

MISE EN PLACE D'UN MOTEUR DE RECHERCHE POUR UN PORTAIL GÉOGRAPHIQUE DONT LA PERTINENCE TIENT COMPTE DE L'ÉLOIGNEMENT GÉOGRAPHIQUE

**SERVICE DE L'INFORMATIQUE DE LA POLYNÉSIE FRANÇAISE, CELLULE SIG
PROJET ENCADRÉ PAR EMMANUEL BOUINIOT (SIPF-CELLULE SIG)**

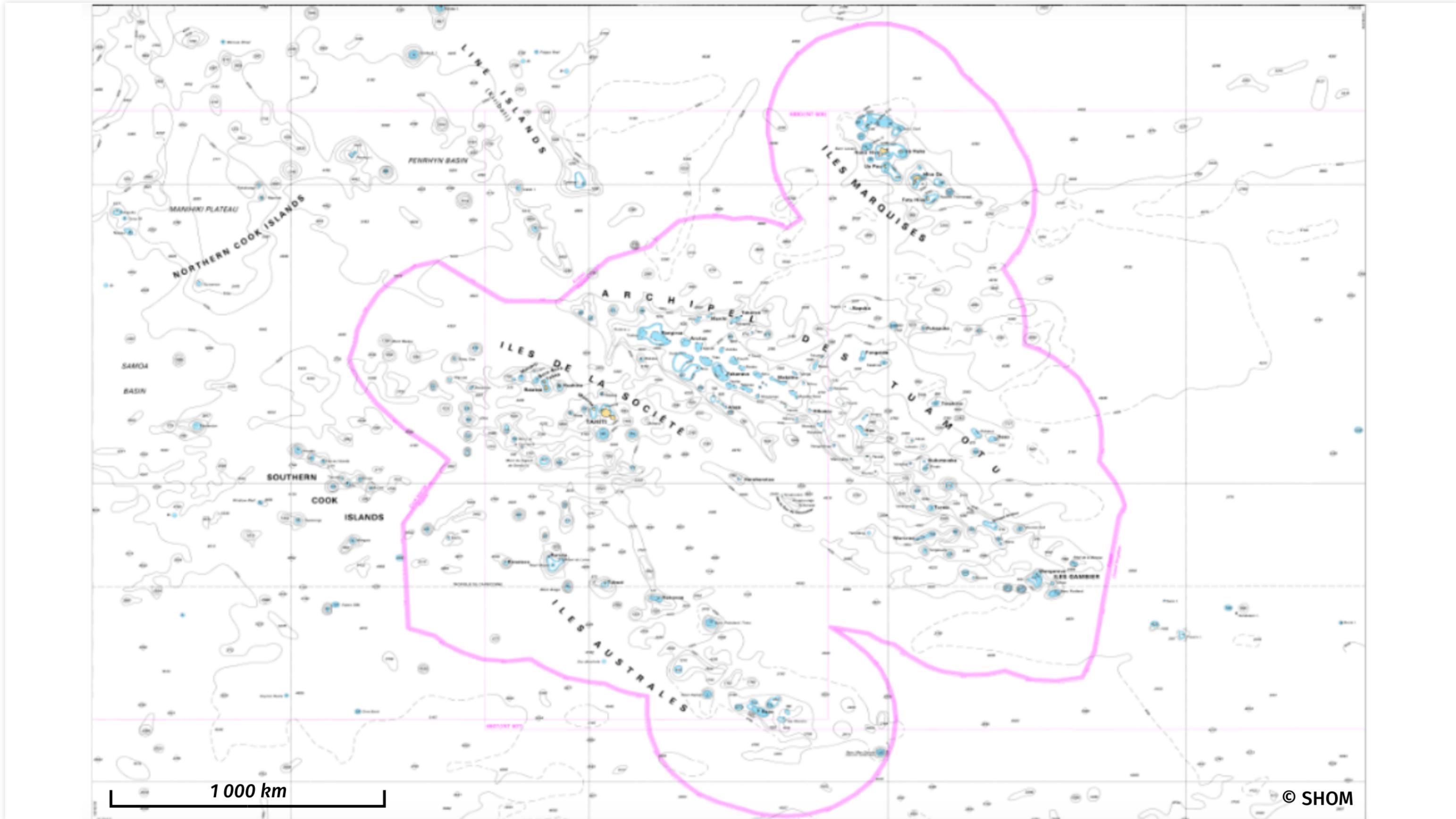
**EMMANUEL BOURASSIN, CYCLE DES INGÉNIEURS DE L'ENSG
STAGE PLURIDISCIPLINAIRE, MAI-AOÛT 2018**

