

# Pérennisation et mise à disposition des données de l'Observatoire de Recherche Méditerranéen de l'Environnement

© Arnaud Martin



# Plan

- Contexte
- Méthodologie
- Construction des bases de données
- Diffusion des données (si temps)
- Bilan



# Contexte



# L'OSU OREME

- Unité Mixte de Service



UNIVERSITÉ  
DE MONTPELLIER

- Structure transverse

- 40 Tâches d'Observation
- 17 Services d'Observation
- 6 laboratoires



# Données d'observation : Spécificités

## ▪ Persistance critique

Impossible de re-générer les données

→ Stocker, sauvegarder, archiver

→ Surveiller (interruption / dysfonctionnement chaîne de traitement)

## ▪ Interprétation et traitement de la donnée

Chaine de traitement : calibration, correction, validation, ...

→ Décrire les protocoles, les instruments, les traitements ..

→ Qualification de la donnée

## ▪ Besoin de valoriser, partager, rendre visible

→ Cataloguer, référencer, identifier

→ Décrire les jeux de données

→ Diffuser : visualiser & exporter



# Données d'observation : Spécificités

## ▪ Besoin d'interopérabilité

- Au niveau applicatif (traitements, modèles, visualisation, ...)
- Avec d'autres bases de données (croisement, appartenance)
  - Respecter les normes et standards (OGC, formats, métadonnées ...)

## ▪ Besoin d'outils de visualisation & navigation

- Interpréter, corriger, corréler, surveiller
- 2 dimensions : temps & espace
  - Espace : outils SIG
  - Temps : graphiques (séries temporelles)



# Un contexte hétérogène

- Disciplines variées

Hydrologie, biologie, physique, ...

- Types de données

Analyses génétiques ou physico-chimiques, observations visuelles, photos, données carto ..

- Méthodes d'acquisition

Manuelle / visuelle, capteurs, ..

- Fréquences

Qq fois / an  $\rightarrow$  10 Hz

- Volumes

Mo  $\rightarrow$  To

- Besoins, méthodes

Qualité, visibilité, outils, ...



# Et des moyens restreints

- Personnel OSU OREME
  - 1 IR permanent : chef de projet, BDD, admin système
  - 1 IE contractuel : BDD, SIG, web
  - ½ IE contractuel : design web
- Contacts dans les laboratoires





# Méthodologie



# Des bases de données

- Relationnelles

- Structuration
- Persistence
- Intégrité
- Robustesse
- Requêtage, vues
- Interopérabilité

- Spécifiques mais interopérables

- Pas de solution générique
- Outils adaptés à chaque spécificité
- Capitalisation (savoirs-faire, outils, ...)



# Mutualisation & Pragmatisme

- Mutualisation

- Des développements : BDD, outils, web
- Des compétences : expériences, connaissances ...
- De la gestion : admin serveurs, SAN

- Pragmatisme

- Ré-utilisabilité des développements à moyen terme



# Identifier ...

- Les contacts
- Les données d'intérêt
  - Types, formats
  - Acquisition : méthodes, fréquences, volumes
  - Standards du domaine (référentiels, vocabulaires)
- Les traitements
  - Filtres, corrections, traitements ...
  - Niveaux de données (corrigées, traitées, validées, ...)
- Les méthodes de diffusion
  - Graphiques, export numérique, cartes, ..
  - Ouverture des données (embargo, conditions d'utilisation ..)



# Développer ...

- Les bases de données
  - Conception, description
  - Alimentation
  - Traitements
- Les outils de diffusion
  - Visualisation, export
  - Métadonnées, catalogage



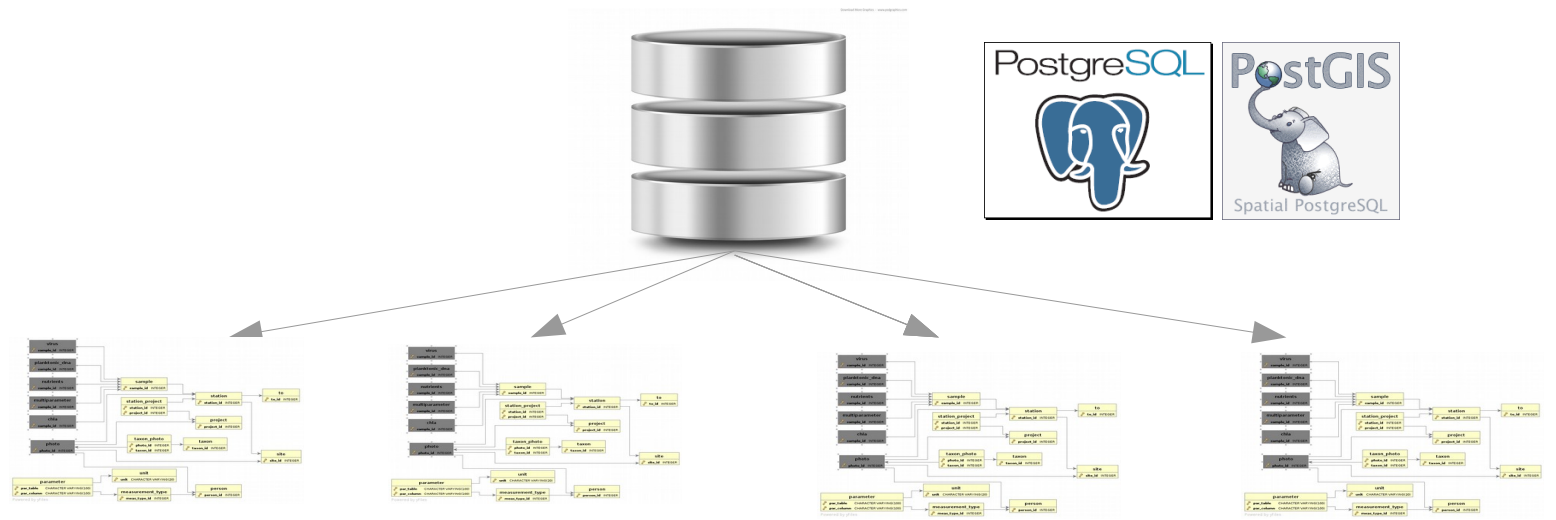
A  
L  
L  
E  
R  
-  
R  
E  
T  
O  
U  
R  
S



