

Offre de données et de services

Un portail pour les données in situ

I. Braud (UR RiverLy, Irstea, Lyon)
S. Galle, V. Chaffard, P. Juen, C.
Cousot (IGE, Grenoble)

Présentation générale et objectifs

Pourquoi un portail de données in situ sur les surfaces continentales?

- De nombreux observatoires acquièrent de la données in situ et la mettent à disposition via des portails dédiés mais sans harmonisation
- Des données in situ sont aussi acquises pour le calage/validation de données satellites

=> Ces données doivent être visibles et disponibles pour tous

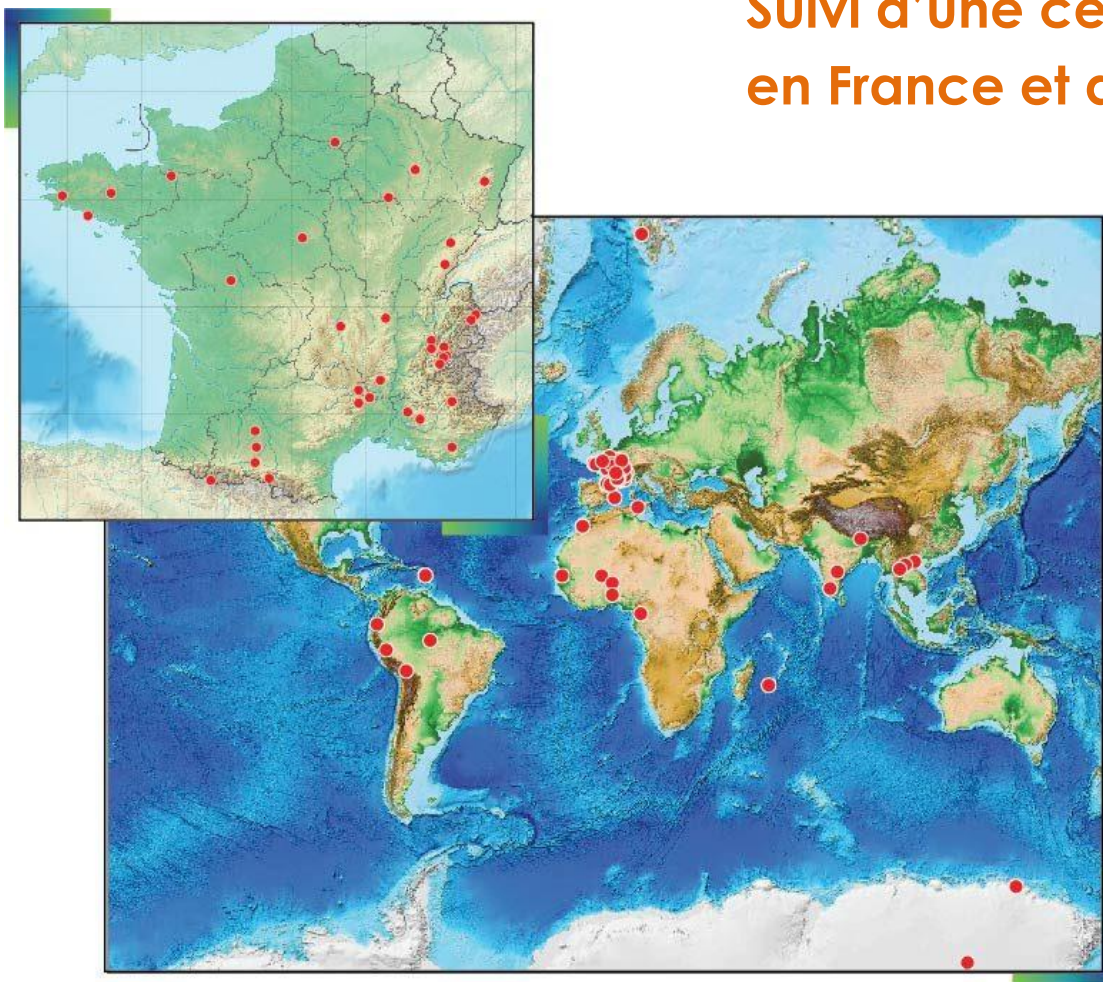
Objectifs

- **Construire un portail unique** d'accès aux données in situ d'observation des surfaces continentales qui permette un accès **transparent** aux utilisateurs
- Ne pas dupliquer les données qui sont déjà archivées par ailleurs mais les rendre visibles
- Respecter les **critères d'accessibilité et d'interopérabilité** (données FAIR – Findable Accessible Interoperable Reusable). Utiliser un vocabulaire contrôlé. Encourager la déclaration de **DOI de données**
- Permettre l'animation des communautés pour la création de **nouveaux produits** (CES) à partir des observations de base in situ et satellitaires

