

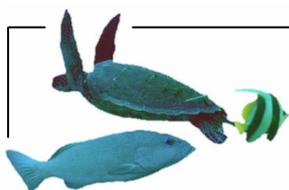


**ZoNéCo**

PROGRAMME D'ÉVALUATION DES RESSOURCES MARINES  
DE LA ZONE ÉCONOMIQUE DE NOUVELLE-CALÉDONIE



Institut de recherche  
pour le développement



## **INVENTAIRE ET INTEGRATION DES DONNEES SUR LES RESSOURCES**

**Proposition d'un outil informatique pour l'inventaire, la  
consultation et la restitution des données**

**RAPPORT**

Guénohé BOUVET

## REMERCIEMENTS

Je tiens à remercier l'UR-CoRéUs de l'IRD et le programme ZONECO de m'avoir confié ce travail extrêmement intéressant. Je remercie toutes les personnes qui m'ont aidé dans cette tâche à différentes étapes : Jocelyne Ferraris, Richard Farman, Jean-Pascal Torréton, Yves Lafoy, Julien Barde, Lionel Loubersac, Sonia Dossantos, Sébastien Dupont, Damien Buisson, Olivier de la Pommeraie, Alexandre Le Turc, Marion Amand, Denis Poignonnec, Serge Andréfouet, Michel Kulbicki, Isabelle Jollit, Sébastien Sarramegna, Nathalie Pari, Nicolas Boiteux, Michel Ménezo, Philippe Tirard, Pierre Desfontaines et Adrien Rivaton. Je tiens à m'excuser auprès des personnes et des organismes (DRN, Falconbridge, IRD, UNC) visités ou appelés à plusieurs reprises et parfois avec insistance au cours de mon travail afin de récupérer les informations. Que ces personnes qui se reconnaîtront sachent que c'était pour la bonne cause !

Je tiens également à remercier Christian Colin de m'avoir accueilli au centre IRD de Nouméa pendant la durée de ce contrat.

# SOMMAIRE

<b>LISTE DES TABLEAUX</b> .....	<b>5</b>
<b>LISTE DES FIGURES</b> .....	<b>5</b>
<b>INTRODUCTION</b> .....	<b>6</b>
<b>1. DEROULEMENT DU TRAVAIL</b> .....	<b>8</b>
<b>2. LES BESOINS DES UTILISATEURS ET PRODUCTEURS DE DONNEES</b> .....	<b>9</b>
2.1. GENERALITES .....	9
2.2. LES BESOINS DES UTILISATEURS .....	9
2.3. LES BESOINS DES PRODUCTEURS .....	9
<b>3. LES CHOIX DE DEVELOPPEMENT</b> .....	<b>10</b>
3.1. DESCRIPTION DU SYSTEME PRECEDENT .....	10
3.2. DEVELOPPEMENTS CHOISIS .....	11
<b>4. CONSTITUTION D'UN MODELE DE METADONNEES</b> .....	<b>12</b>
4.1. LES METADONNEES : DEFINITION .....	12
4.2. ADAPTATION DU MODELE A UN STANDARD .....	13
4.2.1. <i>Pourquoi s'aligner sur une norme ?</i> .....	13
4.2.2. <i>Description de la norme ISO 19115</i> .....	14
4.2.3. <i>Adaptation de la norme au contexte</i> .....	16
<b>5. DESCRIPTION DE L'OUTIL INFORMATIQUE DE GESTION ET CONSULTATION DES METADONNEES</b> .....	<b>18</b>
5.1. GENERALITES .....	18
5.2. DESCRIPTION DE L'OUTIL INFORMATIQUE .....	19
5.2.1. <i>Généralités</i> .....	19
5.2.2. <i>Description de la base de données</i> .....	19
5.2.3. <i>Organisation du site web</i> .....	20
5.2.4. <i>Hébergement du site et administration</i> .....	24
5.2.5. <i>Les langages informatiques utilisés</i> .....	26
5.2.6. <i>Quelques précisions sur la saisie des fiches de MD</i> .....	27

<b>6. TRANSFERT, ECHANGE ET PROTECTION DES DONNEES .....</b>	<b>28</b>
6.1. GENERALITES .....	28
6.2. PROPOSITION .....	29
6.2.1. <i>Une charte</i> .....	29
6.2.2. <i>Une accès contrôlé aux données</i> .....	32
<b>7. PERSPECTIVES.....</b>	<b>33</b>
7.1. INTEGRATION DES DONNEES ET RESTITUTION .....	33
7.2. DEVELOPPEMENT A PREVOIR .....	34
7.2.1. <i>Profils utilisateur et thésaurus</i> .....	35
7.2.2. <i>Une base bibliographique</i> .....	37
7.2.3. <i>Interactivité et MapServer</i> .....	38
<b>CONCLUSION.....</b>	<b>39</b>
<b>BIBLIOGRAPHIE .....</b>	<b>41</b>
<b>LISTE DES ANNEXES .....</b>	<b>43</b>

## **LISTE DES TABLEAUX**

Tableau 1 : Groupes de MD de la norme ISO 19115

Tableau 2 : MD du cœur de la norme ISO 19115

Tableau 3 : Description des MD retenues et créées

Tableau 4 : Description des MD retenues et créées pour décrire une personne

Tableau 5 : Fichiers pour la visite du site internet

Tableau 6 : Fichiers pour l'administration du site internet

Tableau 7 : Proposition de charte

## **LISTE DES FIGURES**

Figure 1 : Organisation schématique des fiches descriptives du modèle précédent.

Figure 2 : Organisation du système : centralisation et coordination.

Figure 3 : Principe de la norme ISO 19115

Figure 4 : Nouveau découpage géographique de la zone

Figure 5 : Page d'accueil du site WEB

Figure 6 : Module de consultation

Figure 7 : Résultat de la recherche

Figure 8 : Détails sur un résultat

Figure 9 : Module d'ajout d'une fiche

Figure 10 : Page de l'annuaire

Figure 11 : Résumé de l'organisation du site internet

figure 12 : Contrôle d'accès aux données

figure 13 : Thésaurus et relations ontologiques

## INTRODUCTION

Ce rapport dresse le bilan du travail effectué dans le cadre de l'opération ZONECO-2003 « Inventaire et intégration des données sur les ressources vivantes lagunaires ».

Il fait suite au premier travail d'inventaire des données réalisé par S. Bernard-Peyre en 1999 (Inventaire des données en milieu récifo-lagonaire. Ressource vivante et Méthode de travail, contrat ZONECO). En parallèle, une réflexion sur l'intégration de données multisources et la constitution d'une base de données géographiques étaient menées (Ferraris, Garrigue, Virly, 1998 ; Bouvet, 1999 ; Bourderie, 2000 ; Pautret, 2001, Ferraris, 2002) dans le programme ECOTROPE (Ecosystème Côtier Tropicaux sous influences terrigènes et anthropiques) puis dans l'Unité de Recherche CoRéUs (Communautés Récifales et Usages), au centre IRD de Nouméa. Le présent travail prend la suite logique de ces actions.

Les termes initiaux de référence de l'opération sont les suivants : « Poursuite de l'inventaire des données disponibles sur les lagons de Nouvelle-Calédonie ; analyse des besoins des partenaires, valorisation des fiches descriptives sur les données ; proposition d'un protocole de mise en forme, transfert et restitution de l'information ; application à une étude de cas sur une ressource-type ».

Etant donné le temps imparti pour mener à bien ces actions, et les besoins des producteurs et utilisateurs des données, l'opération a été réorientée en mettant l'accent sur la continuité du premier travail et la valorisation des fiches descriptives des données par un changement de support.

Le travail de S. Bernard Peyre ayant montré la diversité des données autant sur leur forme que sur leur mode d'acquisition, aucun protocole de mise en forme et de transfert des données n'a pu être proposé. Ces notions sont par ailleurs discutées en regard des développements récents réalisés dans le domaine des technologies des systèmes d'information. Des protocoles simples de restitution des données sont proposés. En revanche, les aspects mise en forme, transfert et restitution ont été traités au niveau des informations descriptives des données. Les besoins des partenaires ont également été étudiés au niveau des informations descriptives.

L'application à une étude de cas qui devait se faire sur une ressource type, les données de poissons, représentatives car diverses et variées, n'a pas pu se faire en raison des difficultés de récupération de l'information auprès des différents propriétaires de données, difficultés en

partie liées à une mauvaise compréhension du devenir et de l'utilisation du Système d'Information à élaborer.

La première phase des travaux réalisés s'articule essentiellement autour de la conception d'un outil pour constituer une base de **métadonnées**. Sous format informatique (type web), l'outil permet de reprendre les fiches descriptives des données en standardisant leur archivage et propose une solution pour poursuivre l'inventaire des données plus récentes.

Afin de répondre aux besoins des partenaires, notamment l'accès facilité aux données, un module de consultation de l'information et de téléchargement des données a été développé et mis en ligne sur internet.

Enfin, un protocole de récupération des données par les utilisateurs est proposé. Il devra faire l'objet d'une charte entre les partenaires ZONECO et son application devra être assurée par un administrateur, tel que le Service de Gestion et de Valorisation des données (SGVL) du programme au Service des Méthodes Administratives et de l'Informatique (SMAI), partenaire de ZONECO.

## **1. DEROULEMENT DU TRAVAIL**

L'opération a été effectuée entre mai 2003 et février 2004 dans le cadre d'un contrat de 10 mois à mi-temps.

Elle a commencé par une analyse du travail effectué par S. Bernard Peyre en 1999 et par la prise de contact avec des organismes confrontés à des besoins similaires en matière d'archivage et de restitution des données. L'IFREMER de Brest (programme REBENT) et de Sète (programme SYSCOLAG) ont été contactés et ont apporté leur conseils, notamment en suggérant l'utilisation d'un standard international pour optimiser l'archivage des informations descriptives des données.

Le modèle d'archivage des informations descriptives proposé en 1999 a alors été revu. A partir d'un nouveau modèle s'appuyant sur le standard international choisi, un outil informatique destiné à faciliter la collecte de ces informations a été développé. Il a ensuite été étendu pour la consultation par des utilisateurs.

Une maquette de l'outil informatique a été présentée lors d'une réunion avec les partenaires de ZONECO et plusieurs aspects ont été discutés. L'outil dans sa version de démonstration, installé sur les serveurs de l'IRD, a pu être testé en ligne par les personnes ayant participé à la réunion.

Grâce aux commentaires recueillis, l'outil a notamment été enrichi d'un module de récupération des données basé sur une charte d'utilisation des données rédigée en collaboration avec la SGVL du programme.

L'étude de cas concernant les protocoles de mises en forme, transfert et restitution de l'information qui devait se faire à partir des données de comptages visuels de poissons en plongée acquises dans des contextes financiers différents et par des personnes différentes n'a pas été réalisée puisqu'il a été impossible de récupérer ces données.

## 2. LES BESOINS DES UTILISATEURS ET PRODUCTEURS DE DONNEES

### 2.1. Généralités

La première étape de toute stratégie de gestion de l'information consiste à commencer par documenter les ressources d'information-clé et les rendre visibles tant au sein qu'en dehors de l'organisation (programme MADAME<sup>1</sup>, 2000). L'ensemble des renseignements nécessaires à l'**archivage** et à la **diffusion** des données doit être collecté.

### 2.2. Les besoins des utilisateurs

**Les utilisateurs** ont des profils différents : chercheurs, étudiants, décideurs, professionnels, grand public, et donc des besoins différents. Néanmoins, le point commun de leurs exigences reste **l'accès facilité à l'information**. Il implique une prise de connaissance et une récupération des données simple et rapide.

En revanche, la forme des données récupérées peut différer : de l'information brute à l'information « digérée » et présentée. En ce qui concerne les informations sur les données, il semble également nécessaire de proposer plusieurs niveaux en fonction des besoins de l'utilisateur.