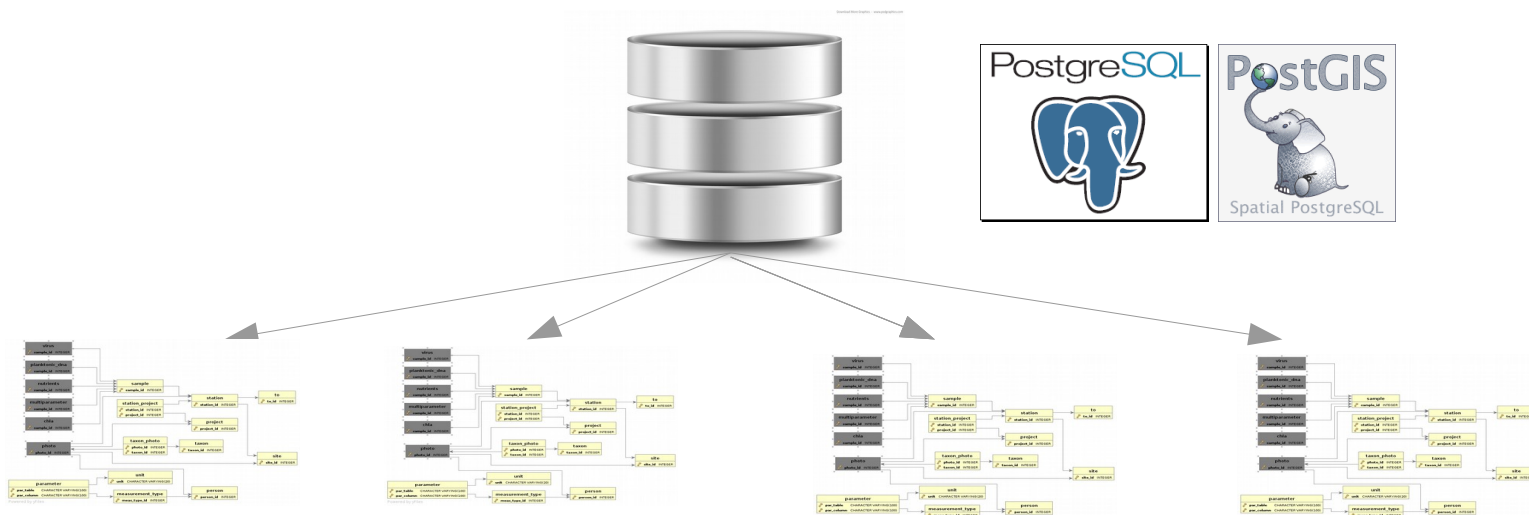
# Construction des BDD



# Outils et modèles



# Outils et modèles



DONNEES

DESCRIPTEURS

STANDARDS (référentiels, thésaurus, ontologies, ..)

