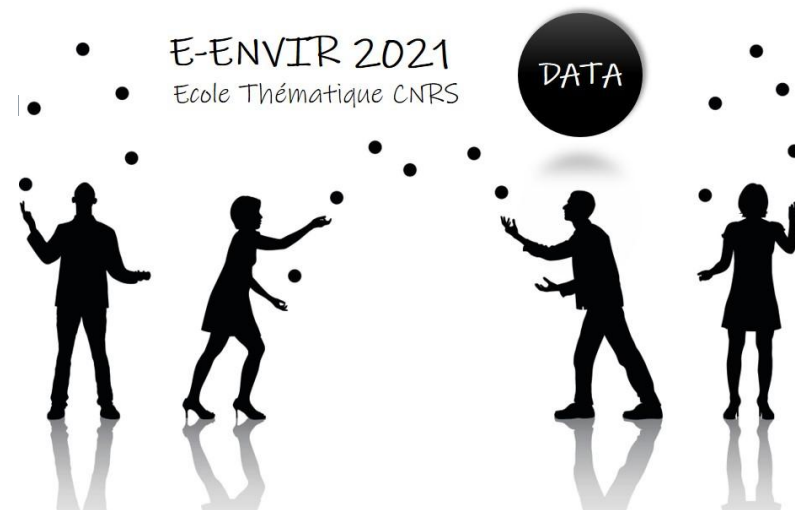




INTEROPERABLE & REUSE

INTRODUCTION AUX CONCEPTS CLEFS ET IMMERSION



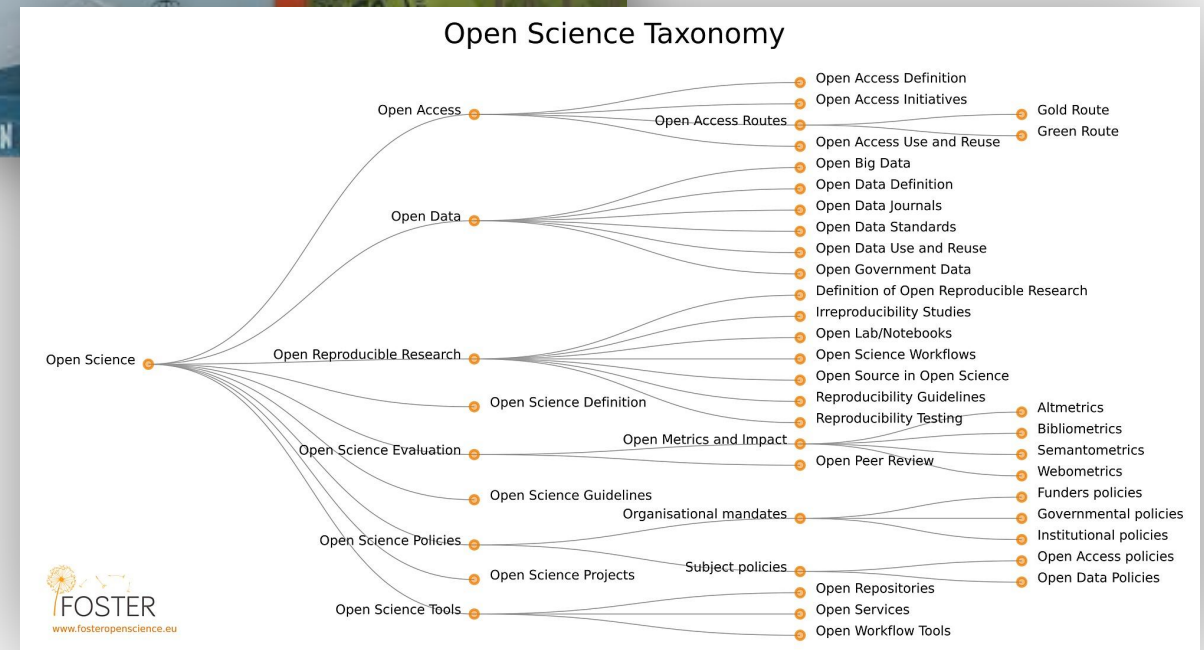
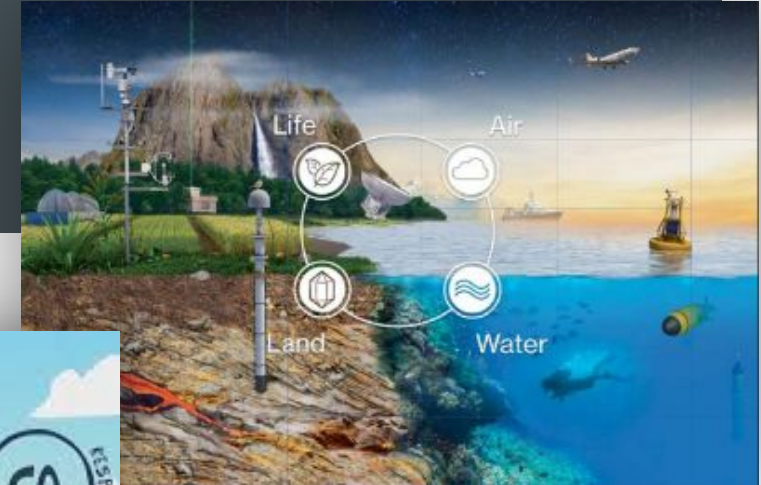
ENJEUX DU PARTAGE DES DONNÉES

- Favoriser une approche plus intégrée interdisciplinaire du Système Terre
 - en permettant de combiner des données multi-sources et multi-disciplinaires

= rendre les données
INTEROPERABLES

- Science ouverte
 - Capitaliser sur les résultats précédents (qualité),
 - éviter la duplication des efforts (efficacité)

en REUTILISANT des données
acquises par d'autres





INTEROPÉRABILITÉ

DONNÉES & MÉTADONNÉES

DÉFINITION (1/3)

Interopérabilité kesako

- "L'interopérabilité est la capacité que possède un produit ou un système, dont les interfaces sont intégralement connues, à fonctionner avec d'autres produits ou systèmes existants ou futurs et ce sans restriction d'accès ou de mise en œuvre. Il convient de distinguer « interopérabilité » et « compatibilité »."
- Intéropérabilité entre 2 systèmes veut dire via l'utilisation d'une norme externe à ces systèmes. Une norme pouvant être un format de données, un standard de métadonnées, ...



DÉFINITION (2/3)

Interopérabilité et l'importance de la métadonnée

- Elles permettent d'avoir des informations sur des données (ou autre objet numérique) permettant leur accès, leur lecture, leur compréhension. Ainsi, même si des données sont complètement différentes en termes de formats de contenus, les métadonnées permettent d'y accéder, de les lire



DÉFINITION (3/3)

Un cas précis : les formats de données ouverts

- Ils sont interopérables, c'est à dire qu'ils peuvent être créés, lus et modifiés par tous les logiciels destinés à traiter le type du fichier (image, texte, audio, etc.). Les formats ouverts sont souvent des standards ouverts, comme le format .html qui sert à construire des pages Web

=> Exemple, format opendocument = format pour bureautique. Tous les logiciels l'utilisant peuvent lire et écrire ce format sans restrictions d'accès ou de mise en œuvre"



INTEROPÉRABILITÉ DES DONNÉES ET DES SYSTÈMES

Interopérabilité des données

- standards de description de (méta)données + format
- vocabulaire contrôlé et partagé sur le Web (FAIR)
- métadonnées doivent inclure des références vers d'autres métadonnées

Interopérabilité des systèmes

- protocole d'échange de (méta)données standardisés actionnables par des machines



INTEROPÉRABILITÉ PAR LES METADONNEES

Interopérabilité repose sur une norme externe (dans le sens où cette norme peut être utilisée par tous les systèmes non directement liés le souhaitant).

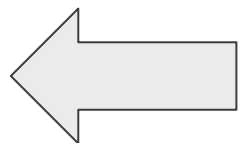
Dans les types de normes les plus utilisés pour permettre l'interopérabilité de systèmes d'informations, viennent notamment le **standard de métadonnée**. Un tel standard permet l'accès, la lecture et la compréhension des données d'un système d'information.

INTEROPERABILITE DES DONNEES

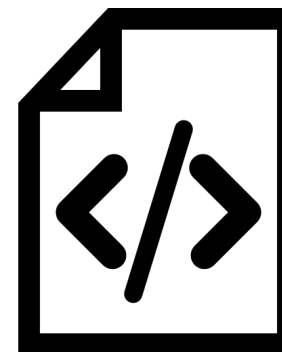
XML un format compréhensible par les machines

Pourquoi ????

que veut dire le
terme
"ecosystem" ?



dans la fiche de métadonnées



INTEROPERABILITE DES DONNEES

Pourquoi ????

que veut dire le
terme
“ecosystem” ?

ma définition correspond à ce qu’un groupe
d’expert a indiqué ici :

http://purl.obolibrary.org/obo/ENVO_01001110 et portant le code “ENVO:01001110”



dans la fiche de métadonnées
auteur de la fiche



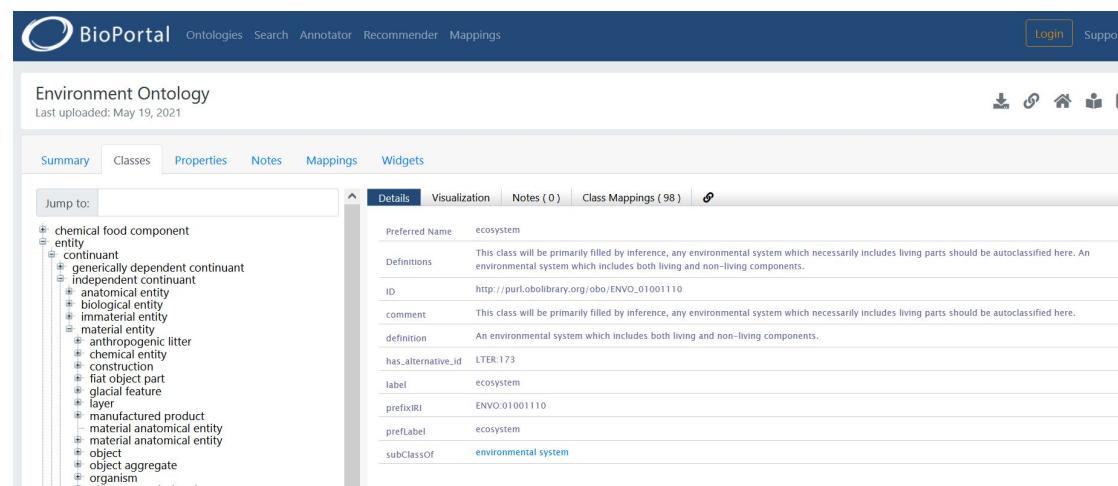
INTEROPERABILITE DES DONNEES

Pourquoi ???

que veut dire le
terme
“ecosystem” ?

ma définition correspond à ce qu'un groupe d'expert a indiqué ici :

http://purl.obolibrary.org/obo/ENVO_01001110 et portant le code “ENVO:01001110”



“An environmental system which includes both living and non-living components.”



INTEROPERABILITE DES DONNEES

Pourquoi ????

que veut dire le
terme
“ecosystem” ?

c’est vrai que je peux mettre à jour la fiche
de métadonnées pour indiquer ce code
“ENVO:01001110”...

“ecosystem” /
http://purl.obolibrary.org/obo/ENVO_01001110



dans la fiche de métadonnées

“An environmental system which includes both living and
non-living components.”



INTEROPERABILITE DES DONNEES

Pourquoi ????

que veut dire le terme
"ecosystem" décrit dans
la métadonnée?

que veut dire le terme
"ecosystem" décrit dans
la métadonnée?

que veut dire le terme
"ecosystem" décrit dans
la métadonnée?

que veut dire le terme
"ecosystem" décrit dans
la métadonnée?

Pour que tous les lecteurs / réutilisateurs soient informés de la signification des termes employés

"ecosystem" /
http://purl.obolibrary.org/obo/ENVO_01001110



dans la fiche de métadonnées

"An environmental system which includes both living and non-living components."

