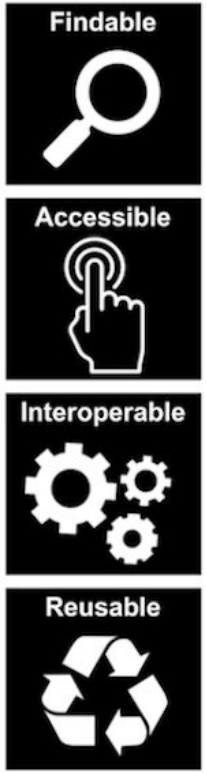
# Utiliser des vocabulaires contrôlés pour FAIRiser les données

## Partie 2 : FAIRiser son vocabulaire

*Victoria Agazzi (CNRS, UAR CPST)*
*Véronique Chaffard (IRD, UMR IGE)*
*Charly Coussot (IRD, OSUG)*
*Jean-Christophe Desconnets (IRD, ESPACE-DEV)*

# Fairiser son vocabulaire

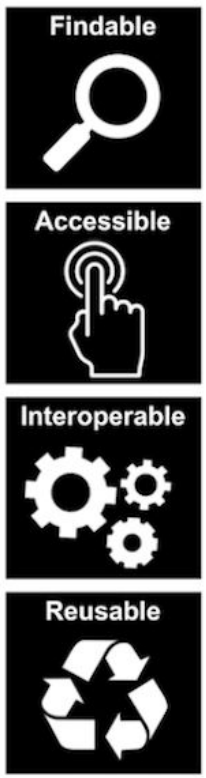



## Les attentes

- Construire et partager les terminologies au delà de son domaine et de ses bases de données
- Fournir une représentation lisible et accessible pour les machines

- Contrôler l’annotation sémantique des données et des métadonnées
  - Interroger les vocabulaires pour faire de l’enrichissement et de l’harmonisation de données
  - Stocker les libellés et les URI pour exploiter les propriétés d’un terme

# Fairiser son vocabulaire




## Les promesses

- Indexation des données sur un vocabulaire contrôlé
- Découverte des données via un système à facettes maîtrisé
- Intégration de données pluri-disciplinaires (climat, océan, atmosphère, surfaces continentales)

# Fairiser son vocabulaire

**“Ten simple rules for making a vocabulary FAIR”**(S. Cox et al., 2020)

- **Rule 1.** Determine the governance arrangements and custodian of the legacy vocabulary
- **Rule 2.** Verify that the legacy-vocabulary license allows repurposing, and agree on the license for the FAIR vocabulary
- **Rule 3.** Check term and definition completeness and consistency in the legacy vocabulary
- **Rule 4.** Establish a traceable maintenance environment for the FAIR vocabulary content
- **Rule 5.** Assign a unique and persistent identifier to (a) the vocabulary and (b) each term in the vocabulary
- **Rule 6.** Create machine readable representations of the vocabulary terms
- **Rule 7.** Add vocabulary metadata
- **Rule 8.** Register the vocabulary
- **Rule 9.** Make the vocabulary accessible for humans and machines
- **Rule 10.** Implement a process for maintaining the FAIR vocabulary

# SKOS: Simple Knowledge Organisation System

Des bases documentaires au web de données

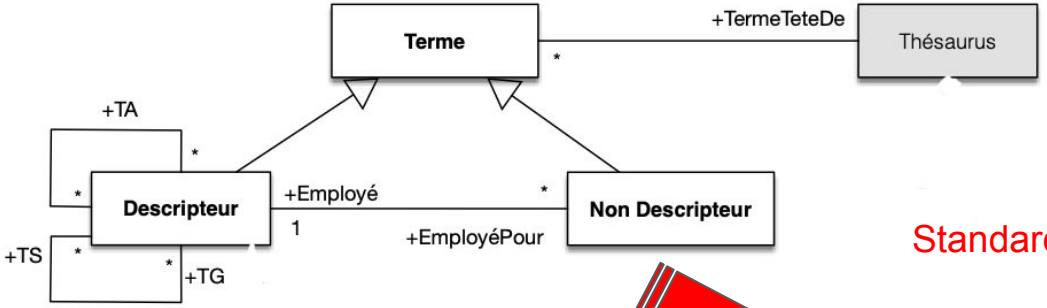



*Diagramme de classes formalisant la notion de Thesaurus*

Standardisation pour le web : SKOS (W3C, 2008)

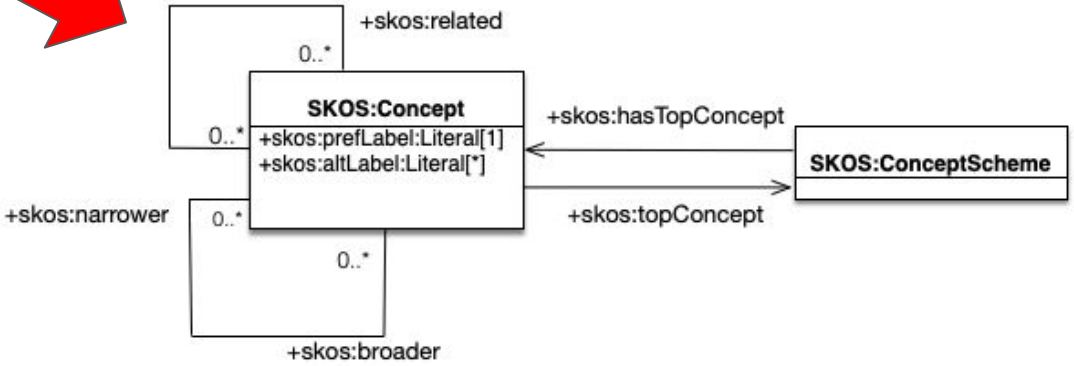



*Diagramme de classes de l'ontologie SKOS Core*

# SKOS : Simple Knowledge Organisation System

**Pour chaque terme:**

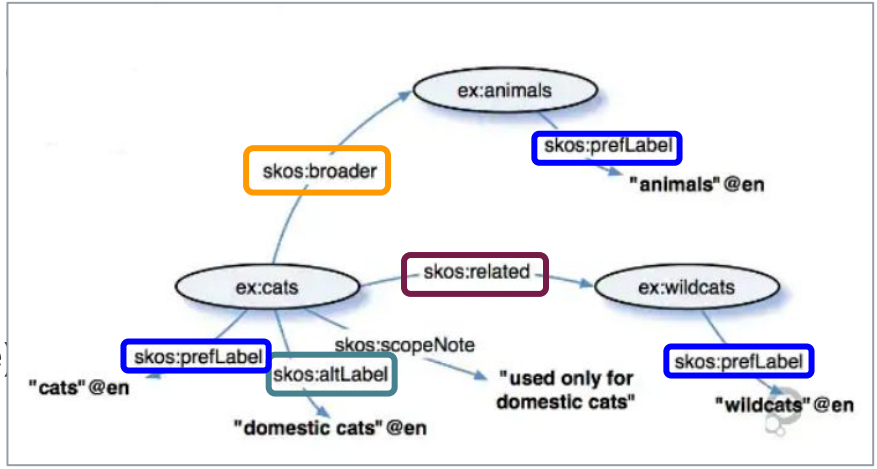
- Un **identifiant** pérenne = URI
  Ex: https://dbpedia.org/resource/Strasbourg et https://sws.geonames.org/2973783 représente la ville de Strasbourg

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

**Entre termes:**

- **Relations** sémantiques **hiérarchiques** :
  - estPlusGénérique: skos:broader
  - estPlusSpécifique: skos:narrower

- **Relations** sémantiques **associatives** (similarité)
  - estSimilaireA: skos:exactMatch,
  - EstProcheDe: skos:closeMatch,
  - estAssociéA: skos:related

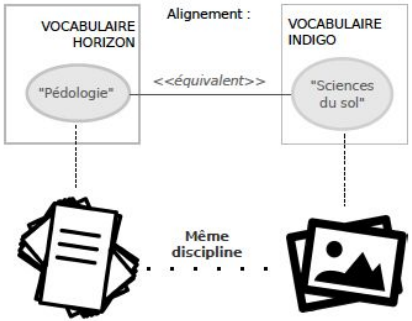


# Fairiser son vocabulaire

Imposer un vocabulaire standard existant est difficilement envisageable

Chaque discipline doit adapter son vocabulaire à ses besoins

- ~~Solution 1 : imposer un vocabulaire commun~~
- Solution 2 : aligner les vocabulaires



# Fairiser son vocabulaire

**en s’appuyant sur les principes suivants**

- **Réutiliser** l’existant: **les thésaurus de référence sont une brique de base pour la construction de vos thésaurus**
- **Enrichir** pour adapter à ses besoins
- **Publier** ses propres vocabulaires dans le respect des principes FAIR
- **Mettre en correspondance** (si possible) avec d’autres vocabulaires

# Comment fairiser son vocabulaire ?

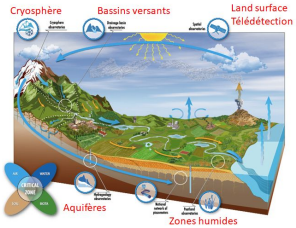
1. Identifier les **thésaurus dit de “référence” et les concepts à réutiliser / compléter**

2. **Organiser** les concepts afin d’avoir une organisation **adaptée à sa communauté**

Exemples: Theia/OZCAR, Data Terra

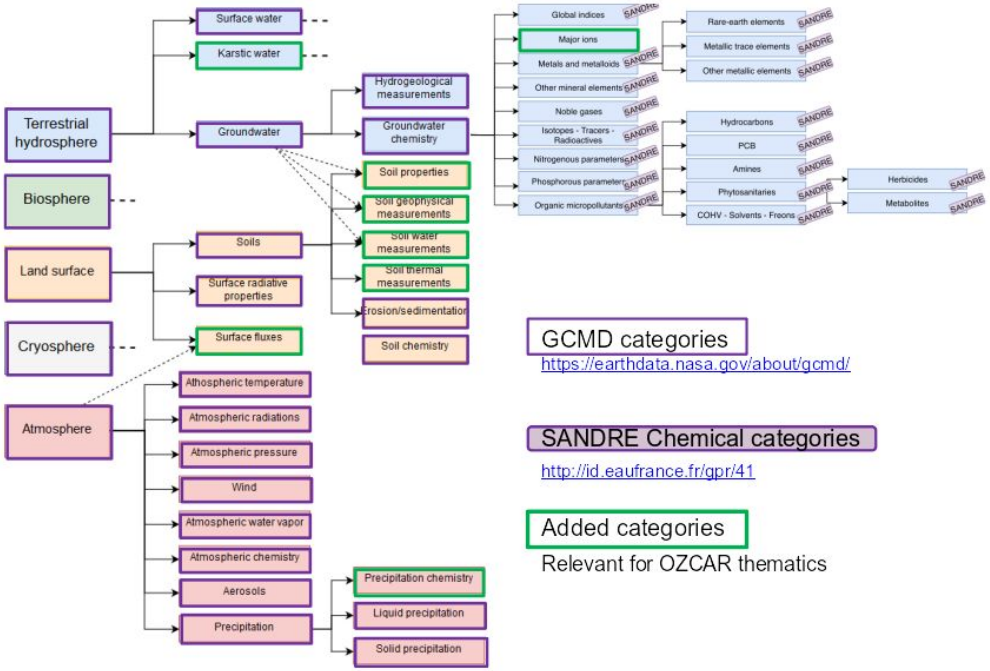
# Thésaurus Theia/OZCAR

## Thésaurus pour les données in-situ de la zone critique et des sciences de l'environnement




- Vocabulaire commun à l'ensemble des observatoires de l'IR OZCAR / pôle Theia in-situ.
- Basé sur NASA GCMD Earth Science