PROFESSEUR RÉFÉRENT : BENOIT COSTES (IGN/ENSG/DE/DEI)

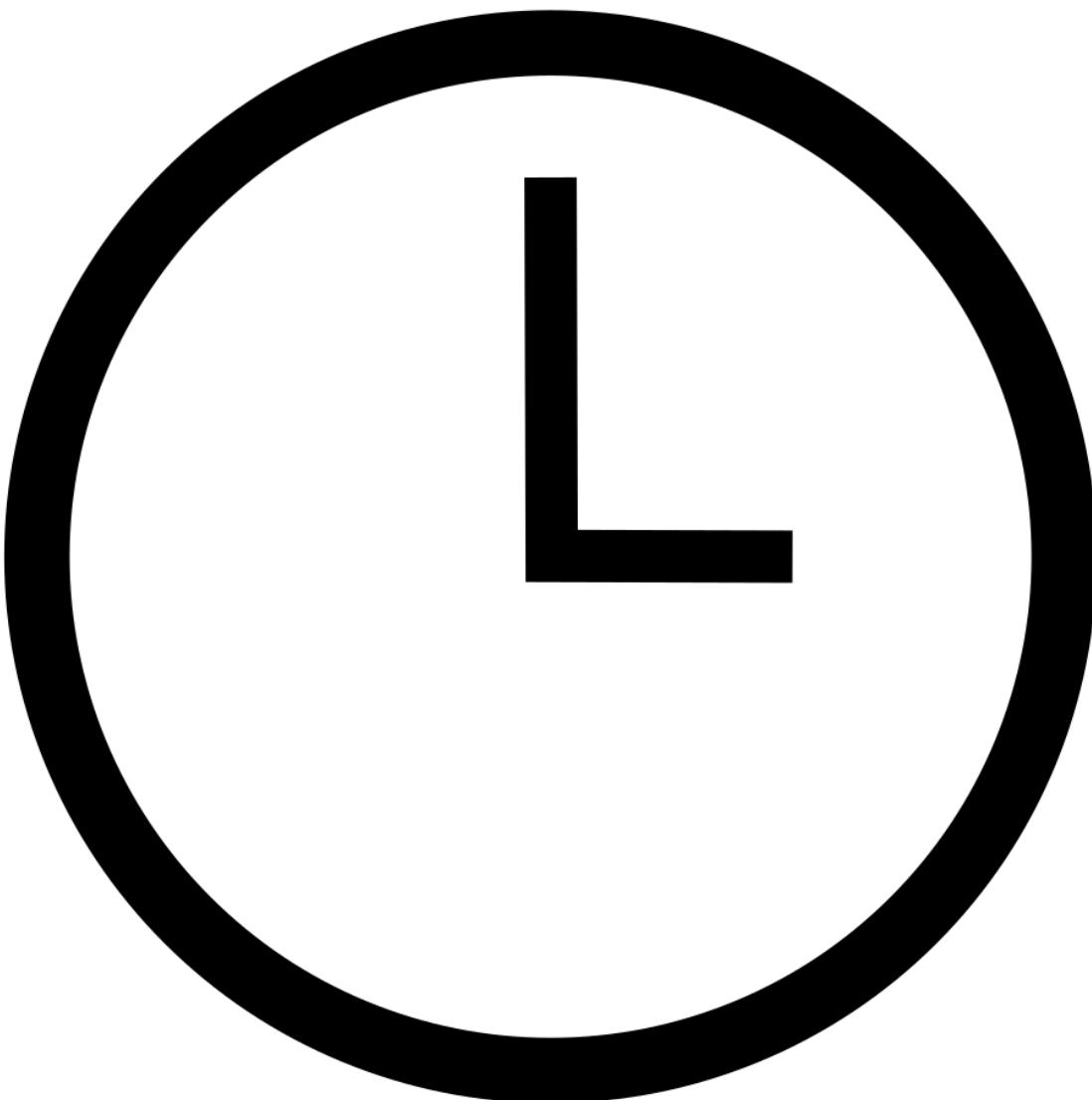


CONTEXTE DU STAGE

Contexte



Rapidité