INTEROPERABILITE DES DONNEES

Pourquoi ????

Pour que les machines puissent également utiliser ces liens “webs” et pourquoi pas “reconcilier” seules les informations hétérogènes

que veut dire le terme “ecosystem” décrit dans la métadonnée?

que veut dire le terme “ecosystem” décrit dans la métadonnée?

que veut dire le terme “ecosystem” décrit dans la métadonnée?

que veut dire le terme “ecosystem” décrit dans la métadonnée?

“ecosystem” /
http://purl.obolibrary.org/obo/ENVO_01001110



dans la fiche de métadonnées



“An environmental system which includes both living and non-living components.”

INTEROPERABILITE DES DONNEES

Pourquoi ????

Pour que les machines puissent également utiliser ces liens “webs” et pourquoi pas “réconcilier” seules les informations hétérogènes

que veut dire le terme “ecosystem” décrit dans la métadonnée?

que veut dire le terme “ecosystem” décrit dans la métadonnée?

que veut dire le terme “ecosystem” décrit dans la métadonnée?

que veut dire le terme “ecosystem” décrit dans la métadonnée?

“ecosystem” /
http://purl.obolibrary.org/obo/ENVO_01001110



dans la fiche de métadonnées

“An environmental system which includes both living and non-living components.”



General

Annotations ⓘ is about biodiversity is about marine mammals is about Natural area

Dataset *is about* biodiversity

Find more information about this term at http://aims.fao.org/aos/agrovoc/c_33949

Find more datasets with measurements of biodiversity

VOCABULAIRE CONTRÔLÉ

Thèmes

☒ Show only valid items

Filter Etiquette

Etiquette

Adresses

Altitude

Bâtiments

Caractéristiques géographiques n

Caractéristiques géographiques c

Conditions atmosphériques

Dénominations géographiques

Géologie

Habitats et biotopes

Hydrographie

- **Liste “à plat” de termes.**
Ex: Registre des thèmes Inspire

Niveau
d'organisation
des termes
(sémantique)

- **Thesaurus:** ensemble organisé de termes avec des relations sémantiques hiérarchiques et associatives entre les termes
- **Ontologie:** ensemble organisé de termes/concepts avec des relations sémantiques variées décrivant un domaine de connaissance. Ex: FOAF (FriendOfAFriend)

+++

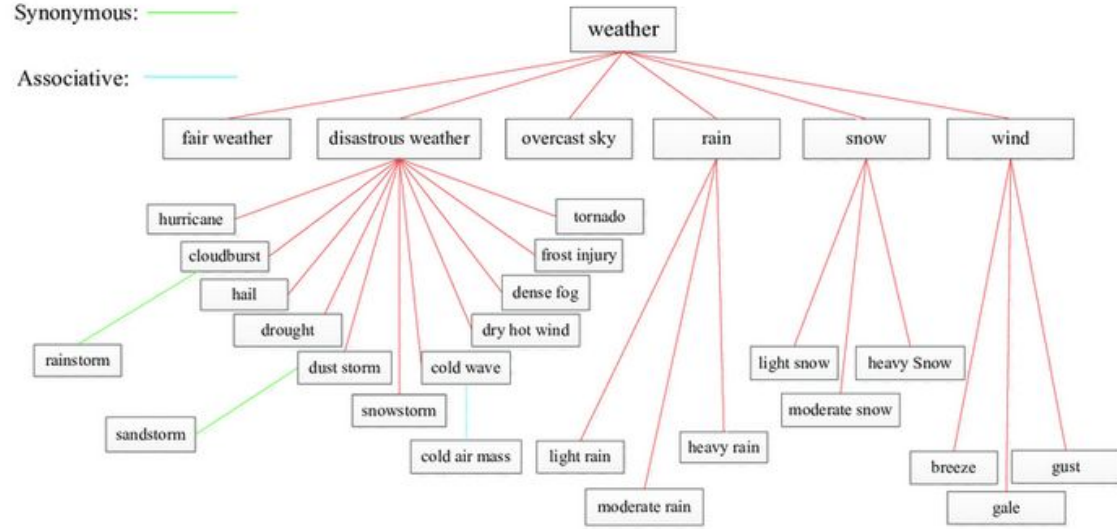
Relation Legend:

Hierarchical: —

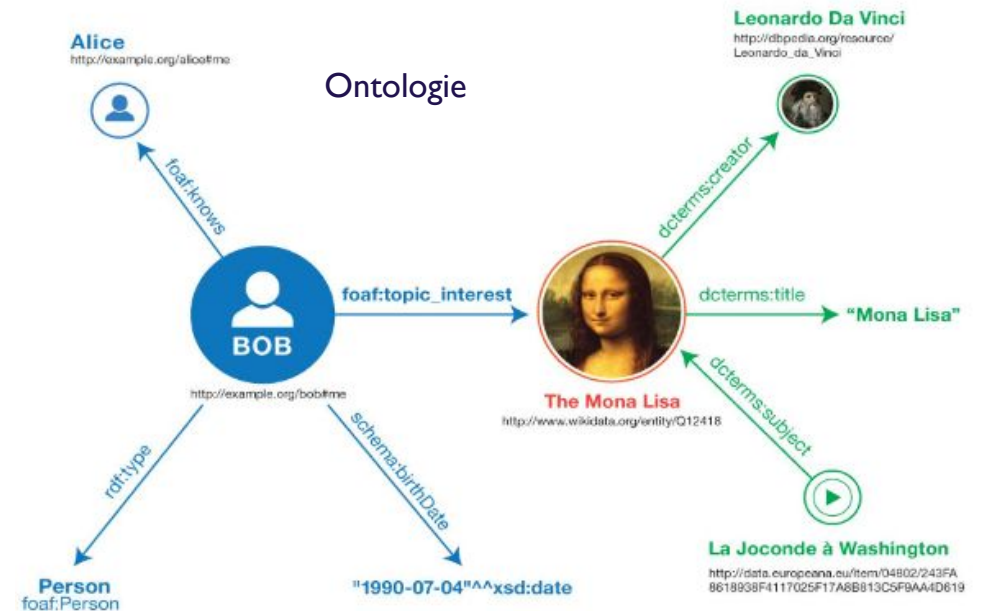
Synonymous: —

Associative: —

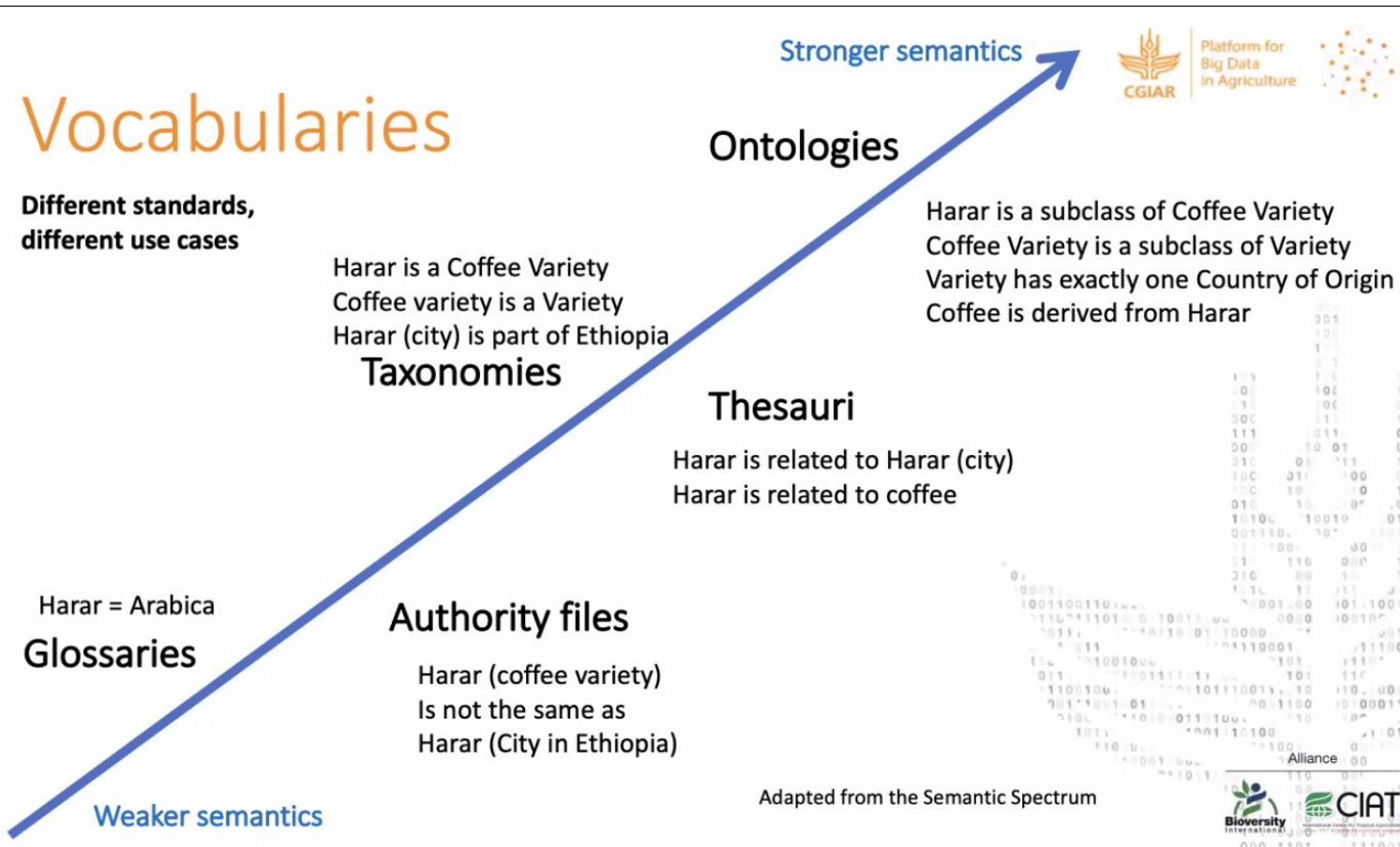
Thesaurus



Ontologie



VOCABULAIRE CONTRÔLÉ



THESAURUS FAIR

Publié sur le Web et décrit formellement selon les standards du **web sémantique**

standard SKOS (Simple Knowledge Organization System)

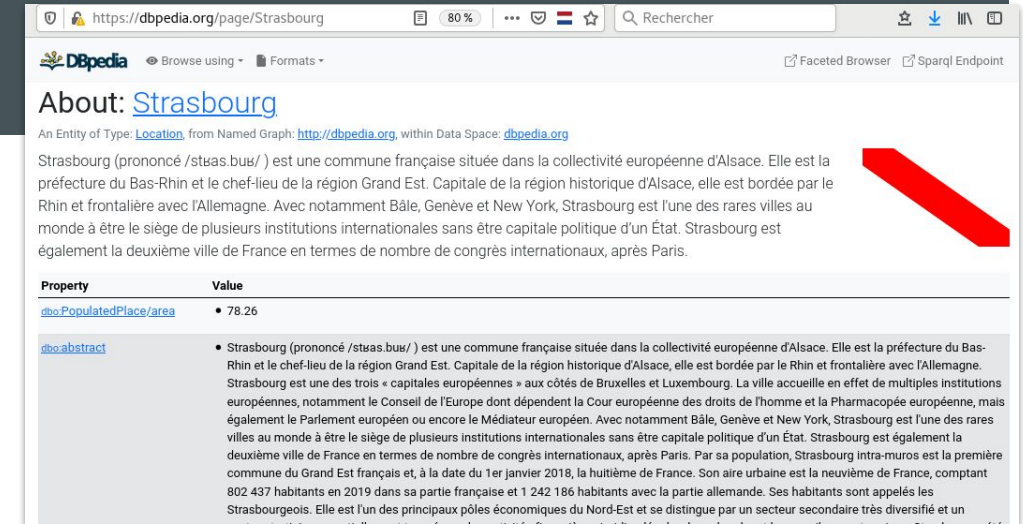
Pour chaque terme:

- Un **identifiant** pérenne = URI
Ex: <https://dbpedia.org/resource/Strasbourg> et <https://sws.geonames.org/2973783> représente la ville de Strasbourg

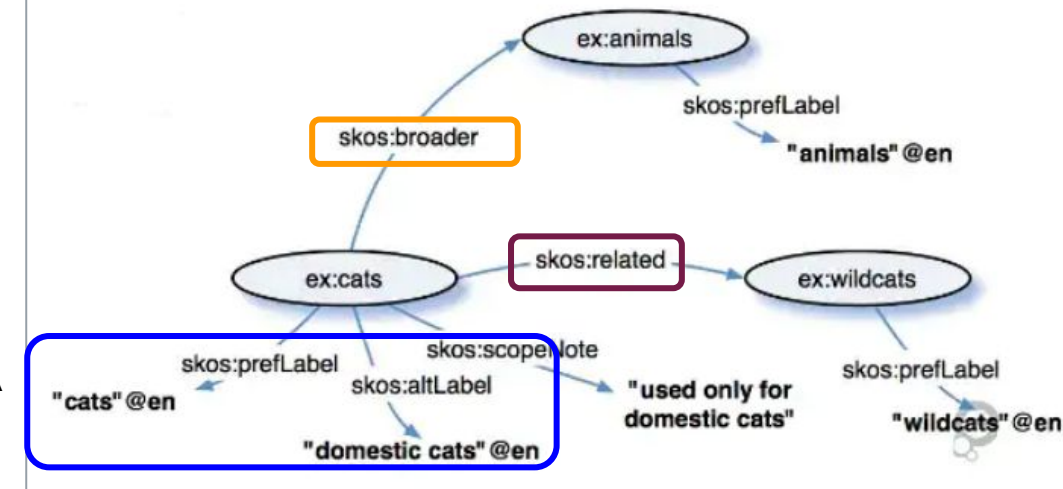
- Des **propriétés**: **terme préférentiel** (skos:prefLabel) ou alternatif (skos:altLabel), + équivalent dans d'autres langues (@en, @fr, ...), définition (skos:definition), ...

Entre termes:

- **Relations sémantiques hiérarchiques** estPlusGénérique (skos:broader), estPlusSpécifique (skos:narrower)
- **Relations sémantiques associatives** (similarité): estSimilaireA (skos:exactMatch), EstProcheDe (skos:closeMatch), estAssociéA (skos:related)



The screenshot shows the DBpedia page for Strasbourg. The URL is <https://dbpedia.org/page/Strasbourg>. The page title is "About: Strasbourg". The text describes Strasbourg as a commune in the Grand Est region of France, known for its international institutions. A table with properties and values is visible, including "dbo:PopulatedPlace/area" with the value "78.26".



PRÉSENTATION DE THESAURUS THÉMATIQUES

Généraliste (+environnement)

- AGROVOC (FAO): agriculture, foresterie, alimentation, environnement
- GACS : agronomie, regroupe 3 thesaurus (AGROVOC, NALT et CABT)

Environnement

- EnvThes - Environmental Thesaurus (eLTER)
- GCMD - Earth Science Thesaurus (NASA)
- GEMET - GEneral Multilingual Environmental Thesaurus (European Env. Agency)
- EARTH - Environmental Applications Reference Thesaurus
- OZCAR-Theia thesaurus

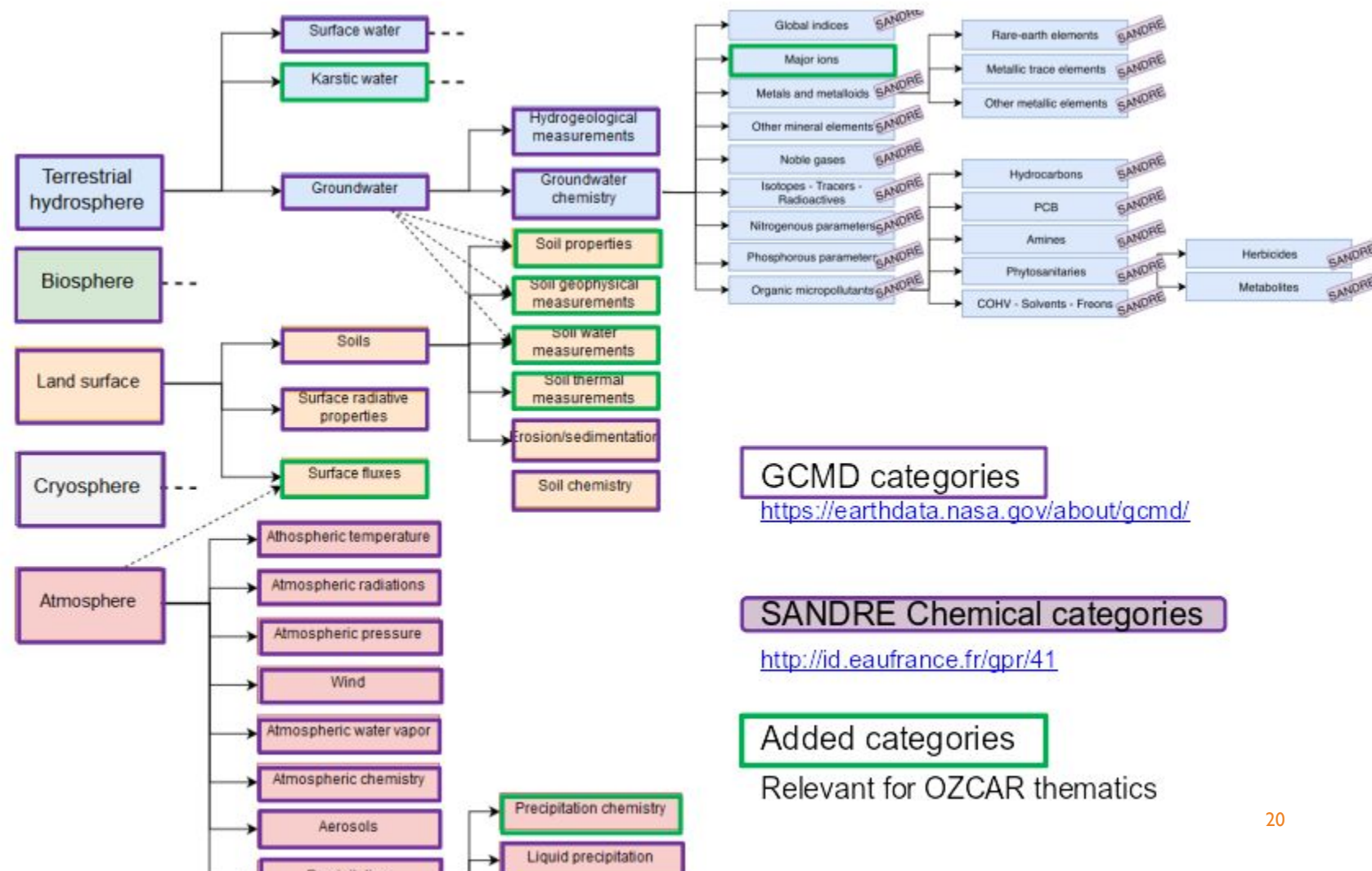
Ecologie & Biodiversité

- Thésaurus de la Biodiversité (Inist-CNRS)
- Thésaurus AnaEE (AnaEE-France)
- Référentiel taxonomique TAXREF (GBIF)
- Classification EUNIS des habitats (European Env. Agency)
- Thesaurus Of Plant characteristics - TOP (CEFE-CNRS)

THESAURUS OZCAR-THEIA

Thesaurus Zone Critique

- Noms & catégories de variables, objets d'intérêts observés (à venir)
- basé sur thésaurus NASA GCMD + SANDRE (chimie eau)
- anglais
- Alignements avec thésaurus internationaux: EnvThes, AGROVOC, GEMET, ANAEE, GACS, Earth



THESAURUS OZCAR-THEIA

Skosmos

Vocabularies

OZCAR-Theia thesaurus

Content language English ▾

Alphabetical Hierarchy Groups

[Nitrogenous parameters \(surface water\)](#)
[Noble gases \(surface water\)](#)
[Organic micropollutants \(surface water\)](#)
[Other mineral Elements \(surface water\)](#)
[Phosphorous parameters \(surface water\)](#)

Variables

[3-4-dichloroaniline](#)
[Absolute humidity](#)
[Absorbance at 254nm](#)
[Absorbance at 280nm](#)
[Aclonifen](#)
[Actinothermal index](#)
[Actual evapotranspiration](#)
[Aerosols chemistry](#)
[Aerosols radioactive isotope](#)
[Aerosols size distribution](#)
[Air pressure](#)
[Air temperature](#)
[Albedo](#)
[Atkatimity](#)
[Aluminium \(Al\)](#)

[... > Surface water > Surface water chemistry > Metals and metalloids \(surface water\) > Metallic trace elements \(surface water\) > Antimony \(Sb\)](#)
[Variables > Antimony \(Sb\)](#)

PREFERRED TERM	Antimony (Sb)
BROADER CONCEPT	Metallic trace elements (surface water) Variables
BELONGS TO GROUP	Variables
URI	https://w3id.org/ozcar-theia/variables/antimony/
Download this concept:	RDF/XML TURTLE JSON-LD

EXACTLY MATCHING CONCEPTS	Antimony antimony antimony http://aims.fao.org/aos/agrovoc/c_510	AnaEE Thesaurus EnvThes EnvThes aims.fao.org
---------------------------	---	---

- Terme préférentiel
- Liens hiérarchiques plus génériques
- Identifiant du terme (URI)
- Alignements avec thésaurus internationaux:

EnvThes, AGROVOC, GEMET, ANAEE, GACS, Earth

THESAURUS ENVTHES (eLTER)

Environmental Thesaurus

- Noms de variables, objets d'intérêt observés, méthodes de mesure, ...
- anglais majoritaire, multilingue en dev.
- Aligné avec AGROVOC, GEMET, EARTH, US-LTER, EuroVoc, DBpedia

<http://vocabs.lter-europe.net/EnvThes/10000>

ET E-ENVRI2I – 2 AU 5 NOVEMBRE 2021

The screenshot displays the EnvThes Semantic web interface. On the left, a sidebar contains navigation links: Home, Taxonomies, Ontologies, Help, and Server administration. The main content area is titled 'Concept Hierarchy' and shows a tree structure of concepts. A red box highlights the 'diatoms' concept, which is under 'algae' > 'chlorophyta'. To the right, the 'diatoms (Concept)' details are shown. The unique identifier 'http://vocabs.lter-europe.net/EnvThes/20588' is circled in red and labeled 'Identifiant unique'. The details include labels, description, metadata, standard relationships, and notes.