  + SANDRE pour la chimie des eaux




Des catégories ont été rajoutées pour le besoin de la communauté IR OZCAR

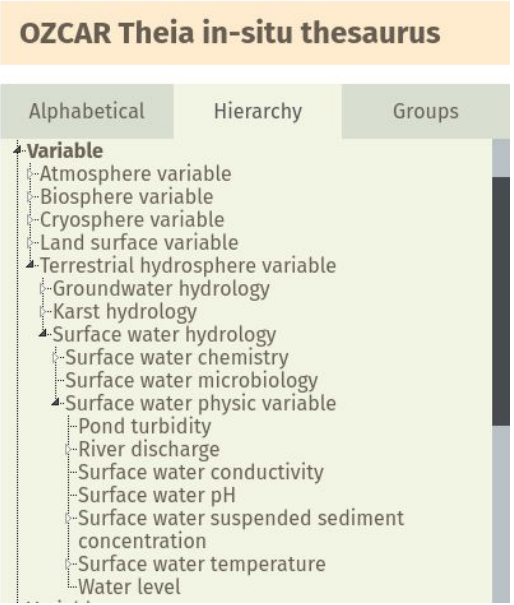
# Thésaurus Theia/OZCAR

## Thésaurus pour les données in-situ de la zone critique et des sciences de l'environnement

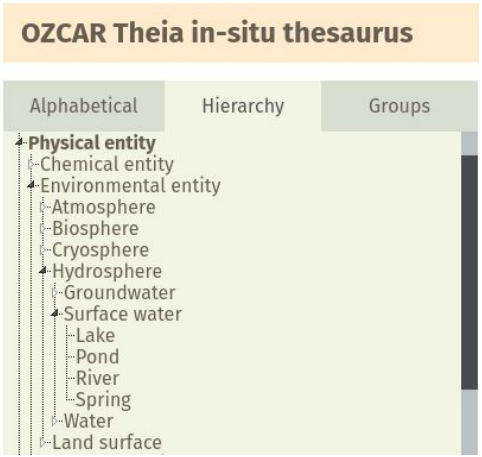
- Organisé en 1 seul thésaurus contenant:
  - variable
  - objet d'intérêt

- Anglais

Variable



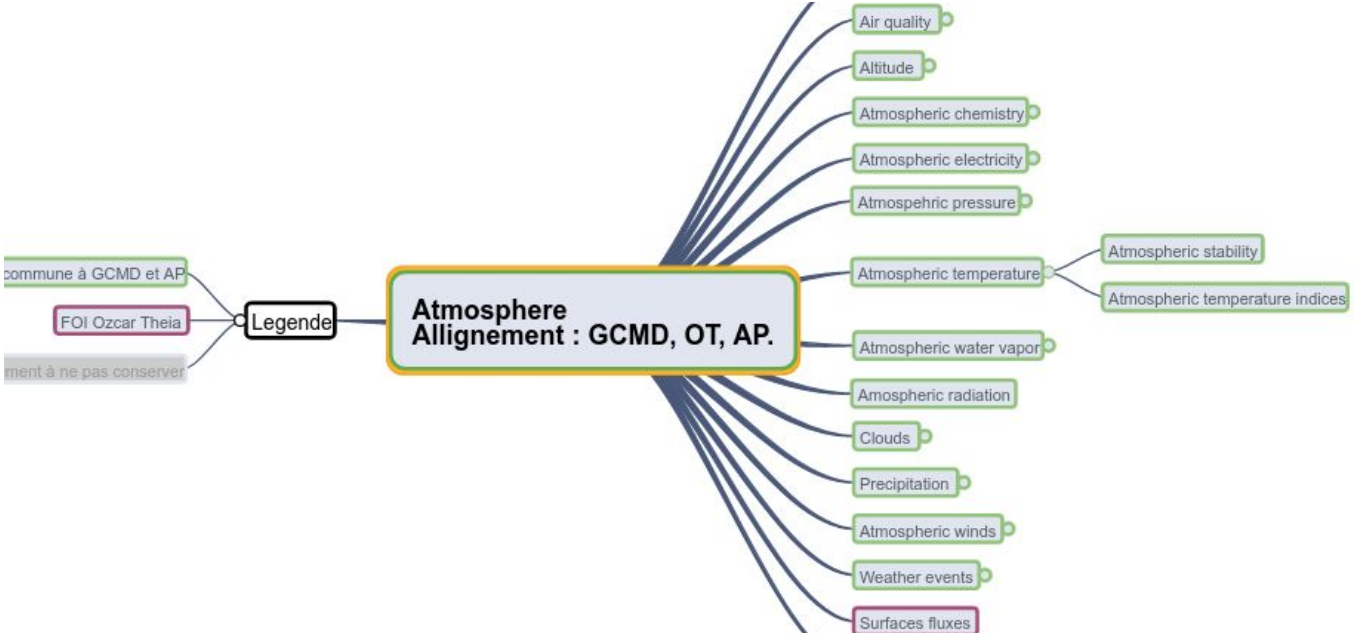

Objet d'intérêt



# Thésaurus Data Terra

## Thésaurus pour les données multi source et multi thématiques du système Terre

- Vocabulaire commun à l'ensemble des 4 pôles
- Basé sur NASA GCMD + thésaurus des pôles



# Thésaurus Data Terra

## Thésaurus pour les données multi source et multi thématiques du système Terre

- Organisé en 4 thésaurus:
  - Type d'Objet d'intérêt: Earth Science Feature
  - Variables: Earth Science Property
  - Type de plateforme: Earth Science Platform
  - Type d'instrument: Earth Science Instrument

- Anglais

**Earth Science Feature**

Alphabetical | Hierarchy | Groups

- Atmosphere
- Biosphere
- Cryosphere
- Land surface
  - Crops
  - Land cover
  - Landscape
  - Rock
  - Sediment
  - Soil
  - Topography
  - Urban areas
- Oceans
- Solid Earth
- Terrestrial hydrosphere

**Earth Science Platform**

Alphabetical | Hierarchy | Groups

- Earth Observation Airbornes
- Earth Observation Satellites
  - ADEOS (Advanced Earth Observing Satellite)
  - ALOS
  - ALOS-2
  - Aqua
  - Aura
  - COSMO-SKYMED
  - ENVISAT
  - FY (Feng-Yun)
  - GCOM-W1
  - GOES (Geostationary Operational Environmental Satellite)
  - Green-house gas Observing Satellite
  - Green-house gas Observing Satellite - 2
  - Himawari
  - JASON-2
  - JASON-3
  - LANDSAT-8
  - METEOSAT
  - METOP
  - MT1
  - NASA Earth System Science Pathfinder
  - PARASOL
  - PLEIADES
  - RADARSAT
  - RAPIDEYE
  - SENTINEL-1
  - SENTINEL-2
  - SENTINEL-3
  - SENTINEL-5
  - SPOT
  - Terra
  - TerraSAR-X
  - VENuS
- In Situ Land-based
  - Accelerometric station
  - Air monitoring Stations
  - Fixed Observation Stations
  - Geodetic station
  - Gravimetric station
  - Ground Stations
  - Meteorological Stations
  - Mobile stations
  - Seismological station
  - Weather Stations
- In Situ Ocean-based
- Models
- Navigation systems

**Earth Science Property**

Alphabetical | Hierarchy | Groups

- **Essential Ocean Variables**
  - **Biochemical ocean variables**
    - Alkalinity
    - Colored dissolved organic matter (CDOM)
    - Dissolved oxygen
    - Water turbidity
  - **Biological and Ecosystems ocean variables**
    - Zooplankton abundance
    - Zooplankton biomass
    - Zooplankton size
  - **Cross-disciplinary ocean variables**
    - Acoustic velocity
  - **Physical ocean variables**
    - Air temperature
    - Conductivity
    - Heat flux
    - Humidity
    - Ice depth distribution
    - Ocean heat content
    - Radiative flux
    - Salinity
    - Sea surface height (SSH)
    - Sea surface temperature
    - Speed profile
    - Surface water temperature
    - Water depth
    - Water pressure
    - Wave direction
    - Wave height
    - Wave period
    - Wave speed

**Earth Science Instrument**

Alphabetical | Hierarchy | Groups

- Altimeters
- Chemical Meters/Analyzers
- Current/Wind Meters
- Gauges
- Imaging Radars
- Magnetic/Motion Sensors
- Photon/Optical Detectors
- Positioning/Navigation
- Pressure/Height Meters
- Probes
- Profilers/Sounders
- Recorders/Loggers
- Samplers
- Spectrometers/Radiometers
- Tachometers
- Temperature/Humidity sensors
- Thermal/Radiation detectors

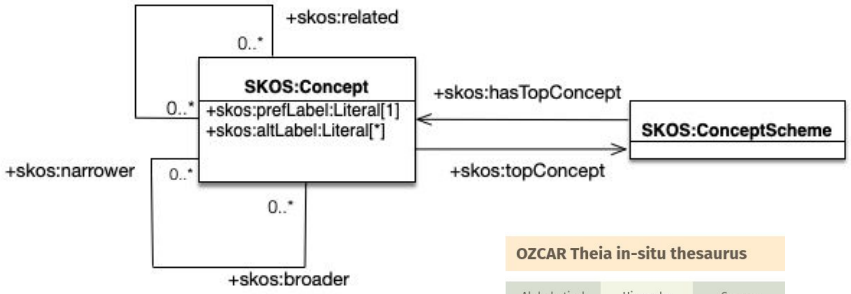
# Comment fairiser son vocabulaire ?

**3. Décrire son vocabulaire en SKOS**

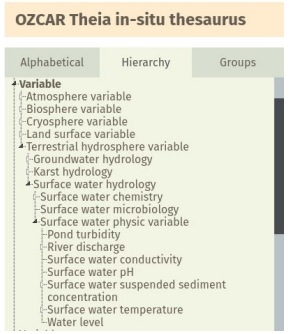
4. Formaliser les correspondances avec les thésaurus identifiés (alignements)

Exemples: Theia/OZCAR

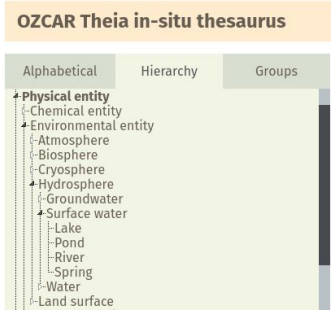
# Thésaurus Theia/OZCAR




Top concept: Variable




Top concept: Objet d'intérêt




URI: https://w3id.org/ozcar-theia

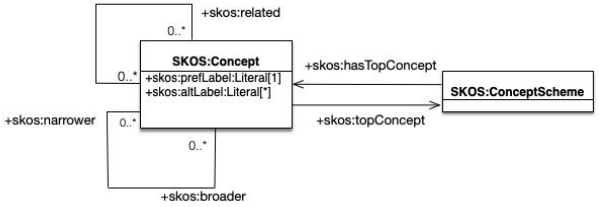
**:ozcarTheiaThesaurus** a **skos:ConceptScheme**;

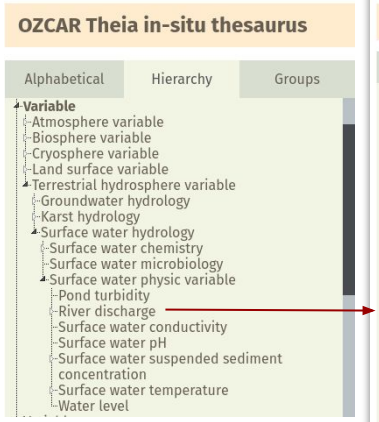
**skos:hasTopConcept** :c_024afb0a, :c_06813aff, **:c_25ff936c**, **:c_341ce66a**, :c_55ebb111,:c_b456a9b4, :c_c23c9bb6, :c_d0bd8b1c, :c_d4be84c4 .

