

## ➤ PGD structure

Atelier distant sur la formation aux PGD - 25 juin 2020

Dominique L'Hostis, Sylvie Cocaud

# > Sommaire

- Qu'est ce qu'un PGD structure ?
- Les PGD structure à INRAE
  - un contexte favorable
  - la démarche
  - le modèle PGD structure INRAE
- Mise en œuvre et accompagnement
  - Rédiger un PGD structure
  - Accompagner les unités
  - Quelques exemples
- Où en est-on ?

# ➤ Qu'est ce qu'un PGD structure ?



## ➤ Qu'est ce qu'un PGD structure ?

- Document de référence sur la gestion des données de la structure
  - Mêmes objectifs que le PGD de projet : décrire comment une structure gère ses données de recherche et les ressources qu'elle y consacre.
- Structures concernées :
  - Unité (mixte) de recherche,
  - Infrastructure Scientifique Collective :
    - Unité ou Installation expérimentale,
    - Observatoire de Recherche en Environnement (ORE),
    - Centre de ressources biologiques ou Génétiques (CRB),
    - Plateforme analytiques et numériques, plateau technique,
  - Infrastructure de recherche



# ➤ PGD de projet vs PGD structure

## PGD de projet

- initié le plus souvent à la demande d'un financeur,
- porte sur le périmètre thématique et temporel du projet,
- organisé par jeu de données (ou autre "produit de recherche") dont le PGD formalise les étapes du cycle de vie

## PGD structure

- démarche volontaire de la structure ou demande de la direction (instrument de pilotage, enjeux patrimoniaux)
- périmètre plus large, durée plus longue
- organisé par type de données, type d'activité de la structure ou par mode de gestion ou processus.

- Distanciation sur les pratiques et identification des points d'amélioration
- Élément de la démarche qualité
- Élément-clé dans le cadre de l'Open Science et le respect des principes FAIR

## ➤ Les PGD structure à INRAE



# ➤ Les PGD structure à INRAE

Un contexte favorable

- Cadre institutionnel
  - Charte pour le libre accès aux publications et aux données
  - Charte des infrastructures de recherche
  - Campagne de labellisation des ISC de l'INRA
- Principes FAIR
- Départements pro-actifs

## ➤ Cadre institutionnel (1/3)

Charte pour le libre accès aux publications et aux données (2016)



- Outre le libre accès aux publications, l'engagement de l'Inra en faveur de l'ouverture des données de la recherche s'inscrit dans une double dynamique
  - Mise à disposition des données dans un objectif de transparence et de preuve en appui aux publications scientifiques
  - La réutilisation des données, dans un objectif de création de valeur et d'accélération de l'innovation
- « Les données hébergées dans les centres de données gérés par l'Inra ou ses prestataires sont préservées et archivées selon les prescriptions des **plans de gestion des données des projets ou des plateformes** »



# ➤ Cadre institutionnel (2/3)

## Charte des infrastructures de recherche (2016)



- Reprend beaucoup d'éléments intégrés dans un PGD sans y faire référence explicitement

### Accès aux infrastructures de recherche

- Les infrastructures de recherche (IR) de l'Inra sont ouvertes à toutes les unités de recherche de l'Inra ainsi qu'aux autres acteurs de la recherche publique ou privée. Des modalités de régulation peuvent être instaurées selon des combinaisons de principes de la charte.
- L'accès aux services offerts par les IR est fourni sous condition d'une double évaluation scientifique par les pairs et technique et économique par l'IR.
- Les données collectées dans le cadre des IR sont accessibles selon des procédures qualifiées.
- Les IR sont dotées d'instances collectives de gouvernance reconnues : celles-ci sont notamment en charge du respect de ces principes d'ouverture, de durabilité économique et d'accessibilité aux dispositifs ou aux données produites.

### Information

- Les IR apportent, via un dispositif unique et transparent, les informations sur les services et les matériels qu'elles fournissent, les procédures de soumission, les conditions d'usage et la gestion des données.
- Les IR disposent d'un règlement intérieur et d'un guide d'usage des infrastructures et dispositifs.
- Les utilisateurs apportent aux IR une information complète sur les expérimentations et les observations qu'ils comptent conduire dans les dispositifs, sur les matériels physiques ou biologiques qu'ils pourraient introduire ou utiliser ainsi que sur l'utilisation des données fournies, en accord avec les règles de confidentialité fixées d'un commun accord.
- Les IR et les utilisateurs s'assurent que les IR sont dûment citées dans les productions (publication, brevet etc.) auxquelles les IR ont contribué.

### Conditions économiques et financières

- L'accès aux services apportés par les IR respecte les règles de libre concurrence.\*
- Les charges à payer pour accéder aux services d'une IR permettent sa soutenabilité économique. La détermination de ces charges — variant des coûts marginaux aux coûts complets — fait l'objet de règles claires et explicites, modulées selon la nature des partenaires utilisateurs de l'IR et la nature des services fournis.

### Règlements, accords et engagements

- Les IR et leurs utilisateurs s'assurent que les projets sont menés dans le respect de l'éthique et de la déontologie ainsi que de la législation et des réglementations, notamment relatives à la protection du patrimoine scientifique et technologique, et à la sécurité des personnes, des biens, des données et de l'environnement.
- Les IR et les utilisateurs concluent ex ante des accords formels intégrant notamment les éléments de propriété intellectuelle et d'utilisation des données et résultats obtenus.
- Les données produites par les IR et financées par des fonds publics entrent dans le domaine public ou sont disponibles en accès libre, sous réserve de principes éthiques ou d'une éventuelle période d'usage restrictif liée aux dispositions légales et réglementaires (respect des libertés individuelles, secret industriel et commercial, propriété intellectuelle, protection du patrimoine scientifique et technique, etc.).
- Selon les accords, les IR garantissent tout ou partie de la confidentialité des données financées sur fonds privés.

