Building the Information System of the OZCAR French Critical Zone Observatories network: principles and first prototype

Isabelle Braud, C. Coussot, V. Chaffard, S. Galle, P. Juen

ESSI2.1 - Metadata, Data Models, Semantics, and Collaboration
Poster X1.73

Building the common data portal of the OZCAR French Critical Zone Observatories network: principles and first prototype

Isabelle BRAUD[1], Charly Coussot[2], Véronique Chaffard[2], Sylvie Galle[2], Patrick Juen[2] and The Theia/OZCAR ITTeam
1 : IRSTEA Riverly Lyon, France; 2: IGE Grenoble, France

1. Context and objectives

Context: the OZCAR Research Infrastructure

- OZCAR: a network of 22 observatories monitoring various compartments of the critical zone all over the world (about 60 sites, Fig. 1) (Gaillardet et al., VZJ, 2018)
- Data presently scattered in various portals or data repositories
- A large diversity of observed variables:time series, gridded and vector 2D data
  - ➢ Water, energy, sediment and matter transport, geochemistry in catchments
  - ➢ Transport of water, solutes and reactive elements in groundwater, fractured or karstic aquifers
  - ➢ Glaciers, snow and permafrost processes
  - ➢ Soil profiles/carrots and sampling specimen
  - ➢ Geophysical data (2D)
  - ➢ Vector GIS data
  - ➢ Surveys (agricultural practices)
  - ➢ Raster data and remote sensing products: Land use, land cover, high resolution DTM

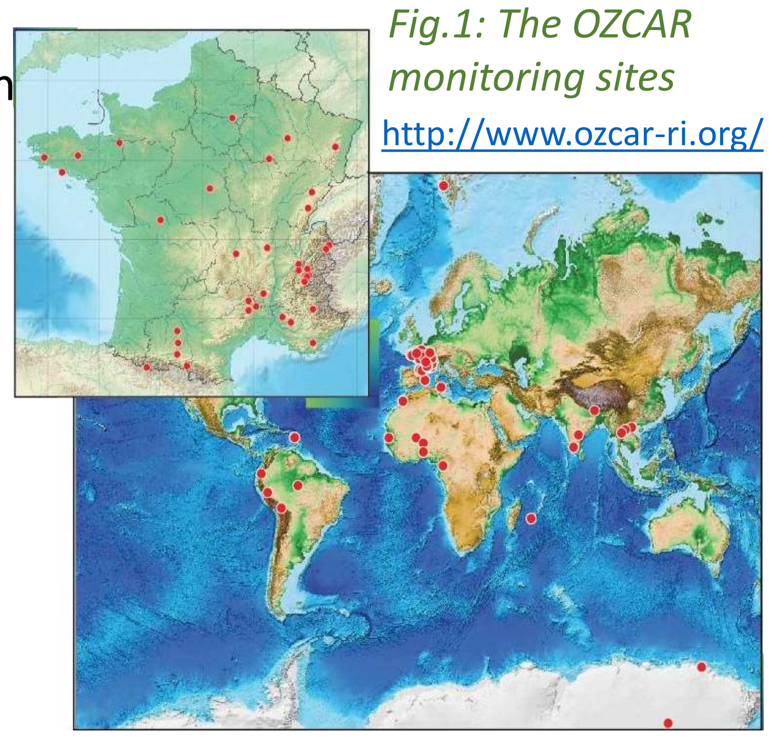


Fig.1: The OZCAR monitoring sites
http://www.ozcar-ri.org/

Objectives of the Theia-OZCAR Information System (IS)

- Theia: one of the French data pole dedicated to continental surfaces, exposing presently mainly remote sensing data
- A need to extend the Information System (IS) to in situ data (Galle et al., 2018)