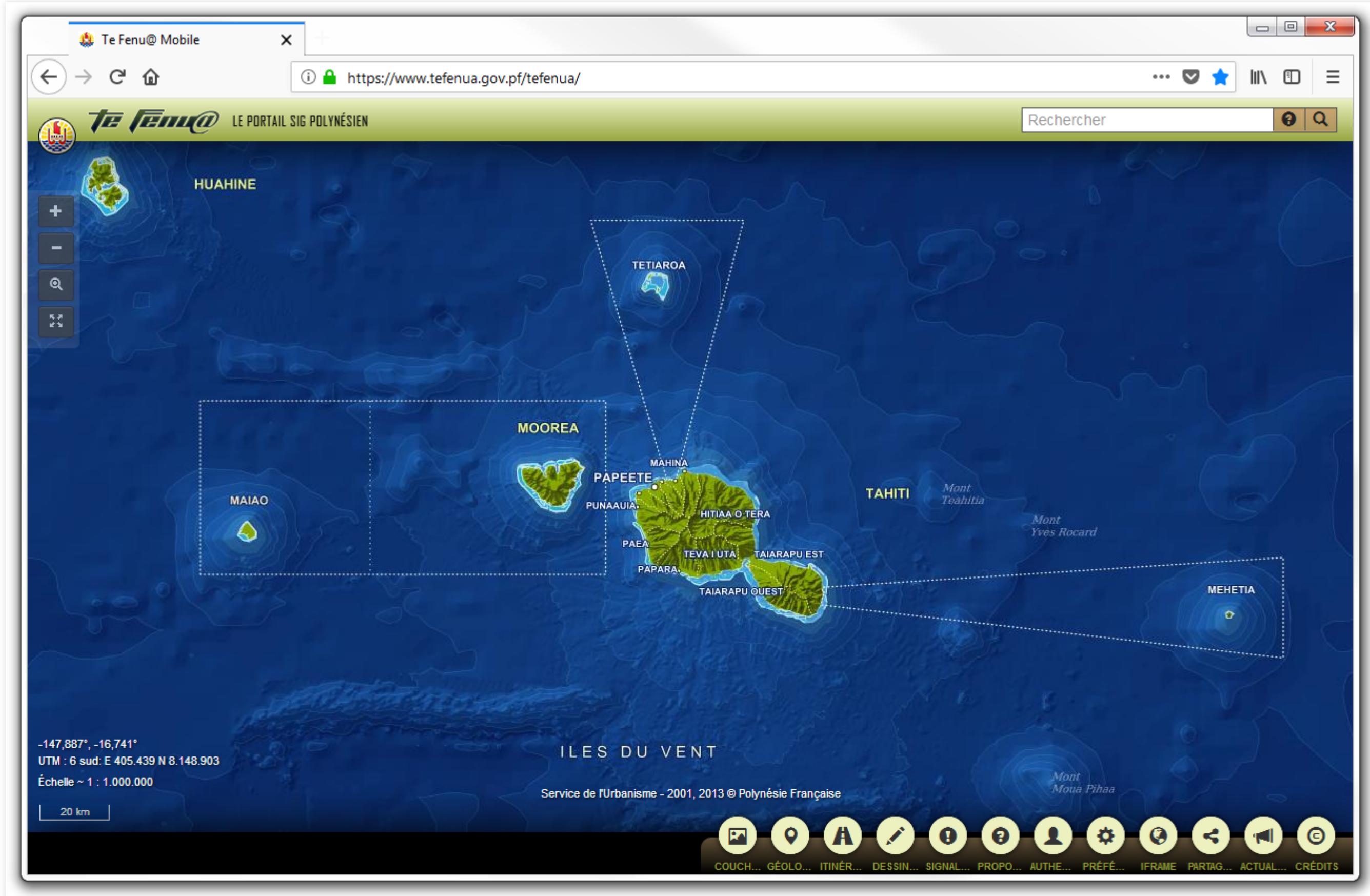
Pertinence



- 1. Te Fenua**
- 2. Elasticsearch**
- 3. Aspect sémantique**
- 4. Aspect géospatial**

