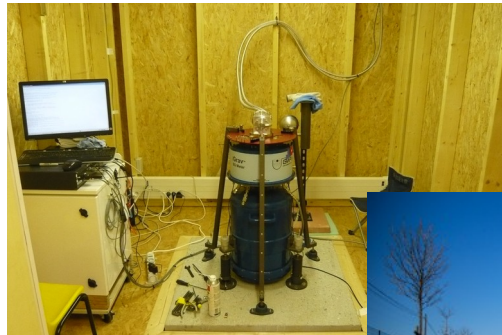


Pérennisation et interopérabilité des données d'observation

Exemple de l'Observatoire de
Recherche Montpelliérain de
l'Environnement

Ecosystème des données d'observation au CNRS INSU

1. Services Nationaux d'Observation (SNO)

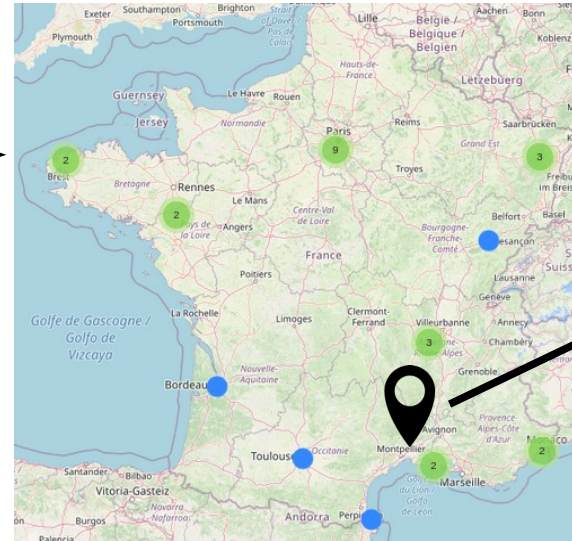


Gravimètre
SNO H+

Station météo
SNO KARST



2. Observatoires des Sciences de l'Univers (OSU)



OREMÉ

1 UAR
8 UMR



INRAE
Institut de Recherche
pour le Développement
FRANCE

3. Infrastructures de Recherche Pôles de données



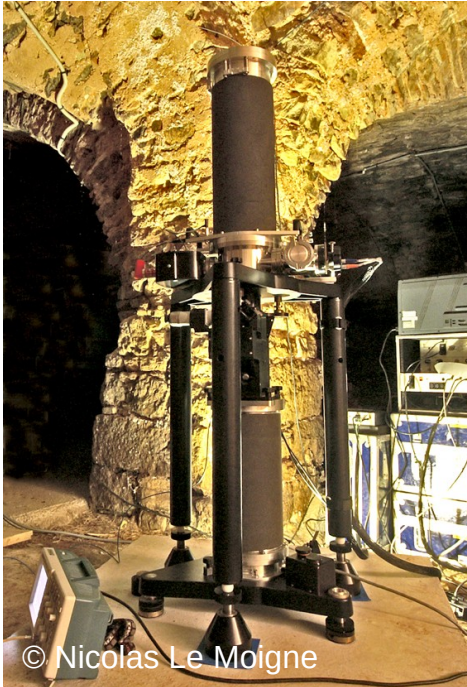
Réseau technologique
Séries Interopérables et
Systèmes de Traitement

Veille technologique & Formation



Des données d'observation hétérogènes

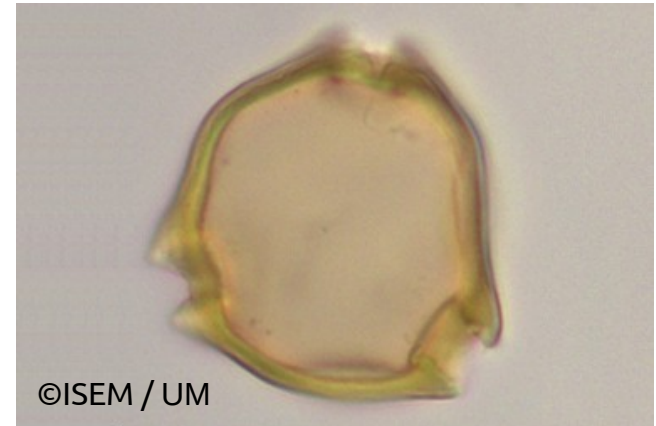
- Disciplines, types, méthodes d'acquisition, fréquences ..