**dct:title** "***OZCAR Theia in-situ thesaurus"@en***;
**dct:created** "2018-01-01"^^xsd:date;
**dct:creator** "Charly Coussot https://orcid.org/0000-0002-0544-4802", "Isabelle Braud https://orcid.org/0000-0001-9155-0056",
"Sylvie Galle https://orcid.org/0000-0002-3100-8510", "Véronique Chaffard https://orcid.org/0000-0003-2823-7117";
**dct:description** "Thesaurus of the Theia in-situ information system : https://in-situ.theia-land.fr/"@en;
**dct:license** <http://creativecommons.org/licenses/by/4.0/>;
…

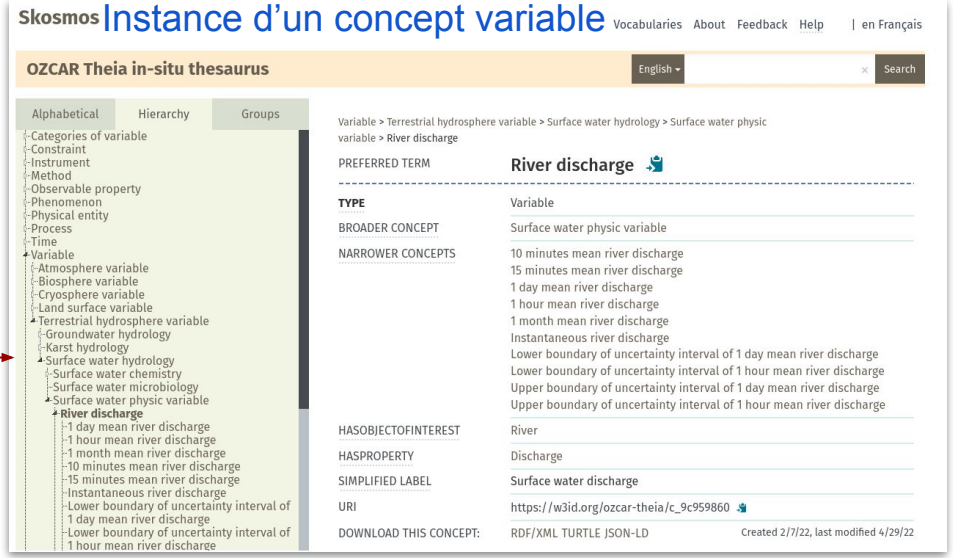
# Thésaurus Theia/OZCAR

## Top concept: Variable




## Instance d'un concept variable




**URI**: https://w3id.org/ozcar-theia/**c_341ce66a**

**URI**: https://w3id.org/ozcar-theia/**c_9c959860**

**:c_341ce66a** a **skos:Concept;**

**skos:inScheme** :ozcarTheiaThesaurus;
**skos:narrower** :c_0d8719e8, :c_8b34f633, :c_d9824eec, :c_db51a9f2, :c_fb3683f0;

**skos:prefLabel** "***Variable"@en***;

**skos:topConceptOf** :ozcarTheiaThesaurus .

**:c_9c959860** a **skos:Concept**;

**skos:inScheme** :ozcarTheiaThesaurus;
**skos:narrower** :c_27bfea84, :c_29b79f1a, :c_7aece067, :c_9391db4b, :c_963cc0b5, :c_9c9f688b, c_d76a5276, :c_da4a2284, :c_dabd2d39;
skos:broader :c_b23989c9;

**skos:prefLabel** "***River discharge"@en;***

# Thésaurus Theia/OZCAR

## Variable : instance

Skosmos

Vocabularies About Feedback Help | en Français

OZCAR Theia in-situ thesaurus

English Search

Alphabetical Hierarchy Groups

- Categories of variable
- Constraint
- Instrument
- Method
- Observable property
- Phenomenon
- Physical entity
- Process
- Time
- Variable
  - Atmosphere variable
  - Biosphere variable
  - Cryosphere variable
  - Land surface variable
  - Terrestrial hydrosphere variable
    - Groundwater hydrology
    - Karst hydrology
    - Surface water hydrology
      - Surface water chemistry
      - Surface water microbiology
      - Surface water physic variable
        - **River discharge**
          - 1 day mean river discharge
          - 1 hour mean river discharge
          - 1 month mean river discharge
          - 10 minutes mean river discharge
          - 15 minutes mean river discharge
          - Instantaneous river discharge
          - Lower boundary of uncertainty interval of 1 day mean river discharge
          - Lower boundary of uncertainty interval of 1 hour mean river discharge

Variable > Terrestrial hydrosphere variable > Surface water hydrology > Surface water physic variable > River discharge

| | |
|---|---|
| PREFERRED TERM | **River discharge** |
| TYPE | Variable |
| BROADER CONCEPT | Surface water physic variable |
| NARROWER CONCEPTS | 10 minutes mean river discharge<br>15 minutes mean river discharge<br>1 day mean river discharge<br>1 hour mean river discharge<br>1 month mean river discharge<br>Instantaneous river discharge<br>Lower boundary of uncertainty interval of 1 day mean river discharge<br>Lower boundary of uncertainty interval of 1 hour mean river discharge<br>Upper boundary of uncertainty interval of 1 day mean river discharge<br>Upper boundary of uncertainty interval of 1 hour mean river discharge |
| HASOBJECTOFINTEREST | River |
| HASPROPERTY | Discharge |
| SIMPLIFIED LABEL | Surface water discharge |
| URI | https://w3id.org/ozcar-theia/c_9c959860 |
| DOWNLOAD THIS CONCEPT: | RDF/XML TURTLE JSON-LD |

Created 2/7/22, last modified 4/29/22

## Catégorie de variable: instance

Alphabetical Hierarchy Groups

- Turbidity
- Volume
- Volumetric mass density
- Water holding capacity
- Water potential
- Water retention curve
- Phenomenon
- Physical entity
- Process
- Time
- Variable
  - Atmosphere variable
  - Biosphere variable
  - Cryosphere variable
  - Land surface variable
  - Terrestrial hydrosphere variable
    - Groundwater hydrology
    - Karst hydrology
    - Surface water hydrology
      - Surface water chemistry
      - Surface water microbiology
      - **Surface water physic variable**
        - River discharge
        - Surface water conductivity
        - Surface water level
        - Surface water pH
        - Surface water suspended sediment concentration
        - Surface water temperature
        - Surface water turbidity
          - Pond turbidity

Variable > Terrestrial hydrosphere variable > Surface water hydrology > Surface water physic variable

| | |
|---|---|
| PREFERRED TERM | **Surface water physic variable** |
| TYPE | Category of variable |
| BROADER CONCEPT | Surface water hydrology |
| NARROWER CONCEPTS | River discharge<br>Surface water conductivity<br>Surface water level<br>Surface water pH<br>Surface water suspended sediment concentration<br>Surface water temperature<br>Surface water turbidity |
| SIMPLIFIED LABEL | Surface water physic |
| URI | https://w3id.org/ozcar-theia/c_b23989c9 |
| DOWNLOAD THIS CONCEPT: | RDF/XML TURTLE JSON-LD |

Created 2/7/22, last modified 4/29/22

**URI**: https://w3id.org/ozcar-theia/**c_9c959860**

```
:c_9c959860 a skos:Concept;

  skos:broader :c_b23989c9;
  skos:inScheme :ozcarTheiaThesaurus;
  skos:narrower :c_27bfea84, :c_29b79f1a, :c_7aece067, :c_9391db4b,
:c_963cc0b5, :c_9c9f688b, c_d76a5276, :c_da4a2284, :c_dabd2d39;
  skos:prefLabel "River discharge"@en;
```

**URI**: https://w3id.org/ozcar-theia/**c_b23989c9**

```
:c_b23989c9 a skos:Concept;

  skos:broader :c_af623670;
  skos:inScheme :ozcarTheiaThesaurus;
  skos:narrower :c_1862b46d, :c_6ea2dbd7, :c_8b23ac2b,
:c_9c959860, :c_a6bdee26, :c_c04c5c87, :c_c159113b;
  skos:prefLabel "Surface water physic variable"@en;
```

# Comment fairiser son vocabulaire en pratique ?

SKOS Play!: permet de produire un fichier skos rdf à partir d'un fichier Excel "template"

| | | | |
|---|---|---|---|
| ConceptScheme URI | http://ontology.deic.org/cv/vocab-name/ | | General setup |
| PREFIX | vocab-name | http://ontology.deic.org/cv/vocab-name/ | |
| PREFIX | pav | http://purl.org/pav/ | |
| PREFIX | dct | http://purl.org/dc/terms/ | |
| PREFIX | owl | http://www.w3.org/2002/07/owl# | |
| PREFIX | xsd | http://www.w3.org/2001/XMLSchema# | |
| PREFIX | skos | http://www.w3.org/2004/02/skos/core# | |
| dct:title | | | Controlled vocabulary metadata |
| dct:description | | | |
| dct:creator | | | |
| dct:rights | | | |
| pav:version | | | |
| pav:createdOn | | | |
| pav:lastUpdatedOn | | | |
| Identifier | skos:prefLabel@en | skos:altLabel(separator=",") | Definition of terms (and optionally properties) |
| vocab-name: | | | |
| vocab-name: | | | |
| vocab-name: | | | |
| vocab-name: | | | |
| vocab-name: | | | |
| vocab-name: | | | |
| vocab-name: | | | |
| vocab-name: | | | |
| vocab-name: | | | |

*Nikola Vasiljevic, & John Graybeal. (2021, April 22). FAIR DATA COLLECTIVE: FAIR Made Easy. Zenodo. https://doi.org/10.5281/zenodo.4705970*

# Comment fairiser son vocabulaire en pratique ?

SKOS Play!: permet de produire un fichier skos à partir d'un fichier Excel "template"

- Relations SKOS hiérarchiques et associatives
- Collections SKOS : regroupements utiles dans les cas où plusieurs concepts partagent un aspect commun
- Gestion des concepts dépréciés
- Possibilité d'ajouter des relations différentes de celles définies par SKOS et basés sur d'autres ontologies
- …

Quelques exemples de templates dans l'outil SKOS play!

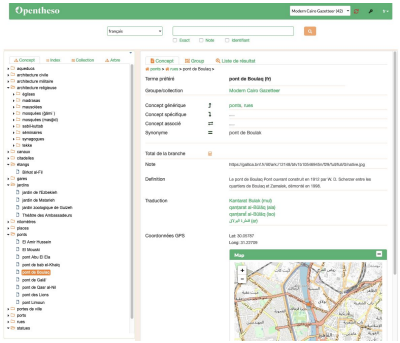
**Avantage**: outil permettant de commencer un thésaurus assez rapidement et offrant de nombreuses possibilités.
**Inconvénient: gestion du thésaurus très compliquée** lorsque le vocabulaire grossit et lorsque l'on complexifie les relations entre concepts!!