=> Construire un premier prototype en s'appuyant sur la diversité des observatoires de l'Infrastructure de Recherche OZCAR (Observatoires de la Zone Critique. Applications et Recherche)

Un réseau de 21 observatoires

Suivi d'une centaine de sites
en France et dans les pays du Sud



De quelles données parlons-nous ?

Quels observatoires ?

- Observatoires labellisés par les organismes de recherche (CNRS, IRD, Irstea, INRA, BRGM, Universités) + réseau de suivi de la qualité de l'eau souterraine (ADES)
- Observatoires initialement construits pour répondre à des questions de recherche en lien avec les territoires où ils sont implantés: crues méditerranéennes, pollutions agricoles en Bretagne, pluies acides dans les Vosges, érosion des sols, changements globaux, etc...
- Structuration de l'ensemble des observatoires dans des infrastructures de recherche nationales (IR OZCAR), mais aussi européennes (eLTER), favorisant les recherches croisées
- La labellisation est conditionnée à la mise à disposition des données

De quelles données parlons-nous ?

De quelles variables parlons-nous?

- Etablissement de bilans et calcul de flux d'eau, d'énergie, de sédiments et de matière (éléments chimiques, contaminants) dans les bassins versants
- Transport d'eau, de solutés, d'éléments réactifs dans les nappes, les milieux fracturés, et les karsts
- Monitoring des processus des glaciers, de la neige et du permafrost
- Bilans d'eau, d'énergie, de gaz à effet de serre dans des territoires agricoles avec un focus sur l'utilisation de la télédétection haute résolution
- Caractérisation des milieux et du sous-sol: MNT haute résolution, occupation du sol, pratiques agricoles



Stratégie de construction du portail

Mise en route du projet

- Mise en place de l'équipe projet en Juin 2017
- Tour de France des 21 observatoires pour échanger avec les scientifiques et les développeurs des bases de données/portails autour de leurs attentes
- Atelier en avril 2018 sur les fonctionnalités attendues de l'interface du portail web
- Contacts avec les responsables du site Hydroweb du CNES mais pas d'interactions en pratique sur la construction du portail Theia/OZCAR

