

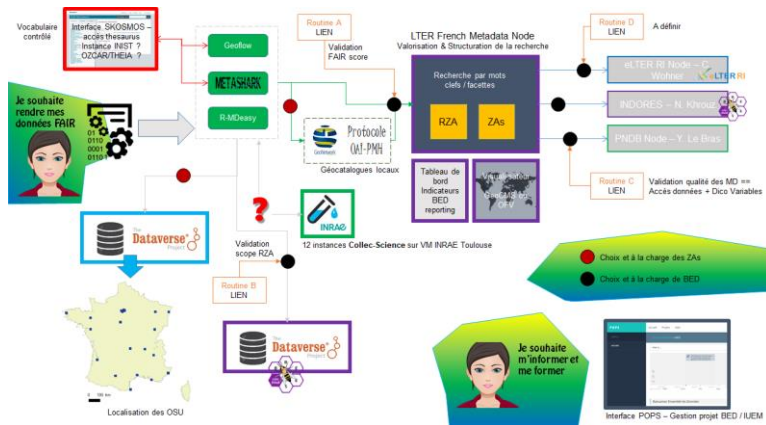
Flux de données et Outils

Présentation de l'écosystème autour des données environnementales



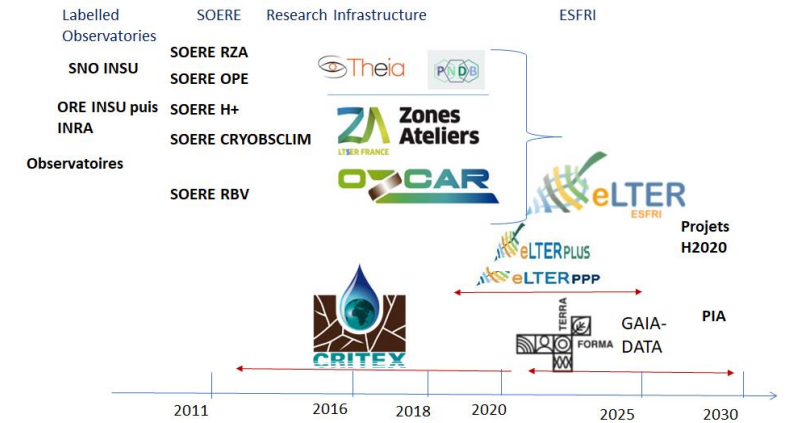
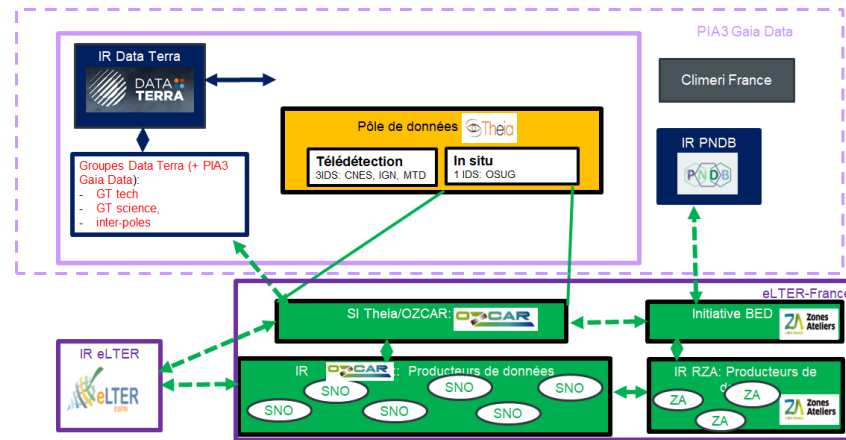
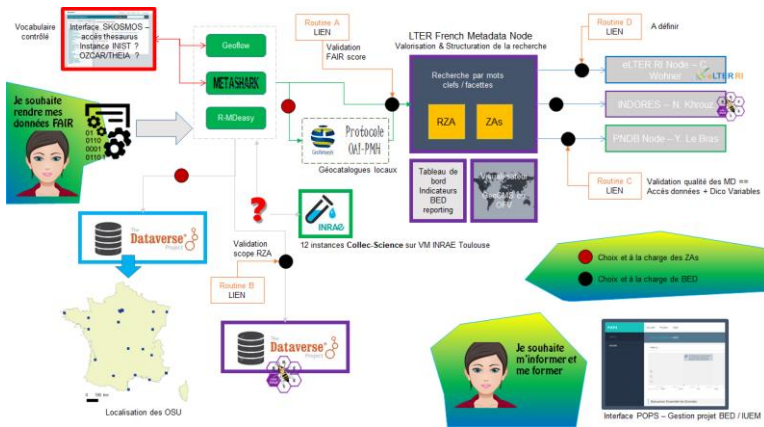
Vue d'ensemble

Schéma général avec LTER, GBIF, PNDB, OZCAR Theia, CNRS, INRAE, futur RDG ...



