

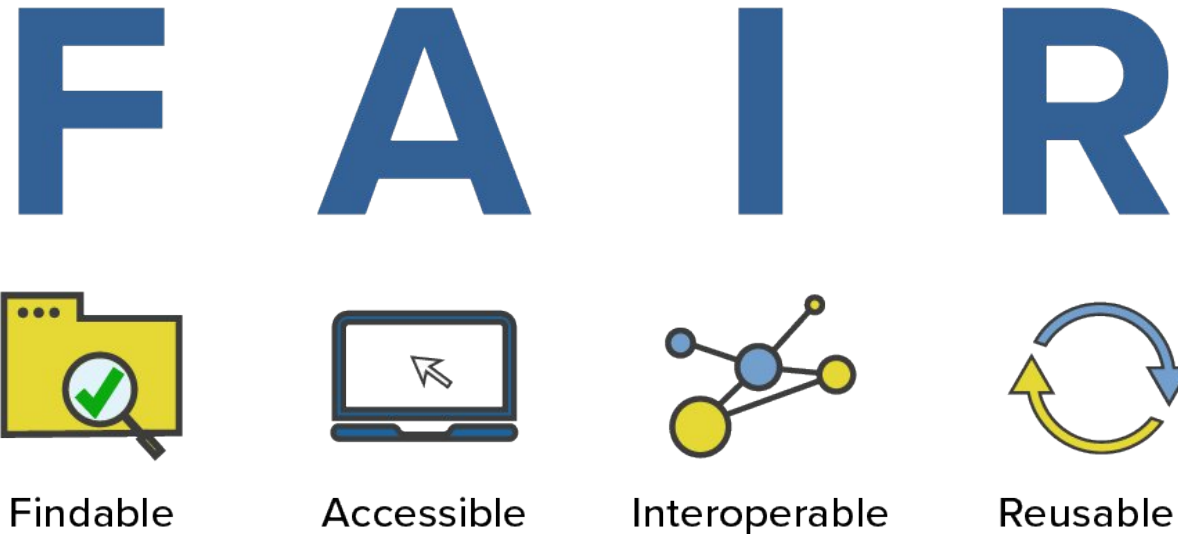
FAIRiser les données du réseau d'observatoires de la zone critique Français (OZCAR) : cas d'utilisation de l'ontologie I-ADOPT

Charly Coussot (1), Isabelle Braud (2), Véronique Chaffard (3), Brice Boudevillain (3)

1. OSUG - Observatoire des Sciences de l'Univers, Grenoble, France
2. INRAE, RiverLy, Villeurbanne, France
3. IGE, Grenoble, France

Principes FAIR

Le but des principes FAIR est d'optimiser la réutilisation des données.



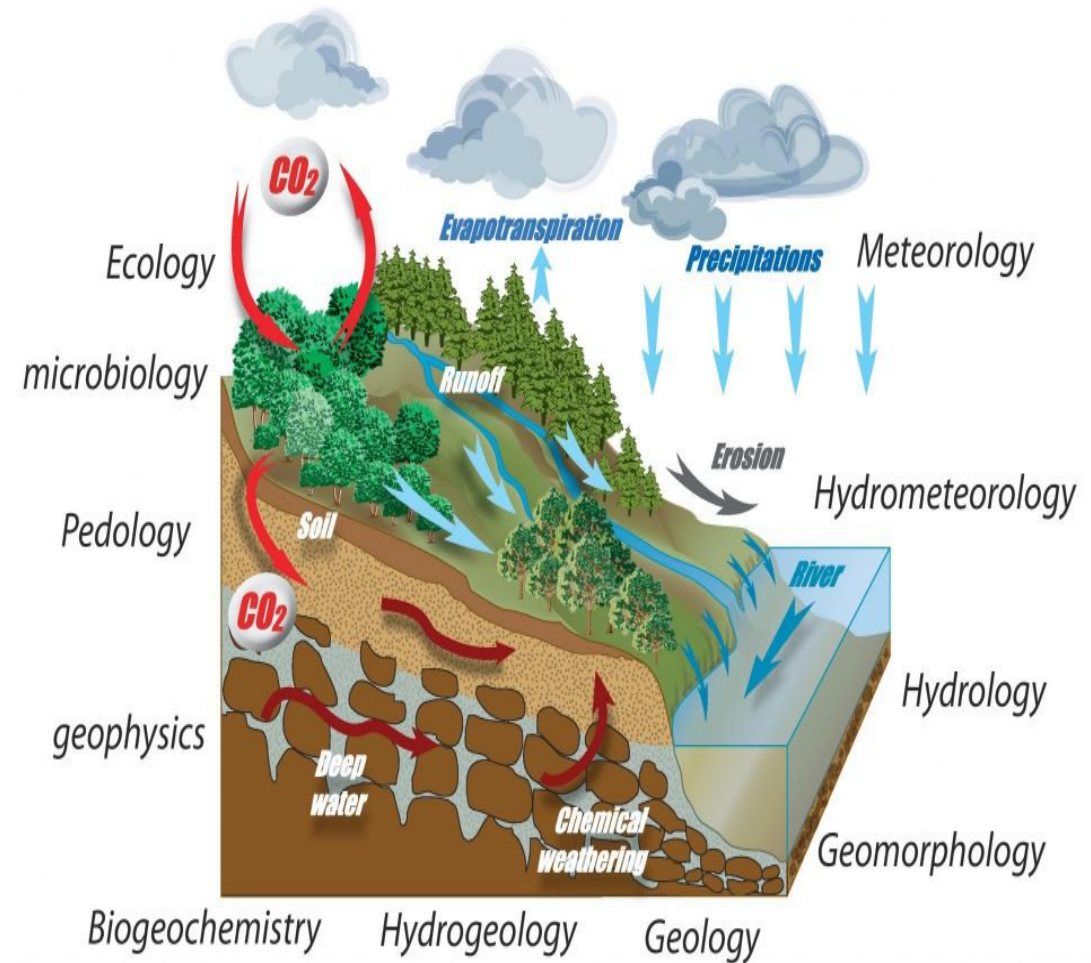
Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

Objectif de la présentation

- Le contexte de l'observation de la zone critique en France
- Le système d'observation Theia/OZCAR pour FAIRiser les données d'observation de la zone critique
- Les standards d'interopérabilité utilisés **avec un focus sur I-ADOPT**

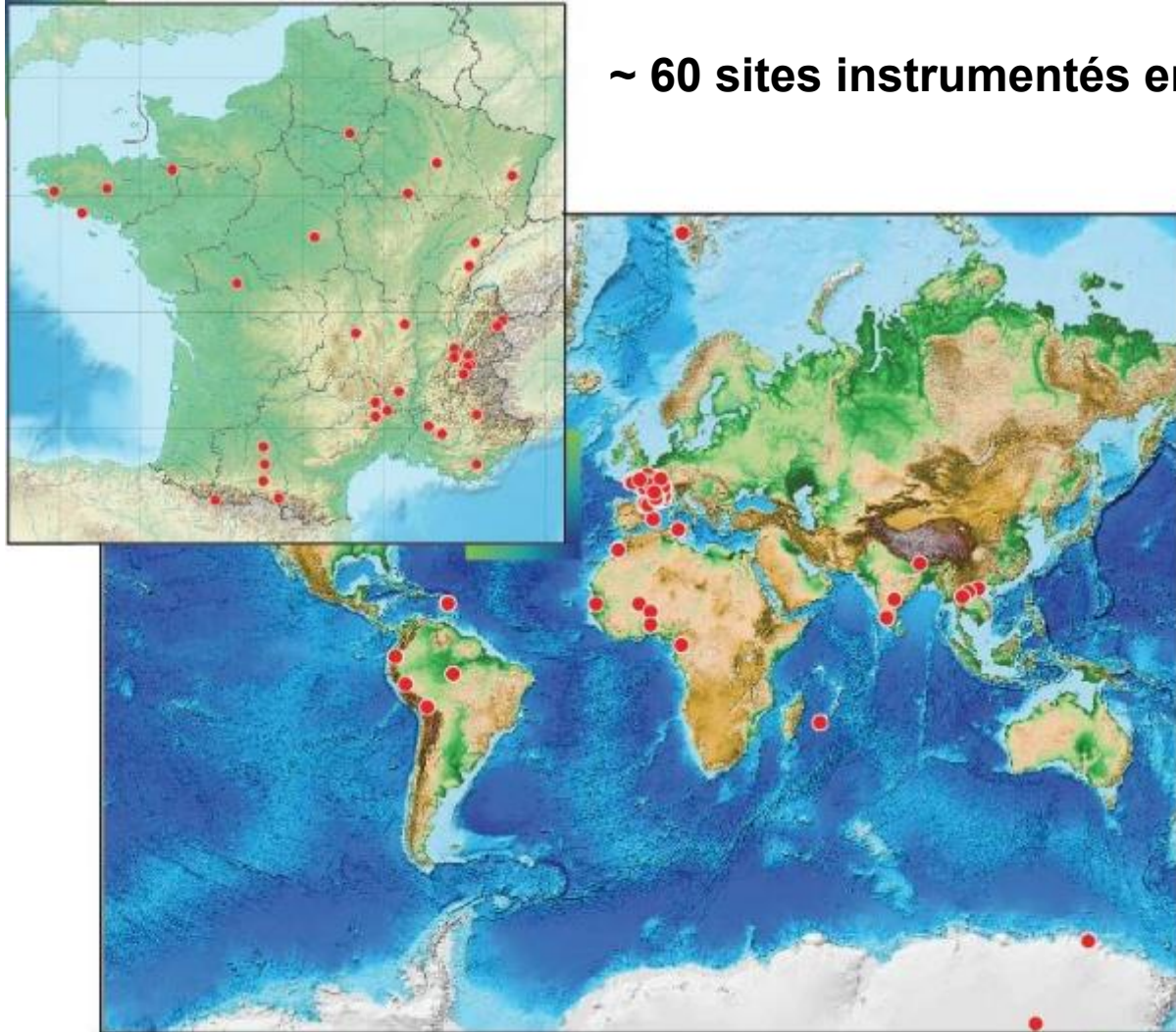
Qu'est ce que la zone critique ?

