

Auteurs : Florent Millet, Benoît Toutain

Relecteurs : Nicolas Saby, Manuel Martin, Jean-Baptiste Paroissien

L'objectif de ce document est de fournir les informations nécessaires pour pouvoir utiliser l'entrepôt de données de l'unité Infosol.

## Table des matières

1.Connexion à l'entrepôt.....	1
2.Organisation des données.....	1
3.Données disponibles.....	2
Résultats d'analyses du programme RMQS.....	2
Résultats d'analyses du projet OPE ANDRA.....	3
Résultats d'analyses du projet AGRINNOV.....	4
4.Documentation de l'entrepôt.....	4
5.Métadonnées des tables.....	5
6.Bugs et nouvelles demandes.....	5

### 1. Connexion à l'entrepôt

C'est la base Dela qui contient les données de l'entrepôt. A l'occasion de la sortie de cette nouvelle version, **Dela** a été déplacée vers un nouveau serveur de base de données appelé Mayari.

Pour vous connecter à Dela, vous avez besoin :

- 1) de disposer d'un client PostgreSQL à jour (PgAdminIII par exemple)
- 2) d'avoir un compte PostgreSQL sur le serveur Mayari (les comptes existants sur le serveur Baracoa ont été copiés sur le serveur Mayari ; si vous n'avez pas de compte, adressez vous à Jean-Philippe ou à Benoît),
- 3) d'utiliser les paramètres de connexion suivants :
  - Hôte : **mayari.orsleans.inra.fr**
  - Port TCP : **5432**
  - Base : **dela**
  - Nom d'utilisateur : **votre identifiant LDAP**
  - Mot de passe : **votre mot de passe LDAP**

### 2. Organisation des données

Les data marts (magasins de données) sont les données de l'entrepôt qui sont directement utilisables dans des traitements. Ces data marts se présentent sous la forme d'une ou plusieurs tables classées dans des schémas Postgresql distincts au sein de la base de données **Dela**. Le nom de ces schémas est préfixé par « dm\_ ». Le tableau suivant dresse la liste des data marts disponibles :

Nom du schéma	Description du contenu
dm_dimensions	tables de dimensions de l'entrepôt, elle peuvent être utilisées comme axe de recherche dans les différentes tables de faits (temporelles, spatiales, profondeur, etc.)
dm_donnees_ponctuelles	data mart contenant les données ponctuelles de l'entrepôt (résultats d'analyses du RMQS et de ses dérivés, de la BDAT, de la BDETM, etc.)

dm_georefersols	data mart contenant les données de la nouvelle version de l'application de recherche des études pédologiques Georefersols
dm_rasters_covariable	grilles de covariables (MNT) au format PostGIS utilisables pour différents travaux
dm_rasters_prediction	grilles au format PostGIS issues de traitements de données
dm_vecteurs_covariable	différentes couches vecteurs au format PostGIS (grilles météo, Corine Land Cover, GeoFLA, PRA), pouvant entre autre être utilisées comme covariables dans des travaux de traitement de données
dm_webservices_bdetm	couches de données utilisées pour la publication des services web d'InfoSol en rapport avec la BDETM
dm_webservices_geosol	couches de données publiées dans la nouvelle version de l'outil GeoSol
dm_webservices_rmqs	couches de données utilisées pour la publication des services web d'InfoSol en rapport avec le RMQS
meta_webservices	métadonnées utilisées pour la diffusion des services web de l'unité InfoSol

Les schémas déjà existants dans la base **Dela** ont été copiés dans la nouvelle version :

Nom du schéma	Description du contenu
bdgsf	données élaborée issues des données de la BDGSF
data	données élaborées à partir des données brutes des bases de données de l'unité
igcs	données préparée à partir des données des programmes IGCS de la base Donesol
meta	métadonnées pour la base de données élaborées d'InfoSol

### 3. Données disponibles

La base de données **Dela** s'est enrichie d'un certain nombre de nouvelles données. Elle conserve aussi les données déjà présentes dans la première version de l'entrepôt.

Afin de pouvoir réutiliser un jeu de données ancien, les tables et vues des data marts sont versionnées à la date du jour (les noms des tables sont suffixées avec la date au format AAAAMMJJ).

