

Démarche qualité appliquée à la préparation et gestion des données

pour le département des Sciences Sociales, Agriculture &
Alimentation, Espace & Environnement (SAE2)



INTRODUCTION

SAE2, département des sciences économiques et sociales de l'INRA.

Les recherches qui y sont menées visent à comprendre :

- Le fonctionnement et l'évolution économiques et sociales de l'agriculture, des industries agro-alimentaires et de l'alimentation en lien avec les enjeux environnementaux
- se donne pour vocation d'éclairer les débats publics et décisions publiques et privées

Le département SAE2 regroupe 200 chercheurs et ingénieurs réunis en 19 unités dans toute la France.

Les données (enquêtes, statistiques ou de sources administratives et de gestion) sont une composante stratégique des recherches menées dans le département SAE2.

Ces données font l'objet de prétraitements afin de les consolider et les rendre utilisables dans des travaux de recherches.

Le Conseil de Gestion du Département SAE2 a donc souhaité mettre en place une démarche qualité afin d'assurer la fiabilité et la traçabilité de ces processus de traitements au sein des unités.

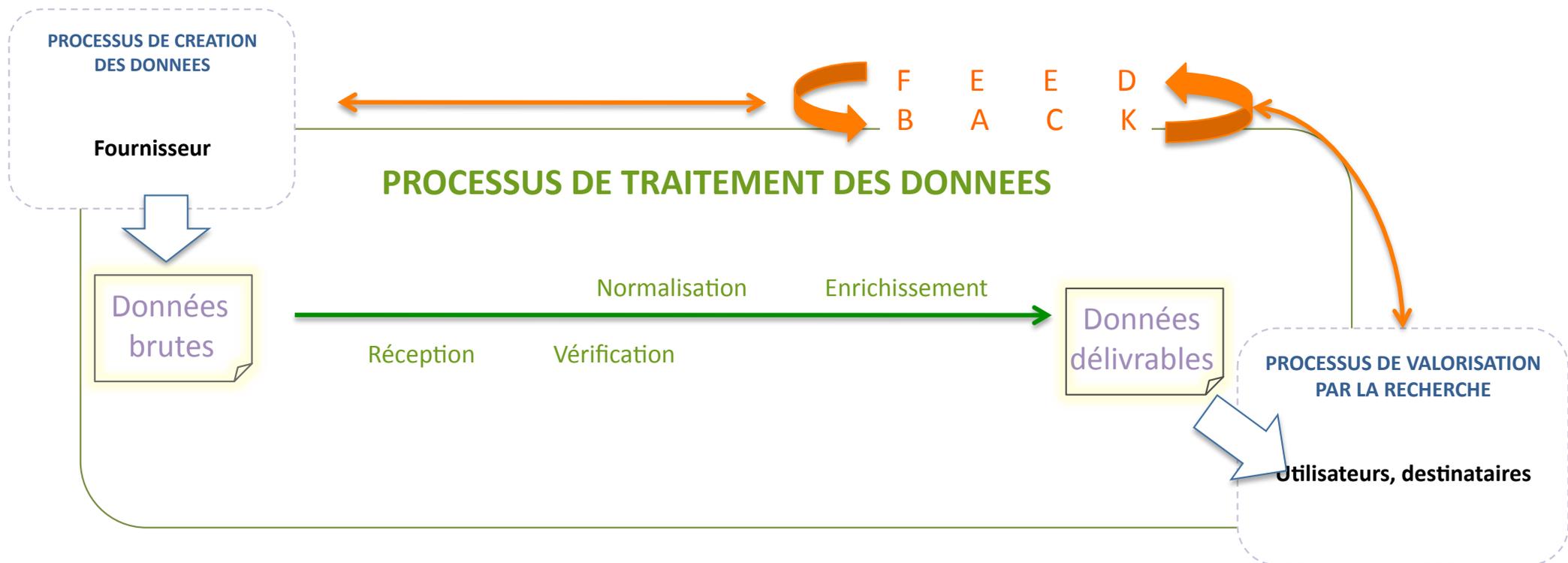
Dans ce cadre un groupe de travail constitué d'ingénieurs du département à élaborer des outils pour promouvoir et diffuser une démarche qualité visant à homogénéiser les pratiques et garantissant la qualité des processus de traitements des données.