# Comment fairiser son vocabulaire en pratique ?

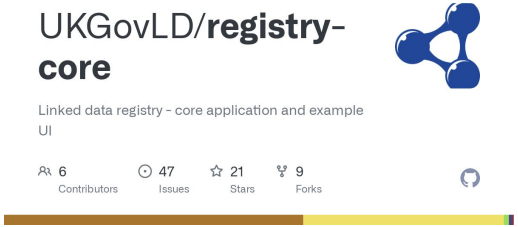
**Des éditeurs plus avancés:**

Gestion facilitée, certaines opérations automatisées (exemple de la date mise à jour renseignée automatiquement), historique, collaboration…

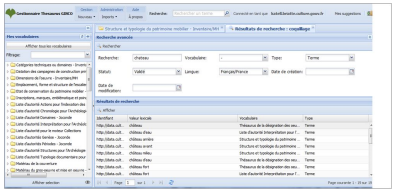
Courbe d’apprentissage plus importante, déploiement de l’outil



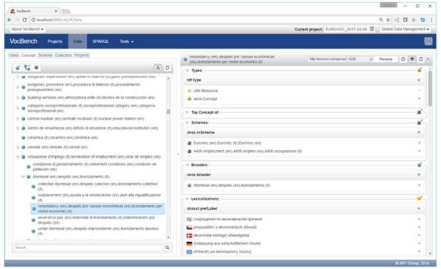

OpenTheso




UKGovLD Registry



Ginco



Vocbench

# Comment fairiser son vocabulaire en pratique ?

## Choix de l'uri du thésaurus:

- Importance de choisir **une URI persistante dans le temps!** Si le nom de domaine ne nous appartient pas, les services purl ou w3id qui permettent d'attribuer des uri persistantes.
- **Choisir le format de l'URI:** URI de base + pattern pouvant être généré automatiquement par les outils d'éditions (bonne pratique !)

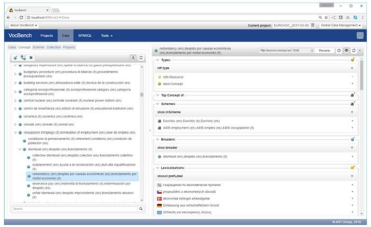
Exemple : https://w3id.org/ozcar-theia/c_d9824eec
URI de base du thesaurus + ID du concept

- Le format de l'URI peut dépendre de la plateforme de publication du thésaurus visée!!

  Exemple des URI Loterre avec l'URI de base Loterre et un identifiant ARK :
  http://data.loterre.fr/ark:/67375/37T-C767L3Z0-Q

# Comment fairiser son vocabulaire en pratique ?

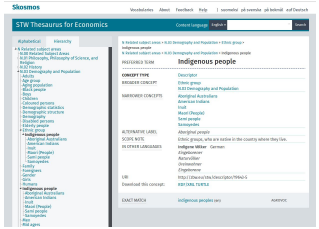
**Déployer sa propre infrastructure pour gérer, stocker, diffuser son vocabulaire** : exemple du thésaurus Theia/OZCAR



RDF4J compliant API →

rdf4j

← SPARQL



← HTTP

Editeur : Vocbench

Triple store : RDF4J

Visualisateur : Skosmos

↑ HTTP 301 Redirection

ID persistent + Reverse proxy

← HTTP

# Comment fairiser son vocabulaire en pratique ?

**Plateformes de publication institutionnelles pour accueillir un thésaurus**

LOTERRE

V.O.INRAE VOCABULAIRES OUVERTS

IRD

**IRD Référentiels**

EcoPortal

AgroPortal

BioPortal

**Recommandation** : favoriser l'hébergement d'un thésaurus au sein d'une plateforme de publication institutionnelle, disciplinaire ou thématique

# Comment fairiser son vocabulaire ?

3. Décrire son vocabulaire en SKOS

**4. Formaliser les correspondances avec les thésaurus identifiés (alignements)**

Exemples: Theia/OZCAR

# Alignement avec d’autres thésaurus

*Objectifs et intérêts:*

- *Faire des passerelles entre les différents silos de données dont le contenu peut être décrit de manière différente (usage de terminologie différente) et/ou mal documenté (sans définition explicite)*
- *Réutilisation et interopérabilité accrue des données*

*Concepts:*

- *Interopérabilité sémantique des données*

*Mise en oeuvre:*

- *Description formelle de liens sémantiques vers des termes exacts ou similaires d’autres thésaurus à l’aide des propriétés SKOS (skos:exactMatch, skos:closeMatch, skos:relatedMatch)*

# Propriétés SKOS pour décrire les alignements

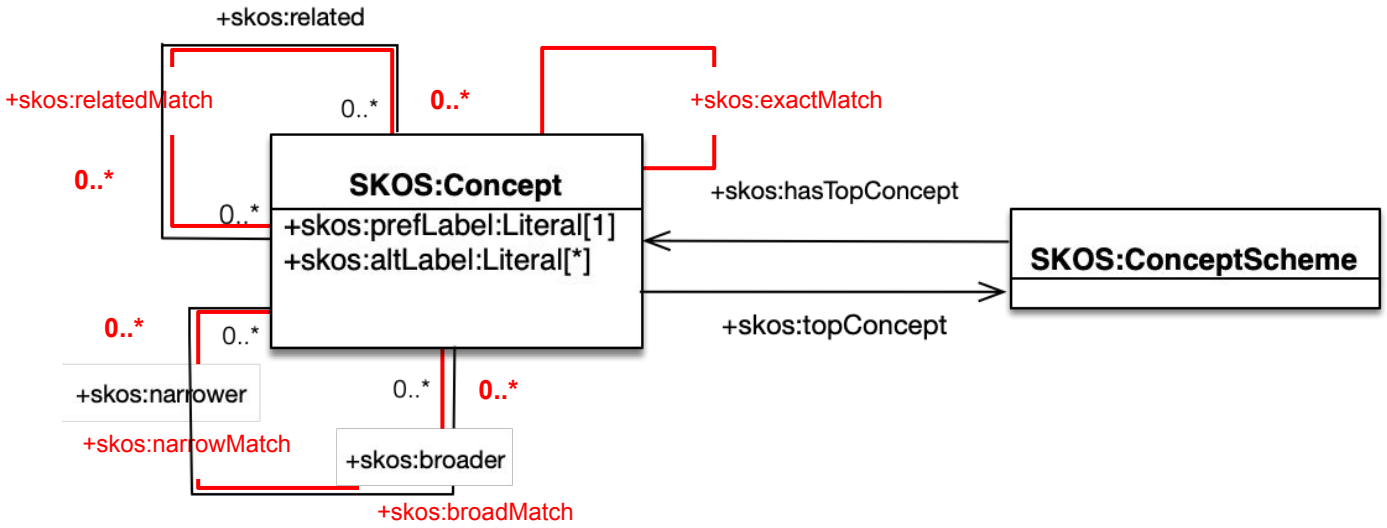
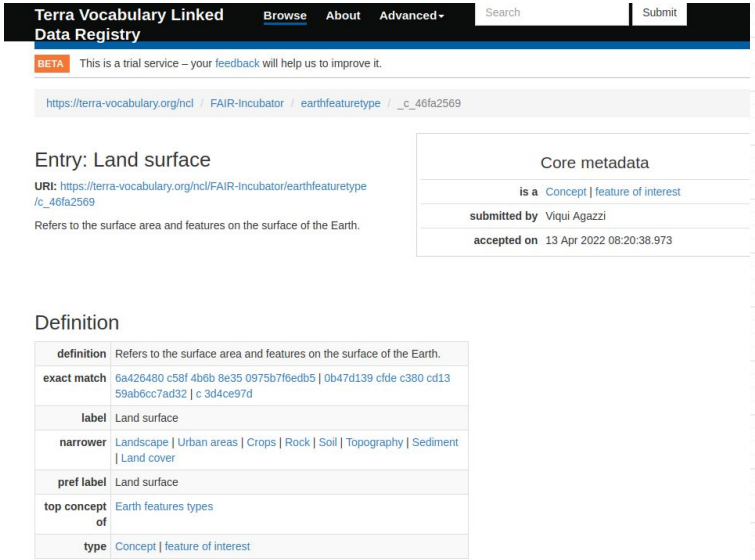
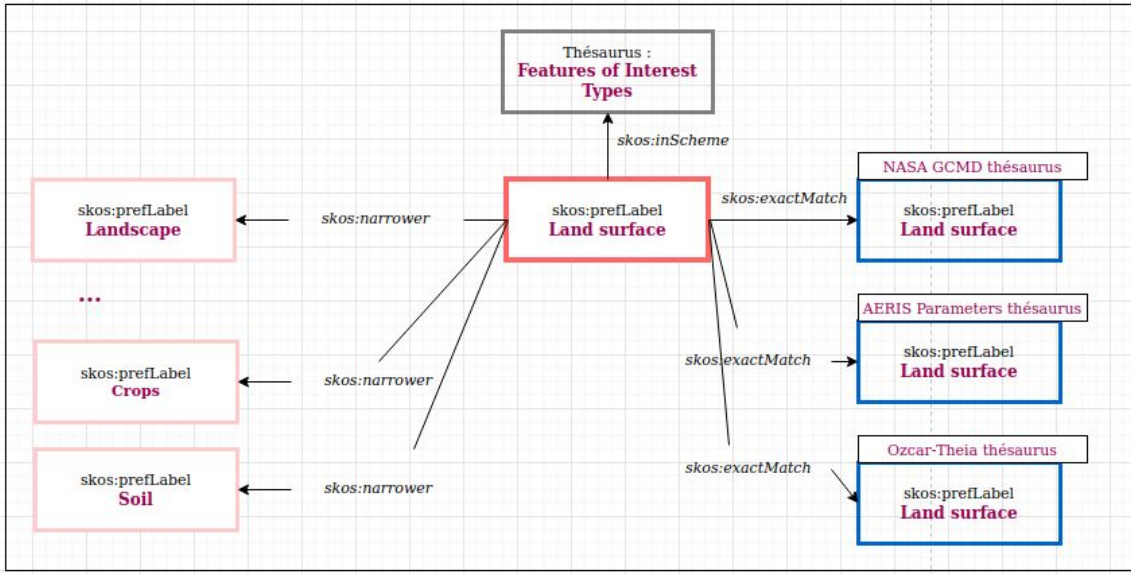



*Diagramme de classes de l'ontologie SKOS*

# Exemples d’alignements entre thésaurus

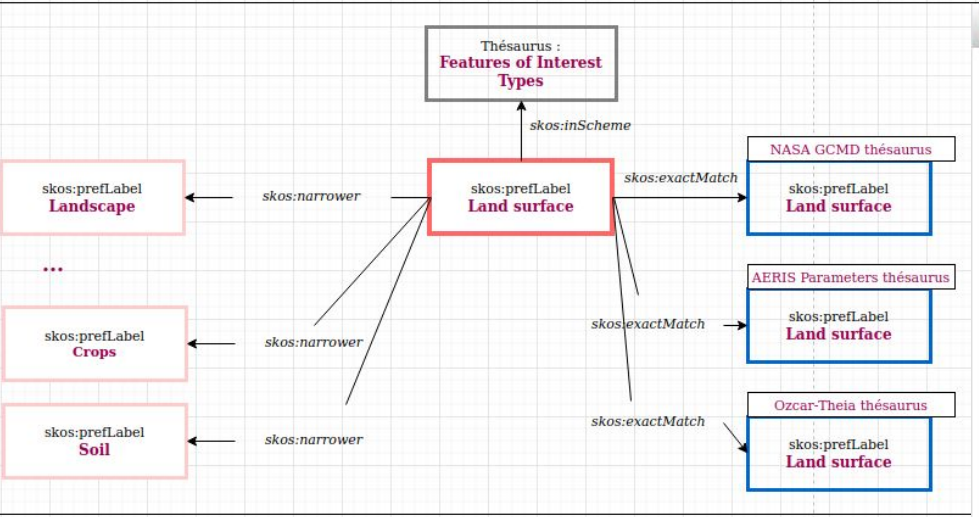
**Thésaurus Data Terra :** Alignements avec les thésaurus : GCMD + pôles