### 2.3. Les besoins des producteurs

**Les producteurs** de données ont eux aussi leurs exigences :

- respect de la propriété intellectuelle et financière ;
- respect de la confidentialité ;
- souhait de connaître les utilisateurs de leurs données.

Et leurs craintes :

---

<sup>1</sup> MADAME : Méthodes d'Accès aux Données et Aux Métadonnées en Europe

- vol ;
- ou mauvaise utilisation de leurs données.

➔ Ces besoins impliquent donc le développement d'un système d'archivage et de consultation :

- souple, évolutif pour pouvoir s'adapter à l'utilisateur ;
- simple, rapide et facile d'accès ;
- contrôlé et protégé.

### 3. LES CHOIX DE DEVELOPPEMENT

#### 3.1. Description du système précédent

Le recensement des données sur les ressources vivantes lagunaires effectué en 1999 par Stéphane Bernard-Peyre a permis la constitution de 118 fiches descriptives des données (Bernard-Peyre S. 1999)

Ces fiches sont de deux types : *fiche descriptive de l'étude* et *fiche descriptive du lot de données*. La première décrit le cadre général : l'étude, et la seconde décrit les données elles-mêmes : le lot. A une étude peuvent se rapporter plusieurs lots de données. L'organisation schématique est exposée sur la figure 1.

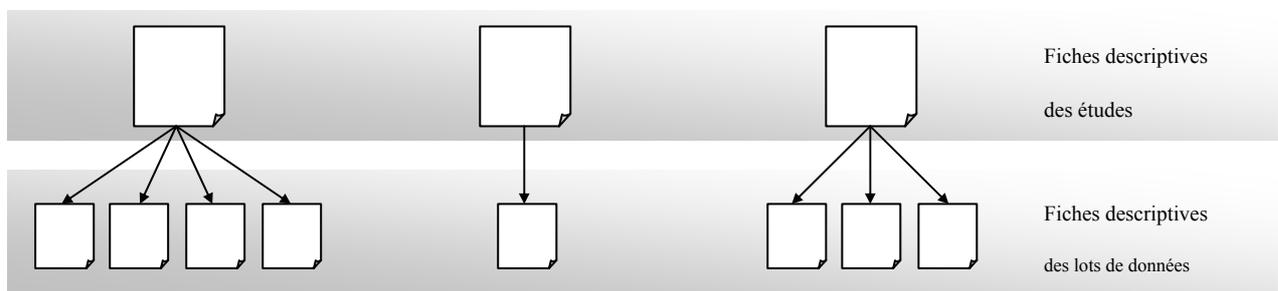


Figure 1 : organisation schématique des fiches descriptives du modèle précédent.

Un exemple de ces fiches est donné en annexes 1 et 2.

Elles ont été constituées sous un format Microsoft Word qui ne répond pas aux besoins des utilisateurs en matière d'accès facilité à l'information (accès à distance, souplesse, évolutivité).

Un nouveau système doit donc être développé.

### **3.2. *Développements choisis***

Une **base de données contenant les informations descriptives des données** permet l'archivage.

Une **interface de type web** permet :

- un accès interactif, simple et rapide aux informations,
- une évolutivité du système ;
- un contrôle et une protection des données.

➔ Le développement retenu est donc le suivant : **une base de métadonnées accessible sur internet.**

Le fonctionnement du système proposé peut se résumer par le schéma suivant (fig. 2). La base de métadonnées en est la clé de voûte :

- elle permet la visibilité des données à l'extérieur, valorisant ainsi le travail des producteurs de données et permettant aux utilisateurs de prendre connaissance des données existantes ;
- elle contient les informations nécessaires à la bonne utilisation des données ;
- elle contient les informations quant au droit d'utilisation des données.

Le système est géré par un organisme centralisateur. Il coordonne les relations entre les producteurs et les utilisateurs.

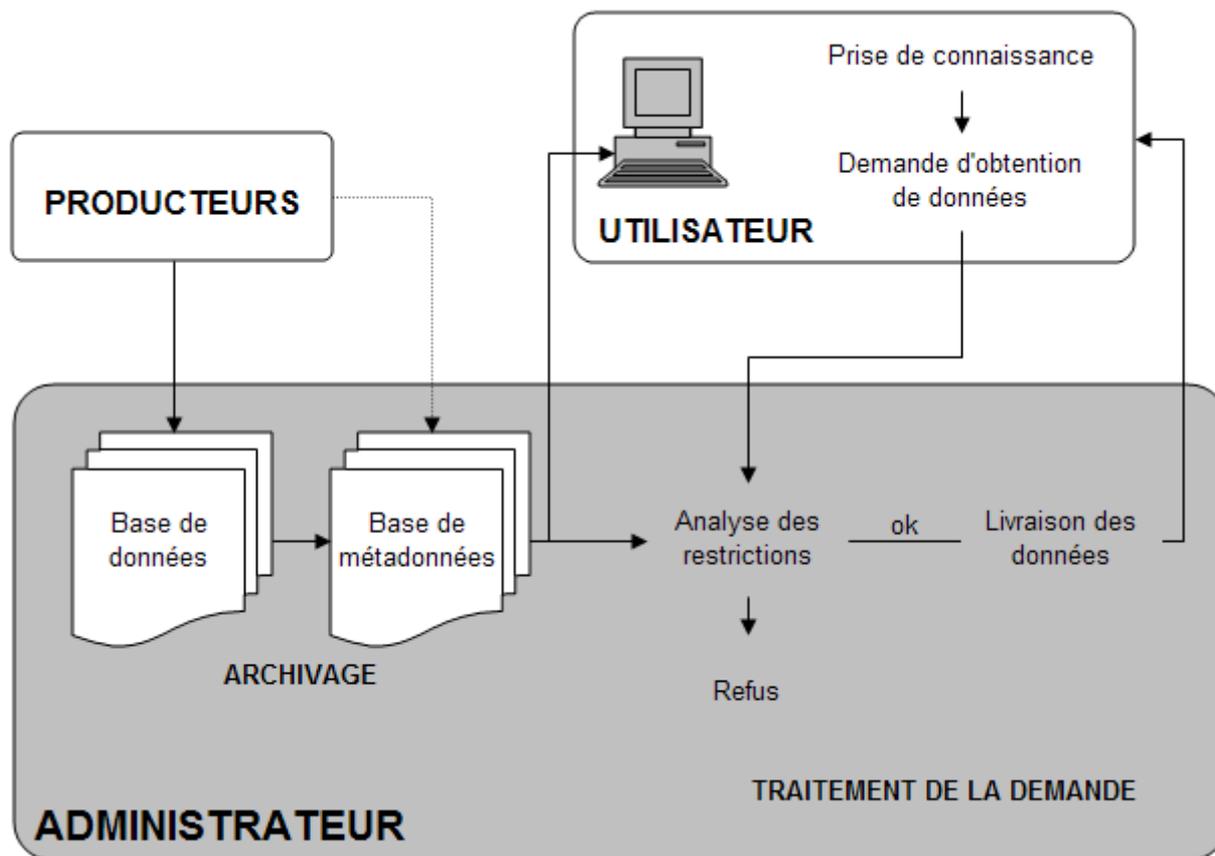


Figure 2 : Organisation du système : centralisation et coordination.

Les différents éléments du système sont décrits dans les parties 4 (modèle de métadonnées), 5 (interface utilisateur et base de métadonnées) et 6 (centralisation et coordination entre producteurs et utilisateurs).

## 4. CONSTITUTION D'UN MODELE DE METADONNEES

### 4.1. les metadonnées : définition

Dans son acception première, le terme métadonnée signifie « données sur les données ou données qui renseignent sur certaines données et qui permettent ainsi leur utilisation pertinente » (Bergeron, 1993). Les métadonnées ont un rôle d'identification et de spécification de la donnée « référence » et peuvent ainsi servir de base aux moteurs de recherche (Desconnets *et al.*, 2001).

Il s'agira par exemple de définir la date à laquelle les données ont été récoltées, par qui, comment, dans quelle zone géographique, dans quel but, qui en sont les propriétaires intellectuels, les propriétaires financiers, etc.

Les métadonnées correspondent à l'information et à la documentation qui rendent les données compréhensibles et partageables pour les utilisateurs dans le temps.

Nous pouvons distinguer différents niveaux de métadonnées diversement détaillées :

- les métadonnées pour l'**inventaire** : internes à une organisation et souvent rudimentaires, il s'agit le plus souvent d'un code ;
- les métadonnées pour la **découverte** : celles nécessaires aux utilisateurs pour trouver des données. En général, celles correspondant aux questions « qui ? », « quand ? », « quoi ? », « où ? » ;
- les métadonnées pour l'**usage** : celles qui permettent une utilisation correcte des données parce qu'elles documentent les données de façon complète et précise. Dans le cas des ressources vivantes lagonaires, il s'agira par exemple des informations sur la stratégie d'échantillonnage des données, la qualité des données, etc.

Il est important de constituer une base de métadonnées qui pourra répondre à ces trois besoins.

## ***4.2. Adaptation du modèle à un standard***

### **4.2.1. Pourquoi s'aligner sur une norme ?**

Il a été reproché à l'auteur du système précédent sur les métadonnées des ressources lagonaires de Nouvelle-Calédonie de ne pas s'être inspiré des normes existantes (Ferraris, com. pers.).

Les normes ISO en matière de métadonnées, récemment constituées, sont aujourd'hui utilisées par un nombre croissant d'organismes. La NASA, le BRGM et l'IFREMER notamment en font partie. La norme ISO 19115 apparaît comme étant un standard.

Le guide des bonnes pratiques en matière d'élargissement d'accès aux informations du secteur public (programme MADAME) préconise l'utilisation de cette norme.

Il semble donc assez judicieux de s'aligner sur cette tendance afin notamment de pouvoir assurer par la suite des connexions (import/export) avec d'autres bases de données.

#### 4.2.2. Description de la norme ISO 19115

La norme ISO 19115 comporte 407 métadonnées (MD), soit 407 renseignements différents pour décrire un maximum d'aspect du lot de données. La norme ISO laisse la possibilité de créer de nouvelles MD, en fonction des besoins. La figure 3 résume le principe de la norme. La norme ISO 19115 (draft) est téléchargeable à l'adresse suivante :

[www.ird.nc/COREUS/documents/iso19115.pdf](http://www.ird.nc/COREUS/documents/iso19115.pdf)

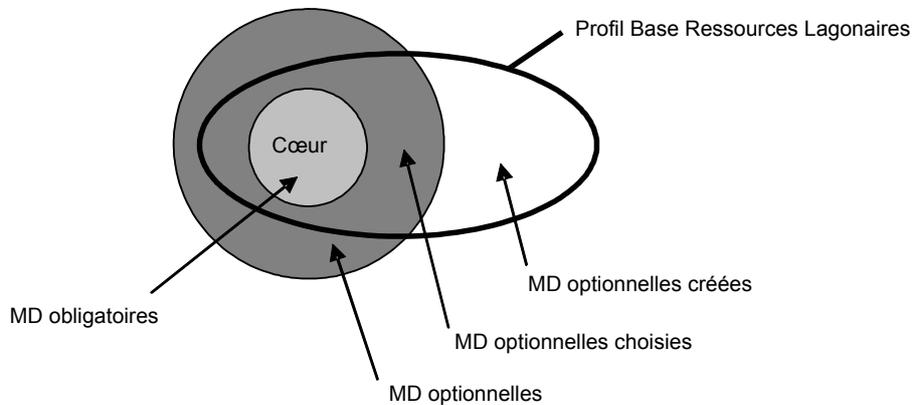


Figure 3 : principe de la norme ISO 19115

Les MD de la norme ISO 19115 sont rassemblées en plusieurs groupes. Le tableau 1 liste ces groupes et décrit leur intérêt.

