


Pourquoi créer un axe thématique « Cycle de vie des données » au CEFÉ ?

Marie-Claude Quido



Ce(tte) œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Partage dans les Mêmes Conditions 4.0 International.

Vous êtes autorisé à :

Partager — copier, distribuer et communiquer le matériel par tous moyens et sous tous formats

Adapter — remixer, transformer et créer à partir du matériel

Selon les conditions suivantes :



Attribution — Vous devez mentionner le nom de l'auteur de la manière suivante :
« Marie-Claude Quidoz, CEFÉ-CNRS, 2021 »



Partage dans les Mêmes Conditions — Si vous modifiez, transformez ou adaptez cette œuvre, vous n'avez le droit de distribuer votre création que sous une licence identique ou similaire à celle-ci.

Le CEFE , c'est ... (1/4)

- Centre d'Écologie Fonctionnelle et Évolutive (CEFE)
 - ✓ Institut écologie et environnement du CNRS (INEE)
 - ✓ Créé en 1961
- Un grand laboratoire en écologie
 - ✓ 294 personnes (dont 94 chercheurs ; 49 IT ; 76 doctorants/post doctorants)
- Un projet à large spectre
 - ✓ Comprendre la dynamique, le fonctionnement et l'évolution du vivant, de « la bactérie à l'éléphant » et « du génome à la planète »
- Un grand producteur de données
 - ✓ D'observation / expérimentales / computationnelles



Le CEFE , c'est ... (2/4)

- 4 départements scientifiques
 - ✓ 16 équipes de recherche
- 6 plateformes techniques
 - ✓ Documentation, Bibliothèque
 - ✓ Génomique, Écologie Moléculaire et Evolution Expérimentale (GEMEX)
 - ✓ Analyses Chimiques (PACE)
 - ✓ Programmes à Long Terme (PLT)
 - ✓ Terrain d'Expérience (TE)
 - ✓ Système d'Information en Écologie (SIE)
- 5 axes scientifiques transversaux (création officielle en 2020)
 - ✓ Écologie et agronomie
 - ✓ Evolution expérimentale
 - ✓ Sciences et sociétés
 - ✓ Biodiversité numérique
 - ✓ Cycle de vie des données



Le CEFE , c'est ... (3/4)

- **Établissement(s) de rattachement**

- ✓ Principal : CNRS
- ✓ Principal : Université Montpellier (UM)
- ✓ Principal : IRD
- ✓ Principal : EPHE
- ✓ Principal : Université Paul Valéry Montpellier 3 (UPVM3)
- ✓ INRAE, SupAgro

- **Institut(s) de rattachement**

- ✓ INEE
- ✓ Secondaire : INSHS

- **Contrat de recherche financés par des institutions publiques (HCERES 01/01/2014 au 30/06/2019)**

- ✓ 23 contrats européens (dont 6 ERC) en tant que porteur
- ✓ 111 contrats nationaux (dont 50% ANR) en tant que porteur

- **400 publications par an**



Le CEFE , c'est ... (4/4)

- De nombreux modèles de plan de gestion de données
 - ✓ INRAE, IRD, UM, UMPV3
 - ✓ ANR, ERC
- De nombreux entrepôts de données
 - ✓ Institutionnel : INRAE, IRD, TGIR Huma-Num, UM (en 2021), INEE (en 2021)
 - ✓ Thématique : GenBank
 - ✓ Multidisciplinaire : Zenodo, DRYAD
- De nombreux documents officiels
 - ✓ CNRS : feuille de route pour la science ouverte / plan données de la recherche
 - ✓ UMPV3, EPHE : charte science ouverte
 - ✓ ...



Pourquoi créer un axe transversal ?

- Mieux valoriser les données de la recherche
- Faire monter en compétence le personnel
 - ✓ Face aux déluges de mots clefs / sigles / concept
- Fournir un point d'entrée unique
 - ✓ Des zones blanches / chevauchements entre plateformes
- Améliorer la qualité des données
 - ✓ Sur toutes les étapes du cycle de vie des données
- Amplifier la dynamique déjà initiée
 - ✓ «Libérez vos Publications !» 2018



Animatrices de l'axe thématique

- Véronique Arnal (biologiste moléculaire)
 - ✓ Équipe BEV / plateforme GEMEX
 - ✓ Plateforme ADN dégradé
 - ✓ Référente données expérimentales et terrain
 - ✓ DU «Gestion des données de la science – Scientific Data Management»
- Anne Gorgeon (documentaliste)
 - ✓ Plateforme Documentation, Bibliothèque
 - ✓ CIST, Réseau Doccitanist, Réseau Renatis
 - ✓ Référente publications et HAL
- Marie-Claude Quidoz (gestionnaire de données)
 - ✓ Plateforme Système d'Information en Écologie
 - ✓ Réseau rBDD, GT Atelier Données
 - ✓ Référente données d'observation et préservation des données

Des réalisations (1/4)

(pour répondre à des questions des producteurs de données)

La sauvegarde des données

Pour remplir cette rubrique des plans de gestion des données, vous aurez besoin d'informations qui dépendent de l'infrastructure mise en place au CEFE.



