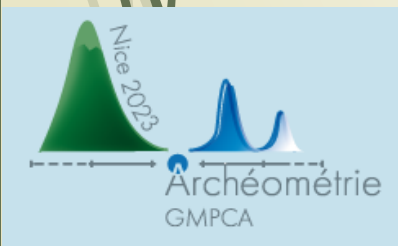
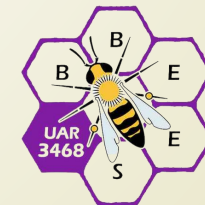


Gestion des jeux de données et Science Ouverte : l'implication des unités d'appui à la recherche



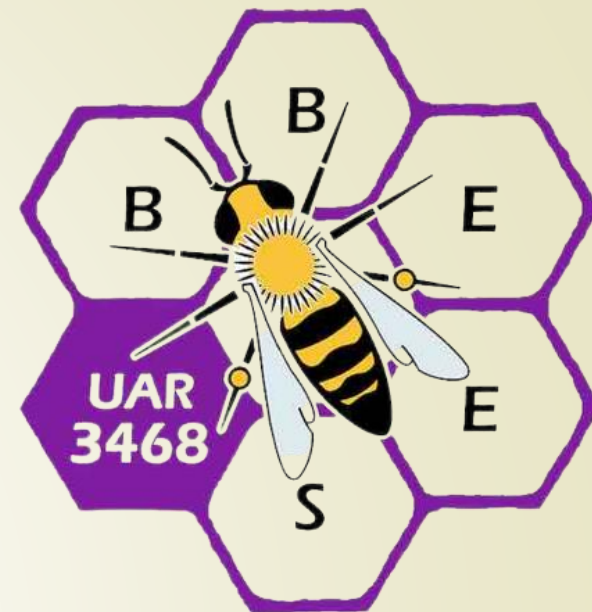
Isabelle Baly, Amandine Hénon et Cécile Callou



L'UAR BBEES : un point d'entrée pour les données environnementales et de biodiversité

Double tutelle : Muséum national d'Histoire naturelle et CNRS-INEE
Notre périmètre pour l'archéologie : la bioarchéologie (implication notamment dans le GDR BioArchéoDat)

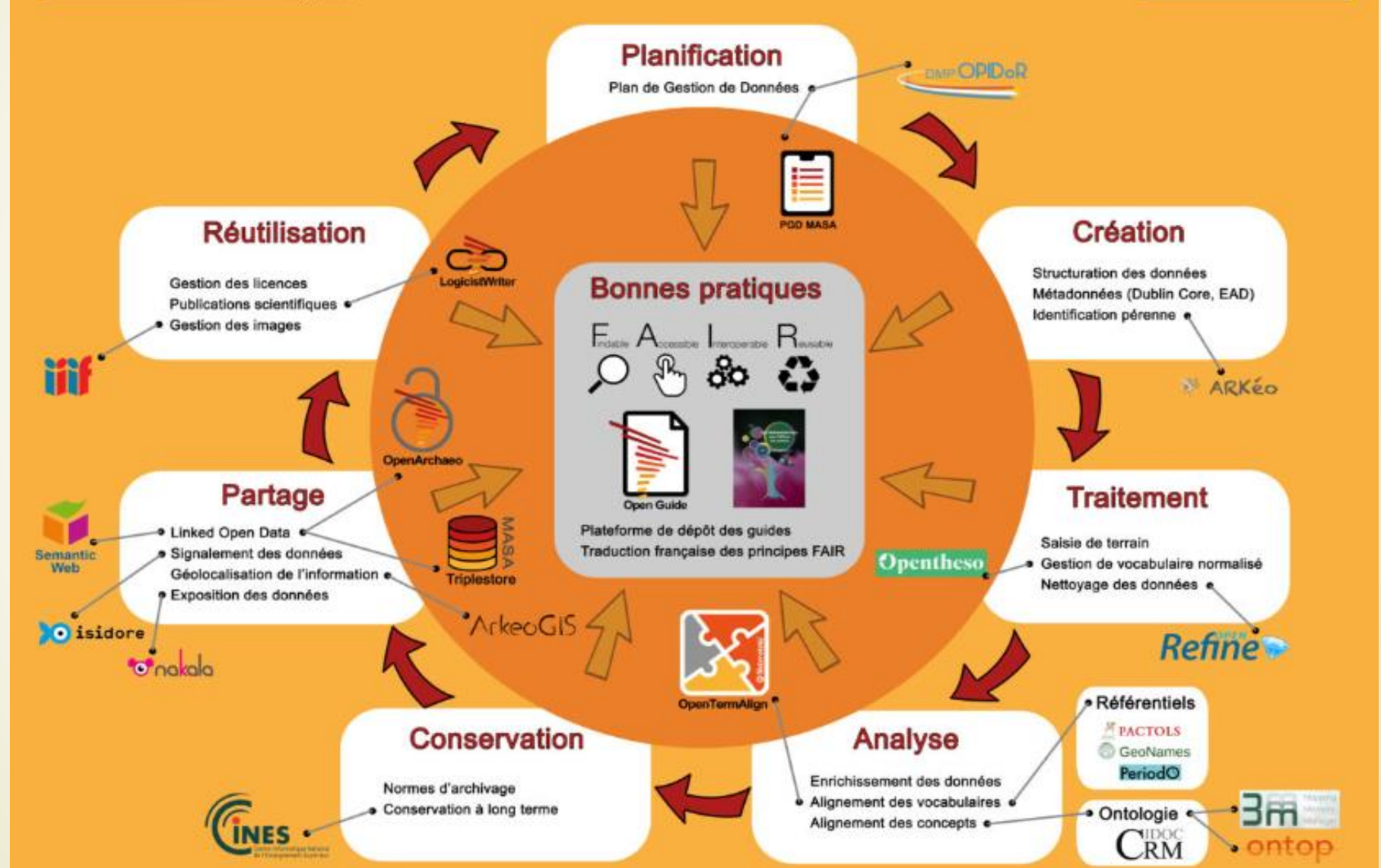
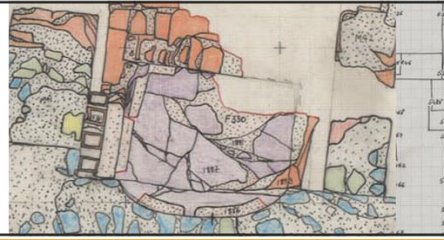
- **Assistance** : conseil à la maîtrise d'œuvre et d'ouvrage
- **Accompagnement** : formation et enseignement
- **Pérennisation** et **diffusion** : conception de sites web ou mise à disposition de l'entrepôt data.InDoRES



Point d'entrée privilégié par les tutelles pour les données



Petite équipe de seulement 8 personnes



Planification

Plan de Gestion de Données



Création

Structuration des données
Métadonnées (Dublin Core, EAD)
Identification pérenne



Traitement

Saisie de terrain
Gestion de vocabulaire normalisé
Nettoyage des données



Analyse

Enrichissement des données
Alignement des vocabulaires
Alignement des concepts



Bonnes pratiques

F A I R
Findable Accessible Interoperable Reusable



Plateforme de dépôt des guides
Traduction française des principes FAIR

Réutilisation

Gestion des licences
Publications scientifiques
Gestion des images



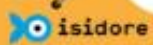
Conservation

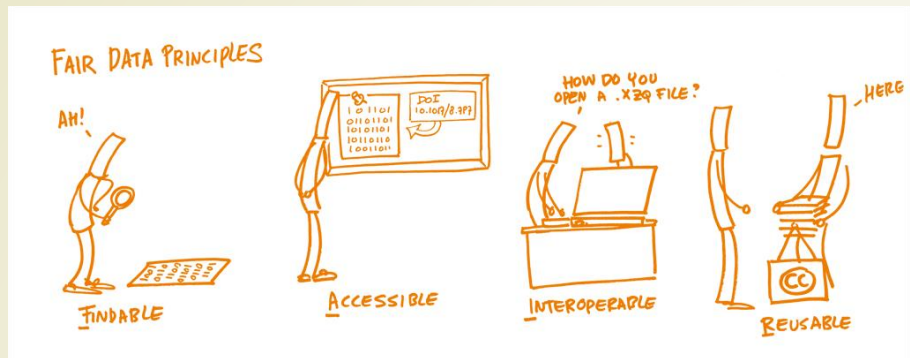
Normes d'archivage
Conservation à long terme



