Un Plan de Gestion de données

Pourquoi ?

Hanka Hensens, IST-IRD
Hanka.hensens@ird.fr

Supports :
https://tinyurl.com/2p8ts5pp

Un PGD, qu'est-ce que c'est ?

En quelques mots, un formulaire :

- Document structuré **et** évolutif
- Décrit l'ensemble des données produites ou utilisées dans un projet, une structure,…
- Liste toutes les questions à se poser
- Sur leur obtention, gestion, conservation, partage,…
- Tout le long de leur cycle de vie

LACHEZ LES FREINS !

Instructions du Data Excuse Bingo PGD

Enumérer sur des post-its à coller
dans les cases du tableau
Toutes les raisons qui peuvent pousser à ne PAS rédiger de DMP/PGD

A vous de jouer !

Data Excuse Bingo PGD

Avez-vous coché toutes les cases du Data Excuse Bingo ?

Je n'ai pas le temps / Ce n'est pas la peine / On le fera plus tard…	Je n'y connais rien / C'est trop compliqué / Ce n'est pas à moi de le faire…	Mon financeur/mon institut/mon université ne m'y oblige(nt) pas…
J'ai trop de partenaires / J'ai des partenaires privés/étrangers / Je ne suis pas propriétaire des données/ Mon partenaire s'occupe de tout…	Je ne veux pas qu'on sache comment je gère mes données / On va voir que je ne respecte pas les règles/les standards…	Mes données sont sensibles, personnelles / Elles concernent les ressources génétiques ou la biodiversité / Elles sont Top secret !
Je veux publier / Je veux déposer un brevet / Je veux les utiliser après la fin du projet…	Je n'ai pas encore de données pour ce projet / Mes données ne sont pas finalisées	Mes données ne regardent que moi / On va me copier, me spolier,… ne pas me citer !

Lâcher les freins…

Je n'ai pas le temps… A faire dès le début, ça fait gagner beaucoup de temps après [1] A faire en équipe : se répartir la tâche [2] Long la 1ère fois, beaucoup moins long avec l'expérience [3]	*Je n'y connais rien…* A faire en équipe [2] Assistance des juristes, des informaticiens, des documentalistes,...[2] Engager un data manager (et le budgéter) Recommandations en ligne, exemples [4]	*Je ne suis pas obligé…* Financeurs, organismes,... recommandent de plus en plus s'ils n'obligent pas.[8] Avantages : cohésion de groupe, partage des informations, prévision des dépenses, aplanissement des éventuels problèmes juridiques, informatiques,..[1]
Avec mes partenaires c'est compliqué… Justement, c'est dans ces cas là qu'il faut éclaircir les droits de propriété sur les données, qui fait quoi, qui a accès à quoi,..[1]	*Mes données ne sont pas parfaites…* On le soupçonnera d'autant plus s'il n'y a aucun PGD : Assurance de qualité [8] Preuve de sérieux et de savoir faire [5]	*Mes données sont sensibles…* C'est quand les données sont sensibles qu'il faut les gérer au mieux sous peine de problèmes pour les utiliser, les partager, ou même de procès,…[1]
Je veux exploiter mes données… Un PGD n'empêche rien et peut même être publié, et cité, Un PGD n'oblige pas à partager ses données : il suffit d'expliquer pourquoi on ne peut pas ou pas tout de suite [6]	*Mes données ne sont pas finalisées…* Tout ne doit pas être mis dans la v1, versions successives, corrections et ajouts possibles si données non prévues, traitements non prévus, etc. [7]	*Mes données m'appartiennent…* Recherche financée par des fonds publics = obligation d'ouverture si possible [8] Un PGD n'est pas public par défaut, un PGD ne contient pas de

