

**ODR, Bases de données administratives à différentes échelles spatiales**

**Cédric Gendre**  
**Inra, ESR Toulouse**



**2èmes journées de recherches en sciences sociales**

**INRA SFER CIRAD**

11 & 12 décembre 2008 – LILLE, France

## **1. Introduction et contexte**

L'observatoire des programmes communautaires de développement rural (ODR) a pour objectifs de contribuer à la préparation au suivi et à l'évaluation en France des politiques communautaires du développement rural, plus particulièrement des mesures financées par les règlements de développement rural (RDR) de 1999 (programmation 2000-2006) et de 2005 (programmation 2007-2013), et de constituer un outil coopératif pour la recherches en Sciences Sociales sur les politiques agricoles, le développement rural et l'agro-environnement.

L'observatoire est régi par une convention cadre et un comité de pilotage. Les partenaires fondateurs sont le Centre National pour l'Aménagement des Structures des Exploitations Agricoles (CNASEA), l'INRA et le Ministère de l'agriculture et de la pêche (MAP). Il peut accueillir de nouveaux partenaires et des "tiers agréés", susceptibles de fournir ou d'utiliser des données. Il est administré et développé par un chef de projet et une équipe opérationnelle, situé à l'INRA (Toulouse), cette équipe est désignée comme équipe d'administration.

C'est un système d'information partenarial au service de la préparation, du suivi et de l'évaluation des politiques publiques. Accessible par internet (<https://esrcarto.supagro.inra.fr/intranet/>), il s'agit d'un serveur de données, complété d'une plateforme logicielle permettant diverses actions sur les données (importer/exporter, traiter, visualiser). L'observatoire est piloté par des partenaires, liés par une convention. Le développement de l'observatoire s'effectue de façon collaborative par des utilisateurs disposant de droits d'accès au serveur. Le fonctionnement de l'observatoire est régi par des règles : Les données sont soit libres (données publiques), soit réservées avec un accès contrôlé par leur propriétaire.

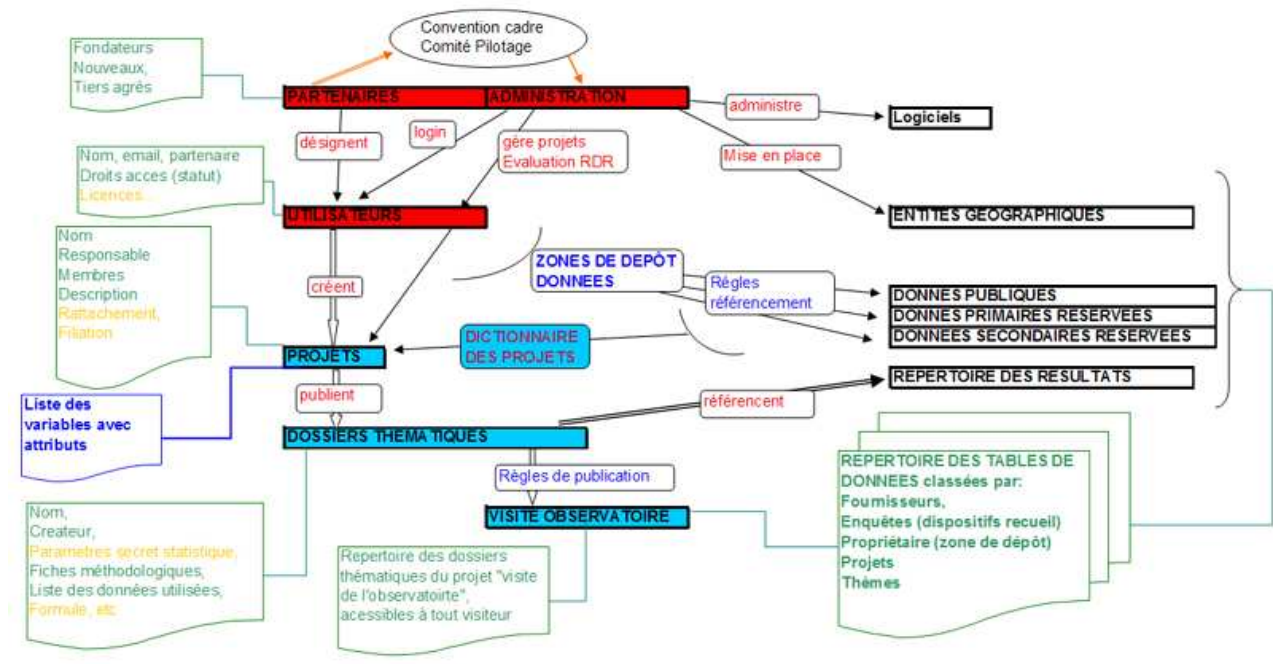
Les données disponibles sont rattachées à des codes géographiques, en particulier aux communes ou groupes de communes, mais aussi à d'autres types d'entités géographiques comme les zones Natura 2000, les programmes d'action territoriaux pour l'eau, les unités de distribution d'eau potable, etc. Par ailleurs l'observatoire donne accès à la documentation nécessaire à la compréhension des données utilisables ou consultables via un serveur collaboratif (cartowiki : <https://esrcarto.supagro.inra.fr/intranet/carto/cartowiki/index.php/Accueil>).