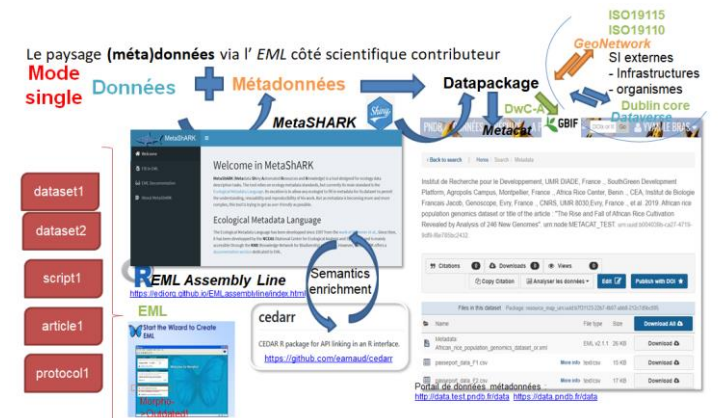
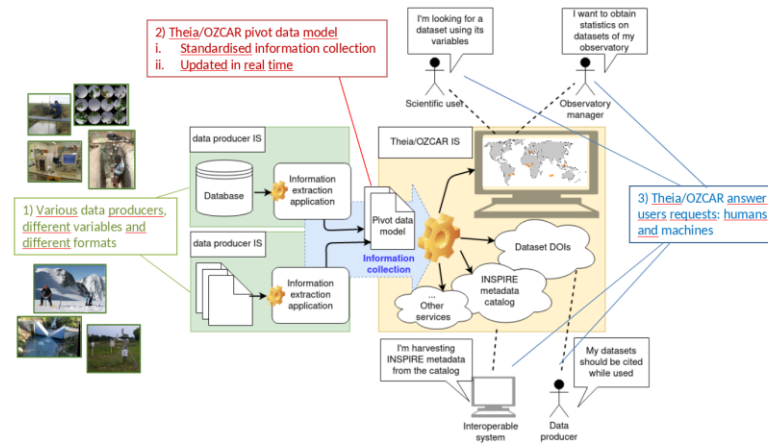
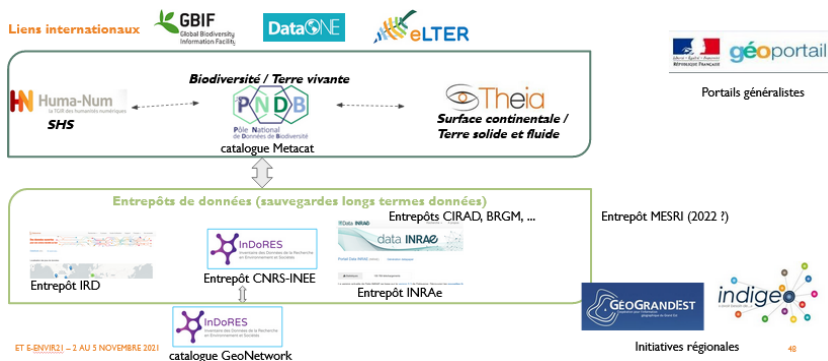
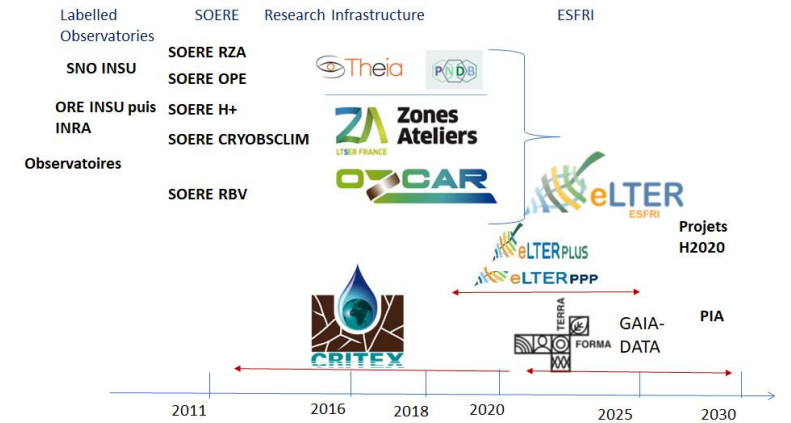
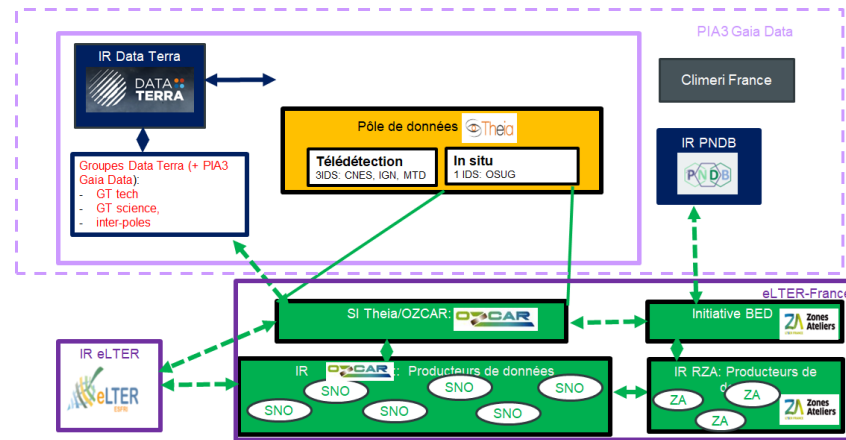
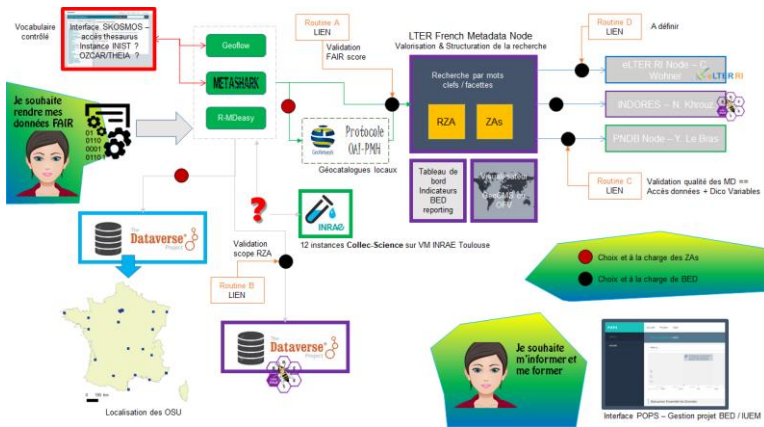
Vue d'ensemble

Schéma général avec LTER, GBIF, PNDB, OZCAR Theia, CNRS, INRAE, futur RDG ...