Partage

Linked Open Data
Signalement des données
Géolocalisation de l'information
Exposition des données

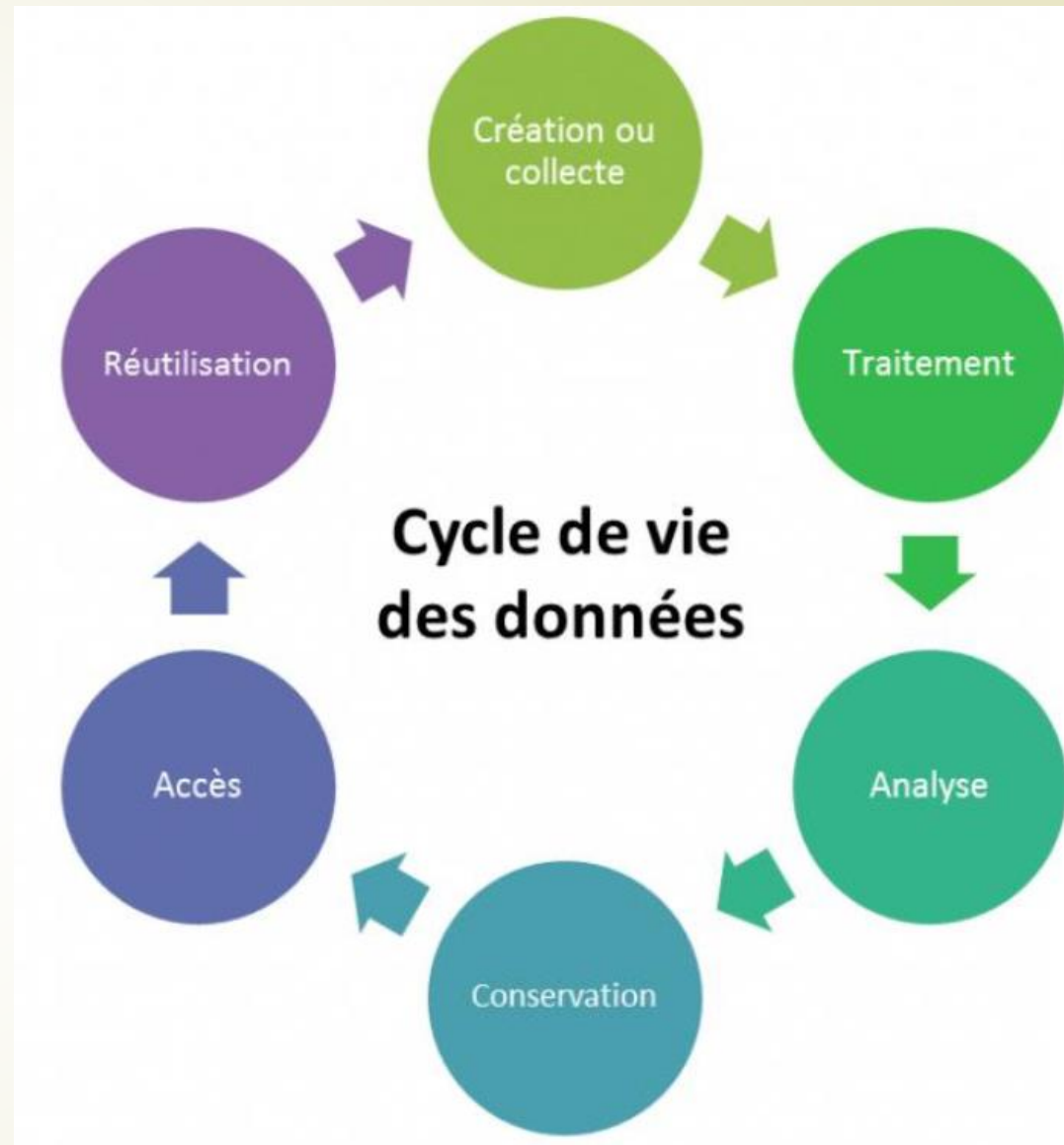




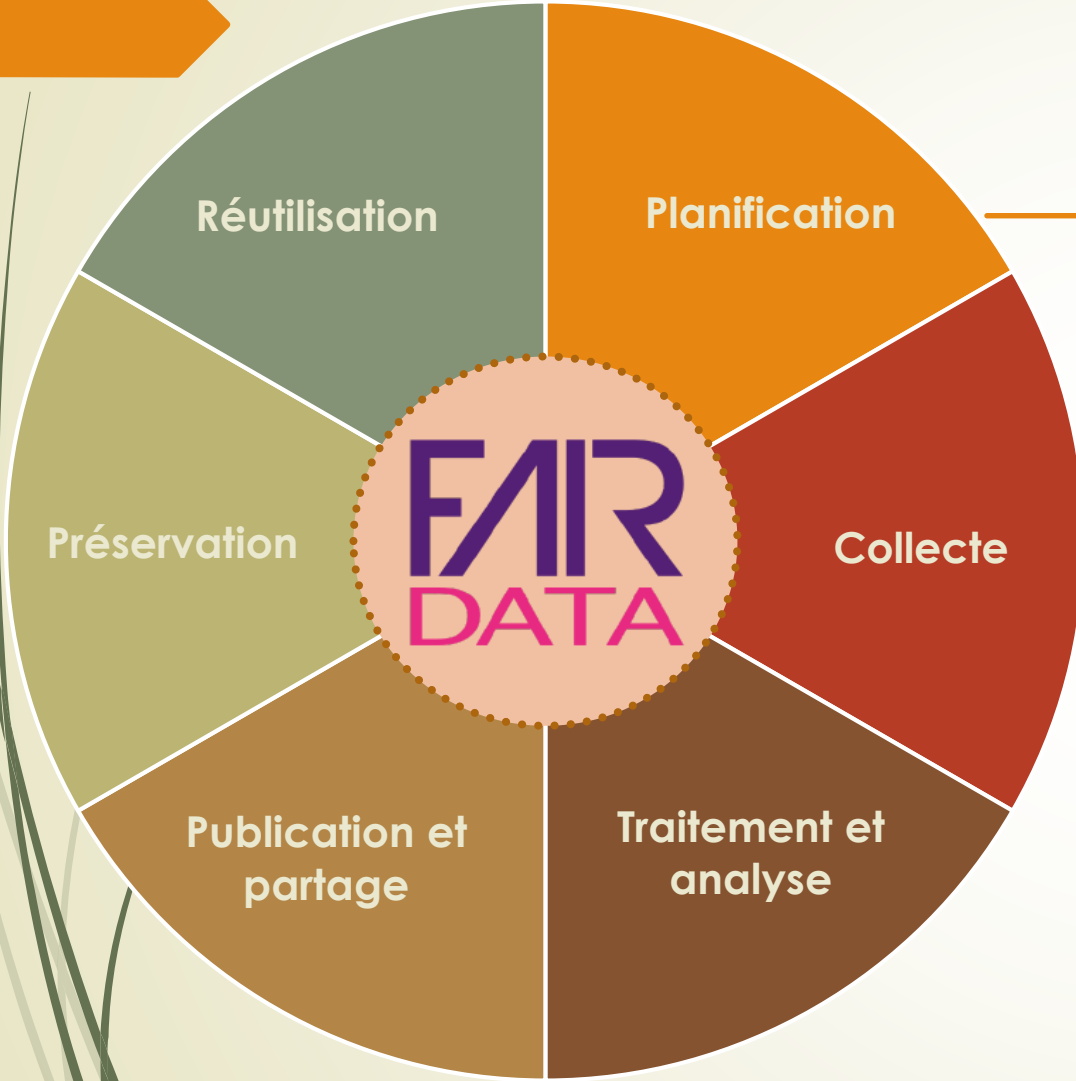
L'INEE et le Muséum national d'Histoire naturelle proposent également des **outils** et **services**, notamment par le biais des UAR (dont BBEES) pour gérer les données de la Recherche selon les **principes FAIR** et la **Science Ouverte**.