\* L'exercice d'une activité économique portée par les IR se justifie, dans la limite de leurs compétences et des missions de l'Inra, par l'intérêt public qui peut notamment résulter de l'expertise unique et particulière proposée. Ces activités économiques s'inscrivent uniquement comme accessoire au service public de recherche ; elles contribuent en ce sens à l'équilibre financier ou permet d'amortir les investissements réalisés.

## ➤ Cadre institutionnel (3/3)

### Labellisation des infrastructures collectives et de recherche de l'Inra

Les critères pour la labellisation des **Infrastructures Scientifiques Collectives**

- ➔ **Pertinence scientifique** et technologique, soutenue par un porteur scientifiques et CD
- ➔ Les **3 missions** d'IR : production de services, développement techno, formation
- ➔ Une **démarche qualité** avérée et reconnue voire une certification
- ➔ Les critères de la **charte** des IR dont la **conformité** avec les règles éthiques (Nagoya, BE) et diverses réglementations en termes de sécurité (PPST, personnel, env., biologique)
- ➔ Les critères de **soutenabilité économique** : compte d'exploitation prévisionnel, tarification, engagement des Départements, stratégie territoriale
- ➔ Une politique claire en matière **d'accès aux données (data plan)**;

The guideline: F.A.I.R.

#### Le dossier

##### ➔ Volet 4 - Données et gestion des données (par fonction quand pertinent)

- ❑ **Stratégie générale** de l'entité en matière de données, notamment sur le concept **FAIR**, actuelle et projetée
- ❑ **Les données** : nature et volumétrie des données produites actuelles et futures, données produites ailleurs, si oui nature et volumétrie
- ❑ **Les prestataires** : prestataire ou partenaire extérieur, pour gérer ses données, nature et volumétrie des données, nature des services (le stockage, la préservation, le calcul, le traitement, la diffusion), CATI impliqué(s) le cas échéant
- ❑ **Modalités de stockage et ou d'archivage** : serveurs internes, ailleurs à l'Inra, à l'extérieur de l'Inra, mesures de préservation utilisées
- ❑ **Capacité de gestion et d'ouverture des systèmes d'information** : Base de donnée utilisée, gestion des droits, production et gestion de métadonnées, capacités à rendre accessibles les données et métadonnées selon les normes ou standards du domaine, exposition des données ...
- ❑ **La formalisation des échanges** : documentations, plan de gestion par l'entité ou par les utilisateurs
- ❑ **Modalités d'accès et de diffusion des données** pour des utilisateurs hors de l'entité : catalogue et de fonctions de recherche, Attribution de DOI aux jeux de données, Utilisation de licences, Production de Bases de données, de logiciel, contraintes à la diffusion le cas échéant
- ❑ **Sécurité** : procédures, limitations et régulations de l'accès, CNIL et données relatives à l'homme
- ❑ **Economie de la donnée** : Coûts des charges opérationnelles (hors salaires permanents) directes liée au traitement et la gestion des données, Temps (ETP) des CATI dédiés à l'entité, personnels affectés aux données, décrire les compétences actuelles et celles projetées (Cf. volet 3)

# ➤ Les PGD structure à INRAE

## Origine de la démarche à INRAE

- Demande de participants aux formations "rédiger un PGD"
- Leurs besoins
  - Organiser la gestion des données de leur structure
    - cartographier les activités autour des données,
    - définir/mettre en œuvre une politique données,
    - identifier les moyens nécessaires,
    - développer les systèmes d'Information,
    - contribuer à l'Open Science/Data.
  - Disposer d'un document de référence, versionné
    - en interne pour l'ensemble des acteurs "données", et pour la direction de la structure
    - en externe pour évaluer les pratiques (adéquation aux principes FAIR) et dialoguer avec les utilisateurs (IR)
  - Faciliter la production des PGD de projets