Concept Hierarchy

- envthes
 - constraint
 - deprecated concept
 - infrastructure
 - méthode
 - object of interest
 - entity
 - matrix
 - organisme
 - algae
 - chlorophyta
 - diatoms**
 - green alga
 - macroalgae
 - marine algae
 - animals
 - birds of prey
 - fungi
 - lichens
 - microbes
 - microorganism
 - part of organism
 - plants
 - procaryote
 - protists
 - organizational unit
 - substance
 - temporal entity
 - event
 - process
 - parameter
 - property
 - research focus

diatoms (Concept)

<http://vocabs.lter-europe.net/EnvThes/20588> **Identifiant unique**

Labels and Description

- preferred label:** diatoms (en)
- alternative label:** Diatomophyceae (en), bacillariophyceae (en)
- type:** [Concept](#)
- definition:** [Henderson's] n. common name for a member of the phylum Bacillariophyta, a group of algae characterized by delicately marked thin double shells of silica. (en)

Metadata

- creator:** herbert.schentz@umweltbundesamt.at
- created:** 2016-03-31T00:00:00Z
- modified:** 2018-12-18T08:53:13.492+00:00

Standard Relationships

- broader concept:** [algae](#)
- related concept:** [bacillariophyceae abundance](#), [bacillariophyceae species richness](#), [cover](#), [diatomophyceae coverage](#), [water](#)

Notes

- scope note:** US LTER controlled vocabulary (en)

Matching Relationships

- has exact match:** [diatoms](#), <https://vocab.lternet.edu/vocab/vocab/?tema=151>

THESAURUS BIODIVERSITÉ (INIST-CNRS)

- Nombreux concepts thématiques, différents regroupements dont variables essentielles de la biodiv (EBV)
- anglais/français
- aligné avec GEMET et AGROVOC

<http://dx.doi.org/10.13143/lotr.4647>

Biodiversity Thesaurus

Content language English ▾ Search

Home / Explore / Browse / Vocabularies / Biodiversity Thesaurus Help | français español

Alphabetical Hierarchy Groups

Grouping by Essential Biodiversity Variable

- Community composition
- Ecosystem function
- Ecosystem structure
- Genetic composition
- Species populations
- Species traits

Grouping by semantic type

- Activity / Behavior / Society
- Idea / Abstraction
- Living being
- Material / Matter / State of matter
- Object
- Phenomenon / Process / Property
- Space
- Structure of living being
- Technique / Method
- Work / Document

Grouping by subject field

- Biological sciences
- Earth and universe sciences
- Multidisciplinary

ecological parameter > ecological indicator > biodiversity indicator > ecological abundance

PREFERRED TERM **ecological abundance**

BROADER CONCEPT(S) biodiversity indicator

RELATED CONCEPT(S) community structure
diversity measure
dominant species
species diversity
species richness

SYNONYM(S) relative species abundance
species composition
species equitability
species evenness

BELONGS TO GROUP Analysis / Measure
Ecology_Miscellaneous
Property / Parameter / Characteristic_Miscellaneous
Species populations

IN OTHER LANGUAGES **abondance écologique** French
abondance relative des espèces
équitabilité des espèces
régularité de la distribution des espèces

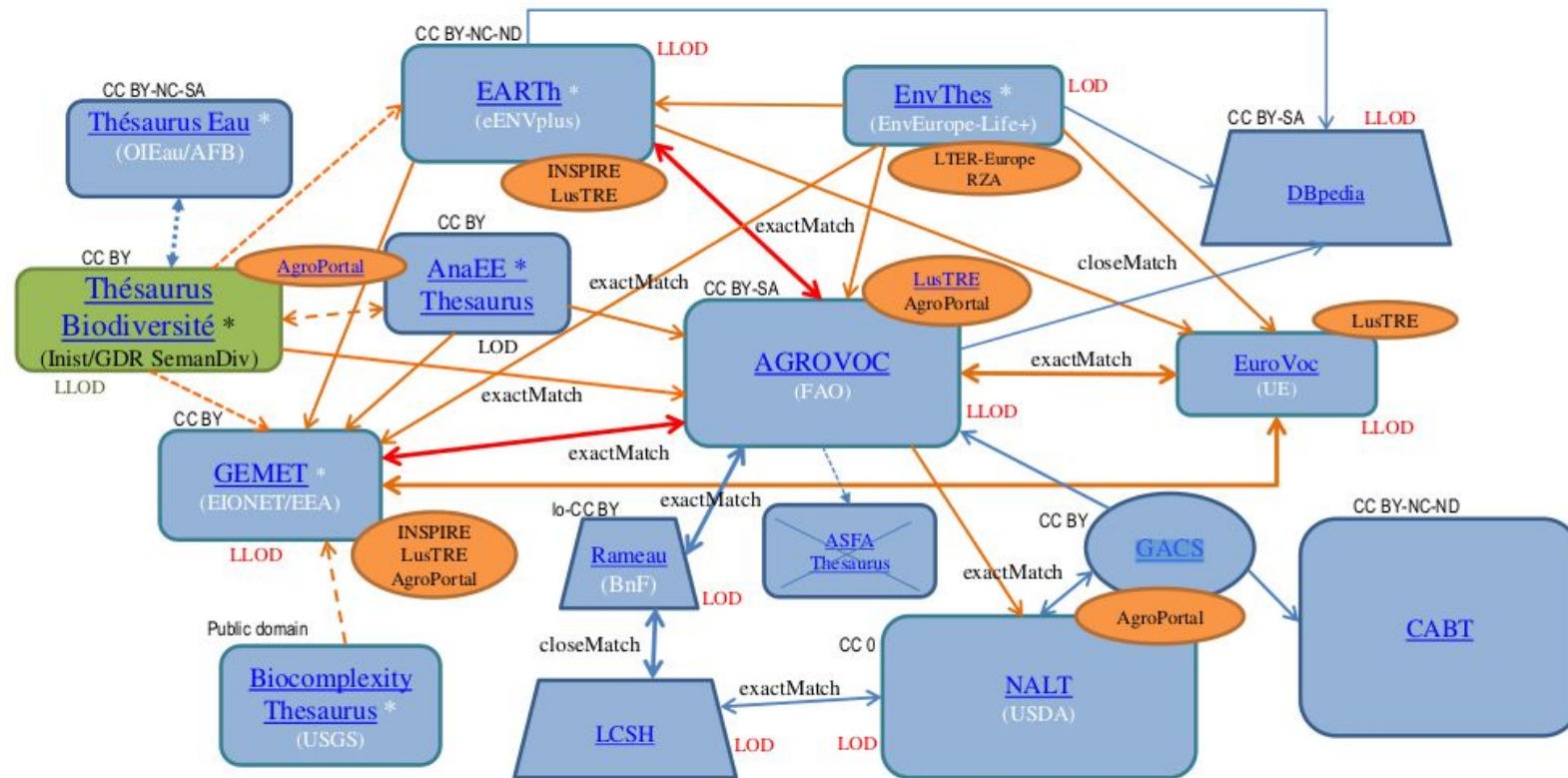
URI <http://data.loterre.fr/ark:/67375/BLH-M6XZHD55-J>

LIENS ENTRE THESAURUS THÉMATIQUES



Concept « Biodiversity » (concept-URI)
among **thesauri** on Environment and Agronomy in the **web of data**

Ex: concept biodiversity



PRÉSENTATION DES CATALOGUES DE THESAURUS



Class Search

Search for a class in multiple semantic

Enter a class, e.g. Shape, Trait, etc...

Include in your search...

- ☐ Property values
- ☐ Obsolete classes
- ☐ Views

Narrow your search to...

- ☐ Exact matches
- ☐ Classes with definitions

Categories

Type here to select categories or le

Semantic Resources

Start typing to select semantic res

The image shows the AgroPortal web interface. At the top is a green navigation bar with the AgroPortal logo and links for Browse, Search, Mappings, Recommender, Annotator, Projects, and Landscape. There are also Login and Support buttons. The main content area is titled "Browse" and shows a list of ontologies. The list includes: AnaEE Thesaurus (ANAEETHES), AGROVOC (AGROVOC), French Crop Usage (Classification des plantes cultivées en France en fonction des usages) (CROPUSAGE), Agriculture and Forestry Ontology (AFO), Agronomy Ontology (AGRO), and OntoBiotope (ONTOBIOTOPE). Each entry shows the ontology name, a brief description, and some statistics like the number of classes and the upload date. On the left side of the interface, there are filters for "Entry Type" (Ontology, Ontology View), "Uploaded in the Last", "Category" (Agricultural Rese..., Animal Science a..., Biodiversity and..., Breeding and Ge...), "Group" (AGBIODATA, AGROLD, CROP, INRAE), and "Format" (OBO, OWL).

ECOPORTAL <http://ecoportal.lifewatch.eu/search?opt=advanced>

AGROPORTAL: <http://agroportal.lirmm.fr/>

INTEROPÉRABILITÉ DES SYSTÈMES

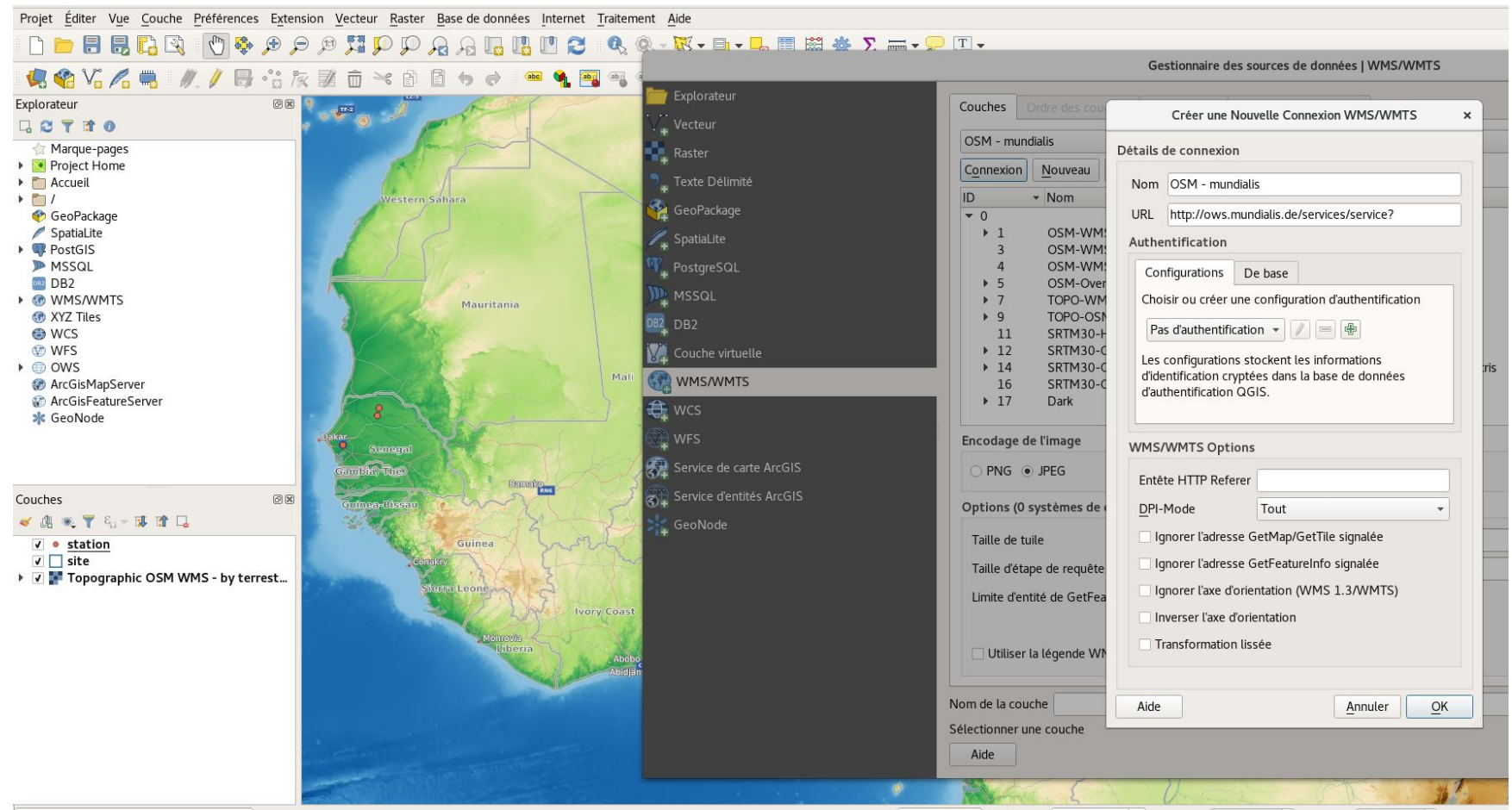
INTEROPÉRABILITÉ DES SYSTÈMES

Standards	Caractéristiques	Structure	Protocole d'accès	Scope	Exemple d'utilisation
Dublin Core	Standard interdisciplinaire relatif aux données numériques générales, avec 15 éléments basiques (titre, sujet, description, auteur, format, droits de propriété,...)	XML, RDF	OAI-PMH, CSW	Non spécifique	https://zenodo.org/record/1145786/export/xd#.YV7tPn06-Hs
DataCite	Standard de métadonnées associées aux DOIs. Pour décrire des objets digitaux.	XML	OAI-PMH	Non spécifique	https://zenodo.org/record/1145786/export/dcite4#.YV7t-n06-Hs
ISO/TC 211 & OGC	Standard de métadonnées pour des données à référence spatiale : ISO 19115/19139, avec dictionnaire des données (19110) et services de données (19119) ISO 19115/19139: s'appuie sur Dublin Core et complété par élmts supp. (référentiel spatial, catégories thématique,) Inspire : profil particulier de ISO 19115; rend obligatoire des métadonnées facultatives de ISO 19115	XML, GML, image	- catalogue de jeux de données: CSW - données vecteur et raster: WMS (image), WFS (vecteur), WCS (raster)	Géomatique et données à référence spatiale, pertinente pour les données env.	Geonetwork, Geoflow, Modèle Theia/OZCAR SI Les normes OGC ont pris une importance grandissante dans le cadre de la directive européenne INSPIRE(2007-2020)
EML	Standard de métadonnées en écologie, basé sur FGDC CSDGM (Content Standard for Digital Geospatial Metadata) et présentant une structure modulaire et extensible	XML, sous forme de "data package" quand association métadonnées/données (voir également codes sources)	OAI-PMH, JSON-LD	Ecologie, Biodiversité, des gènes aux socio-écosystèmes.	https://doi.org/10.48502/8bb5-pk85 Metashark
OGC O&M et SensorML (Sensor Web Enablement framework)	Standard de données et métadonnées pour des données d'observations issues de capteurs	XML	SOS	Données in-situ, capteurs	Theia/OZCAR SI

