



RÉPUBLIQUE  
FRANÇAISE

*Liberté  
Égalité  
Fraternité*



Geoscience for a sustainable Earth

**brgm**

# THE FUTURE GEOLOGICAL SURVEY: A WORK IN PROGRESS

Matt Harrison

Director of Digital Platforms, BRGM

# Outline of the Talk

---

- Downstream engagement
  - Digital Twins
  - SousSol Bien Commun
- Predictive Services
  - MétéoNappe/Predinappe
- Urban Geoscience
  - BIM exists in a heterogeneous environment
- Embracing New Technologies
  - The value of AI technologies in a GSO

# JUNON Program



-4 years long program starting from 2022

-About 15 millions of euros in investments (ARD, CPER and FEDER)

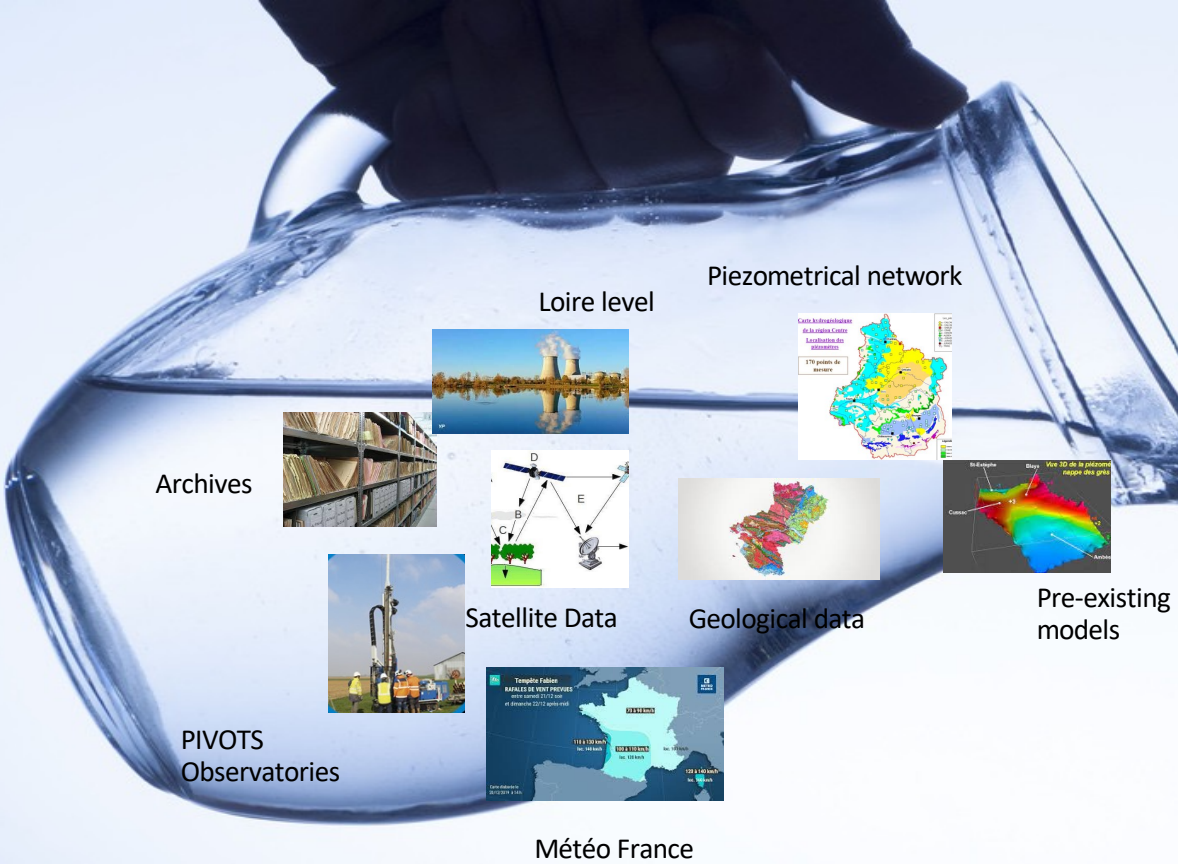
-Some twenty research teams involved (5 research organisms, 4 private companies)

« Create a digital research center on the continental environment (agricultural, urban, forest and fluvial), including each level of the critical zone, and dedicated to research and innovation on all links of the value chain from environmental sensor to plain digital twins. »

## Exceptional Natural resources representative of plain environment!

- Beauce aquifer (biggest European water table)
- Beauce grain belt (second biggest one in Europe)
- Orléans forest
- Loire natural banks and wetlands

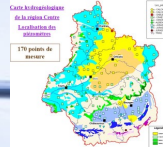




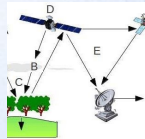
Archives

Loire level

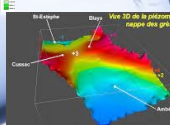
Piezometrical network



Satellite Data



Geological data



Pre-existing models

PIVOTS Observatories



Météo France

Standardization (FAIR)



Deep Learning



# Water Digital Twins (example)

What will be the water table level in 8 months???

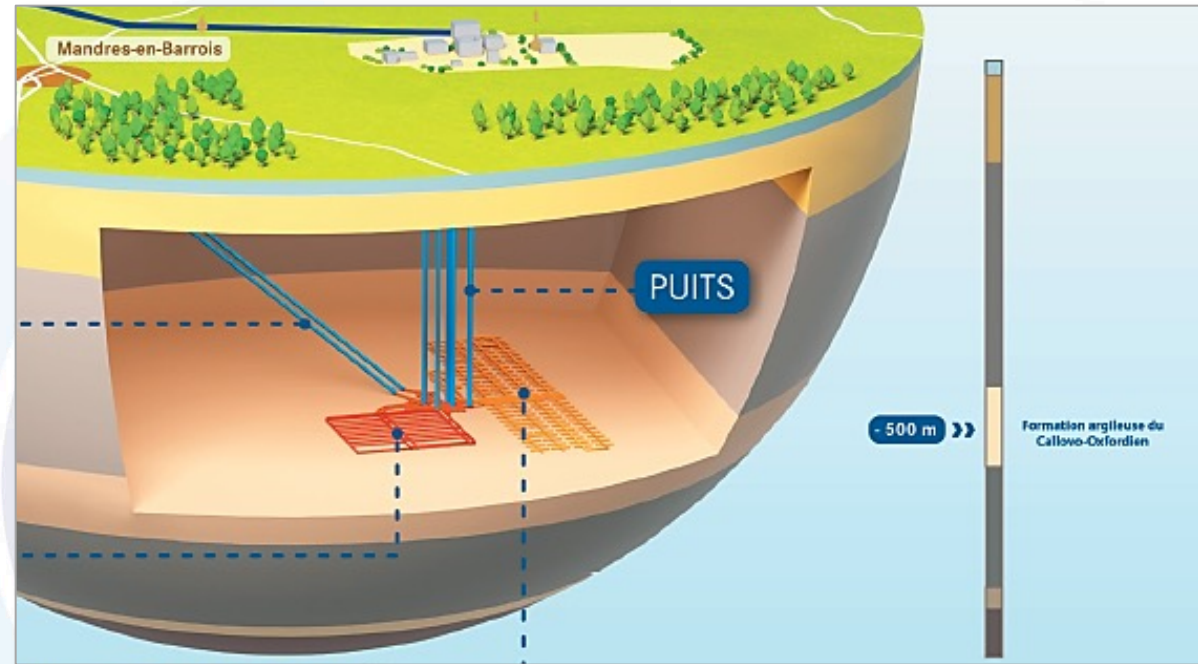
# Digital Twins for surface and underground buildings and infrastructures