Enseignements tirés de ces rencontres

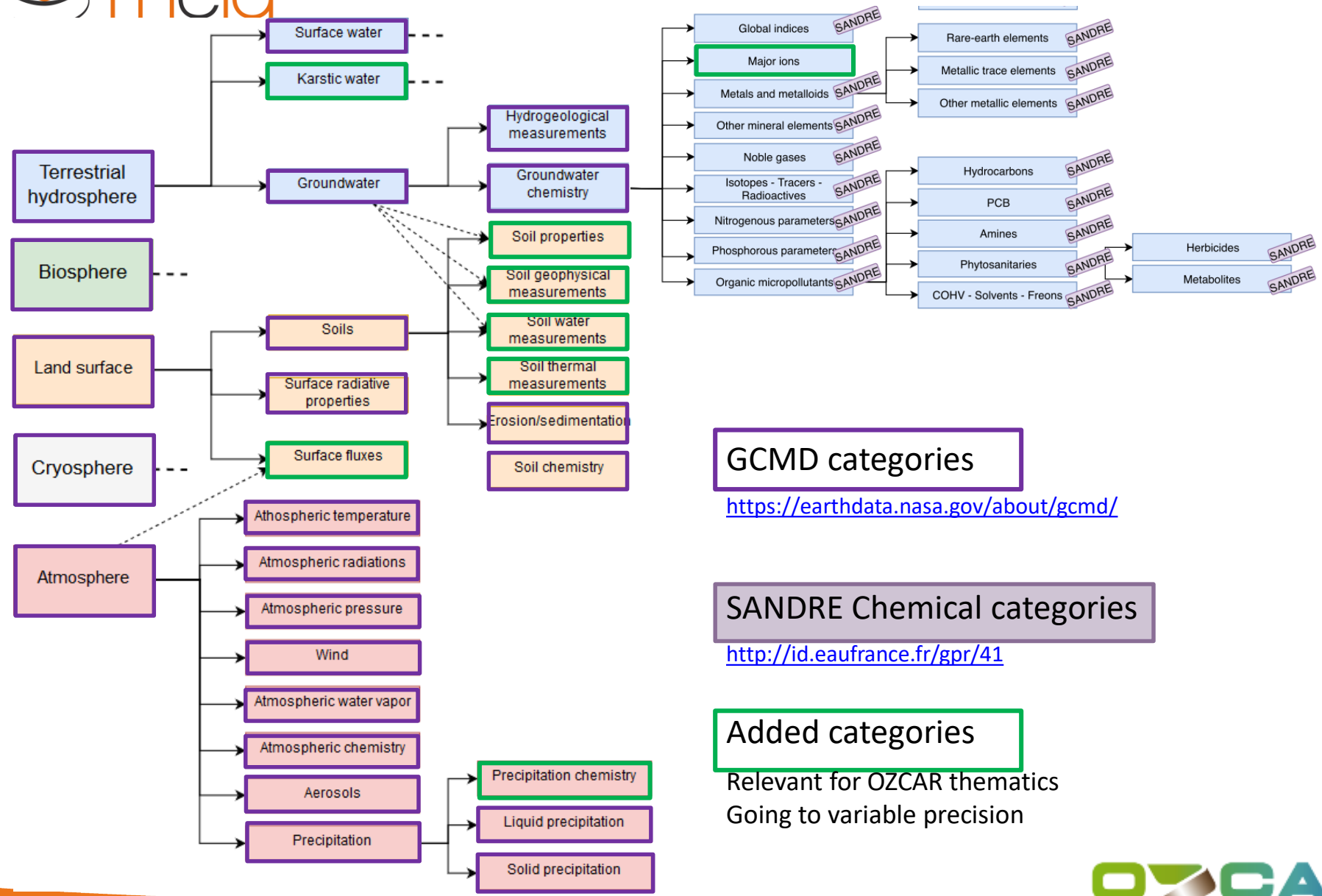
- Grande hétérogénéité dans la manière de gérer les données, définir les jeux de données
- Importance de maintenir la source des données au plus près des producteurs
- Besoin d'accéder aux variables mesurées et rapidement aux données

Avancement du travail

Principes de la construction du portail

- Les observatoires fournissent des métadonnées sur leurs jeux de données, organisés de manière hétérogène, et l'information sur les variables mesurées n'est pas toujours facile à obtenir
 - => d'où un gros travail sur les variables pour la recherche sur le portail
- Inventaire des variables et proposition d'un vocabulaire contrôlé basé sur les catégories du GCMD
 - => Plus de 2200 noms de variables recensés (hors qualité des eaux issues de la base ADES)
 - => Harmonisation des noms en environ 300 variables
- Définition des flux d'informations à mettre en place entre le portail et les observatoires
- Définition de l'architecture du portail et définition d'un format pivot pour le transfert des informations nécessaires à l'affichage des métadonnées respectant les standards (INSPIRE, déclaration de DOI, etc...), mais aussi satisfaisant les souhaits des utilisateurs en termes de critères de recherche (variables, climat, géologie, financeurs, tutelles, etc..)

