

---

## **PGD UEFP**

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

**Créateurs du PGD** : Laetitia Deyris, Thierry Labbé, Céline Meredieu

**Affiliation du créateur principal** : INRAE - Institut national de recherche pour l'agriculture l'alimentation et l'environnement

**Modèle du PGD** : INRAE - Trame Structure

**Dernière modification du PGD** : 29/06/2021

**Financeur** : INRAE

### **Résumé du projet :**

#### **Plan de Gestion des données de l'UEFP**

L'UEFP est une unité expérimentale de l'Institut National de la Recherche pour l'Agriculture l'Alimentation et l'Environnement (INRAE), rattachée au département ECODIV (Ecologie et Diversité). Elle constitue un support aux recherches menées sur les écosystèmes forestiers.

Ses travaux s'effectuent dans le cadre d'infrastructures nationales de recherche : AnaEE-France, ICOS Europe, RARe Pilier Forêt, In-Sylva France.

Toutes les activités de l'UEFP sont effectuées selon le système de management environnemental (SME) certifié ISO 14001.

L'UEFP génère une grande quantité de données, de nature variée et ayant une grande importance patrimoniale. Ces données sont organisées dans différents systèmes d'information.

**Chercheur Principal** : Frédéric BERNIER

**Contact pour les Données** : Frédéric BERNIER

### **Droits d'auteur**

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source.

L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

## Informations sur la structure

UE0570 Unité Expérimentale Forêt Pierroton

Lien vers l'annuaire des structures du Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation

- Unité de recherche, Unité ou Installation Expérimentale
- ISC (Infrastructure Scientifique Collective)

INRAE - UEFP

Site de Recherches Forêt Bois de Pierroton

69 route d'Arcachon

33612 CESTAS Cedex - FRANCE

<https://www6.bordeaux-aquitaine.inrae.fr/ue-pierroton>

L'UEFP contribue aux trois missions dévolues aux infrastructures scientifiques collectives de INRAE (dont 62 % pour la participation et l'appui au processus recherche). Son action s'inscrit au profit de 3 fonctions principales :

- Expérimentation (installation et gestion de dispositifs expérimentaux et de réseaux, phénotypage d'arbres forestiers, récoltes et gestion d'échantillons variés, expérimentation systèmes sylvicoles,...). Les principales essences concernées sont le pin maritime, les chênes, l'eucalyptus, le douglas, le pin taeda mais l'unité travaille également sur de nombreuses autres essences.
- Observation (gestion de sites ateliers & d'ORE/SOERE, installation et maintenance des infrastructures et de l'instrumentation scientifique, pilotage d'une halle biomasse,...)
- Gestion et conservation de ressources biologiques (production de matériel végétal original, gestion et conservation de ressources génétiques y compris les parcs à clones et à pied-mère) pour des espèces forestières variées (pin maritime, chênes, peuplier, douglas, robinier, mélèze, merisier, châtaignier, pins noirs, frêne, tulipier, pin sylvestre, sapins, pin taeda,...).

L'UEFP gère un vaste patrimoine foncier associé aux expérimentations (plus de 700 ha en propre et sous convention), offrant un potentiel d'accueil apprécié. En parallèle, l'UEFP possède une pépinière de 4 ha permettant la production de matériel végétal, ainsi que des locaux techniques dédiés au conditionnement des graines et des pollens et une halle dédiée à la mesure de la biomasse forestière installée dans le cadre du PIA Xyloforest.

Outre les compétences en observation, expérimentation et gestion de matériel végétal au sens large, l'UEFP maîtrise des savoir-faire en création/administration de SI et de BDD ainsi qu'en installation et suivi d'instrumentations scientifiques.

<https://doi.org/10.15454/1.5483264699193726E12>

Nom, Prénom	Courriel	Rôle
Frédéric BERNIER	frederic.bernier@inrae.fr	Directeur d'unité
Patrick PASTUSZKA	patrick.pastuszka@inrae.fr	Directeur d'unité adjoint
Alexia MATHOU	alexia.mathou@inrae.fr	Ingénieure d'étude
Céline MEREDIEU	celine.meredieu@inrae.fr	Chargée de recherche et responsable PGD
Annie RAFFIN	annie.raffin@inrae.fr	Ingénieure de recherche
Laetitia DEYRIS	laetitia.deyris@inrae.fr	Responsable PGD
Thierry LABBE	thierry.labbe@inrae.fr	Responsable PGD

INRAE

- EFPA
- Autre (à préciser dans la zone de commentaires)

**ECODIV**

Ecologie et biodiversité.

