSS permet de s'abonner et de s'informer en temps réel des mises à jour en termes de contenu.
- La possibilité de se créer un compte personnel permet de recevoir par email des notifications concernant les mises à jour du site ainsi que d'envoyer un mail contenant une référence bibliographique depuis son espace personnel. Enfin, La fonction « Panier » permet de sélectionner et sauvegarder des références bibliographiques et de les télécharger.
- Une FAQ est présente pour aider les utilisateurs de la base Etudes afin d'anticiper et de répondre aux questions les plus fréquemment posées.
- La possibilité de proposer un document est offerte à l'utilisateur. Ainsi celui-ci peut soumettre une référence bibliographique au GIP BE à l'aide du formulaire en ligne.

2.2 Le géocatalogue dans le miroir du processus itératif visant à satisfaire un besoin d'information (besoins – sources – collecte – traitement – diffusion)

En mettant de côté pour le moment l'aspect géolocalisation et en prenant pour postulat que les missions de référencement et de centralisation du géocatalogue sont les piliers de fonctionnement du projet, nous pouvons imaginer représenter les processus mises en œuvre sous l'angle du processus itératif visant à satisfaire un

besoin d'information. Cela est notamment beaucoup utilisé pour théoriser le cycle de la veille informationnelle²⁷. Ce cycle a été décrit et représenté de nombreuses fois sous différentes variantes et dans différentes finalités (renseignement militaire, veille concurrentielle, stratégique, etc...), nous l'utiliserons cependant ici étant donné les aspects centralisateur, diffuseur, itératif automatisé du projet. On peut représenter le cycle ainsi :



²⁷ Doc TICE. Veille informationnelle : <http://missiontice.ac-besancon.fr/documentation/index.php/veille-informationnelle/>

2.2.1 Définition des besoins : thématiques et typologie des documents

Au niveau des thématiques développées au sein du site Études celui-ci reprend les grandes thématiques environnementales que sont :

- air et climat
- déchets
- développement durable
- eau
- énergie
- mer et littoral
- patrimoine naturel
- paysages
- risques
- santé (environnement)
- nuisances
- sols
- sous-sol
- activités et usages.

La typologie des documents retenus pour la constitution du catalogue est décrite dans le tableau suivant :

Article scientifique	Article publié dans une revue scientifique nationale ou internationale
Étude thématique	Travail portant sur un thème précis et s'appuyant sur des données, des chiffres ou un corpus bibliographique
Étude bibliographique	Liste structurée de références d'ouvrages ou d'autres documents, notamment d'articles de revues, sur un thème commun
Étude d'impact	Etude technique visant à apprécier les conséquences, notamment environnementales, de la mise en œuvre d'un projet ou d'un programme
Thèse / Mémoire	Document réalisé par un chercheur, rendant compte de ses travaux de recherche et de leurs résultats sous la responsabilité d'un directeur de thèse, soutenu devant un jury universitaire pour l'obtention d'un diplôme universitaire (doctorat) ; travail de fin d'études d'un étudiant de niveau Bac +4 et Bac +5
Diagnostic / étude préalable	Etude liée à un aménagement, un programme d'action, une politique publique
Rapport / bilan	Rapports et bilans de programmes d'action, rapports d'enquête
Suivi / Inventaire	Campagne de mesure pour des séries courtes ou longues
Charte / Guide et fiche technique	Document ayant une portée territoriale.
Atlas	Recueil ordonné de cartes conçu pour représenter un espace donné et/ou exposer un ou plusieurs thèmes
Acte de colloque/congrès	Document recensant la version écrite des conférences prononcées lors d'un colloque ou d'un congrès

Notons que tous ces documents peuvent être considérés comme de la littérature grise mise à part le type de document « article scientifique » qui se différencie du reste du corpus. Pour rappel, la littérature grise est définie par l'ADBS comme tout « *document produit à l'intention d'un public restreint, en dehors des circuits commerciaux de l'édition et de la diffusion, et en marge des dispositifs de contrôle bibliographique. Le développement de la mise à disposition de ces documents sur les réseaux de télécommunication remet en question ce concept* »²⁸.

Alors un article scientifique ne peut être considéré comme de la littérature grise qu'il soit commercialisé par un éditeur scientifique ou disponible sur une archive ouverte. Toutefois, le projet d'état des connaissances vise à référencer tout ce qui est pertinent aux thématiques du projet, même si les références ne sont pas des données publiques ou de la littérature grise. Une ressource soumise à monétisation pour son accès au texte intégral sera donc tout de même référencée et pointera sur la source d'origine où les conditions de consultation seront présentées.

2.2.2 Sources d'informations

2.2.2.1 Les acteurs de l'environnement en Bretagne

C'est à partir de l'annuaire des acteurs de l'environnement en Bretagne²⁹ que le travail de prise de contact avec les sources potentielles a été entrepris. Cet annuaire est le fruit du travail du GIP BE depuis sa création. De nombreux échanges avec ces acteurs ont permis d'établir une cartographie détaillée permettant de connaître pour chacun d'eux les renseignements suivants lorsqu'ils pouvaient être complétés : le territoire d'intervention, le type d'organisme, le contact pour le service de documentation, l'appartenance à une collection HAL, l'adresse du dépôt HAL, l'adresse URL OAI, un descriptif de la structure, les types de documents possédés et le type d'études pour l'intégration au projet, le nombre d'études et les possibilités de partenariat.

²⁸ ADBS. *Littérature grise* : http://www.adbs.fr/litterature-grise-17647.htm?RH=OUTILS_VOC

²⁹ Annuaire des acteurs de l'environnement en Bretagne du GIP BE : <http://www.bretagne-environnement.org/Media/Acteurs/Organismes>

Les organismes peuvent être classés selon la typologie suivante :

- services de l'état
- les collectivités ou organismes qui en émanent ayant des compétences en matière d'environnement (Conseil régional, conseils généraux, ...)
- les associations agréées en matière d'environnement et productrices de données ou les associations reconnues pour leur expertise (ex : Cedre, Coeur, Bretagne vivante, ...)
- les universités (ex : Rennes2, Osur, ...)
- les instituts de recherches, organismes et établissements publics (ex : CNRS, Onema, BRGM...)
- les instances professionnelles fédératives et les chambres consulaires

Au sein de ces acteurs, nous allons retrouver trois profils différents du point de vue de leur gestion documentaire. Un premier ensemble est constitué par les organismes qui ne sont pas structurés documentairement. Il s'agit par exemple de Bretagne Vivante, du Groupe Mammalogique Breton (GMB), de l'AAMP (parc marin d'Iroise), du CEVA etc... Un second ensemble est constitué par les organismes qui ont une base documentaire tels que le Centre Botanique National de Brest (CBNB), de l'ADEME et l'IDELE (Institut de l'élevage). Enfin, le troisième ensemble est constitué par les réservoirs d'archives ouvertes comme HAL, Documentation Eau France et AgroCampus.

