 Procédure	Etapes d'intégration des données du Data Rescue	Rédacteur Bretaudeau Peggy, Pierre Chavance	Date création : 27.07.2005	Référence: 21.010P/1
		Dernière modif. : 13/09/2005 17:05:00	Pages 10	

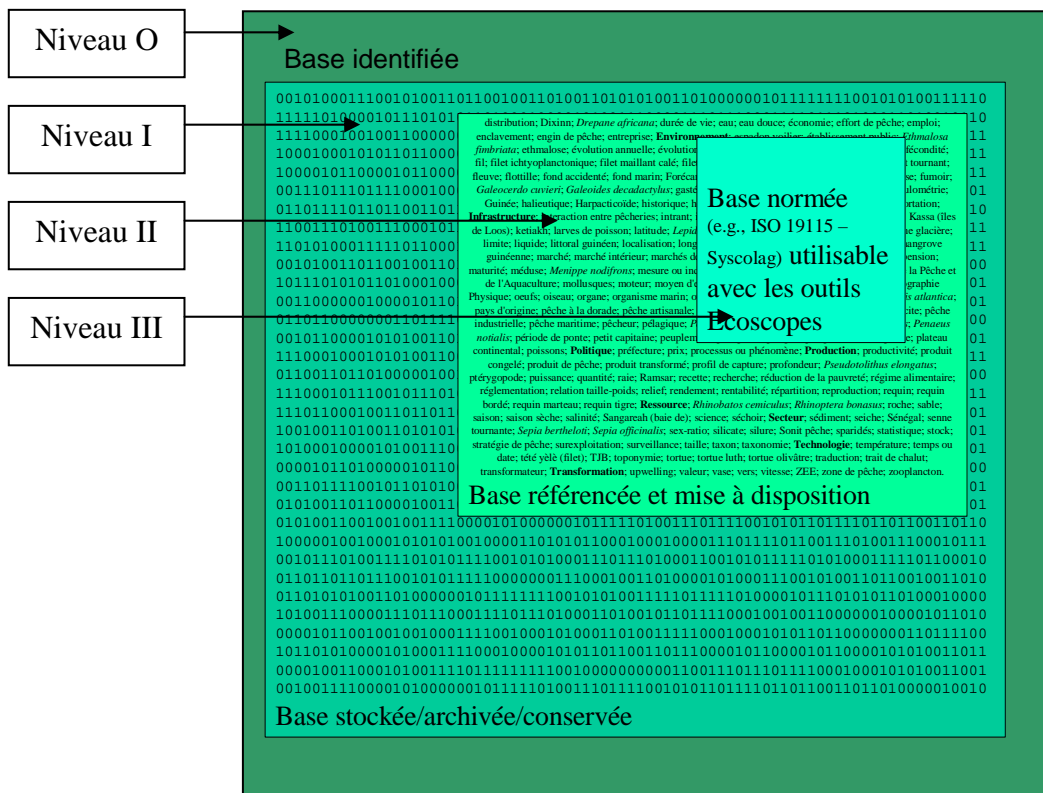
Diffusion : P.Chavance, P.Bretaudeau, J.Le Fur, documents du projet, site web

Le data rescue doit respecter des règles de procédures strictes. Ce document permet de préciser les différentes étapes suivies par les données.

## 1. PRINCIPES

La figure 1 représente la structure des données du data rescue (cf. documents du projet 21CR.reunion\_data\_rescue3) sur lequel on a rajouté la notion de niveau. Ces niveaux correspondent aux étapes qu'auront à franchir les données avant d'être considérées comme stables et susceptibles d'être rendues accessibles.

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



A partir de ce schéma, quatre niveaux d'action sont définis. Ils sont définis par le type de traitement auquel seront soumis les données afin de les rendre plus compréhensibles et utilisables. A chaque étape correspondra une validation avec un certain nombre de critères. Le parcours fera l'objet d'un enregistrement dans un fichier de suivi correspondant aux données.

Le fournisseur du jeu de données devra participer à son référencement (recueil de méta informations sur le jeu de données) et à sa validation avant mise à disposition.

