

# Retour d'expérience sur... la rédaction et l'évaluation des plans de gestion de données

Contexte Inra



# Un service Inra autour des PGD

- ❖ Grille type et un outil de rédaction Opidor,
- ❖ Tutoriels et guides : via le [site Data Partage](#) ou l'aide en ligne

The screenshot shows the INRA 'gestion et partage des données scientifiques' website. The INRA logo is at the top left. The main header reads 'gestion et partage des données scientifiques'. Below the header, there is a navigation menu with 'Accueil', 'Gérer', and 'Plan de gestion'. The 'Plan de gestion' section is active, displaying the title 'Plan de gestion des données' and a brief description: 'Un plan de gestion des données est un document formalisé, établi au démarrage d'un projet de recherche, et qui décrit la façon dont les données produites seront obtenues, documentées, utilisées, gérées et disséminées, pendant la phase de recherche et une fois le projet terminé.' To the right, there is a 'Boîte à outils' sidebar with a list of links: 'Questions/Réponses sur les données de la recherche', 'Rédiger un plan de gestion', 'Obtenir un DOI', and 'Arbre de décision, statut des...'. On the left side of the screenshot, there is a blue-bordered box containing a list of links: 'Plan de gestion', '> Pourquoi rédiger un plan de gestion?', '> Comment rédiger un plan de gestion?', and '> Exemples de plans'.

- ❖ Information / présentation
  - Journées « Données » IST : Nancy, Nantes.
  - Appui / journées Hackathon – CATI CGI
- ❖ Séances de formation – classes virtuelles (infodoc express)
- ❖ Aide, conseils pour la rédaction et la relecture

# Rédaction PGD – Retours d'expérience

## ❖ Porteurs de projets – Rédacteurs PGD

- Projet H2020 DiverImpacts – A. Messéan (Eco-Innov)
- Projet Rapsodyn – S. Durand (URGI)

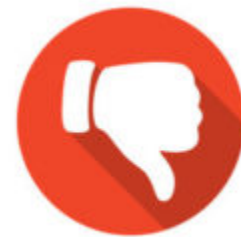
## ❖ Enquête auprès des titulaires de comptes Opidor

- formulaire de 9 questions adressé à ~100 personnes
- 12 retours : 11 rédactions de PGD
  - dont 9 pour un projet, 2 pour une structure,
  - dont 3 suite à l'exigence / financeur, 8 pour une autre raison (test outil : 5, exigence / labellisation : 1, améliorer la gestion des données : 1, programme multiprojets : 1)
  - dont 7 ont utilisé la maquette Inra, 3 la maquette H2020, 3 une autre maquette

## ❖ Exploitation des échanges Q/R lors des séances de classe virtuelle animées sur les PGD

- 5 séances de juin à décembre 2017 ( ~2h – 51 participants)

# Rédiger un PGD : Difficultés...



- ❖ **collecte et assemblage des informations** nécessaires pour renseigner les questions du PGD
  - selon la taille du projet (structure, nombre de jeux de données à prendre en compte)...
- ❖ **mobilisation** des partenaires difficile
  - exercice demandé trop tôt, nécessité d'actualiser régulièrement,
  - personnes encore peu expérimentées,
  - autorité du rédacteur du PGD.
- ❖ **limites de l'outil Opidor**
  - description de l'ensemble des datasets difficile,
  - partage possible mais fonctionnalités encore limitées , (identité non visible, pas de workflow de validation)
  - pas de gestion optimale des versions du PGD actuellement,

# Rédiger un PGD : Bénéfices

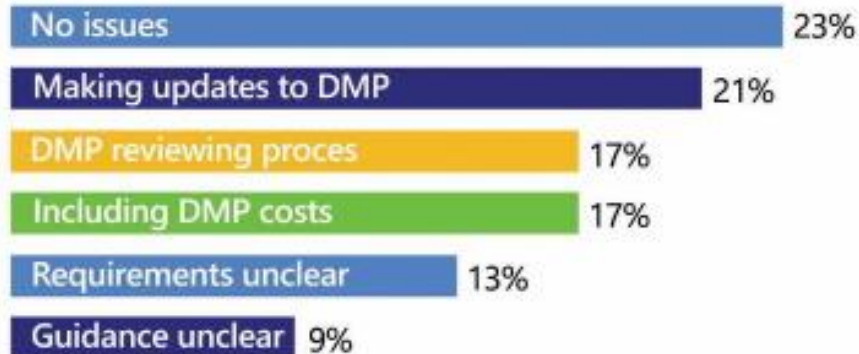


- ❖ **lister les données** qui seront produites et **échanger sur les méthodes** de production, traitement, gestion,
- ❖ **identifier des incohérences** dans le plan de travail et **harmoniser des protocoles**,
- ❖ garantir la **production des méta-données**,
- ❖ anticiper dans un projet les **données qui seront partagées** et mettre en place les bonnes conditions du partage (démarche FAIR : identifiant, entrepôt, métadonnées ...)
- ❖ préparer/faciliter la **publication des données** dans le cadre de la politique d'ouverture des données et anticiper les éventuels problèmes.