Gravimétrie haute fréquence



Suivi de populations d'oiseaux



Collection de photos de pollen

- Formats : fichiers numériques, bases de données, données spatialisées, photos



Les données d'observation à l'OREME

Des besoins communs

Stocker

.. et sauvegarder

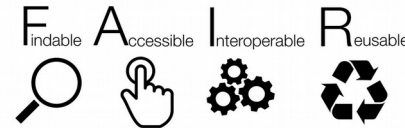
Gérer

=> ajouter, éditer, supprimer

Visualiser

Extraire

Diffuser



Des ressources limitées ..

- 2 ingénieur·es (depuis 2011)
- web : stagiaires IUT + 1 alternant L3 web

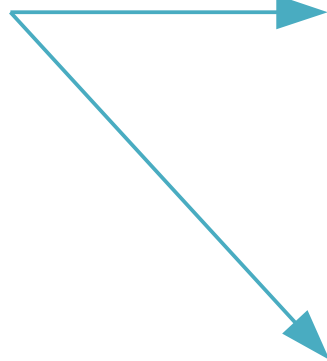
=> Développement d'outils et de méthodes **mutualisées** (tant que possible)

Méthode (1) : Structurer

Objectif : rendre les données interrogeables et cohérentes

Extraire, formater (formats ouverts), nettoyer, modéliser (BDD) ..

R, perl, foreign data wrapper (PG)



1 BDD relationnelle

PostgreSQL, PostGIS

33 schémas



Serveur cartographique



Méthode (2) : Décrire

date	valeur
2018-11-03	12

?

Objectif 1 : Rendre les données **compréhensibles**, non ambiguës



Objectif 2 : **Lier** les données entre elles



Méthode (2) : décrire

Décrire **finement** – utiliser des descripteurs **communs**

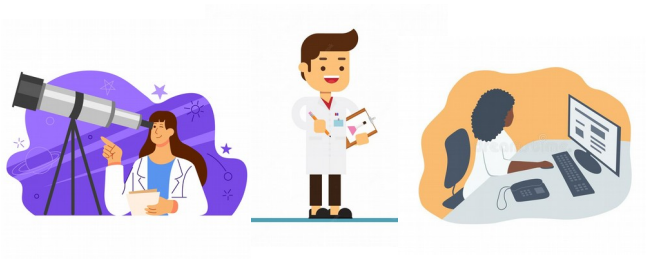
Question	Descripteurs
Qui ?	<ul style="list-style-type: none">• personnes auteures / créatrices de la ressource• personnes collaboratrices• personnes contact• gestionnaires de données• organisations ..
Pourquoi ?	<ul style="list-style-type: none">• titre• résumé• cadre : projet, etc
Quand ?	<ul style="list-style-type: none">• date de début / fin, de création ..
Où ?	<ul style="list-style-type: none">• emprise géographique• pays, ville• géographie, écosystème, ..

Question	Descripteurs
Comment .. a-t-on acquis ce jeu de données ?	<ul style="list-style-type: none">• protocole• méthode• outils / instruments• fréquence d'acquisition• financements
Comment .. peut-on ré-utiliser ce jeu de données ?	<ul style="list-style-type: none">• licence• conditions d'utilisation• lien d'accès à la ressource• ressources en lien
Quoi ?	<ul style="list-style-type: none">• thématique / discipline / domaine• variable mesurée• unité de mesure• taxon, gène, ..• ..

Méthode (2) : Décrire

date	valeur
2018-11-03	12

I. Dupont a mesuré une température de l'eau brute de 12°C le 03/11/2018 dans le Lez avec une sonde CTD (mots-clés : karst, hydrologie, crue)



Méthode (2) : Décrire

Intégrer des **standards** pour **enrichir** les données

- référentiels, vocabulaires contrôlés, thésaurus, ...