PRÉSENTATION

Te Fenua



UTILISATEURS

Te Fenua

15

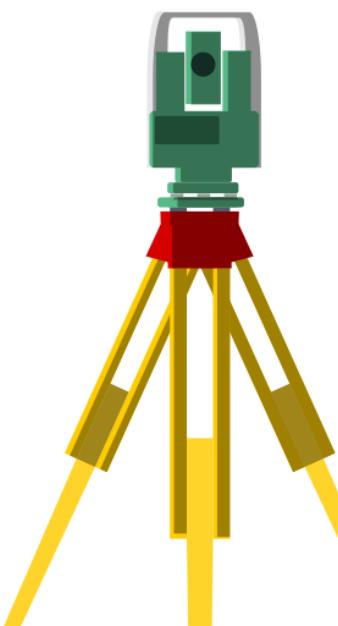
SAMU

17

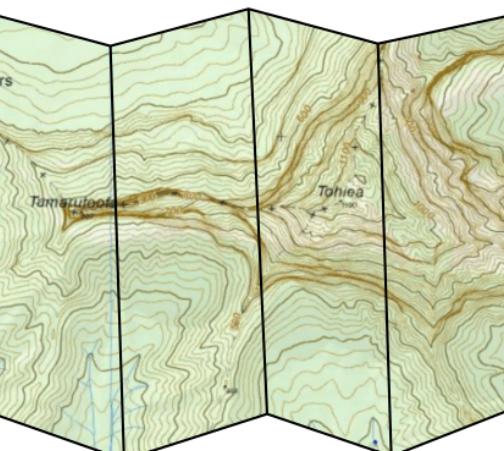
Gendarmerie

18

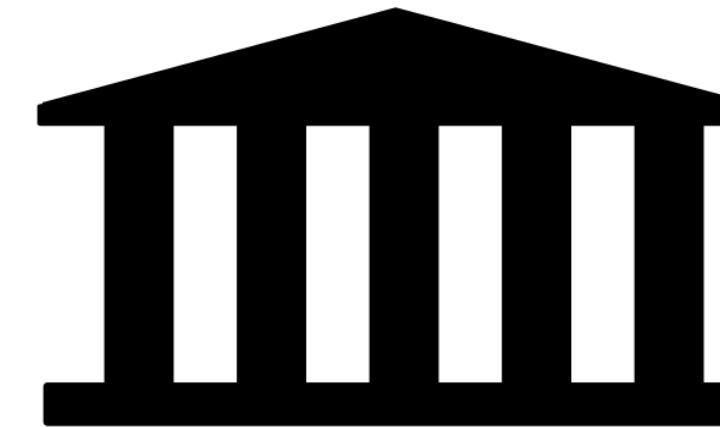
Sapeurs-pompiers



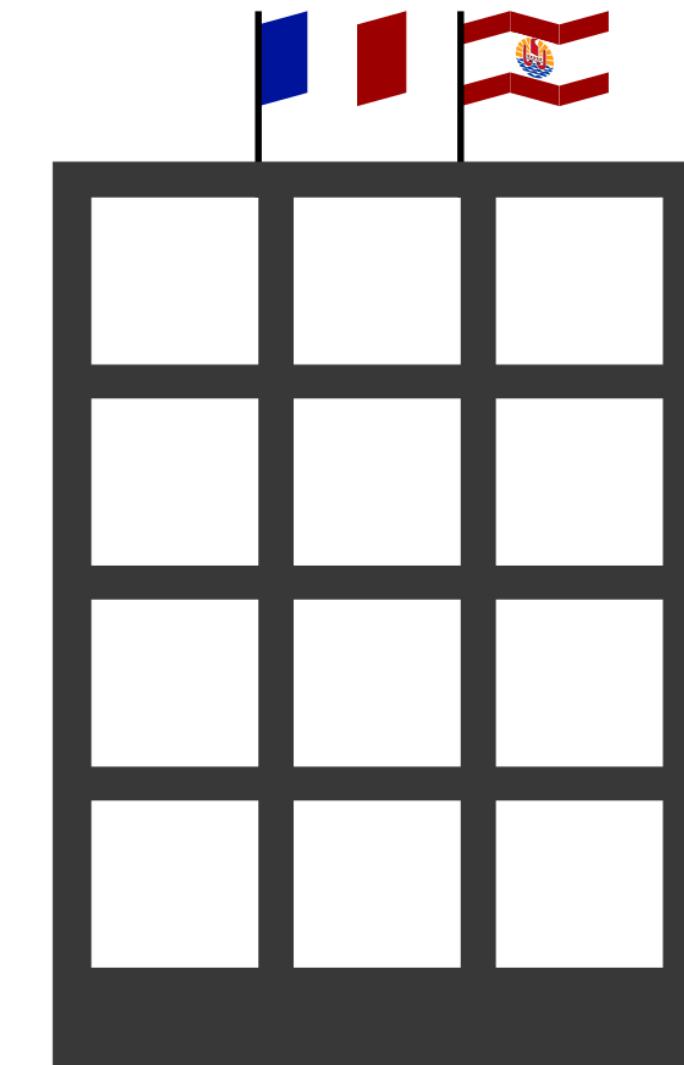