La zone critique est l'environnement terrestre qui s'étend de l'atmosphère, jusqu'aux roches non altérées. Elle englobe la basse atmosphère, la pédosphère, la zone vadose, les nappes d'eau souterraine et les roches altérées. C'est l'environnement de proche-surface hétérogène dans lequel des interactions complexes entre la roche, le sol, l'eau, l'air et les organismes vivants régulent l'habitat naturel et déterminent la disponibilité des ressources vitales. La zone critique abrite presque toute la vie continentale, dont l'humanité.



OZCAR : un réseau d'observatoire de la zone critique

~ 60 sites instrumentés en France et dans les pays du Sud



Un long historique d'observation

- Les observatoires ont été développés de manière indépendante
- Des pratiques de gestion de données hétérogènes

Une diversité d'objets d'intérêts et d'observations

- Bassins versants, rivières, karsts
- Aquifères
- Glaciers
- Permafrost
- Tourbières

OZCAR : un réseau d'observatoire de la zone critique

Une large diversité de variable collectées

- Séries chronologiques ponctuelles : météorologie, hydrologie, hydrogéologie, glaciologie, bilan énergétique de surface, flux de sédiments, éléments géochimiques et concentrations de contaminants.
- Carottes de sol
- Profils géophysiques 2D
- Cartes (matricielles ou vectorielles) caractérisant les sites : occupation des sols, MNT, propriétés physiques du sol
- Enquêtes : rotations des culture



OZCAR : un réseau d'observatoire de la zone critique

The collage illustrates the OZCAR network, which is a network of observatories for the critical zone. It features a map of France on the left, highlighting the locations of various participating institutions and their logos, including INRAE, OSUR, BRGM, Ope, OHGE, EOST, INRAE BDOH, OTHU, Real Collobrier, OSUG, CESBIO, SEDOO, BV Auradé, M-TROPICS, OREME, INRAE, IRD, INGREF, INAT, and OSU-Réunion.

Overlaid on the map are several screenshots of web portals:

- ORE AgrHyS**: A portal for agricultural hydrology data, showing a map of France and a list of observatories.
- Base de Données des Observatoires en Hydrologie**: A portal for hydrological observatory data, featuring a presentation of the network and a list of observatories.
- Réseau National de sites Hydrogéologiques**: A portal for the national network of hydrogeological sites, providing information on the database and the application of Google Earth.
- le système d'information du SNO Tourbières (en construction)**: A portal for the SNO Tourbières system, detailing the collection, organization, and diffusion of data.

Partage des données dans le cadre d'OZCAR

22 observatoires de long terme ayant des historiques et des pratiques de gestion de la données différentes :

- Vocabulaires et descriptions des données (métadonnées) hétérogènes
- Formats des données propres à chaque observatoires

Des difficultés à connaître l'existant
Des difficultés à croiser les données

Les données doivent pouvoir être réutilisées dans un contexte interdisciplinaire dans le cadre de la science de la zone critique et au-delà.



Données du système Terre



Données de l'environnement et des écosystèmes



Données sur l'eau

Nécessité de rendre FAIR les données des observatoires de OZCAR

Système d'informations Theia/OZCAR

Objectifs du SI Theia/OZCAR

Rendre les données de l'IR OZCAR accessibles à la communauté scientifique selon les principes du FAIR data (Findable Accessible Interoperable Reusable) (Braud et al., 2020)

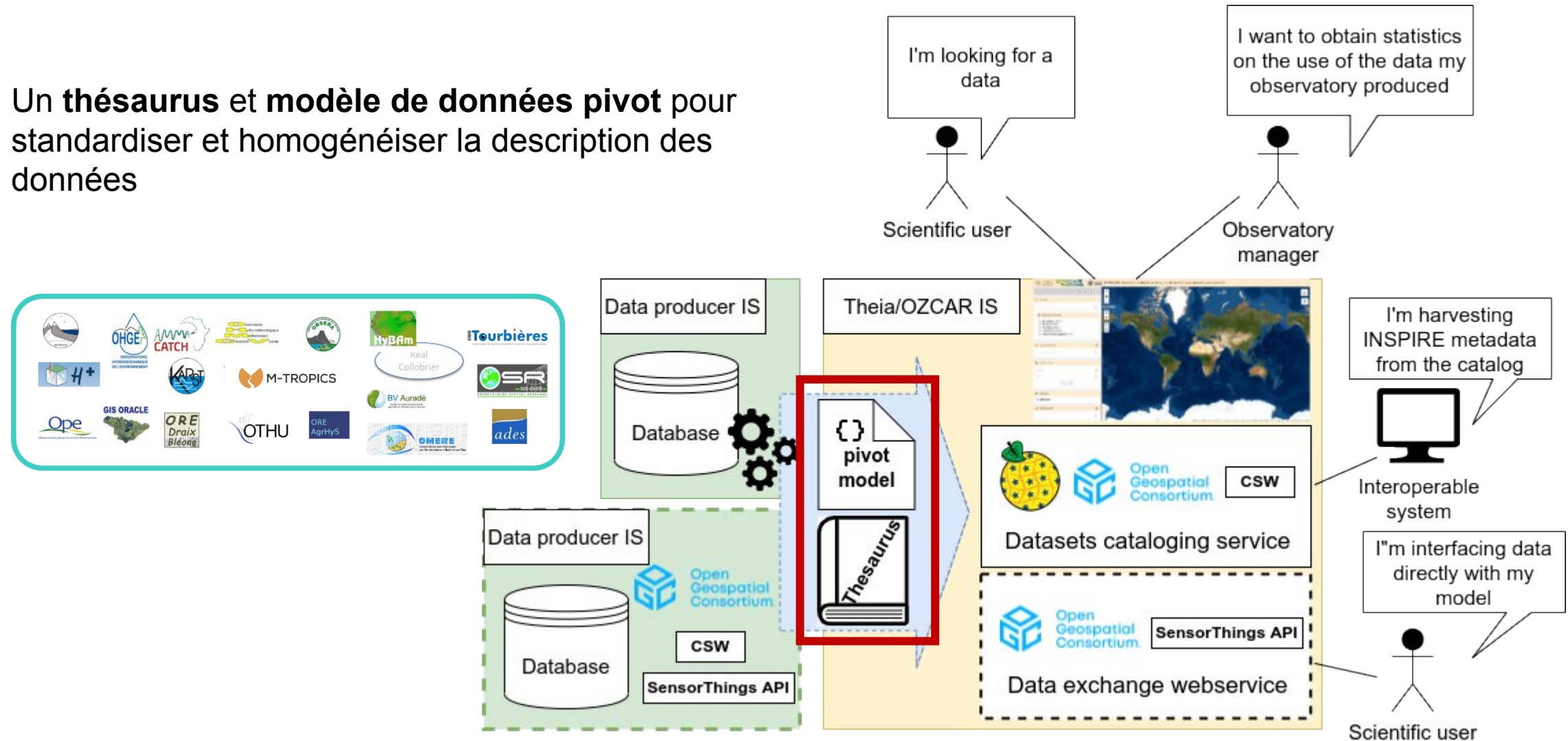


<https://in-situ.theia-land.fr>

- Homogénéisation et enrichissement des métadonnées par un modèle pivot et [un thésaurus](#) de noms de variables et d'objets d'intérêt
- Organisation d'un flux de données entre les SIs des observatoires et le SI Theia/OZCAR qui fournit des services (portail de recherche, service web de catalogage, ..)
- Visualisation et téléchargement des données dans des formats harmonisés (en cours de développement)

