

 compte-rendu	3 <sup>ème</sup> réunion « Data Rescue »	<b>Rédacteur</b> Alain Damiano	<b>Date création :</b> 01.07.05	<b>Référence:</b> 21.007CR/1
			<b>Dernière modif. :</b> 04/08/2005 3:09	<b>Pages</b> 3

**Diffusion :** participants à la réunion, documents du projet, web

Résumé : une réunion restreinte des personnels de l'US Osiris a été réalisée pour faire le point et programmer la suite de l'opération Data Rescue. Une liste d'actions prioritaires a été identifiée.

Lieu de réunion : Sète, CRHMT, Salle Méditerranée

Date : Mercredi 29 juin 2005, de 10H30 à 12H30

Organisateur : Pierre CHAVANCE

Participants : Peggy BRETAUDEAU / Pierre CHAVANCE / Jean LE FUR / Alain DAMIANO

Objectifs : bilan et poursuite de l'opération « Data Rescue »

Ordre du jour (proposé par P. Chavance):

- 1/ Liens de DataRescue avec le projet Ecoscope
- 2/ Bilan de l'avancement des travaux
- 3/ Outils et méthodes à développer
- 4/ Calendrier prévisionnel

## Compte-rendu

### 1/ Lien de DataRescue avec le projet Ecoscope

J. Le Fur a indiqué que le DataRescue était représenté par le module 2.1 du circuit des savoirs (avec une procédure spécifique pour la prise en compte des données). Si cette position est acceptée, il est nécessaire de conserver une trace de la réunion sous forme d'un compte rendu assurance qualité.

P.Chavance confirme l'appartenance de l'action à l'ensemble Ecoscope

Le travail va se dérouler en trois phases :

- Récupération des données,
- Etude en parallèle des procédures – amélioration du circuit des données
- Cas d'étude thématique (avec P. Cury)

Approche du DataRescue (que va-t-on faire pour les données ?)

## 2/ Avancement des travaux

- Tableau Excel élaboré par A.Damiano et P.Bretraudeau (doc n° 21.004) -> il faut maintenant récupérer les données. Sur certaines données il y a des décisions à prendre (récupération de données sur cassettes, sur papier, à saisir, à restructurer ?). Toutes les données vont être récupérées si possible sur le disque dur data rescue (un dossier par ligne dans le tableau)
- Il faut veiller à poursuivre l'alimentation du tableau inventaire (qui devrait être transformé en base access)
- Etude sur le thon germon (P.Bretraudeau) -> avancé mais encore du travail à réaliser. A ce propos la question se pose de l'opportunité d'accepter la base germon sachant que ces données sont localisées dans le Golfe de Gascogne même si on sait que le DataRescue est plus large que l'Ecoscope... Les données Germon ne sont pas prioritaires mais il serait bon de stocker le travail effectué par Peggy. Ces données pourraient servir de travaux pratiques pour illustrer les difficultés rencontrées lors de leur sauvegarde. Peggy rappelle que pour certaines données on a fait de nouveaux traitements, des hypothèses et que cela crée une nouvelle base
- Travail sur les méta-données (P.Bretraudeau avec l'assistance de J.LeFur ?) -> doc lues mais à synthétiser
- Tableau indicateur des données présenté par J.Le Fur. Il propose un graphique destiné à visualiser des indicateurs de données disponibles dans le module Data Rescue :
  - Données stockées
  - Données référencées
  - Données normées

Par ailleurs, P.Bretraudeau s'inquiète que les chercheurs puissent fournir des données de mauvaise qualité (brute, sans information) alors que nous n'avons pas les connaissances thématiques.

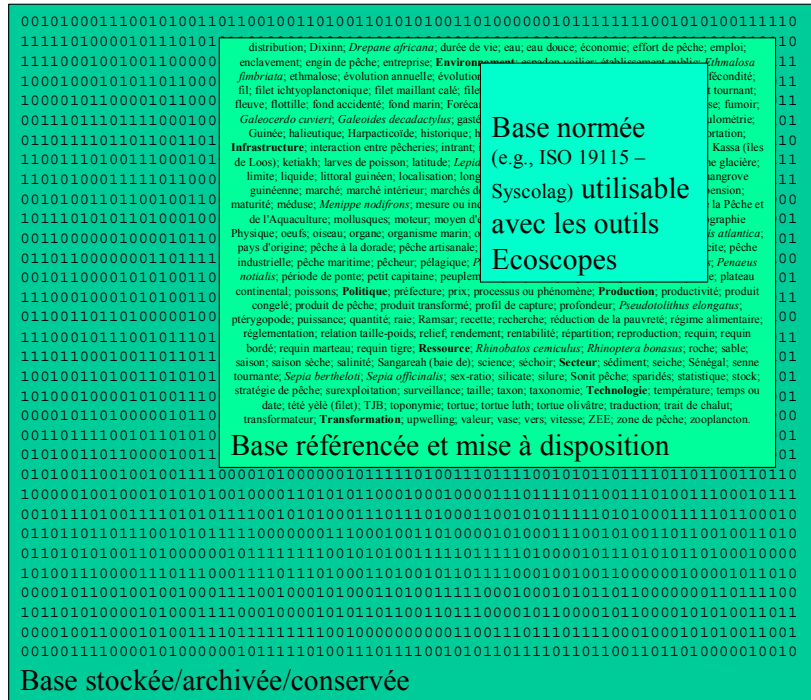
## 3/ Outils et méthodes à développer

Pierre explique qu'il souhaiterait une sorte de « guichet data rescue » où l'on proposerait un service potentiel, mais il faut élaborer un guide d'utilisation de ce guichet. Jean précise que les procédures sont déjà mises au point pour les savoirs et qu'il suffit maintenant d'étendre ces procédures aux données. Entre autres, le « proposant » d'un savoir ou de données est responsable de ce qu'il propose.

En ce qui concerne StatBase et TrawlBase, Pierre pense que nous devons être garants de ces données, quitte à mettre des « flags » sur les données douteuses, mais Jean soulève le problème du « versioning » (difficulté à garder des données remises à jour en permanence). Dans Istam on est censé retravailler sur TrawlBase et si une nouvelle version est créée on décrira la procédure de mise à jour et on fera tourner le nouveau jeu dans le circuit avec la méta-donnée associée.

Grossièrement, selon J. Le Fur, la structure des données de Data Rescue serait la suivante (voir figure ci-après) : une base conservée qui contiendrait les données brutes, enregistrées mais non formatées, une base « référencée » selon les critères du circuit des savoirs et qui pourrait être accessible, utilisable et articulée avec les autres savoirs de l'Ecoscope. Une base « normée », enfin, qui serait formatée selon une norme de méta-données.

Figure 1 Les trois bases constitutives du data rescue Ecoscope et leur taille relative prévisible



## 4/ Calendrier des actions

Les priorités sont définies comme suit :

- 1°) Rapatriation physique et sauvegarde des jeux de données brutes (stockage)
- 2°) Référencement : description de la structure et des variables. Établissement d'une notice (échanges avec le proposant). Dans un premier temps, il s'agit d'étudier un cas simple qui va permettre d'établir les procédures à suivre. Le choix s'est porté sur les sardinelles de Côte d'Ivoire.
- 3°) Norme de métadonnées ISO 19115 : Peggy sollicite l'aide de Jean. L'action est repositionnée en priorité secondaire.

Le plus urgent est d'avoir du tangible et que des données commencent à être stockées dans DataRescue. Peggy et Alain sont chargés de commencer le stockage et l'étude de cas avec l'appui de Jean