Géomètres-
topographes,
architectes



randonneurs



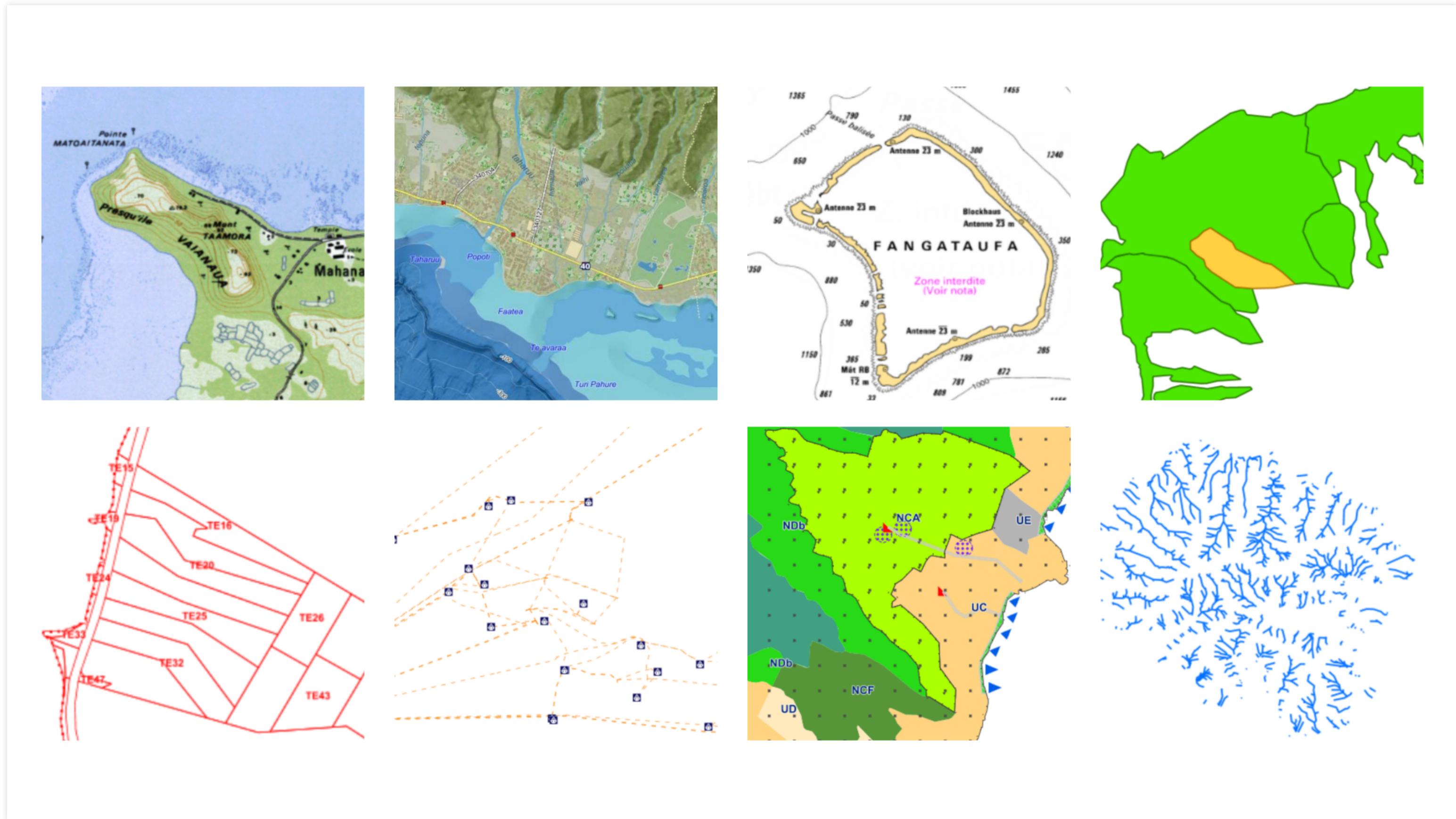
mairies



administrations et
services de la
Polynésie française

DONNÉES

Te Fenua



- 1. Te Fenua**
- 2. Elasticsearch**
- 3. Aspect sémantique**
- 4. Aspect géospatial**