INTEROPÉRABILITÉ AVEC DES SYSTÈMES TIERS: OSM via QGIS

Récupérer des fonds
OpenStreetMap

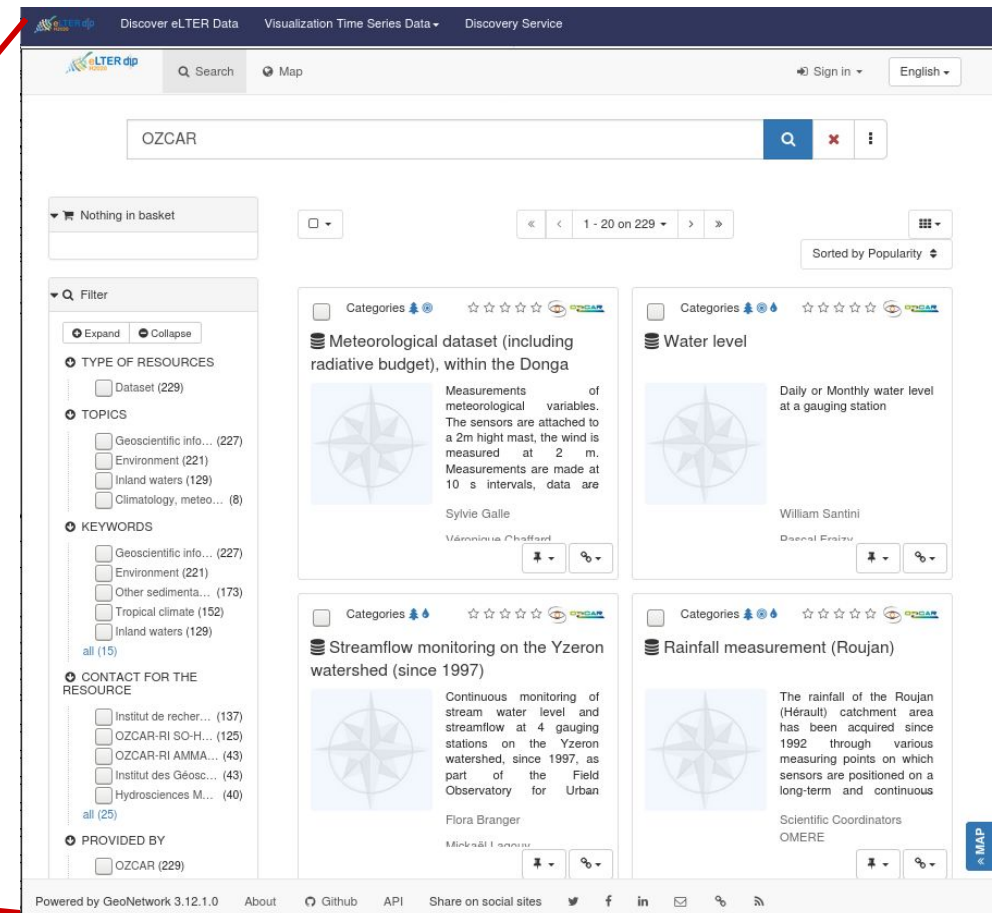
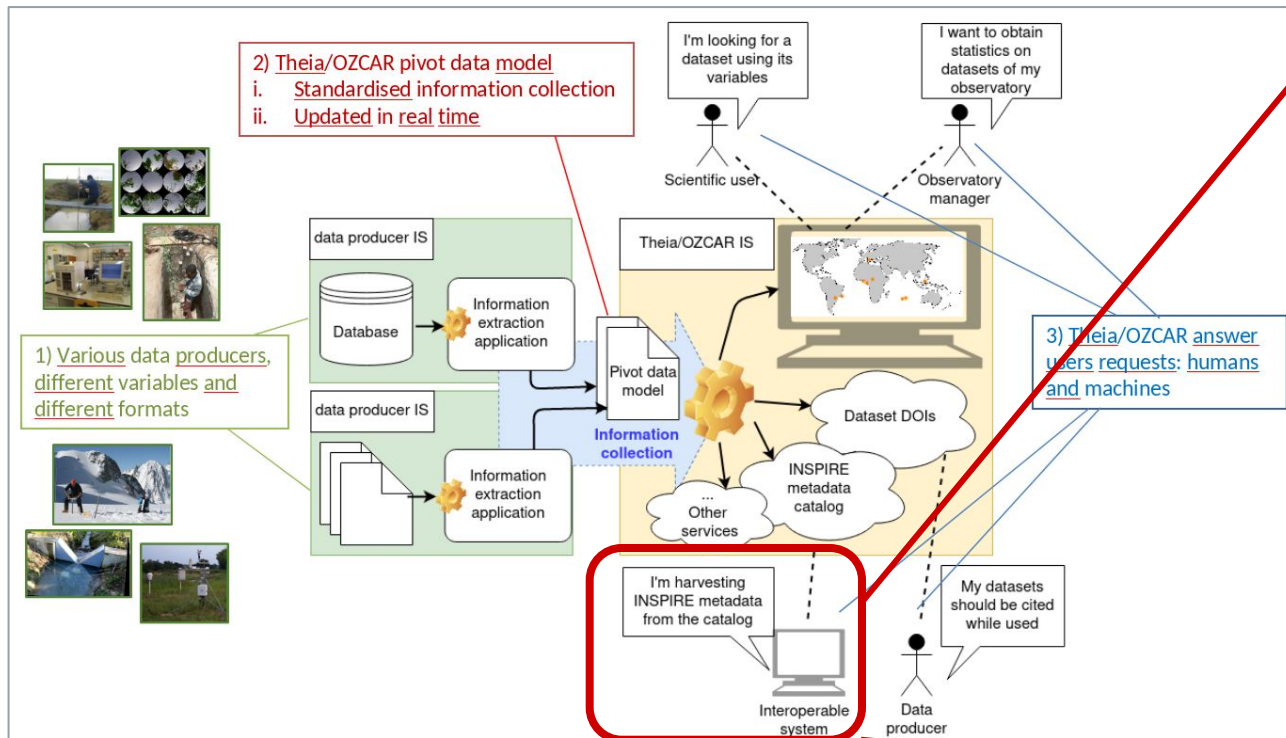
URL du service WMS:
[http://ows.mundialis.de/s
ervices/service?](http://ows.mundialis.de/services/service?)



INTEROPÉRABILITÉ AVEC DES SYSTÈMES TIERS: Theia/OZCAR /eLTER

Les jeux de données Theia/OZCAR sont moissonnables par des systèmes tiers via son web-service CSW

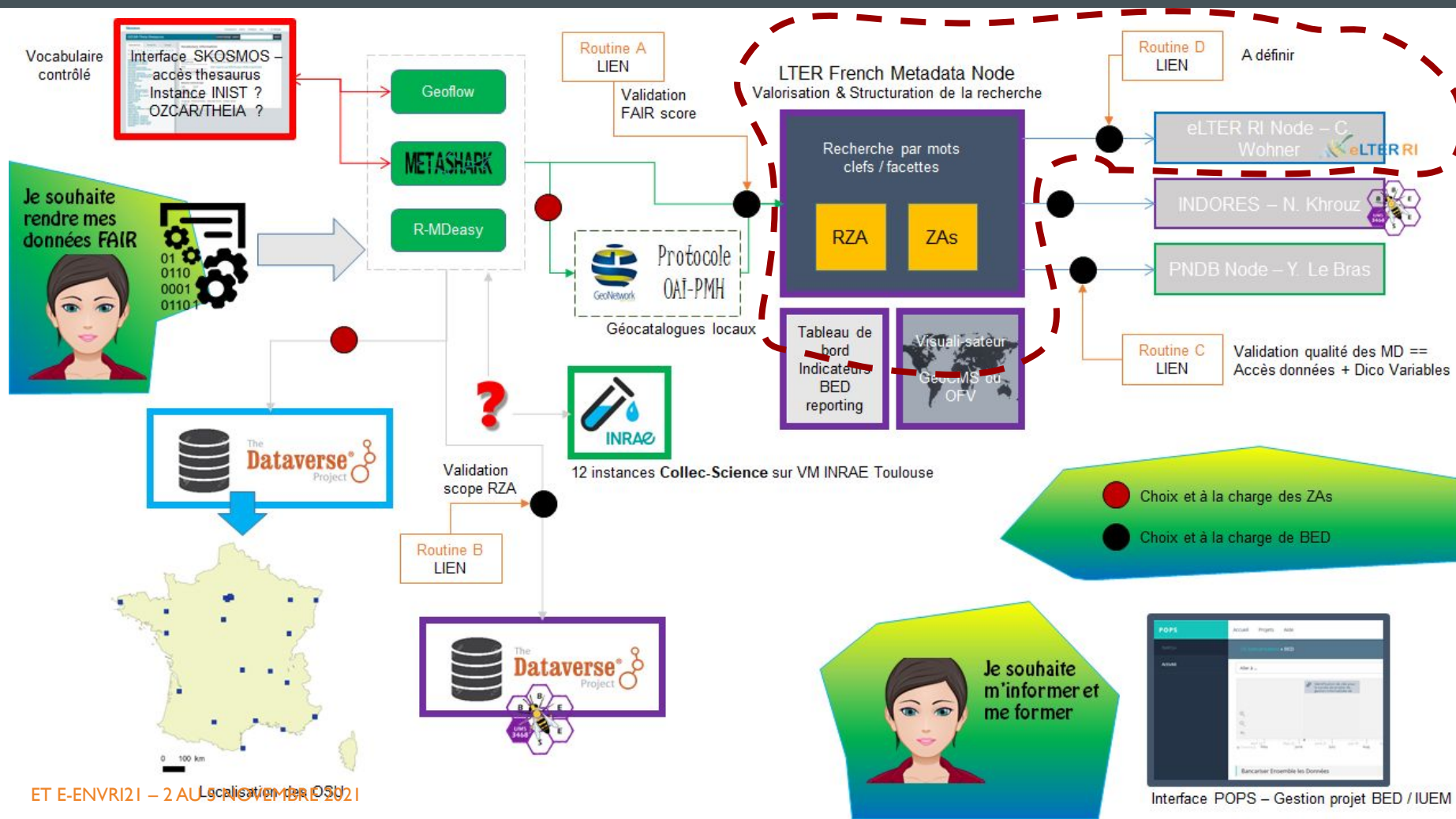
Catalogue eLTER: jeux de données OZCAR



service CSW Theia/OZCAR:

<https://in-situ.theia-land.fr/geonetwork/srv/eng/csw?service=CSW&version=2.0.2&request=GetCapabilities>

INTEROPÉRABILITÉ AVEC DES SYSTÈMES TIERS: RZA/eLTER



Un interfaçage au niveau des catalogues de données / moissonnage du GN-RZA

Des bonnes pratiques qui intègrent des composantes de la structuration des données et des métadonnées

- thesaurus EnvThes
- dictionnaire de variables
- etc.