⇒ OZCAR-RI data: a complex enough case study to design and test the IS

Objectives of the project

- Propose a unique data portal to access in situ data documenting continental surfaces and the critical zone, that are presently scattered in various information systems
- Keep the data close to data producers and make the best use of existing data management systems
- Define information fluxes between observatories and the Theia/OZCAR IS (distributed architecture)
- Be useful to observatories managers (identification of data users, publications using data, putting DOI on data sets, etc…)
- Offer services and interoperability with other portals
- Foster data discovery and exploration, their sharing and reusability, their citation (Fig. 2)

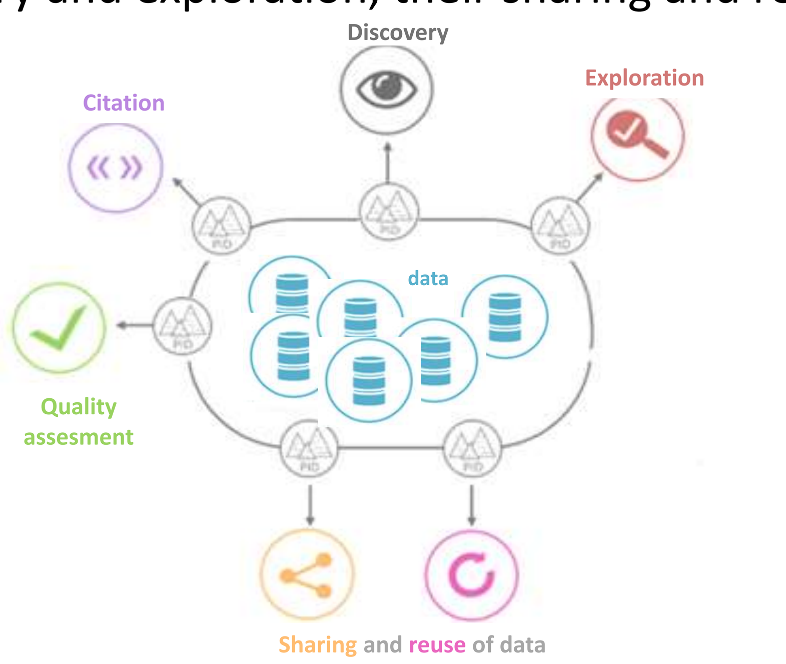


Fig. 2: Illustration of the advantages of open data.