# Exemples d’alignements entre thésaurus

**Thésaurus Data Terra :** Alignements avec les thésaurus : GCMD + pôles




Land surface (en)

```
1  @prefix : <https://terra-vocabulary.org/ncl/EarthFeatureType/> .
2  @prefix grddl: <http://www.w3.org/2003/g/data-view#> .
3  @prefix dct: <http://purl.org/dc/terms/> .
4  @prefix owl: <http://www.w3.org/2002/07/owl#> .
5  @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
6  @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
7  @prefix skos: <http://www.w3.org/2004/02/skos/core#> .
8  @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
9  @prefix dc: <http://purl.org/dc/elements/1.1/> .
10
11 :c_46fa2569 a skos:Concept, <http://www.w3.org/ns/sosa/FeatureOfInterest>;
12   skos:prefLabel "Land surface"@en;
13   skos:exactMatch <https://w3id.org/ozcar-theia/c_3d4ce97d>,
14   <https://vocab.aeris-data.fr/parameter/0b47d139-cfde-c380-cd13-59ab6cc7ad32>,
15     <https://gcmd.earthdata.nasa.gov/kms/concept/6a426480-c58f-4b6b-8e35-0975b7f6edb5>;
16   skos:definition "Refers to the surface area and features on the surface of the Earth."@en;
17   skos:narrower :c_b98c746c, :c_582c60b7, :c_72e9e9bf, :c_5caafe58, :c_9e04ad11, :c_57139860,
18     :c_83cfca32, :c_bdf17c1e;
19   skos:inScheme <https://terra-vocabulary.org/ncl/earthfeaturetype>;
20   skos:topConceptOf <https://terra-vocabulary.org/ncl/earthfeaturetype> .
21
```

# En pratique: alignement avec d’autres thésaurus

**L’alignement avec un autre thésaurus en 4 étapes**

***1 -*** *Identifier les vocabulaires avec lesquels aligner son thésaurus*

***2 -*** *Réaliser les alignements (semi-automatiquement)*

***3 -*** *Les valider (manuellement) avec un spécialiste du domaine*

***4 -*** *Appliquer les alignements à son thésaurus*

# En pratique: Réaliser les alignements

Utilisation de onaGUI pour aligner les concepts de notre thésaurus avec d'autres thésaurus **de manière semi-automatique**

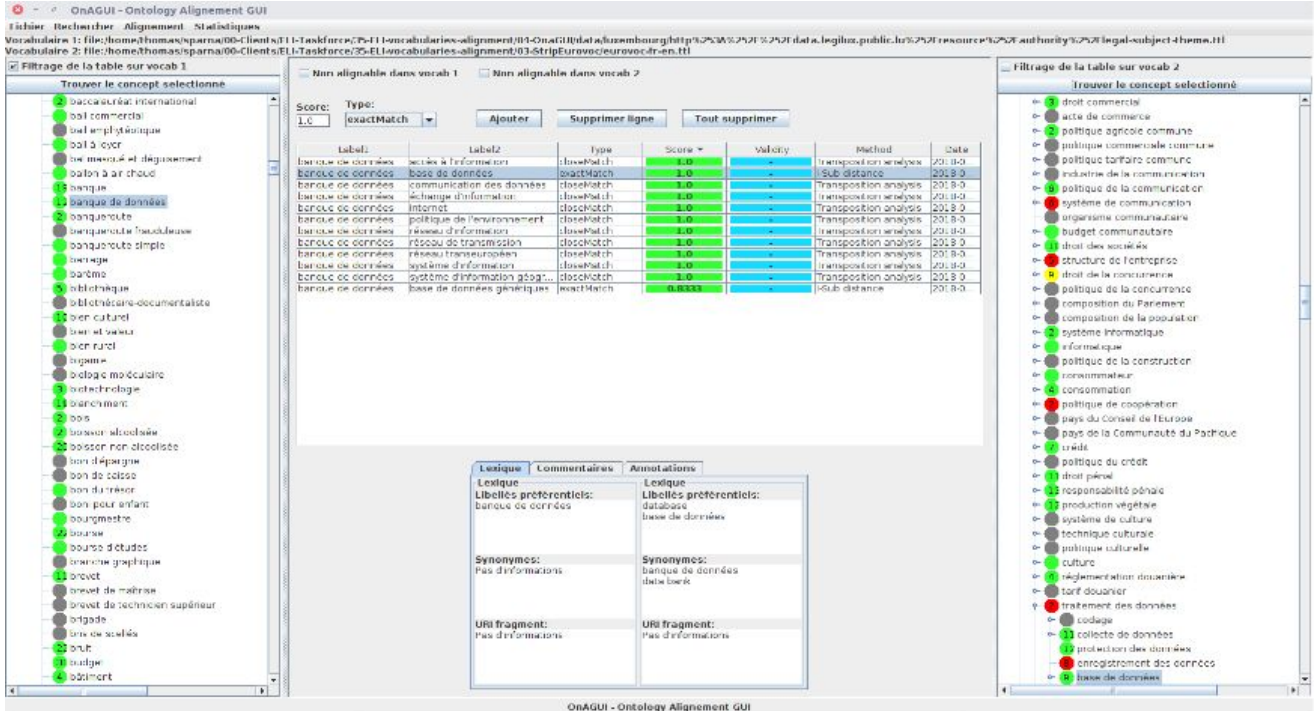
- L'outil produit un fichier d'alignement des thésaurus au format EDOAL ou en utilisant les propriétés d'alignement SKOS directement
- Les alignements peuvent ensuite être appliqués à l'aide d'outil comprenant le format EDOAL

VocBench permet la création de fichiers d'alignements et l'application des alignements sur les thésaurus.

démonstration du process sur un cas simple

# En pratique: alignement avec d’autres thésaurus

onaGUI pour aligner les concepts d’un thésaurus avec d'autres thésaurus **de manière semi-automatique**



# Rappel sur  les bonnes pratiques de création d’un thésaurus

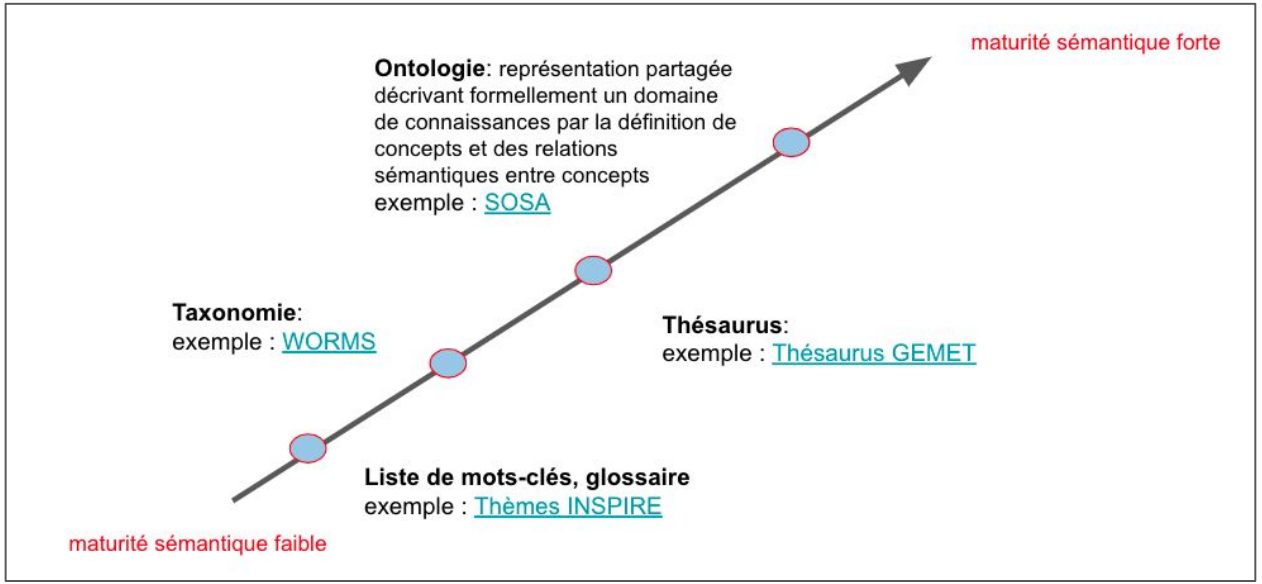
**1** - Utiliser les **thésaurus** disciplinaires **existants** et de **référence**

**2** - Contribuer si possible à les enrichir  - **éviter** le plus possible de produire des **thésaurus “locaux”**

**3** - Utiliser des **URI persistantes** et des identifiants opaques

**4** - Les **enrichir** par des **alignements** et des **nouveaux concepts** en les organisant autour de la vision d’observation : **Variables, Plateforme-capteur, Objet d’intérêt**

# Aller vers une démarche plus ontologique

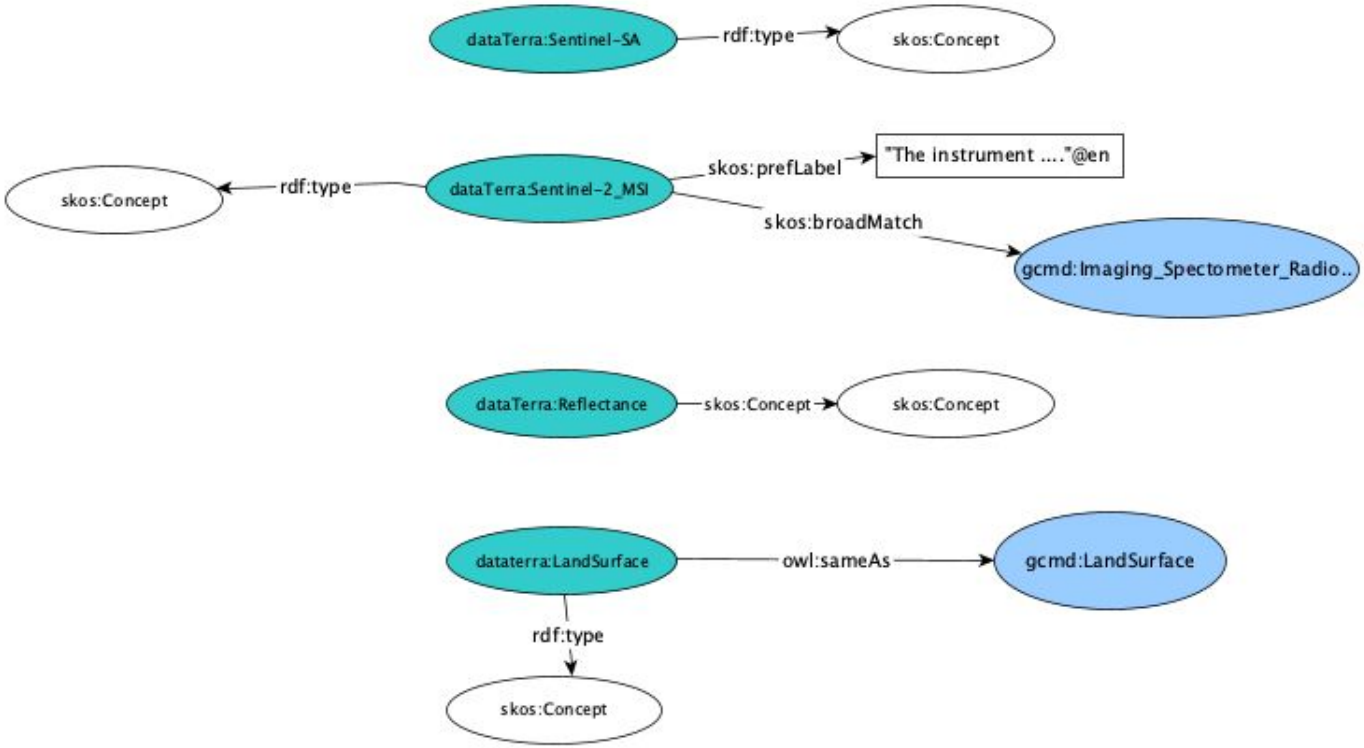



Améliorer la sémantique de la connaissance

Quels apports pour les fonctionnalités de nos dispositifs de gestion ?