I.dupont@cnrs.fr,
<https://orcid.org/0000-0001-9832-0000>

https://gcmd.earthdata.nasa.gov/KeywordViewer/scheme/all/61594015-4ab4-4b38-ae4f-e31a4757b065?gtm_keyword=WATER%20TEMPERATURE>m_scheme=Earth%20Science

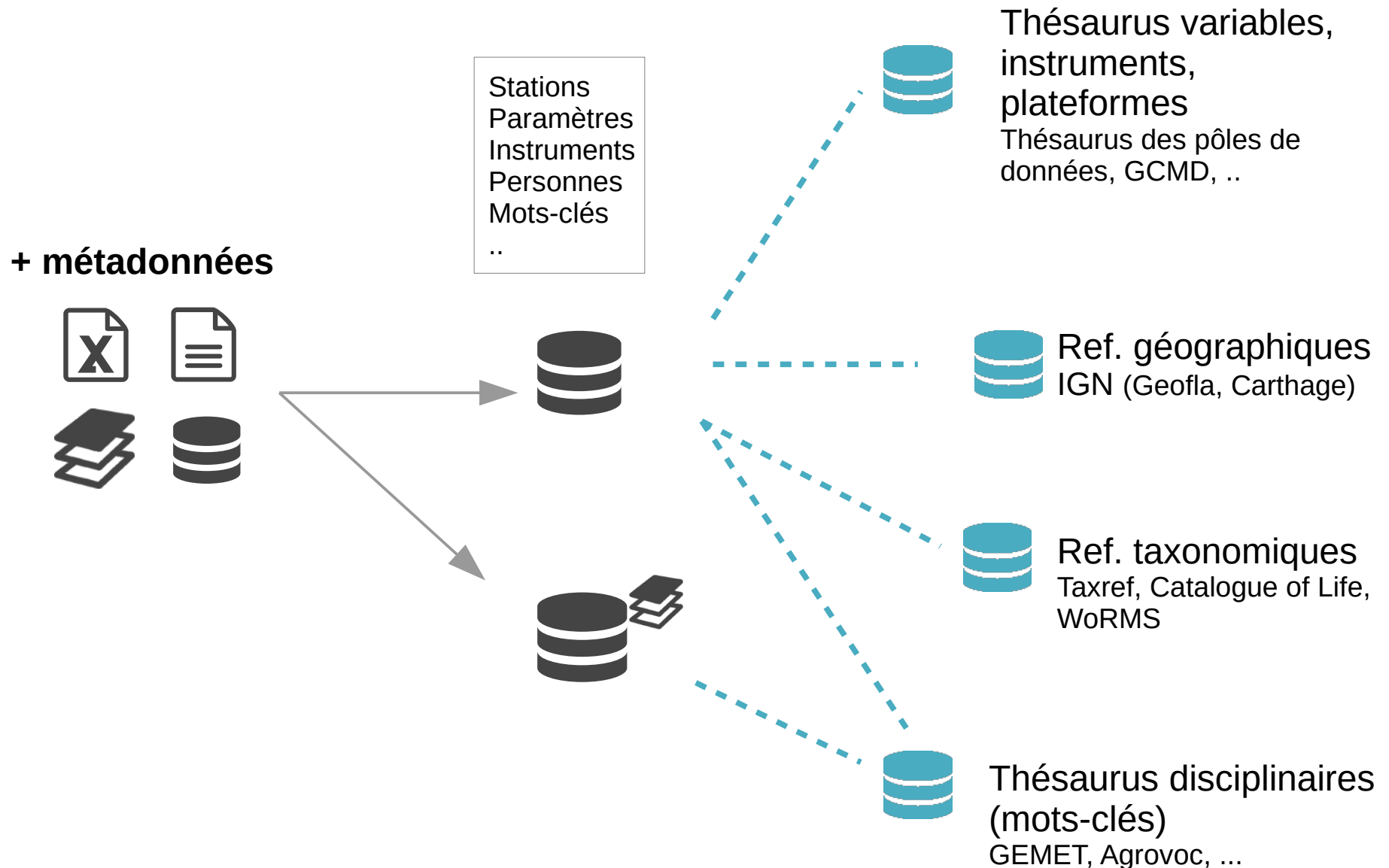
I. Dupont a mesuré une **température de l'eau** brute de 12°C le 03/11/2018 dans **le Lez** avec une **sonde CTD** (mot-clé : **karst**)