Le serveur de données permet soit l'accès à des données pour en réaliser un traitement (calcul et cartographie), soit la simple visualisation de résultats (cartes et tableaux déjà préparés) qui sont des données traitées. Il s'agit d'une visualisation dynamique (actualisation des calculs à chaque affichage), avec de larges possibilités de navigation entre échelles et types de territoires.

## **2. Un système d'information collaboratif géré en partenariat**

L'observatoire ODR est un système d'information, que l'on peut représenter par le schéma ci-dessous. Les relations entre les entités du système sont :

- soit directement gérées par le système informatique et de façon collaborative par les utilisateurs via les interfaces mises à leur disposition, les possibilités d'action dépendant de leur statut,
- soit elles résultent de décisions prises par les partenaires (qui désignent des utilisateurs), le comité de pilotage (qui valide les projets et émet des demandes) ou l'administration (qui attribue les droits utilisateur et les mots de passe). Ces décisions sont mises en oeuvre par l'administration.



Les rectangles du schéma sont des ensembles ou classes d'entités, définies par des attributs listés dans les fiches d'identité associées à ces classes d'entités (en vert sur le schéma). Les rectangles rouges représentent les acteurs et les rectangles bleus les ressources. Les actions sont décrites en rouge. Les principales actions sont : enregistrement (import) de tables de données, référencement de données, traitement de données, publication de résultats.

## 2.1. Les données gérées

Les bases de données sont deux types principaux :

- des données géographiques (géoréférencées) qui fournissent les contours de différents types de découpages géographiques (fond de carte),
- des données localisées (géocodées) qui sont des attributs des entités géographiques connues dans l'observatoire.

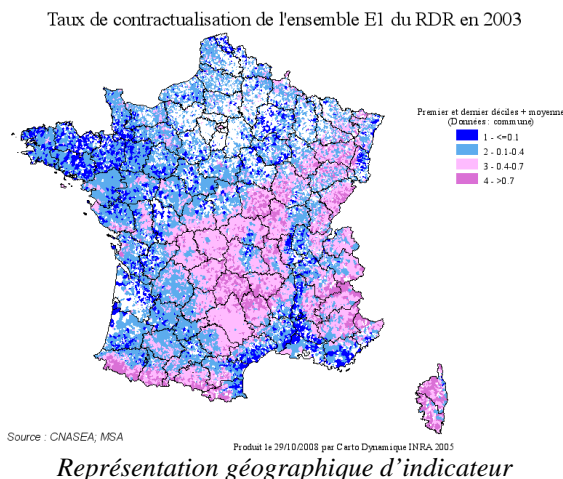
La mise à disposition et la mise à jour des données géographiques (géoréférencées) sont réservées à l'administration.

Les données localisées (ou géocodées) sont déposées par les utilisateurs désignés et agréés par les partenaires. Ces données sont nécessairement localisées au niveau des communes ou d'ensembles de communes, soit "agrégées" (une valeur par entité géographique), soit "individuelles" (plusieurs valeurs par entité géographique). Il existe dans ce dernier cas une ou plusieurs clés permettant de distinguer si nécessaire les individus. Les comptes "utilisateur" des partenaires principaux fournisseurs de données (CNASEA notamment) sont gérés par l'équipe d'administration. Toutefois, tout utilisateur titulaire peut déposer des données dans des zones de dépôts réservées (sous sa seule administration) et peut référencer ces données dans tout projet dont il est membre ou inscrit comme donateur. Les règles de dépôt et de référencement sont les mêmes pour tous les utilisateurs. La création de projet est libre. Une information sur les données (métadonnées) est mise à disposition des utilisateurs via des répertoires accessibles dans le menu accueil.

Certaines ressources sont publiques (accessibles à tout utilisateur) et d'autres sont réservées. Les données "publiques" sont accessibles à tout titulaire dans tout projet créé par celui-ci. Les données

réservées doivent être demandées au propriétaire, ou au comité de pilotage (demande d'agrément de projet) si ces données proviennent des partenaires fondateurs. Certains "résultats" sont également publics, c'est-à-dire accessible à toute personne ayant un accès visiteur. Tout projet peut servir à mettre à disposition des résultats pour des visiteurs autorisés en publiant des dossiers thématiques. Enfin, tout utilisateur titulaire peut contribuer à la documentation de l'observatoire via un module mediawiki appelé Cartowiki.