Photography encouraged

http://doranum.fr/thematique-identifiants-perennes-pid/

2. Methodology

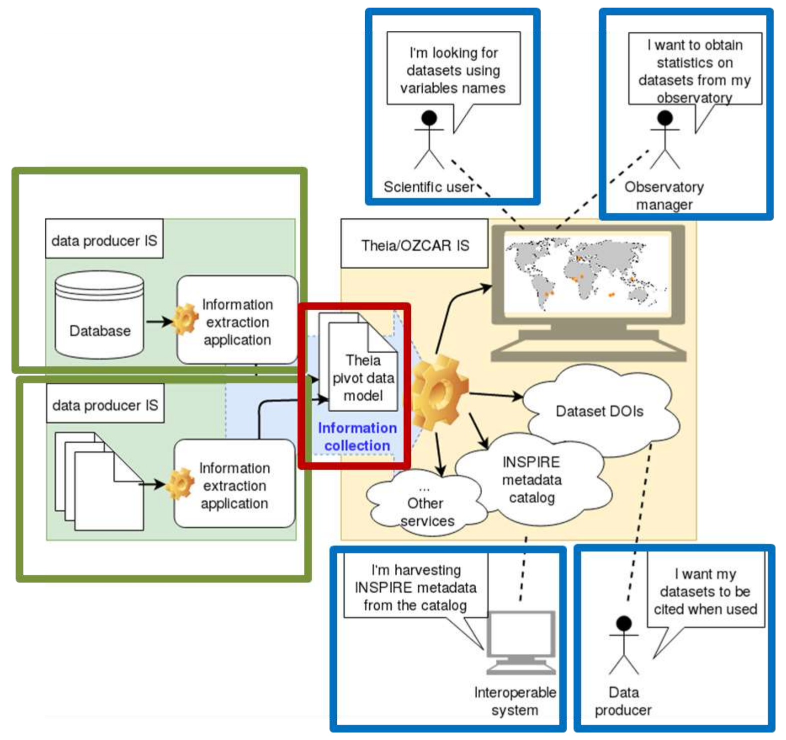


Fig. 3: Scheme of the Theia/OZCAR IS: producers with different databases (green) push and update the information on their data sets using the pivot model (red). The Theia/OZCAR IS responds to human or machine requests and can implement services following standards (blue).

System uses cases

- User-oriented approach: dialog between scientists & IT teams of the source observatories (working groups, meetings)
- Story mapping (AGILE method)

Building strategy

- Definition of common controlled vocabulary based on domain-specific thesaurus
- Open source softwares
- Definition of common metadata based on metadata standards
- Put in practice FAIR principles
- Organization of data fluxes between data producer IS and Theia/OZCAR-IS using a **pivot data model** (Fig. 3)

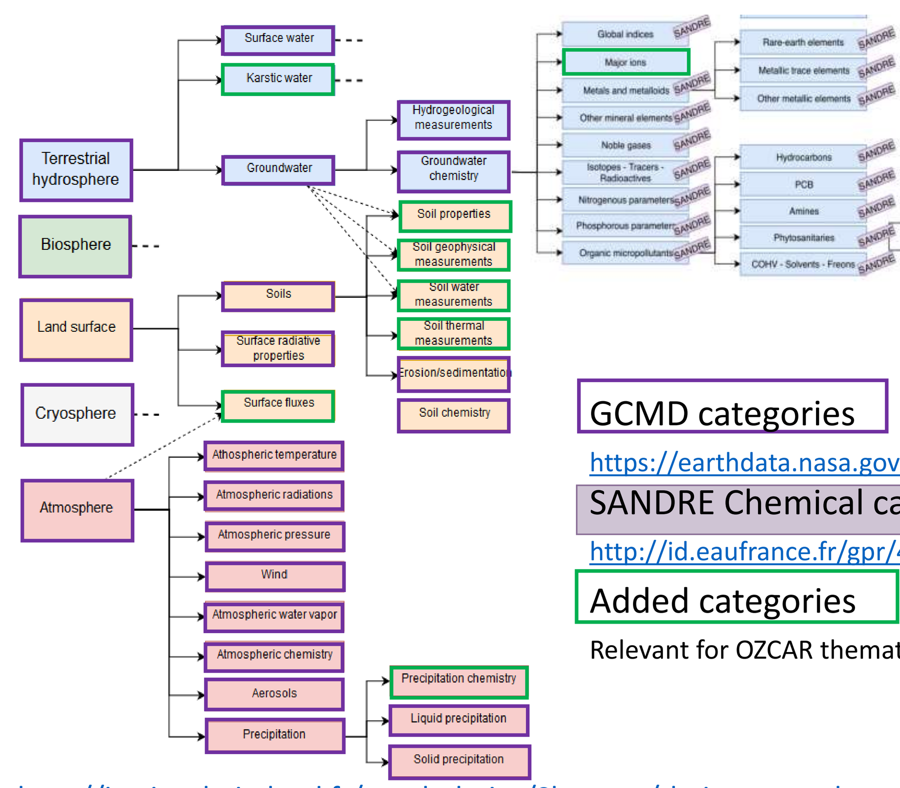
3. Implementation

Choice of a common Theia-OZCAR hierarchized vocabulary (Fig. 4)

- Controlled vocabulary based on hierarchized concepts of GCMD Earth Science keywords thesaurus and Sandre thesaurus for water chemistry
- Harmonized variables names and categories

⇒ From about 3000 data producers names to about 250 controlled names

⇒ Publication in Skos using SKOSMOS



http://in-situ.theia-land.fr/skosmos/theia_ozcar_thesaurus/en/index

Fig. 4: Categories retained in the controlled vocabulary of the Theia/OZCAR data portal.