http://www.sandre.eaufrance.fr/geo/CoursEau_Carthage2017/O04-0400

<https://www.eionet.europa.eu/gemet/en/concept/4545>

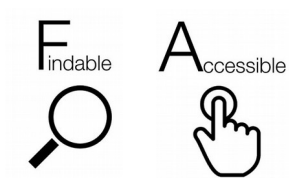
https://gcmd.earthdata.nasa.gov/KeywordViewer/scheme/all/01cc0beb-7c9a-40ed-ad86-0661b41aee53?gtm_keyword=CTD>m_scheme=instruments

Méthode (2) : Décrire



Méthode (3) : Cataloguer, identifier

Objectifs : faire connaître les données, les référencer dans des catalogues, les rendre identifiables

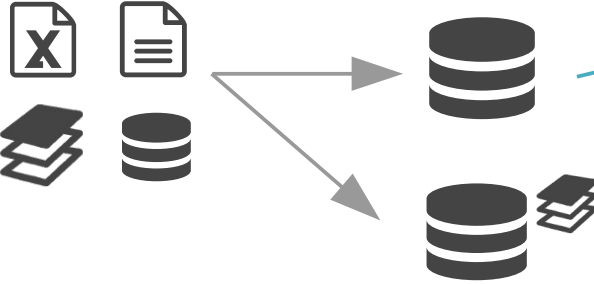


Métadonnées : **formats** et **protocoles** standards, ouverts et opérables par des machines

Identifiants uniques

Méthode (3) : Cataloguer, identifier

+ Métadonnées



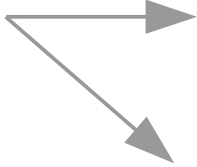
SQL, R


Identifiants : DOI - *API DataCite*



Méthode (3) : Cataloguer, identifier

+ Métadonnées



SQL, R

SQL, R geoflow

Identifiants : DOI - API DataCite

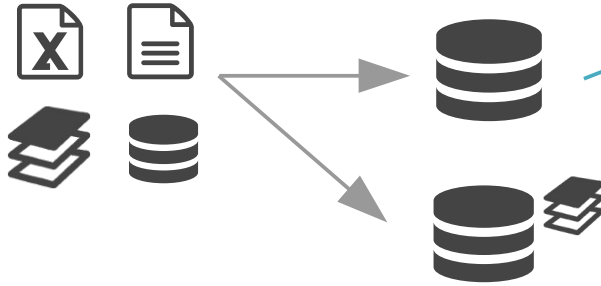


Catalogue de données
GeoNetwork
Norme ISO 19115 / Inspire



Méthode (3) : Cataloguer, identifier

+ Métadonnées



SQL, R geoflow

Catalogue de données

GeoNetwork

Norme ISO 19115 / Inspire



Flux CSW

(Catalogue Service for the Web - OGC)

Catalogues (indores ..)