Le département ECODIV a pour mission de produire et transférer des connaissances sur la structure, le fonctionnement et l'évolution des écosystèmes forestiers, prairiaux et aquatiques. Ses travaux permettent de proposer et co-construire avec ses partenaires des stratégies de gestion afin de répondre aux enjeux multiples de résilience et adaptation de ces écosystèmes aux changements globaux, de biodiversité et de bioéconomie.

- INRAE Budget SE (Subvention d'Etat)
- Recettes

## Informations sur le plan de gestion

Question sans réponse.

Date	n° de version	Status	Auteur	Affiliation de l'auteur (se reporter à l' <i>annuaire Inrae</i> )	Validé par	Validé le
07/02/2020	0.1	Non publié	Laetitia Deyris et Céline Meredieu	<a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J</a> <a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B</a>		
16/04/2020	0.2	Non publié	Laetitia Deyris et Céline Meredieu	<a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J</a> <a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B</a>	Frédéric BERNIER	
10/08/2020	0.3	Non publié	Laetitia Deyris et Céline Meredieu	<a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J</a> <a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B</a>		
17/03/2021	0.4	Non publié	Laetitia Deyris et Céline Meredieu	<a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J</a> <a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B</a>		
15/04/2021	0.5	Non publié	Laetitia Deyris et Céline Meredieu	<a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J</a> <a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B</a>		
27/04/2021	0.6	Non publié	Laetitia Deyris et Céline Meredieu	<a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J</a> <a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B</a>		
29/06/2021	1.0	Non publié	Laetitia Deyris et Céline Meredieu	<a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=18454J</a> <a href="https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B">https://intranet.annuaire.inra.fr/afficherPersonne.action?matricule=115484B</a>		

## Présentation générale des données

- Données générées par la structure

UEFP génère des données, le plus souvent en milieu naturel et le plus souvent sur le long terme en distinguant :

- données « issues d'expérimentations » acquises dans le cadre de dispositifs où certains facteurs sont contrôlés ou manipulés (dans des essais de type agronomique ; caractérisation de collections de type ressource génétique, phénotypage).
- données « issues d'observations » acquises sur des objets/écosystèmes non manipulés

Ces données peuvent être acquises avec ou sans instruments, avec ou sans intervention humaine. Ce sont souvent des données répétées (série temporelles).

- Données de terrain capturées manuellement ou automatiquement; généralement uniques et irremplaçables
  - Données d'observation en forêt
    - relevés sur des arbres en milieu naturel ou dans des expérimentations in situ (croisements contrôlés, plantations comparatives, parcs à clones, ...),
    - mesures/observations de différents organes (feuilles, rameaux, bourgeons, racines, écorces,...)
    - prélèvements (arbre entier, organes, tissus, sols...)
    - relevés pédoclimatiques
    - mesures biophysiques
    - mesures biogéochimiques

- images
  - Données d'observation sur échantillons (végétal, sol)
- Données expérimentales capturées grâce à des équipements de labos ; peuvent être très coûteuses
  - en serres
  - à la halle biomasse
- Métadonnées liées aux dispositifs expérimentaux pour le département ECODIV
- Données issues de la gestion d'échantillon (graines et pollens)
- Codes informatiques / Logiciels (Systèmes d'information, applications web, scripts, logiciels, ...)
- Photos

Les flux de donnée par processus :

- collecte donnée
- traitement et curation des données
- description des données
- stockage sur le poste de travail
- transfert des fichiers vers GDO pour archivage (application hébergée au datacenter Ile de France)
- mise en base de données dans certains cas
- partage à la demande vers d'autres Systèmes d'Information
- droits et propriété via DATA INRAE

- Expérimentation
- Observation
- Simulation, modélisation

Définitions :

- observation (données de terrain uniques, non reproductibles ; avec ou sans instrument, avec ou sans intervention humaine),
- expérimentale (obtenues à partir d'appareils de mesures ou en laboratoire, peuvent être reproductibles),
- simulation (générées par des modèles informatiques ou de simulation, reproductibles)

- Collection
- Dataset
- Image
- Model
- Physical object
- Software

Nature des données pour nos principaux projets.