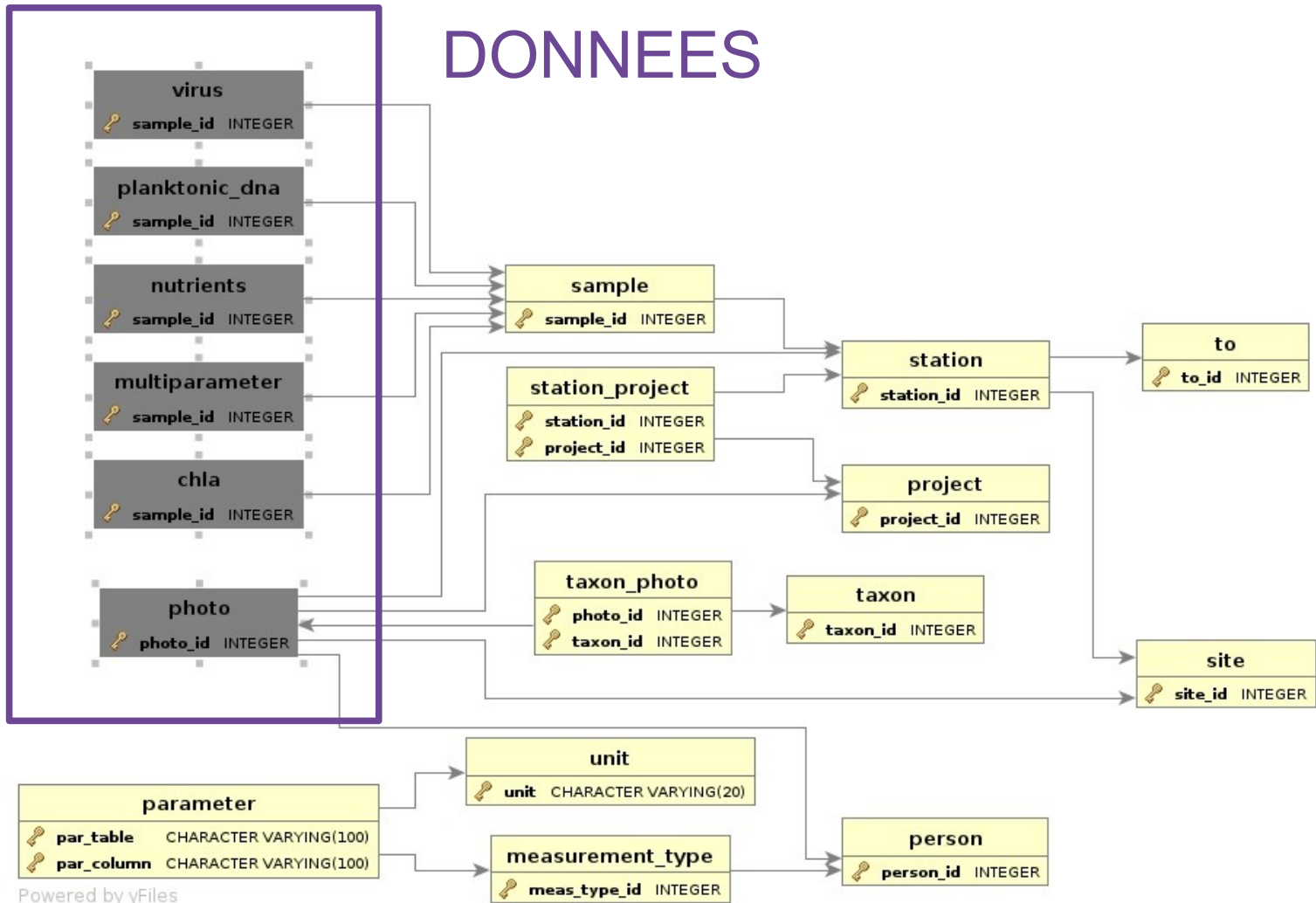
METADONNEES standards / normées

Ré-utilisabilité / Interopérabilité

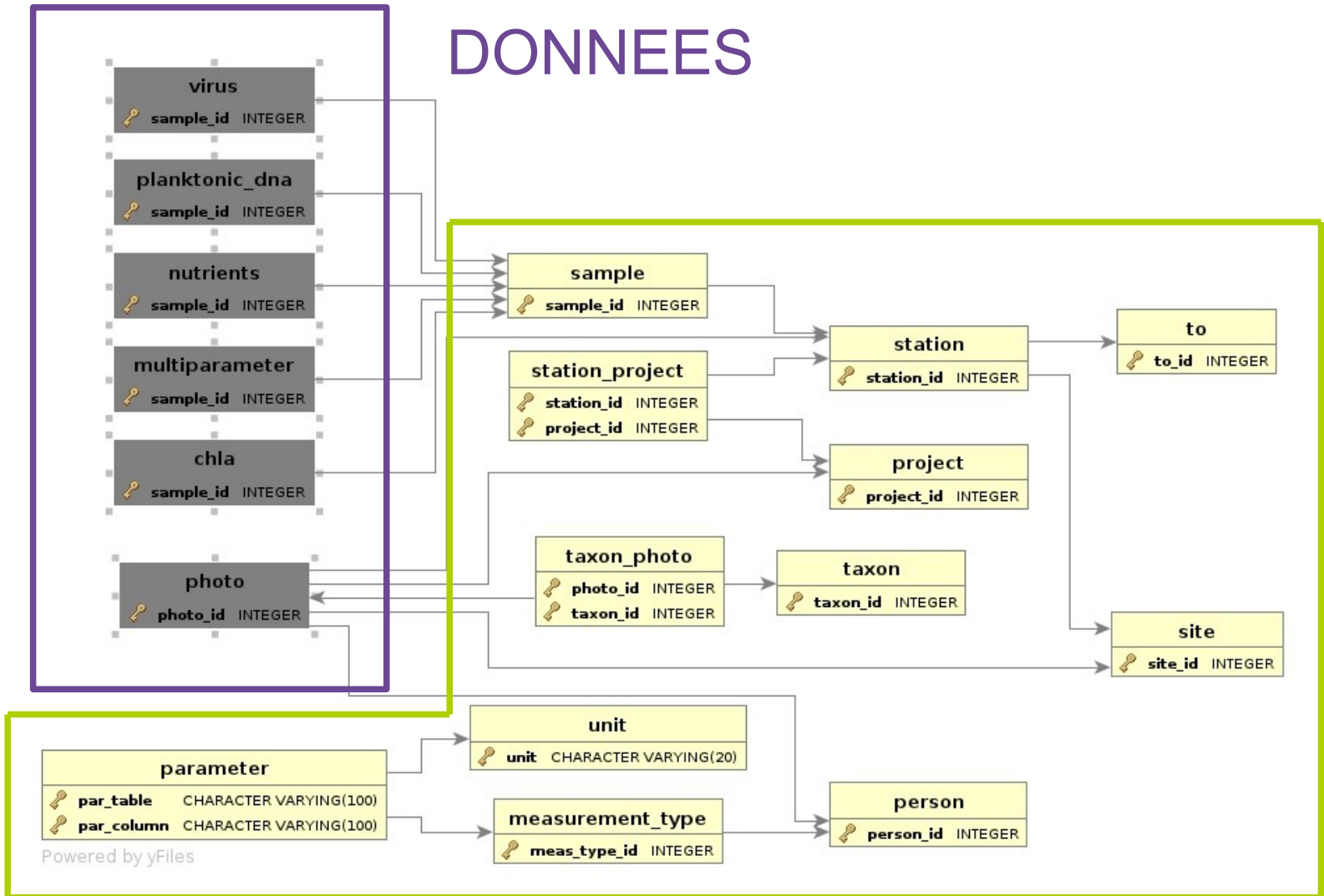




# Outils et modèles

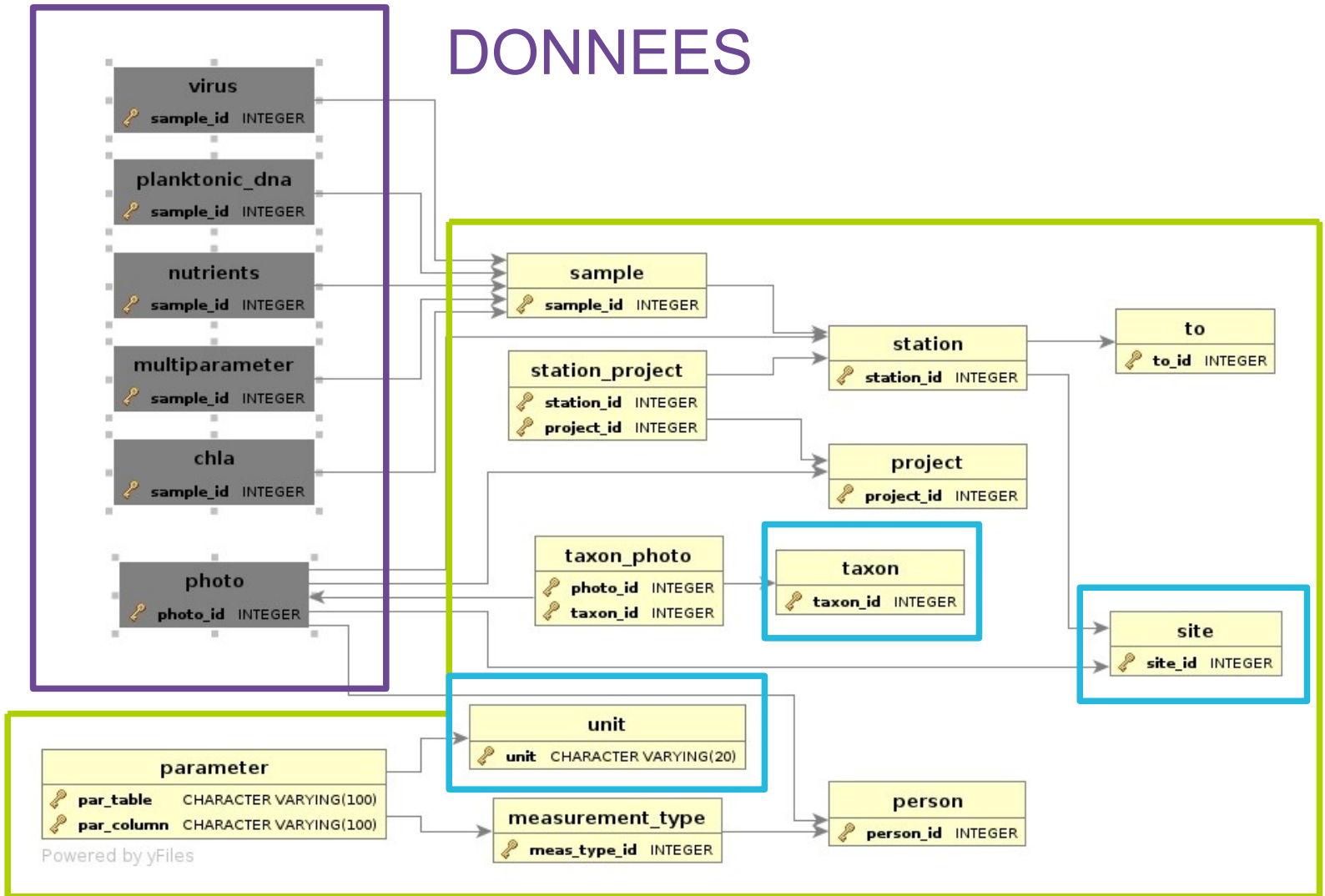


# Outils et modèles



## DESCRIPTEURS

# Outils et modèles



DESCRIPTEURS

STANDARDS

# Quelles données ?



*Niveaux 1, 2 ...*

Corrigées, calculées, ...,  
synthétisées, validées

*Niveau 0*

Données brutes

# Quelles données ?



*Niveaux 1, 2 ...*

Corrigées, calculées, ...,  
synthétisées, validées

- Motivation : Interpréter, analyser
- Publics visés : Métier + public

*Niveau 0*

Données brutes

- Motivation : Pérenniser, re-générer les données traitées
- Public visé : Experts

# Quelles données ?

- Notre choix : stocker tous les niveaux de données en base de données
  - Niveau 0 : tables
  - Niveaux 1, 2, et + : tables, vues matérialisées ou non



# Décrire et enrichir

- Décrire

Qui, pourquoi, où, comment, quoi ?