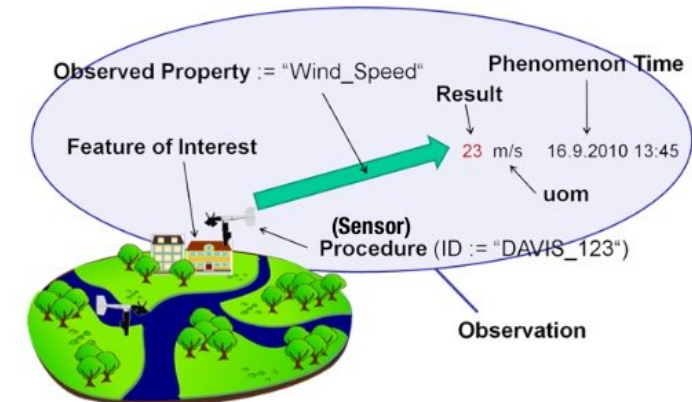
Système d'informations Theia/OZCAR

Un **thésaurus** et **modèle de données pivot** pour standardiser et homogénéiser la description des données



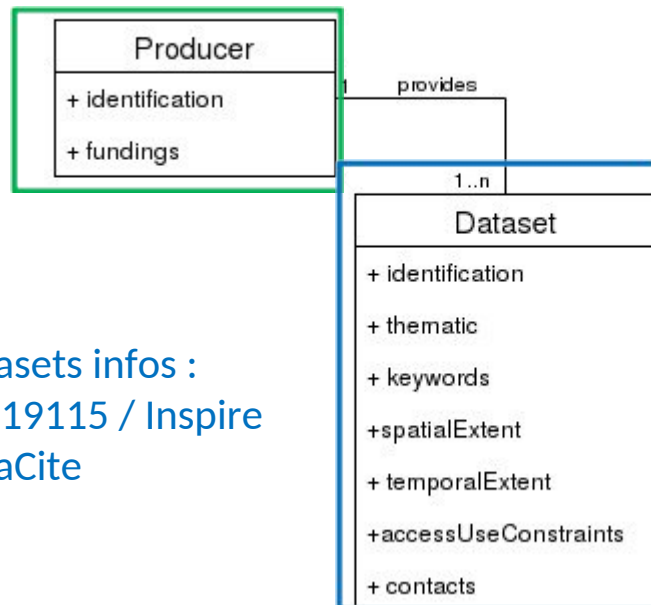
Le modèle pivot Theia/OZCAR

Harmoniser la description des données entre producteurs et permettre de mettre en œuvre des services web standardisés en sortie (CSW, SensorThings)

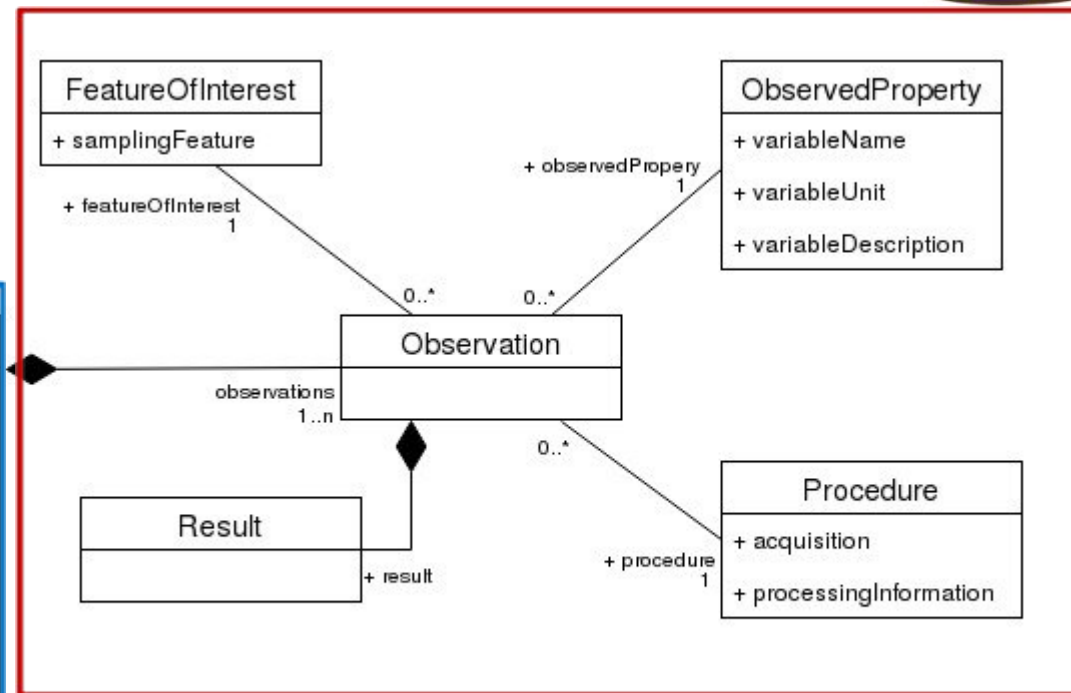


Observations infos: O&M

Producers info (Dublin Core, DataCite)

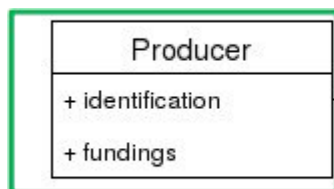


Datasets infos :
ISO 19115 / Inspire
DataCite



Le modèle pivot Theia/OZCAR

Producers info (DataCite, scanR)



Identifiant Organisme



Producer

Identification
Name: OZCAR-RI SNO KARST
Description: Karst is a complex hydrogeologic entity whose study requires specific and original models and approaches. The goal of SNO KARST is to bring out scientific questions and advances, specifically in terms of hydrogeochemical flux modeling at karst hydrosystem outlets and the links between global changes and the physico-chemical composition of water at the interface between hydrologic and hydrogeologic compartments.
Email: snokarst@services.cnrs.fr

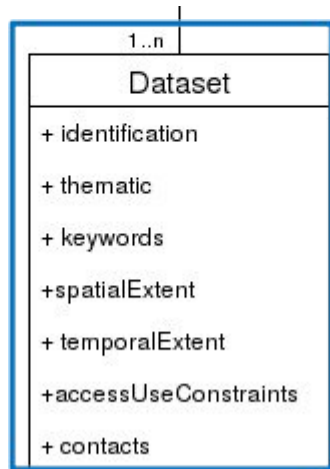
Contacts
Project leader
Hervé Jourde
herve.jourde@umontpellier.fr
Hydrosiences Montpellier (HSM) - France

Online resources
Producer data portal : <https://data.oreme.org/observation/snokarst>
Website : <https://sokarst.org/>
Webservices

Fundings
French universities and schools
Aix-Marseille University Aix-Marseille Université(AMU) - France
Avignon Université- France
Université de Bordeaux- France

Le modèle pivot Theia/OZCAR

Datasets infos :
ISO 19115 / Inspire
DataCite



Dataset

Identification

Title: Time series of type chemistry in Dardennes - Le Las basin - PORT-MIOU observatory - KARST observatory network - OZCAR Critical Zone network Research Infrastructure

Description: Data of type chemistry measured on Dardennes - Le Las basin within the framework of the PORT-MIOU observatory, that is part of the KARST observatory network. The KARST observatory network aims to strengthen knowledge-sharing and to promote cross-disciplinary research on karst systems. In the framework of the

Contacts

Principal investigator

Hervé Jourde

Online resources

Dataset download : <https://doi.org/10.15148/18b4d8bf-aa12-4159-92fc-23ec70d1198e>

Information : <https://sokarst.org>

Spatial extent

43.1

Temporal Extent

From 2012-10-19 to 2019-03-16

Data access and use constraints

Use constraint : These data were gathered within the framework of the PORT-MIOU observation site. PORT-MIOU observation site is part of the KARST observatory network (sokarst.org) initiated by INSU/CNRS, which aims to strengthen knowledge-sharing and to promote cross-disciplinary research on karst systems in the framework of the OZCAR Critical Zone network Research Infrastructure. SNO KARST data policy imposes the following constraints: (i) Formal acknowledgement of the dataset providers, (ii) Opportunity to collaborate on the project using the dataset,

Licence : [Licence ouverte / Open Licence](#)