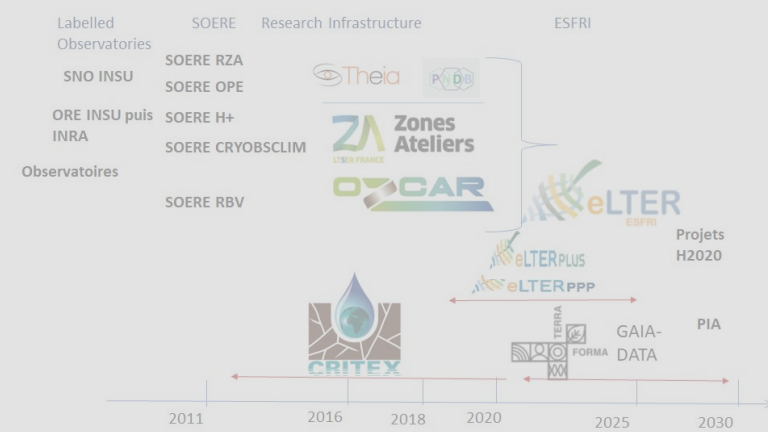
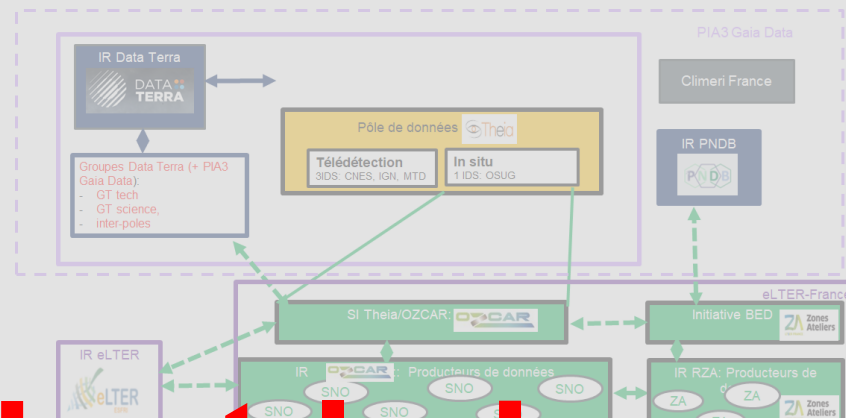
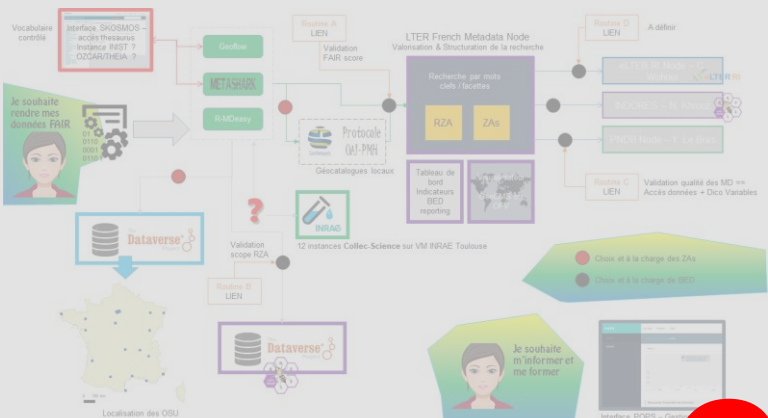
Vue d'ensemble

Schéma général avec LTER, GBIF, PNDB, OZCAR Theia, CNRS, INRAE, futur RDG ...

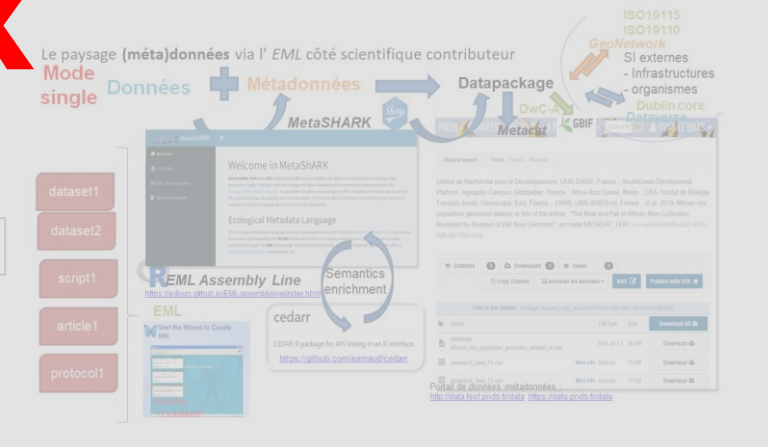
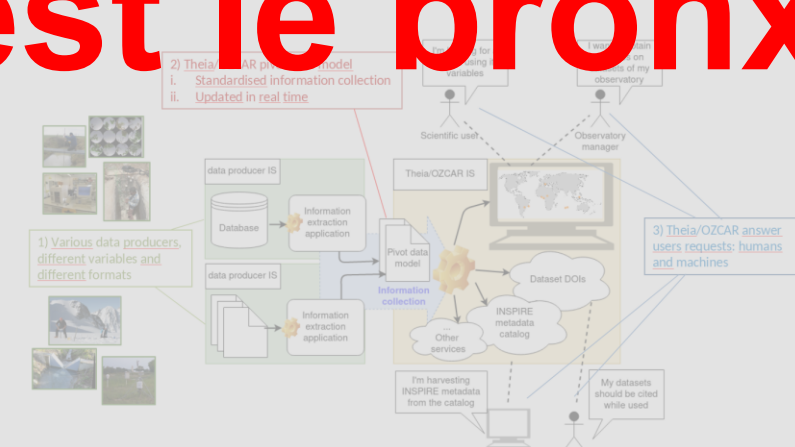
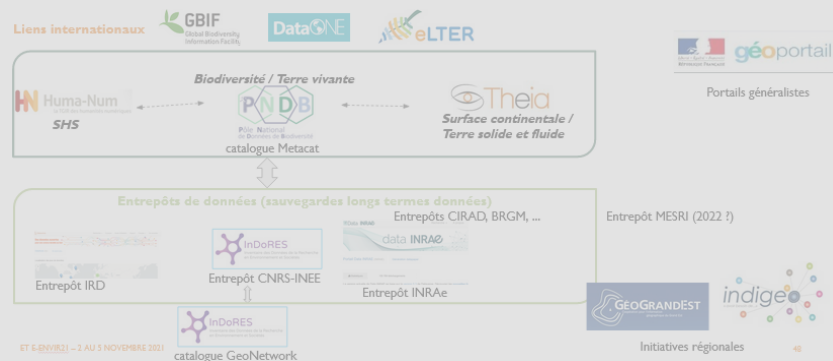


Vue d'ensemble

Schéma général avec LTER, GBIF, PNDB, OZCAR Theia, CNRS, INRAE, futur RDG ...



C'est le bronx



Vue d'ensemble

C'est pourquoi Geoflow et MetaShARK sont là !