Source : DipSO

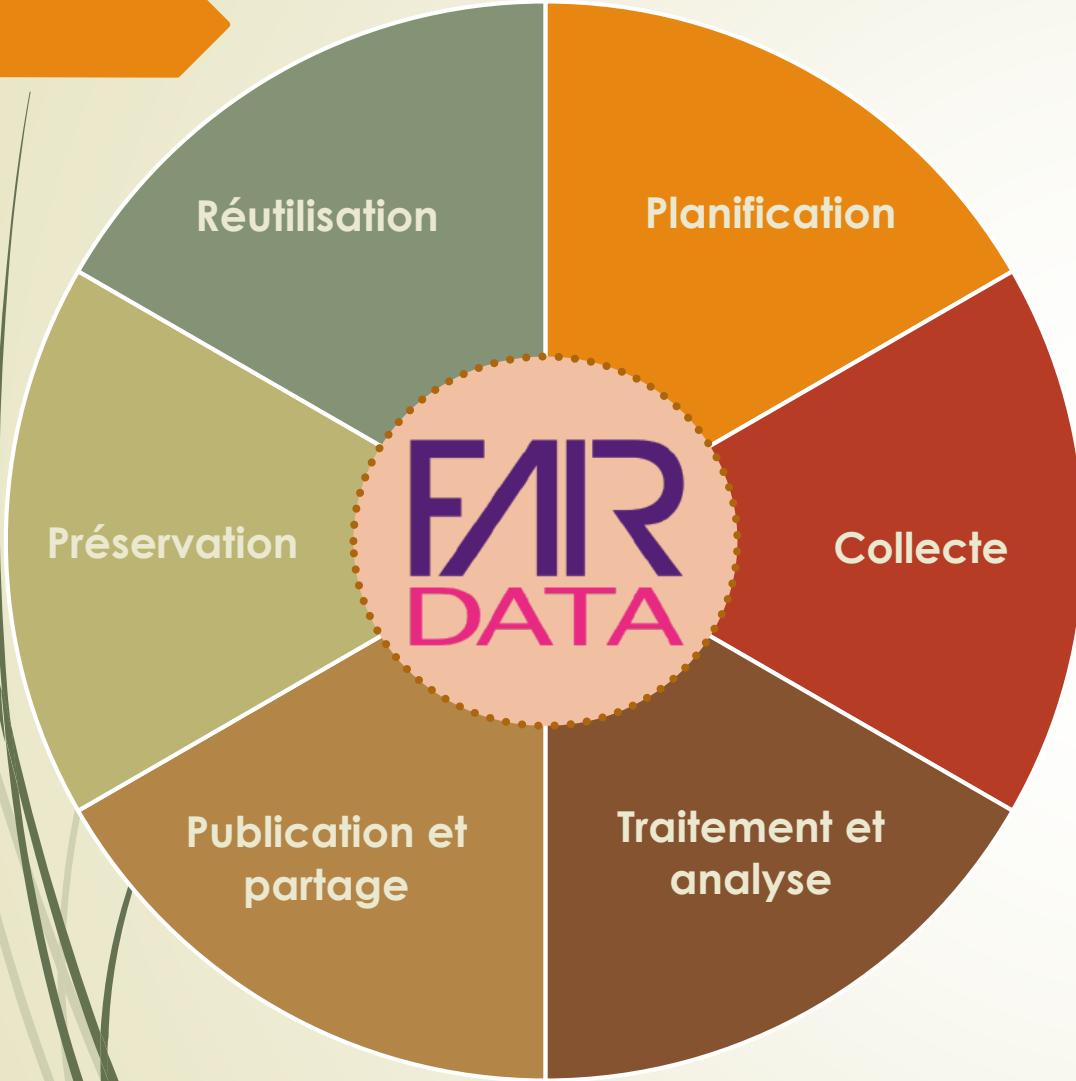


Source : Institut Pasteur



- Aide à la rédaction du **Plan de Gestion de Données**
- Comprendre le **paysage** politique et technologique de la donnée
- Evaluer les **besoins**/rédaction d'un cahier des charges