Le fait de relayer à la fois les connaissances à forte visibilité et mieux équipés au niveau gestion documentaire, tout comme celles moins accessibles et moins bien référencées constitue l'une des richesses du fonds.

2.2.2.2 D'autres sources exploitables ?

Il faisait partie de mes missions d'explorer l'éventualité d'intégrer de nouveaux contributeurs pertinents au projet du géocatalogue des études. J'ai alors été guidé vers la réflexion sur l'utilisation potentielle d'autres types de sources que sont par exemple les portails *RefDoc*³⁰ et *WorldCat*³¹.

³⁰ Site *RefDoc* : <http://www.refdoc.fr/>

³¹ Site *WorldCat* : <http://www.worldcat.org/>

Le premier est un service de référencement documentaire scientifique de l'INIST qui délivre les textes intégraux sur commande ou pointe vers le libre accès lorsqu'il existe, alors que le second est une initiative de l'OCLC visant à créer un point d'entrée unique aux catalogues des bibliothèques mondiales participant au projet.

Au premier abord le constat est que ce type de source offre rapidement des références très nombreuses, faciles d'accès et interopérable. J'ai entrepris de réaliser un inventaire des sources pas encore exploitées pouvant potentiellement s'intégrer au projet (*voir tableau 1 ci-dessous*). Ensuite, j'ai élaboré une typologie de champs pouvant permettre de les caractériser (*voir tableau 2 ci-dessous*). Ce travail exploratoire avait pour but d'offrir une première visibilité ainsi qu'une grille de lecture pour mieux appréhender la compréhension des enjeux liés à l'utilisation de ces nouvelles sources, notamment en terme de redondance, de complémentarité et d'efficacité par rapports aux sources déjà utilisées.

Tableau 1 : Liste des sources potentiellement pertinentes au projet :

Sources	URL
Refdoc	http://www.refdoc.fr/
WorldCat	http://www.worldcat.org/
BASE Lab	http://lab.base-search.net/
i-revues	http://irevues.inist.fr/
Google scholar	http://scholar.google.fr/
Lara	http://lara.inist.fr/
OAlster	http://oaister.worldcat.org/
OpenAIRE	http://www.openaire.eu/
DOAJ	http://www.doaj.org/
bibliovie	http://bibliovie.inist.fr/
BioText	http://biosearch.berkeley.edu/
TEL	http://tel.archives-ouvertes.fr/
DART-Europe	http://www.dart-europe.eu/
theses.fr	http://www.theses.fr/
OpenGray (OG)	http://www.opengrey.eu/
scirus	http://www.scirus.com/
WorldWideScience	http://worldwidescience.org/
Pascal	Accès payant
Erudit	http://www.erudit.org/
Revues.org / open edition	http://www.revues.org/

Tableau 2 : Grille de lecture :

Champs	Exemple : LARA
Nom complet	Libre accès aux rapports scientifiques et techniques
Éditeur	INIST
Thématiques	Sciences de la vie et santé, sciences de l'Homme et de la Société, Sciences Exactes et Sciences Appliquées
Descriptif	LARA vous permet de consulter librement et gratuitement l'intégralité des rapports scientifiques et techniques qui vous intéressent. LARA ne présente pas de rapport à caractère confidentiel. Ce nouveau site est né de l'engagement du CNRS dans le mouvement de l' Open Access et s'inscrit dans le prolongement du rôle de l' INIST-CNRS en tant que guichet national pour la collecte des rapports scientifiques et techniques français. LARA est un site fédérateur et multidisciplinaire, alimenté par différents laboratoires, instituts, centres de recherche et organismes français. Il a pour objet la valorisation de la littérature grise scientifique et technique et plus précisément des rapports, qu'ils soient techniques, de recherche, d'étude, d'activité, institutionnels...
Contributeurs	CERTU, CNAM, INIST-CNRS, INRP, INRS, INSERM, LIP
Proposition de document (ouvert au public)	Oui
Intégré (fédéré) à d'autres bases de données	Information à trouver
Format de métadonnées	Format Dublin Core
Protocoles d'échanges de données (moissonnage/interopérabilité)	Protocole OAI-PMH
Recherche libre ?	Oui
Recherche avancée ?	Oui
Recherche plein texte ? (texte intégral)	Oui
Recherche à facettes	Non
Système de collections (type HAL)	Information à trouver
Alertes / actualisation	Alertes courriels et fils RSS

Texte intégral (<i>full text</i>)	Oui (uniquement), formats textes (.doc, .rdf, .pdf...), URI pour chaque document
Type de documents	Rapports scientifiques et techniques (de recherche, d'étude, d'activité, institutionnels...)
Archive ouverte ?	Oui
Plateforme logicielle	DSpace (libre)
Droits de consultation / licences	Accès libre aux utilisateurs finaux sous réserve du respect des conditions générales d'utilisation
Gratuit / payant	Gratuit
Commentaires	
Contact	http://lara.inist.fr/feedback
URL	http://lara.inist.fr/

L'avantage premier de ces sources est qu'elles offrent des solutions élaborées : interfaces de recherches, jeu de métadonnées, protocole d'échange des données, systèmes d'alertes et d'actualisation.

Une cartographie affinée de ces sources peut s'avérer pertinente et nécessaire pour se repérer dans un univers documentaire très riche et vaste. En effet, de nombreux outils et portails existent avec des caractéristiques et spécificités particulières. Alors, il revêt d'une importance méthodologique et stratégique d'avoir une représentation d'ensemble de cet « écosystème » mais aussi pour chacun, de connaître ce que l'on peut y trouver, comment y effectuer des recherches, comment récupérer des notices bibliographiques et des pleins textes, dans quel projet l'outil prend-t-il forme et s'inscrit-il ?

Aussi nous pouvons constater qu'il existe un phénomène réticulaire, c'est-à-dire où certaines bases de données sont elles-mêmes moissonnées et représentées dans d'autres. Ces sources sont aussi beaucoup plus génériques et moins précises en termes de pertinence de contenu que les partenaires directs de l'environnement en Bretagne. Il peut ainsi il y avoir beaucoup de bruit documentaire et un travail conséquent de sélection et de validation de la pertinence des contenus peut en être la conséquence.

Enfin nous pouvons nous interroger également sur l'identité des documents et références que l'on retirerait d'une utilisation de ces outils en comparaison avec ceux collectés via les partenaires du réseau du GIP BE. Dans quelles mesures les résultats trouvés seront des doublons ? Dans quelles mesures il y aura-t-il des nouveautés et donc une complémentarité ?

2.2.3 Collecte de l'information

2.2.3.1 Aspect collaboratif

Rappelons que l'ambition du projet de géocatalogue des études est de co-construire la base de donnée documentaire et ainsi de responsabiliser les contributeurs en leur proposant des outils et des méthodes afin d'homogénéiser les pratiques documentaires. Les méthodes varient selon les pratiques documentaires des partenaires, la gestion documentaire des études pouvant à une extrémité être inexistante à l'autre extrémité être moissonnable et interopérable dans le respect des normes documentaires internationales. En ce sens on peut distinguer deux grands types de contributeurs chez les partenaires : d'une part les organismes de recherche et les établissements publics qui sont la plupart du temps bien équipés techniquement et humainement, et d'autre part les associations et autres petites structures locales dont la gestion de ressources documentaires structurées est beaucoup moins rationalisée et ne représente pas vraisemblablement une préoccupation ou une priorité.