Keywords

Le modèle pivot Theia/OZCAR

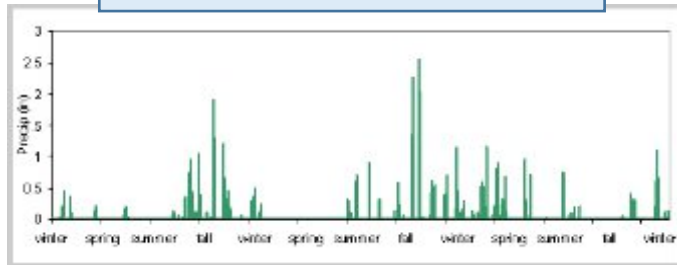


Station Météo at
Laguette peatland

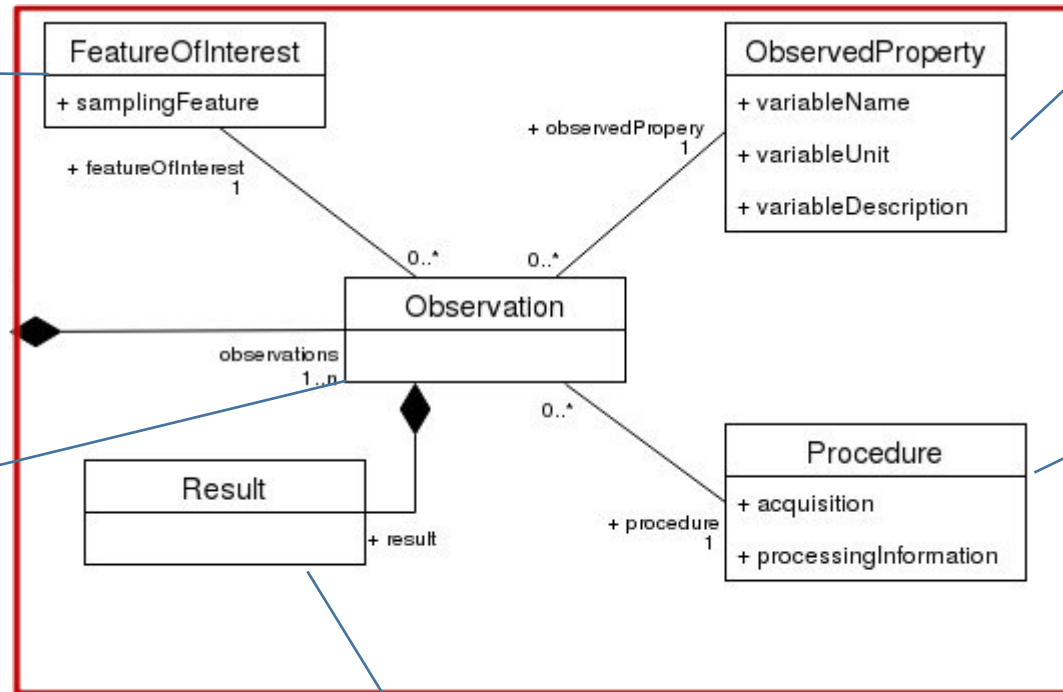
Nom de station (M)

Observation collection

Precipitation timeserie



Observations infos: O&M



Nom variable producteur
+ unité (M)

Precipitation amount



type de capteur (M)

Raingauge



Data file

format csv: 1 fichier/ variable et
/station

```

dateBeg;dateEnd;latitude;longitude;altitude;value;qualityFlags;[liste de valeur additionnelles
séparées par des « ; »];\n
2005-12-31T06:00:00Z;2006-01-01T06:00:00Z;13.5933;2.6544;152;-9999.9;21100.0;11;1,a,25;\n
2006-01-01T06:00:00Z;2006-01-02T06:00:00Z;13.5933;2.6544;152;-9999.9;21100.0;11;1,a;\n
2006-01-02T06:00:00Z;2006-01-03T06:00:00Z;13.5933;2.6544;152;-9999.9;21100.0;11;;
    
```

Le thesaurus Theia/OZCAR

Skosmos

Vocabularies About Feedback Help | en Français

Theia/OZCAR thesaurus

English -

Search

Alphabetical

Hierarchy

Groups

A B C D E F G H I K L M N O P

R S T U V W Y Z 0-9

Aboveground

Aboveground dry vegetation biomass

Aboveground herbaceous plant mass

Absolute humidity

Absorbance

Absorbed radiation

Abundance

Accumulation

Accumulation since last measurement

Accumulation since last raingauge bucket tip

Accumulation since the beginning of the year

Acetochlor

Acclonifen

Acoustic investigation variable

Acoustic velocity

Acoustic wave

Actinotherm index

Actual evapotranspiration

Actual evapotranspiration

Actual evapotranspiration of peatland

Adsorption coefficient

Aerosol variable

Air

Air absolute humidity

Air constraint

Air pressure

Air pressure at 20 meters height

Air pressure at 5 meters height

Air relative humidity

Air relative humidity at 15 meters height

Air relative humidity at 18 meters height

Air relative humidity at 2 meters height

Air relative humidity at 20 meters height

Air relative humidity at 5 meters height

Air specific humidity

Air temperature

Air temperature at 15 meters height

Air temperature at 18 meters height

Air temperature at 2 meters height

Air temperature at 20 meters height

Air temperature at 5 meters height

Alachlor

Albedo

Alkalinity

Altitude

Altitude → Height

Aluminium

Ametryn

Amino acid G → Glycine

Aminomethylphosphonic acid

Aminotriazole

Ammonium

Ammonium mass fraction in soil from 0 to 0.1 meters depth

Ammonium mass fraction in soil from 0.2 to 0.3 meters depth

Ammonium mass fraction in soil from 0.3 to 0.4 meters depth

Ammonium mass fraction in soil from 0.5 to 0.6 meters depth

Amount concentration → Molar concentration

AMPA → Aminomethylphosphonic acid

Amplitude

Andia

Vocabulary information

TITLE

Theia/OZCAR thesaurus

DESCRIPTION

Thesaurus for in situ data from Environmental and Critical Zone Sciences.
Used by Theia/OZCAR information system : <https://in-situ.theia-land.fr/>

CREATOR

Charly Cousot <https://orcid.org/0000-0002-0544-4802>
Véronique Chaffard <https://orcid.org/0000-0003-2823-7117>
Isabelle Braud <https://orcid.org/0000-0001-9155-0056>
Sylvie Galle <https://orcid.org/0000-0002-3100-8510>

LICENSE

<http://creativecommons.org/licenses/by/4.0/>

LANGUAGE

<http://lexvo.org/id/iso639-3/eng>

SOURCE

GCMD Science Keywords: <https://earthdata.nasa.gov/about/gcmd/global-change-master-directory-gcmd-keywords>
Interoperable Descriptions of Observable Property Terminologies (I-ADOPT) framework: <https://w3id.org/iadopt/ont>
Sandre vocabulary: Service d'Administration National des Données et Référentiels sur l'Eau: <http://id.eaufrance.fr/gpr/41>

CREATED

Monday, January 1, 2018 00:00:00

LAST MODIFIED

Monday, January 22, 2024 15:54:29
Monday, January 22, 2024 15:55:03
Monday, January 22, 2024 15:56:24
Monday, January 22, 2024 15:56:48
Monday, January 24, 2022 16:12:45

DC:REQUIRES

<http://purl.org/voc/cpm>
<http://www.w3.org/ns/prov-oi>
<http://www.w3.org/ns/ssn/>
<https://w3id.org/iadopt/ont>

OWL:IMPORTS

<http://purl.org/voc/cpm>
<http://www.w3.org/ns/prov-oi>
<http://www.w3.org/ns/sosa/>
<http://www.w3.org/ns/ssn/>
<https://w3id.org/iadopt/ont>

TYPE

<http://www.w3.org/2002/07/owl#Ontology>
<http://www.w3.org/2004/02/skos/core#ConceptScheme>

URI

<https://w3id.org/ozcar-theia/>

Download this vocabulary: RDF/XML TURTLE

- Faciliter la découverte des données en fournissant des critères de recherche à l'aide de noms de variables harmonisés.
- Améliorer la réutilisation et l'interopérabilité des données en fournissant des informations riches sur les variables permettant une interprétation sans ambiguïté des données par une communauté d'utilisateurs plus large que celle qui les a produites et également par des machines.

Exemple du niveau de détail attendu:

- "Precipitation amount": intégré sur quel pas de temps? Est-ce de la précipitation liquide ou solide (neige)?
- "Water level": sur quelle entité s'applique cette variable ? eau surface ou eau souterraine?

Les variables mesurées par les observatoires sont décrites à l'aide de l'ontologie I-ADOPT.

https://in-situ.theia-land.fr/skosmos/theia_ozcar_thesaurus/en/

Pourquoi utiliser une ontologie pour nommer les variables ?

Problème !!

Plus le concept de la variable est précis, plus il est difficile de trouver des relations de similitude avec des concepts d'autres vocabulaires.

➡ Interoperabilité 

Pourquoi utiliser une ontologie pour nommer les variables ?

Problème !!

Plus le concept de la variable est précis, plus il est difficile de trouver des relations de similitude avec des concepts d'autres vocabulaires.