La majorité de nos fichiers de données sont déposés sur GDO.				
Pour le projet XYLOSYLVE, les données sont acquises avec Excel, ADONIS et en automatique par les capteurs.				
Les données non acquises automatiquement sont stockés sous l'entrepôt XYLOFRONT				
<b>XYLOSYLVE</b>				
	Fréquences	Variables	Arbres	Volumétrie
Capteurs	30 secondes	252	-	360 Mo/jour
Solutés	mensuel	18	-	108 ko/an
Analyse de sol	quinquennal	59	-	0,88 Mo/5 ans
Analyse de végétation	annuel	8	19426	0,25 Mo/an
Dendrométrie	annuel	3	19426	0,3 Mo/an
Trackeur solaire	20 secondes	40	-	40 Mo/jour
Capteurs	50 msec, 10, 20 ou 60 secondes en fonction des variables	311	-	127 Mo/jour
Pour les arboretums, les données sont acquises avec ADONIS. Les arboretums peuvent être équipés de station météorologique.				
<b>REINFORCE et CLIMAQ</b>				
	Fréquences	Variables	Arbres	Volumétrie
Phénologie	ponctuelle	14	850	0.6Mo/an
Dendrométrie	2 ans	2	20000	3Mo/an
Observation Sanitaire	ponctuelle	5	20000	2 Mo/an
Météo	horaire	12	-	24 Kb/jour

Pour le GIS COOP, les données sont acquises avec ADONIS.				
<b>COOP</b>				
	Fréquences	Variables	Arbres	Volumétrie
Phénotypage	Tous les 3 ans	3 à 7 variables	23276	4Mo
Pour le GIS GPMF volet risque -stress hydrique, les données sont acquises automatiquement via des capteurs				
Toutes ces données sont stockées après curation dans le SI PineLine administré par l'UMR BIOGECO				
	Fréquences	Variables	Arbres	Volumétrie
Suivi micro dendromètre	horaire	1	170	8800 valeurs par arbre
Taux d'humidité du sol à différentes profondeurs	horaire	1		64 sondes (16 + 16 +32 )
Hauteur de la nappe phréatique				3 capteurs
Sation météo	horaire	5		3 stations
Pour le GIS GPMF volet génétique, les données sont acquises avec ADONIS.				
<b>GPMF volet génétique</b>				
	Fréquences	Variables	Arbres	Volumétrie
Phénotypage	à 1 an	1	3500	0,3Mo
	à 8 ans	8	14000	2,5Mo
	à 12 ans	2	14000	1,5Mo
Pour le GIS GPMF volet fertilisation, les données sont acquises avec ADONIS.				
<b>GPMF volet fertilisation</b>				
	Fréquences	Variables	Arbres	Volumétrie
Analyse de végétation	Variable	8	-	0,25Mo
Phénotypage	Variable	4	10271	0,3Mo
<b>ICOS Salles Bilos</b>				
	Fréquences	Variables	Arbres	Volumétrie
Analyse de végétation	annuel	8	-	0,25Mo/an
Phénotypage	annuel	3	-	0,3Mo/an
<b>ORPHEE</b>				
	Fréquences	Variables	Arbres	Volumétrie
Phénotypage	annuel	2 à 6	10 000	181 ko pour 10000 arbres/variable
Irrigation	Hebdomadaire	1	-	22 ko/an
<b>AGROCLIM</b>				
	Fréquences	Variables	Arbres	Volumétrie
météo site	Horaire	14	-	3 Kb/jour
<b>Réseau Génétique Moléculaire</b>				
	Fréquences	Variables	Arbres	Volumétrie
Phénotypage	Variables	2 à 4	1000 à 10 000	6 pour 1000 arbres à 60 ko pour 10000 arbres/variable

<b>Réseau GEN4X</b>				
	Fréquences	Variables	Arbres	Volumétrie
Phénotypage	variables	3 à 5	4500 à 10 000	27 pour 4500 arbres à 60 ko pour 10000 arbres/variable
<b>Cybèle</b>				
	Fréquences	Variables	Arbres	Volumétrie
Inventaires RG	variables	4	2208	1818 ko/sauvegarde journalière
stock	annuelle	159 lots de pollen 185 lots de graine	-	
<b>Tests divers (&lt;10)</b>				
	Fréquences	Variables	Arbres	Volumétrie
Phénotypage	variables	2 à 4	500 à 2500	3 pour 500 arbres à 15 ko pour 2500 arbres/variable