moteur
de recherche



Interface REST

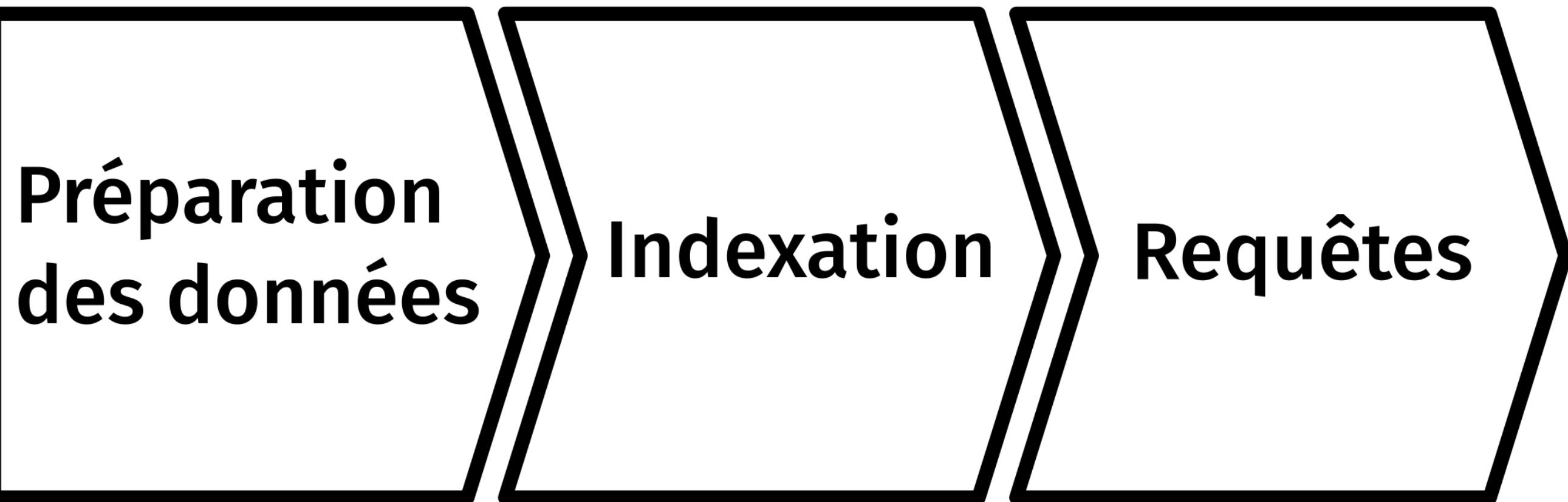
JSON/GeoJSON

opensource

