Manage and distribute data in a french environment observatory: the example of OSU OREME and the KARST observatory network

Juliette FABRE and Olivier LOBRY, OSU OREME, CNRS
The SNO KARST team

OZCAR
CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

eLTER
ESFRI

TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

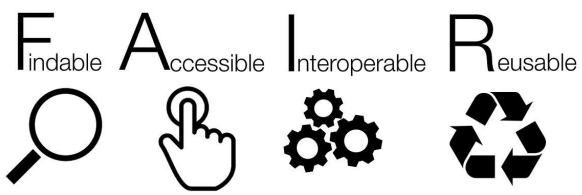
Context

- OSU OREME: Montpellier Research Observatory of Environment
  - 8 labs, 26 Observation Services
  - Support for long-term observation & data dissemination
  - 1 data portal / infrastructure : data.oreme.org with **shared tools**
  - 2 database engineers

- SNO KARST (OZCAR RI)
  - 9 observatories
  - **multidisciplinary time series on karst**
  - 600 stations & 150 parameters
  - 2016 : OREME in charge of the data

Needs and constraints

- SNO KARST
  - needs:
    - data **visualisation** / **download** / **citation** (DOI)
    - integration of time series from the **french hydrological bank**
  - but: **little time** available for data management 🤯

- National requirements : 

- OZCAR RI : **feeding of the data portal** => semantic and syntax **interoperability**

Methods

- Project management
  - Small team : 2 engineers + a scientist + project leaders
  - Iterative development process: development, user feedback, dev .. (still in progress !)
  - Interactions with OZCAR data team (18 months)

Methods

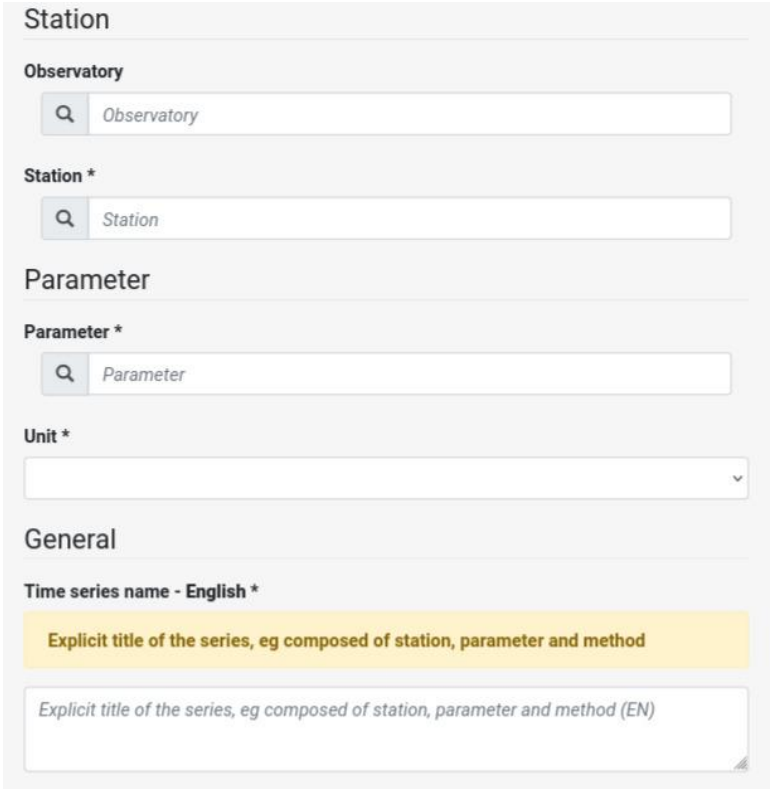
- Database

  - Describe data with **detailed metadata**
    - stations, sensors, acquisition methods ..
  - Link descriptions to **existing repositories** => **enrich** and **link** data
    - see Data Management Plan
  - **Align parameters** with Theia-OZCAR / GCMD thesauri

Example of linked repositories

Methods

- Database feeding
  - 1/ **Metadata** web edition tools - PHP / JS
  - 2/ Generation of Excel/CSV sheet **templates** to submit **data** (FTP server) - PHP / R
  - 3/ **Automatic** data integration from the french hydrological bank - Hub'eau API - R



Time serie edition form

	A	B	C	D	E	F
1	datetime	storm_event	conductivity	discharge	fluorescence	water_level
2	YYYY-MM-DD hh:mm:ss	true or false (defaut false)	µS/cm	m3/s	mV	m
3	UTC	empty values authorized	Epenoy site	Epenoy site	Epenoy site	Epenoy site
4			Quality-controlled data	Derived products	Raw data	Quality-controlled data
5			auto	auto	auto	auto
6			epeCTD	epeCTD	epeGGUN	epeCTD
7						
8			2201	2202	2203	2199
9			Conductivity at Epenoy stati	Discharge at Epenoy station	Fluorescence at Epenoy sta	Water level at Epenoy statio
10						