## 2. DESCRIPTION DES DIFFÉRENTS NIVEAUX ET DES ÉTAPES À PARCOURIR (cf. Figure 2)

### - NIVEAU O

- A ce niveau, les données sont identifiées. C'est-à-dire que l'on a signalé leur existence mais qu'elles n'ont pas été stockées physiquement par l'équipe du data rescue.
- Les informations connues sur les données sont répertoriées dans un inventaire (cf. Annexe 1 et doc AQP 21.004, 21E.suivi\_sauvetage.xls).

### - NIVEAU I

- Pour accéder au niveau I, plusieurs actions doivent être menées mais le préalable est que les données soient dans un format électronique lisible et que les informations nécessaires au remplissage du catalogue soient disponibles.
- Les données sont stockées physiquement sur l'ordinateur du data rescue dans un répertoire propre.
- Le catalogue est rempli et une fiche de suivi est créée. Cette dernière permet de suivre toutes les actions qui sont menées sur les données.
- Un accusé de réception des données est donné au fournisseur pour attester du stockage physique des données.
- Les données peuvent être accompagnées de documentation ou non.
- Le statut de sauvegarde des données passe au niveau I.

### - NIVEAU II

- Pour accéder au niveau II :
  - Toutes les variables ou champs doivent être explicités. Les nomenclatures et unités utilisées dans les données doivent être connues.
  - Dans la mesure du possible, les informations concernant les protocoles de collecte et de mesures doivent pouvoir être explicités.
- Ce travail est fait avec la participation active du fournisseur de données.
- Le recueil de toutes ces informations est réalisé de façon à ce que toute personne ne connaissant pas les données puisse être en mesure de les comprendre à l'aide de la documentation qui est constituée, si elle n'existe pas, ou complétée.
- Des critères de qualité, qui dépendent du type de données, sont mis en place afin de garantir aux futurs utilisateurs des données les plus « saines » possibles.
- Un document est donné au fournisseur, qui atteste de la conformité des données avant éventuelle mise à disposition. Un jeu de données ainsi que la documentation associée est aussi transmis au fournisseur sur CD. Ce dernier servira de référence en cas de réclamation.
- La mise à disposition des données se fait après avoir obtenu l'accord du/des propriétaire(s) du jeu de données.
- Les modalités de mise à disposition sont elles aussi définies par le/les propriétaire(s) du jeu de données.