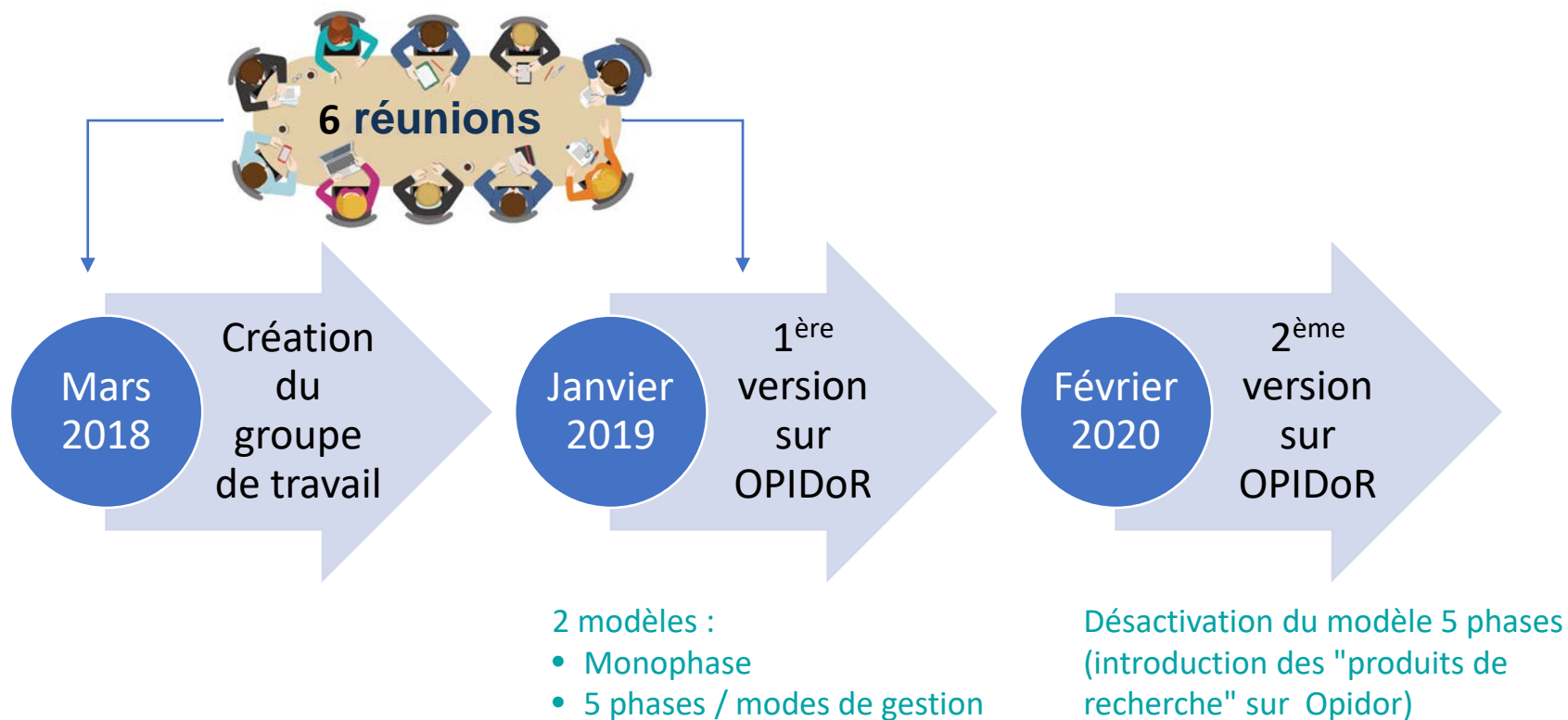
# ➤ Les PGD structure à INRAE

## Groupe de travail

- Animation : Sylvie Cocard et Dominique L'Hostis (DIST)
- 14 membres : 9 structures , 5 départements scientifiques
- UR et UMR
  - Ecologie des Forêts Méditerranéennes (Provence-Alpes-Côte d'Azur)
  - Institut de Génétique, Environnement et Protection des Plantes (IGEPP) (Bretagne-Normandie)
  - Laboratoire Sols et Environnement (Grand-Est-Nancy)
  - Agroressources et Impacts environnementaux (Hauts de France)
  - Ecologie Comportementale et Biologie des Populations de Poissons (Ecobiop) (Nouvelle-Aquitaine Bordeaux)
- UE
  - Ecologie et Ecotoxicologie Aquatiques (Bretagne-Normandie)
- Plateformes et observatoire
  - Plateforme Génobois (Val de Loire)
  - Plateforme Métabolome - Fluxome - Unité Biologie du Fruit et Pathologie (Nouvelle-Aquitaine Bordeaux)
  - Observatoire des Programmes Communautaires de Développement Rural (Occitanie-Toulouse)

# ➤ Les PGD structure à INRAE

## Chronologie



# ➤ Les PGD structure à INRAE

## Mise en œuvre (1/2)

- Des débuts un peu laborieux ...
  - Difficultés initiales de compréhension :
    - entre les participants issus de communautés très variées, travaillant sur des types de données et des environnements très différents, et avec des attentes diverses,
    - entre les participants et les animatrices ...
- Plan de travail
  - Recherche de PGD structure existants dans d'autres organisations
  - Travail collectif autour d'un template pouvant répondre à des objectifs divers :
    - PGD générique d'unité : cadre portant la politique de la tutelle et les moyens
    - PGD par activité (ex. des plateformes réalisant des prestations diverses),
    - PGD par sources de données ... (big data, données d'enquêtes, données des observatoires ...)
  - Articulation PGD projets / PGD structures



# ➤ Les PGD structure à INRAE

Mise en œuvre (2/2)

- Méthodologie : travail collaboratif et itératif à partir du modèle INRA pour les projets
  - Mise à disposition d'une maquette « martyre » sous Opidor (pré-production)
  - Suppression des éléments non pertinents
  - Adaptation des questions, listes fermées pour les réponses si possible
  - Révision de l'aide en ligne
  - Recherche de l'architecture la plus adaptée → 2 maquettes
    - monophasé
    - multiphasés (5 modes de gestion)
- Mise en production des deux modèles de PGD Structure
  - Communication + large
- Désactivation du modèle 5 phases (12/2019)

## ➤ Quelques exemples de PGD structures externes

- **DISSCo** Distributed System of Scientific Collections [DMP](#)
- **EMBRC** European Marine Biological Resource Centre [DMP](#)
- **FGCI** Finnish Grid and Cloud Infrastructure [DMP](#)
- **IDDO** Infectious Diseases Data Observatory [DMP](#)
- **NCI** National Computational Infrastructure [DMP](#)
- **NOAA** National Oceanic and Atmospheric Administration [DMP](#)
- **TERENO** TERrestrial ENvironmental Observatories [DMP](#)





# ➤ Les PGD Structure à INRAE

## Le modèle de PGD structure

Informations sur la structure ( 7 questions )

Informations sur le plan de gestion ( 2 questions )

Présentation générale des données ( 6 questions )

Droits de propriété intellectuelle ( 1 question )

Confidentialité ( 3 questions )

Partage des données ( 6 questions )

Organisation et documentation des données ( 5 questions )

