



OBSERVATOIRE DE RECHERCHE
MÉDITERRANÉEN DE L'ENVIRONNEMENT

Pérennisation et interopérabilité des données de l'OSU OREME

Observatoire de Recherche
Méditerranéen de l'Environnement

Les OSU, l'Oreme

- Des structures transverses du CNRS
 - Observatoires des Sciences de l'Univers, 27 en France
 - Mutualiser des moyens pour l'observation pérenne (équipements, gestion de données, ...)
 - Favoriser l'interdisciplinarité
 - OSU OREME : 4 tutelles, 8 UMR, 20 Services d'observation
- Missions autour des données
 - Pérenniser, gérer, mettre à disposition
 - Alimenter les pôles de données (IR Système Terre), IR, SNO ..
- Des données hétérogènes
 - Disciplines, types, méthodes d'acquisition, fréquences ...
 - => développement d'outils et de méthodes mutualisés (tant que possible)

Problématique

Fichiers, cartes, bases
de données ..



Formats propriétaires,
non descriptifs,
incohérences, ..



?

Pôles, infrastructures
de recherche, SNO ..

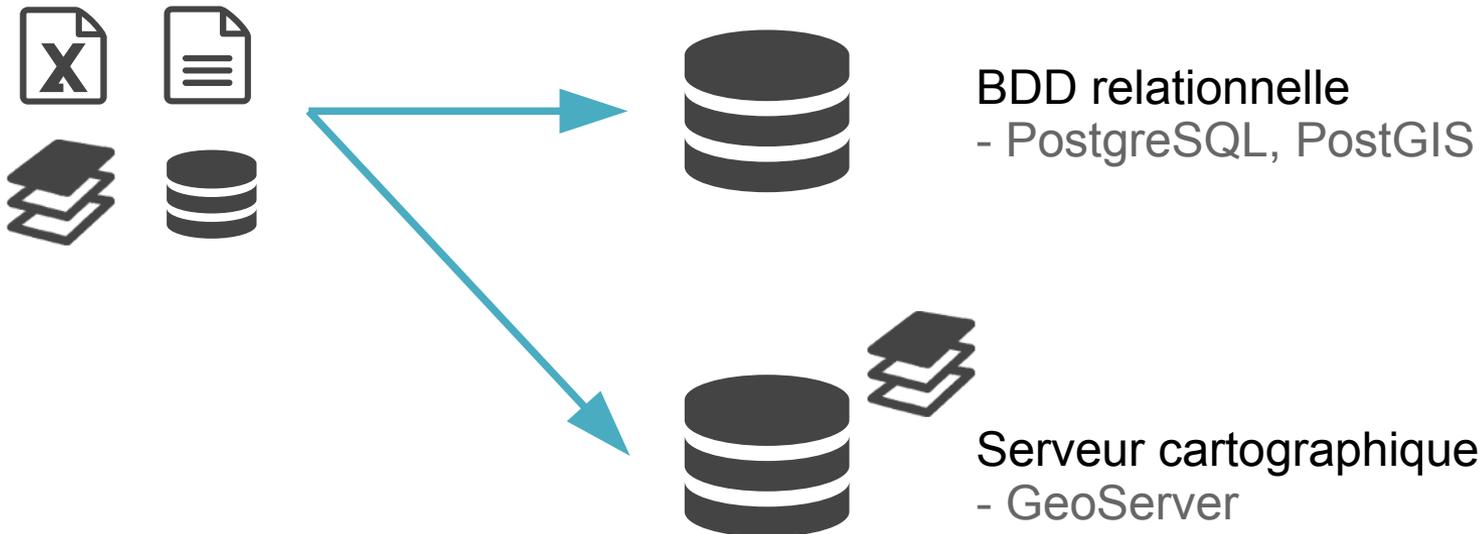


Portails internationaux
(GBIF, ...)

Structurer

- Objectif : rendre les données interrogeables et cohérentes

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



Décrire

date	valeur
2018-11-03	12



- Objectifs :
 - Rendre les données **compréhensibles**, non ambiguës
 - **Lier** les données entre elles

Ré-utilisabilité

Interopérabilité sémantique

- Qui, où, quand, quoi, comment (protocoles, traitements, ...), pourquoi (contexte) ? - Tables SQL
- Intégrer des **standards** pour **enrichir** les données
 - référentiels, vocabulaires contrôlés, thésaurus, ...