Un symbole de cette approche collaborative où la notion de collaborateur est à prendre dans son sens le plus fort qu'il soit, est la possibilité offerte à l'utilisateur, directement depuis le site, de soumettre une référence et/ou un document avec une aide à la saisie des métadonnées (voir **annexe III**). Cette démarche que l'on retrouve par exemple sur les portails LARA ou Documentation Eau France³², va dans le sens d'une vision « durable » de l'information comme précédemment évoqué ainsi que peut être loué comme une « bonne pratique » ou initiative, participant au développement de la « culture documentaire » des acteurs et usagers de l'information.

³² Outil d'aide à la saisie, couverture géographique : <http://www.documentation.eaufrance.fr/le-portail/outils-pour-les-contributeurs/article/outil-d-aide-a-la-saisie>

2.2.3.2 Trois méthodes de collecte

Différentes approches sont possibles pour intégrer les contenus, elles varient en fonction de la volumétrie ainsi que de la gestion documentaire utilisée par le contributeur :

- La saisie manuelle est la solution trouvée pour les petites quantités et pour les références qui ne sont pas d'un point de vue documentaire et informatique structurées. Il s'agit du cas par exemple d'un certain nombre d'associations de taille modeste qui ont ouvert leurs pratiques de diffusion documentaire à la mise en ligne sur le réseau Internet, à travers leurs sites Web, d'une page spéciale dédiée à la documentation publiée (voir l'exemple du site du GMB³³). J'ai pu durant le stage travailler de cette manière sur les références des associations Air Breizh, GMB et Cœur Emeraude. On peut retrouver sur ces pages une description bibliographique des publications, voir le lien sur le texte intégral au format PDF. Toutefois, techniquement, les informations bibliographiques concernées sont du simple texte au sein de pages HTML et ne sont en aucun cas structurées (logique champ / valeur capable d'être échangée informatiquement). Ce travail manuel peut être réalisé par l'organisme lui-même via la transmission par le GIP BE d'une matrice Excel (voir **annexe IV**) et d'une aide à la saisie, ou par le GIP BE lui-même si l'organisme déclare ne pas pouvoir le faire. Cette solution reste adaptée dans un contexte de petites quantités et pour des structures n'investissant pas dans une gestion documentaire. Toutefois, l'expérience montre qu'il est encore compliqué de faire s'approprier la matrice par les structures afin qu'ils la complètent eux-mêmes. Cela peut peut-être parfois paraître encore utopique aujourd'hui, au vu des enseignements ressortis lors d'un entretien téléphonique avec l'un de ces partenaires : *pas les moyens temps / budget, pas le métier, pas les compétences et les connaissances, faible intérêt, prise de conscience de l'intérêt ultérieur que cela peut avoir ?*

³³ Page web des publications de l'association Groupe Mammalogique Breton (GMB) : <http://www.gmb.asso.fr/publications.html>

Pourtant cela rentre néanmoins dans la philosophie et les objectifs du projet ainsi que dans l'idée que l'on peut se faire du développement du Web et du partage des connaissances avec la notion d' « information durable » et l'avènement prochainement du Web des données, symbolisé par toujours plus de partage, de structuration et d'interopérabilité des données. Un travail encore plus développé de sensibilisation ainsi que des formations peuvent être des solutions.

- L'import de fichiers structurés d'un point de vue documentaire est possible lorsque les organismes possèdent et gèrent une base de données bibliographique. Je n'ai cependant pas eu l'occasion de travailler sur ce type de collecte durant mon stage.
- L'intégration avec le protocole d'interopérabilité OAI-PMH quand les contributeurs disposent de cette possibilité, permet de faire remonter des volumes énormes de références documentaires structurées. Ainsi c'est l'entrepôt HAL qui a fait l'objet d'une moisson sur une collection portant sur l'environnement. Le flux OAI de Documentation Eau France a quant à lui été moissonné avant mon arrivée et son traitement a constitué une grosse partie de mon stage, nous y reviendrons après.

Pour revenir plus en profondeur, sur le protocole OAI-PMH, rappelons qu'il est né dans le cadre de l'initiative des Archives Ouvertes que nous avons évoqués précédemment. Il s'agit d'un protocole de collecte et d'échanges de métadonnées créait pour l'interopérabilité entre différents serveurs de publications électroniques scientifiques.³⁴ Or l'interopérabilité est essentielle pour l'accès à l'information scientifique dans « *un panorama informationnel où les ressources sont distribuées et hétérogènes* »³⁵. On peut définir l'interopérabilité des ressources comme « *leur capacité à communiquer ensemble et à échanger des données entre elles suivant un format et une sémantique maîtrisée* »³⁶ alors que la notion de protocole renvoie à l' « *ensemble des règles selon lesquelles les applications des ordinateurs échangent*

³⁴ Aubry Christine & Janik Joanna (Dir.). *Les archives ouvertes enjeux et pratiques*.

³⁵ André Francis. *Z39.50 versus OAI-PMH : une histoire de générations ?* in *Les archives ouvertes : enjeux et pratiques*.

³⁶ Idem

des données »³⁷. Le protocole OAI-PMH, d'une autre génération que le protocole Z39.50 datant de bientôt trente ans, permet de rendre interopérable des archives distribuées, quelque soient leur taille et volume, en les agrégeant virtuellement par une interrogation unique des métadonnées rassemblées. Enfin, précisons encore que le format minimal de métadonnées supporté par le protocole OAI-PMH est le Dublin Core alors que les réponses aux requêtes sont transmises sous la forme de données XML.³⁸ Notons que le protocole Z39.50 est toujours utilisé mais est davantage lié à l'environnement des bibliothèques.

2.2.3.3 La problématique de l'actualisation

Si le projet étant récent, la problématique actuelle consiste avant tout à alimenter la base de manière rétrospective pour la rendre la plus complète et exhaustive possible, il en sera différent dans l'avenir. En effet, le projet pensé dans sa continuité pose la question de son actualisation dans le temps. Pour cela, un planning a été élaboré pour planifier les collectes auprès des différents partenaires. Il servira de tableau de bord pour planifier ces opérations.

Notons que HAL propose de nombreux flux RSS paramétrables pour s'abonner aux nouveautés sur l'archive³⁹. Ainsi il est par exemple possible de s'abonner aux articles déposés par un auteur en particulier, ou bien un laboratoire, une équipe de recherche, un domaine scientifique, etc... Cela ouvre des perspectives pour affiner le paramétrage d'une veille pour le projet études.