➡ Interopérabilité 

Solution

Décomposer le concept de variable complexe en concepts atomiques plus simples et définir les relations de similarité entre ces concepts plus simples et les concepts d'autres vocabulaires.

➡ Interopérabilité 

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

- Propriété (température, hauteur, ...)
- Entité : rôles [ObjectOfInterest, ContextObject, Matrix] (air, eau, sol, ...)
- Contrainte (profondeur, température, longueur d'onde, ...)

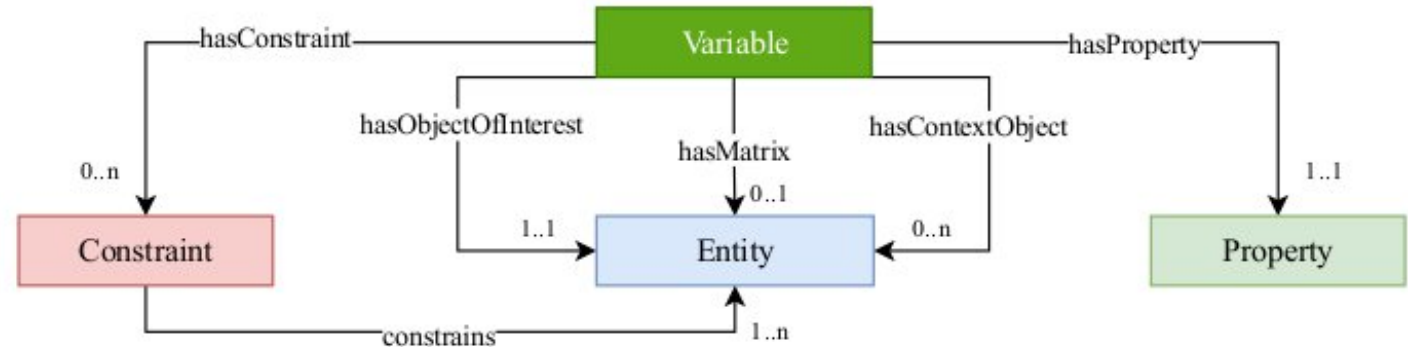
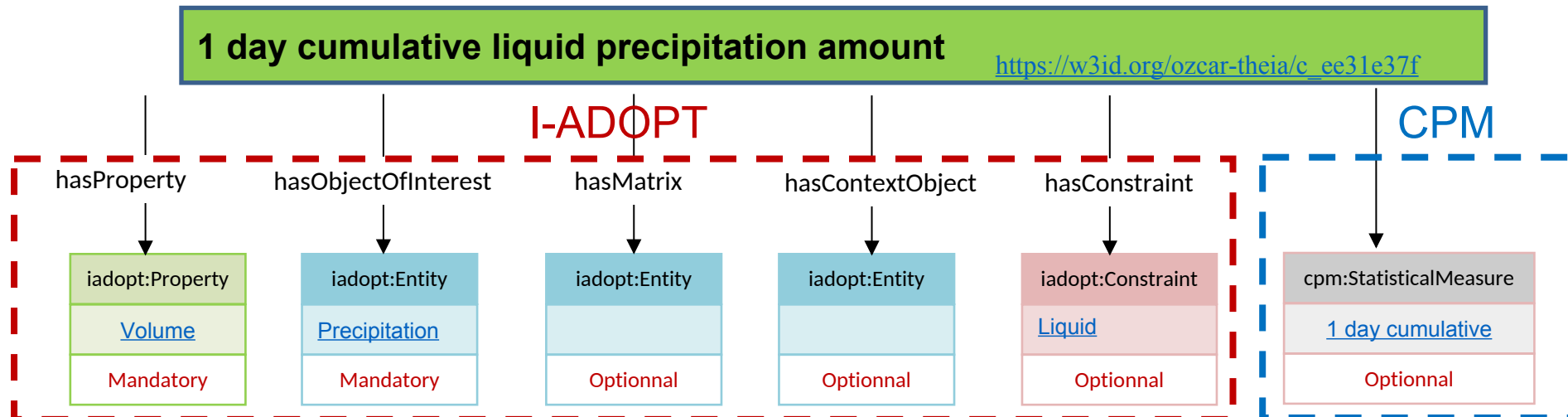


Figure 2: The I-ADOPT Framework.

Magnana et al., S4BioDiv2021, 2021



I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments
2. Identification des rôles
3. Annotation sémantique

Concentration of sulfate in wet flesh of ostrea edulis

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments

2. Identification des rôles

3. Annotation sémantique

Concentration of **sulfate** in **wet flesh** of **ostrea edulis**

- Décomposition de la variable en ses composants essentiels
- Une compréhension claire du phénomène que la variable doit décrire est essentiel pour permettre une modélisation précise

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments

2. Identification des rôles

3. Annotation sémantique

Concentration of **sulfate** in **wet flesh** of **ostrea edulis**

- Décomposition de la variable en ses composants essentiels
- Une compréhension claire du phénomène que la variable doit décrire est essentiel pour permettre une modélisation précise
- Pas d'unité, ni d'éléments décrivant la méthode de mesure ou des conditions d'expérimentation dans la décomposition I-ADOPT

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments
2. Identification des rôles
3. Annotation sémantique

Concentration of **sulfate** in **wet flesh** of **ostrea edulis**



Property

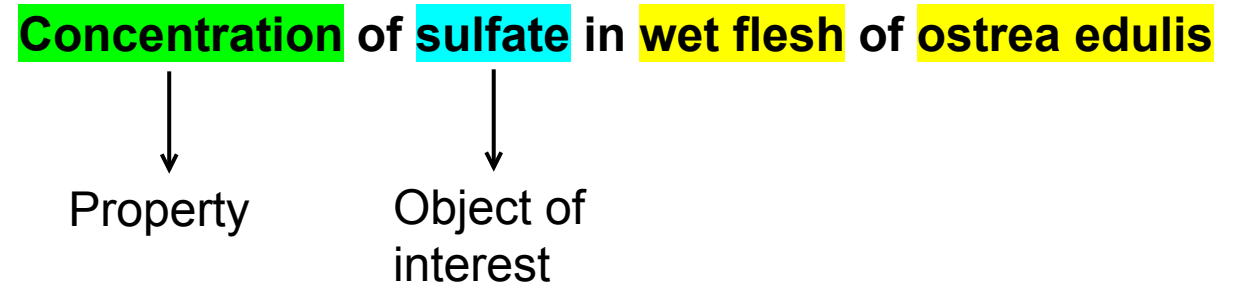
Identification de la propriété observée

- La variable produit-elle une valeur qualitative ? quantitative ?
- A quelle unité sont associées la valeur produite par la variable ?
- [Unit-to-property lookup](#)

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments
2. Identification des rôles
3. Annotation sémantique



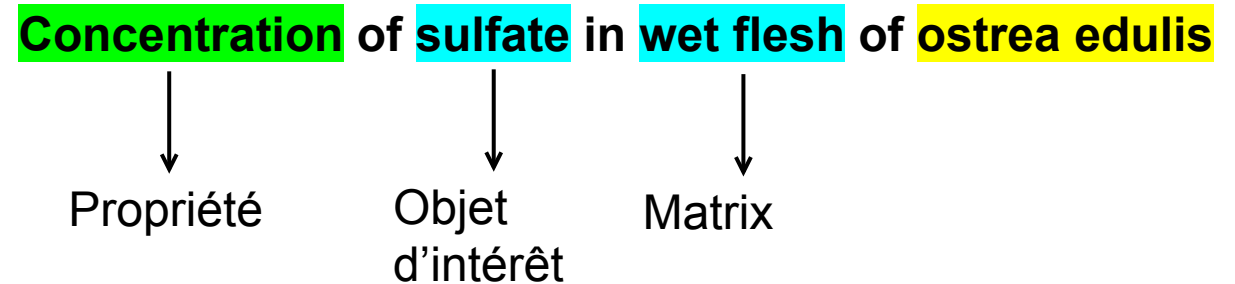
Identification de l'objet d'intérêt

- L'objet d'intérêt est l'entité dont la **propriété** est observée

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments
2. Identification des rôles
3. Annotation sémantique