SOMMAIRE

- ❖ Périimètre du processus entrant dans la démarche qualité
- ❖ Livrables
- ❖ La charte
- ❖ Guide des bonnes pratiques
- ❖ Conclusion

Périmètre du processus entrant dans la démarche qualité.

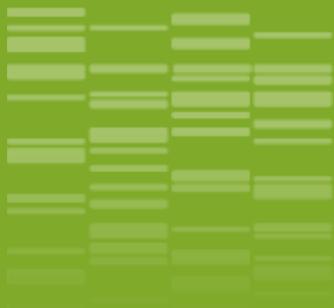


La démarche qualité concerne tous les processus de prétraitement en aval du processus de création des données et en amont de leurs utilisations dans des traitements de recherches.

Livrables

Pour faciliter la diffusion aux unités du département SAE2

- **Trois documents :**
 - Charte
 - Guide des bonnes pratiques
 - Pense-bête
- **Les objectifs :**
 - Educatifs
 - Unification des pratiques
- **Destinataires :**
 - aux plateformes, équipes, projets ou individus.



01

La Charte

5 grands principes

Charte de la démarche qualité

Objectifs

- C'est une introduction aux principes concernant **la fiabilité, la traçabilité et la pérennité** des données.
- C'est un **engagement**
- C'est un **document de référence** qui garantit la mise en place d'une **démarche qualité** au sein d'une entité.

La charte est constituée de **5 grands principes**, appuyés par des exemples de pratiques dans le guide de bonnes pratiques.

Recours aux métadonnées

Charte de la démarche qualité

Premier principe

S'engager à utiliser des normes de métadonnées pour faciliter la diffusion des données

- Ajouter et normaliser des informations sur les données permet une meilleure compréhension de leurs contenus en vue de leurs diffusions
- Destiné aux utilisateurs finaux ou aux systèmes d'informations.

Traçabilité des opérations

Charte de la démarche qualité

Second principe

S'engager à conserver les scripts et les procédures effectuées sur les jeux de données, à ne pas recourir à des opérations manuelles ne laissant pas de traces.

- Capacité de suivre les états d'une donnée tout au long du cycle de vie.
- « *Moins grave d'avoir une donnée erronée que de ne pas savoir pourquoi elle l'est.* »

Organiser le stockage, la conservation

Charte de la démarche qualité

Troisième principe

S'engager à mettre en place des procédures de stockage et de conservation, à respecter des règles de nommage et d'archivage

- But patrimonial, permettre une utilisation ultérieure par d'autres personnes.
- Que doit-on conserver ?
 - Tous les éléments associés à une donnée (logs, script, métadonnées, originaux ...).
- Où et comment doit-on les conserver ?
 - Hardware et Software
 - Plusieurs copies
 - Sécurité/Accès

Respect des règles de diffusion

Charte de la démarche qualité

Quatrième principe

S'engager à appliquer les règles de citation d'auteur, de confidentialité et de conservation des données

- En accord aux conventions passées avec les propriétaires des données ou instances décisionnelles , respecter :
 - Droits d'auteur.
 - Confidentialité.
 - Règles de diffusion du producteur.

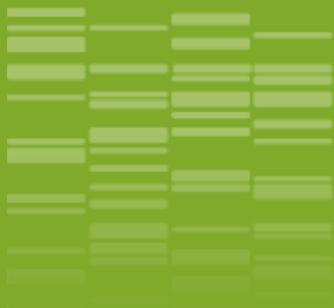
Organiser le « feedback »

Charte de la démarche qualité

Cinquième principe

S'engager à faciliter les retours des utilisateurs sur les données, à signaler les erreurs auprès des utilisateurs et à les intégrer si elles sont validées.