Décrire

<https://gcmdservices.gsfc.nasa.gov/kms/concept/61594015-4ab4-4b38-ae4f-e31a4757b065>

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

<https://www.wikidata.org/wiki/Q25267>

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://services.sandre.eaufrance.fr/Courdo/Fiche/client/fiche_courdo.php?CdSandre=O04-0400

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

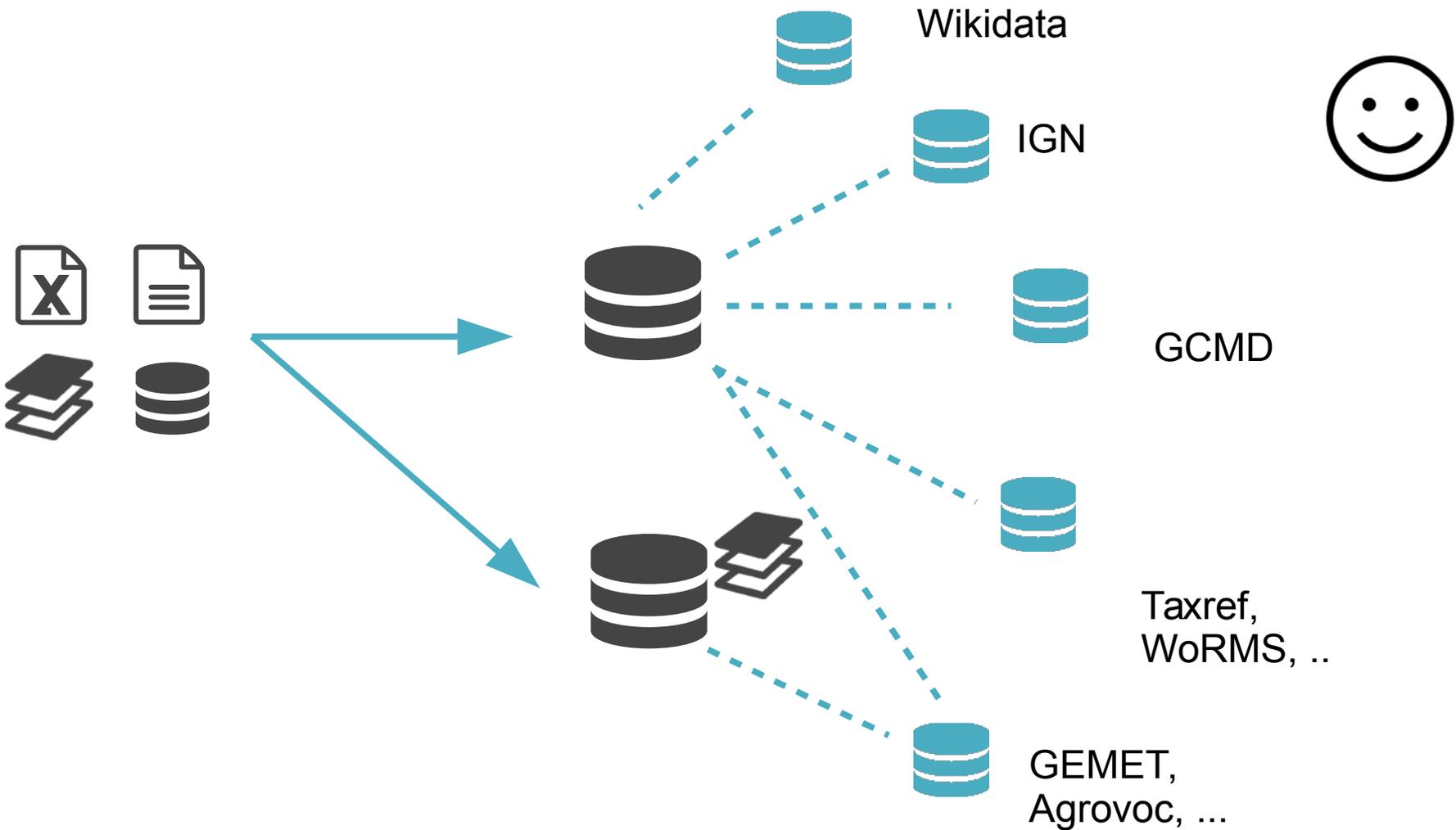
<https://www.wikidata.org/wiki/Q1333963>

<https://gcmdservices.gsfc.nasa.gov/kms/concept/01cc0beb-7c9a-40ed-ad86-0661b41aee53>