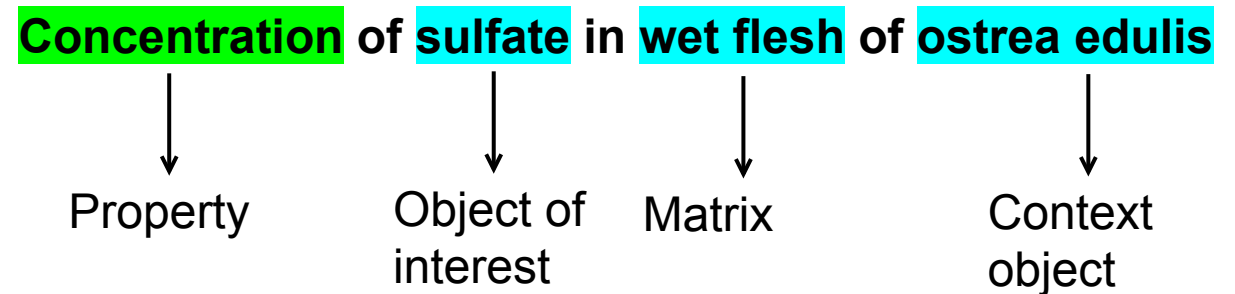
Identification de la matrice

- **La matrice** est l'entité dans laquelle l'objet d'intérêt est inclus

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments
2. Identification des rôles
3. Annotation sémantique



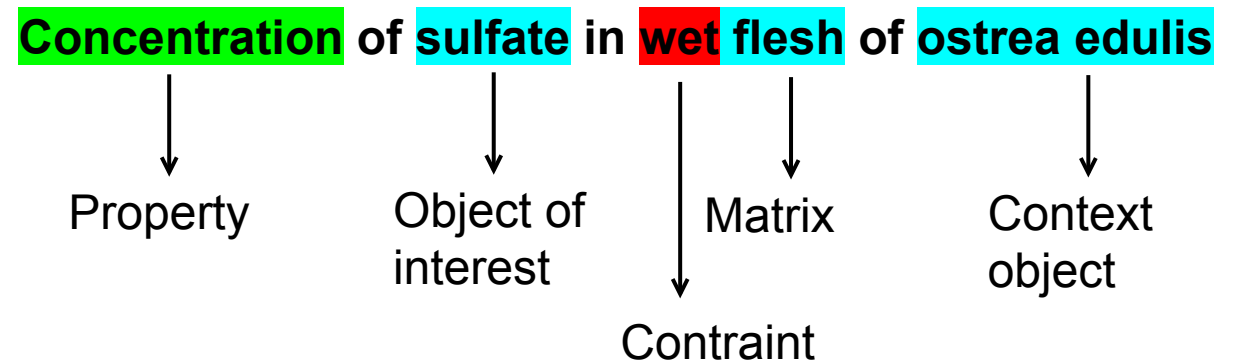
Identification des objets de contexte

- Les **objets de contexte** sont toutes les autres entités nécessaires à la description de la variable

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments
2. Identification des rôles
3. Annotation sémantique



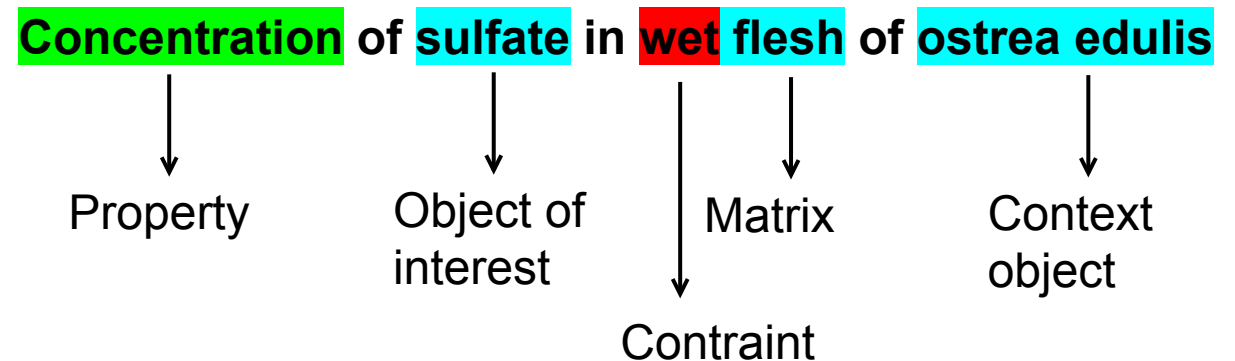
Identification des contraintes

- Les contraintes limitent la portée de la variable
- Passer en revue les entités **ObjectOfInterest**, **Matrix**, et **ContextObject** afin d'identifier si elles peuvent être découper plus finement

I-ADOPT framework ontology

Décomposition des noms de variables en concepts atomiques :

1. Identification des éléments
2. Identification des rôles
3. Annotation sémantique



- Lier chaque composant à un concept disponible depuis une terminologie
- Regarder dans les catalogues de terminologies comme OntoPortal
- [Liste des terminologies permettant de lier les composants I-ADOPT](#)

Les limitations d'I-ADOPT

- Les labels de variables sont souvent trop compliqués pour être utilisés directement pour la découverte de données. Nous avons mis en place notre propre « label simplifié » pour les fonctionnalités de recherche sur le portail web.

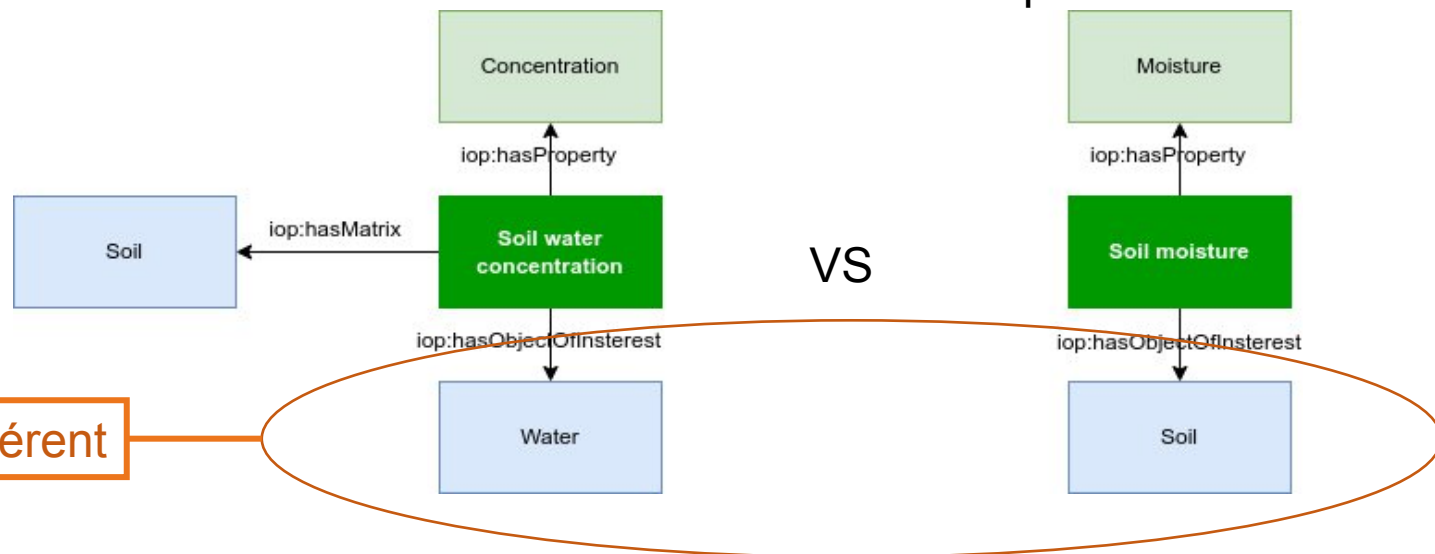
PREFERRED TERM

SIMPLIFIED LABEL

1 day cumulative liquid precipitation amount 

Precipitation amount

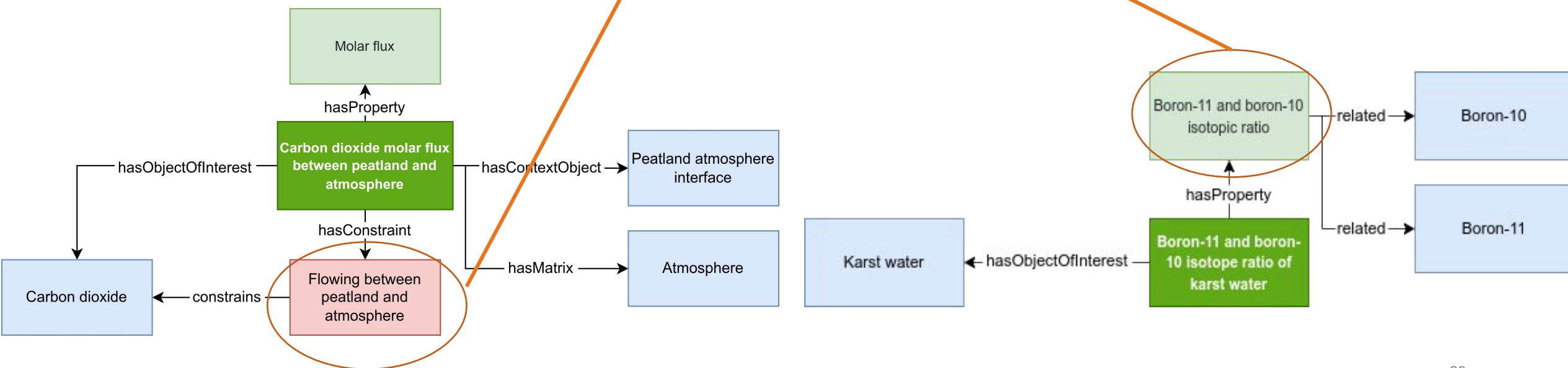
- Pour certaines variables, différentes implémentations sont possibles. Comment choisir l'une ou l'autre ? Comment déduire une relation de similitude entre des variables identiques modélisées différemment ?