Definition of a pivot model to build information flux (Fig. 5)

- Concepts and information based on metadata/data exchange standards (ISO19115/Inspire, DataCite, O&M) and some system functional requirements

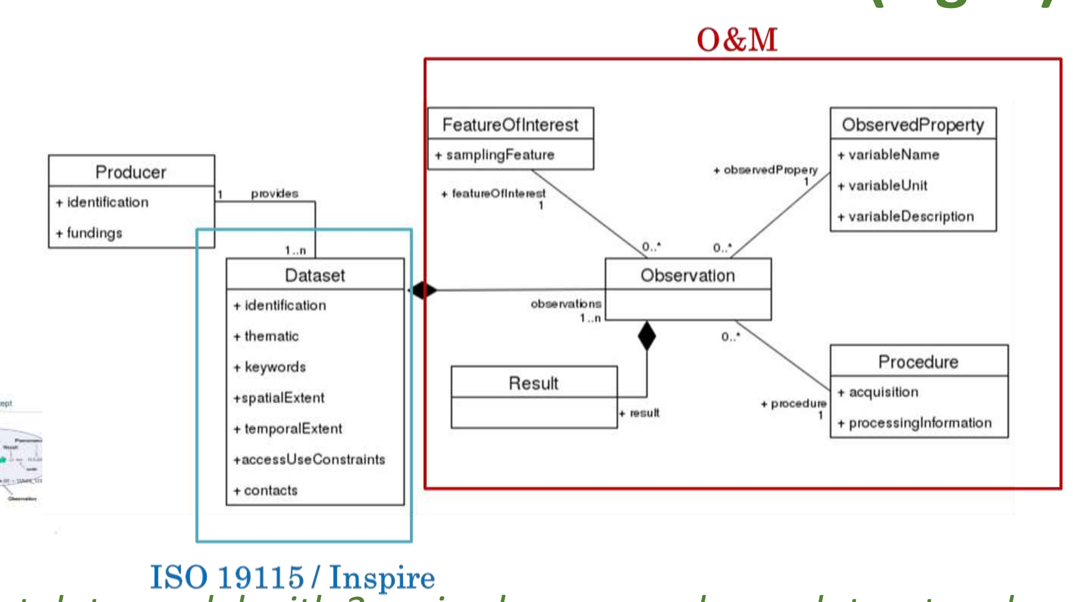


Fig. 5: Conceptual scheme of the pivot data model with 3 main classes: producer, dataset and observation. Information in the blue square corresponds to ISO19115/Inspire information. Information in the red square corresponds to Observation and Measurement (O&M) standard.