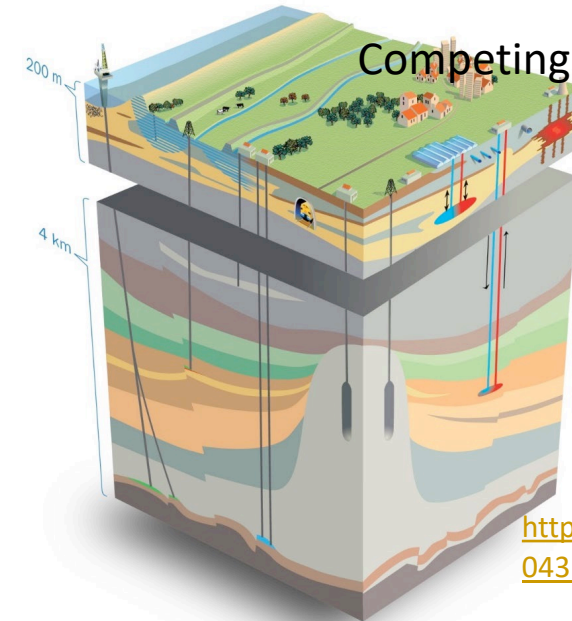
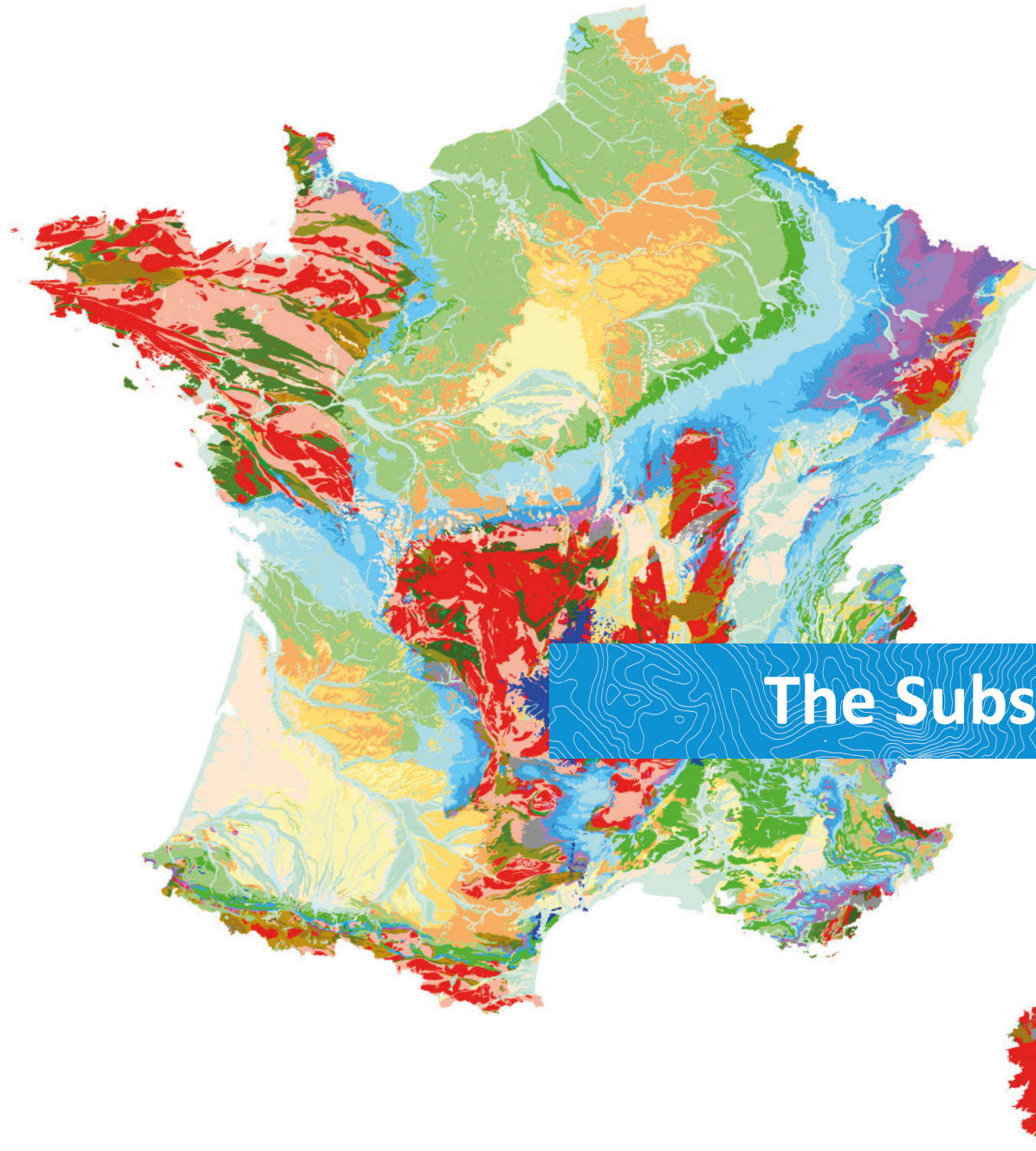
Digital City Synergy project, Helsinki