Stockage et sécurité des données ( 5 questions )

Archivage et conservation des données ( 4 questions )



## Informations sur la structure ( 7 questions )

### Nom de la structure

#### Type de structure

- ☐ Centre
- ☐ Département
- ☐ Unité de recherche, Unité ou Installation Expérimentale
- ☐ Unité de service
- ☐ ORE
- ☐ SOERE
- ☐ CRB (Centre de ressources biologiques), CRG (CR Génétiques)
- ☐ Plateforme, plateau technique
- ☐ ISC (Infrastructure Scientifique Collective)
- ☐ Infrastructure de recherche
- ☐ Autre : à préciser dans la zone "Informations supplémentaires"

#### Identifiant de la structure

*Préciser le fournisseur de l'identifiant (ISNI, VIAF, FundRef, DataCite..)*

#### Responsabilités dans la structure

Nom, Prénom	Courriel	Rôle
		Responsable scientifique
		Responsable technique
		Responsable de la conservation à long terme

**Financier(s) (permettant l'acquisition des jeux de données – hors projet)**

#### Etablissement(s) tutelle(s)

##### Département de rattachement INRAE (ou anciens départements Inra)

- ☐ AGROENV
- ☐ ALIM-H
- ☐ AQUA
- ☐ BAP
- ☐ CEPIA
- ☐ EA
- ☐ ECOFA
- ☐ ECOSOCIO
- ☐ EFPA
- ☐ GA
- ☐ MIA
- ☐ MICA
- ☐ NUMM
- ☐ PHASE
- ☐ SA
- ☐ SAD
- ☐ SAE2
- ☐ SPE
- ☐ TERRA
- ☐ TRANSFORM
- ☐ Autre (à préciser)

##### Département de rattachement INRAE (ou anciens départements Inra)

- ☐ AGROENV
- ☐ ALIM-H
- ☐ AQUA
- ☐ BAP
- ☐ CEPIA
- ☐ EA
- ☐ ECOFA
- ☐ ECOSOCIO
- ☐ EFPA
- ☐ GA
- ☐ MIA
- ☐ MICA
- ☐ NUMM
- ☐ PHASE
- ☐ SA
- ☐ SAD
- ☐ SAE2
- ☐ SPE
- ☐ TERRA
- ☐ TRANSFORM
- ☐ Autre (à préciser dans la zone de commentaires)



**INRAE DipSO**

Atelier distant sur la formation aux PGD

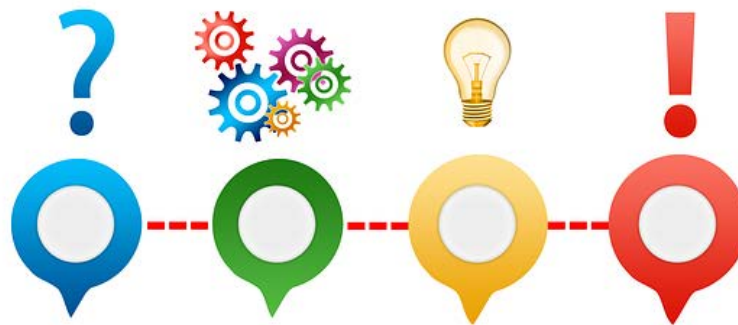
25 juin 2020 / Dominique L'Hostis et Sylvie Cocaud

# ➤ Mise en œuvre et accompagnement

Comment rédiger un PGD structure ?

Accompagnement des unités

Quelques exemples...



# ➤ Mise en œuvre et accompagnement

Comment rédiger un PGD structure ?

- **identification des données** gérées par la structure,
- choix d'un **mode d'organisation** :
  - par type d'activité ou par mode de gestion des données,
  - selon le mode d'obtention, la nature ou le format des données
- **créer un PGD structuré** selon le mode d'organisation choisi.

Exemple : selon la typologie des fonctions utilisées pour la labellisation des ISC

- Expérimentation, Observation,
- Production de matériel biologique / échantillons
- Conservation de ressources biologiques (CRB)
- Données issues de plateaux analytiques
- Données de génomique
- Codes informatiques
- Données issues de modélisation
- Données d'enquête

# ➤ Mise en œuvre et accompagnement

Accompagner les unités

- Des départements pro-actifs
  - Ecologie des Forêts, Prairies et Milieux Aquatiques (EFPA)
    - 2 jours de formation et atelier, pour les représentants des unités du département (juillet 2019)
  - Biologie et Amélioration des Plantes (BAP) - Santé des Plantes (SPE) - Environnement Agronomie (EA) - (février 2020)
    - Action spécifique vers leurs 20 unités expérimentales
      - classes virtuelles dédiées
      - relecture des PGD
      - restitution et bilan lors d'un séminaire des 3 départements
- Un service support

# ➤ Exemple 1 : Genobois (1/3)



## Plateforme de phénotypage des propriétés physico-chimiques du bois et hydrauliques des arbres

**GénoBois = Plateforme de phénotypage des propriétés physico-chimiques du bois et hydrauliques des arbres.**

- Labellisée « Outil collectif INRA » par la CNOC en 2008 et 2013
- Labellisée « ISC » en 2018

• Particularités de GénoBois:

✓ **3 UMR de rattachement**

- BioForA – Orléans
- BIOGECO – Bordeaux
- PIAF - Clermont



✓ **4 sites**

- INRA Orléans
- INRA Bordeaux-Ces
- Université Bordeaux
- INRA Clermont-Crou



### GénoBois produit et gère des données analytiques d'une grande diversité

- **Des données brutes** (en sortie des équipements), formats de fichiers propriétaires, avec **métadonnées techniques**,
- **Des données transformées interopérables**
  - produites à partir des données brutes
  - converties sous des formats numériques non propriétaires (csv, txt, tif).