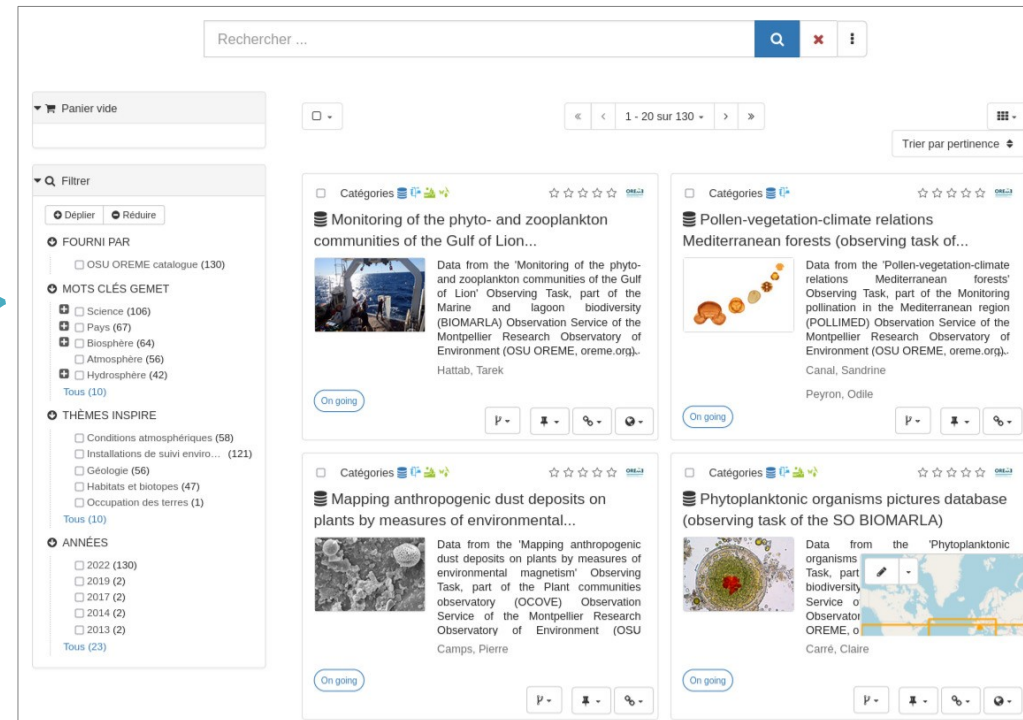
Infrastructures de
données géographiques
GeoCMS, GeOrchestra

data.gouv.fr



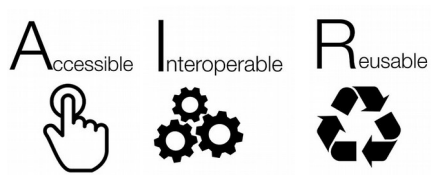
Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

Moissonnage
DCAT



catalogue.oreme.org

Objectifs : ouvrir les données, les rendre accessibles aux humains et aux machines



Données : **Formats** et **protocoles** standards, ouverts et opérables par des machines

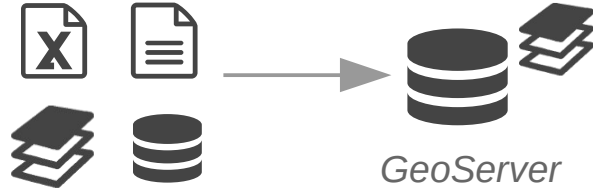
Licences ouvertes :



Méthode (4) – Diffuser - visualiser - extraire

Données cartographiques

+ Métadonnées

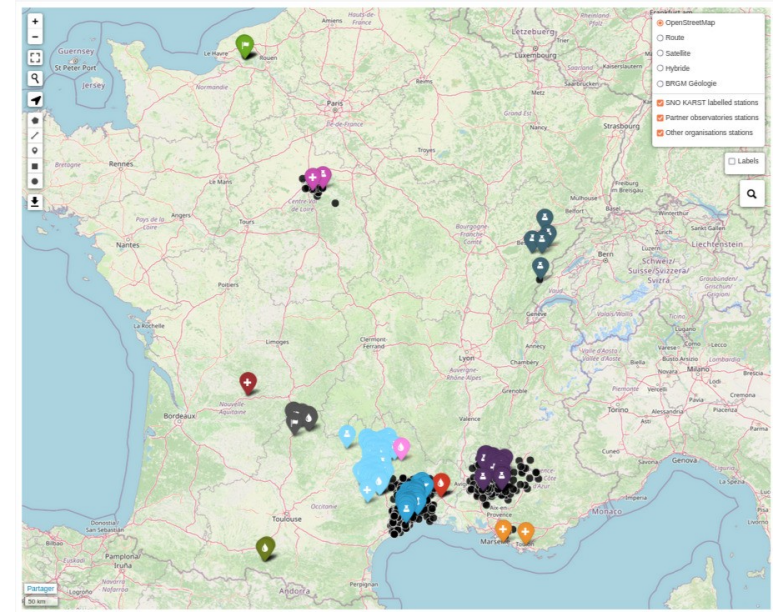


Flux WMS / WFS

(Web Map Service / Web Feature Service
- OGC)