# SURVEY ON THE HORIZON 2020 DMP TEMPLATE

09/01/2018

## Difficultés rencontrées



289 réponses  
50% chercheurs

## 5 priorités

1. Suggest relevant **standards** for my field and data type
2. **Drop-down options** based on good practice per discipline
3. Give more **examples** or suggested answers
4. Include **discipline specific** guidance and tailoring
5. Recommend **repositories or tools** that I can use

## Evolutions à prévoir

- Revise the DMP template structure
- Reduce technical terminology
- Provide discipline-specific guidance
- Offer example DMPs and costings
- Clarify DMP review processes

<https://zenodo.org/record/1120245#.WlhWRzciFaQ>

# PGD : Evaluation

## Expérience de reviewing – Odile Hologne 15 PGD / Commission Européenne

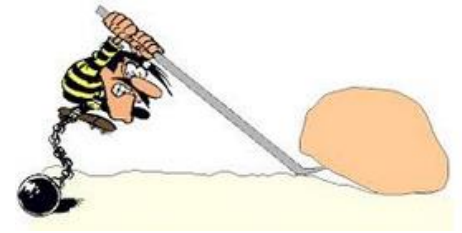
- ❖ Grille d'évaluation : très orientée FAIR data
- ❖ Evaluation sur la prise en compte de la problématique des données au sein du projet.
- ❖ Le PGD permet
  - ❖ de comprendre les types de données produites ou réutilisées,
  - ❖ d'identifier celles qui seront ouvertes et comment ?, celles qui resteront fermées et pourquoi ?
  - ❖ de montrer comment les données produites seront réutilisables

PGD = **REVELATEUR** des procédures mises en œuvre (ou pas) autour des données, et dès le stockage.

Grille	Constat	Conseils
Data Summary	<ul style="list-style-type: none"> <li>• Pas assez d'approche collective, patrimoniale des données</li> </ul>	<ul style="list-style-type: none"> <li>• Bien renseigner la version du PGD (document évolutif)</li> <li>• Ne pas trop fragmenter auprès des partenaires</li> <li>• Evaluer la volumétrie et chiffrer les coûts</li> </ul>
<b>F</b>	<ul style="list-style-type: none"> <li>• Métadonnées souvent mal renseignées</li> </ul>	<p>Dès le <b>stockage</b> :</p> <ul style="list-style-type: none"> <li>• « Rich metadata »</li> <li>• règles de nommage...</li> </ul>
<b>A</b>	<ul style="list-style-type: none"> <li>• Fréquent choix de Zenodo</li> </ul>	<ul style="list-style-type: none"> <li>• Privilégier les entrepôts disciplinaires ou Data Inra</li> <li>• Exigence sur la qualité des métadonnées</li> </ul>
<b>I</b>	<ul style="list-style-type: none"> <li>• Peu renseigné</li> <li>• Un seul projet Linked Data</li> </ul>	<ul style="list-style-type: none"> <li>• Utiliser des schémas de métadonnées et vocabulaires standards</li> </ul>
<b>R</b>		<ul style="list-style-type: none"> <li>• Réfléchir au potentiel de réutilisation des données</li> </ul>
Global	<ul style="list-style-type: none"> <li>• Outil à utiliser dès le montage des projets, aide potentielle à la <b>conception</b></li> <li>• Doit révéler la bonne prise en compte des questions <b>d'ouverture</b>, et <b>réutilisation</b> des données</li> </ul>	

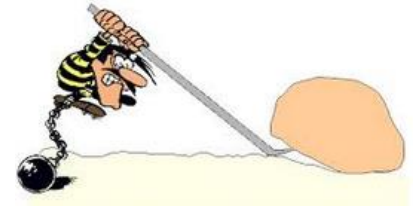


# Freins et leviers (1/2)

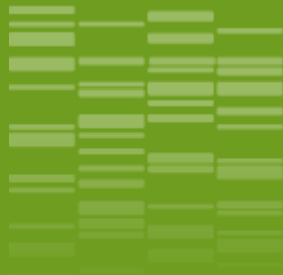


- ❖ opération chronophage, (un document de plus à rédiger !)
- ❖ opération complexe :
  - ampleur des projets,
  - partenariats multiples (données dispersées, outils et formats différents...),
- ❖ nouveauté de l'exercice, encore assez peu connu, et encore peu souvent exigé des financeurs
- ❖ Opidor à optimiser (versionnage, validation, interopérabilité...)
- ❖ intégration des informations dispersées dans différents documents (accord de consortium, documents qualité, cahiers de laboratoire ...),
- ❖ interopérabilité des outils dans le cycle de la recherche (liens entrepôt - PGD par exemple),
- ❖ procédures claires et exigences sur les métadonnées dès le stockage (infrastructures) et lien avec Data Inra
- ❖ échanges au sein de collectifs (exemple CATI ...)

# Freins et leviers (2/2)



- ❖ consolidation du dispositif d'appui
  - en intégrant des référents scientifiques (notion de [“data champions”](#)),
  - en mutualisant les retours d'évaluation des PGD.
- ❖ extension de la démarche à des structures
  - filiation entre PGD unité et PGD projets (ex: unités EFPA)
- ❖ intégration de cette démarche dans le cursus doctorants et jeunes chercheurs,
- ❖ mise à disposition d'un espace dédié au projet (si associé à un PGD) dans l'entrepôt institutionnel Data Inra,
- ❖ faire du PGD un livrable de tout projet produisant des données
- ❖ renforcement de la politique Open Data des financeurs (ANR ?)



Merci de votre écoute !

Dominique L'Hostis & Sylvie Cocard  
Pôle Digitalist