4. System architecture and prototype of the web interface

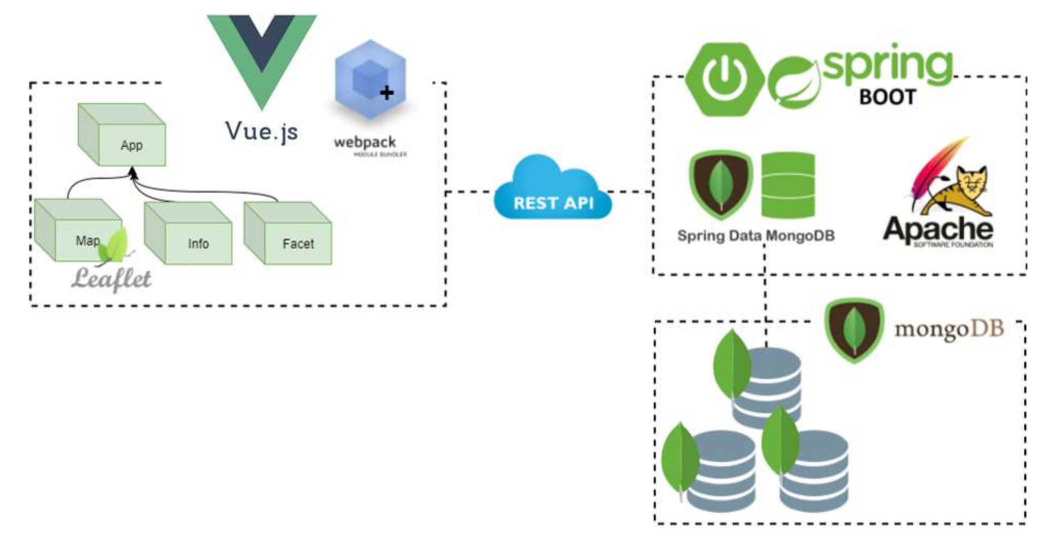
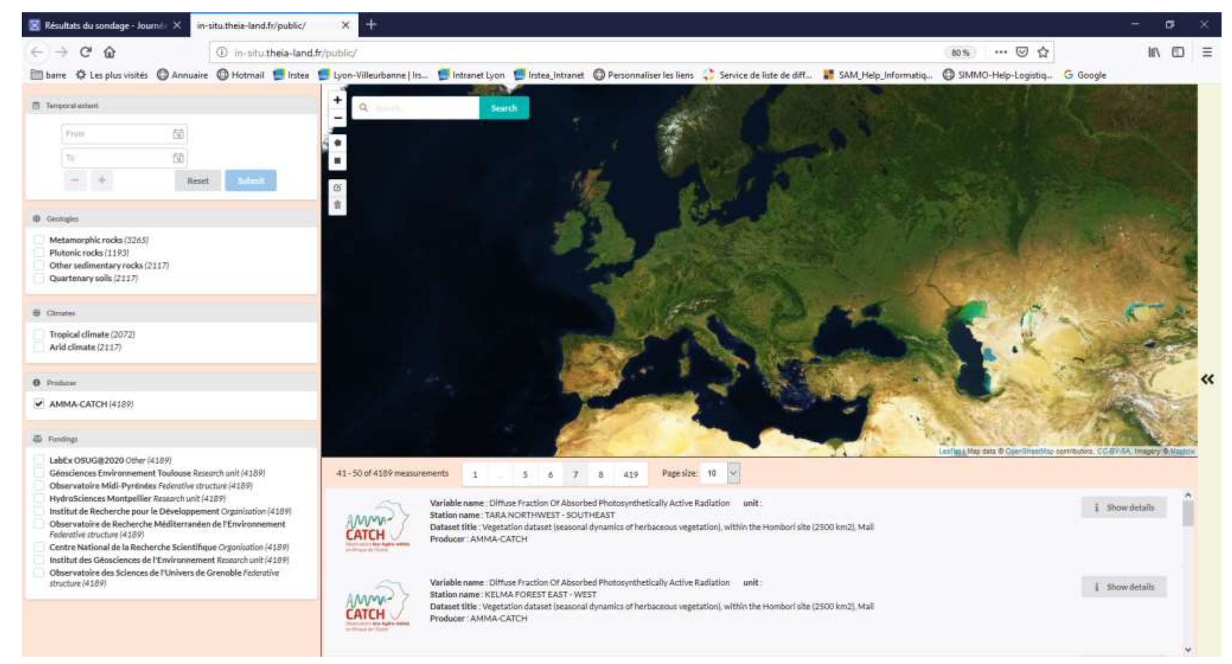
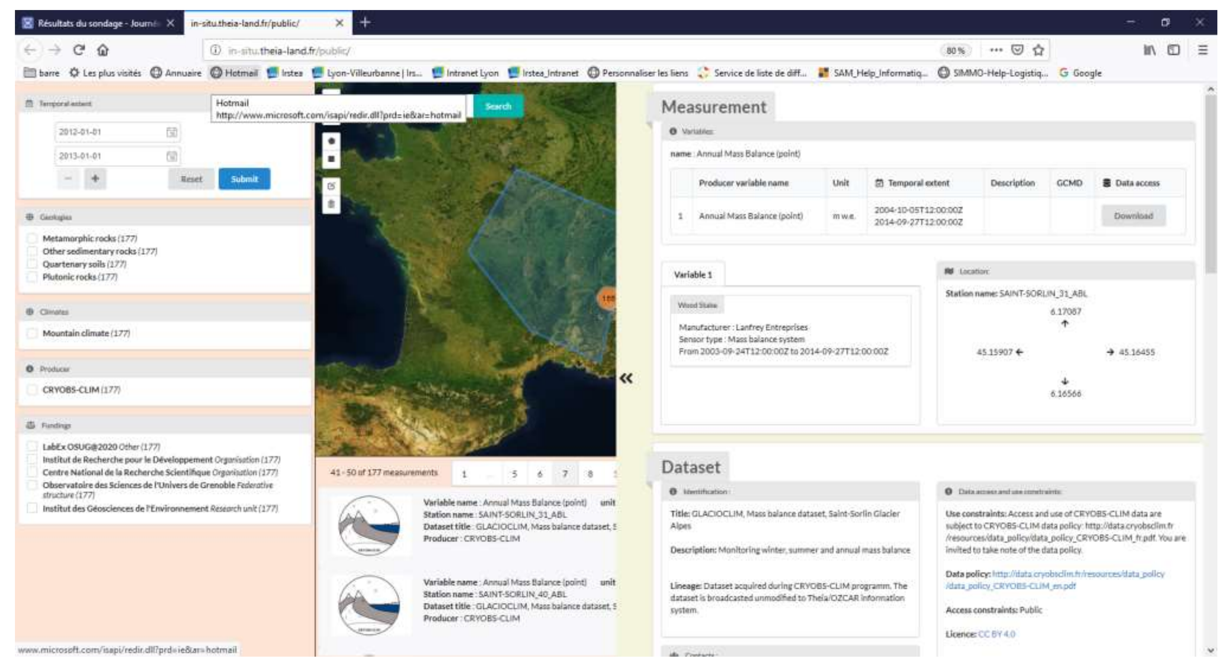