## 2.2. La gestion des résultats



L'interface de traitement permet entre autre de sauvegarder et d'exporter le résultat. Le traitement peut être simplement sauvegardé (comme un brouillon) à la disposition des membres du projet ou sauvegardé comme une "nouvelle variable" qui entrera dans le dictionnaire global des variables ou encore comme "thème" dans un dossier thématique qui sera visible par les visiteurs du projet. La liste de ces dossiers est visible dans le menu visiteur (les menus à gauche de l'écran, cf. photo d'écran page suivante). Au sens de la charte créant l'observatoire ODR, les thèmes et dossiers thématiques sont des publications.

Ainsi les résultats et publications comprennent :

- les **traitements** (sauvegarde formule et paramètres carte) sont des brouillons, ils ne sont accessibles qu'aux membres des projets concernés, ce ne sont pas des publications ;
- les **indicateurs** sont des traitements sauvegardés en tant que nouvelles variables (la valeur pour un niveau géographique donné est sauvegardée dans une base de dépôt du créateur, le lien avec le traitement est conservé (actualisation possible) et notamment les sources de données utilisées, la sensibilité de l'indicateur au secret statistique en est déduite). Ce sont des données secondaires ;
- les **dossiers thématiques** comprennent un ensemble de thèmes correspondant à la cartographie d'un indicateur et (le cas échéant) un "tableau de bord", c'est-à-dire un ensemble d'indicateurs présentés dans des tableaux statistiques géographiques ou des fiches descriptives ad hoc. Ce sont des publications (éventuellement confidentielles si peu de visiteurs inscrits au projet...), soumises notamment au secret statistique ;
- les **thèmes** (traitement validé comme thème avec les paramètres de contrôle affichage) (idem indicateur plus les paramètres d'affichage dans le dossier où cet indicateur est présent comme thème). Ce sont des publications.

Dans le cas des indicateurs sauvegardés comme thèmes ou comme nouvelles variables, plusieurs champs de métadonnées doivent être nécessairement renseignés par les créateurs de ces thèmes ou variables.

Un indicateur devenant une nouvelle variable doit être référencé dans le dictionnaire du projet pour être réutilisable. Son créateur la déclare alors comme personnelle, privée, publiable ou non, ou encore publique (si la formule ne comprend que des données publiques, les propriétaires des données primaires réservées gardent leurs droits sur les données secondaires) ; Il doit fournir un minimum de renseignements méthodologiques sur l'indicateur, qui apparaîtront dans la fiche d'identité de la variable si celle-ci est "publiable".

Enfin une documentation méthodologique peut aussi être attachée au thème, la procédure d'enregistrement précise également les conditions d'application des règles de confidentialité (voir secret statistique). L'interface de visualisation des thèmes ne permet pas d'exporter les données brutes. Toutefois une exportation des valeurs des indicateurs d'un dossier thématique est possible via la mise à disposition des visiteurs de tableaux géographiques et de tableaux de bord statistiques par territoire (quel que soit le découpage). Actuellement un simple titulaire peut seulement créer des dossiers thématiques et y classer des thèmes (cartes). L'insertion d'un tableau géographique ou d'un tableau de bord nécessite l'intervention de l'administration.

The screenshot shows the INRA 'Carto dynamique' interface. The top left features the INRA logo and the user's name 'Utilisateur : allaire'. The main header displays the project name 'évaluation finale rdr' and the user's role 'responsable'. A navigation menu on the left includes 'Le Projet', 'Le Dictionnaire', and 'Dossiers Thématiques'. The main content area shows a table of indicators for France, with a 'Présentation' section containing a 'DOSSIER VALIDE' note and descriptive text about the Rural Development Policy (PAC) and the Rural Development Program (PDRN).

*Exemple de fiche (tableau de bord) dans dossier thématique*

L'ensemble des opérations d'enregistrement et de référencement des données primaires (déposées par les utilisateurs) et des indicateurs forme le circuit des métadonnées. Toutes les variables publiables sont référencées dans le répertoire des données, accessible sur la page d'accueil. La fiche d'identité des données rassemble les métadonnées, dont : les indicateurs statistiques habituels (moyenne, maxi, mini, nombre d'occurrences non null, etc.), le propriétaire dans l'observatoire (l'utilisateur qui a effectué le référencement), la date de création, son lien avec diverses entités (voir ci-après les entités secondaires auxquelles rattacher les données), son statut au regard des règles de diffusion (les « 4 P » : privées, personnelles, publiques, publiables), sa sensibilité au secret statistique.