Si vos données sont sur un espace du serveur Maison

- La sauvegarde a lieu du lundi au vendredi à 23 heures.
- Les sauvegardes sont conservées 30 jours.
- La restauration des dossiers pour lesquels l'utilisateur possède des droits est possible par l'utilisateur lui-même et ce pendant 15 jours grâce à un simple clic droit sur le dossier à restaurer. La restauration des dossiers peut toujours être faite par le service informatique.
- La restauration peut être faite au niveau d'un dossier ou d'un fichier.

Si vos données sont sur votre poste individuel sur le domaine CEFE

- La sauvegarde des postes fixes Windows a lieu du lundi au vendredi à partir de 11 Heures sauf s'ils sont éteints. Pour les portables Windows & Macintosh, la sauvegarde a lieu dès qu'ils se connectent à internet et ensuite toutes les 4 heures. Pour les linux, la sauvegarde n'est pas encore en place.
- 21 sauvegardes sont conservées
- La restauration des dossiers est faite par le service informatique.
- La restauration peut être faite au niveau d'un dossier ou d'un fichier.

Des réalisations (2/4)

(pour se former dans un premier temps et avoir un exemple concret)

PGD 1 : Suivi (fictif) de population de poissons dans le lac du Bourget

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

Créateurs du PGD : Marie-Claude QUIDOZ, VERONIQUE ARNAL, Anne Gorgeon

Affiliation du créateur principal : CNRS

Modèle du PGD : ANR - Modèle de PGD (français)

Dernière modification du PGD : 17/09/2020

Financier : ANR

Numéro de subvention : ANR-2020-000-000

Résumé du projet :

Avertissement : il s'agit d'un PGD destiné à être utilisé à des fins pédagogiques. Le projet est fictif, toute ressemblance avec des projets existants ou ayant existé ne saurait être que fortuite

L'objectif principal est d'observer les conséquences sur la faune marine des actions de dépollution engagées depuis le milieu des années 1970 dans le lac du Bourget qui présentait un phénomène majeur d'eutrophisation.

Le lac du Bourget n'a pour l'instant fait l'objet que de rares campagnes de prélèvements non standardisés.



DMPs publics

Des réalisations (3/4)

(pour essayer de s'y retrouver dans ce déluge de mots)

DEFINITIONS AUTOUR DU CYCLE DE VIE DES DONNEES Sélection de plusieurs définitions sur des termes courants en Science Ouverte



Des termes peuvent être ajoutés ou supprimés : merci de nous faire part de vos suggestions !

Axe : Cycle de vie des données
Auteur : Véronique Arnal
Relecteurs : Anne Gorgeon et Marie-Claude Quidoz
Mise à jour : 28/10/2020

Annuaire / Répertoire d'entrepôts (cf. Re3data, cf. OpenDOAR)

Site recensant des entrepôts de données, permettant de filtrer ses recherches par critère, comme re3data, OAD, OpenDOAR ; voir également l'annuaire d'entrepôts certifiés dans CoreTrustSeal.

Pour en savoir plus :