High Speed 2 Project, London



Geological repository of nuclear waste  
CIGEO Project, France



Competing uses of the subsurface

<https://doi.org/10.1016/j.landusepol.2019.104316>

## The Subsurface as a Common Good

# SURICATE NAT

VIGIE CITOYENNE DES RISQUES NATURELS



Projet soutenu par



Série temporelle ▼

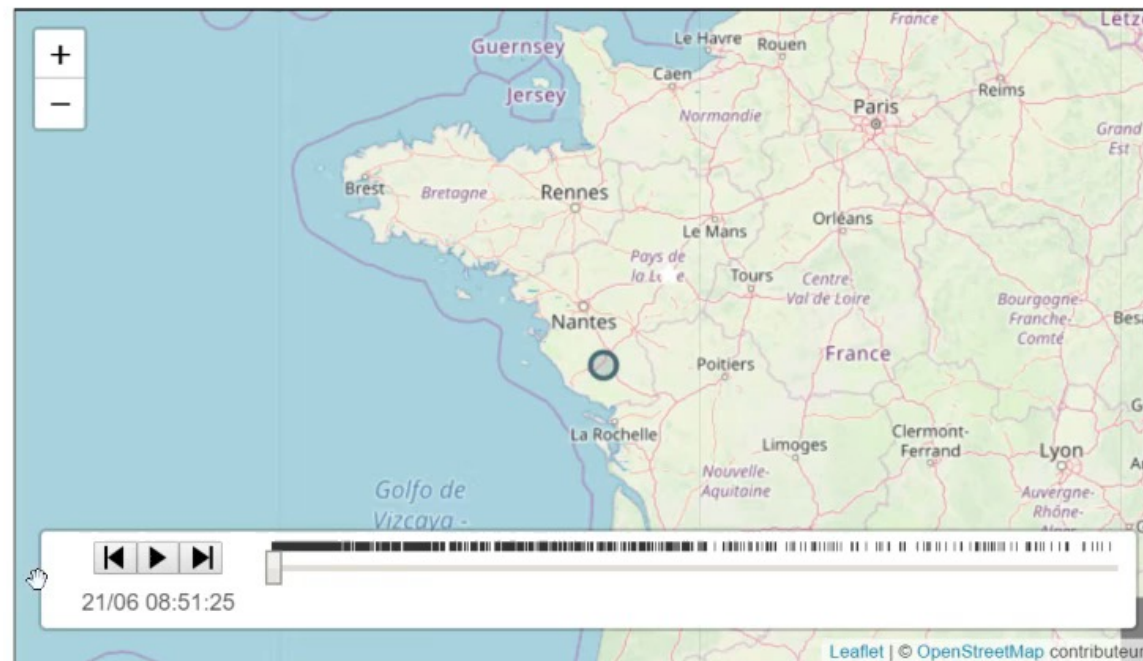
RECHERCHER UN ÉVÉNEMENT

ACTIVITÉ TEMPS-RÉEL TWITTER

FICHES THÉMATIQUES

TOP CONTRIBUTEURS : PODIUM

ADMIN



# 1. SENSOR-BASED DATA-DRIVEN RAPID RESPONSE

- Principle of the SEISAid tool

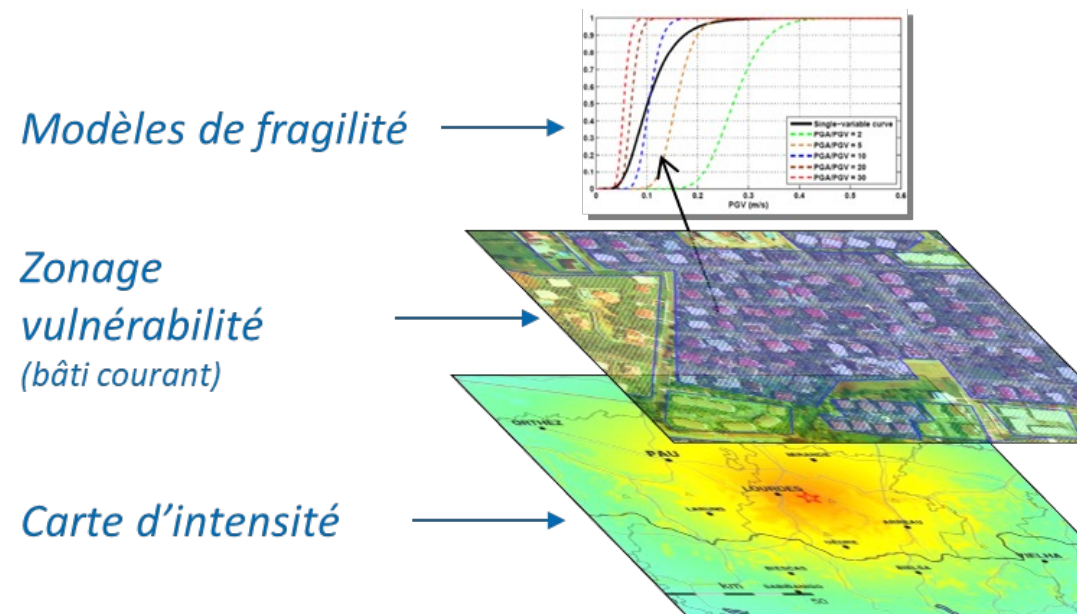
1. Assessment of the intensity of ground motions

- Models PLUS observations (seismometers)

2. Estimated number of buildings affected by intensity level and vulnerability classes

- Statistical model based on the characteristics of buildings (age, constructive mode)

3. Estimation of damages and human losses

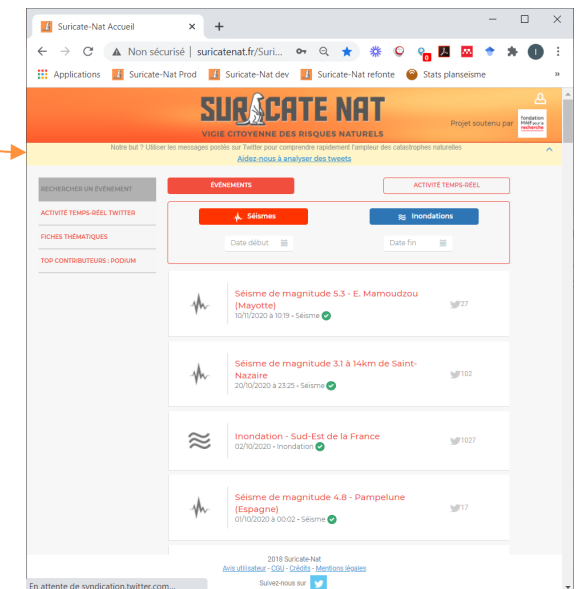
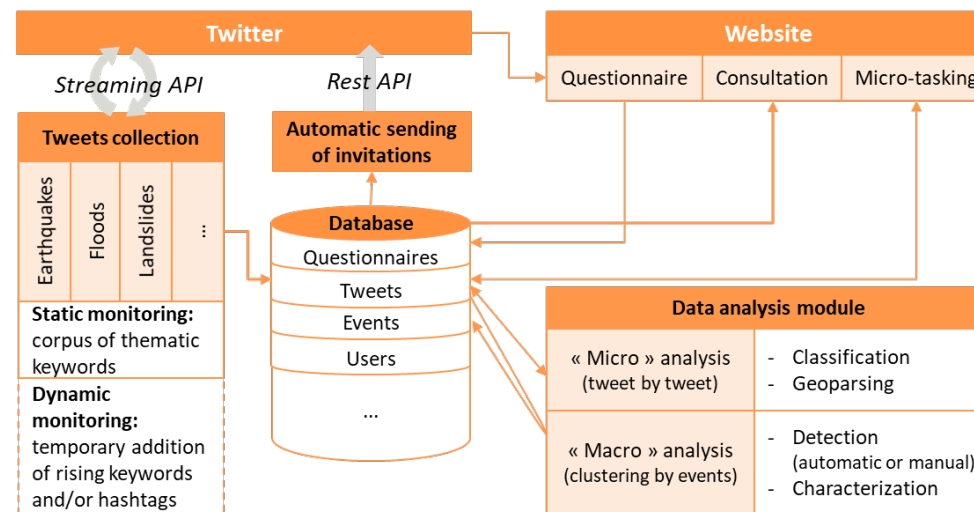




