



DATA
TERRA



ODATIS

Campus de la Mer – 14 mai 2024 /

L'Atelier technique

Joël Sudre (joel.sudre@data-terra.org)



Origine de l'Atelier Technique

Un petit historique:

➤ RESOMAR :

- Mise en place d'un atelier technique pour fédérer les ingénieurs et les chercheurs autour de la données
- Un besoin de plus en plus nécessaire

➤ Création du pôle ODATIS (2016)

- Mise en place d'un atelier technique pour aider les CDS du pôles à travailler en synergie
- **Inform**er **sur** **les** **bonnes** **pratiques** à mettre en œuvre

Campus de la Mer, le mardi 14 mai 2024

Objectifs de l'atelier technique

Réunir les partenaires pour:

- Echanger sur:
 - les pratiques
 - les besoins
 - les attentes
- **Définir les orientations techniques du pôle** (Cahier des Charges des CDS)
- **Informar sur les bonnes pratiques** à mettre en œuvre
- **Homogénéiser et coordonner les CDS** et les producteurs de données par l'émission de recommandations (FAIR, Open Science,...)
- **Aider à la certification CoreTrustSeal**
- **Montrer des mises en pratique** (exercices d'application de logiciels)

Campus de la Mer, le mardi 14 mai 2024

Archive de l'Atelier Technique

- L'ensemble des sessions de l'atelier technique sont disponibles sur le site ODATIS (18 sessions)
- <https://www.odatis-ocean.fr/activites/ateliers-techniques>
- Les 3 derniers ateliers:
 - **Atelier de travail « Catalogue ODATIS »:** Atelier de travail, interne aux membres du bureau exécutif du pôle, clarification du périmètre des catalogues ODATIS et Sextant avec l'entrepôt Seanoe; faciliter la découverte et l'accès aux jeux de données du catalogue ODATIS, assurer la traçabilité & visibilité des catalogues et producteurs des données moissonnées et de faciliter la réutilisation des jeux de données du catalogue ODATIS (identification des fonctionnalités à ajouter pour permettre l'intégration des jeux de données avec des services types VDAP (visualisation/analyse/traitement à distance), création d'un catalogue de services ODATIS).
 - **Atelier données de la Flotte Océanographique Française – FOF:** clarifier le workflow des données depuis la préparation de la campagne jusqu'à la bancarisation et la diffusion de ces données; Aspects réglementaires, documentaires et procédures, Définir un plan d'action pour élaborer un plan de gestion de données commun ODATIS / IR*FOF; Identifier des sujets d'intérêt autour des questions de réglementation et définir un plan d'action; Identifier des sujets d'intérêt autour des services supports aux campagnes et définir un plan d'action.
 - **Atelier « Données bioinformatique de diversité »:** organisé conjointement par les pôles ODATIS et PNDB, composantes de l'IR Data Terra; et en collaboration avec ABIMS, SEBIMER et le MIO: Bonnes pratiques FAIR (standards de données et métadonnées), workflow, outils, web services de bio-analyse.

Campus de la Mer, le mardi 14 mai 2024

Quelques recommandations issues de l'atelier technique (1/3)

- **Métadonnées** de découverte à la norme ISO 19115
Moissonnage CSW (éventuellement OAI/PMH)
- Métadonnées d'usage incluses dans les fichiers de données
- **Format** des données (physique, chimie, biogéochimie, déchets, ...):
 - ASCII TSV (ODV spreadsheet normalisé SDN)
 - NetCDF4 (sans groupes) convention Climate and Forecast (V1.6 ou +)
 - Formats spécifiques :Image (JPEG2000, GeoTIFF), géographique (shape)
- Attributs de **paramètres** (double convention):
 - Convention CF (pour les param. Phy. avec attributs SDN liste P01)
 - Convention SeaDataNet (Biogéochimie, chimie, ...) + attributs supplémentaires « long name » en convention CF
- **Biologie** (en cours, en commun avec PNDB et participation du SAR):
 - *Darwin Core Archive, (EML)*
 - *WORMS (en première instance), nécessaires adaptations « locales »*

Quelques recommandations issues de l'atelier technique (2/3)

- Des recommandations d'implémentation
 - Unité des paramètres (P06 SDN) et la discipline (P08 de SDN)
 - Pour les fiches de métadonnées EV obligatoires (P02 de SDN)
 - Nomenclature d'un paramètre: liste P01 de SDN (si découverte P02)
 - Date et heure entier en « long » avec un offset et un scale factor (UTC)
 - Axes et Orientations à bien préciser (Attention à la profondeur positive!)
 - Positionnement EPSG (European Petroleum Survey Group)
- **Services de données** (une liste adaptable)
 - OpenDAP (Thredds Data Server, ErrDAP)
 - OGC-WMS (visualisation)
 - OGC WCS (données « grilles »)
 - ~~➤ OGC WFS (avec modèles INSPIRE), (?), éventuellement via EMODnet~~

Quelques recommandations issues de l'atelier technique (3/3)

Un point important: la traçabilité des données

- Citer la source des données et des informations (indissociable du travail scientifique)
- **Eviter les doublons (nécessité de données interoperables)**
- Identifier l'origine des données (...qualité)
- Reconnaître l'importance de l'effort d'observation
- Etablir un « constat » de l'état de l'art à un instant T
- Respect de certaines obligations réglementaires (APA, DCSMM 19.3)?
- Utilisation d'identificateurs permanents
 - Choix des **Digital Object Identifiers (DOI)**
 - Notion **d'entrepôts** (copies/ images stabilisées des données / Versions)



Merci pour votre attention