- La qualité des données est un processus continu et itératif.



02

Le guide de bonne pratique

Mise en œuvre de la démarche dans l'unité de service ODR

Guide des bonnes pratiques

Objectifs

- Définit et décrit des **exemples de pratiques** en accord avec la charte qualité.
- Il s'adresse aux **personnes manipulant des données** en amont du processus de recherche.
- Il est composé de **9 fiches**, chacune décrivant un type de pratique entrant dans le processus de prétraitement des données.

Guide des bonnes pratiques

Mise en œuvre dans l'Observatoire du Développement Rural

L'ODR est une unité de service du département SAE2. De par ses missions de centralisation et de diffusion de données, elle participe activement à la mise en œuvre de la démarche qualité.

L'ODR contribue :

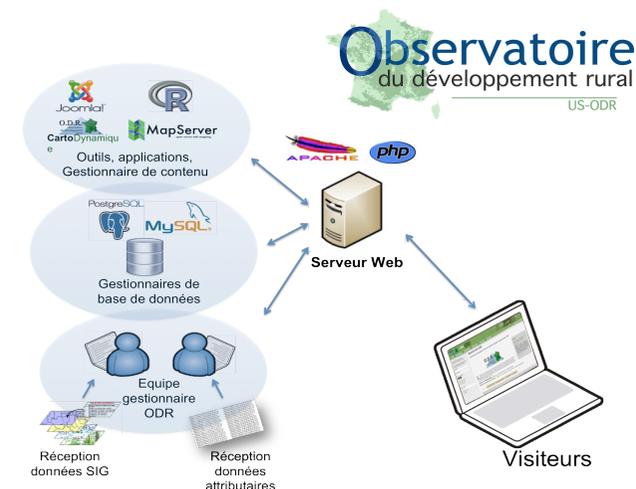
- au développement d'une base d'indicateurs pour le suivi et l'évaluation en France des politiques européennes de développement rural (fournisseur ASP)
- à la centralisation et diffusion du registre parcellaire graphique (RPG) au sein de l'INRA
- à l'élaboration de tableaux emplois agricoles issus des données MSA
- à la centralisation et gestion des données sur les signes officiels de qualité et de l'origine (SIQO, fournisseur INAO)

L'ODR s'appuie :

Sur un ensemble organisé de **ressources matériels** (serveurs), logiciels (CMS Joomla, Carto-dynamique, MySQL etc.), **personnels** (équipe ODR), **procédures** (d'autorisation d'accès, de traitement des données...), pour centraliser et diffuser les données sous forme d'indicateurs dans des *projets (fiches et tableaux)* aux équipes de recherches et autres utilisateurs (Ministère, bureau d'étude, etc.).

Remarque : Les données gérées dans l'ODR ne sont pas la propriété de l'ODR (importance de la définition des droits d'accès aux données).

Les pages suivantes décrivent la mise en œuvre de la démarche qualité dans l'ODR, en se référant aux 9 thèmes du guide des bonnes pratiques.



Utiliser des métadonnées

Objectif : décrire les données en leur associant des documents de type dictionnaire et métadonnées normalisées.

Dans l'ODR

- Métadonnées pour décrire les jeux de données brutes et jeux d'indicateurs.
- Mise en place de la norme Dublin Core[®] et outil web de saisie pour faciliter son utilisation.

Charte

X	Recours aux métadonnées
	Traçabilités des opérations
	Organiser le stockage, la conservation
	Respect des règles de diffusion
	Organiser le feedback

Contenu

Titre [Title] Ensemble des bénéficiaires engagés du PRDH maj janv 2013
 Titre développé de la table de données

Sujets ou mot clés [Subject] engagement;PDRH;rdr2
 Sujets ou mot clés

Source URL [Source] http://odr.supagro.inra.fr/
 URL

Source enquête [Source] RDR2 ensemble PDRH
 Enquête ou dispositif de recueil de l'information