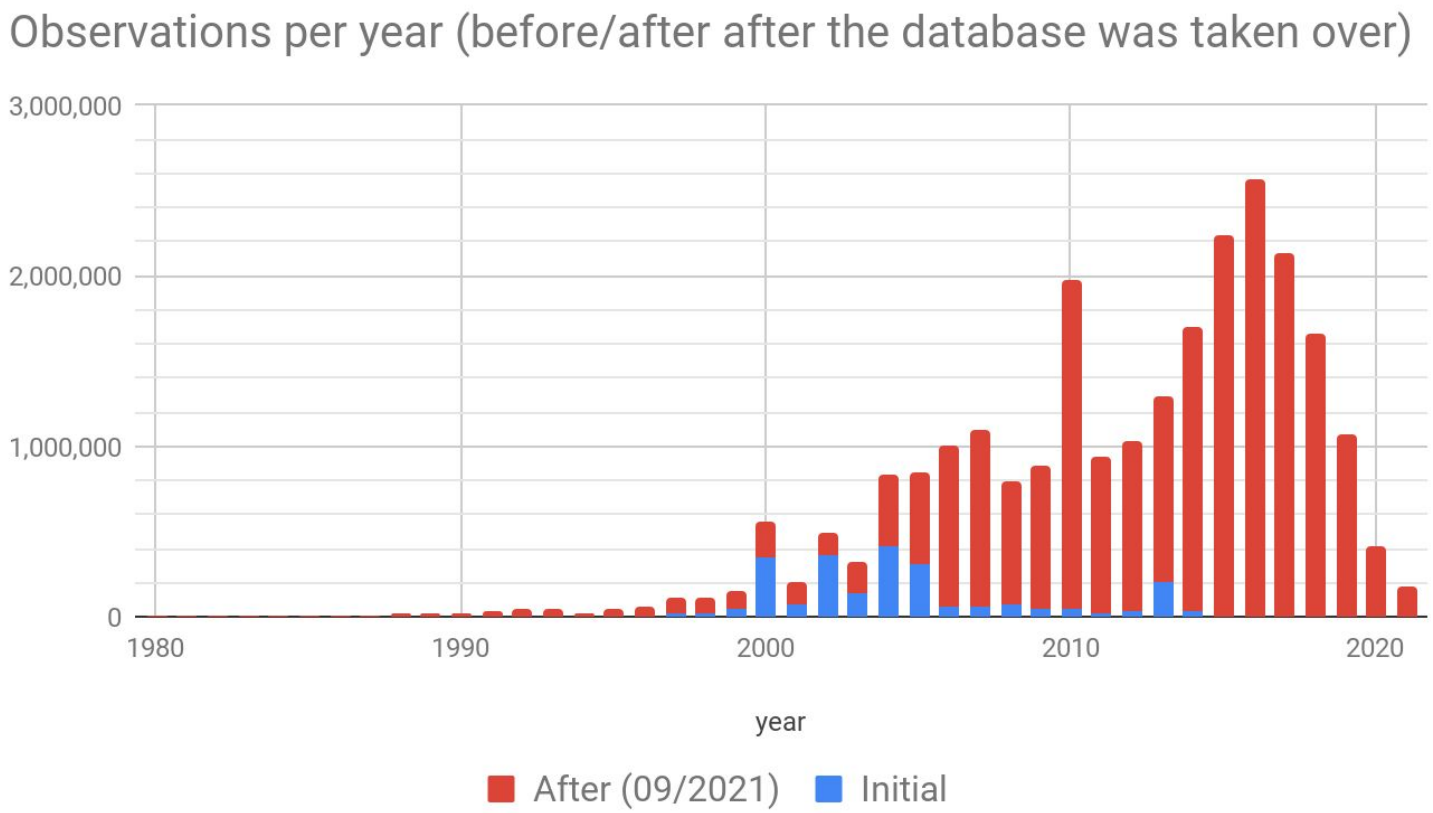
Generated excel / CSV template

TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

OZCAR
CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

Results

- A living database
  - 600 public time series (> 14 M of data records)
  - Thanks to metadata edition / data feeding tools & motivated scientists!