→ Assurer compréhension et ré-utilisabilité

- Utiliser des standards

Référentiels, vocabulaires contrôlés, thésaurus, mots-clés, ontologies

→ Garantir interopérabilité

→ Enrichir les données, déduire de l'information



# Décrire et enrichir

	Concepts	Standards, référentiels, vocabulaires, thésaurus ..
Qui	Personnes, Organisations	
Où	Sites, stations Géographie	BD Geofla BD Carthage
Pourquoi	Projets, Cadres métier	GEMET, EnvThes
Comment	Types de mesure Types d'instruments	GCMD
Quoi	Paramètres / variables Unités	Unit SI
Descripteurs métier	Taxons, Gènes, ...	WoRMS, NCBI





# Décrire et enrichir

	Concepts	Standards, référentiels, vocabulaires, thésaurus ..
Qui	Personnes, Organisations	
Où	Sites, stations Géographie	BD Geofla BD Carthage Geonames ISO3166-1 (pays)
Pourquoi	Projets, Cadres métier	GEMET, EnvThes, AGROVOC
Comment	Types de mesure Types d'instruments	GCMD Unit SI Unit Ontology, OM
Quoi	Paramètres / variables Unités	Ontologies & thésaurus métier / de haut niveau (PO, O&M, OBOE, ..)
Descripteurs métier	Taxons, Gènes, ...	WoRMS, NCBI

Linked Open Vocabularies : **LOV**

# Décrire et enrichir

- Ex1 : Utilisation du référentiel **WoRMS** (Registre mondial des espèces marines)

→ Observations de communautés microbiennes

→ WoRMS : identifiant de taxon  
= AphiaID

→ Fiches descriptives sur WoRMS et autres (MarBEF)

<http://www.marinespecies.org/aphia.php?p=taxdetails&id=106656>

→ Web services : récupération de la classification



# Décrire et enrichir

- Ex1 : Utilisation du référentiel **WoRMS** (Registre mondial des espèces marines)

**Home** | **About** | **Search taxa** | **Taxon tree** | **Literature** | **Distribution** | **Specimens** | **Match taxa** | **Editors** | **Statistics** | **Users** | **Webservice** | **Photogallery** | **Info downloads** | **Sponsors** | **Glossary** | **Manual** | **Log in**

**WoRMS taxon details**

▼ *Oithona similis* Claus, 1866  
AphiaID: 106656

**Classification:** Biota > ▼ Animalia (Kingdom) > ▼ Arthropoda (Phylum) > ▼ Crustacea (Subphylum) > ▼ Maxillopoda (Class) > ▼ Copepoda (Subclass) > ▼ Neocopepoda (Infraclass) > ▼ Podoplea (Superorder) > ▼ Cyclopoida (Order) > ▼ Oithonidae (Family) > ▼ *Oithona* (Genus)

**Status:** accepted  
**Rank:** Species  
**Parent:** ▼ *Oithona* Baird, 1843

**Synonymised names:**  
 ▼ *Oithona helgolandica* (Claus, 1863)  
 ▼ *Oithona pygmaea* Boeck, 1865 (listed as probable synonym by Giesbrecht, 1892)  
 ▼ *Oithona spinifrons* Boeck, 1865  
 ▼ *Oithonella helgolandica* Claus, 1863 (synonym)

**Sources:**  
**basis of record** Boxshall, G. (2001). Copepoda (excl. Harpacticoida), *in*: Costello, M.J. et al. (Ed.) (2001). *European register of marine species: a check-list of the marine species in Europe and a bibliography of guides to their identification*. Collection Patrimoines Naturels, 50: pp. 252-268 (look up in IMIS) [details]

[show all]

**Vernacular Names**

Language	Name
Polish	? ojtona [details]

**Environment:** marine, brackish, fresh, ~~terrestrial~~

**Distribution:** FROM EDITOR OR GLOBAL SPECIES DATABASE

**North Atlantic Ocean**  
 ▼ European waters (ERMS scope) [details]  
 ▼ United Kingdom Exclusive Economic Zone [details]

FROM REGIONAL OR THEMATIC SPECIES DATABASE

(no group)  
 ▼ Cosmopolitan (World Oceans) [details]

FROM OTHER SOURCES

**Arctic Ocean**  
 ? Arctic Ocean [details]

**Baltic Sea**  
 ? Polish Exclusive Economic Zone [details]

**Barents Sea**  
 ? Barents Sea [details]

**Beaufort Sea**  
 ? Canada Basin [details]



# Décrire et enrichir

- Ex1 : Utilisation du référentiel **WoRMS** (Registre mondial des espèces marines)

Filtrer par taxon :

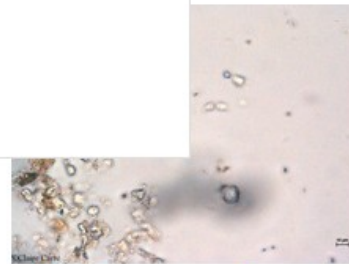
Taxons

120 taxon(s)

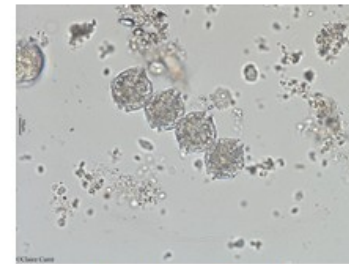
- Acanthoica
- Achnanthes**
- Akashiwo
- Alexandrium
- Amphora
- Ankistrodesmus
- Aphanocapsa
- Aphanothece
- Asterionellopsis



*Acanthoica quattropsina* (Species)



*Akashiwo sanguinea* (Species)



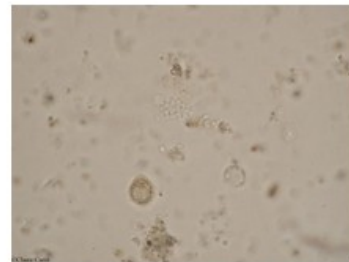
*Alexandrium* (Genus)



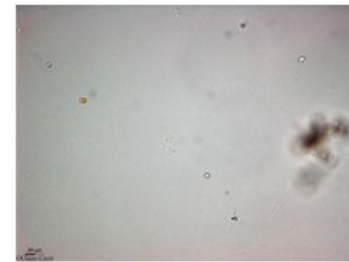
*Amphora* (Genus)



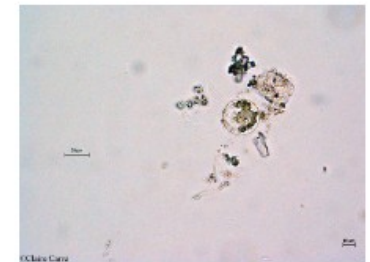
*Ankistrodesmus* (Genus)



*Aphanocapsa* (Genus)



*Aphanothece* (Genus)



*Asterionellopsis glacialis* (Species)



# Décrire et enrichir

- Ex2 : Utilisation du référentiel **Carthage** (Réseau hydrographique, IGN)

→ Echantillonnage de truites

→ BD Carthage : identifiant d'entités hydrographiques

→ Fiche descriptive sur le SANDRE

[http://services.sandre.....fiche\\_courdo.php?CdSandre=Y32-0400](http://services.sandre.....fiche_courdo.php?CdSandre=Y32-0400)