Type de ressource [Type] dataset
 Dataset

Type de données [Type] Données brutes
 Données brutes

Relation [Relation] Remplir ces informations si la table est issue d'une autre table :
 [Zone de dépôt :]
 [Table :]

Description générale [Description] bio [bénéficiaire d'une aide en agriculture dans le cadre du RDR1 ou RDR2]
 annee_engagement [année d'engagement]
 code_departement [Code du département du dossier]
 code_dispositif [Code du dispositif]

Description géographique [Description] commune
 Description géographique

Description individus [Description] bénéficiaires PDRH
 Description individus

Couverture géographique [Coverage] France
 Couverture géographique

Couverture temporelle [Coverage] 2007-2012
 Couverture temporelle

Propriété Intellectuelle

Créateur [Creator] ASP
 Fournisseur pour l'ODR, créateur primaire de la donnée (MSA, CNASEA)

Editeur [Publisher] ODR
 Si pas dans la li

Contributeur [Contributor] ODR
 Si pas dans la li

Droits propriété [Rights] privées
 Si pas dans la li

Droits publication [Rights] publiable
 Si pas dans la li

Instanciation

Date de l'enquête [Date] 2007-2012
 date de l'enquête

Date de mise en ligne [Date] sélectionner :
 Si pas dans la liste : date de mise en ligne

Format [Format] dataset
 MySQL

Identifiant Ressource [Identifier] Identification de la ressource :
 [Zone de dépôt : rdr2_ensemble]
 [Table : engagement_j9]
 Nom complet de la table :
 [Zone de dépôt : nom de la base]
 [Table : nom de la table]

Langue [Language] français
 français

Contacter l'auteur: Gendre cedric
 Voir les membres ayant accès
 Voir la licence de publication
Consulter les métadonnées
 Signaler un bug
 Aide sur les dossiers

Définir et suivre un cahier des charges

Objectif : Décrire les attentes des utilisateurs, dans l'idéal

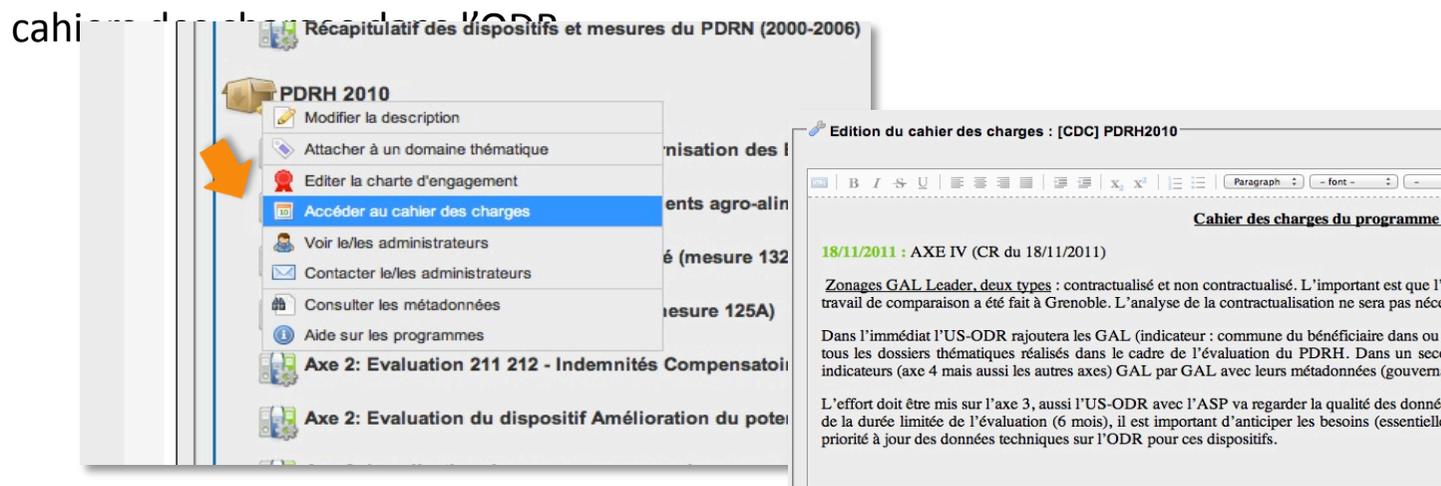
- Format des données à délivrer
- Traitements à effectuer :
 - variables utiles, celles à créer ou modifier
 - vérifications à effectuer
 - corrections à apporter.
- Définit les métadonnées et les statistiques à fournir