OZCAR
CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

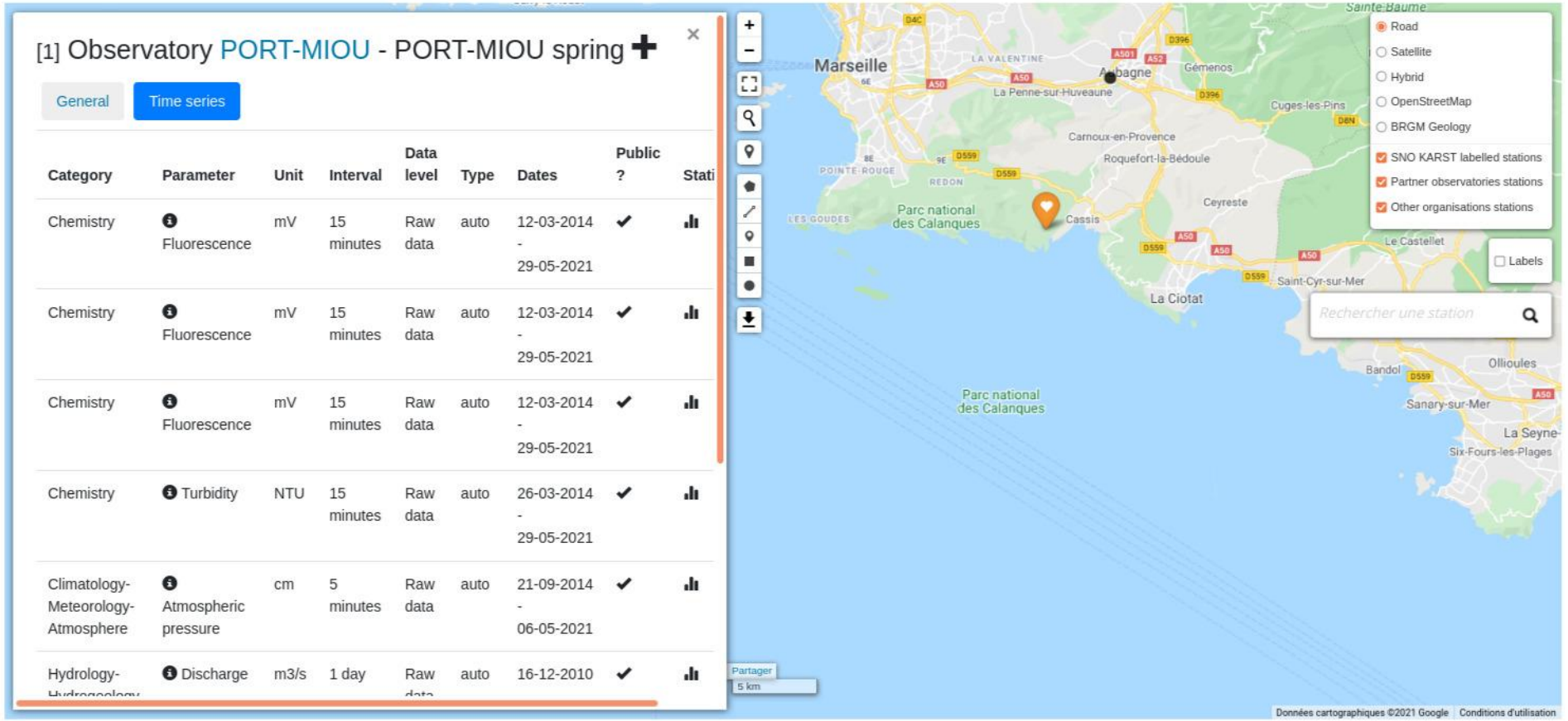
Results

- Accessible data / metadata

Map server

GeoServer

- Standardized map server - WMS / WFS
- Web interactive maps - PHP / leaflet JS

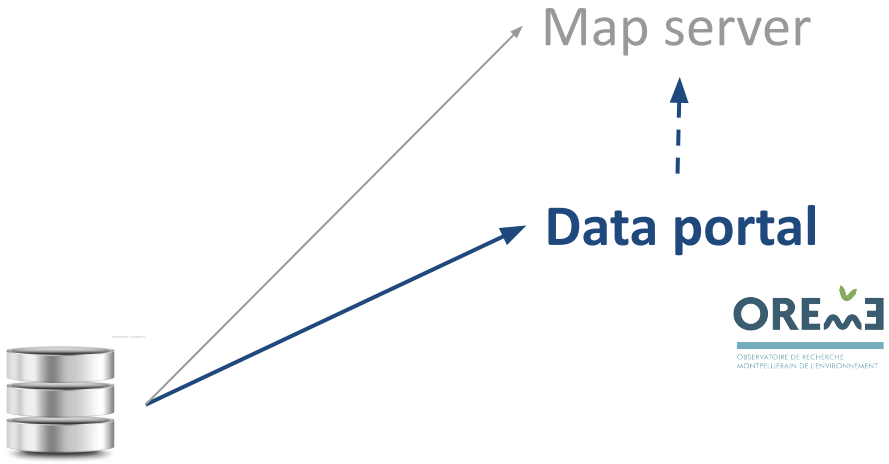


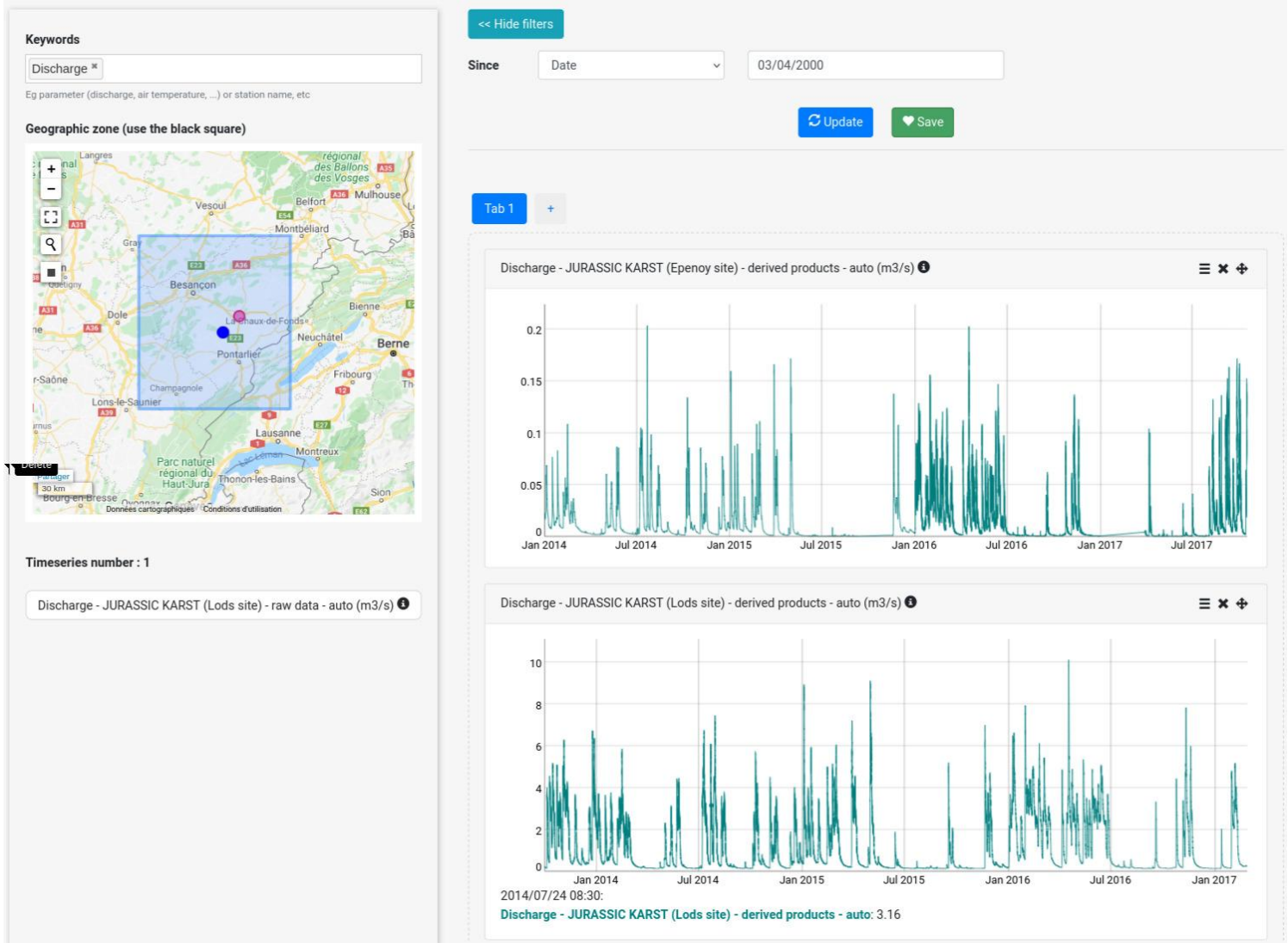
TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

OZCAR
CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

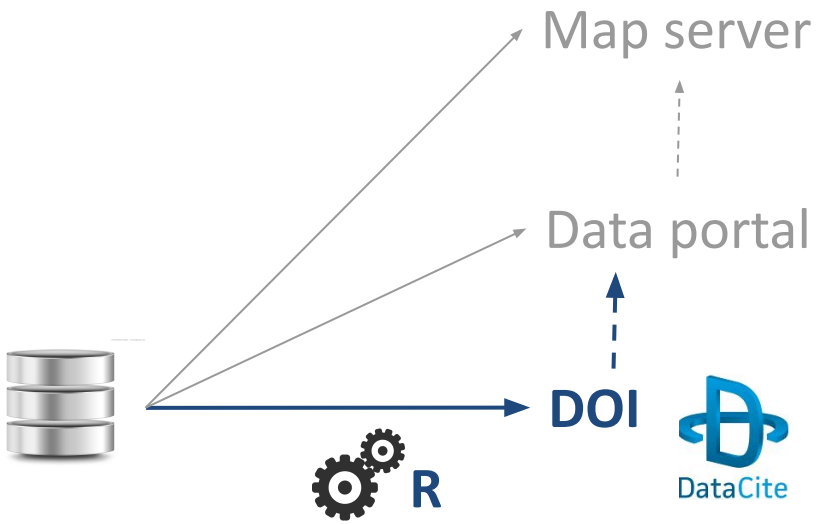
Results

- Web tools: data search / download / visualisation - PHP / JS





Results



- 30 DOI on datasets
- dataset = *category of parameters* watershed*
- Automatic DOI generation - Datacite API - R