<https://www.wikidata.org/wiki/Q1024499>

Cf **présentation SIST18**

Décrire



Diffuser

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

Accessibilité

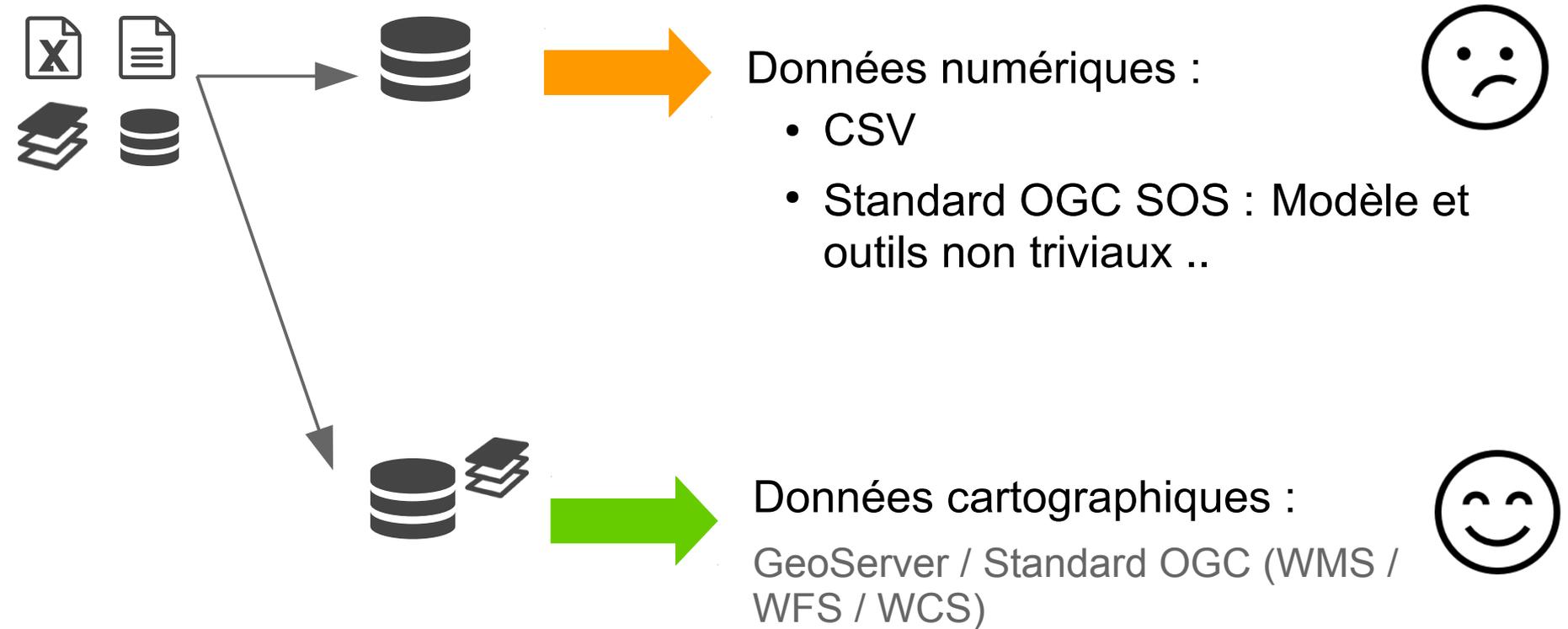
Interopérabilité technique / syntaxique

Ré-utilisabilité

Données : **formats** et **protocoles** standards, ouverts et opérables par des machines – standards OGC

Licences ouvertes :  

Diffuser



Cataloguer, identifier

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

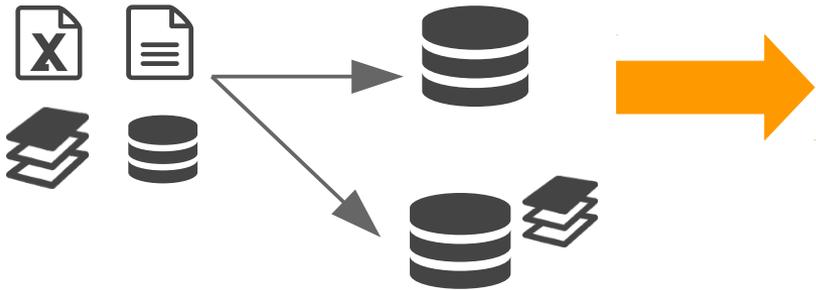
Accessibilité

Trouvabilité

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

Identifiants uniques

Cataloguer, identifier



Métadonnées :

- ISO19115 / Inspire - OGC CSW - GeoNetwork
- Génération auto - R geometa / geonapi
- Mais :
 - Dictionnaires de données ? ISO19110
 - Etre moissonné ..
 - Autres formats ? EML ..



Identifiants : DOI

- Jeux de données dynamiques
- Obligation morale : permettre de retrouver les données dans leur état initial (publication)
- => historiser, versionner .. non trivial



Bilan

- Besoins :
 - Structurer les données
 - Description fines, riches et standardisées **indispensables**
 - Améliorer la standardisation des flux qui sortent des OSU
- Mais .. moyens disparates entre OSU
- Promouvoir les bonnes pratiques : un réseau INSU pour la gestion des données d'observation : **SIST**
- Vers des données liées et intelligentes .. ?
 - Reasonner et créer des connaissances **avec des machines**
 - Web sémantique (RDF, SPARQL, ..)