Multiple formats de données, dépendant de l'origine des données :

- Dataset : csv, xls, xlsx, xml
- Texte : pdf, html, doc, docx
- Image : png, jpeg, tiff
- Formats génériques : rar, zip, xml
- Shapefiles : shx, shp, map
- Bases de données : PostgreSQL
- Programmation : Java, R, Javascript, PHP, Ruby, html

- Biodiversity and Ecology
- Climate
- Computer science
- Forests and Forest Products
- Plant Breeding and Plant Products
- Plant Health and Pathology
- Soils and soil sciences

## Droits de propriété intellectuelle

L'UEFP est en accord avec les principes de l'Open Data et respecte la Charte des Infrastructures Scientifique Collective (ISC) et des Infrastructures de Recherche (IR) de l'INRAE. Les données sont la propriété de INRAE.

## Confidentialité

Un SI local "GIDE" permet de décrire les métadonnées des dispositifs expérimentaux. Il contient des données soumises au RGPD (données géographiques, nom et coordonnées des propriétaires et gestionnaires, conventions de partenariat)

GIDE est en accès restreint au sein de INRAE et les valeurs de ces métadonnées ne sont accessibles qu'aux responsables des dispositifs.

Ces données soumises au RGPD ne seront jamais distribuées.

## Partage des données

L'UEFP est soumise aux règles institutionnelles (formalisées par la [charte pour le libre accès aux données et aux publications](#) ).  
Seules les données acquises dans le cadre d'un partenariat privé peuvent nécessiter une licence particulière.

- Réutilisation pour de nouvelles études ou des études complémentaires
- Aggrégation avec de nouvelles données (méta-analyses)
- Modélisation ou simulation

Tous nos formats de fichiers peuvent être lus par des logiciels libres excepté les fichiers .map où la licence MapInfo est obligatoire.

Publication sur le portail [Data INRAE](#)

- Autre

Pour GIDE : agents INRAE authentifiés et partenaires authentifiés

Pour Data INRAE : open access

- Licence ouverte <https://www.etalab.gouv.fr/licence-ouverte-open-licence> (compatible CC-BY)

## Organisation et documentation des données

Le lien ci-dessous renvoie vers les deux schémas de flux de données :

- Schéma Conservation et production de ressource génétique
- Schéma Observation et Expérimentation
- Schéma Gestion financière

et le tableau récapitulatif contenant l'explication de tous les flux

[https://uefp.pierroton.inra.fr/flux\\_donnees/index\\_flux\\_donnees\\_UEFP.html](https://uefp.pierroton.inra.fr/flux_donnees/index_flux_donnees_UEFP.html)

Métadonnées	Origine, mode de production des métadonnées (ex : saisie manuelle, annotation automatique...)	Standard, Vocabulaires associés	Conditions ou fréquence de la mise à jour (si applicable) (ex : changement de l'accessibilité)
Ontologie des plantes ligneuses	saisie manuelle sur cahier de laboratoire ou ordinateur de terrain (Adonis par exemple)	Woody Plant Ontology <a href="https://forgemia.inra.fr/urgis/ontologies/tree/develop/Tree">https://forgemia.inra.fr/urgis/ontologies/tree/develop/Tree</a>	
EDDY COVARIANCE TURBULENT FLUX RAW DATA FORMAT	Acquisition automatique	ICOS Ecosystem Instructions	
METEOROLOGICAL DATA FORMAT	Acquisition automatique	ICOS Ecosystem Instructions	

Nous avons des protocoles et des modes opératoires décrivant l'ensemble des activités de recueil de données. Toutes ces informations sont également stockées dans GDO.

Tous nos fichiers de données sont déposés sur GDO qui gère le versionnement et l'organisation.

Pour le nommage des fichiers, nous préconisons les mêmes règles que celles fournies par DORANUM. <https://doranum.fr/stockage-archivage/comment-nommer-fichiers/>

Un contrôle qualité est effectué lors de la saisie sur le terrain grâce au logiciel ADONIS.