elasticsearch

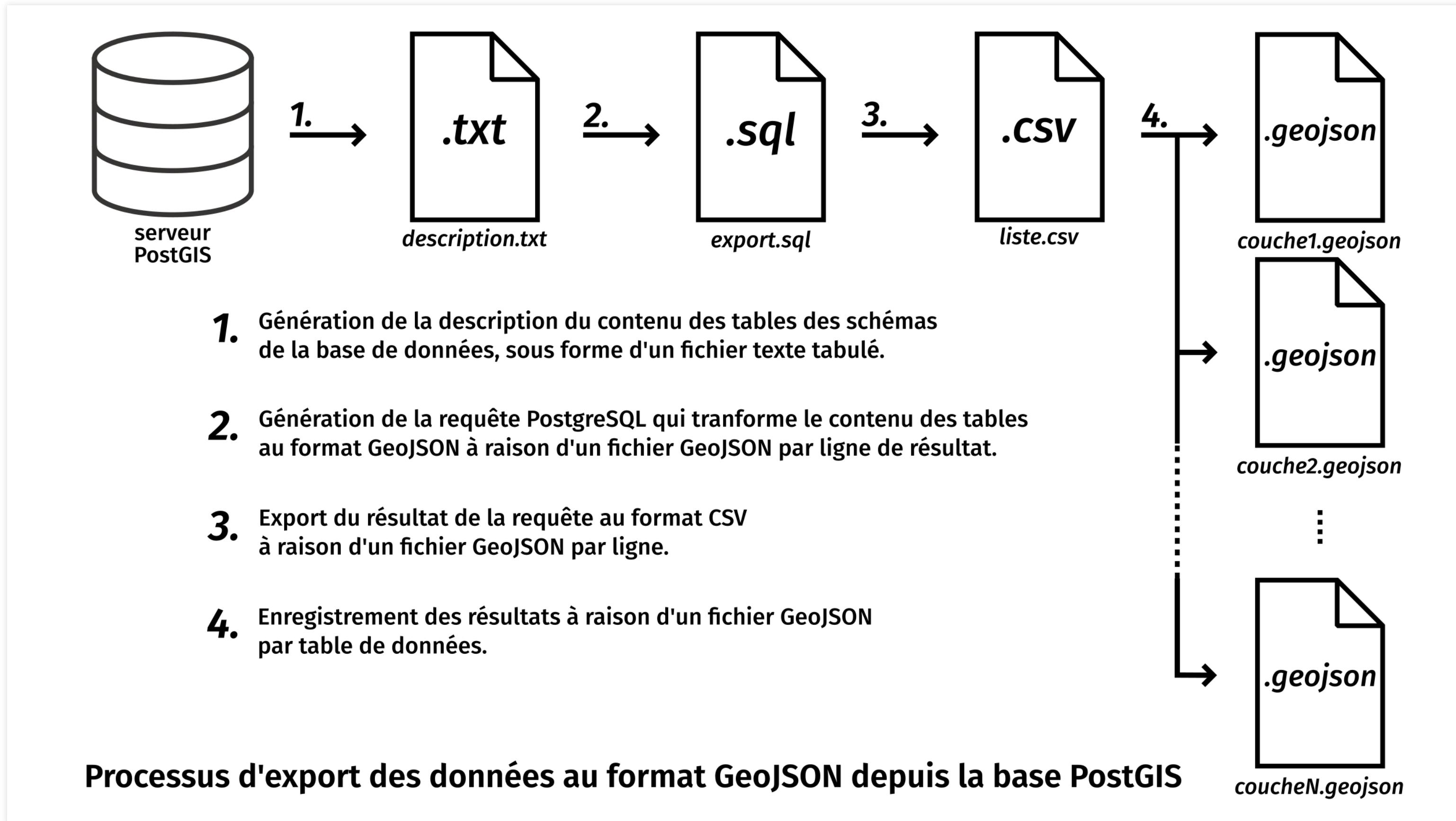
multiplateforme

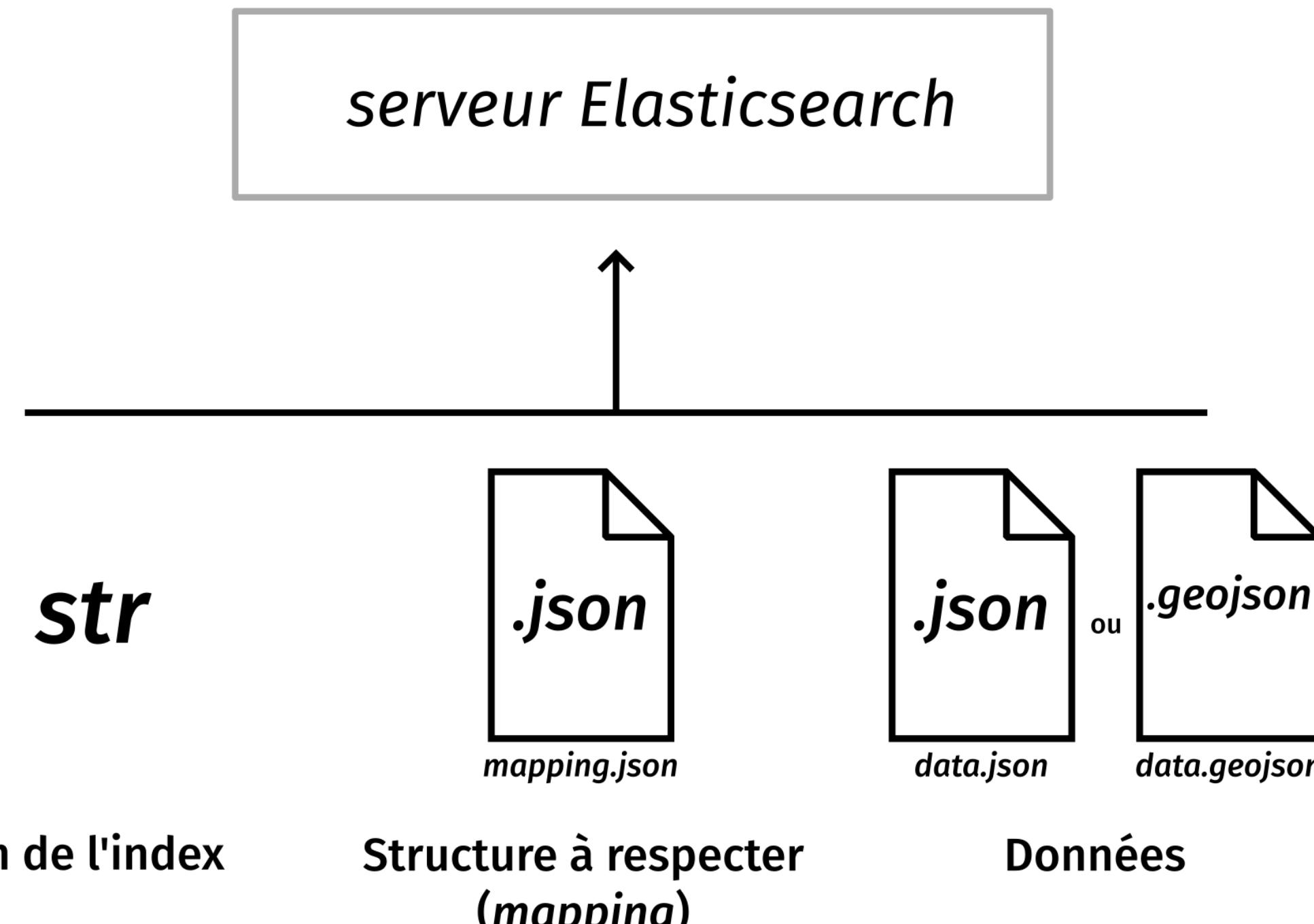