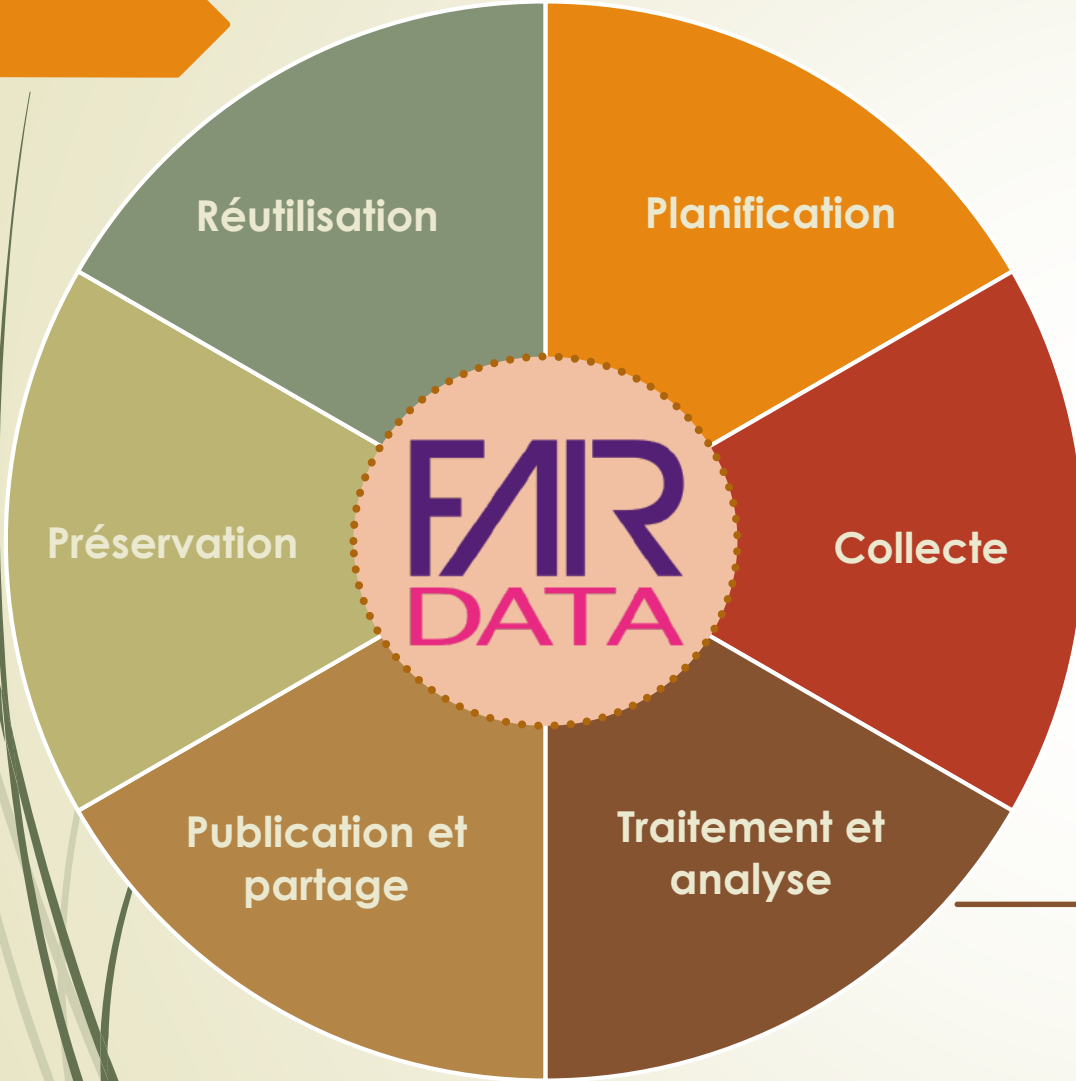




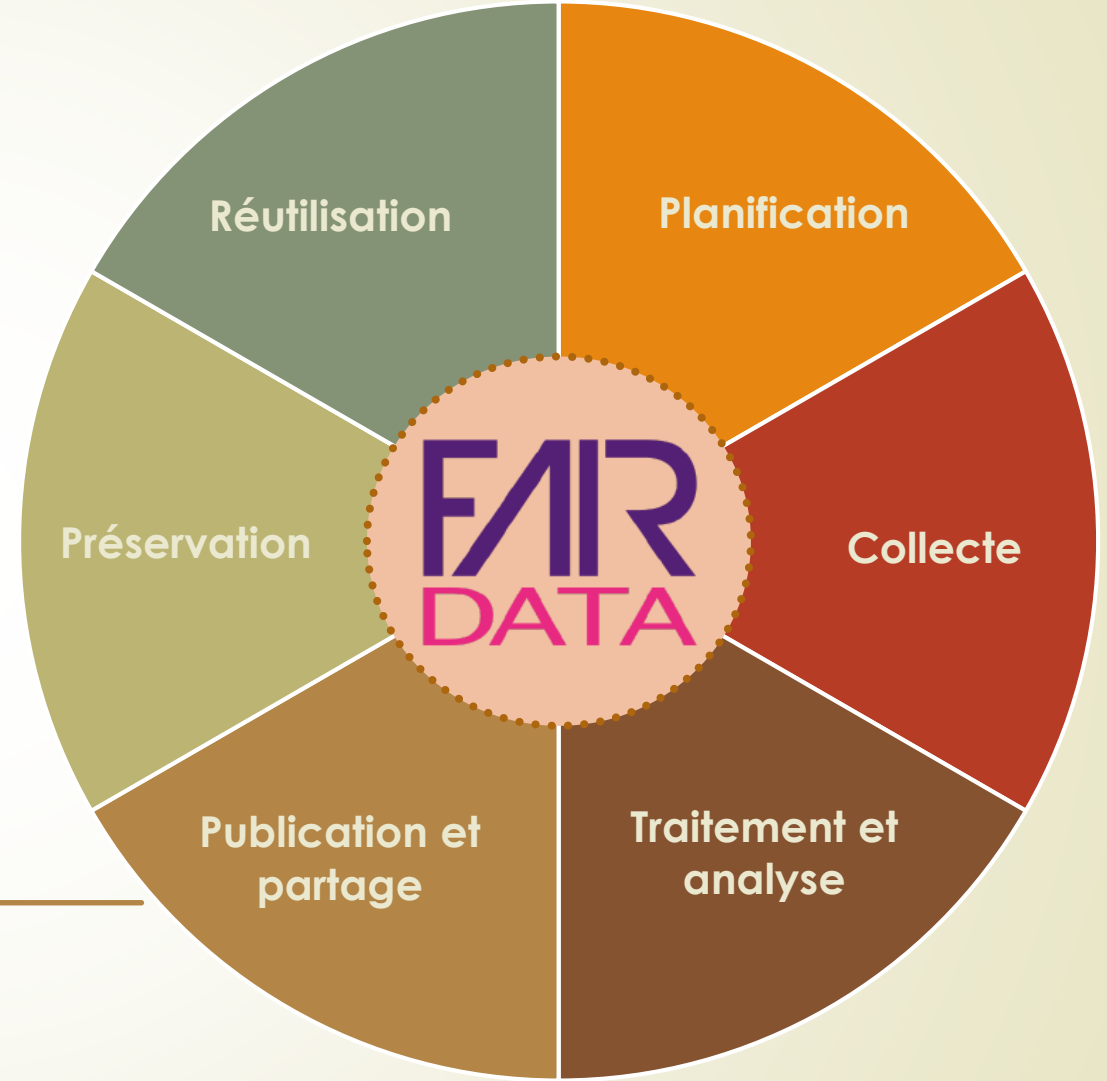
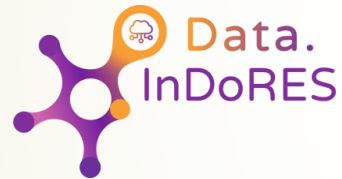
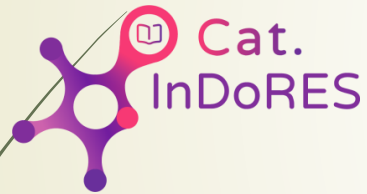
- Conseils sur la **structuration des données**, formats, standards de métadonnées, référentiels et thesaurus
- Conseils sur les **outils**
- Conseils sur la prise en compte des **données personnelles** et/ou sensibles/RGPD
- Tutoriels de **nettoyage** des données avec OpenRefine
- **Stockage**



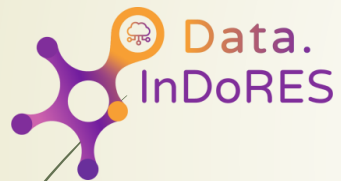
OpenRefine



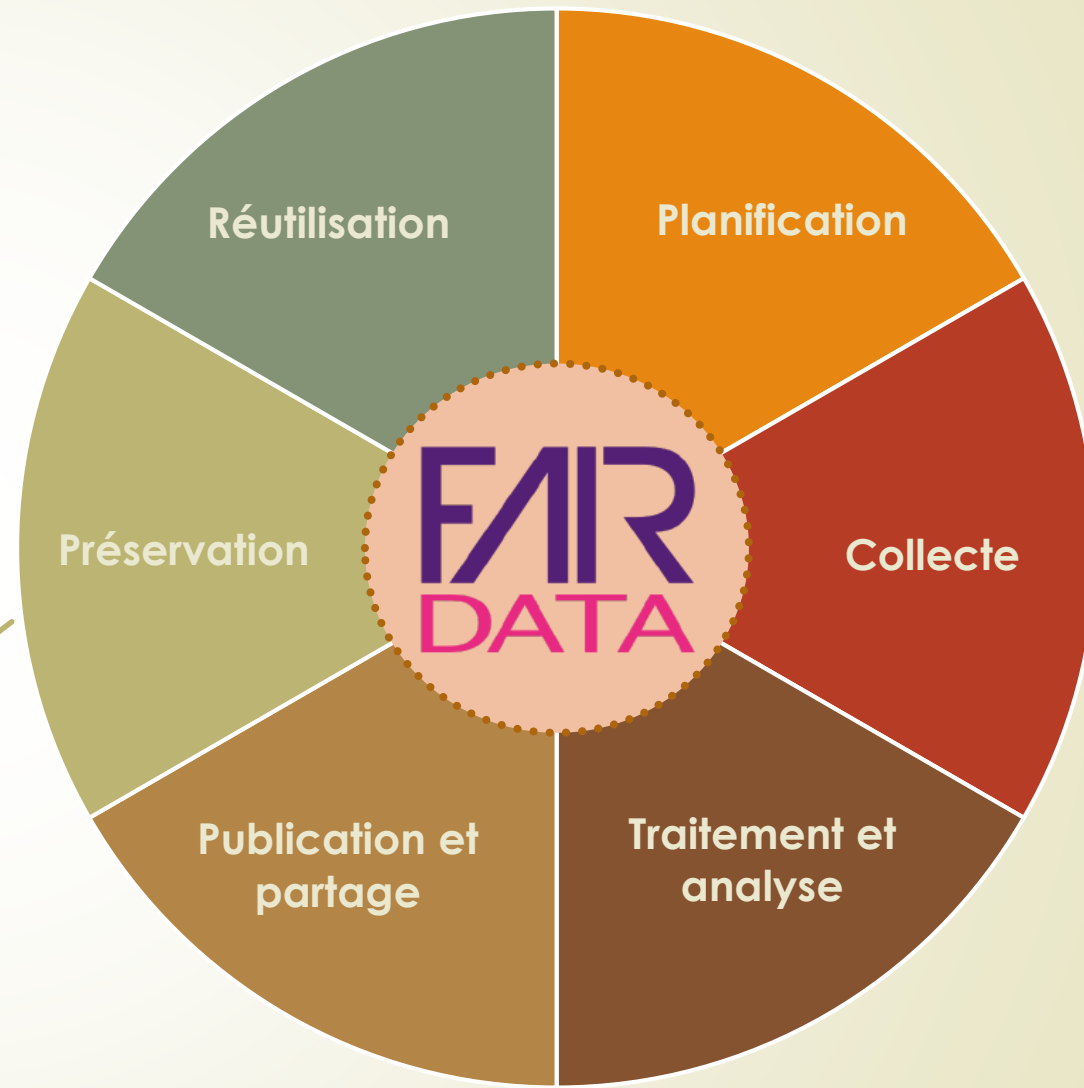
- Accès aux **grilles de calcul** du CC-IN2P3
- **Conseils** et collaborations avec les plateformes analytiques ou avec l'UAR 2AD