Citation

SNO KARST (2019). Time series of type hydrology-hydrogeology in L'Isle basin - KARSTS AQUITAINS observatory - KARST observatory network - OZCAR Critical Zone network Research Infrastructure. OSU OREME. (dataset). DOI: https://doi.org/10.15148/fcf00f41-f22c-47c0-beb4-031255793259

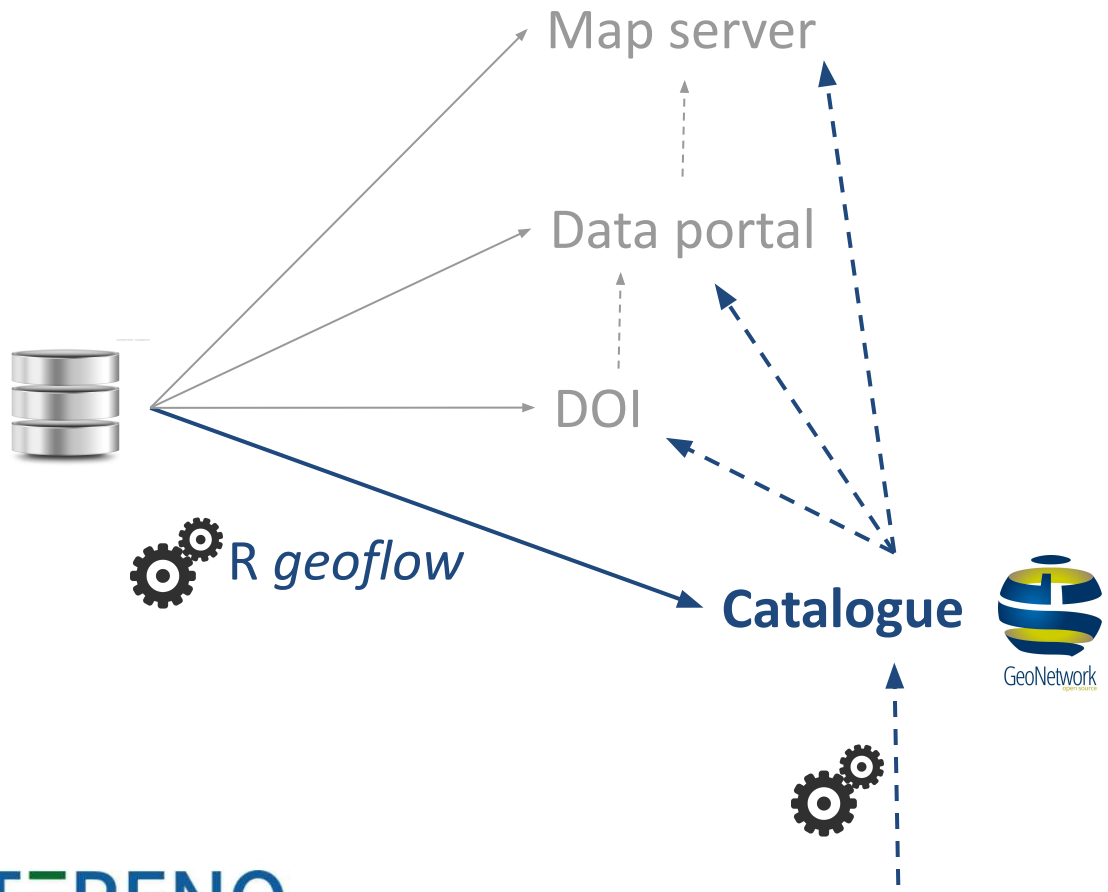
TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