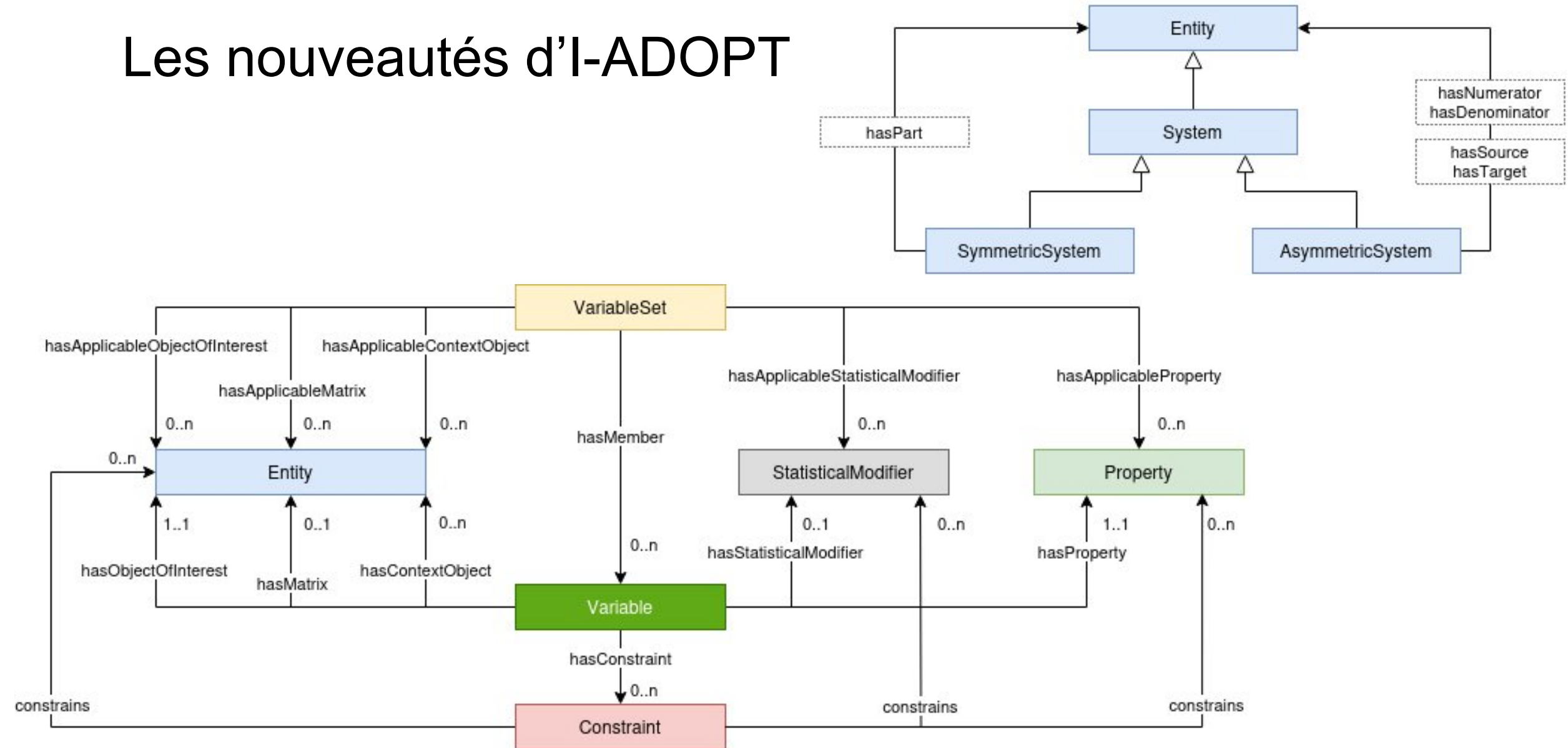
Les limitations d'I-ADOPT

- Difficulté à modéliser certaines variables telles qu'un flux entre deux compartiments (ex : flux de dioxyde de carbone entre la surface terrestre et l'atmosphère).
- Certains concepts doivent être très spécifiques pour décrire la variable. Ce qui peut entraîner une perte de la notion d'atomicité des concepts I-ADOPT.

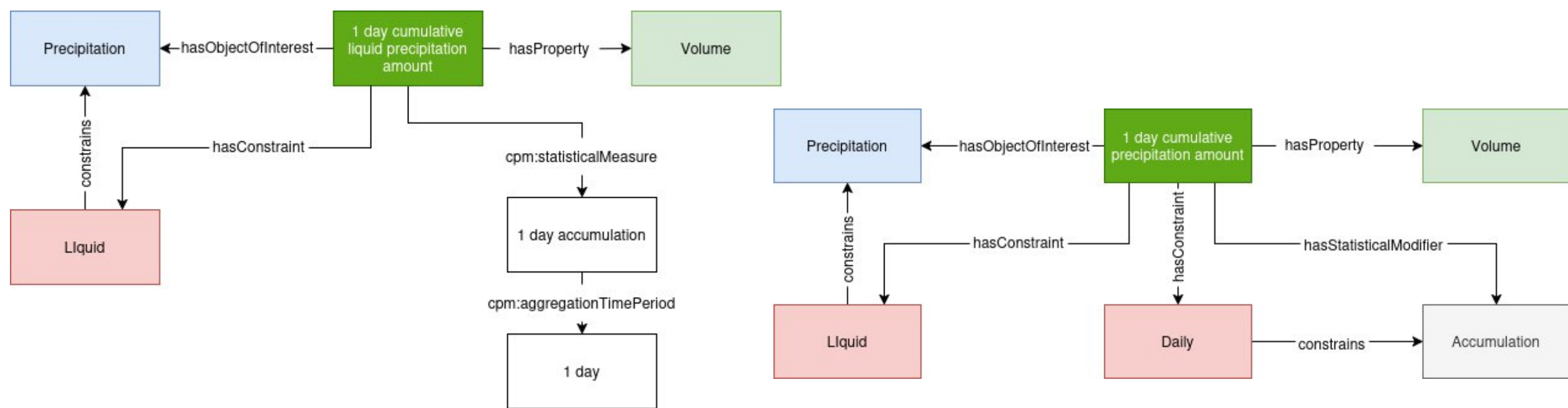
Des concepts très spécifiques remettant en cause la notion initiale de concept atomique