- but de simplifier / faciliter les processus de structuration des données et partage des données et métadonnées
- mais en voulant simplifier / faciliter en créant de nouveaux outils, de nouvelles approches, il y a un risque de complexifier encore plus...

=> c'est pour cela qu'on est là ! ;) C'est pour cela qu'on doit travailler ensemble, échanger

Geoflow

{MetaShARK} 

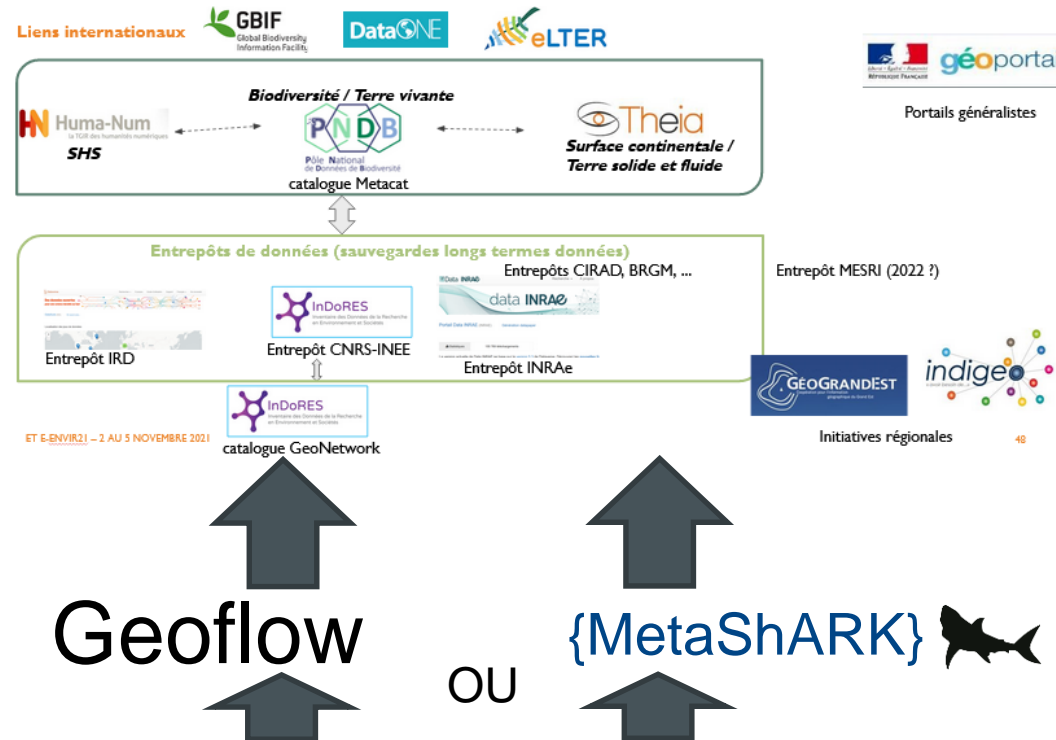
Le but : « Juste » 2 solutions qui permettent de gérer l'ensemble des flux !



Producteurs des données

- Acquisition des données
- Qualification des données
- Bancarisation des données / SI local

Vue d'ensemble



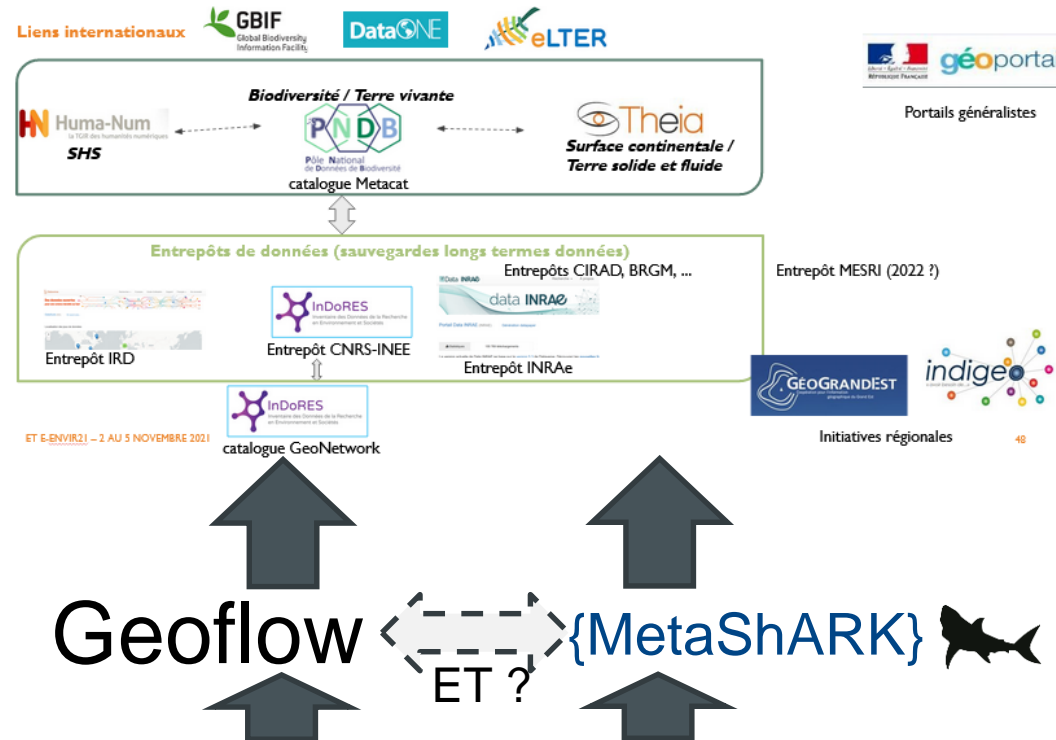
Le but : « Juste » 2 solutions qui permettent de gérer l'ensemble des flux !