- **Cat.InDoRES** (métadonnées riches)
- **Sites** web
- Diffusion et publication des **jeux de données** liés à des articles
- **Bases de données** consultables en ligne

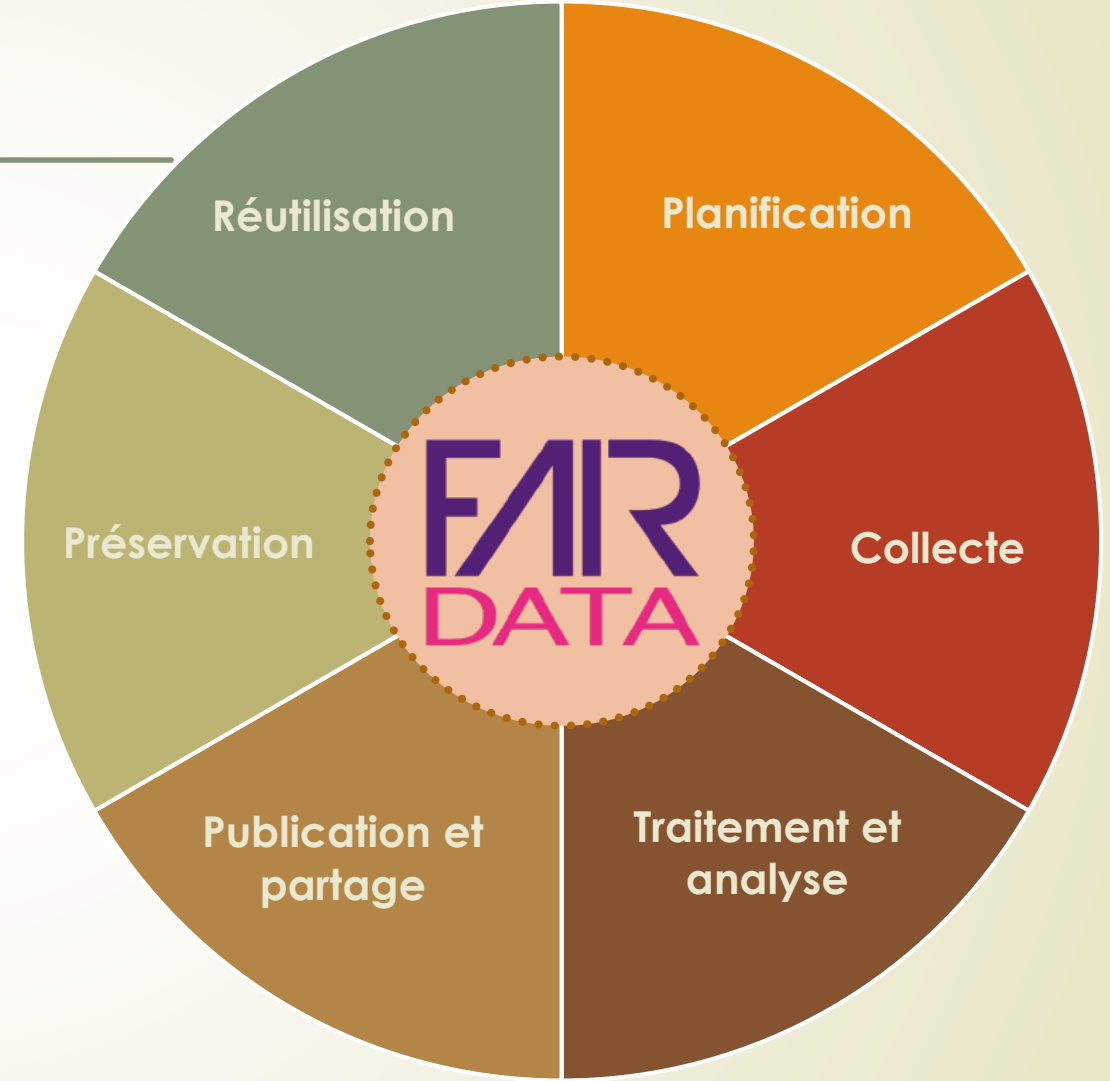


- Data.InDoRES
- Conseils sur les **entrepôts** mais aussi sur l'**archivage**
- DOI





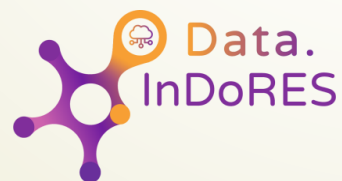
- Choix des **licences**
- Optimisation de réutilisation grâce aux **standards de métadonnées**
- Meilleure citation grâce aux DOI
- Conseils pour la rédaction de **data papers**



Une plate-forme d'aide et d'accompagnement



- Mise à disposition des **ressources documentaires ou numériques** autour des principes FAIR
- Rappel des documents fédérant la Science Ouverte pour les données de la Recherche
- **Accompagnement et conseil** pour la diffusion des données (formats, licence, entrepôts...)
- Accès à des **outils de publication**



<https://www.indores.fr/>



Conformément au Plan national de la Science Ouverte, les données produites dans les organismes de recherche publics doivent être facilement identifiées pour être réutilisées.

Sous la tutelle de l'Institut Ecologie et Environnement du CNRS et du Muséum national d'Histoire naturelle, le portail InDoRES propose :



- de découvrir les bases et jeux de données grâce à une description des projets qui les produisent (catalogue de métadonnées standardisé et interopérable répondant aux directives européennes). Lorsqu'ils existent, des liens sont proposés vers les données à télécharger ou encore vers des publications associées
- d'accéder à l'entrepôt sécurisé data.InDoRES pour publier des données
- de rassembler et faire le lien vers un ensemble de ressources sur les pratiques qui facilitent le partage et la réutilisation des données, conformément aux préconisations des principes FAIR (Faciles à trouver, Accessibles, Interopérables, Réutilisables)

Publier mes données

Explorer la plate-forme InDoRES



Consulter
notre catalogue de
métadonnées



Chercher
Un jeu ou une base de
données



Créer
une nouvelle fiche de
jeu de données



Extraire / publier
un jeu
de données

cat.InDoRES : un catalogue de métadonnées thématique pour la diffusion des bases et jeux de données

- "Annuaire" pour les données en environnement et biodiversité avec **une fiche** par base ou jeu de données
- Catalogue utilisant des **standards** internationaux pour une bonne **interopérabilité** entre les systèmes
- Possibilité de lier des ressources : protocole d'acquisition, lien vers les données, vers des articles (HAL) : **centralisation des informations.**

Plus de 9000 fiches

- Moissonne les catalogues locaux
- Est moissonné par les initiatives nationales et internationales



Q Retour à la recherche < Précédent Suivant >

Base de données assolement sur la ZA Plaine et Val de Sèvre - 2018

Base de données décrivant l'assolement (forme des parcelles agricoles et type de cultures) sur la zone atelier Plaine et Val de Sèvre, au sud de Niort, sur 450 km², depuis 1994. L'assolement est mis à jour deux fois par an par la Zone Atelier, en Avril et en Juin.