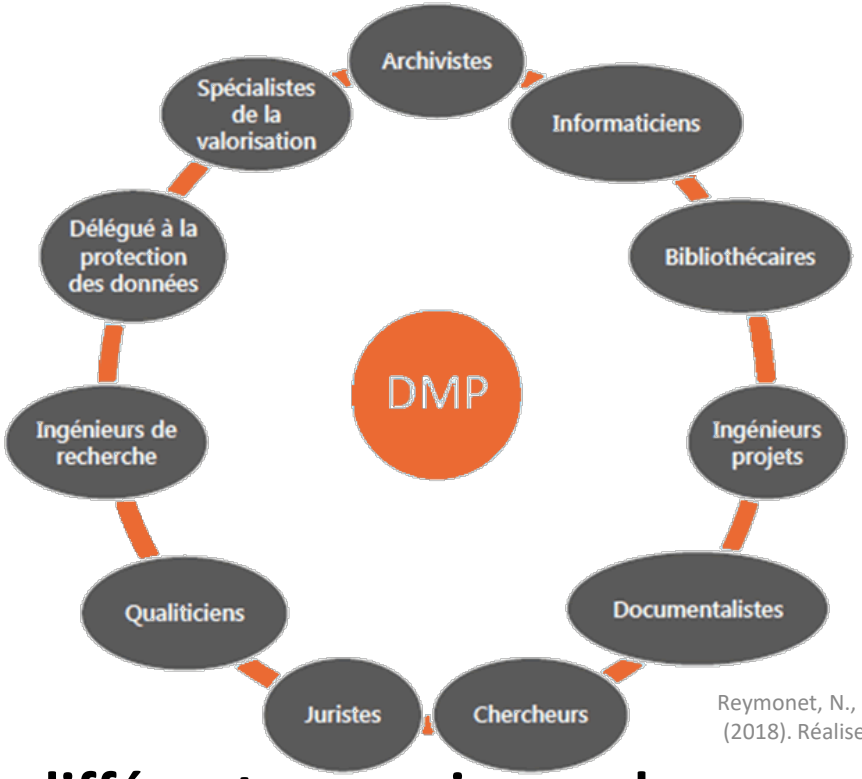
1-Le PGD est une aide à la gestion des données d'un projet

- **Inventaire des données produites ou utilisées dans le projet**
  Vue globale sur l'ensemble des données
- **Check list pour ne rien oublier d'important**
  à chaque étape du cycle de vie des données
- **Permet de rechercher collectivement à l'avance des réponses** spécifiques, précises, réalistes
  - de budgéter et chiffrer les ressources nécessaires
  - De se répartir les rôles
  - De mettre en place de bonnes pratiques communes pour minimiser les risques
- **Améliore la qualité et la reproductibilité de la recherche**
  - Pour des données plus fiables, compréhensibles, disponibles durant le projet
  - Préservées et valorisées si elles le méritent pour une réutilisation future

A commencer le plus tôt possible
(dans l'idéal au démarrage d'un projet)

2-Le PGD est un travail de groupe

Participent à la gestion des données et donc à la rédaction du PGD :



Reymonet, N., Moysan, M., Cartier, A., & Délémontez, R. (2018). Réaliser un plan de gestion de données « FAIR »