Producteurs des données

- Acquisition des données
- Qualification des données
- Bancarisation des données / SI local

Vue d'ensemble



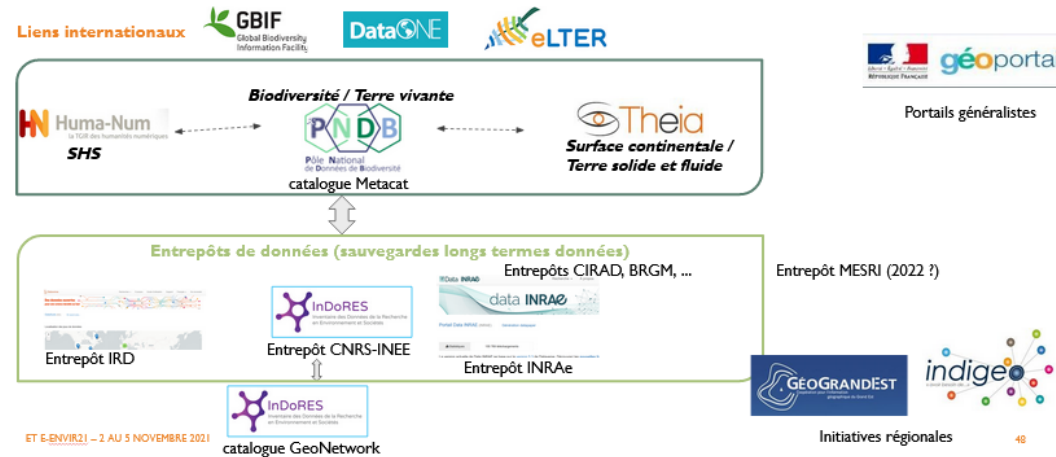
Le but : « Juste » 2 solutions qui permettent de gérer l'ensemble des flux !



Producteurs des données

- Acquisition des données
- Qualification des données
- Bancarisation des données / SI local

Vue d'ensemble



MetaGeoShARKflow ;)

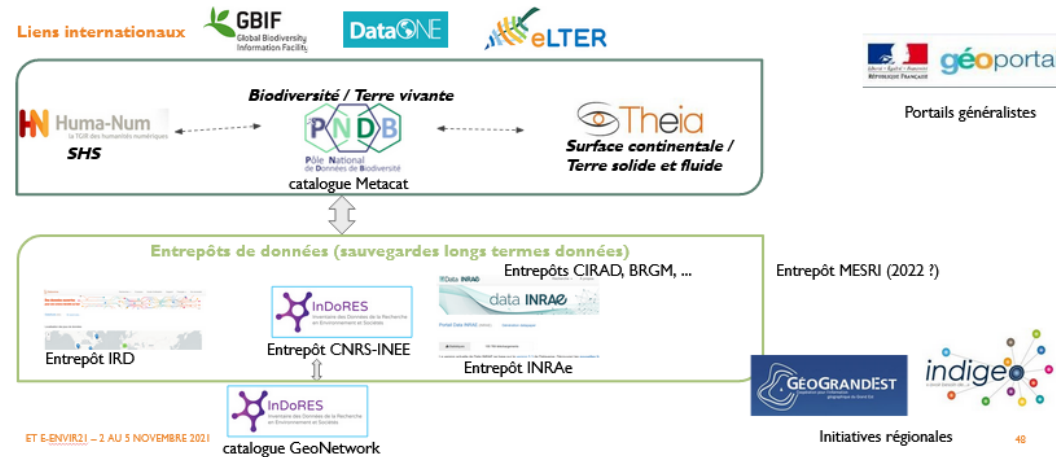
1 solution unique qui permette de gérer l'ensemble des flux ! ?



Producteurs des données

- Acquisition des données
- Qualification des données
- Bancarisation des données / SI local

Vue d'ensemble



MetaGeoShARKflow ;)

1 solution unique qui permette de gérer l'ensemble des flux ! ?

1 solutions unique préférable ?

Pour tous les utilisateurs ?

Types d'utilisateurs différents avec besoins différents !

Producteur de données « solitaire » vs producteurs SI

Producteur vs gestionnaire de données ..



Producteurs des données

- Acquisition des données
- Qualification des données
- Bancarisation des données / SI local

Côté PNDB

Data = **Souhait de ne pas centraliser les données => données téléchargeables directement depuis entrepôt via URL**

- Données brutes et dérivées
- Codes sources
- Protocoles
- Publications
- => Research Objects

Data open uniquement

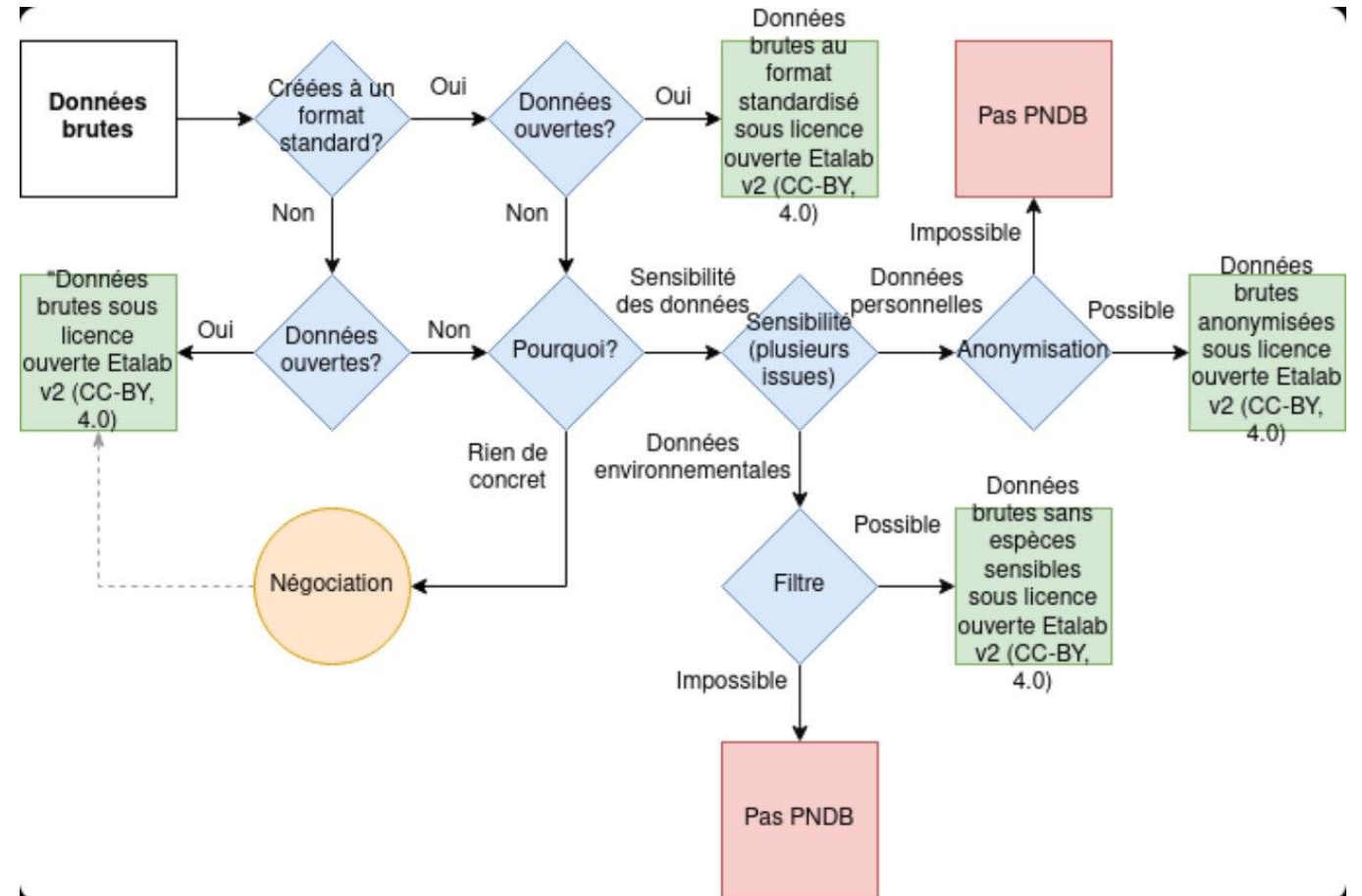
- Licence ouverte Etalab V2 (CC-BY-4.0)
- Téléchargement direct