Dans l'ODR

- Organisation de réunions avec les utilisateurs et fournisseurs
- Production d'un cahier des charges (même succinct).
- Mise en place d'outils web facilitant la rédaction et la lecture de cahiers des charges

Charte

	Recours aux métadonnées
X	Traçabilités des opérations
	Organiser le stockage, la conservation
X	Respect des règles de diffusion
	Organiser le feedback



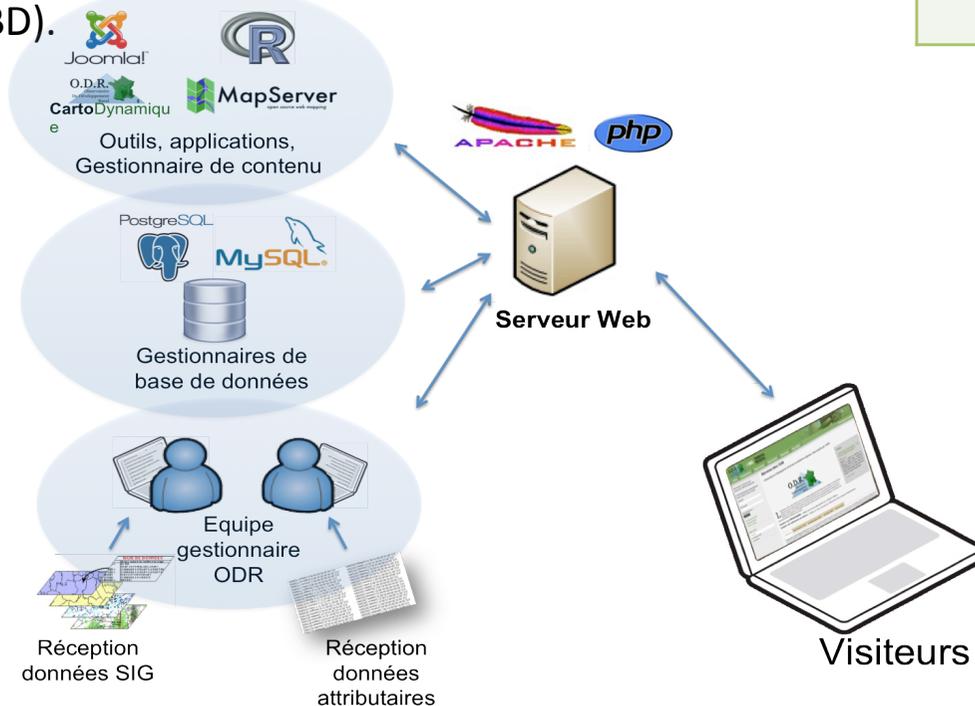
Se fixer des règles de stockage

Et des règles d'arborescences

Objectif : Unicité matérielle et temporelle des données, programmes et procédures.

Dans l'ODR

• Intégration et centralisation des données et traitements sur un serveur (SI) qui embarque un système de gestion de base de données (SGBD).



Charte

	Recours aux métadonnées
X	Traçabilités des opérations
X	Organiser le stockage, la conservation
	Respect des règles de diffusion
	Organiser le feedback

Se donner des règles de nommage

Objectifs : Retrouver et identifier la nature d'un fichier (nommage en lien avec le jeu de données, la version) ou d'une variable (origine ou variable créée).

Dans l'ODR

- Normaliser les noms de champs avec quelques règles simples. Utiliser des minuscules etc.
- Rester en conformité avec les noms des fichiers et variables d'origines fournis par les producteurs de données.
- Suivre les versions de données grâce à leur nommage.
- Rendre les règles de nommages visibles (ex dans cahier des charges).