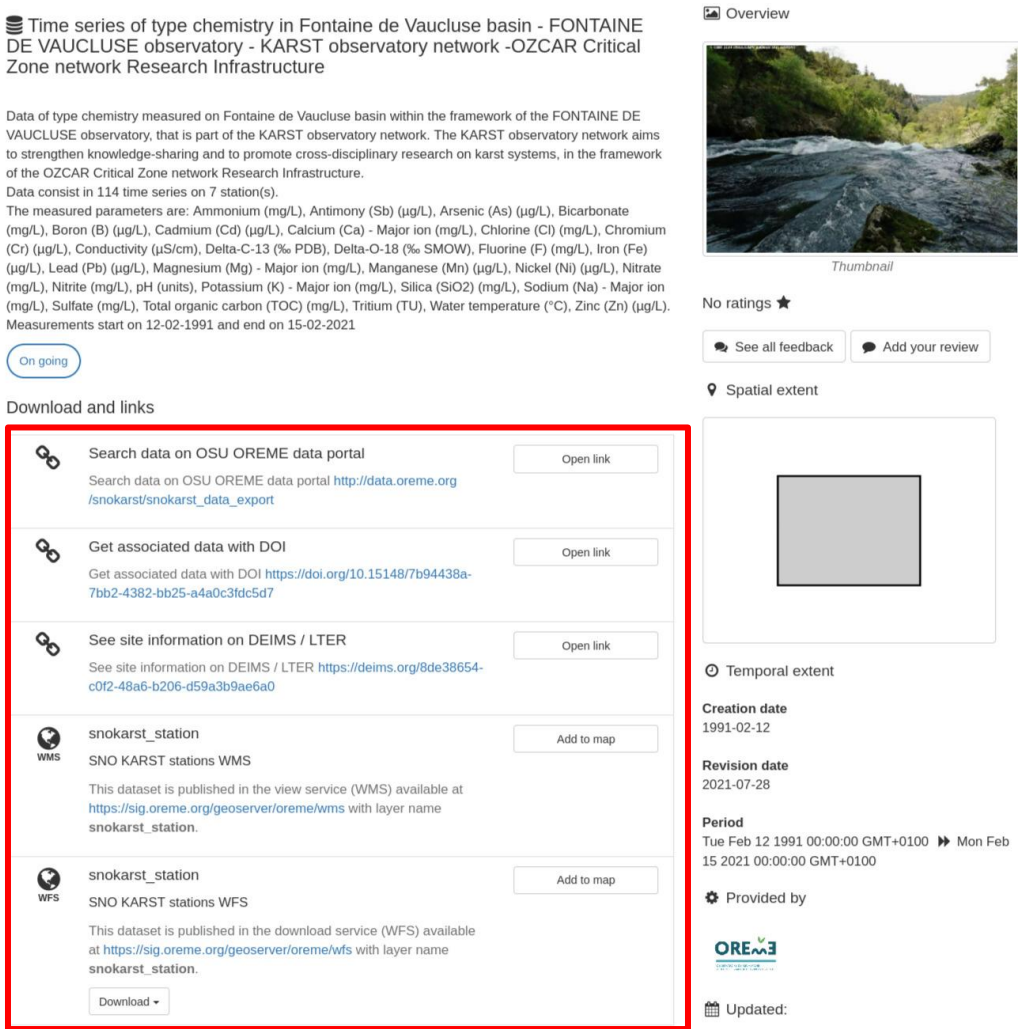
OZCAR
CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