# 2. CITIZEN-BASED DATA-DRIVEN RAPID RESPONSE

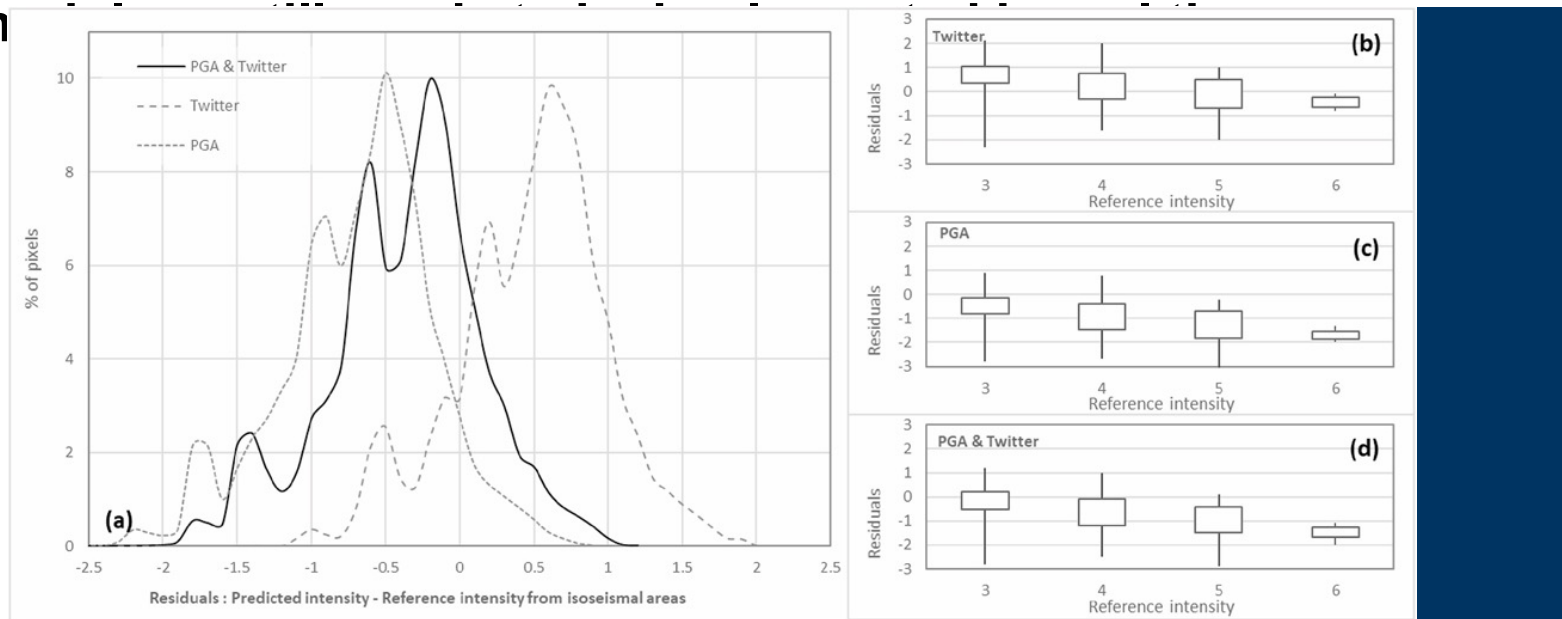
- Principle of the SURICATE-Nat platform
  - Automatic analysis of tweets immediately during/after occurrence of natural disasters
    - Enrich the flow of raw data from Twitter into a human based instrumental stream that can then be analyzed as those coming from technological sensors
    - New version of SURICATE-Nat: event driven architecture

Auclair, S. et al. (2019). SURICATE-Nat: Innovative citizen centered platform for Twitter based natural disaster monitoring. ICT-DM, IEEE. doi: 10.1109/ICT-DM47966.2019.9032950



# 3. multi-sensors DATA-DRIVEN RAPID RESPONSE

- Towards a fusion of sensor-based and citizen-based data
  - Development of a method for coupling data analysis
    - Agnostic spatio-temporal clustering of tweets
    - Combination of data from seismic stations and Twitter with Bayesian Networks
    - This meth



Fayjaloun, R., Gehl, P., Auclair, S., Boulahya, F., Guérin-Marthe, S., & Roullé, A. (2021). Integrating strong-motion recordings and Twitter data for a rapid shakemap of macroseismic intensity. *International Journal of Disaster Risk Reduction*, 52, 101927.