Librairies de webmapping



Librairie *leaflet*, data.oreme.org

SIG “bureau”



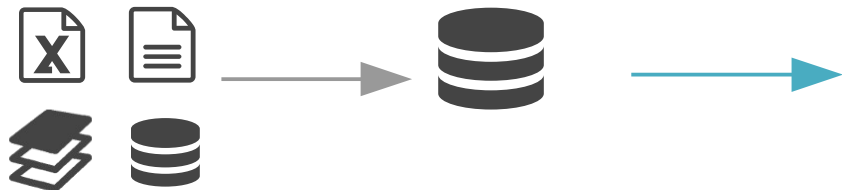
Infrastructures de données
géographiques

GeoCMS, GeOrchestra

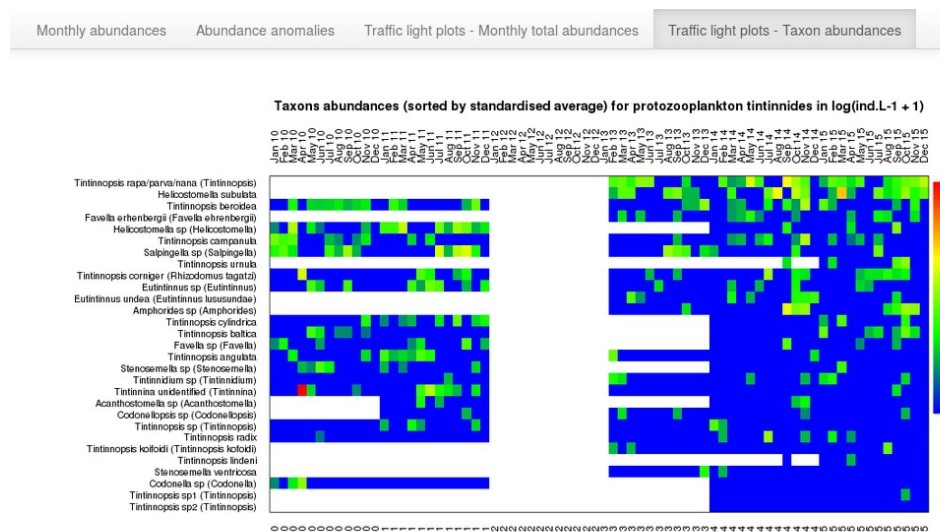
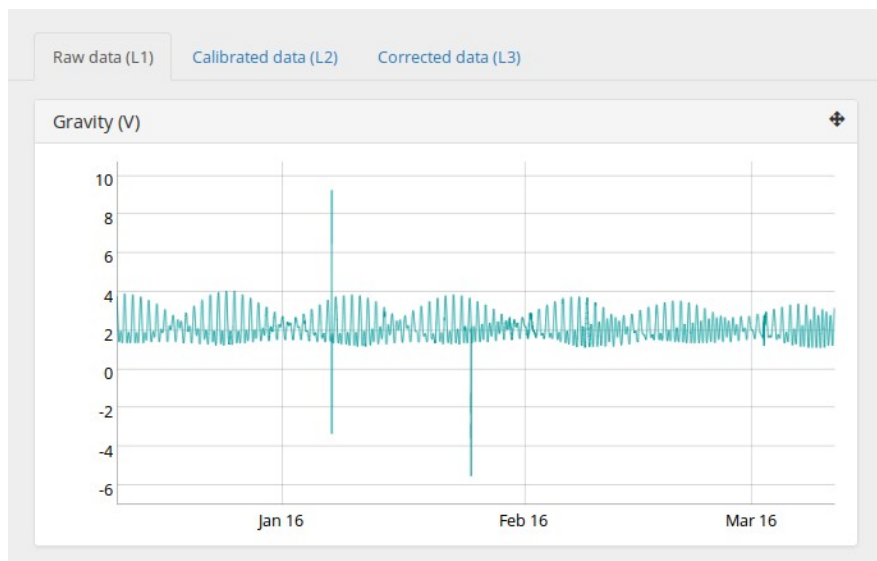
Méthode (4) – Diffuser – visualiser - extraire