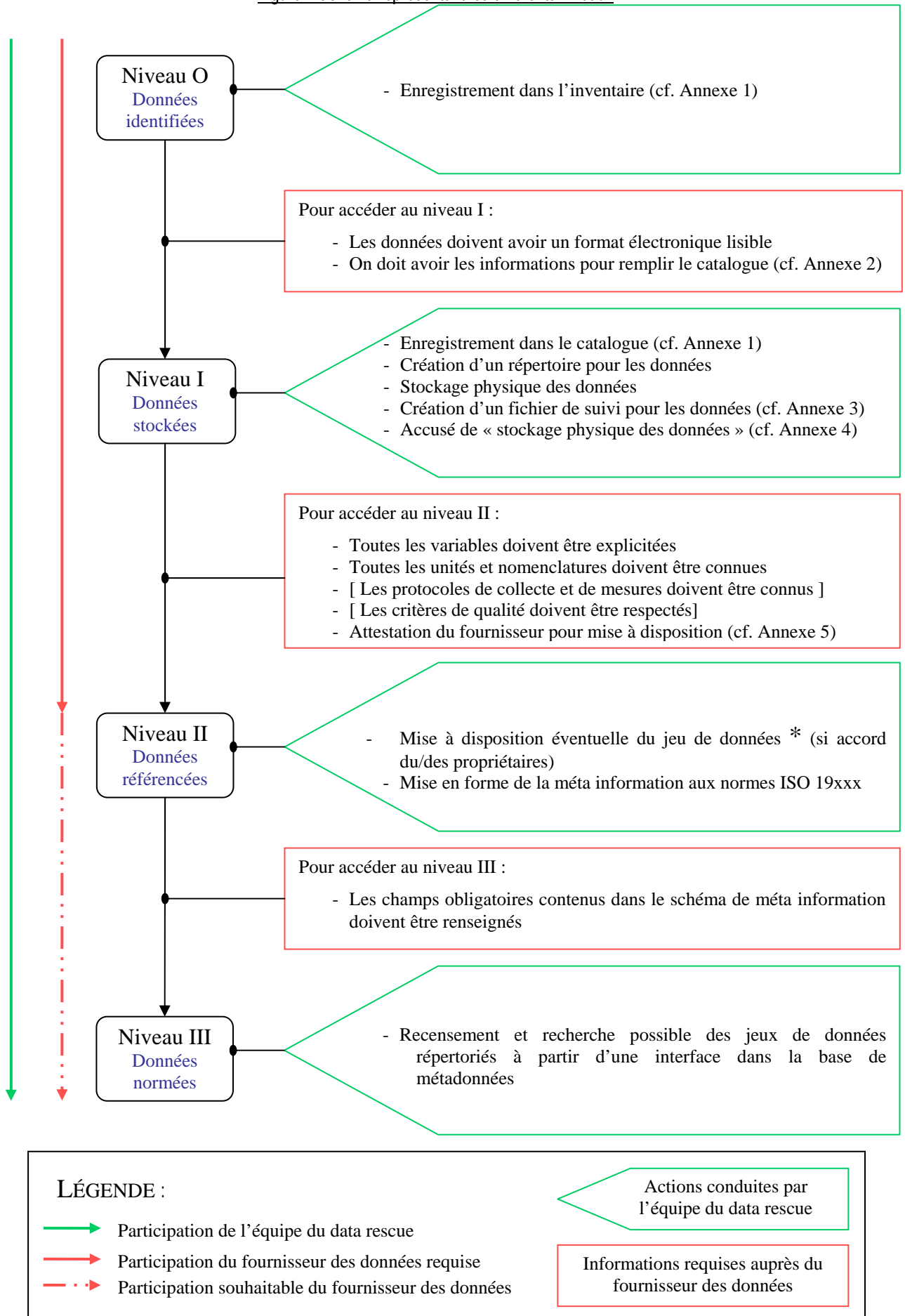
- Le statut de sauvegarde des données passe au niveau II, les données sont donc référencées.
- A ce niveau, on peut inclure les informations sur le jeu de données dans le schéma de métadonnées selon les normes ISO 19xxx. La présence du fournisseur de données n'est plus nécessaire, mais l'équipe du data rescue se réserve la possibilité de lui demander des informations complémentaires.
- La procédure de mise à disposition reste encore à définir mais il est possible qu'elle se fasse dans le centre d'information de l'Ecoscope.

#### - NIVEAU III

- Pour accéder au niveau III, les champs obligatoires contenus dans le schéma de méta informations doivent pouvoir être renseignés. Dans le cas contraire, l'équipe du data rescue sollicite l'aide du fournisseur et du/des propriétaires afin de compléter les champs incomplets.
- Si la méta information est complète, le jeu est référencé et le statut de sauvegarde des données passe au niveau III.
- Les jeux de données ainsi référencés pourront faire l'objet de recherche évoluée à partir des différents champs obligatoires de la norme.

Le niveau III ayant été atteint, le data rescue est terminé. Cependant d'autres actions peuvent avoir lieu au sein de l'Ecoscope (exemple : reformatage complet d'un jeu de données afin qu'il soit utilisable dans un modèle particulier). Ces travaux spécifiques ne constituent pas une généralité.

Figure 2 Schéma représentant les différents niveaux



### 3. PRÉCISION QUANT À LA NOTION DE FOURNISSEUR ET À L'USAGE DES DONNÉES

Le terme de fournisseur définit la personne qui fournit les données à sauvegarder. Le fournisseur n'est pas nécessairement le propriétaire des données qui peut être un organisme ou un groupe d'organisme par le biais d'un projet.

\* Il conviendra de demander l'autorisation au(x) propriétaire(s) des données pour la mise à disposition dans le cadre du data rescue et par la suite de l'Ecoscope. Les conditions de diffusion des jeux de données seront clairement établies entre le propriétaire des données et l'Ecoscope.