La base contient :

- les limites des parcelles, chaque année,
- les limites des microparcelles (construction par intersection de toutes les couches, sans date),
- les OCS des parcelles, chaque année,
- les routes (reclassées et linéaires corrigés par agripop),
- les haies sur la zone, pour 3 années : 2006, 2011, 2014,
- les bandes enherbées sur la zone,
- les exploitants et associations agricoles des parcelles, chaque année,
- les contrats passés entre eux et agripop,
- les secteurs (limites et noms) utilisés par agripop pour l'enquête,
- les noms et limites des communes de la zone.

Téléchargements et liens

Autres ressources : <http://www.za.plainevalsevre.cnrs.fr/index.php/la-zone-atelier/observatoire-de-la-zone-atelier/donnees-dassolement/> Ouvrir le lien

Flux du parcellaire de l'année en cours
Cette donnée est publiée dans le service de téléchargement (WFS) disponible à l'adresse <http://za-geminat.cnrs.fr:8080/geoserver/geminat/wfs>.
Unable_to_connect_to_service Visualiser

Ressources associées

- Bandes enherbées sur la ZA Plaine et Val de Sèvre - 2016 (Fiche enfant)
Bandes enherbées issues de la base de données décrivant l'assolement (forme des parcelles agricoles et type de cultures) sur la zone atelier Plaine et Val de Sèvre, au sud de Niort, sur 450 km², depuis... .. Fiche enfant
- Base de données enquête sur la ZA Plaine et Val de Sèvre - 2018 (Fiche enfant)
Base de données décrivant l'ensemble des Fiche enfant

Télécharger Mode affichage

Aperçu

Part de maïs dans chaque carreau de 1km² enquêté.

Étendue spatiale

- Plaine et Val de Sèvre

Étendue temporelle

Date de création
1994-06-16

Date de publication
2017-02-03

Date de révision
2018-07-17

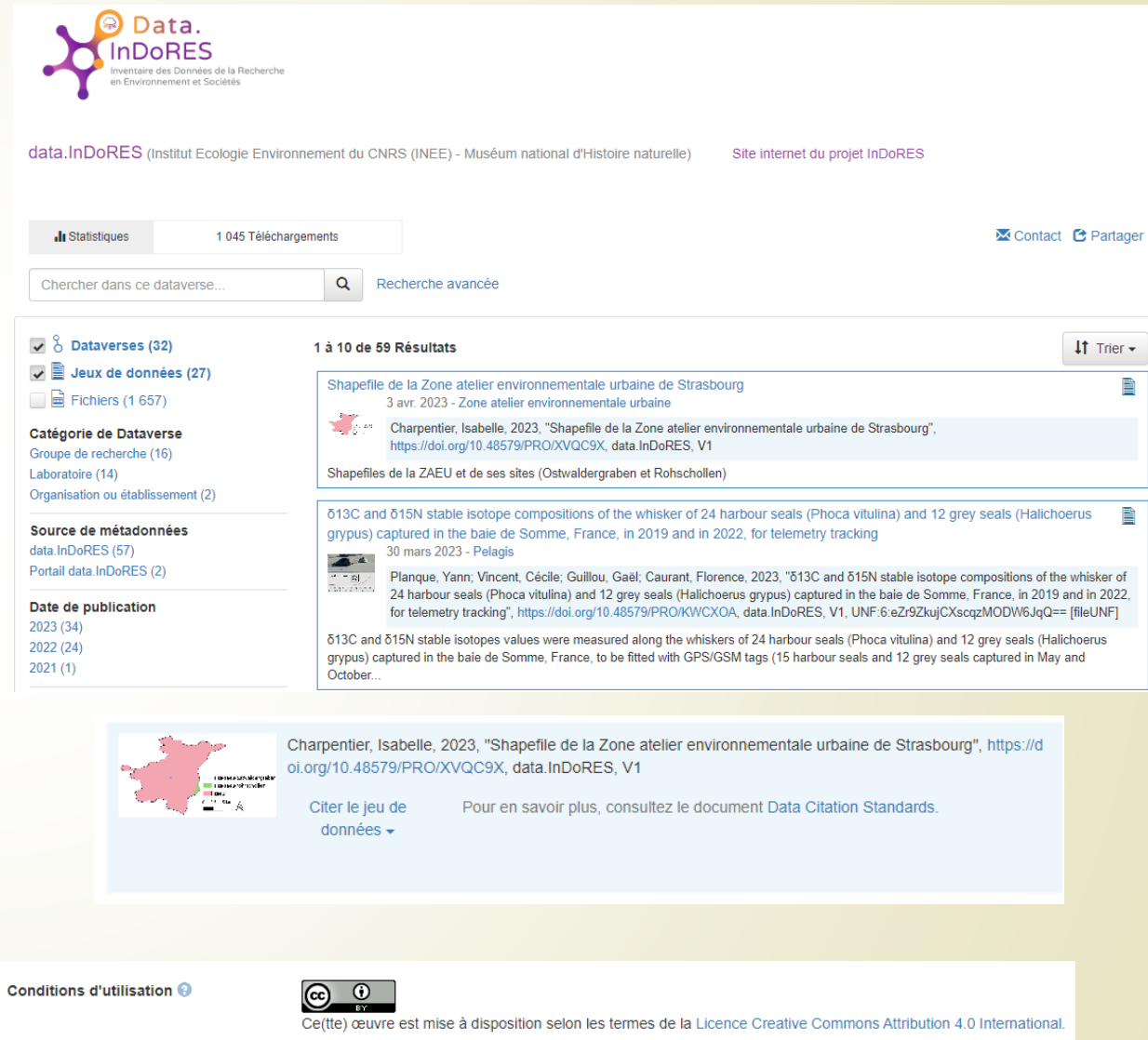
Période
Thu Jun 16 1994 02:00:00 GMT+0200 → Tue Aug 30 2016 02:00:00 GMT+0200

Fourni par

Mis à jour :
il y a 4 mois

data.InDoRES : un entrepôt pour la publication des jeux de données

- Entrepôt **Dataverse** : consensus dans communauté recherche (INRAE, IRD, Recherche Data Gov)
- Données stockées de manière **sécurisée et pérenne** au Centre de Calcul de l'IN2P3 (infrastructure nationale de recherche).
- **Structuration** par unité ou par projet (exemple ANR) pour une visibilité accrue.
- Possibilité **embargo** jusqu'à publication article
- Interface français/anglais
- **DOI** et format de citation pour réutilisation
- **Licence** CC-BY 4.0



The screenshot displays the data.InDoRES website interface. At the top, the logo for 'Data InDoRES' is visible, along with the text 'Inventaire des Données de la Recherche en Environnement et Sociétés'. Below the logo, the website's name and affiliation are listed: 'data.InDoRES (Institut Ecologie Environnement du CNRS (INEE) - Muséum national d'Histoire naturelle)'. A search bar is present with the text 'Chercher dans ce dataverse...' and a search button. To the right of the search bar, there are links for 'Contact' and 'Partager'. Below the search bar, there are filters for 'Statistiques' (1 045 Téléchargements) and 'Recherche avancée'. The main content area shows search results for 'Shapefile de la Zone atelier environnementale urbaine de Strasbourg'. The results are displayed in a list format, with the first result being 'Shapefile de la Zone atelier environnementale urbaine de Strasbourg' by Charpentier, Isabelle, 2023. The result includes a DOI link: 'https://doi.org/10.48579/PRO/XVQC9X, data.InDoRES, V1'. Below the search results, there is a section for 'Conditions d'utilisation' and a Creative Commons Attribution 4.0 International license logo. The text below the logo reads: 'Ce(tte) œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution 4.0 International.'