# Aller vers une démarche plus ontologique

## Représentation de la connaissance proposée par SKOS

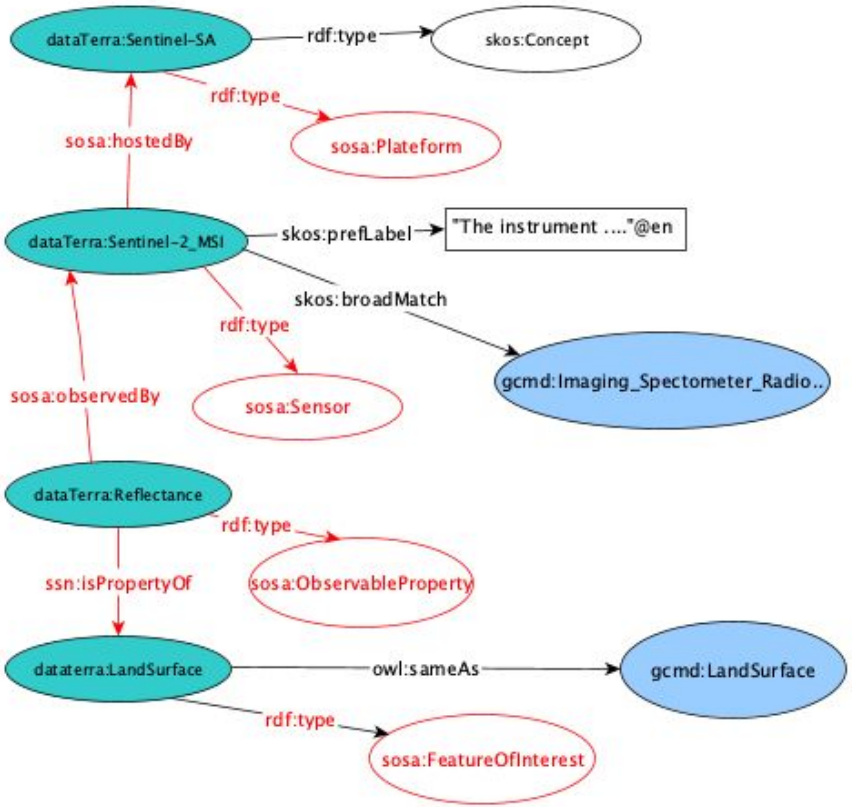



→ Faiblesse (pour les machines) décrire la nature des concepts et leur inter relations

*Graphe portant sur les concepts SKOS : Land Surface, Reflectance, Sentinel-2 (extrait du thésaurus de Data Terra)*

# Aller vers une démarche plus ontologique

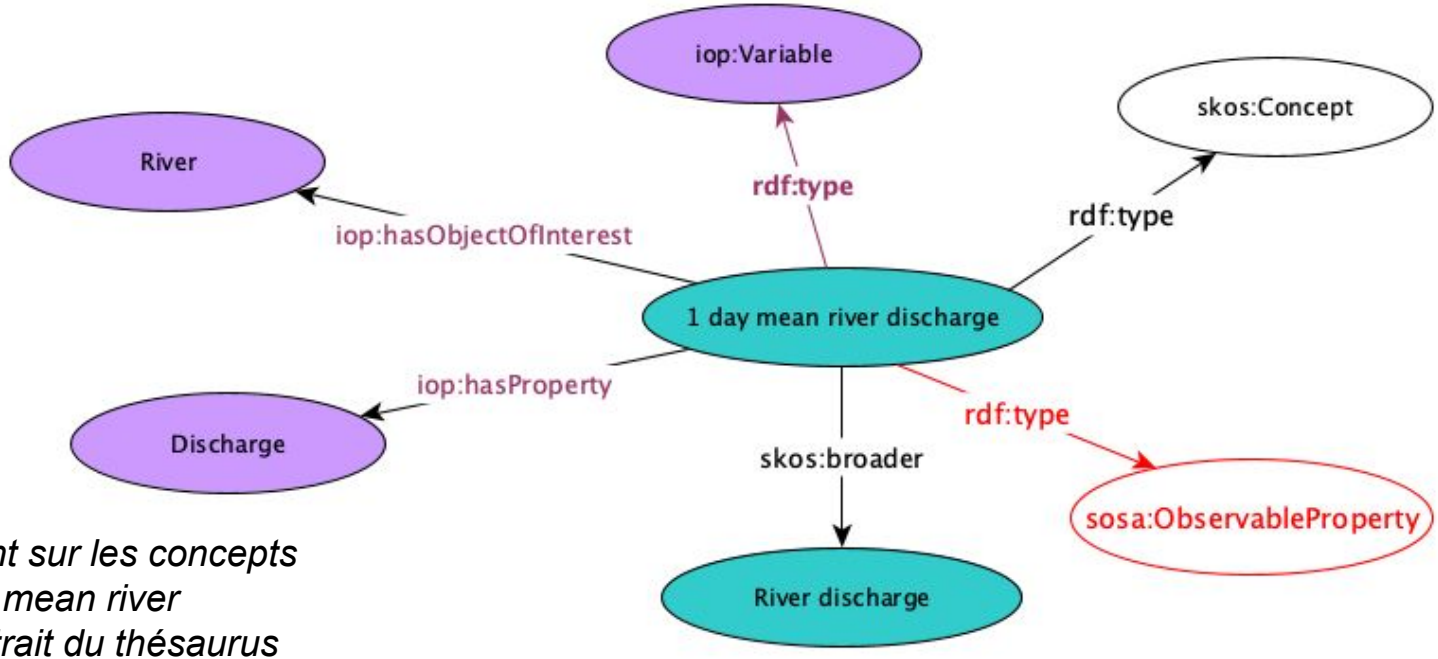
**“Augmenter” les thésaurus de la sémantique du domaine: utilisation de l’ontologie d’observation SOSA**




*Graphe portant sur les concepts SKOS : Land Surface, Reflectance, Sentinel-2 (extrait du thésaurus de Data Terra)*

# Aller vers une démarche plus ontologique

**“Augmenter” les thésaurus de la sémantique du domaine: utilisation de l’ontologie de variables d’observation I-ADOPT**



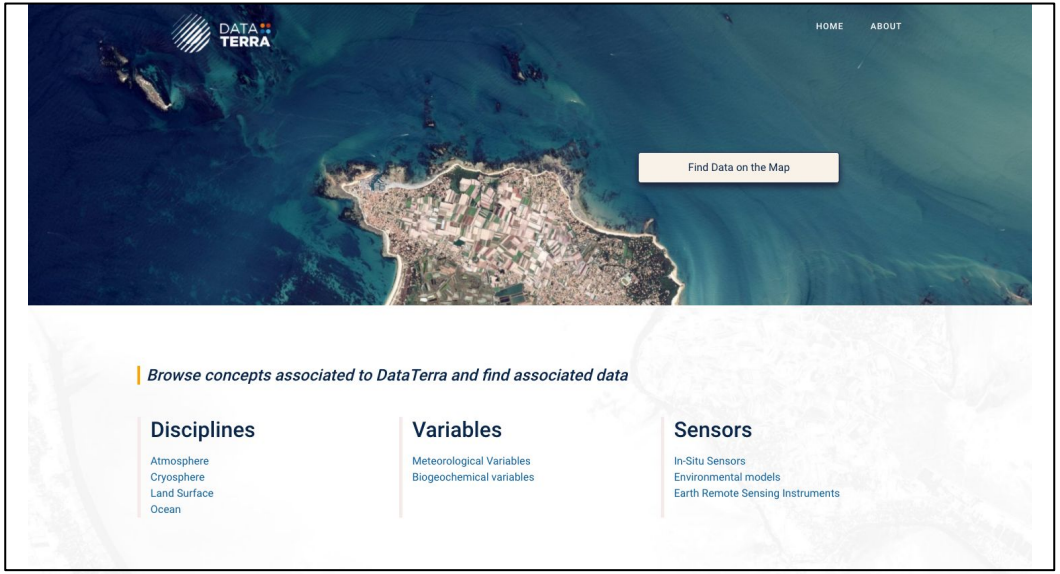

*Graphe portant sur les concepts SKOS : 1 day mean river discharge (extrait du thésaurus de OZCAR - THEIA)*

# Aller vers une démarche plus ontologique

Cas d'usage portail de la connaissance Data Terra: **Naviguer dans les données pluri-disciplinaires, multi-sources du système Terre**

Un climatologue veut réaliser des réanalyses des données climatiques. Il cherche des données de précipitations *in-situ* en afrique subsaharienne