Results

- Standardized catalogue - ISO 19139 / CSW
- Automatic feeding - R *geoflow*



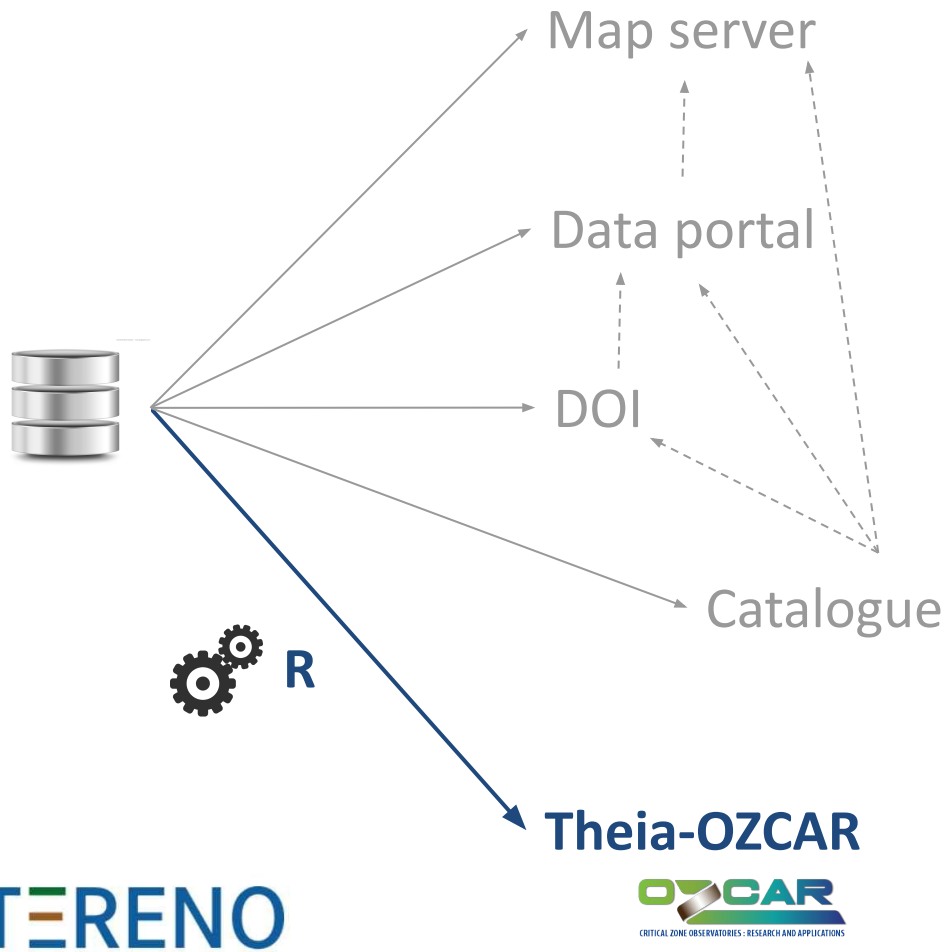


TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

OZCAR
CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

Results



- Automatic feeding of Theia-OZCAR portal - R

TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

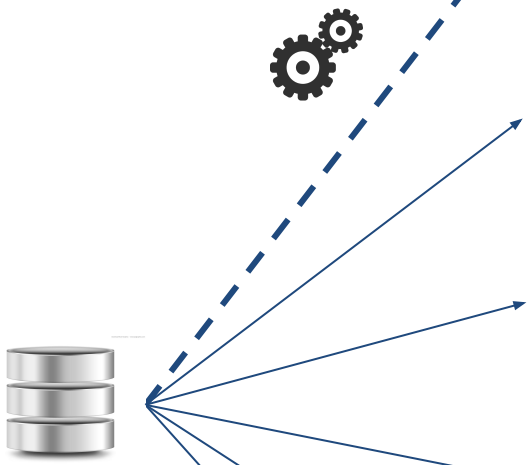
OZCAR
CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

Perspectives

SIST

2 workshops (2021)

Standardized flow for sensor data?



Map server GeoServer

Data portal OREME

DOI DataCite

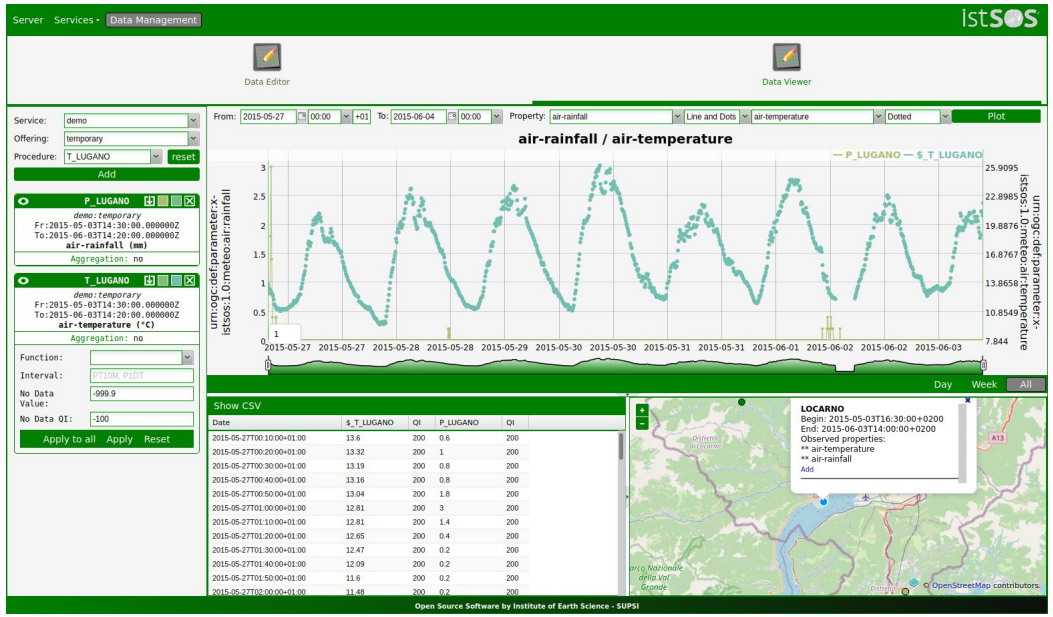
Catalogue GeoNetwork

Theia-OZCAR

TERENO TERRESTRIAL ENVIRONMENTAL OBSERVATORIES

OZCAR CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

- **SOS**: istSOS 52n
  - **clients** for data visualisation and query
  - not sot easy, takes time, interoperability is not so easy (way to model data)
- **SensorThings**:
  - seminar 2022 (BRGM / OFB / Pole INSIDE / SIST)
- Why? who / what will harvest data? 🤔

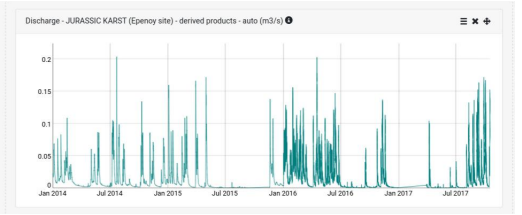


istSOS web interface for data visualisation

Conclusions

- Building an information system takes years 🤓
  => need **permanent staff** for data team and available **time** for scientific team