2.2.4 Traitement de l'information

2.2.4.1 Jeu de métadonnées Dublin Core et matrice Excel

Le jeu de métadonnées utilisé pour la description des représentations bibliographiques est le format Dublin Core. Il fait partie des standards récents, au contraire des formats MARC, et correspond à une description plus généraliste s'appliquant à toute type de ressource et par conséquent plus interopérable. La Bibliothèque Nationale de France (BNF) présente les objectifs du Dublin Core comme étant « *de fournir un socle commun d'éléments descriptifs pour améliorer le signalement et la recherche de ressources au-delà des diverses communautés et des nombreux formats descriptifs propres à chaque spécialité, tout en restant*

³⁷ Idem

³⁸ André Francis. *Z39.50 versus OAI-PMH : une histoire de générations ?* in Les archives ouvertes : enjeux et pratiques.

³⁹ Canaux RSS de HAL : http://hal.archives-ouvertes.fr/index.php?action_todo=rss

suffisamment structuré »⁴⁰. Pour ces raisons, on se rend bien compte que ce format de description simple et générique est adapté aux problématiques du catalogue études.

La matrice utilisée par le GIP BE pour le géocatalogue (voir **annexe IV**) présente quelques modifications du Dublin Core permettant de rendre compte des spécificités du projet. En effet, si la majorité des champs sont repris, la couverture géographique n'est par exemple pas celle entendue par le champ *coverage* du Dublin Core. Ici il ne s'agit pas de la couverture spatio-temporelle au sens large de la ressource mais plutôt celle spécifique aux territoires bretons.

2.2.4.2 Indexation : thésaurus et matrices

Le travail d'indexation se fait avec l'appui de documents assurant l'utilisation d'un langage documentaire contrôlé. Le thésaurus interne qui est plus une liste structurée de mots-clefs sous la forme d'un plan de classement sur quatre niveaux d'affinage, permet d'indexer les références. Le champ « catégorie » correspond alors à l'indexation permettant de retrouver une référence à travers la recherche par le plan de classement du portail. Le champ mots-clefs correspond lui à la recherche à travers l'interface de requête. Le thésaurus épouse toutes les thématiques du catalogue précédemment présentées.

Les matrices sont importantes car elles permettent également de normaliser le vocabulaire à utiliser pour caractériser les types de documents, les types de territoires, ainsi que tous les territoires précisément. Cette rigueur de vocabulaire est importante car la matrice complétée doit être interopérable et donc comprise par le CMS à l'import. Sans quoi la moindre erreur de caractère empêchera la machine de reconnaître un territoire ou un type de document. Aussi, un système de correspondance des territoires avec des codes numériques est indispensable au bon fonctionnement de la fonctionnalité de géolocalisation.

2.2.4.3 Le géo-référencement

La grosse valeur ajoutée du catalogue réside dans sa fonction de géolocalisation qui permet de représenter visuellement, dynamiquement ainsi que de comparer des états des connaissances par types de territoires. Le périmètre de la géolocalisation tout comme celui du corpus retenu et celui de la région Bretagne et

⁴⁰ BNF. Dublin Core : http://www.bnf.fr/fr/professionnels/formats_catalogage/a.f_dublin_core.html

infra, c'est-à-dire les départements, les pays, les SAGE, les bassins versants, les parcs naturels, les EPCI et les communes qui représentent le plus petit maillon à l'échelle locale. Les études menées à plus grande échelle (inter-régionales, nationales, internationales) peuvent être intégrées sur le site, si des éléments contenus dans ces études traitent du territoire breton.

Une étude ne peut être géo référencée que sur un seul et unique type de territoire. Ceci peut parfois représenter un frein au géo-référencement exhaustif d'une ressource. En effet certaines références peuvent prétendre à plusieurs niveaux de territoire. Par exemple une étude réalisée sur les communes de Dinan, St Malo et Dinard dans le cadre du SAGE Rance-Frémur-Baie de Beausseis peut prétendre soit à la géolocalisation par communes, soit à la géolocalisation par SAGE. Un choix devra être fait et dans tous les cas une géolocalisation ne sera pas représentée. A titre de comparaison nous pouvons observer que le catalogue de Documentation Eau France choisit de géo-référencer pour une référence tous les niveaux de territoire (ex : Département Ille-et-Vilaine, bassin versant de la Vilaine, Commune de Rennes).

2.2.4.4 Méthodologie de traitement d'un fichier OAI volumétrique (plus de trois mille références)

Une grosse partie de ma mission a consisté à traiter un fichier OAI extrêmement volumineux correspondant à plus de 3 000 références. Il s'agit du flux OAI de Documentation Eau France. Après avoir été filtré par rapport aux zones géographiques bretonnes, il a été traité pour l'adapter à la matrice du site études, et ensuite il a été trié suivant un score de pertinence, au regard d'algorithmes. Enfin un dédoublonnage avec la base existante a été réalisé.

Si la moisson permet de récupérer beaucoup de métadonnées, il n'en reste pas moins qu'un gros travail de catalogage reste à réaliser pour vérifier, compléter, corriger ou adapter la description bibliographique récupérée aux outils et exigences documentaires du GIP BE. Face à de tels volumes, il s'est révélé qu'un catalogage traditionnel, tel que je le concevais jusqu'à présent n'était pas forcément le plus adapté et efficace. Reprendre notice par notice étant un travail extrêmement laborieux où l'on ne se voit pas avancer dans un volume si dense et où l'on perd énormément de temps pour renseigner ou vérifier les vingt champs. Il est apparu donc que face à des quantités de références volumétriques à traiter, il fallait adopter

des techniques de catalogage volumétriques. Pour ce faire, j'ai progressivement mis en place la stratégie que je vais présenter ci-dessous. La logique consiste à travailler non pas horizontalement référence par référence mais verticalement type de champ par type de champ. Cela considère une approche différente qui induit que l'on ne catalogue et lis jamais une notice dans son intégralité. Quoiqu'il en soit, que cela peut procurer, il est de toute façon question d'un équilibre à trouver entre un travail « volumétrique » et un souci de la rigueur indispensable pour obtenir un catalogue correct et le plus qualitatif possible.