# Geotech IE objectives and Work Packages

(reminder)



## Community oriented goals

- Contribute to federate the geotechnical community around a common position / proposal for geotechnical data,
  - Scientific – IT connection
  - BIM – GIS and more connection
  - Users – Solution providers connection

## Work packages:

- #1: Common conceptual model
- #4: Technical paper
- #5: Implementation Guide for Software Vendors

## Technical oriented goals

- Propose effective solutions to enable digital continuity between GIS and BIM

## Work packages:

- #2: Extension of OGC Geoscience standards,
- #3: Technical documentation on the use of OGC APIs
- #3bis: Implementation forum

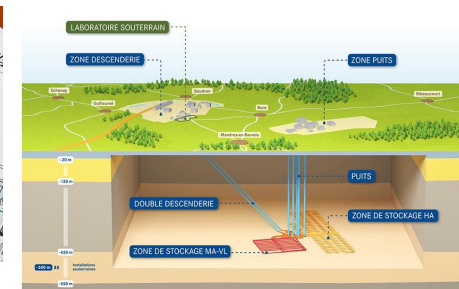
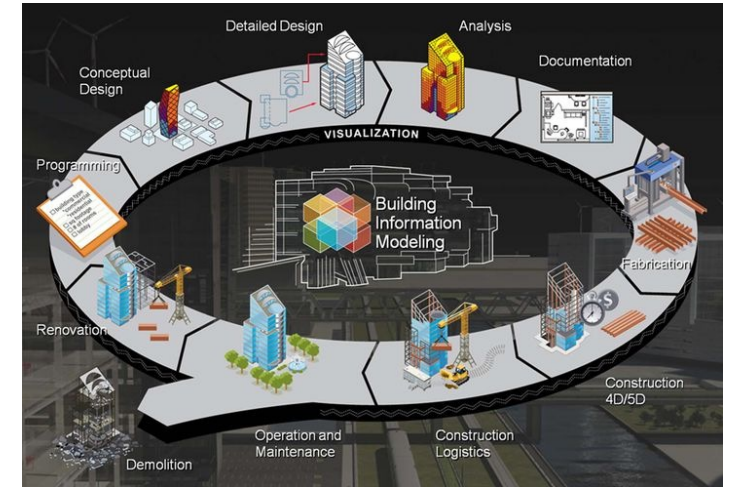
## BIM and its status in France

### Building Information Modeling (BIM)

- Targetting a « Digital Twin » of the construction
- Becoming more and more present for infrastructures (road, rail, bridge, tunnel, earthworks and geotechnics)

### France is among the leaders in BIM

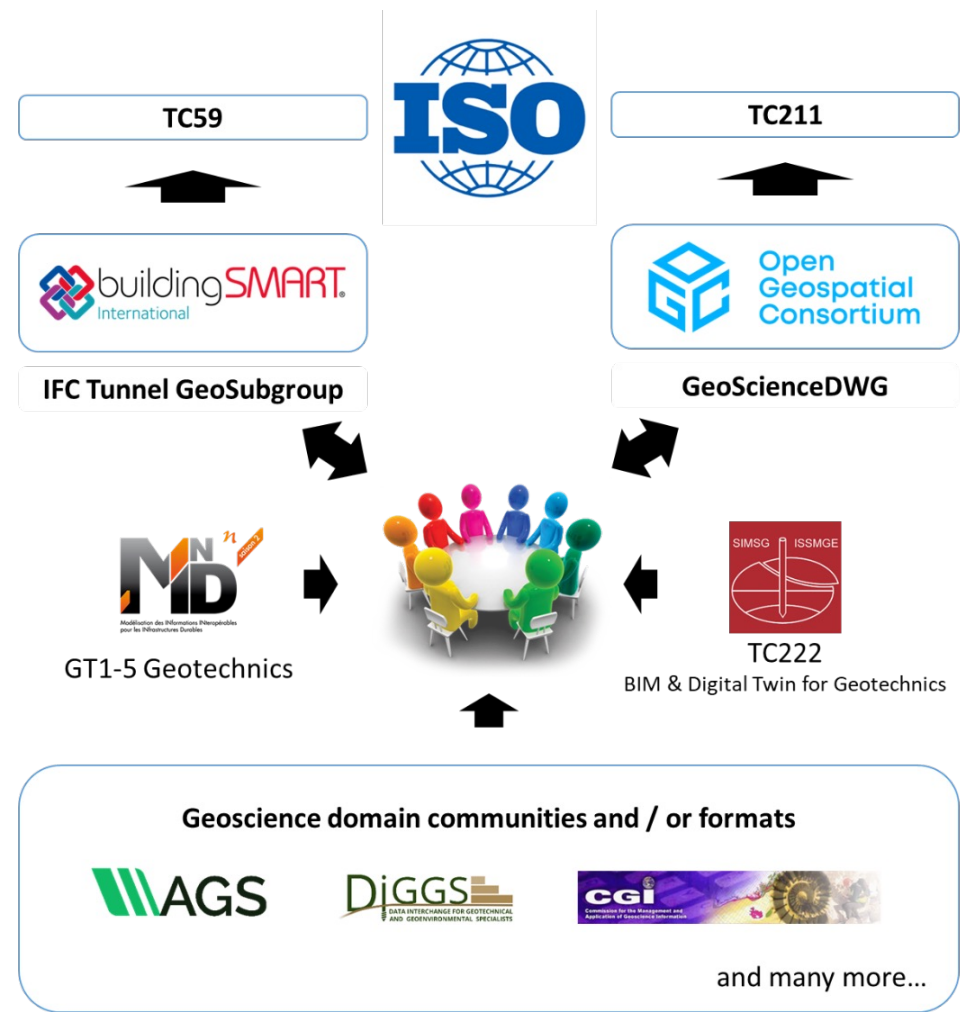
- BIM is not mandatory in France, yet lot of projects rely on it
- Supported by the French Ministry: Building Digital Transition Plan, Plan BIM 2022
- France is very active in the definition of open standards for BIM (OpenBIM) especially for infrastructure
  - The MINnD project federates constructors, engineering, research organizations and owners on that topic
- Several big infrastructure projects in France, especially in the domain of underground infrastructures





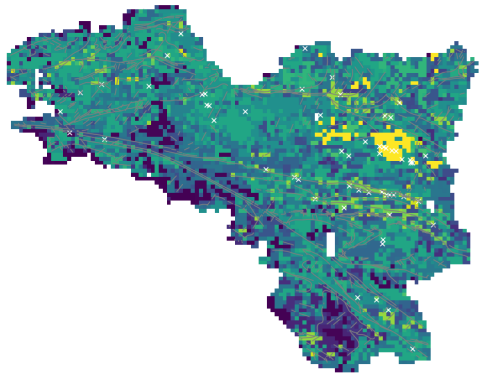
Standardization

Forum

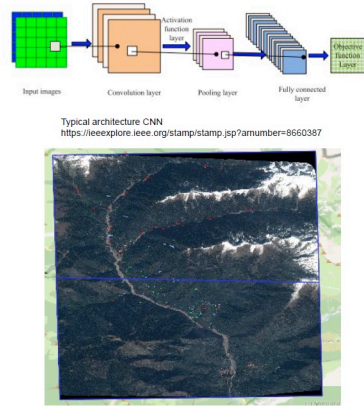


# 4 AI topics of interest for BRGM have been identified:

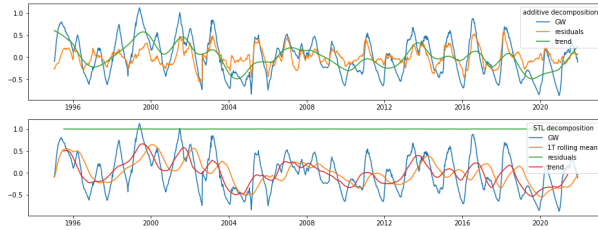
- Predictive mapping



- Image analysis



- Time series forecasting



- Natural language processing



➤ Detail use cases of these 4 topics

IA projects at BRGM

# Predictive mapping

## Développement de la Recherche vers l'Opérationnel en Prédicativité (DROP)

Vincent Labbé, Bruno Tourlière, Axel Rousseau, Guillaume Bertrand (DNG/DGR)