INTEROPÉRABILITÉ AVEC DES SYSTÈMES TIERS: PNDB

MOISSONNAGE des systèmes d'information "Terre vivante" / Biodiversité par le PNDB

Via OAI-PMH

- ANAee
- RZA
- RARe
- Dataverse INRAe
- SEANOE
- ...

Récupération

- URL données
- licence
- métadonnées

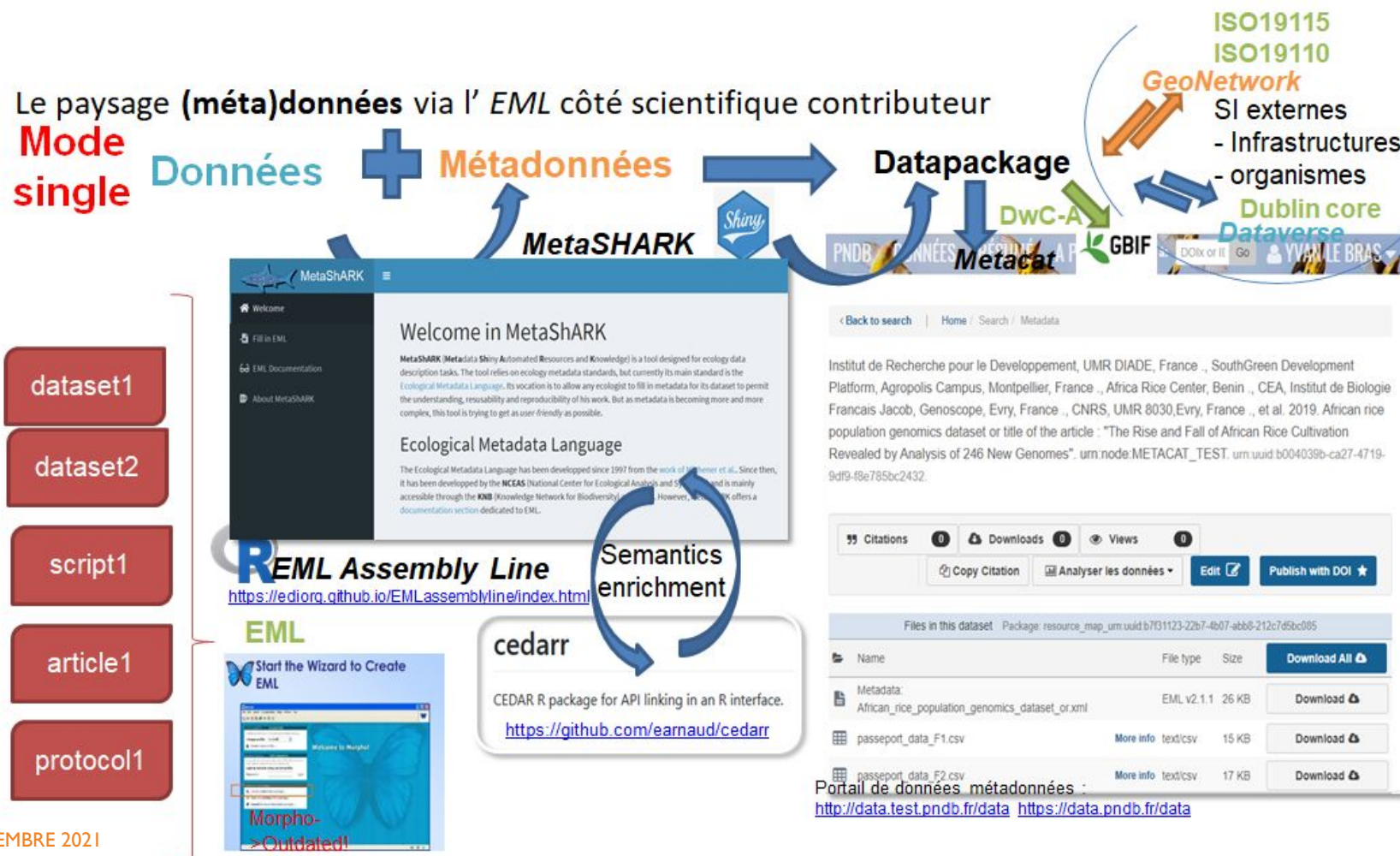
-

Les données PNDB ont vocation à être stockées sur entrepôts institutionnels (Dataverse CNRS-INEE, Dataverse INRAe,..) ou à défaut générique (Entrepôt MESRI, Zenodo)

L'URL de téléchargement direct des fichiers est alors renseignée dans la métadonnée.

-

INTEROPÉRABILITÉ AVEC DES SYSTÈMES TIERS: MetaSHARK



INTEROPÉRABILITÉ AVEC DES SYSTÈMES TIERS: MetaSHARK

Le paysage (méta)données via l' Ecological Metadata Language



Portail de données métadonnées :
<http://data.test.pndp.fr/data> <https://data.pndp.fr/data>

9



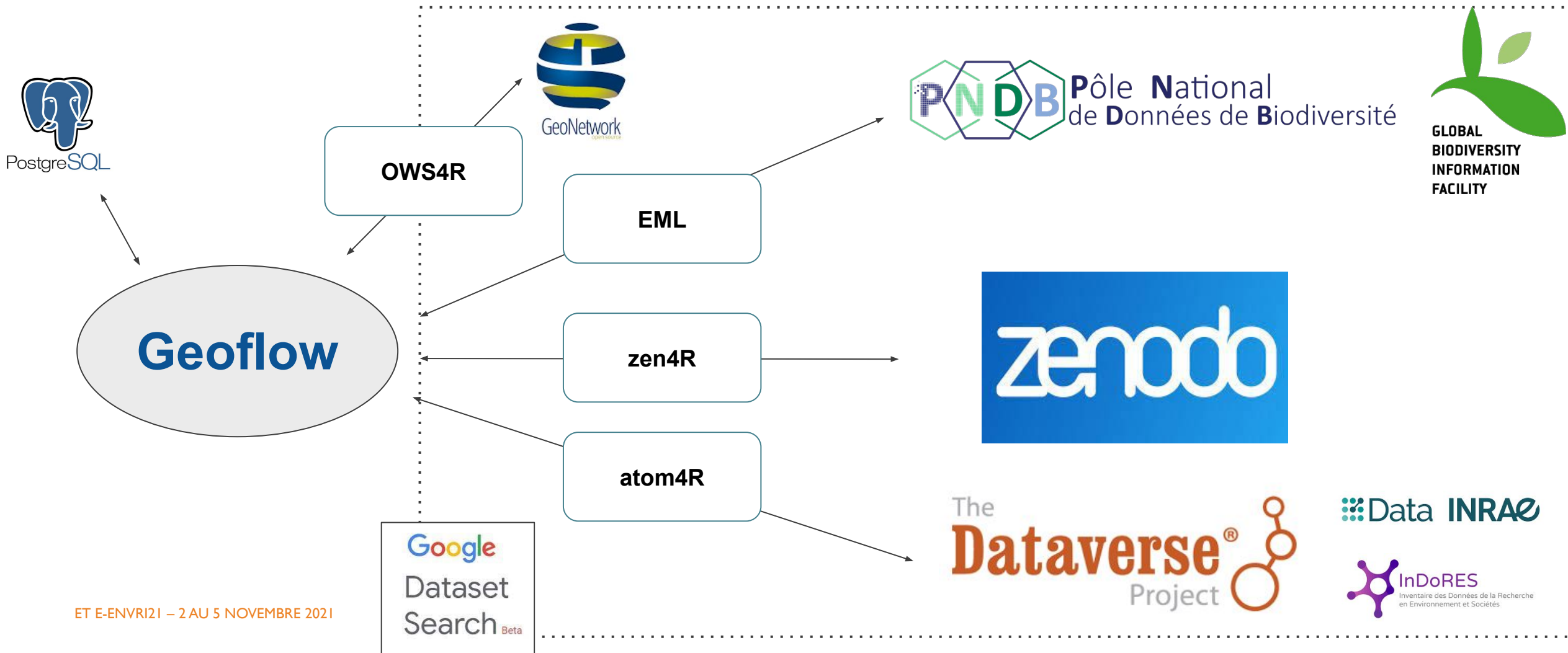
MetaSNAKE



OAI-PMH



INTEROPÉRABILITÉ AVEC DES SYSTÈMES TIERS: GeoFlow



A RETENIR

- **Interopérabilité technique:** à la charge des catalogues, entrepôts et systèmes d'information dans lesquels vous diffusez vos données.

Rend possible le moissonnage par des système tiers

Plus value:

- visibilité accrue de vos jeux de données, citation ()
- pas de duplication des efforts

- **Interopérabilité sémantique:** à la charge du producteur de données en décrivant ses données en utilisant des **vocabulaires contrôlés FAIR**

Plus value:

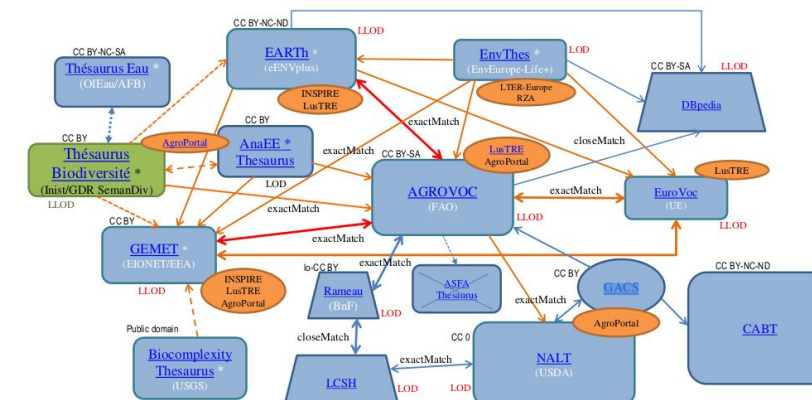
- rend les données réutilisables par des tiers



InDoRES
Inventaire des Données de la Recherche
en Environnement et Sociétés



PANGAEA.





REUTILISATION DES DONNEES



INTRODUCTION

Les données sont dites “Réutilisables” si elles contiennent assez d’information pour pouvoir être comprises et exploitables par des tiers

1. Données suffisamment décrites: quoi ? quand ? où ?
2. Provenance détaillée: comment ? qui ?
3. Conditions d’utilisation: licence de données
4. Données dans des formats standards du domaine et auto-documentés



Données suffisamment décrites: quoi ? quand ? où ?

- **Métadonnées sur les jeux de données**

- Quoi ? : titre, résumé, objectif, mots clés, thématiques
- Quand ? : extension temporelle
- Où ? : extension spatiale

- **Métadonnées sur les observations**

- Quoi ? : variables mesurées + unité
- Quand ? : fuseau horaire des dates UTC
- Où ? : coordonnées des mesures + Système de projection utilisée (code EPSG 4326 = coord. GPS)

Q Retour à la recherche

Télécharger Mode affichage

Les inondations du Rhône en héritage - entretiens auprès des gestionnaires des inondations de 1993 - 94 et 2003

Six entretiens semi-directifs ont été réalisés auprès des principaux gestionnaires locaux présents en 1993 -94 et en 2003 et toujours en fonctions actuellement (le maire d'Arles, le directeur général des services techniques de la mairie d'Arles en charge de la gestion des risques majeurs sur la commune, une représentante de la sous-Préfecture d'Arles, le directeur général adjoint du SYMADREM - syndicat mixte de gestion des digues- et la présidente du Comité de Défense des riverains du Rhône et du Viguerat).