Tableau 1 : Groupes de MD de la norme ISO 19115

Groupes	Exemples de contenu	Intérêt
1- information sur l'ensemble des MD	Date de création des MD   nom de l'auteur des MD   info de base sur les données	Permet de : - savoir qui a saisi les MD et quand - connaître le contexte global de provenance des données
2- identification des données	Résumé des données   aperçu en image   mots-clés   échelle maxi de représentation   utilisation des données   moyen de communication avec les utilisateurs des données   info	Permet d'identifier les données de façon unique
3- Protection des données	Limitation d'utilisation   limitation d'accès   classe de limitation sur les données ou les MD	Permet de : limiter l'utilisation ; ex : « ne pas utiliser pour la navigation », de respecter limiter l'accès aux données ou aux MD (propriété intellectuelle et confidentialité)
4- Qualité des données	Spécification de l'élément auquel s'applique la qualité des données   information sur la transformation des données (événement, date, auteur)   note de qualité sur : données manquantes, précision de positionnement dans le temps et dans l'espace, logique spatiale, adéquation entre les valeurs et leur domaine, logique temporelle, données thématiques, données quantitatives, données non quantitatives...   système d'évaluation de la qualité	Permet de connaître : - la qualité des données à tous les niveaux et le système d'évaluation de cette qualité - les modifications que les données ont subies
5- Mise à jour	Fréquence de mise à jour des données   domaine mis à jour	Permet de connaître la fréquence de mise à jour des données et l'étendue des données concernées
6- Représentation spatiale	Mécanisme de représentation   info sur les couches raster   info sur les couches vectorisées   info sur géoréférencement (point de calage, ...)   possibilité de géoréférencement   type d'objet géométrique	Permet d'avoir des info sur les couches raster, sur les couches vectorisées, sur le géoréférencement
7- Système de référence	Nom du système   projection   ellipsoïde   datum   azimuth   système de projection	Permet de connaître le système de référence
8- Contenu	Localisation du catalogue   description des attributs	Permet de connaître en détail la signification des attributs
9- Catalogue de présentation	Localisation du catalogue (contenant une légende pour un plan par exemple), directives de présentation	Permet de connaître les directives de présentation des données
10- Obtention des données	Personne à contacter   format   version	Permet de connaître la personne à qui s'adresser pour obtenir les données   sous quelle forme on peut obtenir les données
11- Information sur les MD optionnelles	Nom   définition   type   auteur	Permet de savoir qui a saisi les MD optionnelles et quand.
12- Information sur la chaîne de traitement	Langue utilisée dans la chaîne   schéma explicatif de la chaîne   fichier de développement logiciel pour la chaîne	Permet de savoir par quel traitement informatique les données ont été obtenues
13- Etendue spatiale et temporelle	Jeu de points définissant l'emprise géographique des données sous forme de polygone   emprise géographique des données sous forme de rectangle   date et heure d'acquisition des données   extension verticale   unité	Permet de connaître l'étendue spatiale et temporelle des données
14- Références bibliographiques et personne ressource	Titre   abstract   revue   date, etc. et téléphone   adresse   e-mail   site-web   ...	Permet de connaître les publications existantes et les personnes à contacter pour obtenir toute information complémentaire

Les domaines de valeurs pour chaque MD sont libres (entier, texte, etc.), codés par des listes de la norme ou régis par d'autres normes auxquelles il convient donc de se reporter.

En pratique, il est très difficile de mettre en place une base de données constituée de la totalité des MD proposées par la norme. En revanche, il est indispensable de conserver les informations de base qui constituent le cœur de la norme (tableau 2).

Tableau 2 : MD du cœur de la norme ISO 19115

Nom du jeu de données
Date de référence du jeu de données
Personne ou organisme responsable du jeu de données
Localisation du jeu de données par 4 coordonnées
Langue utilisée dans le jeu de données
Jeu de caractères utilisé dans le jeu de données
Domaine traité par le jeu de données
Résolution spatiale du jeu de données
Résumé décrivant le jeu de données
Format de diffusion du jeu de données
Etendue verticale et temporelle des données
Type de données (vecteur, raster...)
Système de référence
Chaîne de traitement
Adresse web pour se procurer le jeu de données en ligne
Identifiant du fichier de MD
Nom du standard de MD utilisé
Version du standard de MD utilisé
Langue utilisée pour documenter les MD
Jeu de caractères utilisé pour documenter les MD
Personne responsable de la documentation des MD
Date de création des MD

### 4.2.3. Adaptation de la norme au contexte

La norme ISO 19115 a été adaptée au contexte des ressources vivantes lagunaires et aux exigences des utilisateurs et producteurs de données. Le cœur de la norme a été conservé.

Le nombre de MD par fiche sera au maximum de 129 dont 11 sont des commentaires/précisions libres pour d'autres MD.

Le tableau 3 décrit ces MD. Les MD précédées d'un cœur (♥) sont celles listées dans la norme ISO 19115. Celles précédées d'un plus (+) ont été créées. Elles sont organisées en 19 groupes. Les groupes « emprise temporelle », « domaine de profondeur » et « nom des personnes ressources » sont constitués de MD créées (+) mais elles trouvent leur équivalent proche dans la norme ISO. Elles ont été adaptées au contexte local. Afin de décrire les personnes ressource (propriétaires intellectuels, financiers, utilisateurs des données, etc.) associées à un jeu de données, 14 MD sont proposées (tableau 3).

Tableau 3 : description des MD retenues et créées

<p><b>Identifiant</b></p> <ul style="list-style-type: none"> <li>♥ fileIdentifiant : identifiant (numéro)</li> </ul> <p><b>Titre</b></p> <ul style="list-style-type: none"> <li>♥ dataTitle : titre du jeu de données</li> </ul> <p><b>Informations sur les métadonnées</b></p> <ul style="list-style-type: none"> <li>♥ MDlanguage : langue utilisée pour renseigner les MD</li> <li>♥ MDcharacterSet : jeu de caractère utilisé</li> <li>♥ MDcontactnom : nom de la personne qui a saisi les MD</li> <li>♥ MDcontactprenom : prénom de la personne qui a saisi les MD</li> <li>♥ MDdateStampDay : date de saisie des MD (jour)</li> <li>♥ MDdateStampMonth : date de saisie des MD (mois)</li> <li>♥ MDdateStampYear : date de saisie des MD (année)</li> <li>♥ MDstandartName : nom du standard de MD utilisé</li> <li>♥ MDstandartVersion : nom de la version du standard de MD utilisé</li> </ul> <p><b>Grand catégorie de domaine concernée</b></p> <ul style="list-style-type: none"> <li>♥ TopicCategory : grand domaine concerné</li> </ul> <p><b>Domaine concerné</b></p> <ul style="list-style-type: none"> <li>+ domaineListe : domaine concerné</li> <li>+ domaineAutre : autre domaine (si non présent dans la liste)</li> </ul> <p><b>Ressource concernée</b></p> <ul style="list-style-type: none"> <li>+ ressourceListe : ressource concernée</li> <li>+ ressourceAutre : autre ressource (si non présente dans la liste)</li> <li>+ ressourcePrecise : précision sur la ressource</li> </ul> <p><b>Informations principales</b></p> <ul style="list-style-type: none"> <li>♥ year : année de 1ère publication</li> <li>♥ citation : titre de la référence bibliographique principale</li> <li>♥ abstract : résumé de la référence bibliographique principale</li> <li>♥ description : description du jeu de données</li> <li>♥ purpose : buts recherchés lors de la récolte des données</li> <li>+ keywords1 : mot-clé n°1</li> <li>+ keywords2 : mot-clé n°2</li> <li>+ keywords3 : mot-clé n°3</li> <li>+ keywords4 : mot-clé n°4</li> <li>+ keywords5 : mot-clé n°5</li> <li>+ keywords6 : mot-clé n°6</li> <li>+ keywords7 : mot-clé n°7</li> <li>+ keywords8 : mot-clé n°8</li> <li>+ keywords9 : mot-clé n°9</li> <li>+ keywords10 : mot-clé n°10</li> <li>+ illustration : existence d'une image d'illustration (oui/non)</li> <li>♥ denominator : dénominateur de la fraction correspondant à l'échelle</li> <li>♥ distance : distance de référence (pour les images satellites notamment : taille du pixel)</li> <li>♥ distanceUnit : unité de la distance de référence (cm / m / km)</li> <li>♥ spatialRepresentationType : type de données (vecteur, raster, texte, tableau, images stéréoscopique, vidéo)</li> <li>♥ language : langue utilisée dans le fichier de données</li> </ul> <p><b>Informations sur l'échantillonnage</b></p> <ul style="list-style-type: none"> <li>+ outil : outil d'investigation</li> <li>+ outilCommentaire : commentaires / précisions sur l'outil d'investigation</li> <li>+ peche : type de pêche</li> <li>+ pecheCommentaire : commentaires / précisions sur le type de pêche</li> <li>+ frequenceValue : fréquence de l'échantillonnage</li> <li>+ frequenceDatum : unité de la fréquence de l'échantillonnage</li> <li>+ frequenceCommentaire : commentaires / précisions sur la fréquence de l'échantillonnage</li> <li>+ planValue : type de plan d'échantillonnage</li> <li>+ planCommentaire : commentaires / précisions sur le plan d'échantillonnage</li> <li>+ observationValue : unité d'observation (transect, point, zone, autre)</li> <li>+ observationCommentaire : commentaires / précisions sur l'unité d'observation</li> <li>+ localisationValue : technique de localisation</li> <li>+ localisationCommentaire : commentaires / précisions sur la technique de localisation</li> <li>+ statutCreateur : statut des personnes qui ont récolté les données (scientifique, bénévoles, etc.)</li> <li>♥ status : état de finition du jeu de données</li> </ul> <p><b>Programme et mission</b></p> <ul style="list-style-type: none"> <li>+ missionName : nom de la mission</li> <li>+ program : nom du programme</li> <li>+ organisation : organisme</li> </ul> <p><b>Emprise spatiale</b></p> <ul style="list-style-type: none"> <li>+ administrativeArea : zone administrative</li> <li>+ administrativeAreaCommentaire : commentaires / précisions sur la zone administrative</li> <li>♥ system : système de référence</li> <li>♥ preciseSystemName : nom précis du système de référence</li> <li>♥ northBoundLatitude : limite nord</li> <li>♥ westBoundLongitude : limite ouest</li> </ul>	<ul style="list-style-type: none"> <li>♥ eastBoundLongitude : limite est</li> <li>♥ southBoundLatitude : limite sud</li> <li>+ polygonAddress : localisation du fichier du polygone d'emprise</li> <li>+ nbreDeSommet : nombre de sommets du polygone d'emprise</li> <li>+ sommet01long : longitude du sommet n°1</li> <li>+ sommet01lat : latitude du sommet n°1</li> <li>+ sommet02long : longitude du sommet n°2</li> <li>+ sommet02lat : latitude du sommet n°2</li> <li>+ sommet03long : longitude du sommet n°3</li> <li>+ sommet03lat : latitude du sommet n°3</li> <li>+ sommet04long : longitude du sommet n°4</li> <li>+ sommet04lat : latitude du sommet n°4</li> <li>+ sommet05long : longitude du sommet n°5</li> <li>+ sommet05lat : latitude du sommet n°5</li> <li>+ sommet06long : longitude du sommet n°6</li> <li>+ sommet06lat : latitude du sommet n°6</li> <li>+ secteur : secteur(s) géographiques concernés</li> </ul> <p><b>Emprise temporelle ♥</b></p> <ul style="list-style-type: none"> <li>+ timeMinJMADay : date minimum dans le jeu de données (jour)</li> <li>+ timeMinJMAMonth : date minimum dans le jeu de données (mois)</li> <li>+ timeMinJMAYear : date minimum dans le jeu de données (année)</li> <li>+ timeMinHMSHour : heure minimum dans le jeu de données (heures)</li> <li>+ timeMinHMSMin : heure minimum dans le jeu de données (minutes)</li> <li>+ timeMinHMSSecond : heure minimum dans le jeu de données (secondes)</li> <li>+ timeMaxJMADay : date maximum dans le jeu de données (jour)</li> <li>+ timeMaxJMAMonth : date maximum dans le jeu de données (mois)</li> <li>+ timeMaxJMAYear : date maximum dans le jeu de données (année)</li> <li>+ timeMaxHMSHour : heure maximum dans le jeu de données (heures)</li> <li>+ timeMaxHMSMin : heure maximum dans le jeu de données (minutes)</li> <li>+ timeMaxHMSSecond : heure maximum dans le jeu de données (secondes)</li> <li>+ dateCommentaire : commentaires / précisions sur les dates et heures</li> </ul> <p><b>Domaine de profondeur ♥</b></p> <ul style="list-style-type: none"> <li>+ profUnit : unité de profondeur</li> <li>+ profDatum : référence des profondeurs</li> <li>+ profMax : profondeur maximum</li> <li>+ profMin : profondeur minimum</li> <li>+ profCommentaire : commentaire sur les profondeurs</li> </ul> <p><b>Précision de la localisation spatiale de données</b></p> <ul style="list-style-type: none"> <li>♥ absoluteExternalPositionAccuracyValue : précision sur 1</li> <li>♥ relativeInternalPositionAccuracyValue</li> <li>+ positionCommentaire</li> </ul> <p><b>Forme des données</b></p> <ul style="list-style-type: none"> <li>♥ dataFormat : format des données (numérique, papier, etc.)</li> <li>+ formatCommentaire : commentaires / précisions sur le format des données</li> </ul> <p><b>Chaîne de traitement</b></p> <ul style="list-style-type: none"> <li>♥ lineage : chaîne de traitement qui a conduit aux jeu de données archivés</li> </ul> <p><b>Mise en garde et protection des données</b></p> <ul style="list-style-type: none"> <li>♥ specificUsage : utilisation spécifique pour le jeu de données</li> <li>♥ useLimitation : limitation d'utilisation</li> <li>♥ accessConstraints : contrainte d'accès</li> <li>♥ useConstraints : contrainte d'utilisation</li> <li>♥ classification : classification de sécurité nationale (libre, top secret, etc.)</li> </ul> <p><b>Dictionnaire des données</b></p> <ul style="list-style-type: none"> <li>♥ includedWithDataset : présence d'un dictionnaire des données dans le jeu de données</li> <li>♥ featureCatalogueAdress : localisation du dictionnaire des données</li> </ul> <p><b>Nom des personnes ressources ♥</b></p> <ul style="list-style-type: none"> <li>+ intellectualIdnom1 : nom du propriétaire des données n°1</li> <li>+ intellectualIdprenom1 : prénom du propriétaire des données n°1</li> <li>+ intellectualIdnom2 : nom du propriétaire des données n°2</li> <li>+ intellectualIdprenom2 : prénom du propriétaire des données n°2</li> <li>+ intellectualIdnom3 : nom du propriétaire des données n°3</li> <li>+ intellectualIdprenom3 : prénom du propriétaire des données n°3</li> <li>+ financialOrgId : nom du propriétaire financier</li> <li>+ userIdnom1 : nom de l'utilisateur des données n°1</li> <li>+ userIdprenom1 : prénom de l'utilisateur des données n°1</li> <li>+ userIdnom2 : nom de l'utilisateur des données n°2</li> <li>+ userIdprenom2 : prénom de l'utilisateur des données n°2</li> <li>+ userIdnom3 : nom de l'utilisateur des données n°3</li> <li>+ userIdprenom3 : prénom de l'utilisateur des données n°3</li> <li>+ distributorIdnom : nom du distributeur officiel des données</li> <li>+ distributorIdprenom : prénom du distributeur officiel des données</li> </ul> <p><b>Protocole de récupération des données</b></p> <ul style="list-style-type: none"> <li>+ obtenir : classification pour l'obtention des données</li> <li>♥ onLine : adresse du site web où les données peuvent être téléchargées (si autre que ZONECO)</li> <li>♥ distributorFormat : format de distribution des données</li> </ul>
--	---