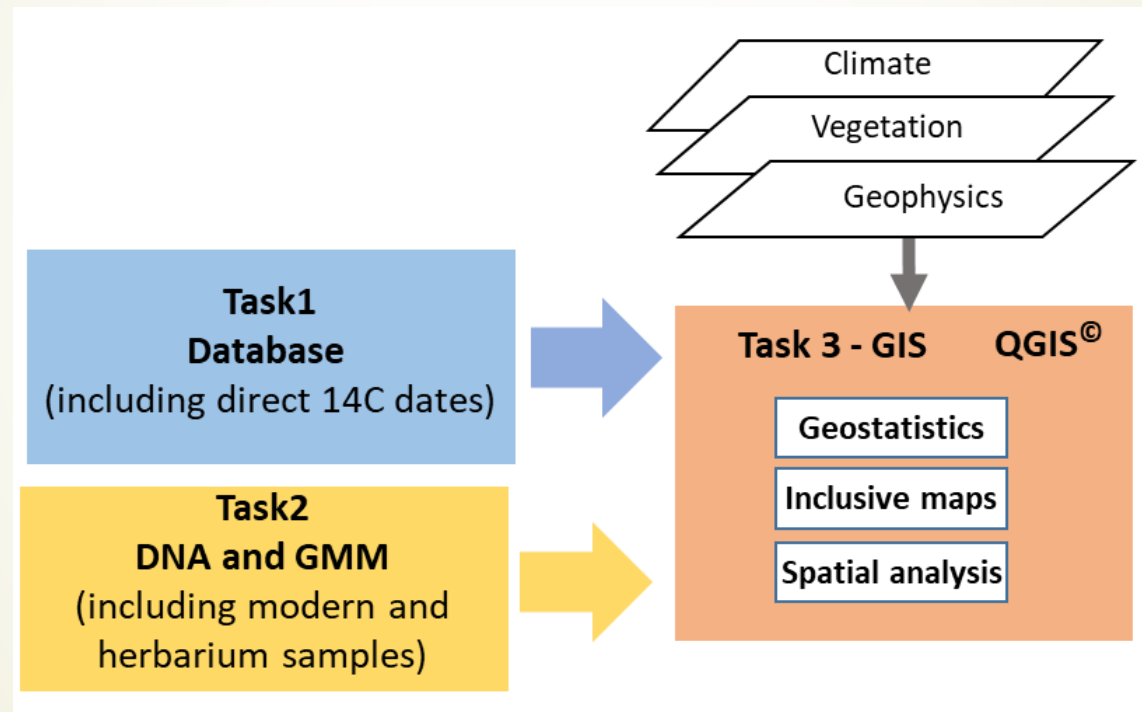
POPPY

Illustration de la mise en œuvre des bonnes pratiques dans la gestion des données d'un projet ANR dans le domaine de la biodiversité

ANR POPPY

(dir. A. SALAVERT MC MNHN, UMR 7209)

- Explore l'origine et la diffusion ancienne du pavot à opium en Europe et dans l'ouest de l'Asie entre 6000 et 50 avant notre ère.



BBEES a été impliqué dès le dépôt du projet ANR ce qui a permis d'anticiper le budget pour recruter un CDD IE pour les tâches 1 et 3

CDD POPPY

(Linda ANGULO LOPEZ)

- Conception de la base de données et d'un système d'information
- Prise en compte des principes FAIR
- Publication de jeux de données
- Dans un second temps, développement de services de visualisation

The screenshot shows the popArch website interface. At the top, there is a navigation bar with links for 'popArch', 'Team', 'Collaborate', 'Explore', and 'Visitor', along with a 'Connection' button. Below the navigation bar, the main content area is titled 'Website Guidelines'. A message states: 'Please login to take advantage of the privileges you have been granted.' Below this, a note mentions that registered users with view privileges can read data from the popArch database using the Visitor connection tab. The 'Featured Data' section displays a map titled 'Simple Bee Farming Demo' with a blue highlighted area. Below the map is a button labeled 'Explore the Interactive Map'. To the right of the map, the 'About the popArch Database' section provides information about the Opium Poppy remains from archaeological excavations, their historical significance, and the database's location at CCINZP3. It also lists the crop as 'Opium Poppy', the species as 'Papaver somniferum', and provides keywords such as 'Database', 'Opium Poppy', 'Botany', 'Archeology', 'Plant Domestication', and 'Species Dispersal'. A 'Crop History' button is also present. The 'References' section indicates there are 250 references. At the bottom of the page, there are logos for CCINZP3, CNRS, and ANR, along with a footer containing contact information and legal notices.

Aide à la rédaction du Plan de Gestion de Données

- Aide à l'utilisation de l'outil  (outil collaboratif développé par l'INIST)
 - <https://dmp.opidor.fr/>
- Conseils pour rédiger les parties techniques sur le stockage notamment

3. Storage and backup during the research process

[tout développer](#) | [tout réduire](#)

Arch DB

GMM Data


DNA Data

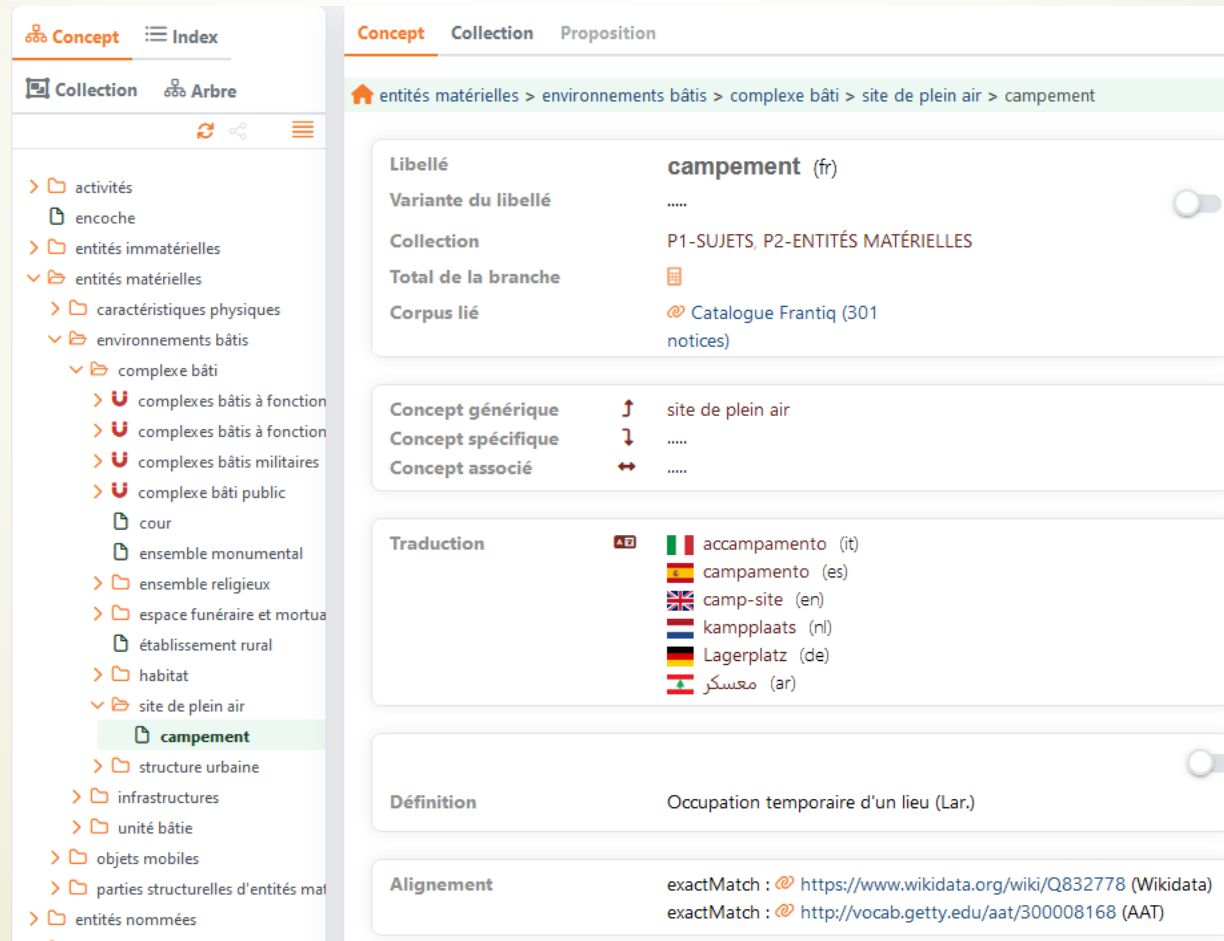
Les réponses de cette section sont communes à tous les produits de recherche