Finalisé

Téléchargements et liens

Document	URL	Ouvrir le lien
Is Rivers 2015 - Présentation de l'action	http://www.graie.org/ISRivers/docs/papers/3C71-49504HON.pdf	Ouvrir le lien
http://ohm-vr.org/fiche_de_synthese/APR_2013/8.Honegger_ficheCOS.pdf	http://ohm-vr.org/fiche_de_synthese/APR_2013/8.Honegger_ficheCOS.pdf	Ouvrir le lien

À propos de cette ressource

Thèmes INSPIRE

Usage des sols

Catégories

LTSER France ZABR

GEMET

- Gestion de crise
- Lutte contre les inondations

GEMET - INSPIRE themes, version 1.0

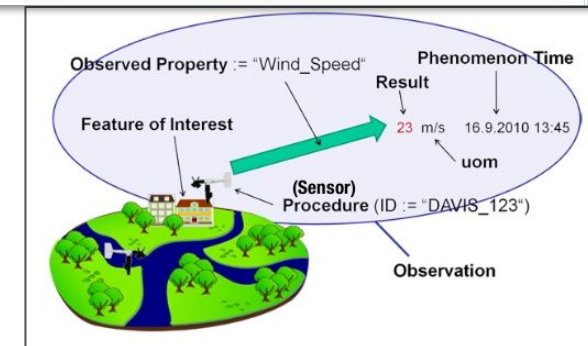
Usage des sols

Langue

Français

Aperçu

Étendue spatiale



Provenance détaillée: comment ? qui ?

Importance de bien documenter les protocoles et les méthodes ayant conduit aux données

- **Métadonnées sur les jeux de données**

- Comment ? : niveau qualité des données (brutes, validées, élaborées), description du protocole, logiciels/codes utilisés
- Qui ? : contacts - différents rôles: PI, data manager, data collector, ...

- **Métadonnées sur les observations**

- Comment ? : instruments ou traitements opérés, incertitude sur les valeurs

Conditions d'utilisation: licenses pour les données

- A quoi ça sert ?
- Licences libres et gratuites pour les données

Conditions d'utilisation: licenses pour les données

Permet de formaliser les conditions de réutilisation et de rediffusion des données

- + Offre un cadre juridique (contrairement aux data policy)

Répondre aux questions:

- Est ce que je veux être cité comme source?
- Est-ce que je veux autoriser les usages commerciaux ou pas ?
- Est-ce que je veux autoriser la création de données dérivées ?

NB1 : la licence ne peut être modifiée que pour aller vers une licence plus ouverte

NB2 : pas de notion de période d'embargo

Conditions d'utilisation: licences pour les données

Licences libres et gratuites pour les données

- **Open Data Base Licence (ODBL)**: licence ouverte (Open Knowledge Foundation). Liberté d'utiliser, de modifier et de redistribuer en gardant la même licence. S'applique aux base de données= structure + données mais pas les données isolément.
- **Etalab v2**: licence ouverte avec attribution. Création dans le cadre de la mission Etalab d'ouverture et de partage des données publiques des administrations françaises + collectivités territoriales (data.gouv.fr).
- **Creative Commons 4.0**: licence ouverte CC0
+ licence ouverte avec attribution (CC-BY) avec des variantes limitant les droits d'utilisation, de modification ou de rediffusion (CC-BY +/- NC, ND et SA). Utilisé largement à l'international

Cadre juridique français

- **loi pour une République numérique, dite Loi Lemaire (octobre 2016)** . Décret d'application avril 2017 (2017-638) : **obligation Etalab et/ou Open Data Base Licence (ODbL), les réutilisations commerciales ne peuvent pas être empêchées**

Conditions d'utilisation: license Creative Common



Creatives Common (autre que CC0): 4 clauses combinables

- **BY - Attribution:** reconnaissance de paternité
- **ND - No derivative works:** restreint la réutilisabilité car oblige de réutiliser dans son intégralité -> déconseillé
- **NC - No commercial :** restreint la réutilisabilité. Pas compatible licence ouverte Etalab v2.0
- **SA - Share Alike :** oblige de rediffuser l'œuvre ou la création dérivée dans les mêmes conditions de licence. Réduit les possibilités de compilations de données car cette clause pourrait empêcher de combiner 2 jeux de données sous des licences différentes -> déconseillé

Exemple du **GBIF**

- CC0: qui rend les données disponibles pour toute utilisation sans restriction
- CC-BY: qui rend les données disponibles pour toute utilisation avec condition d'attribution
- CC-BY-NC: qui rend les données disponibles pour toute utilisation non commerciale avec condition d'attribution

Conditions d'utilisation: phrase de remerciement

Phrase de remerciement

Exemple 1. OHMCV

Data access and use are ruled by the OHMCV data policy.

The following acknowledging sentence should appear in publications using OHMCV data and products:
"OHMCV is funded by the Institut National des Sciences de l'Univers (INSU/CNRS) and the Observatoire des Sciences de l'Univers de Grenoble (OSUG / Université Grenoble Alpes)".

<http://dx.doi.org/10.17178/OHMCV.SMO.CLA.13-14.1>

Conditions d'utilisation: phrase de remerciement

Phrase de remerciement

Exemple 2. AMMA-CATCH

The AMMA-CATCH data is licensed under the [CC-BY 4.0 licence](#) (Creative Commons Attribution 4.0 International License).

For any publication using AMMA-CATCH data, authors are asked to :

- **Attribution:** cite in the article references
Mandatory: the DOI of the AMMA-CATCH observatory. Suggested citation :
AMMA-CATCH (1990): AMMA-CATCH: a hydrological, meteorological and ecological observatory on West Africa. IRD, CNRS-INSU, OSUG, OMP, OREME. [doi:10.17178/AMMA-CATCH.all](https://doi.org/10.17178/AMMA-CATCH.all)
Optional: the DOI of each dataset used
- **Co-authorship:** depending on the contribution of the data to the scientific results obtained, data users should either propose co-authorship to the data providers or at least acknowledge their contribution.

Conditions d'utilisation: phrase de remerciement

Rappel pour les UTILISATEURS

Comment citer des données dans une publication ?

- dans un format proche des formats de citation courants des publications : Creators. Year. Title. Publisher. DOI

Ex: AMMA-CATCH (2002): Meteorological dataset (including radiative budget), within the Donga watershed (600 km²), Benin. IRD, CNRS-INSU, OSUG, OMP, OREME. [doi:10.17178/AMMA-CATCH.AL.Met_Od](https://doi.org/10.17178/AMMA-CATCH.AL.Met_Od)

Où citer un DOI de données dans une publication ?

- Section Acknowledgements (Ex. Science)
- Section dédiée aux données Data Availability (Ex. PLUS ONE)
- Dans les Références

Citation des auteurs

- prénom et nom
- un nom de programme/projet, un nom d'un organisme...dite liste d'une collectivité
- une liste de contributeurs en complément d'une liste d'une collectivité



FORMATS POUR LES DONNÉES

FORMATS POUR LES DONNÉES

Recommandations GT format interpôle IR data Terra

- Formats **ouverts, normalisés & auto-documentés, standards** du domaine.
- Contenant à la fois **données** et **métadonnées**
- Pour format ASCII, choix de l'**encoding de caractères = UTF-8** (et non ISO 8859-1 ou ISO 8859-15)

Les plus:

- Format permettant la parallélisation
- Format permettant la compression

Format recommandés:

- CSV
- NetCDF + convention pour les métadonnées

Format CSV

- Accompagné d'entêtes de métadonnées
- Encoding de caractères UTF-8 OBLIGATOIRE
- Utilisation de caractère standardisé pour:
 - Séparateur: Tab \T
 - Entête commençant par -> #
- NASA AMES (non recommandé)

Format NetCDF et convention de métadonnées

- Adapté pour stocker des **tableaux de nombres multidimensionnels**. Très utilisé en météorologie et en océanographie
- **Portable**
- **Normalisé et auto-documenté**: description via attributs standardisés (noms et valeurs possibles)
- Différentes **conventions de métadonnées**: CF, Coards, ...
- Bibliothèques + logiciels tiers libres et gratuits pour manipuler ce format (créer, lire, modifier, découper, assembler, faire des moyennes, visualiser,...)

<https://www.unidata.ucar.edu/software/netcdf/>

```
netcdf tos_01_2001-2002 {
dimensions:
    lon = 180 ;
    lat = 170 ;
    time = UNLIMITED ; // (24 currently)
    bnds = 2 ;
variables:
    double lon(lon) ;
        lon:standard_name = "longitude" ;
        lon:long_name = "longitude" ;
        lon:units = "degrees_east" ;
        lon:axis = "X" ;
        lon:bounds = "lon_bnds" ;
        lon:original_units = "degrees_east" ;
    double lon_bnds(lon, bnds) ;
    double lat(lat) ;
        lat:standard_name = "latitude" ;
        lat:long_name = "latitude" ;
        lat:units = "degrees_north" ;
        lat:axis = "Y" ;
        lat:bounds = "lat_bnds" ;
        lat:original_units = "degrees_north" ;
    double lat_bnds(lat, bnds) ;
    double time(time) ;
        time:standard_name = "time" ;
        time:long_name = "time" ;
        time:units = "days since 2001-1-1" ;
        time:axis = "T" ;
        time:calendar = "360_day" ;
        time:bounds = "time_bnds" ;
        time:original_units = "seconds since 2001-1-1" ;
    double time_bnds(time, bnds) ;
    float tos(time, lat, lon) ;
        tos:standard_name = "sea_surface_temperature" ;
        tos:long_name = "Sea Surface Temperature" ;
        tos:units = "K" ;
        tos:cell_methods = "time: mean (interval: 30 minutes)" ;
        tos:_FillValue = 1.e+20f ;
        tos:missing_value = 1.e+20f ;
        tos:original_name = "sosstsst" ;
        tos:original_units = "degC" ;
        tos:history = " At 16:37:23 on 01/11/2005: CMOR altered the data in the" ;

// global attributes:
    :title = "IPSL model output prepared for IPCC Fourth Assessment SRES A2 e
    :institution = "IPSL (Institut Pierre Simon Laplace, Paris, France)" ;
    :source = "IPSL-CM4 v1 (2003) : atmosphere : LMDZ (IPSL-CM4_IPCC, 96x71x19
    :contact = "Sebastien Denvil, sebastien.denvil@ipsl.jussieu.fr" ;
    :project_id = "IPCC Fourth Assessment" ;
    :table_id = "Table 01 (13 November 2004)" ;
    :experiment_id = "SRES A2 experiment" ;
    :realization = 1 ;
    :cmor_version = 0.96f ;
    :Conventions = "CF-1.0" ;
    :history = "YYYY/MM/JJ: data generated; YYYY/MM/JJ+1 data transformed At
    :references = "Dufresne et al, Journal of Climate, 2015, vol XX, p 136" ;
    :comment = "Test drive" ;
}
```

Format NetCDF et convention de métadonnées

Convention CF

<https://cfconventions.org/>

Standardisation des métadonnées:

- **liste d'attributs**
- définit des attributs obligatoires pour bien décrire la donnée: `standard_name`, `long_name`, `units`, `missing_value` ou `_FillValue`, `valid_min` et `valid_max`, ...
- **valeurs prises** par certains attributs:
Ex: `standard_name`, `units`,

```
float tos(time, lat, lon) ;  
    tos:standard_name = "sea_surface_temperature" ;  
    tos:units = "K" ;  
    tos:long_name = "Sea Surface Temperature" ;
```

// global attributes:

:Conventions = "CF-1.9" ;

:**title** = A succinct description of what is in the dataset ;

:**institution** = "Specifies where the original data was produced." ;

:**source** = "The method of production of the original data. ;

:**history** = "Provides an audit trail for modifications to the original data" ;

:**references** = "Published or web-based references describing the data" ;

:**comment** = "Miscellaneous information about the data" ;