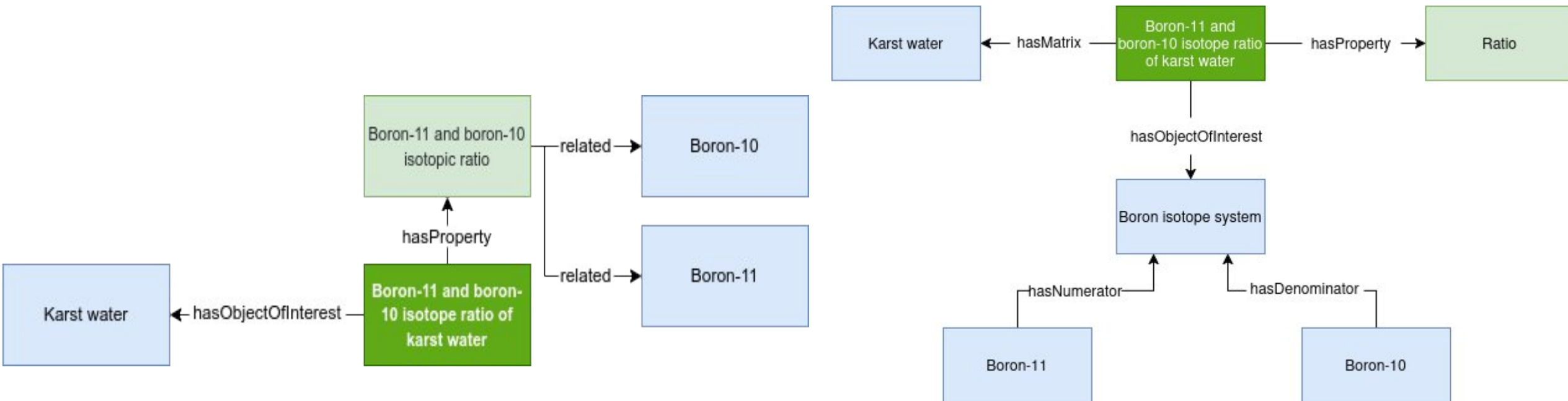
Les nouveautés d'I-ADOPT



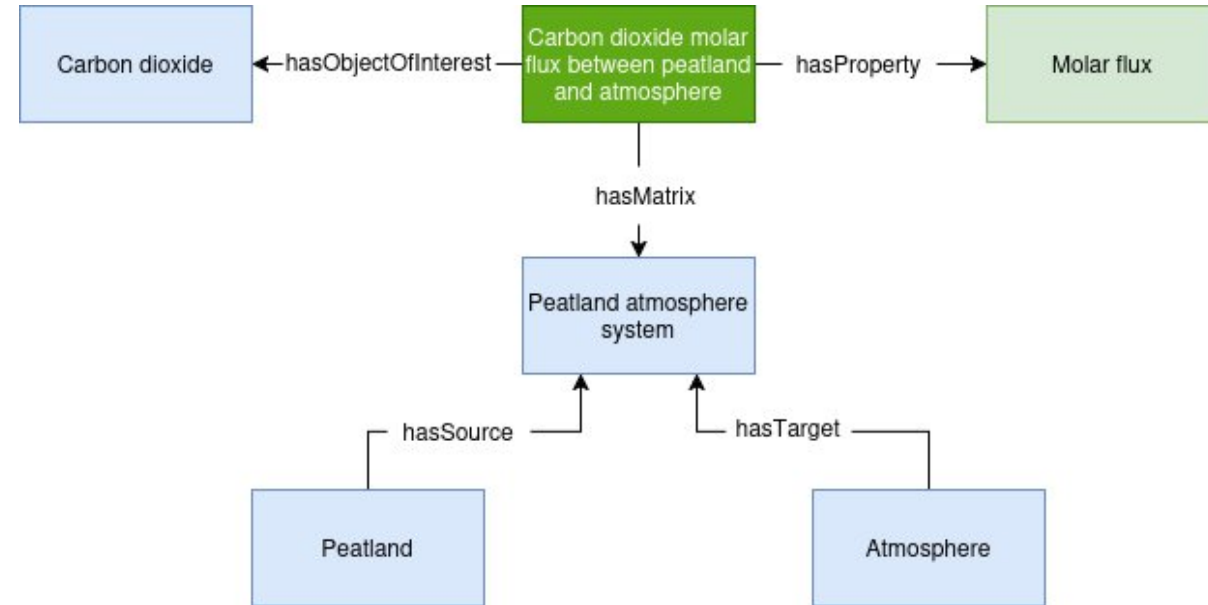
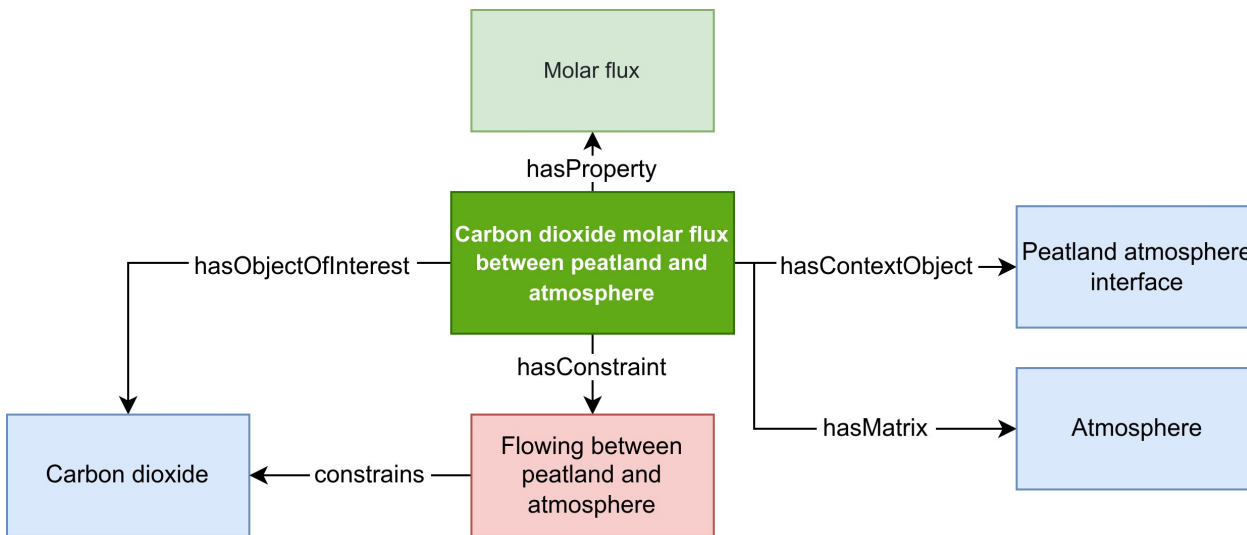
Les nouveautés d'I-ADOPT



Les nouveautés d'I-ADOPT



Les nouveautés d'I-ADOPT



Les bénéfices de l'utilisation d'I-ADOPT

- Permet une interprétation sans ambiguïté des données, et donc une meilleure réutilisation
- Favorise l'alignement des thésaurus internationaux dans le domaine.
- Favorise une meilleure interopérabilité sémantique des données au niveau national/international.

Permet d'ajouter des informations précises sur les variables mesurées par un jeu de données, en utilisant les champs « keyword » des standards de métadonnées qui n'incluent pas de description des variables mesurées.

The screenshot displays the eLTER Data portal interface. At the top, there is a dark blue navigation bar with the eLTER logo and links for 'Discover eLTER Data', 'Visualization Time Series Data', and 'Discovery Service'. Below this is a white search bar with a magnifying glass icon and the text 'Search', followed by a 'Map' button. A 'Back to search' button is also visible. The main content area shows search results for 'OZCAR-Theia variables thesaurus'. It includes a section for 'GEMET - INSPIRE themes, version 1.0' with a link to 'Environmental monitoring facilities'. Below this, a list of variables is shown, each with a magnifying glass icon: 'Dissolved aluminium mass concentration per unit volume in karst water', 'Dissolved arsenic mass concentration per unit volume in karst water', 'Dissolved barium mass concentration per unit volume in karst water', and 'Dissolved hydrogenocarbonate mass concentration per unit volume in karst water'. A dashed orange box highlights the 'OZCAR-Theia variables thesaurus' section and its list of variables.

Discover eLTER Data Visualization Time Series Data Discovery Service

Search Map

Back to search

Time series of type chemistry in Le Lez (Méditerranée) basin - MEDYCYSS observatory - KARST observatory network - OZCAR Critical Zone network Research Infrastructure

GEMET - INSPIRE themes, version 1.0

- [Environmental monitoring facilities](#)

OZCAR-Theia variables thesaurus

- [Dissolved aluminium mass concentration per unit volume in karst water](#)
- [Dissolved arsenic mass concentration per unit volume in karst water](#)
- [Dissolved barium mass concentration per unit volume in karst water](#)
- [Dissolved hydrogenocarbonate mass concentration per unit volume in karst water](#)

En savoir plus sur le projet:

Braud, I., Chaffard, V., Coussot, C., Galle, S., et al., 2020. Building the Information System of the French Critical Zone Observatories network: Theia/OZCAR-IS, Hydrological Sciences Journal, special issue “Data: opportunities and barriers”, <https://doi.org/10.1080/02626667.2020.1764568> .
Coussot et al., 2024. Implementing a new Research Data Alliance recommendation, the I-ADOPT framework, for the naming of environmental variables of continental surfaces. Earth Science Informatics. <https://dx.doi.org/10.1007/s12145-024-01373-9>

Pour accéder au portail et au thésaurus:

[Data portal : https://in-situ.theia-land.fr/](https://in-situ.theia-land.fr/)

[Thesaurus: https://w3id.org/ozcar-theia/](https://w3id.org/ozcar-theia/)

[Cataloguing CSW webservice: https://in-situ.theia-land.fr/geonetwork/srv/eng/csw?service=CSW&version=2.0.2&request=GetCapabilities](https://in-situ.theia-land.fr/geonetwork/srv/eng/csw?service=CSW&version=2.0.2&request=GetCapabilities)

Contacts:

charly.coussot@ird.fr

Veronique.Chaffard@univ-grenoble-alpes.fr

Isabelle.braud@inrae.fr

Sylvie.galle@ird.fr

Remerciements : Le projet contribue au pôle de données de surface continentale Theia et à l'infrastructure de recherche Data Terra. Il bénéficie du soutien de l'infrastructure de recherche OZCAR, de l'IRD, du CNRS et a été partiellement financé par le projet ANR FAIRTois (subvention ANR-19-DATA).

Merci pour votre attention:
Questions ?