<https://www.re3data.org/about>

http://oad.simmons.edu/oadwiki/Data_repositories

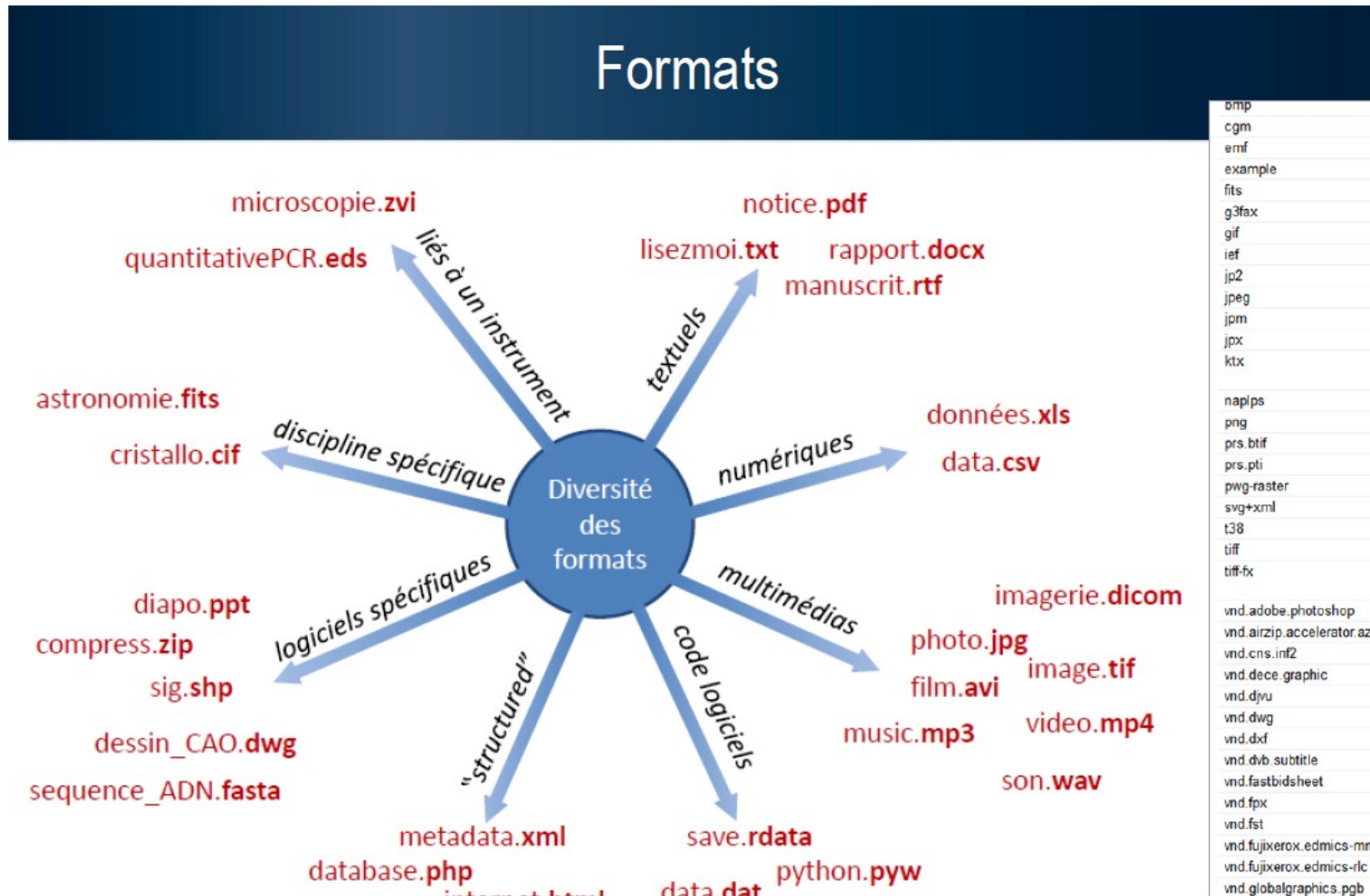
<http://v2.sherpa.ac.uk/opensoar/>

<https://www.coretrustseal.org/why-certification/certified-repositories/>

Pour trouver un entrepôt, exemple de moteur de recherche Pangaea, Data Publisher for Earth & Environmental Science : <https://www.pangaea.de/>

Des réalisations en cours (4/4)

(pour apporter des conseils aux producteurs de données)



Source : (10)



Méthode de travail

- Réunions mensuelles
 - ✓ Point sur les formations / séminaires suivis
 - ✓ Demandes internes
 - ▶ PGD, entrepôts de données, RGPD
- Communication
 - ✓ Intranet
 - ✓ Lettre Hebdomadaire
 - ✓ Intervention à des séminaires / formations
- Objet d'étude fictif
 - ✓ Suivi de population de poissons dans le lac du Bourget



Bilan au bout d'un an

- Ne pas se démoraliser
 - ✓ Le sujet est vaste
 - ✓ les compétences à acquérir sont nombreuses
 - ✓ Le sujet est en pleine structuration
 - ✓ Le personnel n'est pas fan
- Ne pas se faire déborder
 - ✓ On pourrait passer sa journée à suivre des webinaires
- Se fixer une route et essayer de la suivre
 - ✓ Développer l'acquisition de bonnes pratiques au CEFE
- Respecter les passages obligatoires
 - ✓ DMP Projet