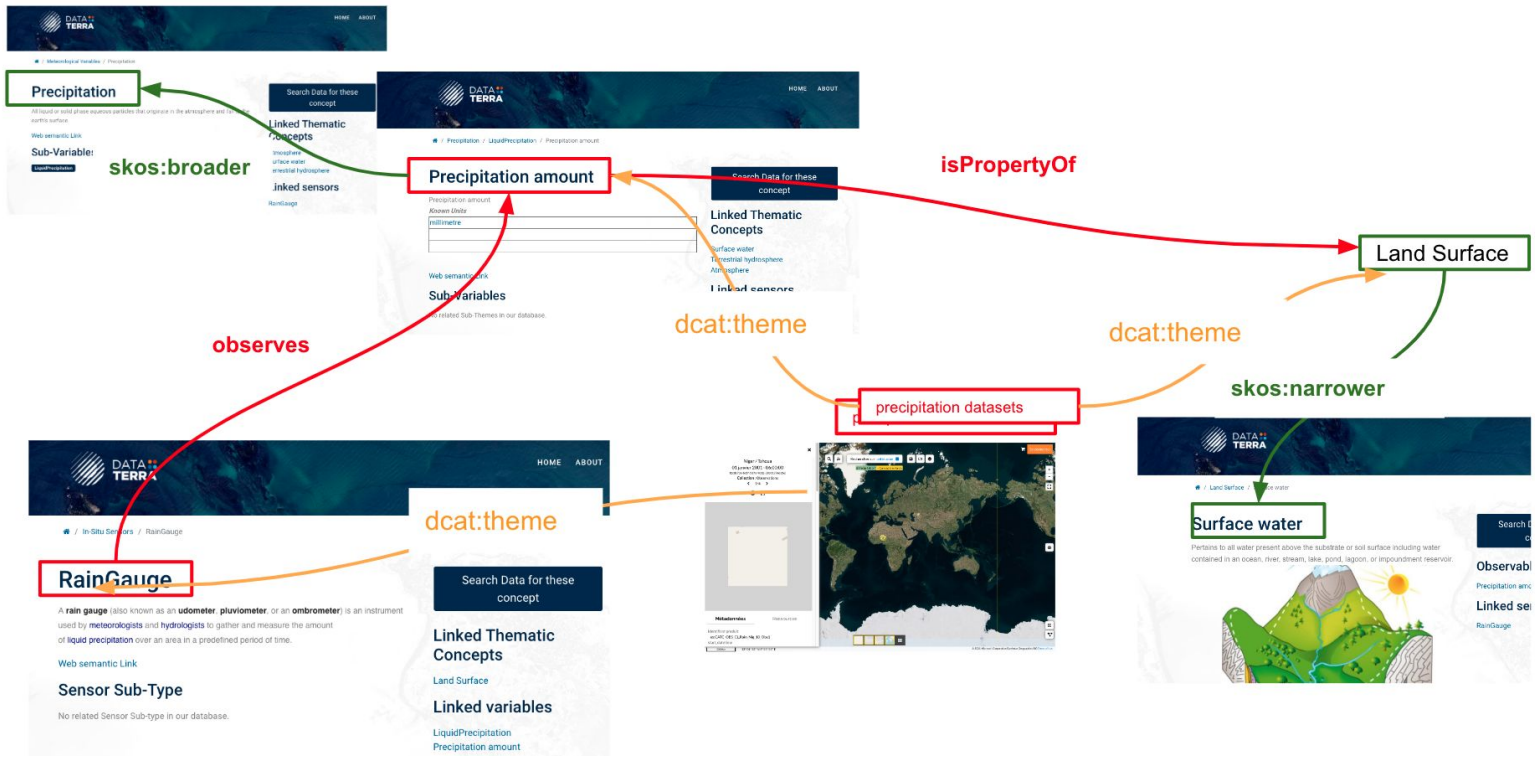
https://dataterra.geomatys.com/




1 - Il interroge le catalogue ou
**2 - il part à la découverte des données**

# Aller vers une démarche plus ontologique

Cas d'usage portail de la connaissance Data Terra: un graphe sémantique pour **naviguer dans les données pluri-disciplinaires, multi-sources du système Terre**



# Thésaurus Theia/OZCAR

https://w3id.org/ozcar-theia

## Variable

Skosmos

Vocabularies About Feedback Help | en Français

OZCAR Theia in-situ thesaurus

English Search

Alphabetical Hierarchy Groups

- Categories of variable
- Constraint
- Instrument
- Method
- Observable property
- Phenomenon
- Physical entity
- Process
- Time
- Variable
  - Atmosphere variable
  - Biosphere variable
  - Cryosphere variable
  - Land surface variable
  - Terrestrial hydrosphere variable
    - Groundwater hydrology
    - Karst hydrology
    - Surface water hydrology
      - Surface water chemistry
      - Surface water microbiology
      - Surface water physic variable
        - **River discharge**
          - 1 day mean river discharge
          - 1 hour mean river discharge
          - 1 month mean river discharge
          - 10 minutes mean river discharge
          - 15 minutes mean river discharge
          - Instantaneous river discharge
          - Lower boundary of uncertainty interval of 1 day mean river discharge
          - Lower boundary of uncertainty interval of 1 hour mean river discharge

Variable > Terrestrial hydrosphere variable > Surface water hydrology > Surface water physic variable > River discharge

| | |
|---|---|
| PREFERRED TERM | **River discharge** |
| TYPE | Variable |
| BROADER CONCEPT | Surface water physic variable |
| NARROWER CONCEPTS | 10 minutes mean river discharge<br>15 minutes mean river discharge<br>1 day mean river discharge<br>1 hour mean river discharge<br>1 month mean river discharge<br>Instantaneous river discharge<br>Lower boundary of uncertainty interval of 1 day mean river discharge<br>Lower boundary of uncertainty interval of 1 hour mean river discharge<br>Upper boundary of uncertainty interval of 1 day mean river discharge<br>Upper boundary of uncertainty interval of 1 hour mean river discharge |
| HASOBJECTOFINTEREST | River |
| HASPROPERTY | Discharge |
| SIMPLIFIED LABEL | Surface water discharge |
| URI | https://w3id.org/ozcar-theia/c_9c959860 |
| DOWNLOAD THIS CONCEPT: | RDF/XML TURTLE JSON-LD — Created 2/7/22, last modified 4/29/22 |

## Catégorie de variable

Alphabetical Hierarchy Groups

- Turbidity
- Volume
- Volumetric mass density
- Water holding capacity
- Water potential
- Water retention curve
- Phenomenon
- Physical entity
- Process
- Time
- Variable
  - Atmosphere variable
  - Biosphere variable
  - Cryosphere variable
  - Land surface variable
  - Terrestrial hydrosphere variable
    - Groundwater hydrology
    - Karst hydrology
    - Surface water hydrology
      - Surface water chemistry
      - Surface water microbiology
      - **Surface water physic variable**
        - River discharge
        - Surface water conductivity
        - Surface water level
        - Surface water pH
        - Surface water suspended sediment concentration
        - Surface water temperature
        - Surface water turbidity
          - Pond turbidity

Variable > Terrestrial hydrosphere variable > Surface water hydrology > Surface water physic variable

| | |
|---|---|
| PREFERRED TERM | **Surface water physic variable** |
| TYPE | Category of variable |
| BROADER CONCEPT | Surface water hydrology |
| NARROWER CONCEPTS | River discharge<br>Surface water conductivity<br>Surface water level<br>Surface water pH<br>Surface water suspended sediment concentration<br>Surface water temperature<br>Surface water turbidity |
| SIMPLIFIED LABEL | Surface water physic |
| URI | https://w3id.org/ozcar-theia/c_b23989c9 |
| DOWNLOAD THIS CONCEPT: | RDF/XML TURTLE JSON-LD — Created 2/7/22, last modified 4/29/22 |

## Objet d'intérêt

Alphabetical Hierarchy Groups

- Categories of variable
- Constraint
- Instrument
- Method
- Observable property
- Phenomenon
- Physical entity
  - Chemical entity
  - Environmental entity
    - Atmosphere
    - Biosphere
    - Cryosphere
    - Environmental interface entity
    - Hydrosphere
      - Groundwater
      - Surface water
        - Lake
        - Pond
        - **River**
        - Spring
      - Water
    - Land surface
  - Man-made object
  - Structure
  - Subatomic particle
- Process
- Time
- Variable

Physical entity > Environmental entity > Hydrosphere > Surface water > River

| | |
|---|---|
| PREFERRED TERM | **River** |
| TYPE | Entity |
| DEFINITION | [GEMET] A stream of water which flows in a channel from high ground to low ground and ultimately to a lake or the sea, except in a desert area where it may dwindle away to nothing. A river and all its tributaries within a single basin is termed a drainage system. |
| BROADER CONCEPT | Surface water |
| URI | https://w3id.org/ozcar-theia/c_97bb7b91 |
| DOWNLOAD THIS CONCEPT: | RDF/XML TURTLE JSON-LD — Created 2/7/22, last modified 4/29/22 |

# Thésaurus Theia/OZCAR

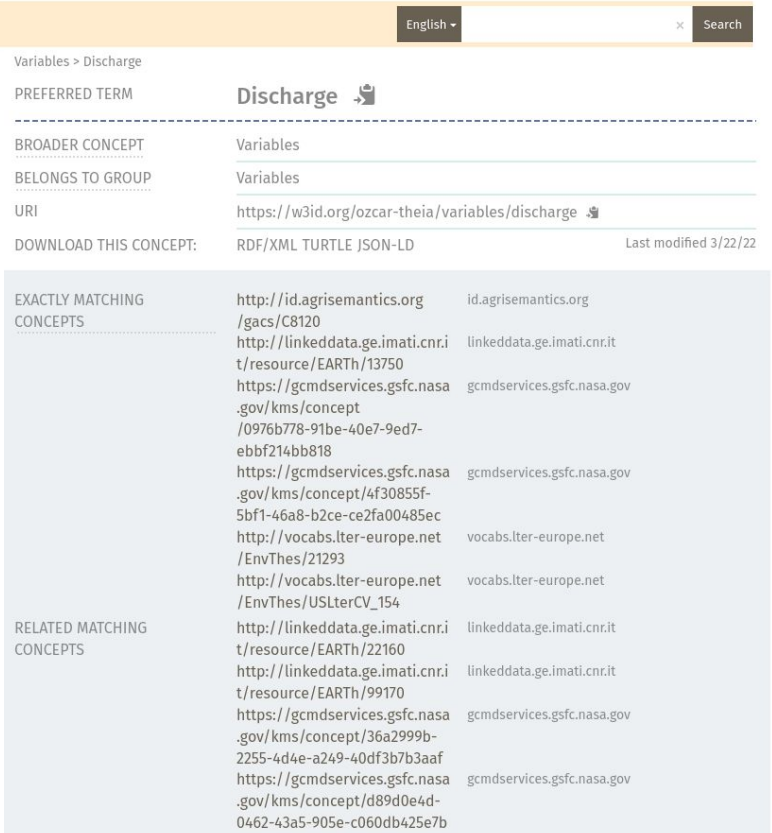
Alignements avec les thésaurus: GCMD, EnvThes, AGROVOC, GEMET, AnaEE, …

<https://w3id.org/ozcar-theia/variables/discharge> a skos:Concept;

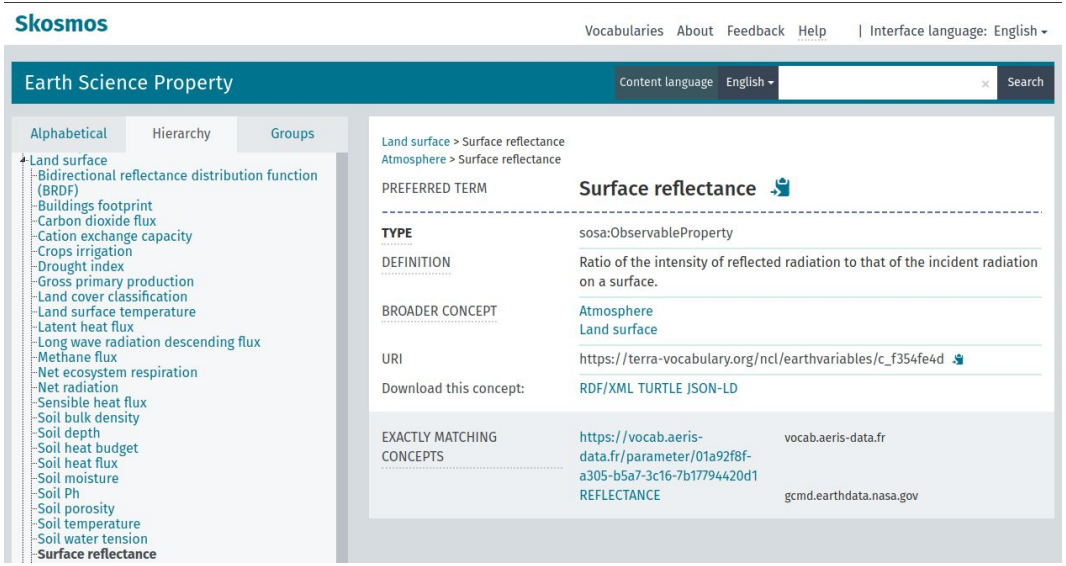
**skos:exactMatch** <http://id.agrisemantics.org/gacs/C8120>,
<http://linkeddata.ge.imati.cnr.it/resource/EARTh/13750>,
<http://vocabs.lter-europe.net/EnvThes/21293>,
<http://vocabs.lter-europe.net/EnvThes/USLterCV_154>,
<https://gcmdservices.gsfc.nasa.gov/kms/concept/0976b778-91be-40e7-9ed7-ebbf214bb818>,
<https://gcmdservices.gsfc.nasa.gov/kms/concept/4f30855f-5bf1-46a8-b2ce-ce2fa00485ec>;
skos:inScheme :ozcarTheiaThesaurus;