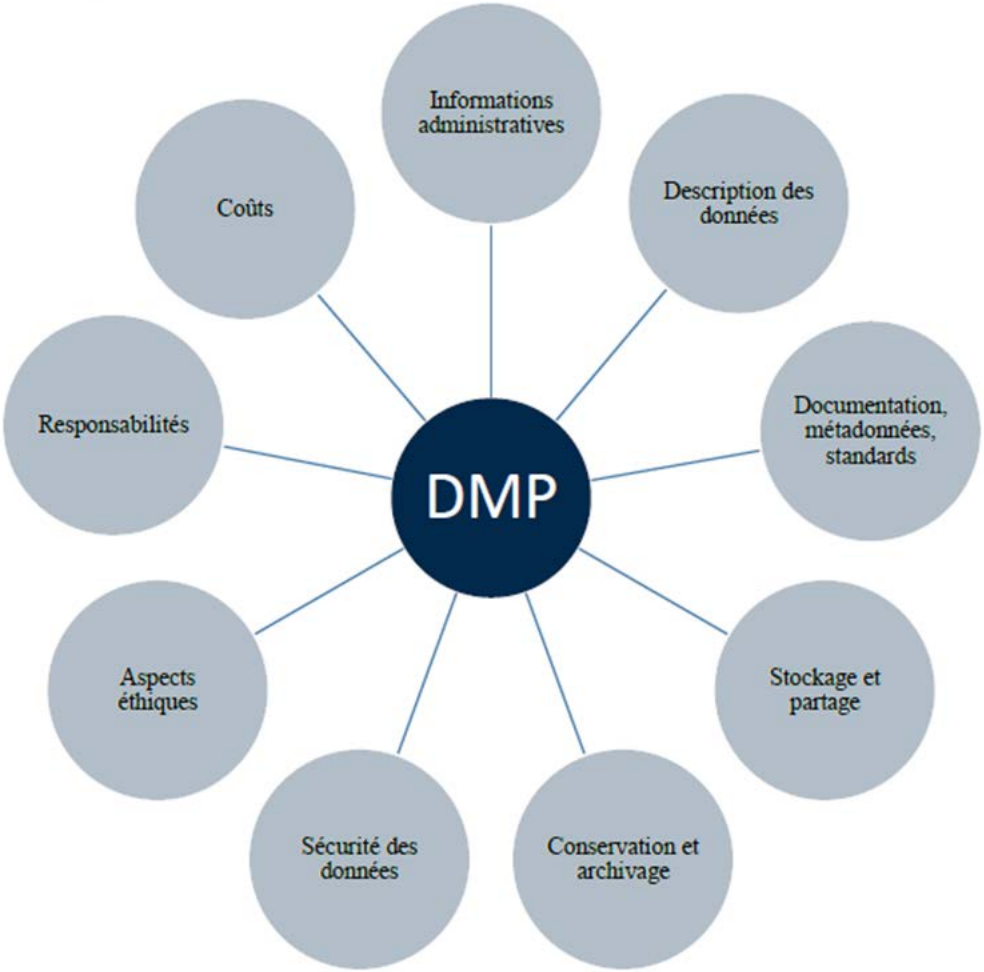
des différents organismes de tutelle

Voir les contacts IRD sur la [Bouée de sauvetage]

\+ **TOUTE L'EQUIPE**

\+ **TOUS LES PARTENAIRES** producteurs ou fournisseurs des données

3-Le PGD a toujours les mêmes rubriques



Fiche synthétique Doranum

Et on peut copier-coller certains paragraphes d'un PGD à l'autre (sécurité, partage,…)

4-Beaucoup d'aides en ligne *(voir la Bouée de sauvetage)*

En particulier, **des exemples réels dans votre domaine**, à trouver sur :

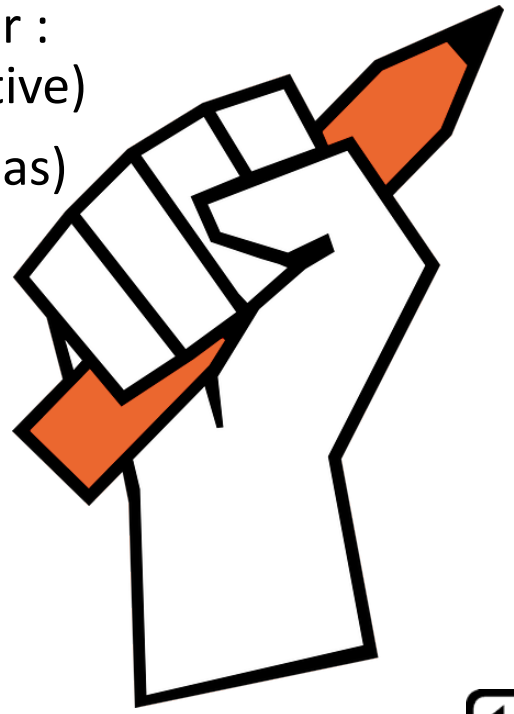
- ✓les plateformes de rédaction de PGD (rubrique public plans) comme DMP Opidor
- ✓les sites de conseil de référence comme DATApartage Inrae
- ✓les entrepôts de données comme Zenodo
- ✓les moteurs de recherche académiques comme OpenAIRE Explore
- ✓Et même certaines revues comme Research Ideas & Outcomes (RIO)