Charte

	Recours aux métadonnées
X	Traçabilités des opérations
X	Organiser le stockage, la conservation
	Respect des règles de diffusion
	Organiser le feedback

Maîtriser l'enchainement des étapes

Objectif : Maitrise de l'ordre des traitements et leurs contenus

Dans l'ODR

Regroupement des différents outils de prétraitement dans une seule interface pour faciliter le suivi et leur utilisation :

- *Variable active* : définition des champs référentiels individus et géographique. Définition des droits sur chaque variable (disponibilité, privée/publique).
- *Métadonnées* : saisie des métadonnées.
- *Libellé* : saisie des libellés des variables et informations (type d'unité etc.)
- *Pense bête* : traitements SQL d'ajout ou de correction des données et Log de ces traitements.
- *Comparer* : Normalisation sur une référence (géographique) pour les rapprocher ultérieurement à d'autres jeux données ayant la même référence.
- *Statistique* : premier contrôle des données (statistiques descriptives).

Charte

	Recours aux métadonnées
X	Traçabilités des opérations
	Organiser le stockage, la conservation
	Respect des règles de diffusion
	Organiser le feedback

	Nom de la Table	Variables actives	Metadonnées	Libellés des Variables	Pense bête	Comparer les géocodes importés aux géocodes référence	Statistiques de base	Export	Refaire l'import	Suppression en série de la liste	Fichier Source
C	Voir la table: referentiel_financeur_fra	Retenir	Créer metadonnees	Créer Libellés	Modifier script	Comparer	-Statistiques ou -Utiliser R	Menu export	▪ Voir script	Supprimer	Table : paiement_prepare_j9

Assurer la sécurité et la sauvegarde

Objectif : mettre en place des procédures (système physique/software) pour assurer sécurité et pérennité de toutes les données. Se prémunir des risques de détérioration ou perte de données

Dans l'ODR

- serveur sur Montpellier.
- sauvegarde journalière sur le serveur et hebdomadaire à Toulouse.



Charte

	Recours aux métadonnées
	Traçabilités des opérations
X	Organiser le stockage, la conservation
	Respect des règles de diffusion
	Organiser le feedback

Assurer la confidentialité

Objectif : Respect des règles (utilisation/accès/conservation/citation)

Dans l'ODR

- Gestion fine des profils utilisateurs et de leurs accès aux données (indicateurs dans l'ODR).
- Outils pour le contrôle du secret statistique, et agrégation des données à des niveaux géographiques exportables.
- Clause d'engagement des utilisateurs avant le téléchargement d'un jeu d'indicateurs (signature numérique).

Charte

	Recours aux métadonnées
	Traçabilités des opérations
X	Organiser le stockage, la conservation
X	Respect des règles de diffusion
	Organiser le feedback

Etude stratégique pastoralisme pyrénées

- Editer les règles de publication
- Configurer les profils d'accès
- Contacter l'auteur: Gendre cedric
- Voir les membres ayant accès
- Voir la licence de publication
- Consulter les métadonnées
- Signaler un bug
- Aide sur les dossiers

122A	Amélioration des forêts	S	
125A	Desserte forestière	29	
125B	Soutien retenues collinaires	S	
131	Identification Ovin Caprin	S	S
132	Qualité alimentaire	77	17204

Demande de confirmation

Vous allez procéder au téléchargement des données présentes sur cette fiche, cette action sera enregistrée sur nos bases de données à des fins statistiques.
Vous pourrez à tout moment consulter vos téléchargements dans la rubrique *Vos informations* du menu utilisateur.

Vous avez déjà signé la charte d'engagement (le 2012-03-28 11:03:08)

Votre inscription sur l'ODR est personnelle, votre login et le mot de passe ne doivent pas être transmis.

Il est conseillé de lire attentivement la note méthodologique associée au dossier.