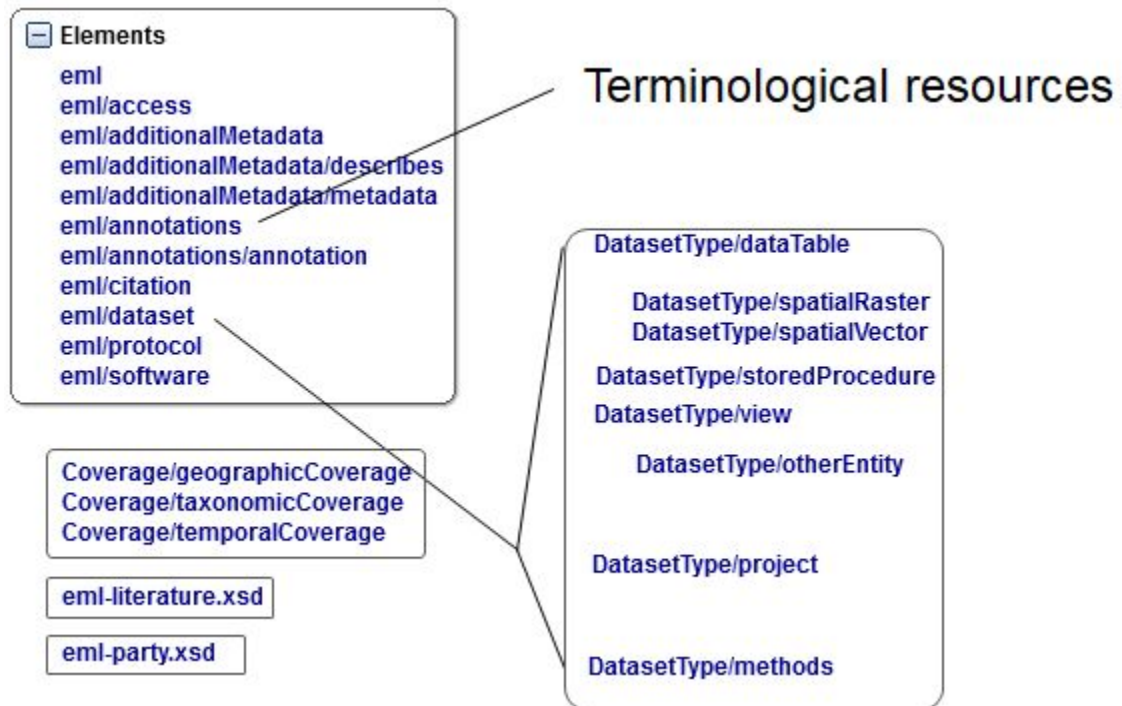
View by Category

Atmospheric Chemistry	Atmosphere Dynamics	Carbon Cycle	Cloud	Hydrology
Ocean Dynamics	Radiation	Sea Ice	Surface	

Standard Name	Canonical Units
age_of_surface_snow	day
atmosphere_mass_content_of_liquid_precipitation	kg m-2
atmosphere_mass_content_of_snow	kg m-2
canopy_snow_amount	kg m-2
canopy_water_amount	kg m-2
change_over_time_in_amount_of_ice_and_snow_on_land	kg m-2
change_over_time_in_canopy_water_amount	kg m-2

Format EML et DarwinCore

EML <https://eml.ecoinformatics.org/>

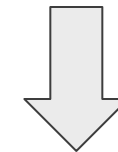
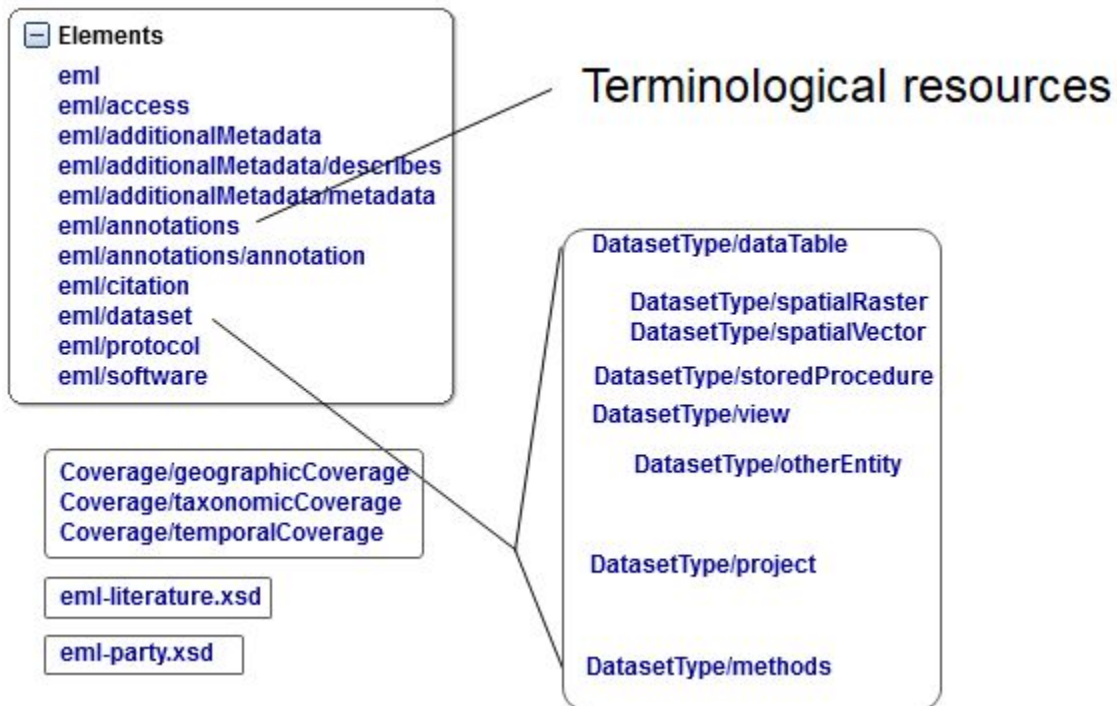


Format EML et Data Package DataOne

EML <https://eml.ecoinformatics.org/>

+

Fichiers de données (et codes sources et protocoles...)



Data Package



Archive Darwin-core et GBIF EML Profile

GBIF EML profile

[https://github.com/gbif/eml-profile/blob/](https://github.com/gbif/eml-profile/blob/bif-profile.xsd)

[bif-profile.xsd](#)

☐ Elements

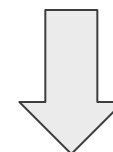
- eml
- eml/access
- eml/additionalMetadata
- eml/additionalMetadata/describes
- eml/additionalMetadata/metadata
- ~~eml/annotations~~
- ~~eml/annotations/annotation~~
- eml/citation
- ~~eml/dataset~~
- eml/protocol
- ~~eml/software~~

Coverage/geographicCoverage
Coverage/taxonomicCoverage
Coverage/temporalCoverage

[eml-party.xsd](#)

+

Fichiers de données (et codes sources et protocoles...)



Data Package



GBIF

Global Biodiversity
Information Facility

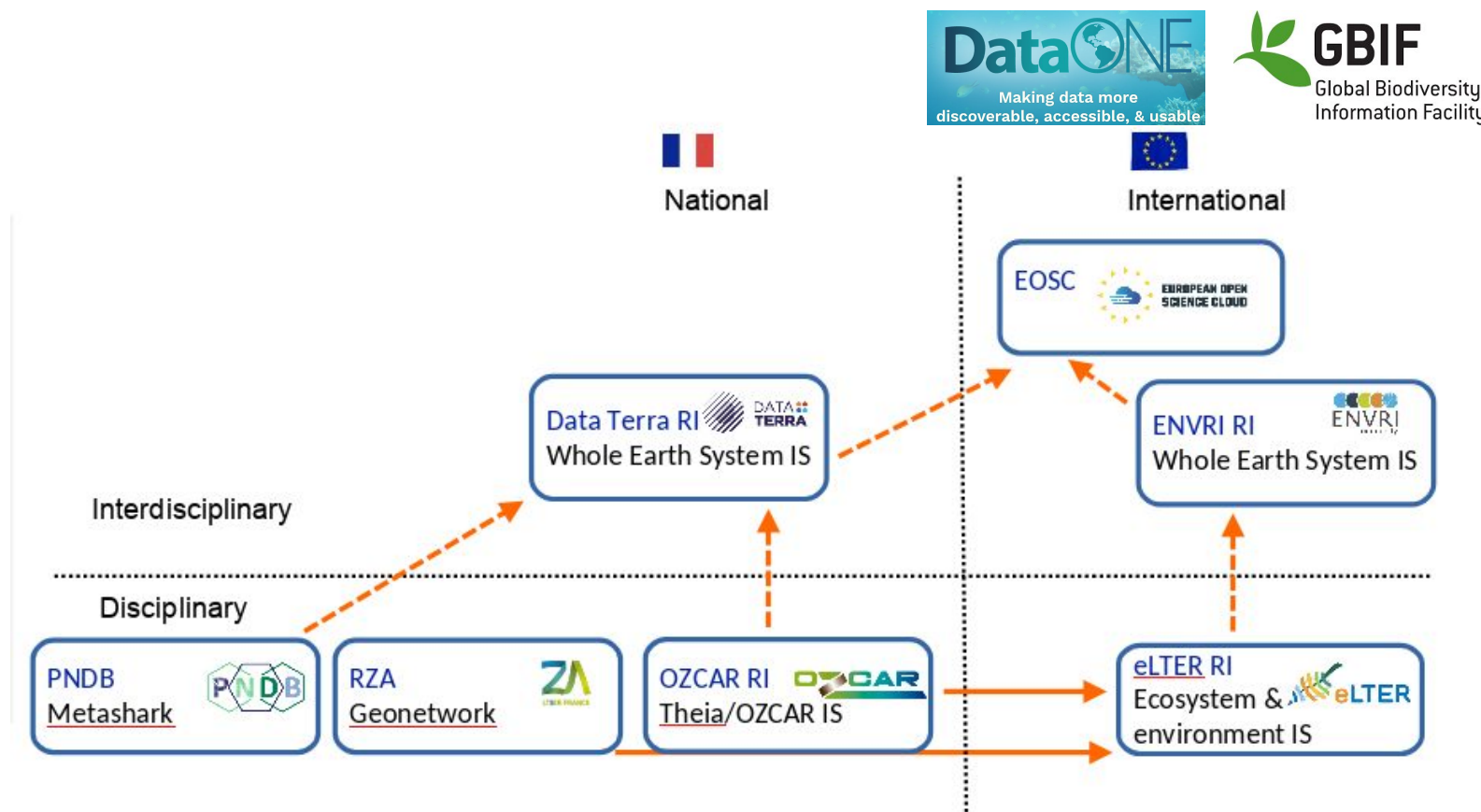


PERSPECTIVES

INTEROPERABILITE DES INFRASTRUCTURES DE DONNEES

Vers une infrastructure virtuelle et fédérée de données de machine à machine :

- Principe de subsidiarité
- Favorise des approches interdisciplinaires



REMERCIEMENT

cnrs

- CNRS soutien financier
 - R. OMNES (conseillère formation, DR10) suivi de la formation
 - P. RUPP (ICUBE), C. DIDIERJEAN (EDYTEM) gestion administrative et financière
 - Le comité organisateur : BED/RZA, OZCAR/THEIA, UMS BBEES, PNDB, avec la participation de la MISHA
 - Centre Saint-Thomas, hébergement
-
- Retrouvez les informations sur : <https://e-envir-21.sciencesconf.org/>



Maison Interuniversitaire des
Sciences de l'Homme
- Alsace -





ANNEXES



Références

Standard de métadonnées:

- E-envir 2019: Les principaux standards de métadonnées à l'échelon international : comment s'y retrouver et comment les exploiter pour la recherche en sciences de l'environnement - Julien BARDE

Vocabulaire contrôlé

- Dominique Vachez. Comparative study of thesauri in Environmental Sciences - Best practices in thesaurus design and FAIRification. 2021.                                                          

Format

- Recommendation GT format interpôle IR Data Terra:
<https://cloud-dataterra.omp.obs-mip.fr/index.php/s/vFHMN8i1F1fsQqp>