### GénoBois transmet à ses utilisateurs ...

- **Des données transformées interopérables**
- **Une documentation** sous forme de compte-rendu de travaux (fichier txt ou PDF) afin de faciliter leur compréhension et leur réutilisation.

... Pour qu'ils produisent des **données élaborées valorisables**

# ➤ Exemple 1 : Genobois (2/3)

## Retour d'expérience

### - Participation **GénoBois** aux classes virtuelles PGD début 2018

- > besoin d'un PGD associé à une PF
- > intégration du GT « **PGD Structure** »

infodoc @n-line  
express

### - **Elaboration de la V1 du PGD GénoBois**

- > modèle INRA - Trame Structure
- > remplissage des rubriques à partir des éléments du volet 4 du dossier de labellisation ISC
- > rédaction collaborative facilitée par l'outil DMP Opidor

The screenshot displays the DMP Opidor web interface. At the top, there is a navigation bar with links: 'Tableau de bord', 'Créer des plans', 'DMPs publics', 'Modèles de DMP', and 'Aide'. Below this, the INRA logo is visible, along with the text 'SCIENCE & IMPACT'. The main content area is titled 'Plan de gestion de données de la Plateforme GénoBois v1'. It features a sidebar with a list of sections, each with a plus icon for expansion. The sections are: 'Informations sur la structure (6 / 7)', 'Informations sur le plan de gestion (1 / 2)', 'Présentation générale des données (6 / 6)', 'Droits de propriété intellectuelle (1 / 1)', 'Confidentialité (3 / 3)', 'Partage des données (6 / 6)', 'Organisation et documentation des données (5 / 5)', 'Stockage et sécurité des données (5 / 5)', and 'Archivage et conservation des données (4 / 4)'. The top of the main content area shows 'Renseignements sur le projet', 'Vue d'ensemble', 'Rédiger', 'Partager', and 'Télécharger'. A progress bar indicates '37/39 réponse(s)'.

INRAE DipSO

Atelier distant sur la formation aux PGD

25 juin 2020 / Dominique L'Hostis et Sylvie Co

## ➤ Exemple 1 : Genobois (3/3)



### Difficultés rencontrées pour la construction du PGD GénoBois

Travail important en amont pour remplir les rubriques

-> impliquer les IP et motiver les utilisateurs de la PF (scientifiques des UMRs d'adossement)

Confidentialité/ droits de propriétés intellectuels/ durée de conservation des données

-> manque de connaissances pour pouvoir remplir ces rubriques

-> pas de politique claire au niveau de la CNOC pour homogénéité des pratiques entre les PF

Qualité du document exporté via DMP Opidor à améliorer -> mise en page/insertion d'images



### Les avantages d'avoir construit notre PDG

Répondre aux critères des chartes Open Access et IR

Améliorer nos pratiques -> produire et gérer nos données selon principes FAIR

Suivre la politique INRA *Pilotage par le Management Qualité en Recherche*

Donner de la valeur ajoutée aux données produites par GénoBois

Pouvoir être partenaire de projets ANR, H2020, ..., pour lesquels la production d'un PGD est un prérequis

Faire évoluer et formaliser nos modes de stockage / archivage en local (UMRs)

-> V6 du cycle de vie



## ➤ Exemple 2 : U3E (1/4)

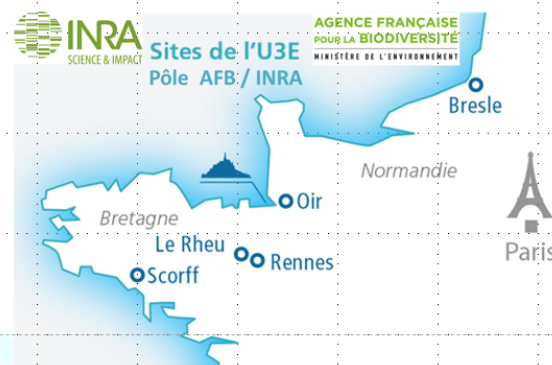
EXAMPLE

### Unité expérimentale

#### U3E, Unité Expérimentale d'Ecologie et d'Ecotoxicologie aquatique

Rattachée au Dpt INRA EFPA, en partenariat étroit avec l'AFB (Agence Française pour la Biodiversité)

Une unité multi-sites



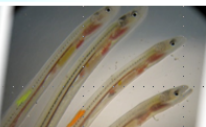
#### 3 fonctions principales autour des milieux aquatiques

- Appui à la recherche
- Formation
- Développements méthodologiques et technologiques

#### 3 missions principales

- Expérimentation / manipulation
  - Systèmes expérimentaux et milieux naturels (+ production de matériel biologique)
- Suivis à long terme
  - Dans les milieux naturels (sites ateliers)
- Centre de Ressources Biologiques (CRB)
  - Colisa (Collection of Ichthyological Samples)