Rédaction : Peggy Bretaudeau, Pierre Chavance / saisie : PB

## Annexe 1

Liste des champs à remplir dans le catalogue :

Liste des champs :	Modalités <sup>1</sup> possibles mais non exhaustives :
Écosystème	<ul style="list-style-type: none"> <li>- Méditerranée/Golfe du Lion</li> <li>- Écosystèmes hauturiers</li> <li>- Écosystèmes côtiers ouest africains</li> <li>- Zones d'upwelling mondiales</li> <li>- Golfe de Gascogne (...)</li> </ul>
Nom du fournisseur des données	[Libre]
Partenaires copropriétaires ou concernés	[Libre]
Origine des données (Programme)	[Libre]
Thématique	<ul style="list-style-type: none"> <li>- Statistiques de pêche</li> <li>- Biologie</li> <li>- Écologie</li> <li>- Environnement</li> <li>- Socio économie (...)</li> </ul>
Thème	<ul style="list-style-type: none"> <li>- Capture/Débarquement/Rejets</li> <li>- Effort</li> <li>- Flottille/Recensement</li> <li>- Contenus stomacaux</li> <li>- Paramètres hydrologiques</li> <li>- Marquage</li> <li>- Structure de taille</li> <li>- Plancton</li> <li>- Campagnes de pêche</li> <li>- Campagnes d'échointégration</li> <li>- Capture/effort</li> <li>- Données de pêche/mensurations/biologie (...)</li> </ul>
Libellé court	[Libre]
Libellé long (description)	[Libre]
Support Matériel	Exemple : papier, informatique... (...)
Type de données	<ul style="list-style-type: none"> <li>- Observations</li> <li>- Données compilées</li> <li>- Données traitées (...)</li> </ul>
Structure des données	<ul style="list-style-type: none"> <li>- Bases de données</li> <li>- Fichiers (ASCII, xls, dbf...) (...)</li> </ul>
Support logiciel	- Exemple : Excel, Word, Access, MySQL (...)
Couverture temporelle	- [Libre]
Résolution temporelle	<ul style="list-style-type: none"> <li>- Observations ponctuelles</li> <li>- Heure</li> <li>- Jour</li> <li>- Quinzaine</li> <li>- Mois</li> <li>- An (...)</li> </ul>
Couverture spatiale	Exemple : ZEE pays, zone Statistique FAO(...)
Résolution spatiale	<ul style="list-style-type: none"> <li>- Ponctuelle (ville, port...)</li> <li>- Zones statistiques (divisions COPACE, préfectures administratives...)</li> <li>(...)</li> </ul>
Accès/Utilisation	<ul style="list-style-type: none"> <li>- Libre/publique</li> <li>- Contrôlée</li> <li>- Confidentielle</li> </ul>
Observations	- [Libre]
Statut de sauvegarde des données	- niveau 0

<sup>1</sup> La liste des modalités évoluera en fonction des données répertoriées dans le data rescue.  
16/09/05 – Ecoscope - 21P.circuit\_data\_rescue.1d.doc

## Annexe 2

Liste des champs à remplir dans le catalogue :