La méthode que j'ai finalement suivie est la suivante :

1. Validation des références et éligibilité au site études. (utilisation principale des champs « titre », « descriptif », « couverture géographique » et mots - clefs », voir « page internet sur l'ouvrage », les autres champs étant masqués)
2. Filtre sur les oui afin de ne plus travailler que sur les notices retenues comme éligibles
3. Il est ensuite possible de s'atteler au renseignement des champs d'indexation (type de bibliographie, catégories, mots-clefs) et de géo-référencement à des techniques purement volumétriques. Cela consiste à réaliser des filtres sur les champs concernés avec certains mots clefs redondants du flux afin d'indexer toutes les références présentant l'occurrence en question par le mot-clef ou le géo-référencement approprié dans les langages contrôlés du GIP BE. Les filtres peuvent être réalisés sur le titre, les mots-clefs ou la couverture géographique. Un document de travail (voir **annexe V**) m'a permis de lister tous ces mots-clefs utilisés comme filtre afin de me repérer dans ma progression. Ce type de catalogage que j'aurai envie d'appeler « inversé » car on ne part pas de la référence (son titre, sa description, son texte intégral) pour l'indexer avec plusieurs mots-clefs mais on part plutôt d'un mot-clef pour indexer plusieurs références. Précisons que ce travail est à voir dans une optique et un effort de dégrossissement du traitement à réaliser mais qu'une relecture ligne par ligne restera indispensable dans l'étape suivante pour assurer une qualité de catalogue documentaire optimale. Par ailleurs, il m'est

apparu qu'un logiciel de type TAL, et plus précisément de lexicométrie, serait d'une aide précieuse à cette entreprise. En effet mon idée était de récupérer les mots-clefs de toutes les références du flux OAI pour en réaliser un traitement statistique des occurrences. Ainsi nous pourrions savoir que le mot-clef index le plus utilisé est « gestion de l'eau » (320 occurrences), suivi de « poissons migrateurs » (289 occurrences), et ainsi de suite... Ce type de renseignement permettrait d'établir un tableau de bord des mots-clefs à traiter en priorité car les plus répétitifs. Alors nous pourrions laisser conserver les plus faibles occurrences pour la saisie plus traditionnelle. Je n'ai pu malheureusement réaliser ce type de traitement lexicométrique de manière satisfaisante à l'aide des rares logiciels libres que j'ai trouvé. Cependant il est fort probable qu'une solution commercialisée tel que le logiciel *Lexico3*⁴¹, permettrait avec un bon paramétrage d'arriver à nos fins.

4. La dernière étape consiste reprendre les références cette fois une par une, « horizontalement » et qualitativement afin de compléter les champs manquants (ex : organismes associés) ou insuffisamment catalogués.

2.2.5 Diffusion de l'information (et export)

Nous passerons rapidement sur cette étape, car s'agissant d'un catalogue en ligne, c'est de l'interface ouverte au public que nous avons à traiter ici. Or, Nous avons déjà parlé de ces aspects dans la partie sur les fonctionnalités et services du géocatalogue.

Nous pouvons par contre appuyer le fait que toutes les références du géocatalogue des études sont moissonnables et interopérables avec le portail national dédié à l'environnement : *tout sur l'environnement*⁴². Cette communication de machine à machine, permet ainsi d'exporter vers une base plus large et ainsi diffuser et donner une visibilité nationale au fonds du géocatalogue des études sur l'environnement en Bretagne. Le protocole utilisé est aussi OAI-PMH.

⁴¹ Site Web - présentation Lexico3 : <http://www.tal.univ-paris3.fr/lexico/>

⁴² Portail Tout sur l'environnement : <http://www.toutsurlenvironnement.fr/>

3 Le géocatalogue des études sur l'environnement en Bretagne : gestion de grosse volumétrie, automatisation, rigueur et qualité de catalogage, l' « information durable » au cœur d'un équilibre difficile à trouver

3.1 Un équilibre à trouver entre les procédures automatiques et la rigueur ainsi que la qualité du catalogage

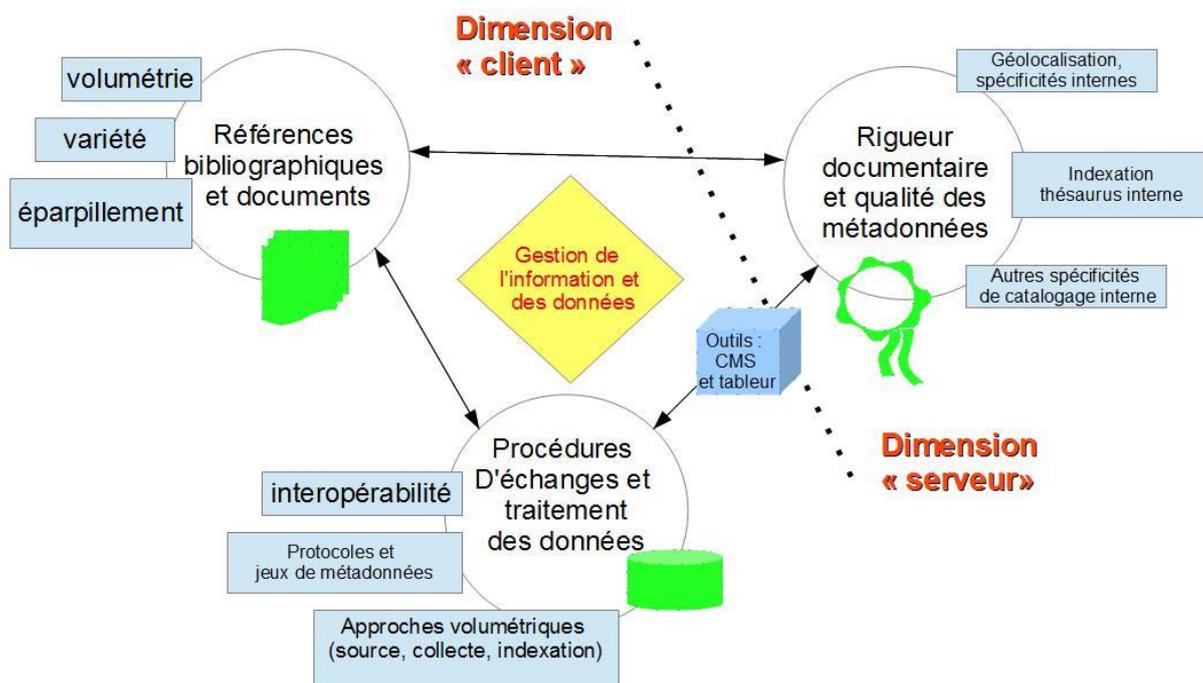
Pour ce type de projet tel que celui du géocatalogue des études, on peut constater qu'il existe une tension entre la manipulation de gros volumes de références à traiter et la nécessaire rigueur et qualité documentaire indispensable à la pertinence et au bon fonctionnement du catalogue. Rappelons que la qualité ou inversement la mauvaise qualité du catalogage et notamment du volet indexation impactera directement les accès et donc la possibilité de retrouver facilement ou non, voire pas du tout, les références. Cependant cette rigueur sera très chronophage. Nous avons vu que, alors que les progrès techniques et de normalisation, ont permis l'utilisation d'outils allant vers toujours plus d'interopérabilité : protocole OAI-PMH, jeu de métadonnées Dublin Core, permettant de récupérer tout ou partie et de traiter des volumes toujours plus importants et variés de références. Il n'en reste pas moins que dans la pratique, les particularités et spécificités de chacun nécessitent une intervention humaine d'un professionnel de l'information – documentation.