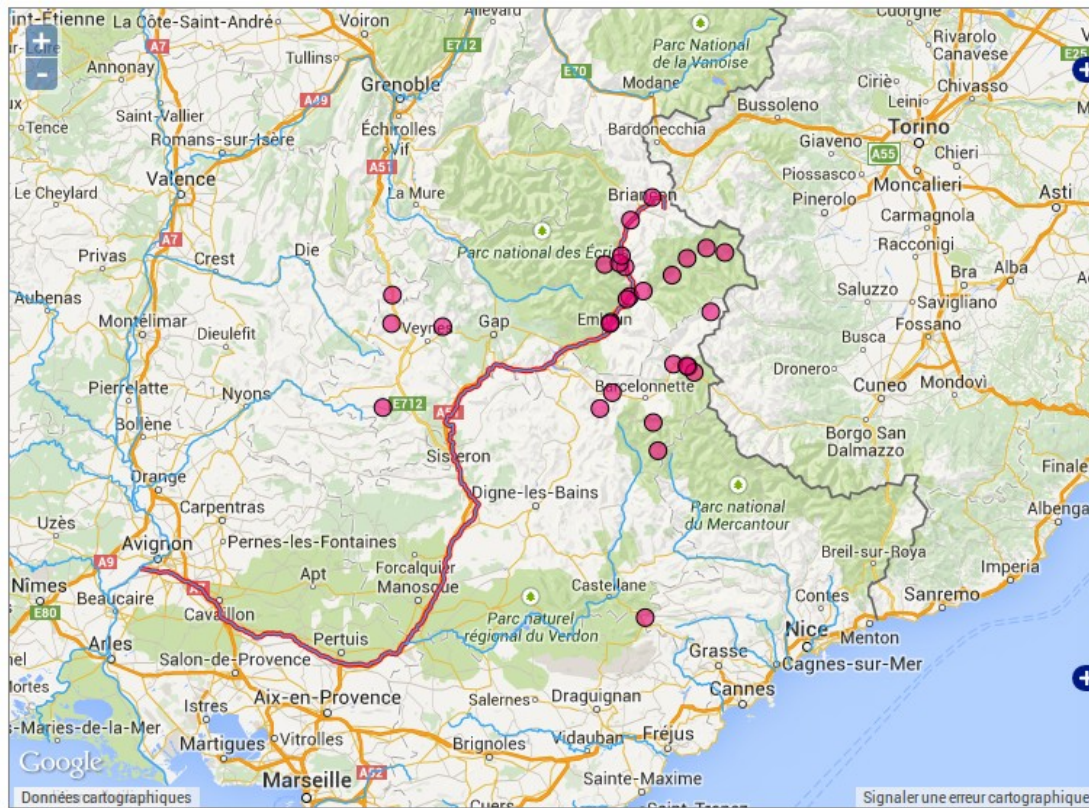
→ Web services géographiques, couches spatiales

→ Arbre confluence / affluence d'un cours d'eau



# Décrire et enrichir

- Ex2 : Utilisation du référentiel **Carthage** (Réseau hydrographique, IGN)



▲ **Styler les stations**

▼ **Critères de sélection des stations**

**PAYS :**

La Durance
L'Asco
L'Aude
L'Aujon
Le Bevinco
La Cagne
Le Chassezac
Course ou Canche
la Dore
Le Doubs
La Dourbie
Le Doux
La Dranse
La Dronne
<b>La Durance</b>
L'Esteron
L'Eyrieux
Le Fango
Le Fium Alto
Le Fium'Orbu
La Salve

La Durance

# Alimenter les BDD

- Données en continu (fichiers texte)
  - Acquisition depuis serveurs / PC distants
  - Vérification et insertion en continu
  - Systèmes d'alerte par mail
- Données ponctuelles (fichiers texte / Excel, photos, ..)
  - Dépôt de fichiers sur FTP puis insertion via site Web
  - Ou upload et insertion en ligne



- Carnets de terrain électroniques



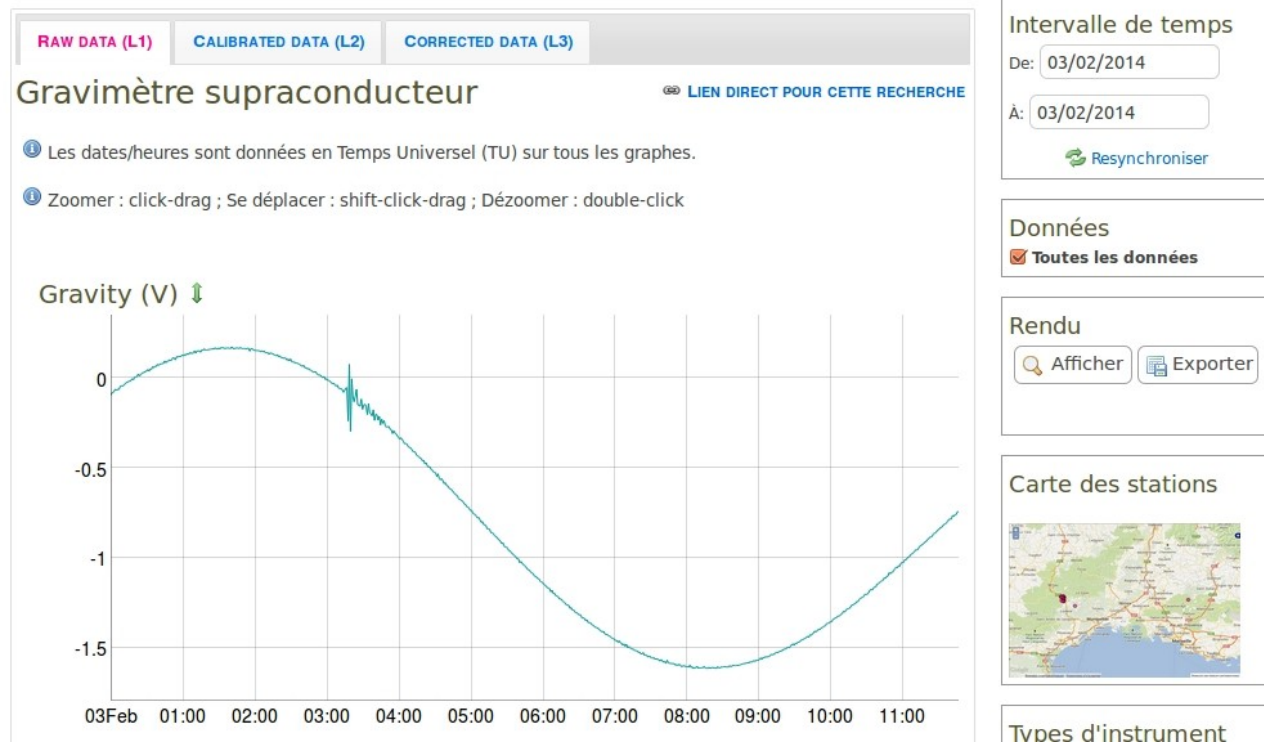
# Diffusion des données





# Graphiques dynamiques

- Séries temporelles interactives (lib. **dygraphs**)



Modèle de description des graphiques  
thèmes, groupes, graphiques, variables, discriminants

=> Pages web génériques



# Graphiques dynamiques

- Séries temporelles interactives (lib. **dygraphs**)