## Stockage et sécurité des données

Le lien ci-dessous renvoie vers les deux schémas de flux de données par processus :

- Schéma Conservation et production de ressource génétique

- Schéma Observation et Expérimentation

et le tableau récapitulatif contenant l'explication de tous les flux

[https://uefp.pierroton.inra.fr/flux\\_donnees/index\\_flux\\_donnees\\_UEFP.html](https://uefp.pierroton.inra.fr/flux_donnees/index_flux_donnees_UEFP.html)

La volumétrie actuelle et provisionnelle est présentée dans un tableau de ce PGD (Voir plus haut rubrique "Présentation générale des données >> Nature des données")

L'infrastructure informatique utilisée pour le système d'information est hébergée en datacenter institutionnel. Elle bénéficie ainsi de la politique de sécurité nationale.

Les offres de services souscrites concernent aussi bien la partie matérielle (offre serveurs virtuels nus avec option sauvegarde des espaces disques de la VM), que les parties data tant non structurées (stockage de données individuelles ou collectives = "*stocker des données de recherche ou d'appui dans des conditions de sécurité adaptées aux enjeux*") ou structurées (base de données sur la base de logiciels libres).

La sécurisation en datacenter se fait aussi bien sur le volet de l'intégrité des données hébergées (plan de sauvegarde, maintenance matérielle et logicielle) que sur le volet du contrôle d'accès aux données et services (pare feu).

L'UEFP, toujours dans le cadre de la sécurisation des données, a été particulièrement attentive à la sauvegarde des postes utilisateurs qui est effectuée quotidiennement (avec rétention sur un mois) sur un dispositif, propre à l'unité, géré par l'Informaticien de Proximité du site de Pierroton (UMR BIOGECO). Une vigilance s'impose du fait que ce matériel est vieillissant et qu'il atteint ses limites en termes de capacité. Plutôt qu'un renouvellement, il est prévu de prendre en considération l'offre institutionnelle de sauvegarde de postes de travail utilisateurs qui devrait prochainement être proposée (en espérant une prompte mise en place, cette offre étant encore à l'état de chantier mais bien avancée selon les retours de la DSI obtenus dernièrement).

Nous pouvons considérer que l'organisation du stockage est une activité contrôlée. Concernant le nommage des fichiers, l'UEFP a diffusé les bonnes pratiques préconisées par DoRANum. L'utilisation du système de classement GDO associant systématiquement des métadonnées permet de palier certaines erreurs et le versionnement des fichiers.

Toutes nos données sont déposées sur notre GDO ainsi que les métadonnées. La confidentialité dépend de la nature des données (ex: données géographiques soumises au RGPD).

Lors de la diffusion/publication des données nous utilisons le dataverse de l'unité.

Les partenaires de l'UEFP signent une charte de citation de l'unité.

### Sécurité - Intégrité - Traçabilité :

Pour l'acquisition des données, nous utilisons la plupart du temps le logiciel ADONIS.

Tous nos ordinateurs sont sauvegardés quotidiennement et sont protégés contre les virus et les intrusions.

Des restrictions sont mises en place sur les droits d'accès à nos BDD ou nos SI.

Toute personne faisant appel à l'UE Forêt Pierroton valide une charte de citation et de remerciement.

Nous utilisons le DOI fourni par le DATA INRAE pour tracer l'utilisation des données par des tiers.

## Archivage et conservation des données

Toutes les données brutes sont à conserver à long terme ainsi que toutes les informations utiles à leur exploitation.

A chaque démarrage de projets, l'UEFP s'engage à établir une réflexion particulière avec les porteurs du projet pour co-construire le PGD spécifique à ce projet prenant en compte les flux de données de l'UEFP vers la(les) autres entités impliquées dans le projet, leur stockage dans le cadre du projet et leur préservation.

Toutes les données sont déposés sur GDO.

Certaines des données pour leur mise à disposition sont également déposées sur le DATA verse INRAE.



La plupart des données générées par l'UEFP sont des données brutes non reproductibles et sont donc conservées sur le long terme sans date limite. Nous avons besoin de conserver nos données sur des dizaines d'années, c'est-à-dire bien au-delà de la révolution des écosystèmes forestiers sur lesquels l'UEFP travaille.

Les coûts associés à la conservation à long terme des données seront financés sur les fonds propres de l'Unité en tant qu'ISC.