En effet, comment un ordinateur peut-il comprendre qu'une notice ne parle pas de la ville de Rennes mais des rennes en Scandinavie ? Cette exemple d'ambiguïté sémantique que nous avons retrouvé dans la pratique, est symbolique des précautions qu'il est utile de prendre aujourd'hui lorsque l'on traite de gros volumes à l'aide de l'outil informatique. Les progrès au niveau de la sémantique vont dans ce sens avec le développement des ontologies. En effet, comme l'écrit Olivier Ertzscheid, les usagers expriment de plus en plus fortement le besoin de « *disposer de fonctionnalités de recherche « intelligente », lesquelles ne peuvent être imaginées sans que soit mis en place un formalisme ontologique minimal* »⁴³. Toutefois cette technologie est encore trop jeune pour imaginer son application directe dans le cadre du projet même si après le temps des doutes, ce secteur est en pleine progression actuellement.

⁴³ Ertzscheid Olivier. *Moteurs de recherche : des enjeux d'aujourd'hui aux moteurs de demain*. In Métadonnées : mutations et perspectives.

Dans la figure ci-dessous, vous trouverez une tentative de représentation et de schématisation des tensions ainsi évoquées entre de grosses volumétries, le besoin d'automatiser et la nécessité de rigueur de catalogage.

Tensions volumétriques / automatisation / rigueur



3.2 Croisement et ouverture avec le concept d'information durable

Revenons plus en profondeur sur le concept précédemment évoqué d'information durable. L'article de Documentaliste-Science de l'information coordonné par Christian Bourré⁴⁴ considère qu'à l'heure de la société de la connaissance, le capital immatériel est devenu un moteur de la croissance économique. Le capital immatériel est à considérer comme la bonne utilisation que l'on fait de l'information professionnelle provenant des organisations privées et publiques. Les auteurs expriment l'idée que l'on peut « *considérer l'information comme un bien durable* », sa valeur ajoutée évoluant en fonction du contexte. Ce

⁴⁴ Bourret Christian *et al.*, « Capital immatériel et information professionnelle. L'émergence d'un concept nouveau : l'information durable »

contexte repose sur « *une meilleure exploitation des gisements d'information accumulés par les générations qui nous ont précédés, exploitation créatrice de richesse pour les générations futures* »⁴⁵.

Pour arriver à cet objectif d'information durable qui renvoie aux concepts de partage, d'équité, d'innovation et de responsabilité, les auteurs posent trois conditions à atteindre :

- « *La prise de conscience de la valeur de l'information accumulée, véritable gisement de richesse à condition de savoir identifier les bonnes informations, les évaluer, les trier et les sélectionner* ». Cela correspond à la raison d'être du projet de géocatalogue des études du GIP BE ainsi qu'à la réflexion menée dans ce travail sur l'importance des étapes de *sourcing*, de collecte et de traitement de l'information ;
- « *une organisation capable d'exploiter rationnellement et intelligemment ces gisements* ». Le travail mené par le GIP BE pour le géocatalogue des études avec les outils, procédures et partenariats mis en place ainsi que l'environnement global technologique et normatif actuel vont dans le sens de cette condition ;
- « *une gestion consciente de ces gisements qui doivent être alimentés en permanence, actualisés et validés afin d'éviter qu'ils ne perdent leur pertinence et leur valeur ajoutée* ». La problématique de l'actualisation des contenus dans le cadre du géocatalogue a été soulevée dans ce travail. Nous pourrions également ajouter qu'une nécessaire appropriation progressive des outils et technologies de structuration de l'information par l'ensemble des acteurs est un enjeu tout aussi important. Dans l'idée du concept de « culture de l'information », la culture de l'information « structurée » devient un enjeu majeur pour parvenir aux objectifs de l'information durable.

⁴⁵ Bourret Christian *et al.*, « Capital immatériel et information professionnelle. L'émergence d'un concept nouveau : l'information durable », *Documentaliste-Sciences de l'Information*

Conclusion

Ce mémoire a permis d'essayer d'éluder les tenants et les aboutissants du projet de géocatalogue des études sur l'environnement en Bretagne. Il en ressort que le domaine de l'environnement est un domaine particulier qui possède ses propres lois et pratiques. Toutefois, on ne peut pas non plus le séparer totalement d'autres domaines comme les archives ouvertes par exemple. Ainsi le projet d'état des connaissances du géocatalogue ne se limitant pas à la littérature grise et aux données publiques, les initiatives des archives ouvertes ou bien certaines plateformes relayant de la littérature blanche, peuvent s'imbriquer dans le *sourcing*, la collecte ou le traitement de références pour le géocatalogue.

Si nous avons vu aussi que le mouvement actuel menait inéluctablement à la création d'outils et de procédures de gestion et de partage de l'information bibliographique de plus en plus structurés et tournés vers les usagers (protocole OAI-PMH, métadonnées Dublin Core, ontologies), reste alors encore la question de l'appropriation de ces outils qui peut être débattue. Cette question orientée usages est un champ de recherche très intéressant déjà étudié dans les SIC, tout comme un réel enjeu pour le fonctionnement optimal du géocatalogue des études à l'avenir.

Bibliographie et Webographie

ADBS. *Littérature grise*. Consulté le 03/06/2013, URL : http://www.adbs.fr/litterature-grise-17647.htm?RH=OUTILS_VOC

ANDRE, Francis. *Libre accès aux savoirs (Open access to Knowledge)*. Paris : Futuribles perspectives, 2005.

ANDRE, Francis. « Z39.50 versus OAI-PMH : une histoire de générations ? » in *Les archives ouvertes : Enjeux et pratiques : guide à l'usage des professionnels de l'information*. ADBS édition, 2005.

AUBRY, Christine & JANIK, Joanna (Dir.). *Les archives ouvertes : Enjeux et pratiques : guide à l'usage des professionnels de l'information*. ADBS édition, 2005.

BOURRET, Christian et al., « Capital immatériel et information professionnelle. L'émergence d'un concept nouveau : l'information durable », *Documentaliste-Sciences de l'Information*, 2008/4 Vol. 45, p. 4-12. DOI : 10.3917/docsi.454.0004

BNF. *Dublin Core*. Consulté le 03/06/2013, URL : http://www.bnf.fr/fr/professionnels/formats_catalogage/a.f_dublin_core.html

Doc TICE. *Veille informationnelle*. Consulté le 03/06/2013, URL : <http://missiontice.ac-besancon.fr/documentation/index.php/veille-informationnelle/>

ERTZSCHEID, Olivier. « Moteurs de recherche : des enjeux d'aujourd'hui aux moteurs de demain ». In *Métadonnées : mutations et perspectives*. INRIA 29 Septembre – 3 octobre 2008 – Dijon. ADBS éditions.