Metadonnées détaillées

- Minima détails des attributs de données

Liens

- Data Terra et CLIMERI
- Politiques publiques SI fédérateurs OFB
- International DataOne et GBIF



Côté PNDB

WIP

Data = **Souhait de ne pas centraliser les données => données téléchargeables directement depuis entrepôt via URL**

- Données brutes et dérivées
- Codes sources
- Protocoles
- Publications
- => Research Objects

Github / Gitlab

Protocols.io

Software Heritage



Entrepôt MESRI (2022 ?)

ZENODO
Si pas autre

DRYAD
Déconseillé

URLs direct download

Data open uniquement

- Licence ouverte Etalab V2 (CC-BY-4.0)
- Téléchargement direct

EML metadata file

Metadonnées détaillées

- Minima détails des attributs de données

Liens

- Data Terra et CLIMERI
- Politiques publiques SI fédérateurs OFB

Download Links + Metadata conversion EML

Plateforme Galaxy-E



Pôle National de Données de Biodiversité catalogue Metacat



Pôle National de Données de Biodiversité catalogue Metacat

Utilisateurs français

Réplication



Utilisateurs étrangers

Conversion Dwc



Global Biodiversity Information Facility

- International DataOne et GBIF

Côté PNDB

WIP

OAI-PMH used by metaSNAKE to « link » French Biodiversity Information Systems to PNDB

MetaSNAKE, « A Python version of MetaShARK »
Only for PNDB team for all PNDB system backend management

metaSNAKE Metacat Moissonnage Stats Rechercher Edito PROD Mode (it's da baby!) Recherche Yvan Le Bras

Document (via OAI-PMH)

Les schemas de metadonnees suivants:

- OAI_DC
- PANSIMPLE

Filtrage sur un sous-ensemble de documents suivants:

- ANAEE (GeoNetwork) (via OAI-PMH)
- GBIF (via OAI-PMH)
- INDORES (GeoNetwork) (via OAI-PMH)
- INRAE (dataverse) (via OAI-PMH)
- LTER France (via OAI-PMH)
- PNDB PrePROD (via OAI-PMH)
- SEANOE (via OAI-PMH)