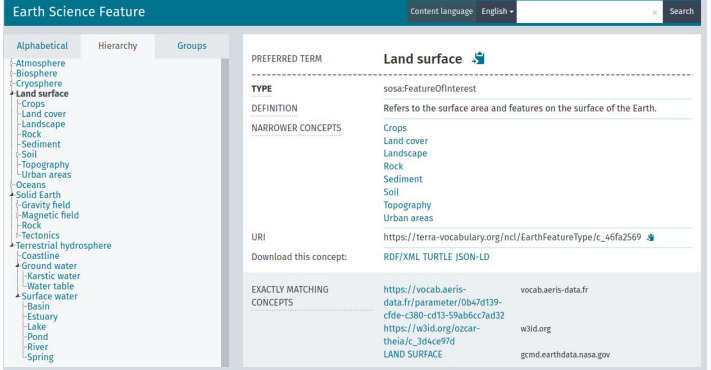
skos:prefLabel "**Discharge"@en**;

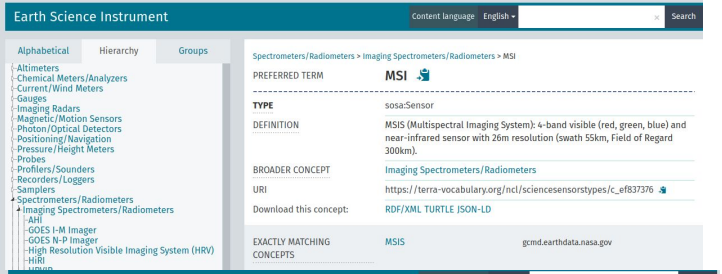
**skos:relatedMatch** <http://linkeddata.ge.imati.cnr.it/resource/EARTh/22160>,
<http://linkeddata.ge.imati.cnr.it/resource/EARTh/99170>,
<https://gcmdservices.gsfc.nasa.gov/kms/concept/36a2999b-2255-4d4e-a249-40df3b7b3aaf>,
<https://gcmdservices.gsfc.nasa.gov/kms/concept/d89d0e4d-0462-43a5-905e-c060db425e7b> .

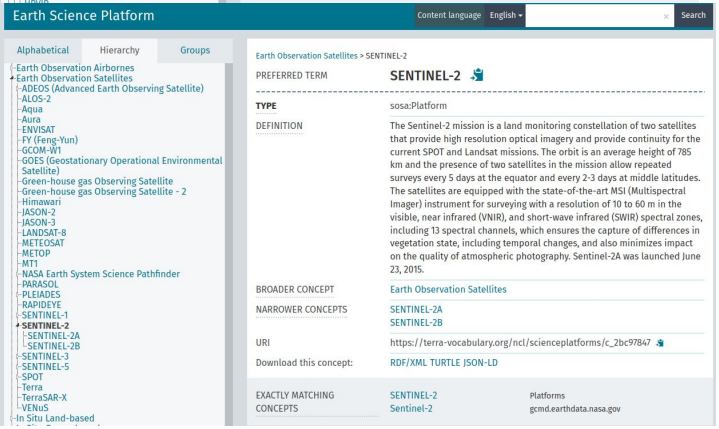


# Thésaurus Data Terra



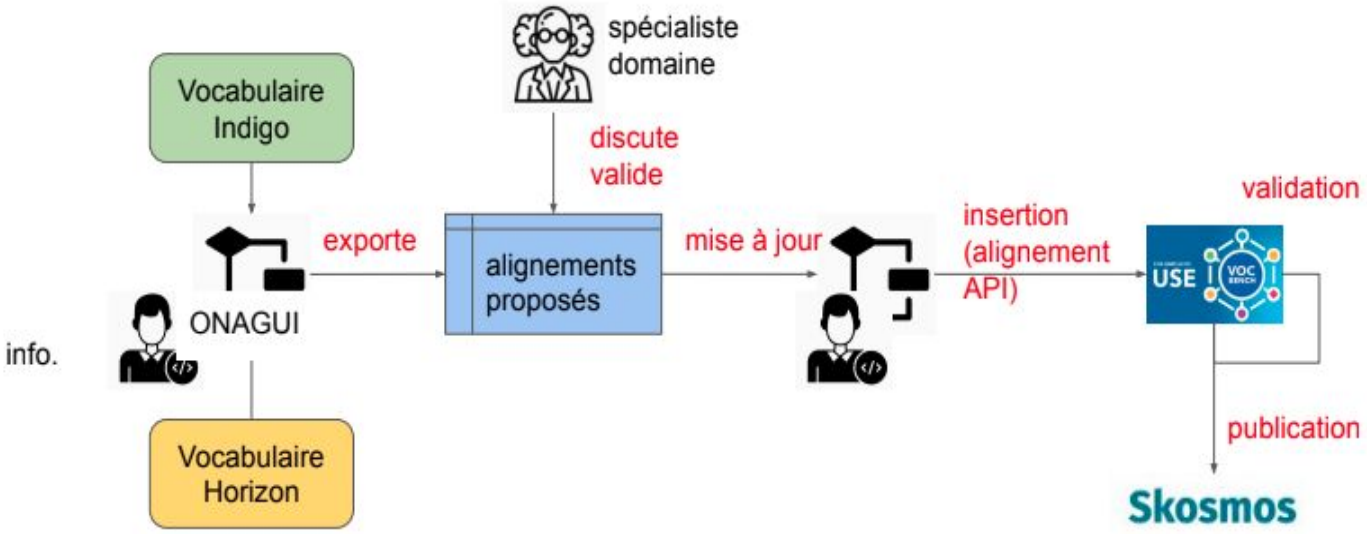



# Exemple d’un processus d’alignement : IRD référentiels

**L’alignement avec un autre thésaurus en 4 étapes**



# Limites de la description des thésaurus avec SKOS

## I-ADOPT

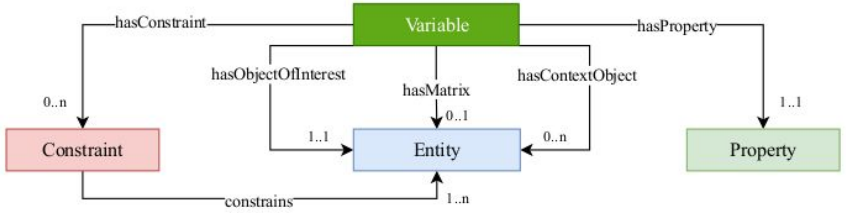
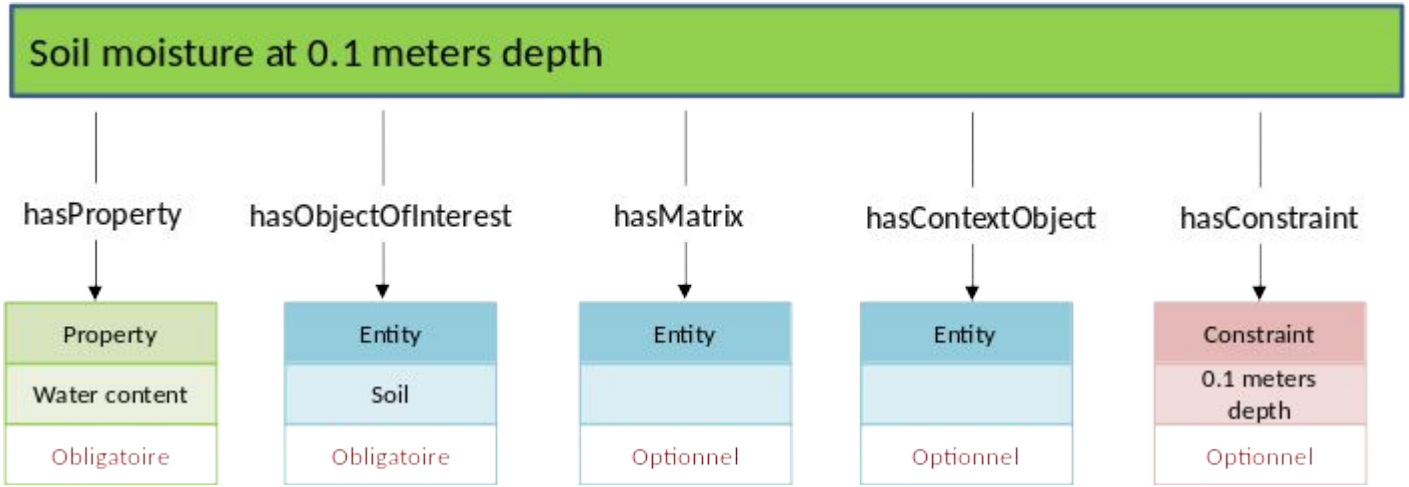


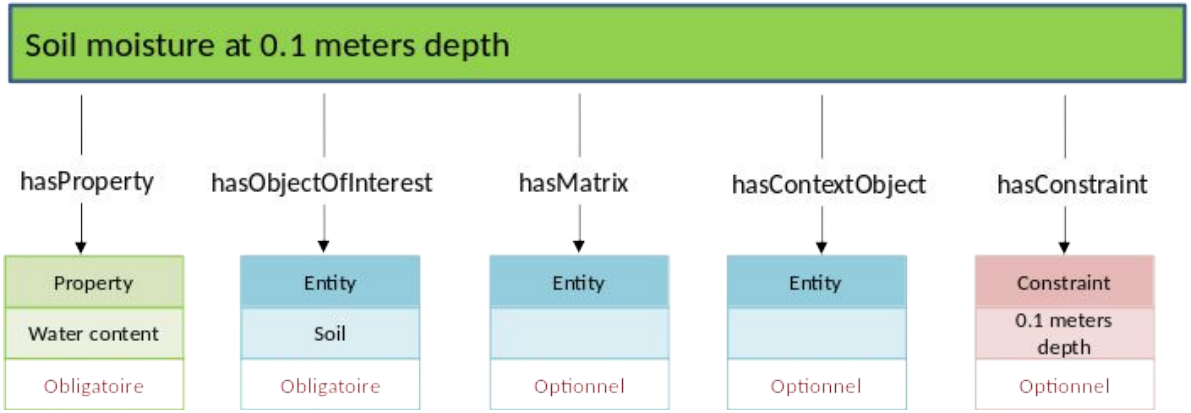

**Figure 2:** The I-ADOPT Framework.

*Magnana et al., S4BioDiv2021, 2021*



# Limites de la description des thésaurus avec SKOS

I-ADOPT




```
:c_72e50d7b a skos:Concept, iadopt:Variable;
  dct:created "2022-02-07T17:00:00"^^xsd:dateTime;

  skos:inScheme :ozcarTheiaThesaurus;
  skos:prefLabel "Soil moisture at 0.1 meters depth"@en;

  iadopt:hasConstraint :c_95689cb4;
  iadopt:hasObjectOfInterest :c_dc6163a1;
  iadopt:hasProperty :c_3b665477  .
```