Tableau 4 : description des MD retenues et  
créées pour décrire une personne

<b>Description d'une personne ressource</b>	
♥	id : identifiant
♥	surname : prénom
♥	name : nom
♥	function : statut
♥	organisation : organisme
♥	voice : n° de téléphone
♥	facsimile : n° de fax
♥	email : adresse électronique
♥	linkage : adresse du site internet de l'organisme
♥	deliveryPoint : adresse postale
♥	city : ville
+	administrativeArea : province
♥	country : pays
♥	commentaire : commentaire

## **5. DESCRIPTION DE L'OUTIL INFORMATIQUE DE GESTION ET CONSULTATION DES METADONNEES**

### **5.1. Généralités**

L'outil informatique a été développé dans un premier temps pour faciliter la saisie de l'information sur les données (condition *sine qua non* à leur archivage et leur intégration). Il a ensuite été développé pour permettre la consultation des informations.

Afin de répondre aux besoins des utilisateurs de données, l'outil développé est un site internet couplé à une base de métadonnées. Le format site internet est devenu familier à de nombreux futurs utilisateurs. Le site a bien sûr pour finalité d'être utilisé à distance.

## 5.2. Description de l'outil informatique

### 5.2.1. Généralités

Le site internet est disponible à l'adresse suivante : [www.data.zoneco.nc](http://www.data.zoneco.nc). Il est constitué de deux modules principaux :

- un pour la consultation des fiches de métadonnées,
- un pour l'administration : ajout, modification, suppression de fiches.

Il est couplé à une base où est stocké l'ensemble des métadonnées.

Ce site web a été présenté une première fois en octobre 2003 et modifié en fonction des remarques. La recherche de fiche avec comme critère la zone administrative (Province Nord, Province Sud, Province des Iles) a été ajoutée. La recherche géographique sur carte a été détaillée en remplaçant un découpage de l'espace en 4 secteurs par un découpage en latitude/longitude en 20 secteurs (fig 4).

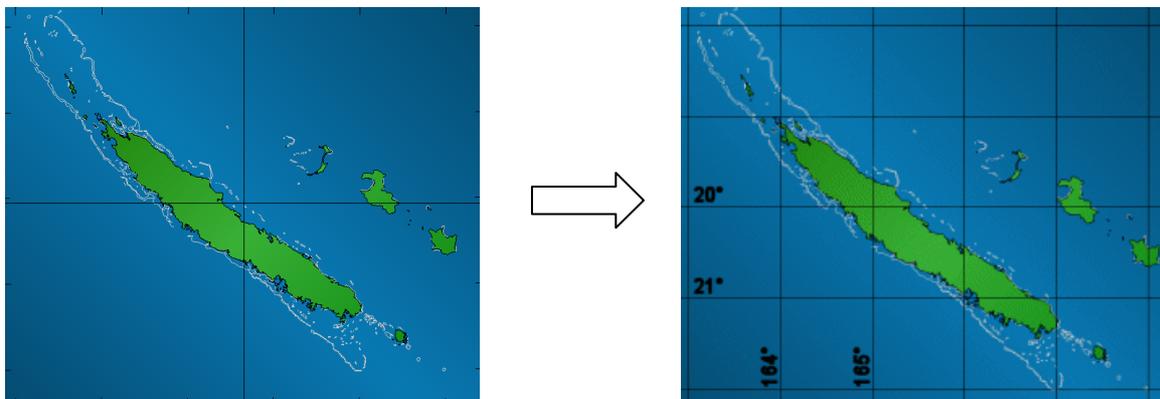


Figure 4 : nouveau découpage géographique de la zone

Enfin, suite à l'élaboration de la charte d'utilisation des données (chapitre 6.2.1), la possibilité de récupération des données en ligne a été implémentée.

### 5.2.2. Description de la base de données

La base de métadonnées est constituée de 2 tables : « infos\_generales » et « personnes », dans lesquelles sont respectivement stockées les métadonnées concernant le jeu de données

renseigné et un annuaire des coordonnées des personnes qui « gravitent » autour des données : propriétaires intellectuels, bailleurs de fonds, etc.

La table « infos\_generales » contient 129 champs, la table « personnes » en contient 14.

### 5.2.3. Organisation du site web

La page d'accueil propose à l'utilisateur (figure 5) :

- la recherche de fiches de MD
- la consultation de l'annuaire

Un bandeau vertical, sur la gauche de l'écran, permanent pendant la navigation, reprend les différentes icônes et permet de retrouver facilement la page d'accueil, le module de consultation des fiches, l'annuaire des personnes, le module d'administration.

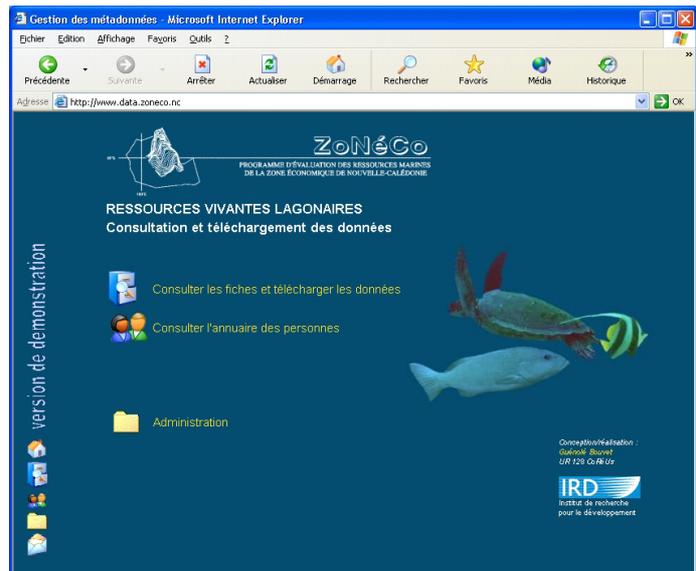


Figure 5 : page d'accueil

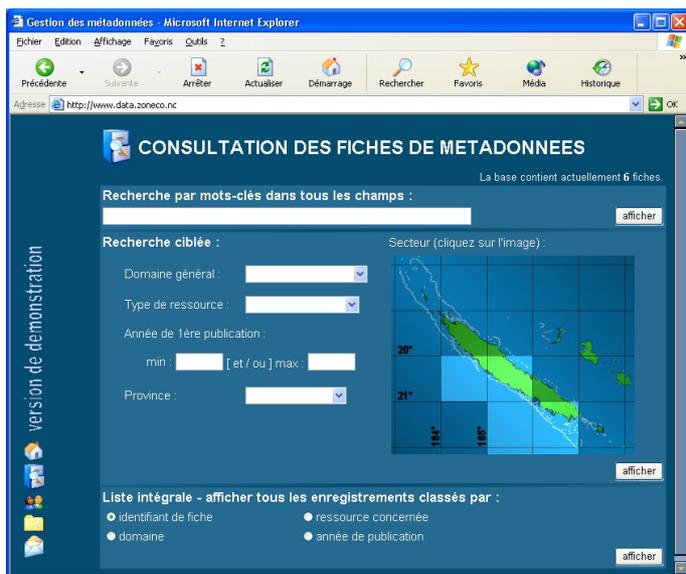


Figure 6 : module de consultation

fiches et de choisir un mode de classement.