LiberTIC, « l'open data ». In *Actions Open data* : <http://actionsopendata.org/l-open-data/definition/>

MAISONNEUVE, Marc & TOUITOU Cécile. « Une nouvelle famille d' Opac ». *BBF* 2007, t. 52, n. 6 (dossier : nouveaux outils pour la recherche). Consulté le 05/06/2013, URL : <http://bbf.enssib.fr/consulter/bbf-2007-06-0012-002>

Wikipédia. *Open Government Initiative*. Consulté le 05/06/2013, URL : http://fr.wikipedia.org/wiki/Open_Government_Initiative

Wikipédia. *Donnée publique*. Consulté le 08/06/2013, URL : http://fr.wikipedia.org/wiki/Donn%C3%A9e_publique

Wikipédia. *Open data*. Consulté le 08/06/2013, URL : http://fr.wikipedia.org/wiki/Open_data

Sigles

ADBS : Association des professionnels de l'information et de la documentation

ADEME : Agence De l'Environnement et de la Maîtrise de l'Energie

BRGM : Bureau de Recherches Géologiques et Minières

BBF : Bulletin des Bibliothèques de France

CADA : Commission d'Accès aux Documents Administratifs

CEVA : Centre d'Etude et de Valorisation des Algues

CMS : Content Management System (système de gestion de contenu)

CSV : Comma-separated values

CNRS : Centre National de la Recherche Scientifique

DIST : Direction de l'Information Scientifique et Technique

EPCI : Établissement public de coopération intercommunale

FAQ : Foire Aux Questions

GIP BE : Groupement d'Intérêt Public Bretagne Environnement

HAL : Hyper Article en Ligne

HTML : Hypertext Markup Language

INIST : Institut de l'Information Scientifique et Technique du CNRS

IST : Information Scientifique et Technique

MARC : Machine Readable Cataloging

OAI-PMH : Open Archives Initiative Protocol for Metadata Harvesting

OCLC : Online Computer Library Center

RSS : Really Simple Syndication

SAGE : Schéma d'Aménagement et de Gestion des Eaux

SIC : Sciences de l'Information et de la Communication

TAL : Traitement Automatique du Langage naturel

URI : Identifiant Numérique Unique

Sommaire des annexes

Annexe I : visualisation statistique cartographiée des études par communes

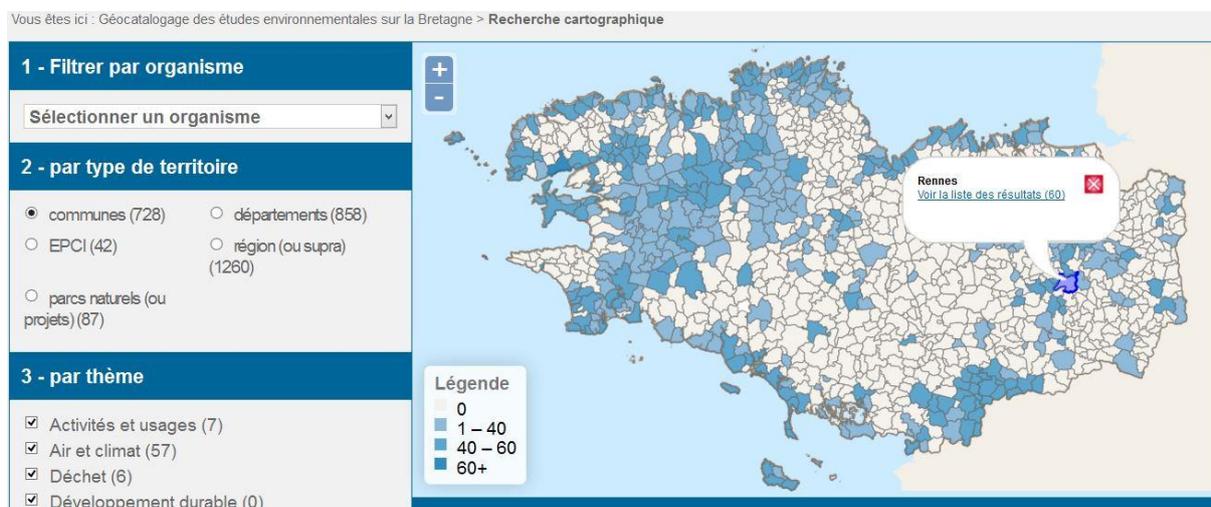
Annexe II : page d'accueil et fonctionnalités du géocatalogue

Annexe III : visualisations du service de proposition de référence documentaire au géocatalogue

Annexe IV : matrice Excel pour la saisie manuelle des références bibliographiques

Annexe V : document de travail pour catalogage volumétrique avec filtres sur un flux OAI Documentation Eau France

Annexe I



Légende : les nuances de couleurs correspondent aux nombres de documents géo-référencés par territoires

Annexe II

Visitez aussi : [portail régional](#) | [observatoire biodiversité](#) | [observatoire déchets](#) | [observatoire eau](#) | [observatoire énergie](#) | [cartographie dynamique](#) | [ma commune](#)

 **Géocatalogue des études sur l'environnement en Bretagne**

Rechercher

[recherche textuelle](#) | [recherche cartographique](#) | [services](#) | [qui sommes-nous ?](#) | [nos contributeurs](#)

Bienvenue sur la base documentaire experte dédiée à la connaissance environnementale sur la Bretagne.

Nouveautés	Recherche cartographique	Statistiques
<p>Synthèse régionale de la qualité de l'eau des bassins versants bretons rapport 2012</p> <p>Contrat Nature : Gestion durable de l'activité récréative de pêche à pied et préservation de la biodiversité littorale. Rapport final 2007-2011 rapport 2012</p> <p>Bilan 2011 du Centre régional opérationnel de surveillance et de sauvetage (Cross) Corsen rapport 2012</p> <p>Le changement climatique dans le Morbihan - Impacts, vulnérabilité et adaptation étude thématique 2012</p>	<p>Cherchez à partir d'une zone géographique et localisez les résultats sur une carte.</p>  <p>Accès à la recherche cartographique</p>	<p>3195 documents 128 organismes</p> <ul style="list-style-type: none">1881 articles scientifiques8 études bibliographiques0 études thématiques136 thèses / mémoires0 études d'impacts69 diagnostics / études préalables270 rapports123 suivis / inventaires17 chartes / guides ou fiches techniques20 atlas239 actes de colloques/congrès

Recherche textuelle

Annexe III :

Proposer un document

Titre *

Renseigner le titre exact du document.

Type de bibliographie *

article scientifique

Logo

Description

En quelques lignes, décrire le sujet traité et les publics visés.

Format **B** *I*                                              

Chemin:

Auteurs

Renseigner les noms et prénoms des auteurs. (S'il y a plusieurs auteurs, séparer leur nom par un point-virgule)

Éditeur

Renseigner la maison d'édition ou l'organisme éditeur (ne pas utiliser d'acronyme).

Fichier

Joindre le fichier PDF correspondant au document (maximum 100 Mo).

Page Internet sur l'ouvrage

Nombre de pages

Langue

Français ▲
Anglais
Espagnol
Allemand ▼

Année d'édition**Catégories ***

activités et usages ▲	<<		▲
agriculture	<		
cultures			
élevage	>		
aquaculture	>>		
algoculture			
conchyliculture			
pisciculture			
économie et industrie			
loisirs ▼			▼

Disponibilité

(en ligne, CD-ROM, en consultation,...)