Liste des champs :	Modalités <sup>2</sup> possibles mais non exhaustives :
Écosystème	<ul style="list-style-type: none"> <li>- Méditerranée/Golfe du Lion</li> <li>- Écosystèmes hauturiers</li> <li>- Écosystèmes côtiers ouest africains</li> <li>- Zones d'upwelling mondiales (...)</li> </ul>
Nom du fournisseur des données	[Libre]
Partenaires copropriétaires ou concernés	[Libre]
Origine des données (Programme)	[Libre]
Thématique	<ul style="list-style-type: none"> <li>- Statistiques de pêche</li> <li>- Biologie</li> <li>- Écologie</li> <li>- Environnement</li> <li>- Socio économie (...)</li> </ul>
Thème	<ul style="list-style-type: none"> <li>- Capture/Débarquement/Rejets</li> <li>- Effort</li> <li>- Flottille/Recensement</li> <li>- Contenus stomacaux</li> <li>- Paramètres hydrologiques</li> <li>- Marquage</li> <li>- Structure de taille</li> <li>- Plancton</li> <li>- Campagnes de pêche</li> <li>- Campagnes d'échointégration</li> <li>- Capture/effort</li> <li>- Données de pêche/mensurations/biologie (...)</li> </ul>
Libellé court	[Libre]
Libellé long (description)	[Libre]
Type de données	<ul style="list-style-type: none"> <li>- Observations</li> <li>- Données compilées</li> <li>- Données traitées (...)</li> </ul>
Structure des données (initial – niveau I)	<ul style="list-style-type: none"> <li>- Bases de données</li> <li>- Fichiers (ASCII, xls, dbf...) (...)</li> </ul>
Support logiciel (initial – niveau I)	- Exemple : Excel, Word, Access, MySQL (...)
Structure des données (final – niveau II)	<ul style="list-style-type: none"> <li>- Bases de données</li> <li>- Fichiers (ASCII, xls, dbf...) (...)</li> </ul>
Support logiciel (final – niveau II)	- Exemple : Excel, Word, Access, MySQL (...)
Couverture temporelle	[Libre]
Résolution temporelle	<ul style="list-style-type: none"> <li>- Observations ponctuelles</li> <li>- Heure</li> <li>- Jour</li> <li>- Quinzaine</li> <li>- Mois</li> <li>- An (...)</li> </ul>
Couverture spatiale	- Exemple : ZEE pays, zone Statistique FAO(...)
Résolution spatiale	<ul style="list-style-type: none"> <li>- Ponctuelle (ville, port...)</li> <li>- Zones statistiques (divisions COPACE, préfectures administratives...)</li> <li>(...)</li> </ul>
Accès/Utilisation	<ul style="list-style-type: none"> <li>- Libre/publique</li> <li>- Contrôlée</li> <li>- Confidentielle</li> </ul>
Observations	- [Libre]
Statut de sauvegarde des données	<ul style="list-style-type: none"> <li>- niveau I</li> <li>- niveau II</li> <li>- niveau III</li> </ul>
Localisation physique sur le PC data rescue	- Exemple : C:\data_rescue\libellé_court

<sup>2</sup> La liste des modalités évoluera en fonction des données répertoriées dans le data rescue.  
16/09/05 – Ecoscope - 21P.circuit\_data\_rescue.1d.doc

## **Annexe 3**

La fiche de suivi est un état construit à partir d'Access où toutes les informations et les actions sur les données sont conservées :

L'état est en cours de construction...



## Annexe 4

### ACCUSÉ DE « STOCKAGE PHYSIQUE DES DONNÉES » - NIVEAU I

Les données doivent être dans un format électronique lisible

Personne du Data Rescue chargée du dossier : .....

Nom du fournisseur : .....

Date de réception des données : ...../...../.....

#### Informations sur les données fournies :

- Support des données : .....
- Nombre de fichiers : .....
- Taille globale des fichiers fournis : .....

Les données sont-elles fournies avec de la documentation : OUI / NON

Type de documents : .....

.....

Localisation physique des données reçues : .....

Commentaires :

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

Date : ...../...../.....

Nom et signature du fournisseur

Nom et signature  
De la personne chargée du dossier

## Annexe 5

### ATTESTATION DE CONFORMITÉ DES DONNÉES FAITE PAR LE FOURNISSEUR POUR MISE À DISPOSITION DES DONNÉES – NIVEAU II

Personne du Data Rescue chargée du dossier : .....

Nom du fournisseur : .....

Tous les champs ou variables sont explicités ? OUI / NON

.....  
.....

Les nomenclatures et unités utilisées sont connues ? OUI / NON

.....  
.....

Les protocoles de collecte et de mesures sont explicités ? OUI / NON

.....  
.....

Niveau de qualité des données :

.....  
.....

A propos du / des propriétaires des données :

- Nom du / des propriétaires : .....

.....

- Accord de mise à disposition : .....

.....

- Modalités de mise à disposition des données :

- Libre/publique
- Contrôlée
- Confidentielle

Commentaires :

.....  
.....  
.....

Date : ...../...../.....

Nom et signature du fournisseur

Nom et signature  
De la personne chargée du dossier