**Objective :** Ease and increase visibility of predictive mapping inside BRGM

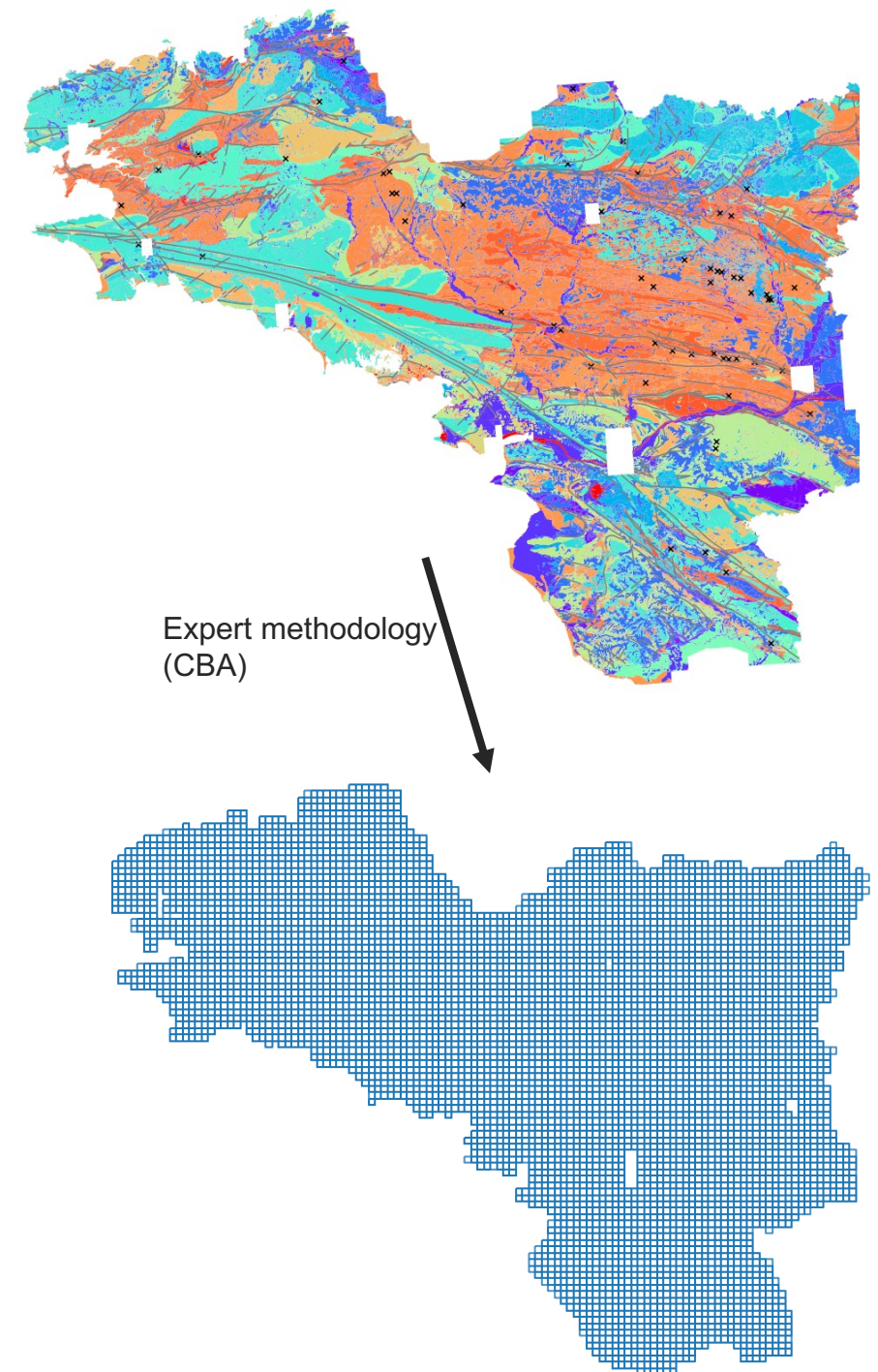
### Realisation :

- Python library: CBA (*Tourlière et al., 2015*) grid creation, data preprocessing, generation and evaluation of the model (random forest), generation of the predictive map
- Creation of a functional prototype of a QGIS plugin

### Perspectives :

- From POC to production
- Technical advances and new features
- Dissemination (open source ?) and provision (web API?)

➤ Combination of expert methodologies and AI techniques



# Time series forecasting

**Demonstrate synergy to the state-of-the-art approaches for groundwater applications**

## 1. Modeling and forecasting

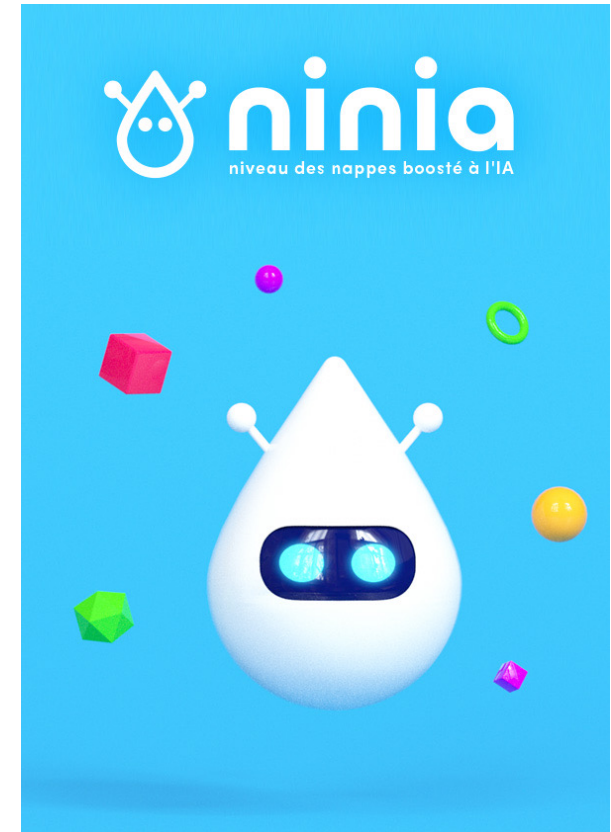
Adding **new approaches, methods and tools** (data driven) to all reading existing one (physically based) for GroundWaterLevel prediction **to face new challenges** :

- \* in real time prediction
- \* massiv target predictions (20k wells)
- \* with new predictors (climatic models/satellites,...)
- \* and so many more !