Fig. 6 Architecture of the Theia/OZCAR Information System. The interface is built using Vue.js JavaScript framework. The map view is implemented using Leaflet JavaScript library. The backend is a SpringBoot application embedding a Apache Tomcat server to publish the REST API on the web and Spring Data MongoDB solution to link the database. The data persistence is ensured using MongoDB replica-set.





5. Next steps

- Continue the implementation of the information flux between the observatories and the Theia/OZCAR IS (5 observatories included up to now)
- Continue the implementation of search functionalities on the web interface: adding search based on variable name thesaurus
- Test if the web interface is user-friendly with future users
- Put the system in production for metadata (end of 2019) and extend it to data (2020)

Acknowledgements: OZCAR-RI: French Ministry of Research, French research institutions and universities. Theia/OZCAR IS is part of the Theia land data center and Earth System Research Infrastructure. CNRS/INSU and IRD for funding human resources for the project.

References: Braud, I., Chaffard, V., Coussot, C., Galle, S., Juen, P. and the Theia/OZCAR ITTeam, 2019.: Building the Information System of the OZCAR French Critical Zone Observatories network, Hydrological Sciences Journal, to be submitted; Galle S., et al., 2018 : « Joint portal for THEIA-OZCAR in-situ data »; Theia bull. n°9; http:// theia-land.fr; Gaillardet J., Braud I, Hankard F. and the OZCAR team, 2018 : « OZCAR: the French network of Critical Zone Observatories », Vadose zone J., 2018, doi:10.2136/vzj2018.04.0067.

Contact: Isabelle.braud@irstea.fr