Plan de gestion  
de données



DÉPARTEMENT ECOLOGIE DES FORÊTS, PRAIRIES ET MILIEUX AQUATIQUES

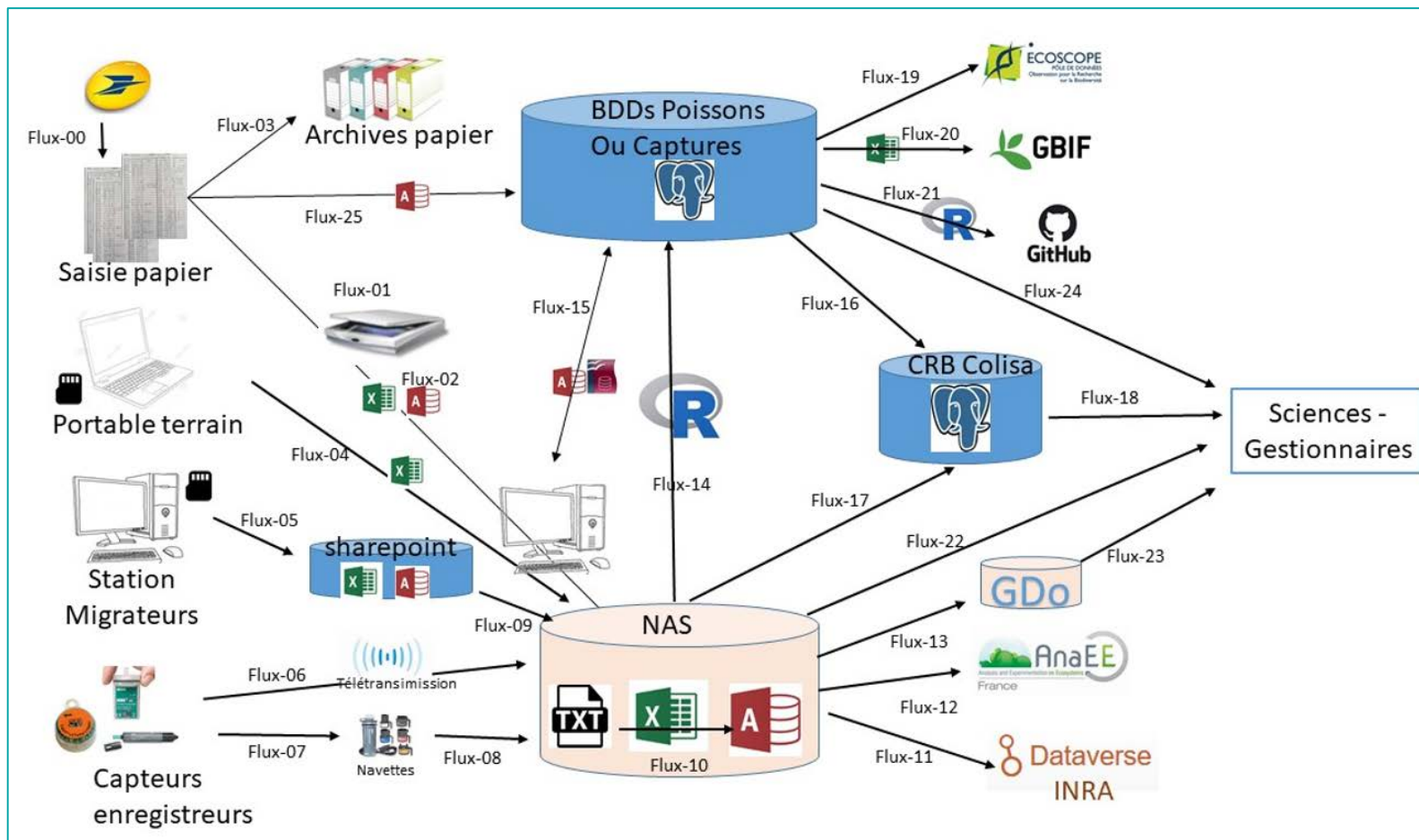
INRAE DipSO

Atelier distant sur la formation aux PGD

25 juin 2020 / Dominique L'Hostis et Sylvie Cocaud

## ➤ Exemple 2 : U3E (2/4)

### Représentation des flux de données



## ➤ Exemple 2 : U3E (3/4)

Les modes de gestion dans DMP-Opidor (produits de recherche)

### U3E

[Renseignements sur le projet](#)[Produits de recherche](#)[Modèle choisi](#)[Rédiger](#)[Partager](#)[Télécharger](#)

**Produit de recherche : jeu de données, logiciel, workflow, échantillon, protocole...**

A renseigner séparément pour des produits de recherche nécessitant une gestion spécifique à leur nature ou discipline.


\* **Nom abrégé (20 caractères max.)** ORE DiaPFC

\* **Nom complet** Observatoire de Recherche en Environnement sur les Diadromes dans les Fleuves Côtiers

\* **Type** ⓘ Other

\* **Autre type** Observatoire de Recherche en Environnement

**Identifiant Pérenne** ORE DiaPFC






\* **Nom abrégé (20 caractères max.)** PEARL

\* **Nom complet** PEARL

\* **Type** ⓘ Dataset

**Identifiant Pérenne**






\* **Nom abrégé (20 caractères max.)** COLISA

\* **Nom complet** COLlection of Ichthyological SAmples

\* **Type** ⓘ Collection

**Identifiant Pérenne** <https://doi.org/10.15454/D3ODJM>






\* **Nom abrégé (20 caractères max.)** CNICS

\* **Nom complet** Centre National d'Interprétation des Captures de Salmonidés

\* **Type** ⓘ Dataset

**Identifiant Pérenne**






\* **Nom abrégé (20 caractères max.)** Expérimentations

\* **Nom complet** Expérimentations

\* **Type** ⓘ Dataset

**Identifiant Pérenne**



[Ajouter un produit de recherche](#)

[Sauvegarder](#)