## 2. Data Analysis

Entering the era of BigData for **new insight** on existing dataset (regulatory, monitoring, scientific reports,...) to :

- \* drive mitigation actions
- \* optimize monitoring network
- \* ...



*made with <3 by V. Godard*



# Time series forecasting

## 1. Modeling and forecasting

Needs identified all over France (if not Europe/World)

### Objectives :

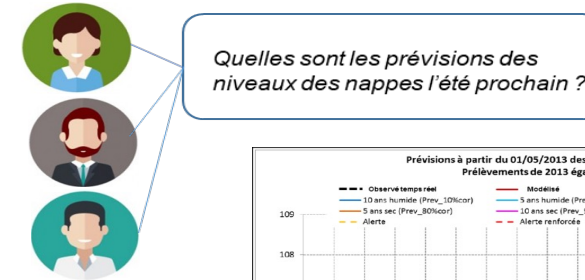
- Deliver a Proof of Concept
- Set up tools and technics for next step
- Gain in expertise

### Achievements :

- Jupyter Notebooks (Python)
- Variety of RNN tested
- End to end project : from data collection to predictions
- Code versioning and documentation on Gitlab

### Perspectives :

- A first web service in end of 2022 for fully automated predictions



Physical modeling (GARDENIA)

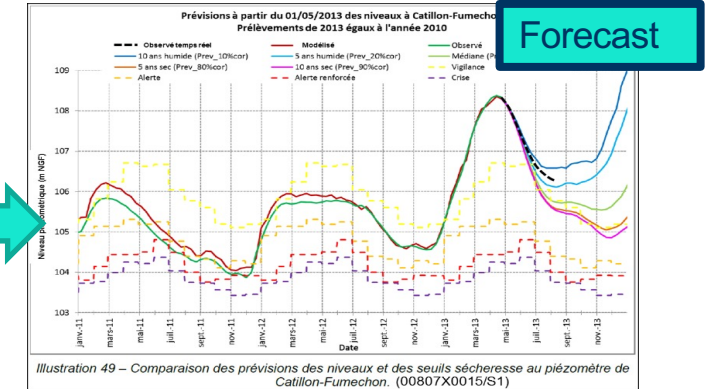
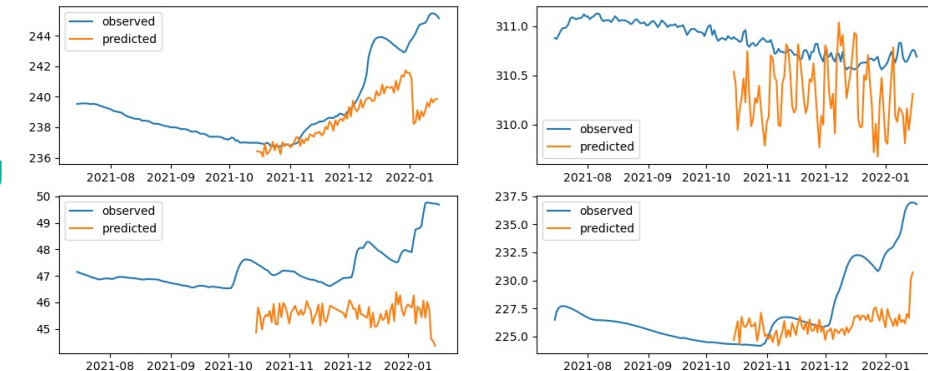


Illustration 49 – Comparaison des prévisions des niveaux et des seuils sécheresse au piézomètre de Catillon-Fumechon. (00807X0015/S1)

Prévisions de niveaux établies à partir d'une modélisation (Gardenia) Bault et al. 2013

Illustration : E. Rouxel

AI modeling (LSTM)



# Natural language processing

## Tools developed to automatically exploit the growing volume of textual documents

- Textual data are complex : ambiguity, implicit and heterogenous
- Explosion of the amount of data available with the internet

- **Objectives :**

1. Improve the exploitation of traditional BRGM data
2. Leverage new data sources



eNzo caMping  
@EnzoCamping

Toi qui lit ce tweet 🤔 ... oui toi 🤔 ... j'ai un truc à te demander 🤔 ..... pk tu révises pas tes partiels ? ils vont s'apprendre tout seul?



WIKIPÉDIA  
L'encyclopédie libre

Rechercher dans Wikipedia

### Traitement automatique des langues

Pour les articles homonymes, voir Tal.

Le traitement automatique du langage naturel (abr. TALN), ou traitement automatique de la langue naturelle<sup>(1)</sup>, ou encore traitement automatique des langues (abr. TAL) est un domaine multidisciplinaire impliquant la linguistique, l'informatique et l'intelligence artificielle, qui vise à créer des outils de traitement de la langue naturelle pour diverses applications. Il ne doit pas être confondu avec la linguistique informatique, qui vise à comprendre les langues au moyen d'outils informatiques.

Le TALN est sorti des laboratoires de recherche pour être progressivement mis en œuvre dans des applications informatiques nécessitant l'intégration du langage humain à la machine<sup>(2)</sup>. Aussi le TALN est-il parfois appelé **ingénierie linguistique**<sup>(3)</sup>. En France, le traitement automatique de la langue naturelle a sa revue, *Traitement automatique des langues*, publiée par l'Association pour le traitement automatique des langues (ATALA).