Le résultat de la recherche (figure 7) s'affiche sous forme d'une liste sommaire (titre, auteur, année du jeu de données). Le nombre de fiches trouvées est précisé. Une imagerie symbolisant la ressource vivante concernée dans le jeu de données est affichée (ici, les deux premières fiches trouvées concernent les poissons). L'utilisateur peut obtenir des informations plus précises sur une fiche par simple clic-souris sur la loupe :

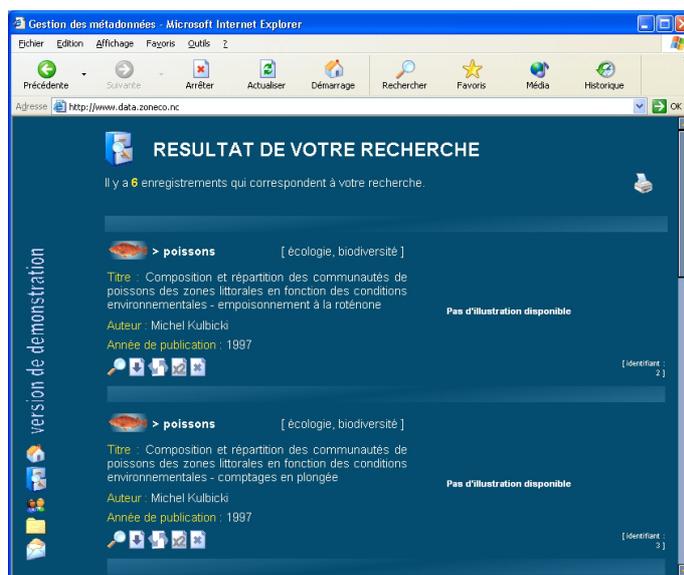


Figure 7 : résultat de la recherche

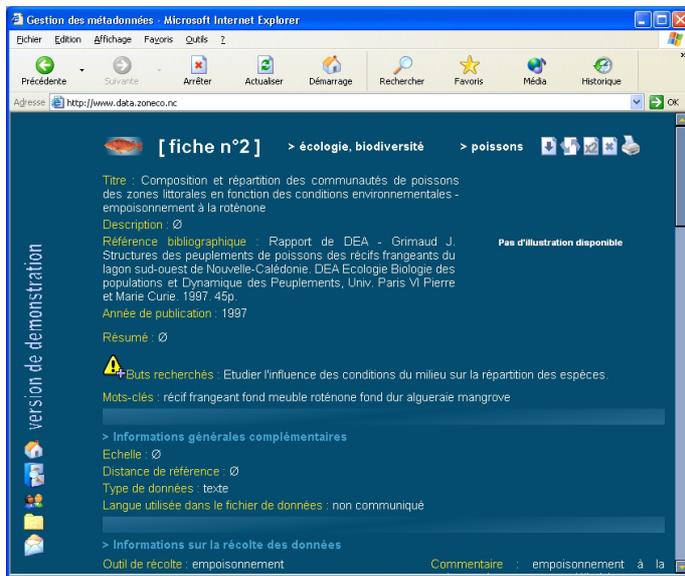


Figure 8 : détails sur un résultat

Il peut alors obtenir l'ensemble des informations au sujet du jeu de données qui l'intéresse (figure 8) : des précisions sur le plan d'échantillonnage, sur les techniques employées pour récolter les données, sur la zone concernée, savoir si d'autres personnes utilisent ces données, connaître les propriétaires, etc.

Les coordonnées du propriétaire ou de toute autre personne citée dans la fiche

sont disponibles en cliquant sur le nom de cette personne.

Si le visiteur veut accéder au jeu de données, il clique sur l'icône : . Selon les restrictions qui concernent le jeu de données, le visiteur est alors dirigé vers différentes pages. Si aucune restriction existe, il peut télécharger les données ; il est néanmoins informé des mises en garde pour une bonne utilisation des données.

L'administrateur peut modifier une fiche, la dupliquer ou la supprimer à partir des pages de consultation en cliquant sur les icônes :   .

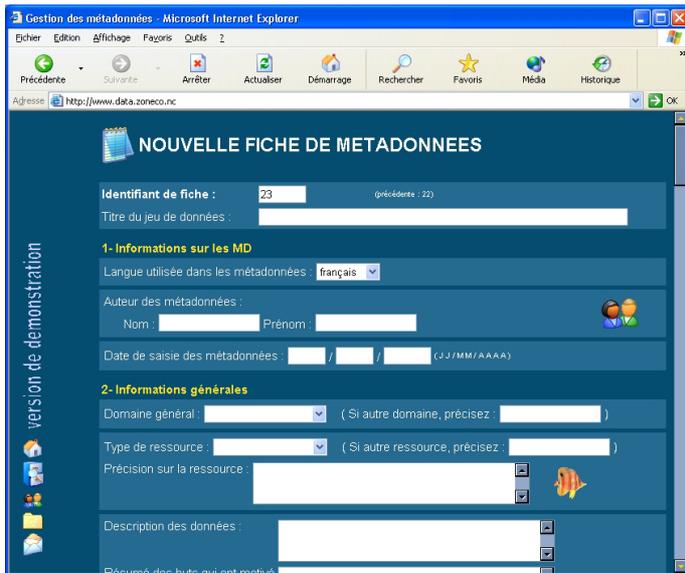


Figure 9 : module d'ajout d'une fiche

Le module d'ajout d'une fiche (figure 9) est disponible depuis la page d'accueil en cliquant sur l'icône : . L'administrateur peut alors alimenter la base de métadonnées et l'annuaire des personnes. La saisie des métadonnées est assistée autant que possible afin de limiter les erreurs de saisie. Des scripts<sup>2</sup> permettent de vérifier que certains champs obligatoires sont bien remplis, que les années comportent bien 4 chiffres, que les mois ne peuvent pas dépasser 12, que les heures ne peuvent pas dépasser 23, etc. Ce module est protégé par un accès restreint via identifiant et mot de passe.

les mois ne peuvent pas dépasser 12, que les heures ne peuvent pas dépasser 23, etc. Ce module est protégé par un accès restreint via identifiant et mot de passe.

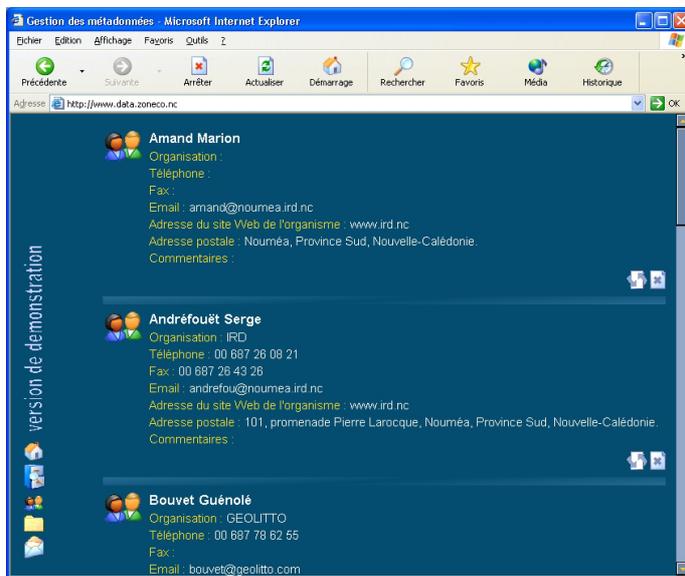


Figure 10 : page de l'annuaire

L'annuaire des personnes (figure 10) est consultable par tous mais la modification des coordonnées d'une personne ou la suppression sont accessibles seulement par l'administrateur.

<sup>2</sup> petits programmes en javascript

On peut résumer l'utilisation du site internet ainsi (cf. figure 11).

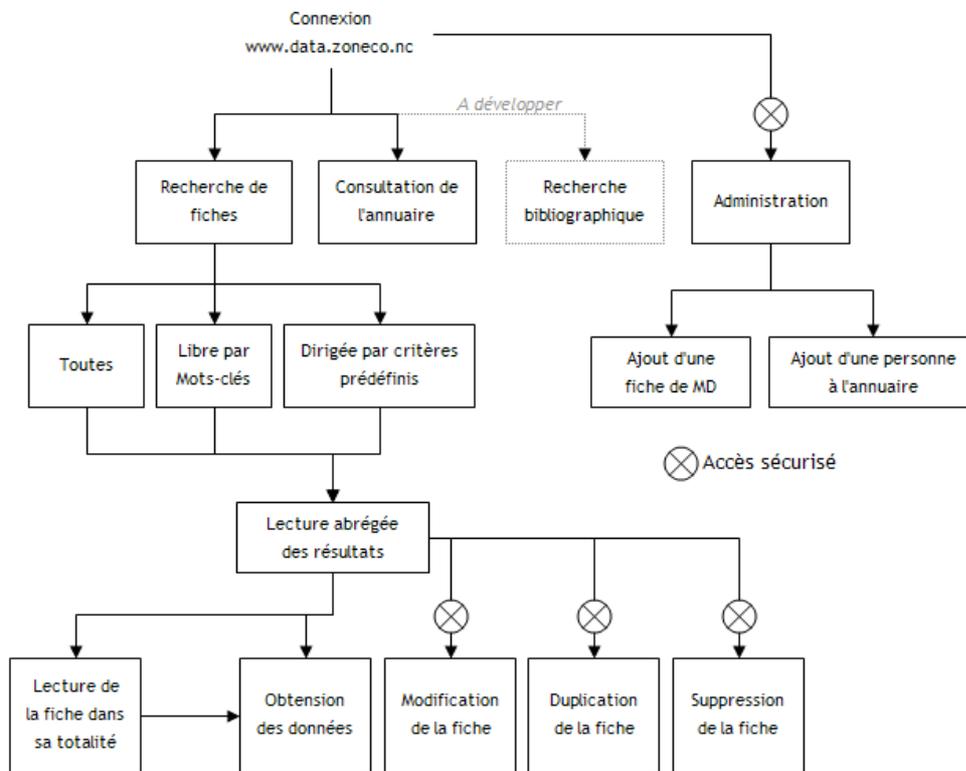


Figure 11 : résumé de l'organisation du site internet

#### 5.2.4. Hébergement du site et administration

Les fichiers informatiques correspondant au site web sont stockés dans un répertoire créé par Laurent Derrien (SMAI) et localisés sur un serveur du SMAI.

Les tableaux 5 et 6 listent ces fichiers. Le tableau 5 concerne les fichiers nécessaires pour l'affichage de l'interface destinée au visiteur et le tableau 6 les fichiers pour l'administration du site.

admin	File Folder
images	File Folder
style.css	2 KB CSS File
accueil.html	2 KB HTML Document
annuaire_personne.html	1 KB HTML Document
documentation.html	1 KB HTML Document
droite.html	1 KB HTML Document
gauche.html	2 KB HTML Document
index.html	1 KB HTML Document
mailto:webmaster.html	1 KB HTML Document
vide.html	1 KB HTML Document
affichage_detail.php	32 KB PHP File
affichage_liste_enregistrement_recherche.php	20 KB PHP File
affichage_liste_personne.php	3 KB PHP File
affichage_liste_totale.php	2 KB PHP File
affichage_personne.php	4 KB PHP File
consult_annuaire.php	4 KB PHP File
consult_enregistrement.php	19 KB PHP File
obtenir.php	8 KB PHP File
telechargement.php	1 KB PHP File

Tableau 5 : fichiers pour la visite du site internet

images	File Folder
style.css	2 KB CSS File
affichage_delete_enregistrement_ok.html	1 KB HTML Document
affichage_delete_personne_ok.html	1 KB HTML Document
affichage_enregistrement_ok.html	1 KB HTML Document
ajout_personne.html	4 KB HTML Document
ajout_enregistrement.php	34 KB PHP File
ajout_personne.php	5 KB PHP File
ajout_personne_sommaire.php	3 KB PHP File
alert_delete_enregistrement.php	1 KB PHP File
modif_enregistrement.php	59 KB PHP File
modif_personne.php	5 KB PHP File
protocole_ajout_enregistrement.php	6 KB PHP File
protocole_ajout_personne_sommaire.php	2 KB PHP File
protocole_delete_enregistrement.php	1 KB PHP File
protocole_delete_personne.php	1 KB PHP File
protocole_duplic_enregistrement.php	6 KB PHP File
protocole_modif_enregistrement.php	7 KB PHP File
protocole_modif_personne.php	2 KB PHP File

Tableau 6 : fichiers pour l'administration du site internet

La base de métadonnées est également stockée sur un serveur du SMAI.