## ➤ Exemple 2 : U3E (4/4)

### PGD finalisé (extraction DMP-Opidor + mise en forme)

#### Plan de Gestion des données

#### Unité Expérimentale d'Ecologie et d'Ecotoxicologie aquatique



Plan de gestion de données créé à l'aide de DMP OPIDoR



Modèle INRA

Créateurs : Frédéric Marchand

Affiliation : INRA - Institut National de la Recherche Agronomique

#### Table des matières

Résumé du projet :	3
U3E-01 - Informations générales :	4
Informations sur la structure :	4
Nom de la structure :	4
Type de structure :	4
Identifiant de la structure :	4
Responsabilités dans la structure :	4
Etablissement(s) tutelle(s) :	4
Financier(s) (permettant l'acquisition des jeux de données - hors projet) :	4
Informations sur le plan de gestion :	4
Présentation des différents modes de gestion proposés par la structure :	4
Historique des versions DMP OPIDoR ne gère pas les versions successives d'un plan. Pour en garder trace, indiquer la version dans le titre et copier le plan pour générer la version suivante :	4
Suivi à long terme :	5
Présentation générale des données :	5
Mode d'obtention des données :	5
Origine des données :	5
Type de données :	7
Nature des données :	7
Format des données :	7
Droits de propriété intellectuelle :	7
Qui détient les droits sur les données et les autres informations créées ? :	7
Confidentialité :	7
Identification des jeux de données contenant des données confidentielles :	7
Quelles sont les mesures prises et les normes auxquelles il est nécessaire de se conformer pour garantir cette confidentialité ? :	7
Partage des données :	8
Y a-t-il une obligation de partage (ou à l'inverse une interdiction ou une restriction) ? :	8
Quelles sont les réutilisations potentielles de ces données ? :	8
La lecture des données nécessite-t-elle le recours à un logiciel ou un outil spécifique ? Si oui, lequel ? :	8
Comment les données seront-elles partagées ? :	8
Avec qui ? :	9
Organisation et documentation des données :	9
Quelles méthodes et outils sont utilisés pour acquérir et traiter les données, depuis leur acquisition jusqu'à leur mise à disposition, leur archivage ou leur destruction ? :	9

Quelles métadonnées seront utilisées pour accompagner le jeu de données ? Quels seront les standards, vocabulaires, taxonomies... utilisés pour décrire et représenter les données et éléments de métadonnées ? Comment les métadonnées seront-elles produites et mises à jour ? .....10

Une documentation complémentaire aux métadonnées est-elle nécessaire pour décrire les données et assurer leur réutilisabilité sur le long terme ? .....11

Comment les fichiers de données sont-ils gérés et organisés : contrôle des versions, conventions de nommage des fichiers, organisation des fichiers .....11

Quelle est la procédure de contrôle qualité des données ? Joindre éventuellement le plan d'assurance qualité .....11

Stockage et sécurité des données .....12

Quels sont les types de flux empruntés par les données et les supports utilisés pour les stocker ? .....12

Quelle est la volumétrie actuelle et prévisionnelle ? .....14

L'entité hébergeant physiquement les données a-t-elle une politique de sécurité pour son système d'information ? (politique locale, charte des infrastructures de recherche) .....14

Sécurité - Confidentialité : les données font-elles l'objet d'échange ou de partage avec de tiers acteurs et selon quelles modalités ? comment sont déterminés les droits d'accès aux données avant leur publication ? .....14

Archivage et conservation des données .....14

Quelles sont les données à conserver sur le moyen ou le long terme et quelles sont les données à détruire ? .....14

Quelle est la durée de conservation des données ? .....15

Quelles garanties de financements couvriront les coûts associés à la conservation à long terme ? .....15

#### RÉSUMÉ DU PROJET :

L'U3E est une unité expérimentale de l'Institut national de la recherche agronomique (INRA), rattachée au département EFPA (Ecologie des forêts, des prairies et des milieux aquatiques). Elle constitue un support aux recherches menées sur les écosystèmes aquatiques. Ses travaux s'effectuent dans le cadre d'infrastructures nationales de recherche : ANAEE-France Analyses et expérimentations sur les écosystèmes continentaux et leur biodiversité ORE DiaPFC Observatoire de recherche en environnement sur les poissons diadromes dans les fleuves côtiers (effet des changements environnementaux, climatiques, liés à l'agriculture). Par ses activités, l'originalité de ses outils et ses compétences élargies (écologie et dynamique des populations, écotoxicologie et qualité des écosystèmes aquatiques, aquaculture), l'U3E intervient dans des programmes interdisciplinaires faisant appel à des organismes de recherche français et étrangers. Toutes les activités de l'U3E sont effectuées selon le système de management environnemental (SME) certifié ISO 14001.