Ci dessous, nous détaillons le schéma **dm\_donnees\_ponctuelles**, qui contient l'ensemble des tables et vues sur les résultats d'analyse ponctuelles du conservatoire :

#### Résultats d'analyses du programme RMQS

##### Tables de fait

Nom table	Description table
analyses_rmqs_composites_hz	Tableau des résultats d'analyse des composites du programme RMQS à la granularité de l'horizon (format horizontal : une ligne égale à l'ensemble des résultats disponibles sur l'horizon considéré)
analyses_rmqs_composites_prelev	Tableau des résultats d'analyse des composites du programme RMQS à la granularité du prélèvement (format horizontal : une ligne égale à l'ensemble des résultats disponibles sur le prélèvement considéré)
analyses_rmqs_fosses_prelev	Tableau des résultats d'analyse des fosses du programme RMQS à la granularité du prélèvement (format horizontal : une ligne égale à l'ensemble des résultats disponibles sur le prélèvement considéré)

## Vues

Nom vue	Table fait source	Description vue
synthese_analyses_rmqs_composites_hz	analyses_rmqs_composites_hz	Sélection des résultats d'analyse des composites de la table analyses_rmqs_composites_hz avec les données en provenance des dimensions (date, localisation, campagne RMQS, etc.)
synthese_analyses_rmqs_fosses_prelev	analyses_rmqs_fosses_prelev	Sélection des résultats d'analyse des composites de la table analyses_rmqs_fosses_prelev avec les données en provenance des dimensions (date, localisation, campagne RMQS, etc.)
reporting_rmqs_2a	analyses_rmqs_composites_hz	Sélection des résultats d'analyse des composites de la table analyses_rmqs_composites_hz pour la campagne test du RMQS (2A) dans le cadre du projet de reporting RMQS
analyses_rmqs1_composites_projet_titia	analyses_rmqs_composites_hz	Sélection des résultats d'analyses des profils composites du RMQS 1 pour le projet de Titia Mulder
analyses_rmqs1_fosses_projet_titia	analyses_rmqs_fosses_prelev	Sélection des résultats d'analyses des fosses du RMQS 1 pour le projet de Titia Mulder
analyses_andrarmqs_composites	analyses_rmqs_composites_hz	Sélection des résultats d'analyse sur les composites RMQS de France Métropolitaine pour les mêmes déterminations que celles utilisées dans le projet ANDRA.
analyses_andrarmqs_fosses	analyses_rmqs_fosses_prelev	Sélection des résultats d'analyse sur les fosses RMQS de France Métropolitaine pour les mêmes déterminations que celles utilisées dans le projet ANDRA.

## Résultats d'analyses du projet OPE ANDRA

### Tables de fait

Nom table	Description table
analyses_andra_composites_hz	Tableau des résultats d'analyse des composites du projet ANDRA à la granularité de l'horizon (format horizontal : une ligne égale à l'ensemble des résultats disponibles sur l'horizon considéré)
analyses_andra_fosses_prelev	Tableau des résultats d'analyse des fosses du projet ANDRA à la granularité du prélèvement (format horizontal : une ligne égale à l'ensemble des résultats disponibles sur le prélèvement considéré)

### Vues disponibles

Nom vue	Table fait source	Description vue
analyses_andra_fosses	analyses_andra_fosses_prelev	Sélection des résultats d'analyse sur les fosses du

		projet OPE ANDRA avec leurs dimensions correspondantes. Ajout de l'identifiant du site tel qu'il apparaît dans le projet OPE ANDRA et des coordonnées correspondantes
analyses_andra_composites	analyses_andra_composites_hz	Sélection des résultats d'analyse sur les fosses du projet OPE ANDRA avec leurs dimensions correspondantes. Ajout de l'identifiant du site tel qu'il apparaît dans le projet OPE ANDRA et des coordonnées correspondantes

## Résultats d'analyses du projet AGRINNOV

### Tables de fait

Nom table	Description table
analyses_agrinnov_composites_hz	Tableau des résultats d'analyse des composites du projet Agrinnov à la granularité de l'horizon (format horizontal : une ligne égale à l'ensemble des résultats disponibles sur l'horizon considéré)

### Vues disponibles

Nom vue	Table fait source	Description vue
reporting_agrinnov_composites	analyses_agrinnov_composites_hz	Sélection des résultats d'analyse sur les composites d'Agrinnov dans le cadre du projet de reporting Agrinnov
synthese_analyses_agrinnov_composites_hz	analyses_agrinnov_composites_hz	Sélection des résultats d'analyse sur les composites du projet Agrinnov avec leurs dimensions correspondantes

## 4. Documentation de l'entrepôt

La documentation de l'entrepôt de données s'articule autour de différents types de documents.

Un **document de cadrage général** décrit le projet de développement de l'entrepôt dans sa globalité (architecture, modèles de données, métadonnées, documentation disponible, etc.).