Export des données + descripteurs + conditions d'utilisation

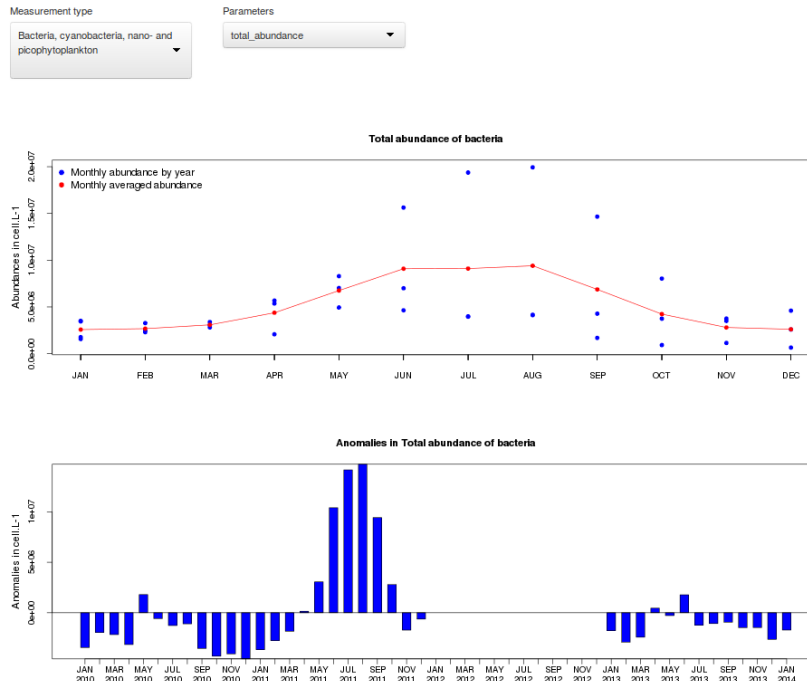


# Graphiques dynamiques

- Graphiques R interactifs (lib. **Shiny**)

OSU OREME Thau Microbial Communities Data

Abundance and anomaly plots of microbial communities of Thau



OSU OREME Thau Microbial Communities Data

Traffic light plots of microbial communities of Thau



# Galleries photo



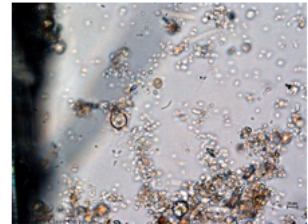
<p><b>Site</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> BIZERTE</li> <li><input type="checkbox"/> MAYOTTE</li> <li><input type="checkbox"/> SCATTERED ISLANDS</li> <li><input type="checkbox"/> SEBKHET IMLILY</li> <li><input type="checkbox"/> THAU LAGOON</li> <li><input type="checkbox"/> TULEAR LAGOON</li> <li><input type="checkbox"/> WHALE ISLAND BAY</li> </ul> <p>Tout désélectionner</p>	<p><b>Projet</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> COMMA</li> <li><input type="checkbox"/> COMMA2</li> <li><input type="checkbox"/> DZAHA</li> <li><input type="checkbox"/> DZIANI</li> <li><input type="checkbox"/> ECOMET</li> <li><input type="checkbox"/> FISHPLANK</li> <li><input type="checkbox"/> GELAMED</li> <li><input type="checkbox"/> ICAR2</li> <li><input type="checkbox"/> IMLILY</li> <li><input type="checkbox"/> IMOCA</li> <li><input type="checkbox"/> LINDANE</li> <li><input type="checkbox"/> MANIP_MASTER</li> <li><input type="checkbox"/> OSU_PLANKTON</li> <li><input type="checkbox"/> RETENUES_COLLINAIRES_DE_MAYOTTE</li> <li><input type="checkbox"/> RISCO</li> </ul> <p>Tout désélectionner</p>	<p><b>Classe</b></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Bacillariophyceae</li> <li><input type="checkbox"/> Bicoecea</li> <li><input type="checkbox"/> Chlorophyceae</li> <li><input type="checkbox"/> Choanoflagellata</li> <li><input type="checkbox"/> Chrysophyceae</li> <li><input type="checkbox"/> Cryptophyceae</li> <li><input type="checkbox"/> Cyanophyceae</li> <li><input type="checkbox"/> Dictyochophyceae</li> <li><input type="checkbox"/> Dinophyceae</li> <li><input type="checkbox"/> Euglenoidea</li> <li><input type="checkbox"/> Prasinophyceae</li> <li><input type="checkbox"/> Pymnesiophyceae</li> <li><input type="checkbox"/> Unknown</li> </ul> <p>Tout désélectionner</p>
--	--	---

Recherche à facettes

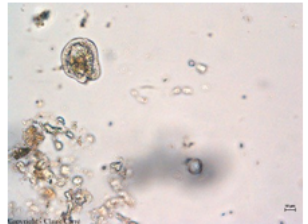
Filtrer par taxon :

## Taxons

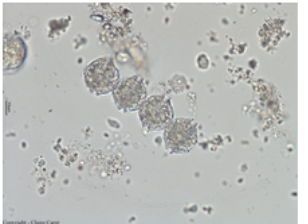
120 taxon(s)



**ACANTHOICA QUATTROSPINA (SPECIES)**



**AKASHIWO SANGUINEA (SPECIES)**



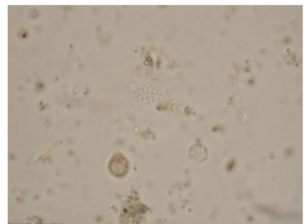
**ALEXANDRIUM (GENUS)**



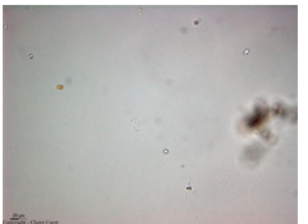
**AMPHORA (GENUS)**



**ANKISTRODESMUS (GENUS)**



**APHANOCAPSA (GENUS)**



**APHANTHECE (GENUS)**



**ASTERIONELLOPSIS GLACIALIS (SPECIES)**

# Cartographie en ligne

- Serveur cartographique **Geoserver**



- Standards OGC
- Web services géographiques WMS, WFS  
=> Diffusion de données interopérables

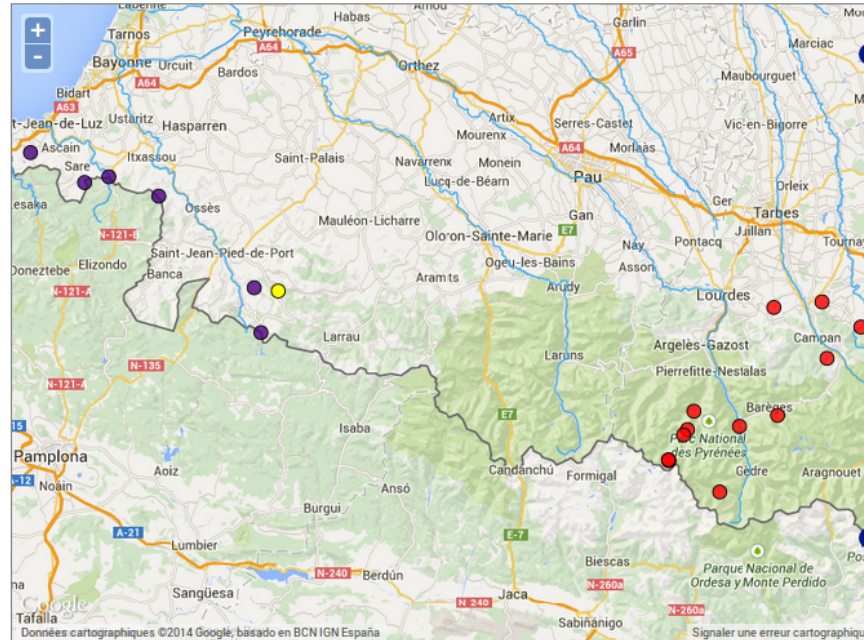
- Clients

- SIG bureau (GQIS)
- OpenLayers
- OpenLayers + Ext JS = **GeoExt**



# Cartographie en ligne

- Cartes “simples” de stations et descriptions



Station Ihounatzeko, MENDIVE, PYRENEES-ATLANTIQUES, FRANCE

Coordonnées (X, Y Lambert II étendu) : 319525, 1.79433e+06

Type : rivière

Entité hydrographique : Indéfini

Nb de truites étudiées : 30

Etude ADOUR1

- Nb de stations : 21
- Rapport : [enl3538-adour1-1997.pdf](#)
- Sponsor(s) : TFP