5-Un PGD est valorisant et valorisable

- ✓ **Nouvelles compétences** recherchées
- ✓ **Nouveau métiers**
  (Data manager, Data Scientist, Data designer,…)
- ✓ Nouvelles publications
  (DataPapers, publications de PGD)
- ✓ Nouvelles collaborations

6-Le PGD ne contraint pas, il explicite

- ✓ Un PGD ne contient pas de données
- ✓ Un PGD n'est pas public par défaut
- ✓ Si on ne peut pas partager des données expliquer juste pourquoi
- ✓ Si on veut se laisser le temps d'exploiter, de publier : décider d'un embargo (avec date limite approximative)
- ✓ On peut signaler ses données et les partager (ou pas) à la demande
- ✓ On peut signaler ses donner en ne publiant que des métadonnées (descriptions)



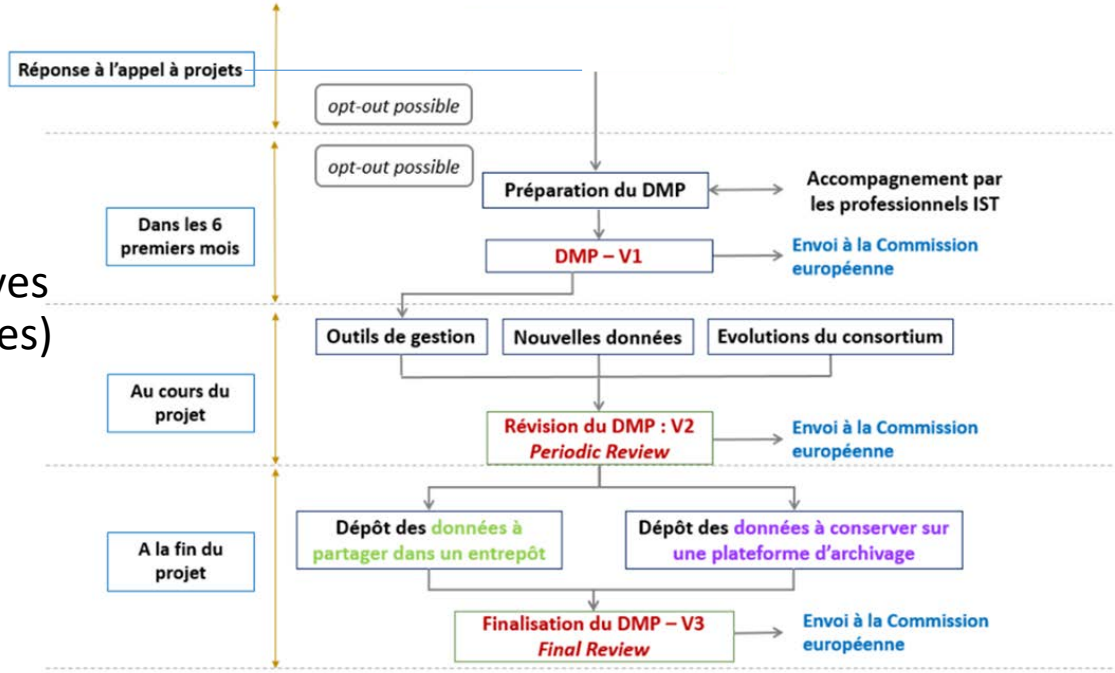
7-Le PGD est un document évolutif

Un PGD n'est pas figé :

- ✓ il doit être mis à jour <u>tout au long du projet</u>
- ✓ Il comporte des versions successives (certaines pouvant être des livrables)

  Exemple Horizon Europe :
  3 versions minimum
  (pas de maximum)

- Pas d'obligation de répondre à toutes les questions dans la V1



A. Cartier, M. Moysan, N. Reymonet, d'après les *Guidelines on FAIR data management*

8-Le PGD est un gage de bonne gestion des données

- ✓ Pour les bailleurs de fond :
  livrables pour l'ANR (2019), l'Europe (2017),…

- ✓ Pour les tutelles :
  recommandations des organismes de recherche (Feuille de route IRD), du gouvernement (PNSO, décret n° 2021-1572), de l'Europe,…

- ✓ Pour les pairs…

Il contribue à la traçabilité et à la reproductibilité des expériences et améliore la crédibilité scientifique.