Afficher 25 lignes par page

747 résultats

Début < 1 2 3 4 5 ... 30 > Fin

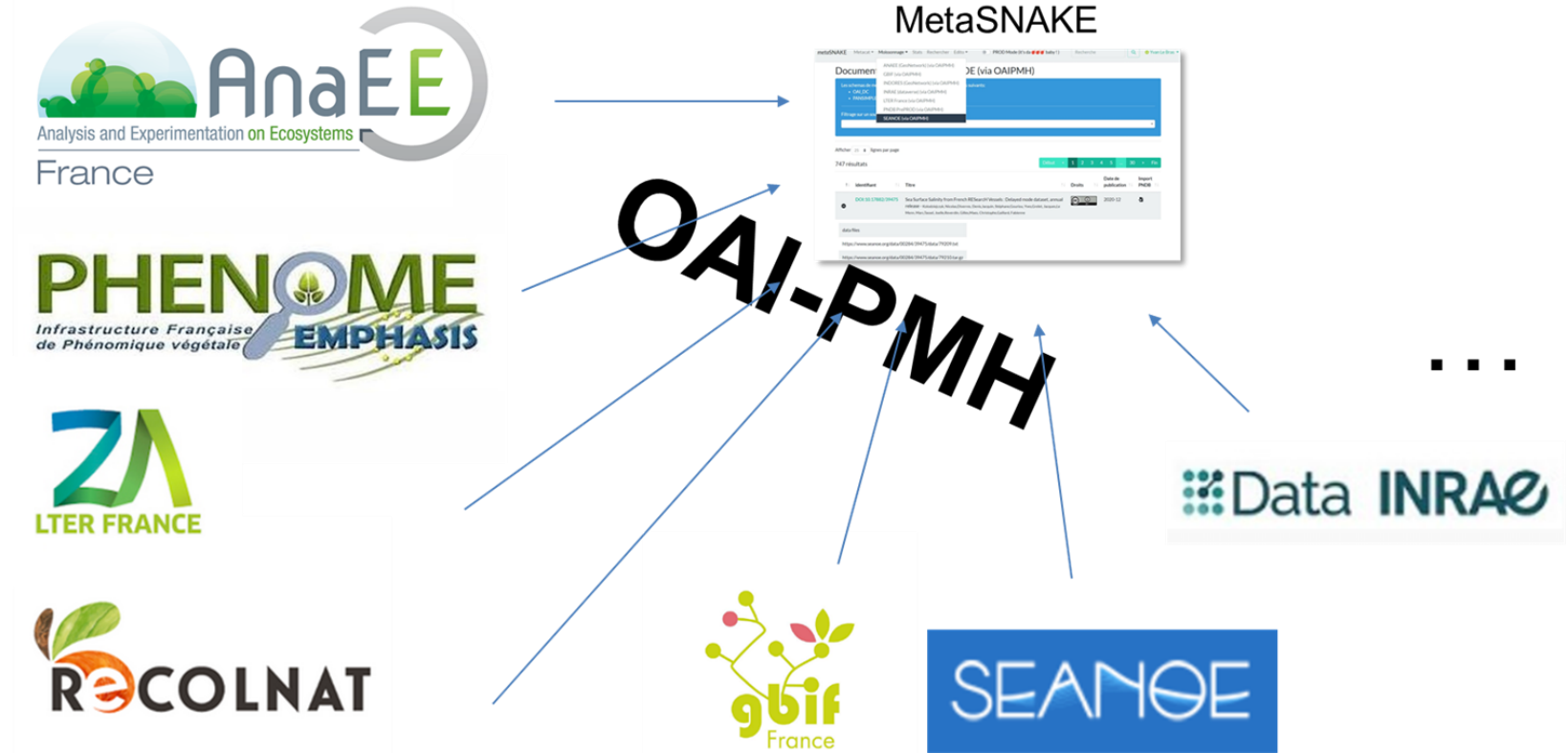
Identifiant	Titre	Droits	Date de publication	Import PNDB
DOI:10.17882/39475	Sea Surface Salinity from French REsearch Vessels : Delayed mode dataset, annual release - Kolodziejczyk, Nicolas,Diverres, Denis,Jacquin, Stéphane,Gouriou, Yves,Grelet, Jacques,Le Menn, Marc,Tassel, Joelle,Reverdin, Gilles,Maes, Christophe,Gaillard, Fabienne		2020-12	
data files				
https://www.seanoe.org/data/00284/39475/data/79209.txt				
https://www.seanoe.org/data/00284/39475/data/79210.tar.gz				

Côté PNDB

WIP

OAI-PMH used by metaSNAKE to « link » French Biodiversity Information Systems to PNDB

MetaSNAKE, « A Python version of MetaShARK »
Only for PNDB team for all PNDB system backend
management



Côté PNDB

WIP

OAI-PMH used by metaSNAKE to « link » French Biodiversity Information Systems to PNDB

MetaSNAKE, « A Python version of MetaShARK »
Only for PNDB team for all PNDB system backend management



Which mtd language ?

- OAI_DC
- PANSIMPLE
- ISO19139
- CSW-RECORD
- ISO19110
- DUBLIN-CORE
- EML
- DATAVERSE_JSON
- OAI_DATACITE
- OAI-DDI
- DATACITE
- ...

9,094



55,979



747



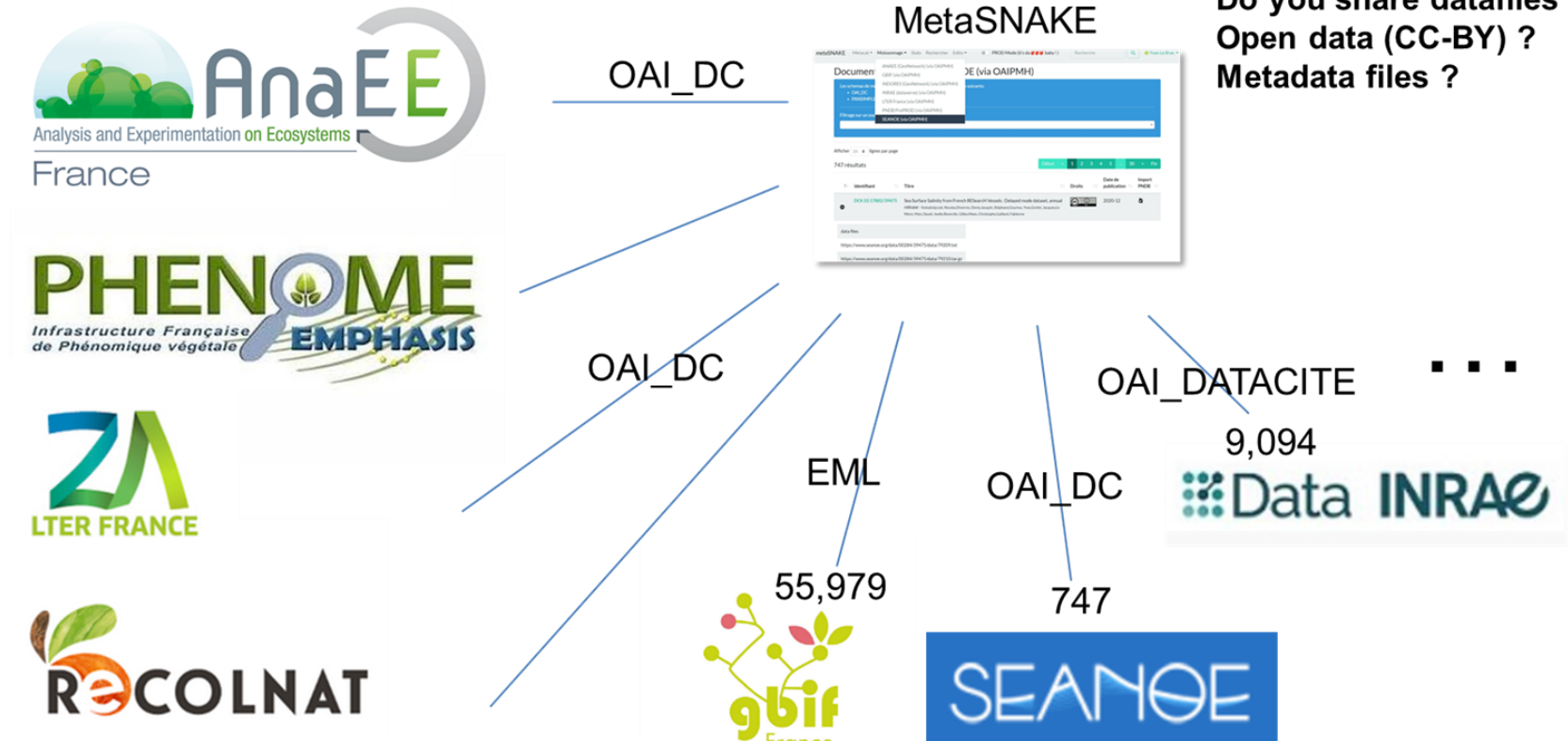
Côté PNDB

WIP

OAI-PMH used by metaSNAKE to « link » French Biodiversity Information Systems to PNDB