#### Sommaire (masquer)

- 1 Histoire
  - 1.1 Années 1950-1980
  - 1.2 Années 1970-1980
  - 1.3 Années 1990-2000
  - 1.4 Depuis 2000
- 2 TALN statistique
- 3 Champs de recherche et applications
  - 3.1 Syntaxe
  - 3.2 Sémantique

Google  
label:unread label:inbox

Gmail

COMPOSE

Inbox (6)

Starred

Sent Mail

Drafts

Notes

More +

Banana

Pinterest	Inbox	witch makeup, engagement rings and other search trends - Come see w
YouTube	Inbox	The Happy Family Show: "S6 E10 "Haunted Dreamhouse!" Outtakes!" -
Google+	Inbox	Banana, a few Google+ posts you may like - Posts from Google+ picked fo
Pinterest	Inbox	Your inspiration for the week - We found some fresh Pins for you From the
REI	Inbox	These Boots Were Made for Winter - Shop Merrell Arctic Grip Footwear   V
Pinterest	Inbox	Love art? Here are popular Pins in art this week - Trending in art this week
REI	Inbox	We're Closing Our Doors Again on Black Friday - Will You Go Out With U
Pinterest	Inbox	Your inspiration for the week - We found some fresh Pins for you Save Fro
Pinterest	Inbox	Love animals and pets? Here are popular Pins in animals and pets this v
Pinterest	Inbox	You've got 25+ new Pins waiting for you - This just in: over 25 new Pins! S
REI Garage	Inbox	Up to 75% Off Great Deals. - Shop the Deals   View With Images REI GAR

Common Mining  
Text segmentation (en)  
Désambiguïsation lexicale  
Résumé automatique de texte  
Conférence

# Natural language processing

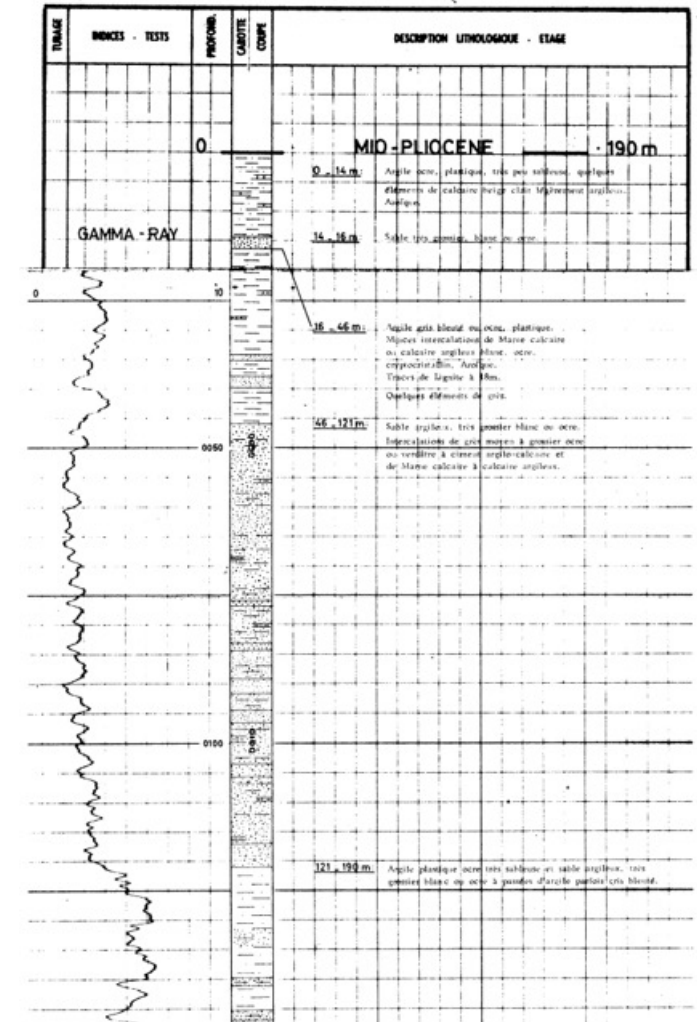
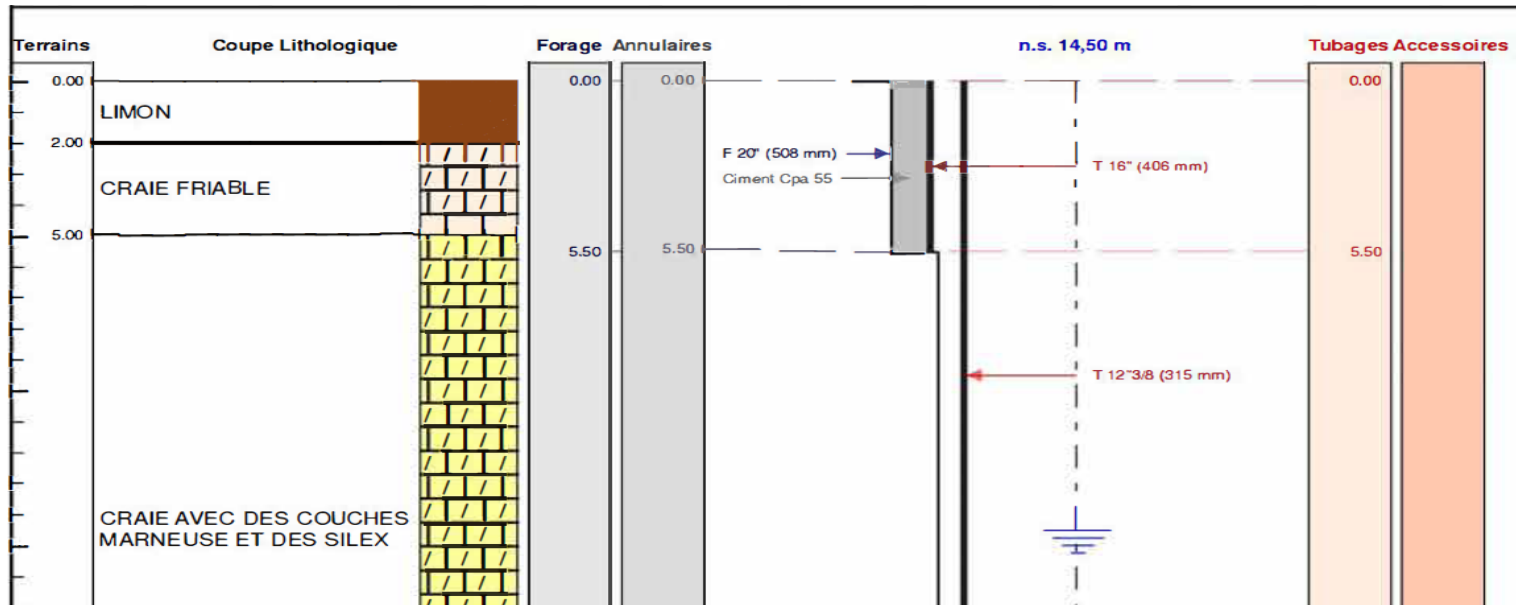
## 1. Improve the quality and accessibility of geological data

Banque du Sous-Sol database : >1 000 000 of boreholes descriptions

→ Most of them are searchable and editable texts

→ But there is also >500 000 scanned documents

→ OCRization has to come first !!



**Merci de votre attention**  
**Thanks for your attention**

Contact Details:

Matt Harrison, Director of Digital Platforms

m.harrison@brgm.fr



*Weathering erosion in an open pit potash mine, DRC – N. Charles*