Pour chaque Data Mart il existe une **fiche d'indicateur** qui lui est propre. Cette fiche est utilisée pour l'analyse des besoins et sert aussi de document qualité décrivant le data mart une fois que celui-ci est mis en œuvre. La fiche d'indicateur décrit ainsi la raison d'être du data mart et la façon dont il a été construit.

Enfin, trois documents d'orientation présentent **les procédures** adaptées aux différents types d'utilisateurs de l'entrepôt de données :

- les administrateurs de l'entrepôt,
- les développeurs des nouveaux jeux de données,
- les utilisateurs des données de l'entrepôt.

Ces documents sont présents sur la forge logicielle de l'unité InfoSol à l'adresse suivante : [forge projet "Entrepôt de données"](#). Le document de cadrage et les fiches d'indicateur sont accessibles *via* l'onglet « Documents » tandis que les fichiers de procédure sont accessibles *via* l'onglet « Wiki ».

## 5. Métadonnées des tables

La signification des champs présents dans les tables de l'entrepôt est consultable grâce à [l'outil PG DaDi](#) en cliquant sur le lien « Consulter » correspondant à la base de données Dela. Vous trouverez également la description des méthodes de création de ces tables.

Il est aussi possible de consulter les commentaires des tables et des champs de la base de données directement grâce au client PostgreSQL PGAdmin.

## 6. Bugs et nouvelles demandes

Les demandes et rapports de bugs sur l'entrepôt sont à effectuer avec l'outil Bugzilla disponible à l'adresse suivante : [Bugzilla Entrepôt](#)

Pour gérer les nouvelles demandes de données et proposer un développement itératif respectant une démarche qualité (suivi des demandes, validation par les décideurs, échanges avec l'utilisateur, etc.), un mode gouvernance de l'entrepôt est mis en place . Les étapes suivies par une demande sont présentées dans l'organigramme suivant (illustration 1).

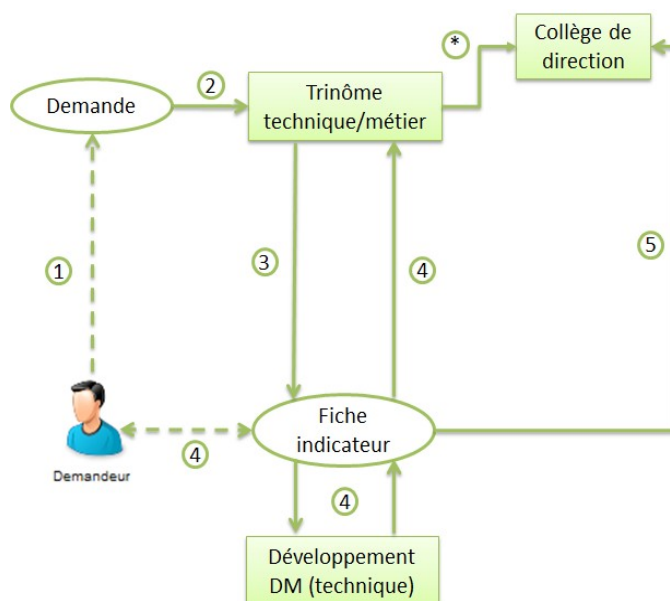


Illustration 1: Organigramme de gouvernance du Système d'Information Décisionnel mis en place à Infosol

1. Une demande de données est adressée à l'unité de la part d'un utilisateur via Bugzilla

2. La demande est remontée au trinôme technique/métier (Nicolas Saby, Manuel Martin, Benoît Toutain) qui estime :

- s'il est possible de répondre à la demande
- dans le cas où le trinôme n'est pas certain de la validité de cette demande et dans le cadre de la stratégie de l'unité, il peut en référer au collège de direction pour que ce dernier valide ou non la demande (étape \*)
- si un nouveau data mart doit être développé ou bien si les data marts existants peuvent répondre à

*la demande*

- *le temps de développement du nouveau data mart le cas échéant*

*3. Le trinôme rédige une fiche d'indicateur en formalisant le besoin de l'utilisateur et les caractéristiques du data mart (quelles données sources, règles métier, agrégations, etc.)*

*4. La fiche est envoyée à l'utilisateur afin qu'il valide si son besoin est bien compris et si la réponse le satisfait. Une fois la fiche validée par l'utilisateur, le data mart est développé. La documentation ainsi que les métadonnées sont complétées*

*5. Une fois que le data mart est opérationnel et validé par l'utilisateur, la fiche d'indicateur est envoyée en tant que document qualité et sauvegardée dans la forge d'Infosol.*

Bonne utilisation !