**Styler les stations**

Aucun style

Selon le département

Selon le nom

Label des stations (ralentit l'affichage !)

---

**Critères de sélection des stations**

**Pays :**

Tous

**Département :**

Tous

01 - AIN

02 - AISNE

03 - ALLIER

04 - ALPES-DE-HAUTE-PROVENCE

**Etude :**

Toutes

ADOUR1

AHP1

AHP2

AHP3

**Cours d'eau et ses affluents :**

Tous

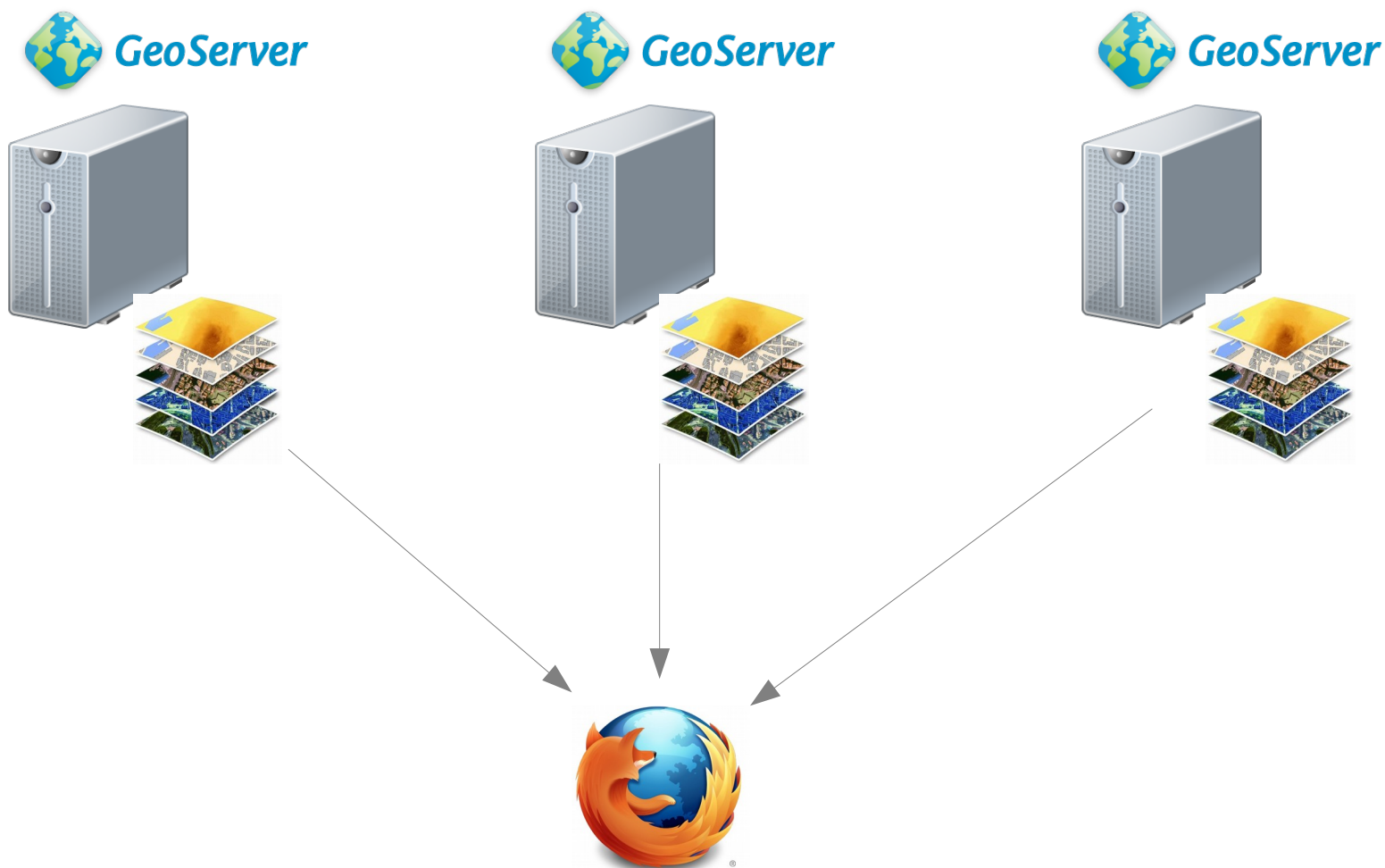
**Station :**

All

Andurentako

# Cartographie en ligne

- Portail cartographique



# Cartographie en ligne

- Portail cartographique : stations + données

The screenshot displays a web-based mapping application interface. At the top, there is a panel for selecting a WMS server, currently set to 'OSU OREME'. Below this is a table of available layers:

Title	Description
SNO DYNLTC - Littoral and shoreline dynamics st...	1- SNO DYNLTC - OSU OREME stations for measurements of littoral and shoreline dynamics
SNO H+ - Boreholes stations for hydrogeophysics	1- SNO H+ - OSU OREME boreholes for hydrogeophysics observations
SNO H+ - GEK stations	1- SNO H+ - OSU OREME GEK ( Geodedy in Karstic Environment) measurement stations
SNO KARST - MEDCYSS stations	1- SNO KARST - OSU OREME MEDCYSS stations (Multi-scale observatory of flood dynamics and underground hydrodynamics in karstic environment)
SNO SOMLIT - REC Thau stations	1- SNO SOMLIT - OSU OREME stations for research in coastal environment of Thau
SO EVOPOP - RIC stations	1- SO EVOPOP - OSU OREME breeding stations for spatio-temporal evolution of insecticide resistance genes of the mosquito Culex pipiens

Below the table is a 'Map layers' panel with a list of layers to be added to the map:

- Base maps
  - Google Sat
  - Google Map
  - OpenStreetMap
- Layers
  - posidonia in Gulf of Lions
  - SHORELINE-1935
  - SNO SOMLIT - REC Thau stations
  - SNO H+ - GEK stations
  - SNO H+ - Boreholes stations for hydrogeop...
  - SNO DYNLTC - Littoral and shoreline dyna...
  - SO Microbial communities - Microbial com...
  - SO EVOPOP - Salmo trutta stations
  - SO EVOPOP - RIC stations

The main map area shows a geographical view of the Languedoc-Roussillon region in France, with various colored markers (red, orange, blue, green, purple) representing different station types. The map includes labels for cities like Nîmes, Montpellier, and Lattes, and geographical features like the Mediterranean Sea. A 'Print map ... (only with OpenStreetMap)' button is visible at the bottom.