Le site et la base sont gérés par un administrateur. Il possède les identifiants et mots de passe permettant :

- l'accès aux pages d'administration (ajout, modification, effacement de fiche de métadonnées) ;
- la modification des fichiers du site web par FTP ;
- l'administration de la base de métadonnées.

Je suis actuellement l'administrateur. Qui prendra la suite ? Qui assurera la maintenance, l'alimentation de la base, la coordination, le contrôle ? Si le système est jugé pertinent, une implication forte des partenaires est nécessaire pour optimiser et poursuivre le développement. Le développement du produit proposé et son utilisation, nécessite d'une part, l'implication d'un informaticien, et, d'autre part, la mise en place d'une dynamique impliquant fournisseurs et utilisateurs de l'information. Le programme ZoNéCo, qui favorise l'émergence d'une telle dynamique, pourrait justifier la mise en opération du système en s'appuyant sur une analyse coût-bénéfice entre l'investissement dans le développement et la maintenance informatique (et donc la rémunération d'un informaticien) et la perte due à une mauvaise gestion de données collectées avec des investissements intellectuels et financiers conséquents.

## 5.2.5. Les langages informatiques utilisés

Le site internet et les interactions avec la base de métadonnées sont programmés par environ 9000 lignes de codes. Quatre langages sont utilisés.

Le **HTML** (Hyper Text Markup Language) permet la programmation de l'aspect du site et de la navigation entre les différentes pages. Un exemple de la syntaxe est donné ci-dessous (affichage du logo ZONECO dans un tableau sans bordure).

```
<body>
<table border=0>
<tr><td rowspan=3 height=20%></td></tr>
</table>
```

Le **Javascript (JS)** sert à rendre plus dynamique l'interface : la sélection des zones géographiques sur la page de consultation des fiches est notamment programmée dans ce langage. Le Javascript sert aussi aux procédures de vérification de syntaxe des métadonnées saisies dans chaque nouvelle fiche. Il permet par exemple de vérifier qu'une date est bien dans un format JJ/MM/AAAA ou que la profondeur maximum saisie est effectivement supérieure à la profondeur minimum et qu'elle ne comporte pas de caractères qui ne seraient pas des chiffres. Un exemple de la syntaxe est donné ci-dessous (fonction de vérification de saisie des champs jour, mois et année).

```
function VerifJour(jour)
{
  with(window.document.saisieMD) //nom formulaire
  {
    if(!VerifNb(jour.value,"Exprimez les jours en chiffres SVP"))
    {
      jour.focus();
      jour.select();
      return false;
    }
    if(jour.value.length != 2 )
    {
      alert("Exprimez les jour avec 2 chiffres");
      return false;
    }
    if(jour.value > 31 )
    {
      alert("Erreur dans le jour");
      return false;
    }
  }
  return true;
}
```

Le **MYSQL** permet d'interroger la base de métadonnées et de l'alimenter avec des nouvelles fiches. Un exemple de la syntaxe est donné ci-dessous (extraction des n° des fiches de métadonnées contenues dans la base).

```
SELECT fileIdentifieur FROM $theTable
```

Le **PHP** permet la construction de page à la volée à partir d'un canevas existant développé en HTML et JS. Les informations issues de la base y sont insérées en fonction des actions effectuées par le visiteur du site web qui peut alors les visualiser. Le canevas PHP est un « squelette » sur lequel les « muscles » (les MD de la base) viennent se greffer sur demande. Un exemple de la syntaxe est donné ci-dessous (tests et formulation de requête pour la base de métadonnées).

```
$k=0;
while($k<20) {
    if ($preconditionSecteur!='' AND $ressource=='' AND $domaine=='' AND $administrativeArea=='') {
        $conditionSecteur=$preconditionSecteur ;
    }
    if ($preconditionSecteur!='' AND ($ressource!='' or $domaine!='' OR $administrativeArea!='' )) {
        $conditionSecteur=" AND ($preconditionSecteur)" ;
    }
    $k++ ;
}
```

Le PHP et le MySQL sont interprétés côté serveur ; le HTML et le Javascript, côté client.

La programmation a été réalisée avec l'éditeur de texte de Windows : Bloc-Note directement en code. Certains codes JS sont inspirés de codes récupérés sur les sites de développement du web :

[www.toutjavascript.com](http://www.toutjavascript.com) et [www.editeurjavascript.com](http://www.editeurjavascript.com).

Le développement des scripts PHP/MySQL a été facilité par l'utilisation du logiciel EasyPHP. Il permet l'interprétation de ces langages et la création/gestion de bases de données sans nécessiter l'utilisation d'une machine serveur.

### **5.2.6. Quelques précisions sur la saisie des fiches de MD**

#### **Opérateur de saisie**

Pour l'instant, la personne en charge de la saisie des MD concernant les jeux de données est l'administrateur. Le produit actuel nécessite donc un intermédiaire informé entre les producteurs de données et la base de MD. Le module de saisie des MD devra être optimisé afin de permettre aux producteurs de données de saisir eux-même les MD concernant leurs jeux. A ce sujet, quelques recommandations figurent dans le chapitre 6 : perspectives.

## **Durée**

Le temps nécessaire à la saisie d'une fiche de MD est d'environ 15 minutes lorsque les données sont bien connues par leur producteur.

## **Recommandation**

Il est important de remplir tous les champs dans la mesure du possible afin de pouvoir ensuite permettre des recherches efficaces via le module de consultation.

# **6. TRANSFERT, ECHANGE ET PROTECTION DES DONNEES**

## **6.1. Généralités**

La suite logique du travail de collecte des MD est l'intégration des données elles-mêmes.

Le transfert des données vers une base centralisée est en train de devenir une notion obsolète. L'évolution des technologies de partage de l'information via Internet permet déjà un regroupement virtuel des données. L'avantage est certain : le maintien de la gestion interne des données par chaque organisme.

Mais ces technologies impliquent le développement de bases de données dans chaque organisme producteur, ce qui constitue un frein important au développement de ce type de système d'information. En effet, l'implication des partenaires doit être très forte. L'investissement peut s'avérer trop coûteux.

La solution d'un regroupement des données sur un serveur commun reste donc d'actualité.

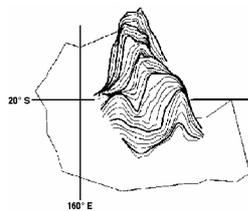
Qu'elles soient transférées vers une base de données communes ou bien qu'elles restent partagées, se pose le problème de leur protection. En effet, personne n'est prêt à partager sa propre information sans garanties (Stoimenov *et al.*, 2003). Les difficultés rencontrées pour la récupération des données de comptages visuels de poissons en plongée témoignent de ce problème. Avant toute chose, il est nécessaire d'élaborer des protocoles officiels de transfert et d'échange de données afin de fournir aux producteurs les garanties dont ils ont besoin.

## 6.2. Proposition

### 6.2.1. Une charte

Il s'agirait de faire signer aux partenaires une charte simple. Un exemple est proposé ci-dessous (tableau 7). Cette charte définit les protocoles de transfert et récupération des données. Elle a été rédigée en collaboration avec Sonia Dossantos (SGVL). Cette charte doit être avalisée par les différentes cellules de ZONECO puis officialisée.

Tableau 7 : proposition de charte



**ZoNéCo**

PROGRAMME D'EVALUATION DES RESSOURCES MARINES  
DE LA ZONE ECONOMIQUE DE NOUVELLE-CALEDONIE

## CHARTRE D'UTILISATION DES DONNEES

### PREAMBULE

La présente charte vise à réglementer la diffusion des données et à protéger leurs propriétaires dans le cadre du programme ZonéCo.

Deux types de données sont définies : les données libres d'accès et les données pouvant faire l'objet de restrictions.

Par cette convention, les signataires se sont engagés à mettre à disposition des autres partenaires leurs données acquises dans le cadre de ZonéCo.

Cette mise à disposition de données impose l'établissement d'un certain nombre de règles d'utilisation.

En vertu de ce qui précède,

**Les signataires de la présente charte, ci-dessous désignées par " partenaires" ont convenu et arrêté ce qui suit :**

#### ARTICLE 1. OBJET DE LA CHARTE

La présente charte a pour objet de définir de fixer les modalités de mise à disposition et d'utilisation des données par les partenaires.

---

## **ARTICLE 2. NATURE DES DONNEES**

### **2.1 LES DONNEES POUVANT FAIRE L'OBJET DE RESTRICTION D'UTILISATION**

#### **2.1.1 Acquisées dans le cadre d'opérations du programme ZoNéCo**

Ces données sont transmises à la SGVL par leur propriétaire intellectuel. Il s'engage à fournir l'ensemble des métadonnées nécessaires à une bonne utilisation des données ainsi que les mises en garde et les restrictions d'utilisation des données.

#### **2.1.2 Acquisées hors ZoNéCo par les partenaires**

Toute personne désirant utiliser ces données en fait la demande au propriétaire et négocie avec lui les termes de leur utilisation.

### **2.2 LES DONNEES LIBRES**

Ces données ont fait l'objet de publication. Elles sont en accès et utilisation libre. Cependant, le demandeur recevra, à leur livraison, les mises en garde nécessaires à une bonne utilisation.

## **ARTICLE 3. RESPONSABILITE**

La responsabilité des partenaires ne pourra être retenue :

- ◆ en cas d'imperfection des données ou de manque de fiabilité du produit,
- ◆ en raison d'imprécisions ou d'absence de certains composants du fichier,
- ◆ pour défaut de compatibilité des fichiers avec le système informatique de l'utilisateur (tiers sous-traitant),
- ◆ pour tout autre motif autre que la défaillance reconnue du support livré (dans cette hypothèse, les partenaires s'engagent à effectuer une nouvelle livraison du même fichier).

## **ARTICLE 4. MISE A DISPOSITION**

### **4.1 MISE A DISPOSITION DES DONNEES POUVANT FAIRE L'OBJET DE RESTRICTION**

#### **4.1.1 Modalités**

##### **4.1.1.1 Mise à disposition entre partenaires**

Les données sont mises à disposition des partenaires qui en font la demande, pour une utilisation en interne et/ou pour ses besoins propres. A cette occasion, les mises en garde pour une utilisation correcte des données sont communiquées.

##### **4.1.1.2 Mise à disposition à des tiers**

Toute personne désirant utiliser ces données en fait la demande à la SGVL qui contacte leur propriétaire financier et intellectuel afin d'obtenir leur accord pour la diffusion. Si le propriétaire intellectuel est favorable, les propriétaires financiers sont consultés.

## Tableau 7 suite : proposition de charte

---

En cas d'avis favorable, les données sont diffusées au demandeur après signature d'un acte d'engagement entre le demandeur et la SGVL.

En cas de refus des propriétaires financiers ou intellectuels, les données ne sont pas diffusées.

### **4.1.2 Droits d'utilisation des données**

Ils sont limités notamment par les propriétaires intellectuels des données et les propriétaires financiers.

## **4.2 MISE A DISPOSITION DES DONNEES ACQUISES HORS ZONECO**

### **4.2.1 Modalités**

Ces données sont mises à disposition du partenaire qui fait une demande auprès du partenaire propriétaire.

### **4.2.2 Droits d'utilisation des données acquises hors ZoNéCo**

Ils sont définis par le propriétaire des données.

## **4.3 FORME DE MISE A DISPOSITION**

Sauf mention contraire, les données composant les fichiers ne font l'objet ni d'un traitement méthodique leur assurant une structure topologique stricte, ni d'un complètement systématique garantissant une image exhaustive et régulière des données partageables.

En fonction de leurs possibilités techniques, les partenaires fixent conjointement les supports informatiques et les formats des fichiers à fournir.

Chaque livraison sera accompagnée d'un courrier mentionnant la référence des données transmises au catalogue des données géographiques.

Lors d'échange de données entre partenaires, la demande et l'acceptation de données entraînent de facto réaffirmation de l'engagement du bénéficiaire à respecter la présente charte ; mention en sera faite dans le courrier cité à l'alinéa précédent.