Toute publication utilisant les données téléchargées du site ODR devra mentionner les sources ainsi que le traitement par l'INRA (US-ODR) et les auteurs, tels que figurant sur le document téléchargé. Tout usage commercial des données est interdit.

Les « dossiers thématiques », « fiches descriptives » et « tableaux d'indicateurs » sont des produits de l'ODR, sous la licence « creative commons ». [Voir la licence](#)

Ok

Organiser le feedback

Objectif : assurer la bonne communication autour des données et améliorer les traitements.

Dans l'ODR

- Contact pour les utilisateurs.
- Utilisation d'outil de suivi des erreurs et corrections.
- Notes de mise à jour dans les publications ODR (fiches, tableaux) permettant de télécharger des jeux d'indicateurs.

Charte

	Recours aux métadonnées
X	Traçabilités des opérations
	Organiser le stockage, la conservation
	Respect des règles de diffusion
X	Organiser le feedback

Introduction 2007 2008 2009

Ensemble des aides du RDR

Etendue géographique: France entière

Présentation

Note de mise à jour (3/05/2013)

Note concernant les engagements et paiements pour l'axe 3 :

- Les montants affichés concernent les paiements effectués par l'ASP ainsi que les paiements nationaux dissociés remontés par l'ASP. Ces paiements dissociés sont différés -le paiement FEADER étant versé en premier par l'ASP- et peuvent être...
- Dans les montants engagés (de l'ASP) les financements dissociés n'apparaissent pas. Ceux-ci sont calculables selon le...

prochaine mise à jour.

Note de mise à jour (2/05/2013)

Données mise à jour RDR2 janvier 2013, source ASP

Etendue par défaut, 6 départements pyrénéens. Possibilité de restreindre l'étendue à la zone massif Pyrénées ou aux anc...
Ventilation par région en cours de modification

Note de mise à jour (17/12/2010)

En 2010 le montant négatif pour les engagements de la PHAE2 (dispositif 214) est dans les faits un désengagement des...
ne sera visible dans l'ODR qu'à la prochaine mise à jour.

Note de mise à jour (04/11/2010)

Les résultats correspondent maintenant à la nouvelle version des engagements et des paiements (2007-2010). La versio...

mantis BUG TRACKER

Logged in as: cedric (reporter)

Main | My

Reporter: any Monitored By: any

Status: any Hide Status: none

Show: 50 View Status: any

Platform: any OS: any

Navigateur: any

Note By: any Sort by: Category Ascending

Search [] Apply Filter [] [Advanced Filters]

Viewing Issues (1 - 50 / 146) [Print Reports] [CSV Export] [Excel Export] [Graph]

P	ID	#	Category	Severity	Status	Updated	
	0000161	1	(CMS Joomla) General	minor	resolved (blaise_developpeur)	2013-05-06	un utilis...
	0000157		(Fiches descriptives et Tableaux d'indicateurs) General	minor	assigned (mledoux)	2013-05-03	fiche de...

Etude stratégique pastoralisme pyrénées

- Editer les règles de publication
- Configurer les profils d'accès
- Contacter l'auteur: Genre cedric
- Voir les membres ayant accès
- Voir la licence de publication
- Consulter les métadonnées
- Signaler un bug
- Aide sur les dossiers

CONCLUSION

La démarche qualité appliquée à la préparation et à la gestion des données

Première étape vers une démarche plus globale de qualité des données intégrant à l'avenir :

- ❖ D'autres étapes et types de données : données expérimentales.
- ❖ Une démarche vers une certification pour les plateformes de données.
- ❖ La conception d'un indice de qualité des données (réflexion à venir au sein de l'ODR).

La démarche qualité est une garantie à faire valoir, lors de réponses à appel à projet et lors de publication d'articles scientifiques.

La démarche décrite ici ne garantit pas la qualité interne (intrinsèque) des données en tant que telle, qui dépend des producteurs / fournisseurs de données, mais elle participe à l'amélioration et au suivi de l'évolution de ces données.

Elle est indispensable pour diffuser les données à un public élargi.