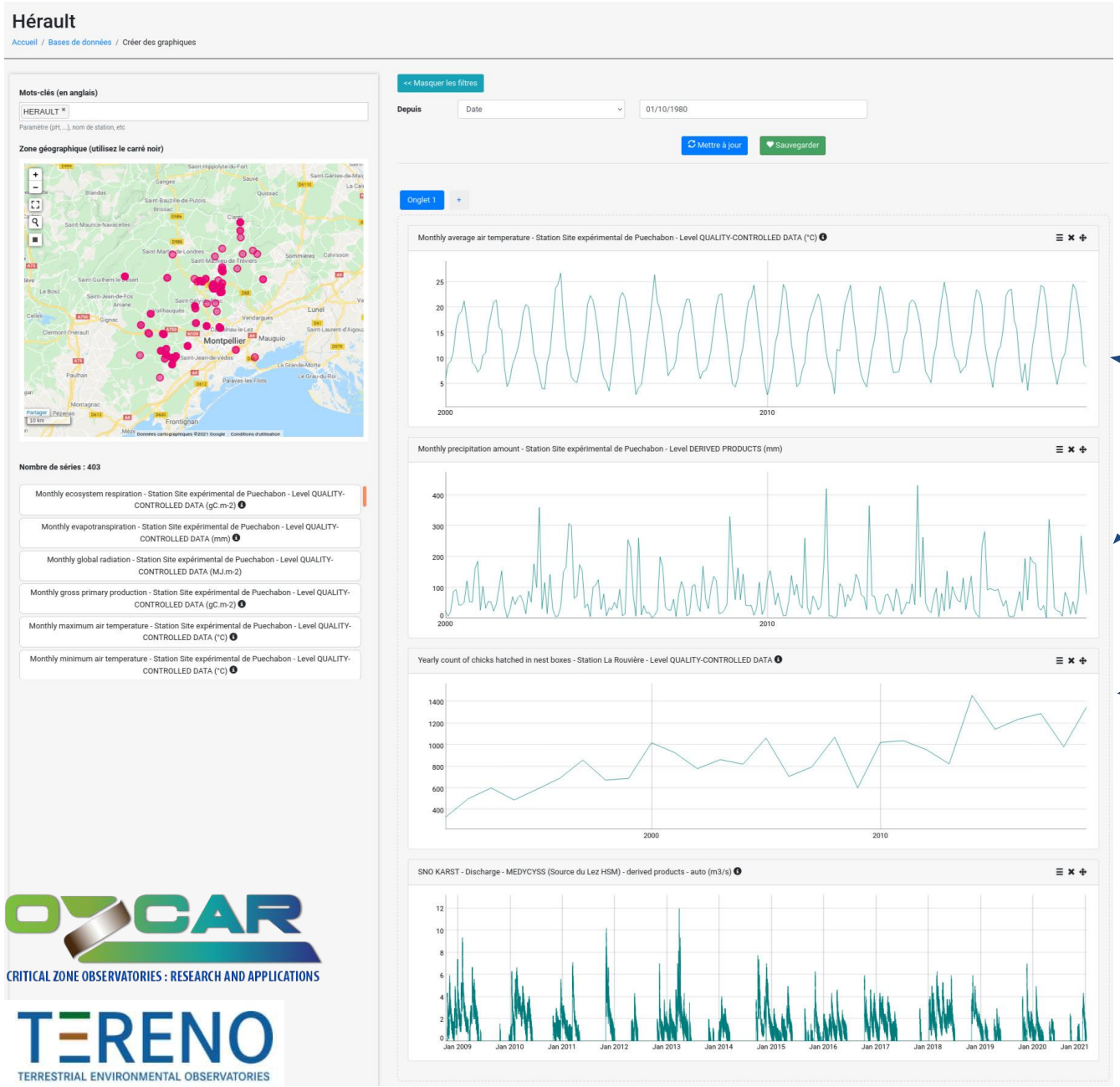
- Database feeding is boring 😭 And only scientists can do it!
  => Develop **tools** with **added values** for scientific team
  => Use generic methods for **database feeding**



- OSU/Observatories rule! 🤩 **Share methods & tools** across databases!
  => less work
  => more interoperability

Cross-reference data from different observation services

OREME
OBSERVATOIRE DE RECHERCHE MONTPELLIÉRAIN DE L'ENVIRONNEMENT



Puéchabon site (TGIR ICOS)

Tits Observation Task (Population Ecology)

SNO KARST

OZCAR
CRITICAL ZONE OBSERVATORIES : RESEARCH AND APPLICATIONS

TERENO
TERRESTRIAL ENVIRONMENTAL OBSERVATORIES