Lors d'une mise à disposition de données à des tiers non-signataire de la présente charte, copies du ou des imprimés remplis et signés, cités à l'article 4, seront jointes au courrier cité à l'alinéa précédent.

**ANNEXE 1 :**

**UTILISATION DE DONNEES**

Je soussigné (nom et raison sociale) m'engage à respecter les termes de la charte d'utilisation des données ZoNéCo pour la protection des données suivantes :

Liste des données transmises

Fait à Nouméa, le

---

**6.2.2. Une accès contrôlé aux données**

Une fois les données recensées et localisées et une charte signée, le développement du site web devra être prolongé afin de permettre aux utilisateurs d'obtenir les données qui les intéressent présentées de la façon la plus compréhensible pour eux. Pour le moment, l'accès aux données brutes est déjà un enjeu de taille.

Les protocoles proposés sur le site web sont issus de la proposition de charte d'échange et de protection des données.

Ils se résument ainsi (fig. 12) :

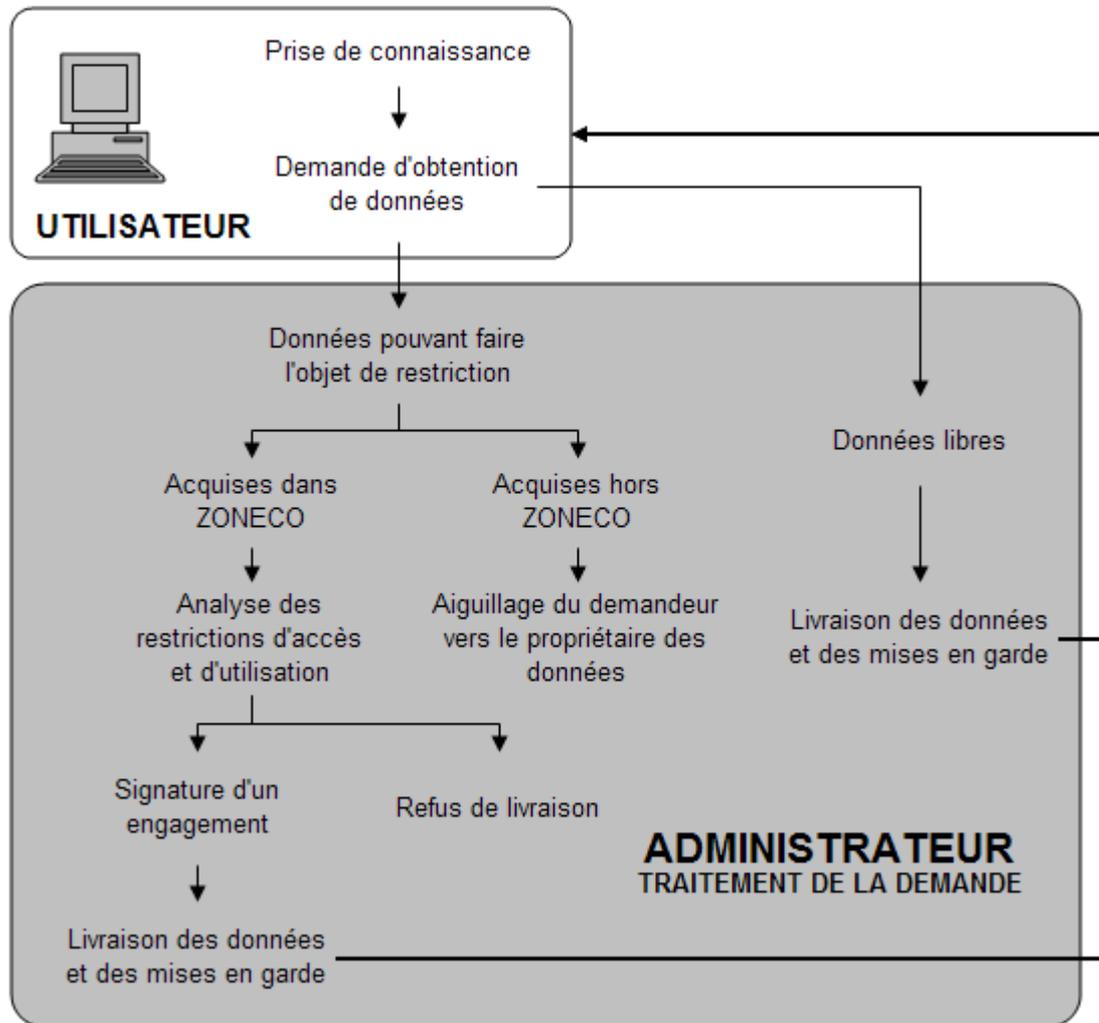


figure 12 : contrôle d'accès aux données

## 7. PERSPECTIVES

### 7.1. *Intégration des données et restitution*

L'outil développé ici ne propose pas de protocole de restitution de données à partir de données brutes ou standardisées. Cette étape est semble-t-il prématurée face aux priorités de constitution d'une base de métadonnées. Il s'agit par ailleurs d'une problématique vaste et qui demande un investissement important de la part des partenaires de ZONECO et de l'ensemble des utilisateurs potentiels du système.

Cependant, pour parer au plus pressé, il est tout à fait possible d'enregistrer dans la base de données, des informations plus compréhensibles que des données brutes : des rapports simplifiés ou des articles de vulgarisation scientifique. On peut envisager également d'utiliser des « moulinettes » spécifiques pour des jeux de données particuliers. Certaines moulinettes existent déjà : c'est le cas du logiciel REACT développé par la CPS pour les comptages visuels de poissons en plongée (Labrosse et al., 2001).

Mais, il est inutile d'essayer de constituer une métabase si avant tout il n'existe pas une base de métadonnées.

Le système imaginé dans le cadre du programme SYSCOLAG fait intervenir plusieurs niveaux de publications des données. Toutes les données référencées ne sont pas disponibles, même si leurs fiches de MD existent. Les niveaux sont les suivants :

- un niveau **interne** permettant la prise de connaissance des données par les personnes fortement impliqués dans la constitution du système d'information ;
- un niveau « **extranet** » permettant l'échange de données en comité restreint (les partenaires de SYSCOLAG) et destiné à susciter les collaborations pluridisciplinaires ;
- un niveau « **internet** » permettant au grand public d'obtenir des données.

Les données passent progressivement de niveau en niveau pour finir sur le niveau « internet ». C'est lors de leur passage sur le niveau « extranet » que des applications sont imaginées et testées. Une fois validées, elles passent sur le niveau « internet ». Dans cette progression, les données évoluent donc d'un niveau « brut » vers une forme plus ergonomique destinée à la communication. Mais cette progression est lente car elle nécessite un investissement fort de la part des partenaires et en tout premier lieu la récupération de l'ensemble des données brutes et de leur archivage.

## ***7.2. Développement à prévoir***

Les développements futurs doivent émaner d'une réflexion profonde sur le système d'information et ses finalités (observatoire des pêches) impliquant tous les acteurs concernés. Néanmoins, quelques développements sont ici proposés.

## 7.2.1. Profils utilisateur et thésaurus

### Optimiser la consultation

Pour améliorer l'outil de recherche, il faut le rendre **intelligent**, c'est à dire capable de s'adapter.

L'interface doit renvoyer une réponse qui correspond à la question que se pose l'utilisateur.

Les dangers sont de deux types différents :

- la réponse est trop complexe pour l'utilisateur, les MD renvoyées sont trop nombreuses et trop techniques ;
- la recherche est mal conduite par le système car le vocabulaire qu'il utilise est trop restreint.

Le premier point nécessite l'utilisation de **profil d'utilisateur**. La réponse du système s'adapte au niveau de l'utilisateur : scientifique, professionnel, grand public, enfant, et pourquoi pas à sa discipline : économiste, écologue, professionnel de la santé, gestionnaire du littoral, etc.

Le second point implique une **maîtrise du vocabulaire**. Un même objet ne s'appelle pas de la même façon pour tout le monde et un même mot ne désigne pas un même objet pour tout le monde. Les conséquences sont très dommageables :

- les réponses du systèmes ne correspondent pas à l'attente de l'utilisateur ;
- le système n'apporte pas de réponse à une question simple car il ne la comprend pas ;

Dans les deux cas, c'est un échec.

Les solutions sont de deux types :

- diriger la recherche en utilisant un découpage « naturel » de la réalité en éléments emboîtés ;
- dans le cadre d'une recherche libre, utiliser ce même découpage pour renvoyer une réponse adaptée. Si l'élément cherché par l'utilisateur ne correspond à aucun résultat direct, on le redirige vers les résultats plus généraux ou associés.

Ce découpage naturel est un **thésaurus**. Il consiste en une représentation de la réalité sous la forme d'une organisation hiérarchique : une arborescence (fig 13). Il a l'avantage de lister de façon exhaustive l'ensemble des éléments qui constitue la réalité.

Mais ce découpage est réducteur : il ne convient pas à tous les utilisateurs car chacun a une vision de l'organisation de la réalité qui dépend profondément du domaine dans lequel il travaille.

Ce découpage peut être amélioré de façon très intéressante par la mise en place de relation de comportement des divers éléments les uns par rapport aux autres.

C'est l'ontologie (fig 13).

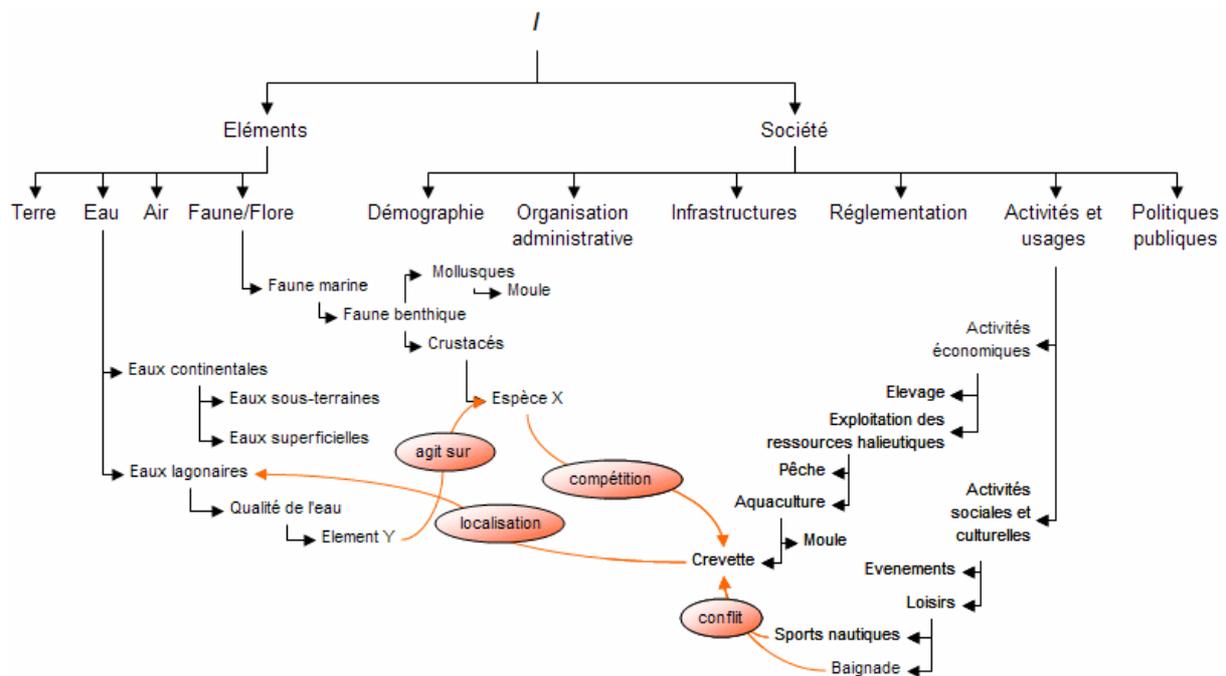


figure 13 : thésaurus et relations ontologiques

Si l'utilisateur est un gestionnaire du littoral et qu'il effectue une recherche libre en tapant le mot « crevette », il obtiendra bien plus que les fiches qui contiennent le mot « crevette » : en suivant les relations ontologiques (« agit sur », « compétition », « conflit », « localisation »), il obtiendra un ensemble de fiches associées qu'il pourra choisir de consulter.