Données numériques / séries temporelles

+ Métadonnées



A destination des humains
Interfaces web “maison”
data.oreme.org



Outil générique de visualisation et
d'export CSV
Librairie *dygraphs*

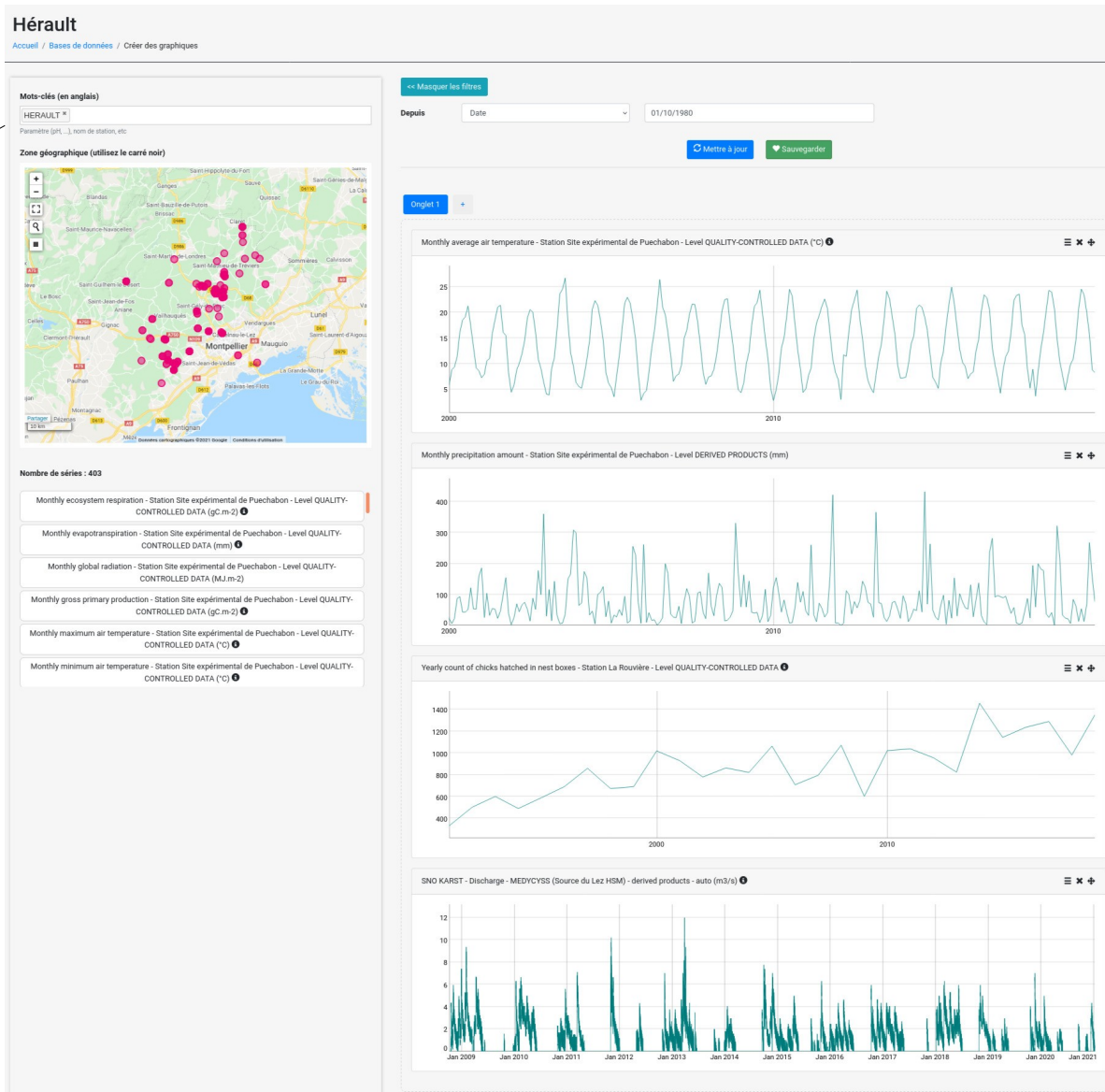


Outils spécifiques de visualisation
R shiny

Méthode (4) – Diffuser – visualiser - extraire

Outil générique de recherche et de visualisation - Librairies *leaflet* + *dygraphs*

Mots-clés
Paramètre
Projet
Thématique
Ville
..