Hierarchised vocabulary



GCMD categories

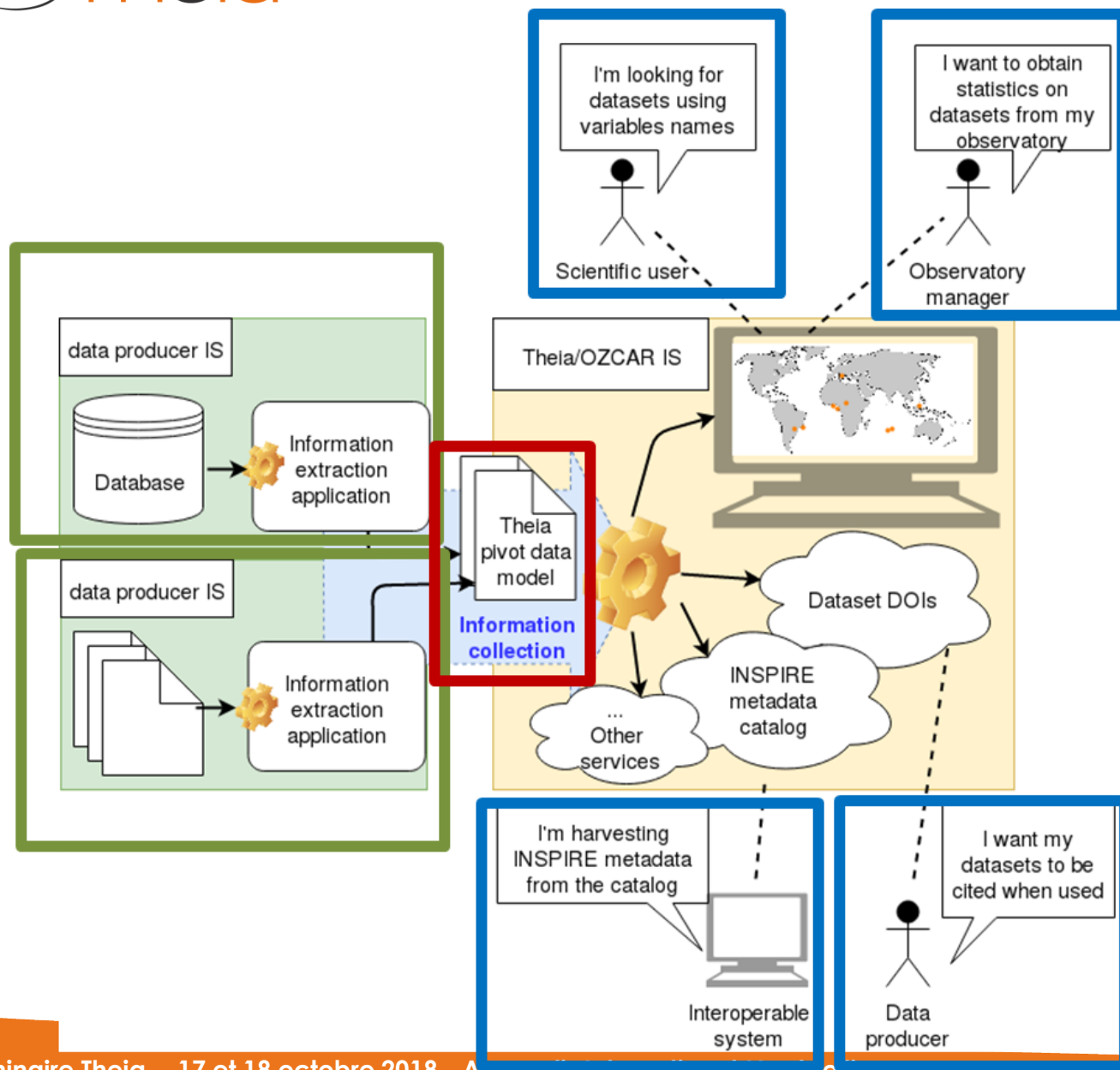
<https://earthdata.nasa.gov/about/gcmd/>

SANDRE Chemical categories

<http://id.eaufrance.fr/gpr/41>

Added categories

Relevant for OZCAR thematics
Going to variable precision



Perspectives

- **Prochaines étapes**

- Mise en place des flux d'informations entre les observatoires et le portail Theia/OZCAR
- Un premier prototype de portail d'ici la fin de l'année
- Atelier prévu fin janvier-début février pour discuter de la granularité des jeux de données pour affichage sur le portail ou déclaration de DOIs

- **Un portail pour quels utilisateurs?**

- Les scientifiques des observatoires eux-mêmes et toute la communauté scientifique
- Une réflexion à engager sur des produits qui puissent aussi être utiles aux gestionnaires (réflexion à inscrire dans la réflexion à venir sur l'infrastructure de recherche européenne eLTER)

Merci pour votre attention