Organismes associés

Cellule d'orientation régionale pour la protection ▲	<<		▲
Cellule économique de Bretagne	<		
Centre d'Etude et de Valorisation des Algues			
Centre d'étude sur le milieu d'Ouessant	>		
Centre de documentation, de recherche et d'ex	>>		
Centre de ressources et d'expertise scientifiq			
Centre Interprofessionnel Technique d'Etudes d			
Centre Permanent d'Initiative pour l'Environnen			
Centre permanent d'initiatives pour l'environner			
Centre permanent d'initiatives pour l'environner ▼			▼

Source**Références**

Droits

(restriction portant sur l'utilisation de la ressource, titulaires des droits, domaine public...)

Couverture géographique *

(modifier la valeur "commune" proposée par défaut si l'étude concerne un EPCI, un pays, un bassin versant, un département ou toute la région)

communes

Communes

Acigné
Allaire
Allineuc
Amanlis
Ambon
Andel
Andouillé-Neuville
Antrain
Arbrissel
Argentré-du-Plessis

Filtres pour la sélection de communes :

Sélectionner un département

Sélectionner un bassin versant

Sélectionner un pays

Sélectionner un EPCI

Votre nom *

Benjamin Duhamel

Votre adresse e-mail *

benjamin.duhamel@bretagne-environnement.org

Annexe IV

The screenshot shows a spreadsheet interface with a yellow warning banner at the top that reads "Avertissement de sécurité Les macros ont été désactivées. Activer le contenu". Below the banner, the spreadsheet has a header row with columns labeled "Titre", "Type de bibliographi", "Logo", "Description", "Auteurs", and "Collectio". A red text box in the first row of the data area contains the instruction: "Pensez à sélectionner les cellules à exporter (y compris la ligne de titres) avant de cliquer sur le bouton !". A button labeled "Exporter en csv" is located in the second row of the data area.

Titre	Type de bibliographi	Logo	Description	Auteurs	Collectio

Champs : Titre - Type de bibliographie - Logo- Description- Auteurs- Collection- Éditeur - Site Internet de l'éditeur- Fichier- Page Internet sur l'ouvrage- Nombre de pages- Langue- Année d'édition - Catégories -Disponibilité- Organismes associés- Source- Références - Droits- Mots clés- Couverture géographique - Communes- EPCI- Bassins versants- Parcs naturels (ou projets)- SAGE- Pays de Bretagne- Département

Annexe V

	A	B	C	D	E	F
1	travail sur le flux OAI					
2						
3	MOTS CLEFS TESTES	SUR TITRE	SUR MOTS-CLEFS OAI	SUR COUVERTURE GEO	STATUT	Correspondance thésaurus
4	boues	x			fini	Déchet organique
5	eau potable	x			fini	Eau potable
6	eau pure (programme)	x			fini	Bretagne Eau Pure
7	SCHEMA DIRECTEUR D'AMENAGEMENT	x			ok, mots clefs et cat.	SDAGE
8	GESTION DE L'EAU		x		ok, mots clefs et cat.	Gestion de l'eau
9	ASSAINISSEMENT		x		à reprendre	Assainissement
10	ASSAINISSEMENT NON COLLECTIF		x		à reprendre	Assainissement individuel
11	RESEAU HYDROGRAPHIQUE				ok, mots clefs et cat.	Réseau hydrographique
12	contrat (restauration et entretien)	x			fini	Contrat Restauration Entretien
13	qualité de l'eau	x	x		à vérifier	Qualité de l'eau
14	saumon (abondance de juvéniles)	x	x		à vérifier	Saumon Atlantique
15	huître	x	x		à vérifier	Mollusques
16	SAGE		x		fini	SAGE
17	bassin versant	x	x		en cours	Bassin versant
18	poissons migrateurs	x	x		en cours	Poissons grands migrateurs
19	zones humides	x	x		à vérifier	Zones humides
20	contrats (de) plan (état région)	x			en cours	
21	coquillage					Mollusques
22	oiseau					Oiseaux
23	sédimentation	x				Sédimentologie
24	vas*					
25	marais	x				Vasières et marais littoraux
26	vas* (cultures)					

MOTS CLEFS TESTES	SUR TITRE	SUR MOTS-CLEFS OAI	SUR COUVERTURE GEO	STATUT	Correspondance thésaurus
communautés de communes..	x			à faire	EPCI
aulne			x	à vérifier	SAGE (Schémas d'aménagement)
22 (département)			x	couverture géo	départements
NA			x	à traiter	
35 BRETAGNE			x (égal à)	Géo ok	départements
22 29 35 56 BRETAGNE			x (égal à)	Géo ok	région (ou supra)
BRETAGNE 22 29 35 56			x (égal à)	Géo ok	région (ou supra)
BRETAGNE 22 29 56 35			x (égal à)	Géo ok	région (ou supra)
22 BRETAGNE			x (égal à)	Géo ok & validation	départements
29 BRETAGNE			x (égal à)	Géo ok	départements
BRETAGNE			x (égal à)	Géo ok	région (ou supra)
56 BRETAGNE			x (égal à)	Géo ok	départements
LOIRE BRETAGNE			x (égal à)	Géo ok	région (ou supra)
COMMUNE	x			Géo ok	communes
LOIRE BRETAGNE	x				région (ou supra)
BASSIN VERSANT					bassins versants
BRETAGNE 22 56 29 35			x		région (ou supra)
35 56 22 29 BRETAGNE					région (ou supra)
BRETAGNE 22 35 29 56					région (ou supra)
BRETAGNE 35					départements

Titre : Enrichissement et amélioration d'une base documentaire dédiée aux études environnementales de Bretagne

Résumé : Mémoire de stage de Master 1 en sciences de l'information et de la communication, parcours documentation, sur les enjeux sous-jacents à la mise en place d'un géocatalogue des études sur l'environnement en Bretagne. Le contexte politique et technologique du projet ainsi qu'une vue plus large de son environnement autour de la diffusion et du partage de l'information, notamment scientifique et publique sera présenté. Les services offerts par le géocatalogue seront mis en valeurs, notamment la géolocalisation permettant de réaliser dynamiquement des états de connaissances. Le fonctionnement de l'alimentation du catalogue sera décortiqué étapes par étapes : *besoins – sources – collecte – traitement – diffusion de l'information*. Les enjeux documentaires seront soulevés comme le recours au protocole OAI-PMH, l'utilisation du jeu de métadonnées Dublin Core, la connexion aux archives ouvertes ainsi que la notion d'information durable. Finalement l'étude montre qu'il existe une tension entre la gestion de grosses volumétries de références, les procédures et pratiques automatiques faisant gagner du temps dans le partage de métadonnées et la nécessité d'un catalogage rigoureux et de qualité, indispensable à la qualité du service proposé (accès, géoréférencement).

Mots-clefs : géocatalogue, catalogage, géo-référencement, partage de l'information, diffusion des connaissances, accès du public à l'information, information durable, navigation à facettes, études, environnement, GIP Bretagne Environnement, métadonnées, Dublin Core, appropriation des TIC, état des connaissances, cartographie de l'information.