MetaSNAKE, « A Python version of MetaShARK »
Only for PNDB team for all PNDB system backend management

Which mtd language
How many records
Do you share datafiles ?
Open data (CC-BY) ?
Metadata files ?



Côté PNDB

WIP

OAI-PMH used by metaSNAKE to « link » French Biodiversity Information Systems to PNDB

MetaSNAKE, « A Python version of MetaShARK »
Only for PNDB team for all PNDB system backend management

OAI_DC / SEANOE example

```
<metadata>
  <oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
    <dc:title>COSTOF, technical data from SeaMoN West, node of the EMSO-Azores observatory, 2013-2014</dc:title>
    <dc:creator>Legrand, Julien</dc:creator>
    <dc:creator>Sarradin, Pierre-marie</dc:creator>
    <dc:creator>Cannat, Mathilde</dc:creator>
    <dc:subject>Mid-Atlantic Ridge</dc:subject>
    <dc:subject>EMSO</dc:subject>
    <dc:subject>Lucky Strike</dc:subject>
    <dc:subject>Time-series</dc:subject>
    <dc:subject>Geophysical monitoring node</dc:subject>
    <dc:subject>MoMAR</dc:subject>
    <dc:subject>COSTOF</dc:subject>
    <dc:description>This dataset contains technical parameters (Voltage in V, current in mA, energy in Ah, internal pressure between September 2013 and February 2014 (n=332) by the COSTOF (COmmunication and STORage Front-end), which is the Electr station of EMSO-Azores dedicated to geophysical observations (data acquired every 6 hours).</dc:description>
    <dc:publisher>SEANOE</dc:publisher>
    <dc:date>2015-10</dc:date>
    <dc:type>dataset</dc:type>
    <dc:identifier>DOI:10.17882/39764</dc:identifier>
    <dc:identifier>http://dx.doi.org/10.17882/39764</dc:identifier>
    <dc:relation>info:eu-repo/grantAgreement/EC/FP7/312463/EU//FIXO3</dc:relation>
    <dc:coverage>North 37.30134, South 37.2888, East -32.275618, West -32.27982</dc:coverage>
    <dc:references>https://www.seanoe.org/data/00286/39764/illustration.gif</dc:references>
    <dc:terms:spatial xsi:type="DCTERMS:Box">37.2888 -32.27982 37.30134 -32.275618</dc:terms:spatial>
    <dc:rights>CC-BY</dc:rights>
  </oai_dc:dc>
```

High level Metadata

MetaShARK
MetaSNAKE → Detailed Metadata

Which mtd language
How many records
Do you share datafiles ?
Open data (CC-BY) ?
Metadata files ?

PANSIMPLE -> Datafiles url

Data



Vue d'ensemble

Schéma général avec LTER, GBIF, PNDB, OZCAR Theia, CNRS, INRAE, futur RDG ...

Les flux de données et métadonnées

Données spatiales :

WMS (Web Map Service)

WFS (Web Feature Service)

WCS (Web Coverage Service)

Métadonnées :

CSW (Catalog Service for the Web)

OAI-PMH (Open Archives Initiative - Protocol for Metadata Harvesting)

REMERCIEMENT

cnrs

- CNRS soutien financier
- R. OMNES (conseillère formation, DR10) suivi de la formation
- P. RUPP (ICUBE), C. DIDIERJEAN (EDYTEM) gestion administrative et financière
- Le comité organisateur : BED/RZA, OZCAR/THEIA, UMS BBEES, PNDB, avec la participation de la MISHA
- Centre Saint-Thomas, hébergement

- Retrouvez les informations sur : <https://e-envir-21.sciencesconf.org/>



Maison Interuniversitaire des
Sciences de l'Homme
- Alsace -

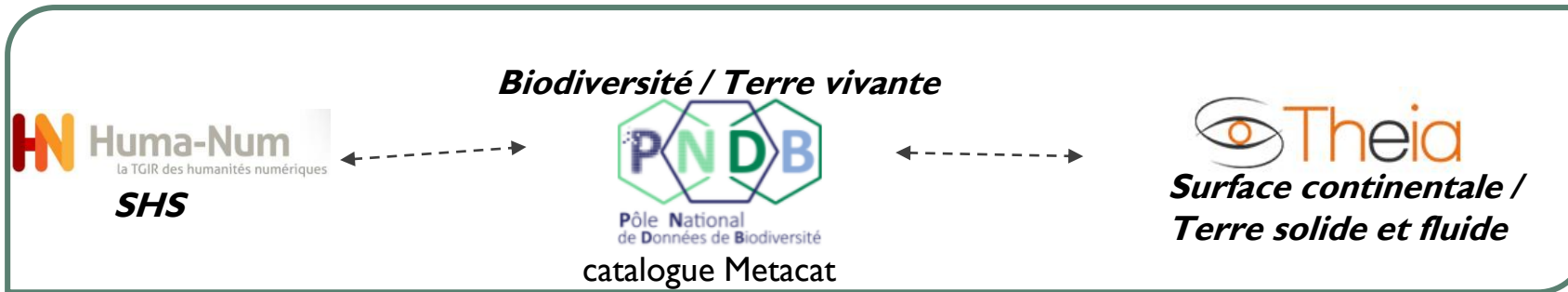




ANNEXES

Vue d'ensemble

Liens internationaux



Portails généralistes



Entrepôt MESRI (2022 ?)



Initiatives régionales