Imaginons maintenant que l'utilisateur effectue une recherche en tapant le mot « moule ». On suppose qu'aucune fiche de MD n'a été recensée à ce sujet. Grâce au thésaurus, le résultat de sa recherche ne sera pas vide. L'utilisateur sera informé qu'aucune fiche ne concerne

directement les moules mais il sera renvoyé sur les fiches qui concernent les mollusques ou l'aquaculture par exemple : des thèmes « parent » dans le thesaurus.

Pisanelli (Pisanelli, 2002) parle de mariage du siècle à propos des ontologies et des systèmes d'informations. Le système développé lors de cette action ZONECO est tout à fait compatibles avec ce genre de développement.

### **Faciliter la saisie**

La saisie des fiches de MD est aujourd'hui réalisée par un administrateur.

Une implication des producteurs de données rendrait le système beaucoup plus vivant par une accélération de sa croissance. Cette implication n'est possible que si la constitution des fiches de MD est améliorée. Elle peut s'avérer pas assez ergonomique pour être acceptée par les producteurs de données.

Les améliorations possibles sont les suivantes :

- utilisation de **profils utilisateurs** pour simplifier le protocole en diminuant le nombre de MD à saisir et en les adaptant.
- adaptation des MD à saisir en fonction du type des données recensées : les champs à remplir diffèrent entre des données de type vidéo, des données de type tableau excel de statistiques de pêche ou des données de type carte.

### **7.2.2. Une base bibliographique**

Les références bibliographiques qui gravitent autour des jeux de données sont nombreuses. Il serait intéressant dans la base de données de créer une nouvelle table qui contiendrait l'ensemble des références afin de pouvoir développer un outil de recherche bibliographique sur le site web.

Une base bibliographique existe sous format Procite. Disponible à la SGVL de ZoNeCo, elle a été constituée par S. Bernard-Peyre lors de l'inventaire des données sur les ressources vivantes et contient toutes les références mentionnées dans ces fiches descriptives. Cette base devrait être actualisées et connectée au système de consultation des métadonnées.

### **7.2.3. Interactivité et MapServer**

#### **Pour la collecte de métadonnées**

La norme ISO 19115 implique la saisie de l'emprise géographique pour chaque jeu de données. Elle oblige une saisie en degrés décimaux. Cette étape est souvent longue et rébarbative. La délimitation directement à l'écran du polygone d'emprise sur un fond de carte serait bien plus efficace. Le développement d'un tel outil est nécessaire.

#### **Pour la consultation des fiches**

L'extraction spatiale de morceaux de jeu de données serait également un outil appréciable par l'utilisateur qui pourrait alors récupérer sur un territoire donné l'information qui l'intéresse et seulement elle. L'Ifremer de Brest a développé un outil de ce type (Sextant). L'utilisation de MapServer serait une solution légère et certainement efficace dans le cadre de ZONECO.

## CONCLUSION

Ce travail réalisé dans le cadre de l'opération « Synthèse des données » constitue une étape obligatoire dans la conduite de la constitution d'un outil efficace de consultation/restitution de l'information.

Le résultat est un site web [www.data.zoneco.nc](http://www.data.zoneco.nc) couplé à une base de métadonnées.

Il doit pouvoir permettre l'archivage des métadonnées concernant les jeux de données sur les ressources vivantes lagunaires en Nouvelle-Calédonie.

Dans l'état actuel des choses, l'intervention d'un administrateur est conseillée pour collecter les métadonnées et coordonner les utilisateurs et les producteurs des données.

Une optimisation est à prévoir afin de permettre :

- une implication directe des producteurs de données dans l'archivage de leur métadonnées,
- une recherche plus efficace des fiches de métadonnées.

Il était envisagé de poursuivre l'inventaire et d'enclencher le rapatriement et la mise en forme des données. Ces objectifs ont été reportés en raison de l'importance de l'étape essentielle et première de la constitution d'un système d'information efficace : la fabrication d'un outil d'archivage et de prise de connaissance de l'information.

Cet outil devrait faire émerger de façon naturelle les besoins en matière de restitution et les collaborations entre disciplines. Mais une implication forte de la part des partenaires est nécessaire afin de faire vivre la base de métadonnées et d'engager rapidement les développements qui contribueront à sa réussite.

Les développements prévus dans le cadre de nouvelles actions ZONECO, notamment la création d'applications à partir de la base de données LAGPLON constituée par les plongeurs – biologistes de l'IRD sur la faune et la flore des lagons de Nouvelle-Calédonie, contribueront à cette suite logique.

L'étape suivante consiste maintenant en la validation du produit proposé et à son appropriation par les partenaires de ZoNéCo et les utilisateurs potentiels de l'information sur les ressources vivantes de la Nouvelle-Calédonie.

## BIBLIOGRAPHIE

- Barde J., Libourel T., Maurel P., 2004. Ontologies et métadonnées pour le partage d'IG en GIZC. Soumis à la revue internationale de géomatique.
- Bernard-Peyre S., 1999. Inventaire des données en milieu récifo-lagonaire. Ressource vivante et Méthode de travail. Rapport ZONECO.
- Bourderie J., 2000. Implémentation de la base de données Ecotrope à l'aide du système d'information géographique Savane et traitement des données bathymétriques et sédimentologiques du lagon sud-ouest de Nouvelle-Calédonie. Rapport Intechmer, 42
- Bouvet G., 2000. Identification de biotopes lagonaires par imagerie Landsat 7 et couplage SIG. Rapport ENSA Rennes, 61 p.
- Desconnets J.C., Libourel T., Maurel P., Miralles A., Passouant M., 2001. Proposition de structuration des métadonnées en géosciences, Géomatique et Espace Rural, Journées CASSINI'01, Montpellier, France, pp. 69-82.
- Ferraris J., Garrigue C., Virly S. Bouvet G., 2000. Some considerations on the use of GIS applied to a tropical marine ecosystem : the southwestern lagoon of New Caledonia (South Pacific). *9th International Coral Reefs Symposium, Bali, October 2000.*
- Ferraris J., 2002. Interactions spatiales des systèmes Ressources-Environnement-Usages du lagon sud-ouest de Nouvelle-Calédonie. Réflexion pour un Observatoire de Recherche sur l'Eco-Système Récifo-Lagonaire. Rapport final de la convention du PEVS/ CNRS « Observatoire de Recherche sur les Anthroposystèmes : Modèles et Méthodes », 24 p.
- Labrosse P., Kulbicki M., Ferraris J., 2001. Comptages visuels de poissons en plongée. Conditions d'utilisation et de mise en œuvre. REAT Outils pour l'évaluation des ressources récifales. *Ed. Secrétariat général de la Communauté du Pacifique, Nouméa, 54 p.*
- Pautret S. 2001 Le SIG sur le lagon sud-ouest de Nouvelle-Calédonie : de la maquette vers un outil opérationnel. Illustration par une application sur les habitats benthiques. Rapport SILAT/IRD-Nouméa, 71 p.

Pisanelli D.M., Gangemi A., Steve G, 2002. Ontologies and Information Systems: the Marriage of the Century? Proceedings of Lyee Workshop, Paris.

University of Sheffield, 2000. Elargir l'accès aux informations du secteur public : guide des bonnes pratiques – Résumé détaillé – MADAME : Méthodes d'accès aux données et aux métadonnées en Europe.

#### **INTERNET :**

<http://dublincore.org/> et <http://dublincore.org/documents/2003/03/04/dcmi-terms/>

Les métadonnées préconisées dans par le DCMI (Dublin Core Metadata Element Set)

<http://www.shef.ac.uk/~scgisa/MADAMENew/french/faqfr.htm>

Questions fréquentes au sujet de métadonnées.

<http://www.lepur.geo.ulg.ac.be/Download/Normes%20de%20meta-donnees.pdf>

Les normes de métadonnées utilisées.

[http://www.pggq.gouv.qc.ca/geoinfo/num/v12n4/v12n4\\_10.htm](http://www.pggq.gouv.qc.ca/geoinfo/num/v12n4/v12n4_10.htm)

Article québécois paru dans GéoInfo présentant l'avancement de la norme ISO 19115 (novembre 2000).

## **LISTE DES ANNEXES**

**ANNEXE 1** : exemple d'une « fiche descriptive de l'étude », S. Bernard-Peyre, 1999.

**ANNEXE 2** : exemple d'une « fiche descriptive du lot de données » , S. Bernard-Peyre, 1999.

# ANNEXE 1

code fiche :  
renvoi sur fiche descriptive d'étude :  
révisée par un expert :  
conception :  
modification :

8  
oui / non  
Bernard-Peyre Stéphane, 1999

1	sujet : ressource : type de pêche : objectif :		poissons vivrière / commerciale / scientifique artisanale / industrielle Descriptif des communautés de poissons		
2	localisation :		Lagon SW		
3	nature de l'étude :		étude scientifique : biologie ; taxonomie ; halieutique ; aquaculture ; biodiversité ; économique ; environnementale ; autre : vente ; autre : / rapport administratif : campagne de pêche, / étude d'impact : ouvrage :		
4	nombre de lots de données de nature différente ?		nbre	critère	brute / traité / analysé
			6	Objet de l'observation	brute
5.1	couverture géographique de l'étude :	verticale : prof moyenne : prof mini : prof maxi : horizontale : latitude mini : latitude maxi : longitude mini : longitude maxi :	0 25  21°55 22°50' 166° 167°		
5.2	biotope associé :		pente externe / récif barrière / pente interne / récif intermédiaire / récif frangeant / fond meuble / fond dur / algueraie / herbier / milieu pélagique / semi-pélagique / mangrove / autre :		
6.1	date de début :		Juin 84		
	date de fin :		Mars 91		
6.2	suivi :		non		
7	contact :	nom : lieu de travail : e-mail : téléphone :	Kulbicki Michel Lab d'Océanographie biologique – IRD - kulbicki@noumea.ird.nc 260821		
8	équipe d'observation :		bénévoles encadrés / bénévoles non encadrés / pêcheurs / professionnels de la mesure / personnel formé / marine marchande / militaire / autre :		
9	commanditaire :		ORSTOM		
10	droit de diffusion :		Public / restreint :		
11	utilisateurs :				
11.1	publication principale :		106- Kulbicki M.; Mou-Tham G.; Bargibant G.; Menou J.-L., and Tirard P. 1987 Résultats préliminaires des pêches expérimentales à la palangre dans le lagon sud-ouest de Nouvelle- Calédonie. Nouméa (NCL): ORSTOM; 104 p. multigr.( Sciences de la Mer. Biologie Marine. Rapports Scientifiques et Techniques (NCL); 49).		
11.2	publication informatisée :		oui / non		
12	carte(s) associée(s) :		schéma / plan de position / plan des stations		
13	publications annexes :		88, 91, 93, 100		
14	lieu(x) de stockage de la donnée :		Kulbicki Michel		
15	plan d'échantillonnage :		temporel / spatial		
16	nombre d'observation :		936		
17	outil d'observation :		Enquête / drague / casier / suceuse / benne / plongeurs (observation, prélèvement manuel) / marquage / poison (spécificité : roténone ) / appareil de mesure (sonar, autre : ), ligne à main / canne / palangre / chalut / filet : autre outil :		
18	remarques :		Un objectif de cette étude est de pouvoir relier les différents modes de prélèvement pour les prises communes entre chaque. La campagne de palangre est l'objet d'une étude préliminaire, il est proposé en 1999 de continuer pour un suivi temporel cette campagne.		
19	biais sur la mesure :				
20	qualité :				