Suivi site
Puéchabon



Suivi mésanges

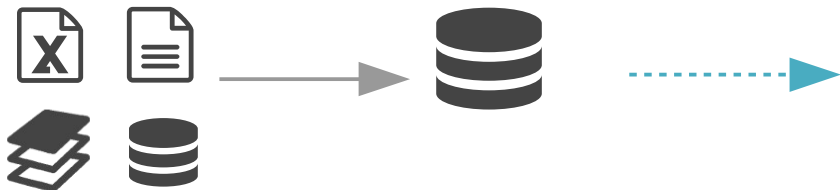


SNO KARST

Méthode (4) – Diffuser – visualiser - extraire

Données numériques / séries temporelles

+ Métadonnées



A destination des humains et des machines

Serveur de séries temporelles

istSOS, 52North, FROST

istSOS²



Flux Sensor Observation Service (SOS)

Flux SensorThings

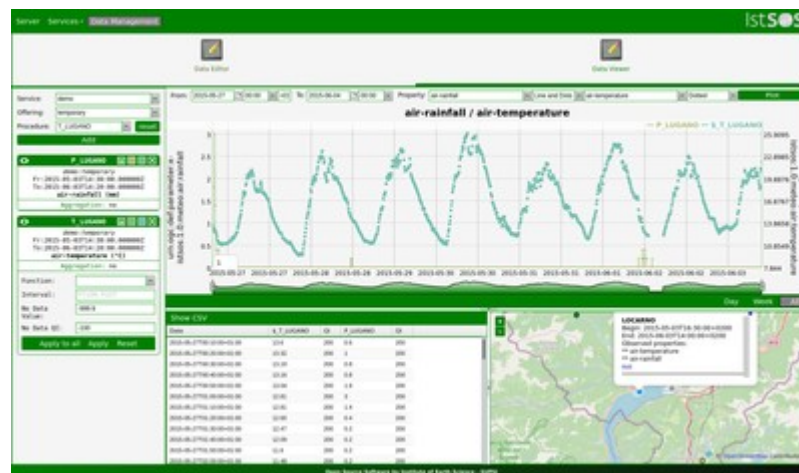
(OGC)

<https://sist.cnrs.fr/2021-istsos/>

<https://sist.cnrs.fr/anf-2021-sos-52north/>

<https://sist.cnrs.fr/webinaire-sensorthings>

?



Ex de client SOS : istSOS

Construire (et maintenir) un SI prend des années

- Besoin de personnel SI **permanent** et de **temps** disponible pour les équipes scientifiques !

Alimenter les BDD, c'est long ..

- Coller au plus près des formats des scientifiques
- Développer des outils à valeur ajoutée pour les scientifiques (visualisations, export auto vers d'autres BDD ..)

Décrire finement les données est une clé de voûte

- Indexation, interfaces de recherche
- Interopérabilité

Les observatoires, c'est cool

- Partager / ré-utiliser des méthodes et des outils => travailler moins pour interopérer plus

Réseau Séries Interopérables et Systèmes de Traitement



Gestionnaires de données
d'observation

OSU, Unités CNRS, Autres
organismes de recherche

Partager des technologies
et savoir-faire

Promouvoir les **bonnes
pratiques** et
l'interopérabilité

- <http://sist.cnrs.fr>
- sist@services.cnrs.fr
- 250 membres

Séminaires et formations

- Catalogage (GeoNetwork), Serveur carto (GeoServer), IDG (GeoCMS), R geoflow
- Plateformes de diffusion Thredds et erddap – format netCDF
- Standards et outils SOS et SensorThing (istSOS, 52North)
- Géomatique
- Vocabulaires et référentiels