Dernière modification : 10/05/2019

## ➤ Exemple 3 : URFM (1/4)

### Unité de recherche

EXAMPLE

#### URFM

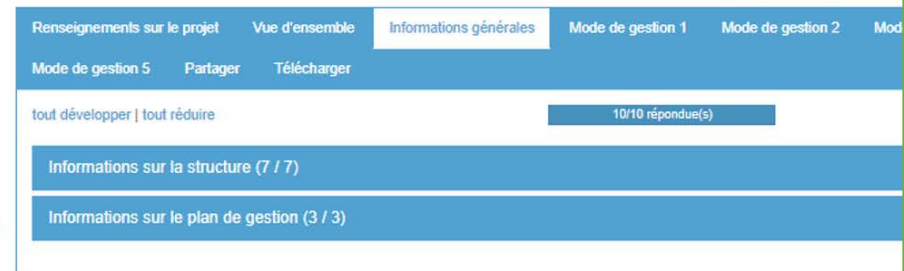
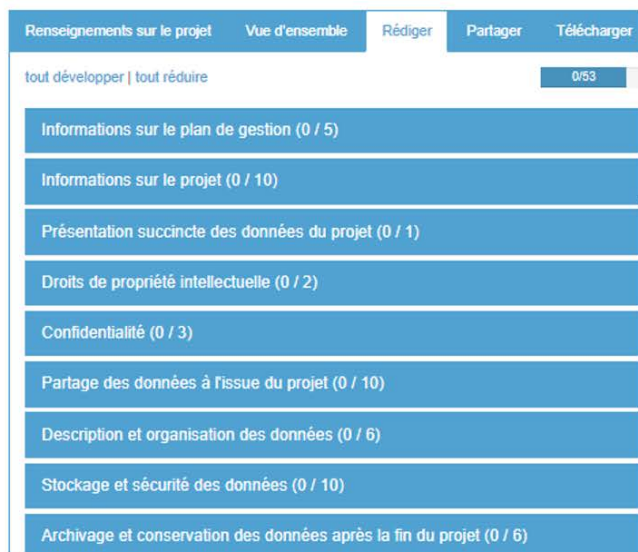
- ▶ Unité URFM (écologie forestière, génétique, environnement)
- ▶ Forte proximité thématique avec l'UEFM
- ▶ Cultures BdD et SI partagées
- ▶ Production de données à flux constant sur long termes (pas de temps > projets)
- ▶ Valorisation vers SI nationaux (Gnpis, AnaEEn, ICOS ...)





## ➤ Exemple 3 : URFM (2/4)

### PGD de projet vers PGD de structure (modèles INRA)



1 ou 5 modalités



## ➤ Exemple 3 : URFM (3/4)

La démarche interne

### Processus d'élaboration

- ▶ Un DU et un Conseil de service « acquis à la cause »
- ▶ Un animateur
- ▶ Choix du modèle multi modalités
- ▶ Identification des fonctions/services
- ▶ Un collectif de contributeurs (délégation)
- ▶ *Un groupe de relecture & validation*

#### Difficultés à prévoir

Identification des fonctions  
Se limiter à l'état de l'art  
Valider/Arrêter le travail

Accompagnement  
Relances

Montpellier Science Ouverte : Plan de Gestion de Données

17 Juin 2019

10

## ➤ Exemple 3 : URFM (4/4)

PGD finalisé – Les modes de gestion

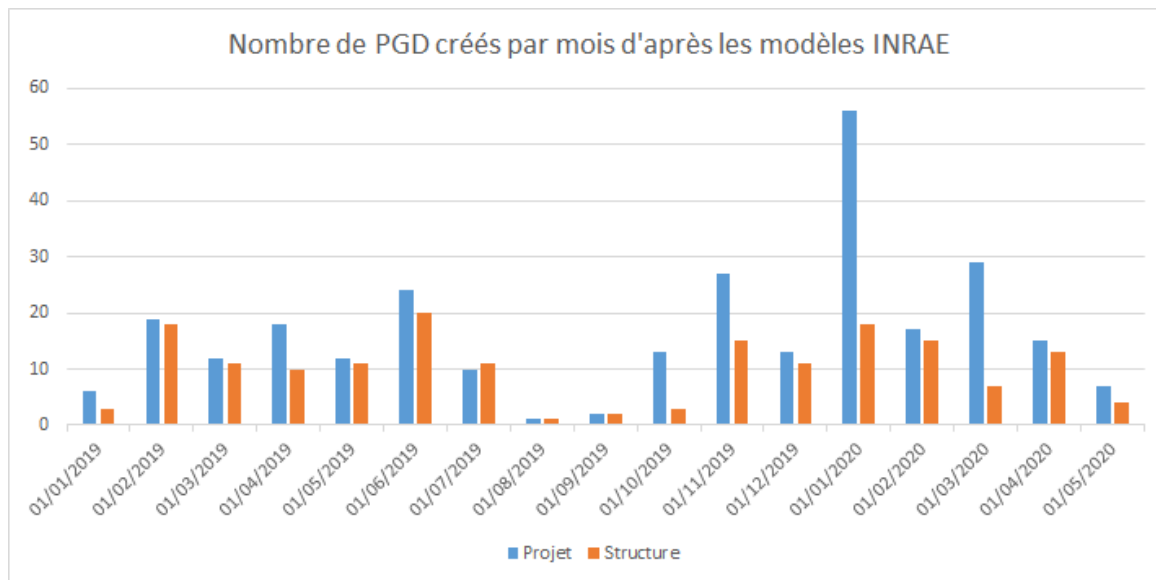
Fonction ou service	Phase PGD
Observations et expérimentations en forêt ou laboratoire (de sylviculture et génétique)	Mode de gestion 1
Observations et expérimentations IR ICOS SOERE F-ORE-T (AnaEE-France)	Mode de gestion 2
Echantillons (combustibles, feuilles, graines)	Mode de gestion 3
Analyses génétiques/génomiques	Mode de gestion 4
Codes/Modèles/Simulations	Mode de gestion 5



## ➤ Où en est-on ?

Un intérêt reconnu pour les PGD Structure

- Demande initiale des scientifiques et utilisation réelle



Depuis 2016 : 409 PGD projets, 173 PGD structures

- Intérêt de certaines communautés et incitation de certaines directions scientifiques

# > Où en est-on ?

Perspectives et questionnements

- Liens possibles entre :
  - PGD Structure et PGD projet : encore à concrétiser
  - PGD Projet-Structure et d'autres documents exigibles :
    - documents d'analyse de risques RSSI
    - certificat de conformité RGPD
- Changement ou évolution des outil
  - Data Stewardship Wizard / DMP-OPIDoR



## > Contacts

### Support PGD INRAE



[sylvie.cocaud@inrae.fr](mailto:sylvie.cocaud@inrae.fr)  
[dominique.lhostis@inrae.fr](mailto:dominique.lhostis@inrae.fr)