# Diffusion des descripteurs

- Outil CRUD : Create, Read, Update, Delete

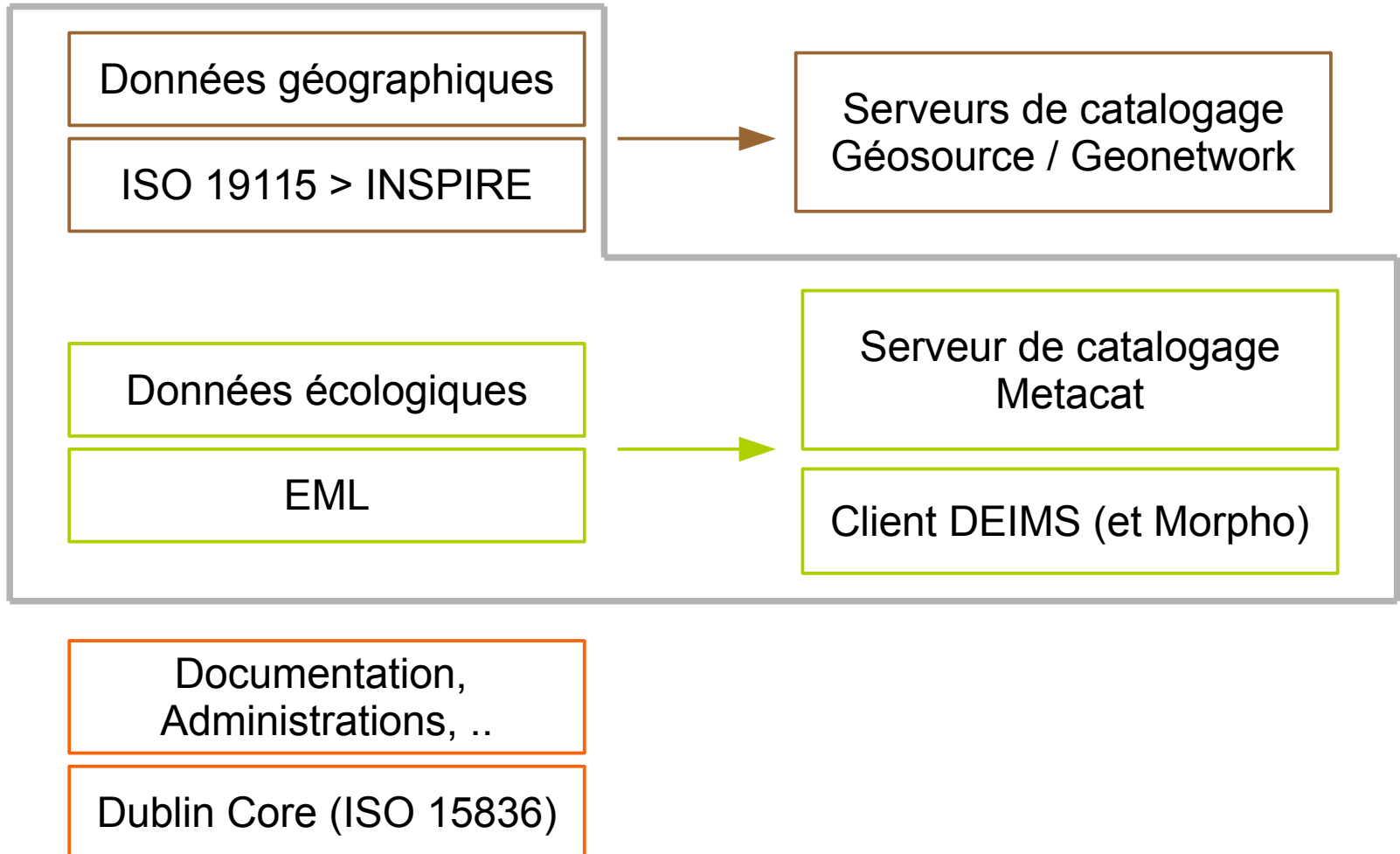
SITES						
STATIONS						
TYPES D'INSTRUMENT						
INSTRUMENTS						
CAMPAGNES DE MESURE						
OPÉRATIONS DE MAINTENANCE						
Ajouter Site						
<a href="#">Exporter</a> <a href="#">Imprimer</a>						
Code	Nom	Description	Fichier de description	Ville	Département	Pays
BAUM	Les Baumelles	Zone of les Baumelles pothole		LA COUVERTOIRADE	AVEYRON	France
BESS	Les Besses	Zone of the pothole		SAINT-MAURICE-NAVACELLES	HERAULT	France
BISE	La Bise	Zone of la Bise pothole		L'HOSPITALET-DU-LARZAC	AVEYRON	France
BLAQ	La Blaquererie	Farm		LA COUVERTOIRADE	AVEYRON	France
CANA	Les Canalettes	Zone of the Canalettes forest ...		NANT	AVEYRON	France
INFR	Les Infruts	Zone of Infruts pothole	<a href="#">infruts.zip</a>	LA COUVERTOIRADE	AVEYRON	France
JASS	La Jasse		<a href="#">jasse_house_map.pdf</a>	L'HOSPITALET-DU-LARZAC	AVEYRON	France
LSBB	Laboratoire Souterrain à Bas-...			RUSTREL	VAUCLUSE	France
MENU	Les Menudes			CORNUS	AVEYRON	France
PORT	La Portalerie	Zone of the pothole	<a href="#">portalerie.zip</a>	LA COUVERTOIRADE	AVEYRON	France

10 Page 1 de 2 Affichage de 1 à 10 de 12 enregistrements



# Catalogage

- Quelques schémas de métadonnées et outils



# Catalogage

- Le choix ?

Décrire en EML

Niveau attributaire :  
paramètres, unités, ...

→ ré-utilisabilité !



Convertir en INSPIRE

Description de haut  
niveau ... mais norme,  
obligatoire

- Outils + ou – matures et ergonomiques ...
- Test de **Metacat**
- **DEIMS** à tester (cf **LTER DEIMS**)



# Bilan



# Les freins

- Disponibilité et motivation des producteurs de données
  - Fournir les descriptions, les infos
  - Formater les données
  - Tester et faire des retours sur les développements
  - Différents degrés de motivation :
    - Respecter les obligations de mise à disposition
    - Valoriser ses travaux
    - Disposer d'outils de travail
- Disponibilité et format des données



# Les freins

- Communication producteur - informaticien
    - Vocabulaire, vulgarisation
    - Compréhension des besoins
  - Coût de montée en compétences
    - Compréhension des thématiques, relationnel
    - Besoins, outils, méthodes, ...
- Temps moyen de constitution des systèmes d'information : 2 ans



# Des clés ?

→ **Motiver** les producteurs : outils, valorisation

→ Plusieurs projets de bases de données en **parallèle**

Si un seul projet :

- Réunions hebdomadaires, voire journalières au début
- Plusieurs personnes contact identifiées

→ Mettre la main dans les données pour accélérer !

→ **Capitaliser** les savoirs-faire : équipe pérenne