3a. How will data and metadata be stored and backed up during the research? >















3b. How will data security and protection of sensitive data be taken care during the research >

Conseils sur l'utilisation de référentiels et thesaurus

- Référentiel  **GeoNames** pour les noms de pays <https://www.geonames.org/>
- Thesaurus PACTOLS pour les types de sites et de structures <https://pactols.frantiq.fr/>

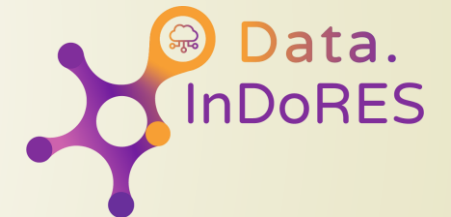
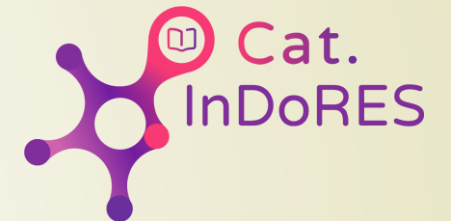


The screenshot displays the PACTOLS interface for the concept 'campement'. The left sidebar shows a hierarchical tree of concepts, with 'campement' selected under 'site de plein air'. The main content area is divided into several sections:

- Libellé**: **campement** (fr)
- Variante du libellé**:
- Collection**: P1-SUJETS, P2-ENTITÉS MATÉRIELLES
- Total de la branche**: 
- Corpus lié**:  Catalogue Frantiq (301 notices)
- Concept générique**:  site de plein air
- Concept spécifique**: 
- Concept associé**: 
- Traduction**:   accampamento (it),  campamento (es),  camp-site (en),  kampplaats (nl),  Lagerplatz (de),  معسكر (ar)
- Définition**: Occupation temporaire d'un lieu (Lar.)
- Alignement**: exactMatch :  <https://www.wikidata.org/wiki/Q832778> (Wikidata), exactMatch :  <http://vocab.getty.edu/aat/300008168> (AAT)

Ouverture des données

- Rédaction de la fiche de métadonnées pour intégrer afin de faire connaître l'existence de la base de données
- Accompagnement pour déposer des jeux de données dans l'entrepôt



P. somniferum subsp. *somniferum*



P. somniferum subsp. *setigerum*

En résumé...

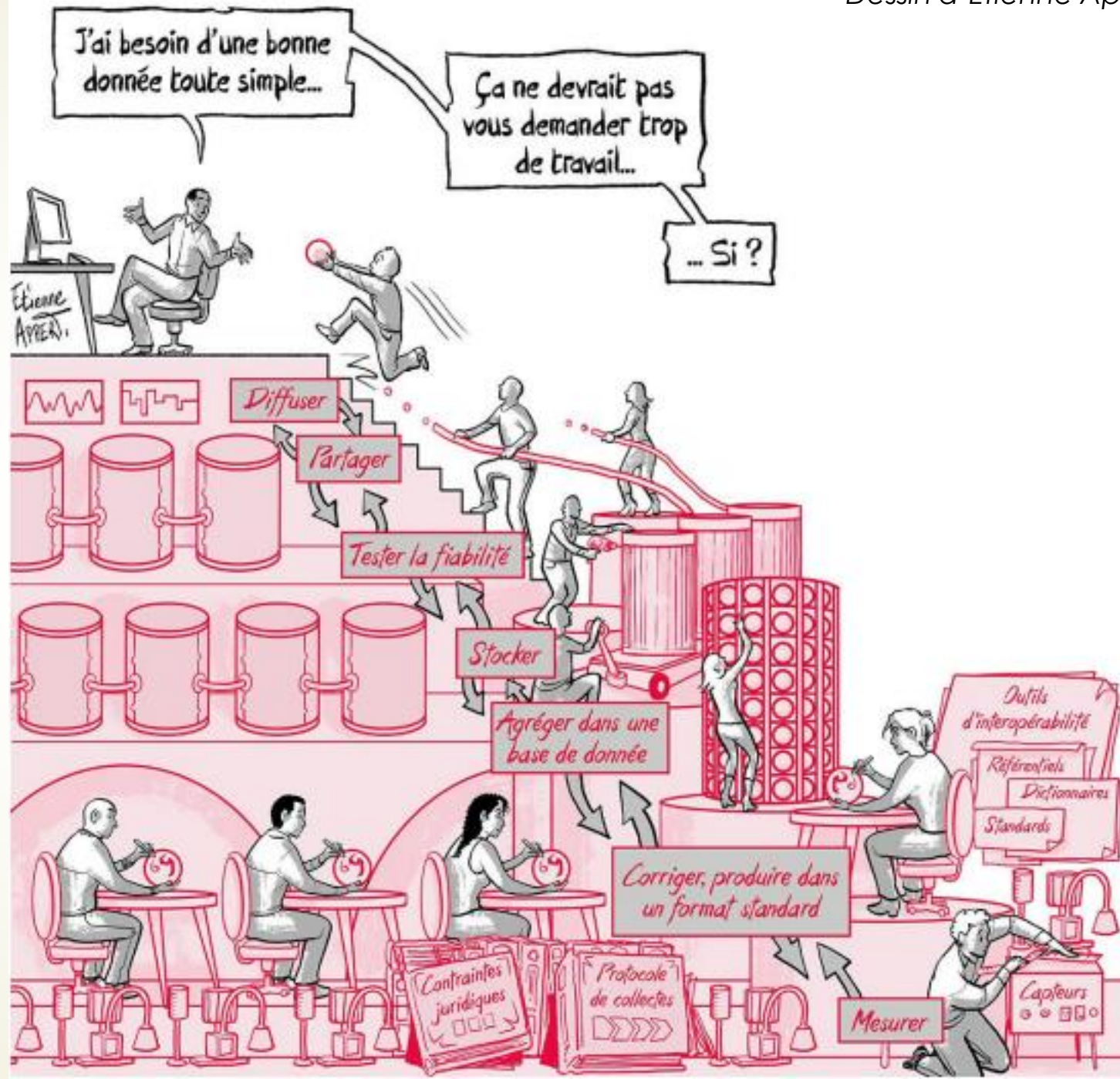
- BBEES : **accompagnement, conseil** tout au long du cycle de vie
- Pas de silos étanches avec Huma-Num : **complémentaires** et **interopérables** (Recherche Data Gouv, Data Terra)

Clé de la réussite : **sensibilisation** (formations de proximité, intervention dans les modules de l'école doctorale, la formation continue au MNHN et au CNRS).



Merci

isabelle.baly@mnhn.fr
amandine.henon@mnhn.fr
cecile.callou@mnhn.fr
contact@indores.fr



CC-BY 4.0