API Python

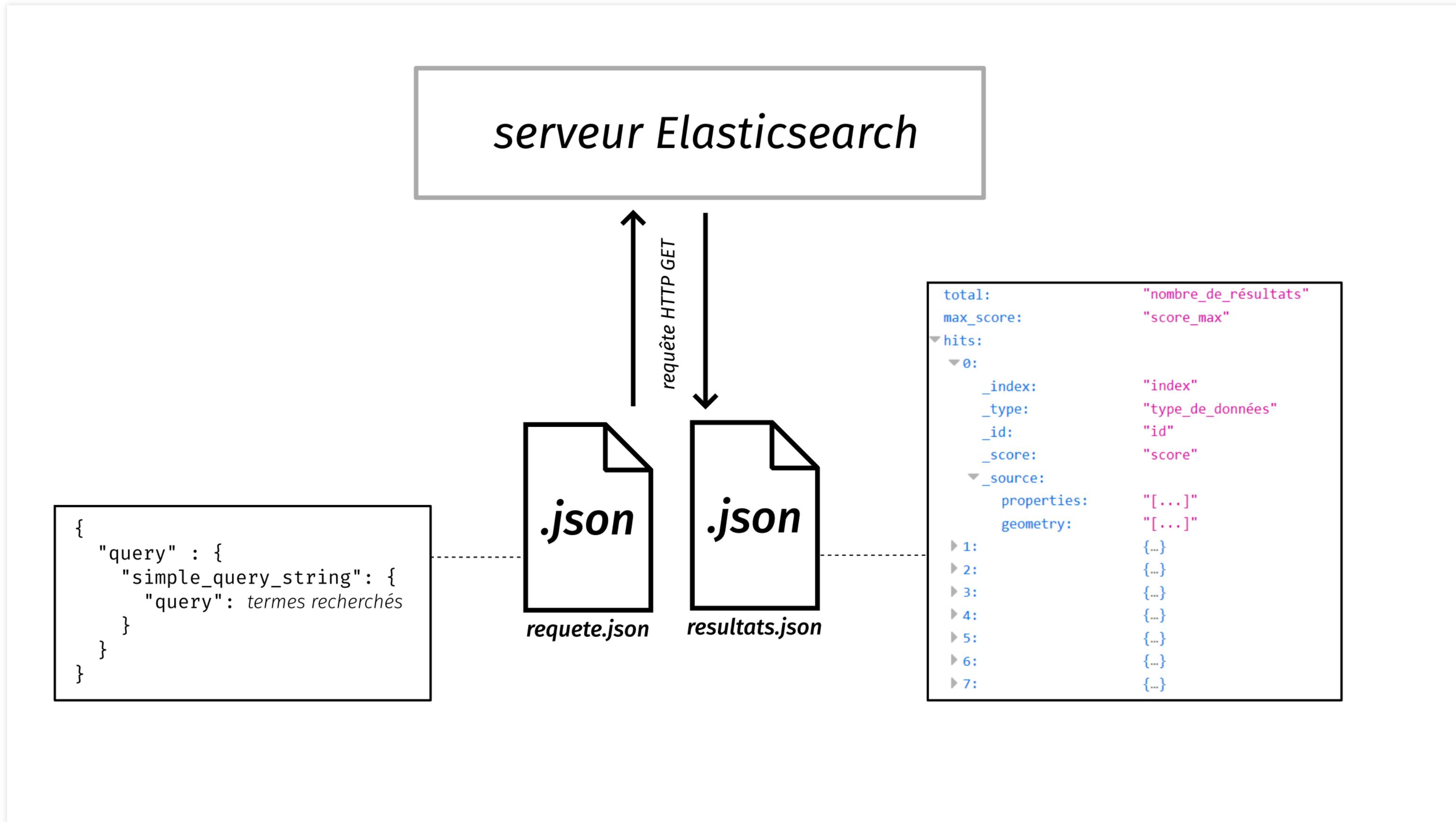


EXPORT PRÉALABLE

Choix techniques

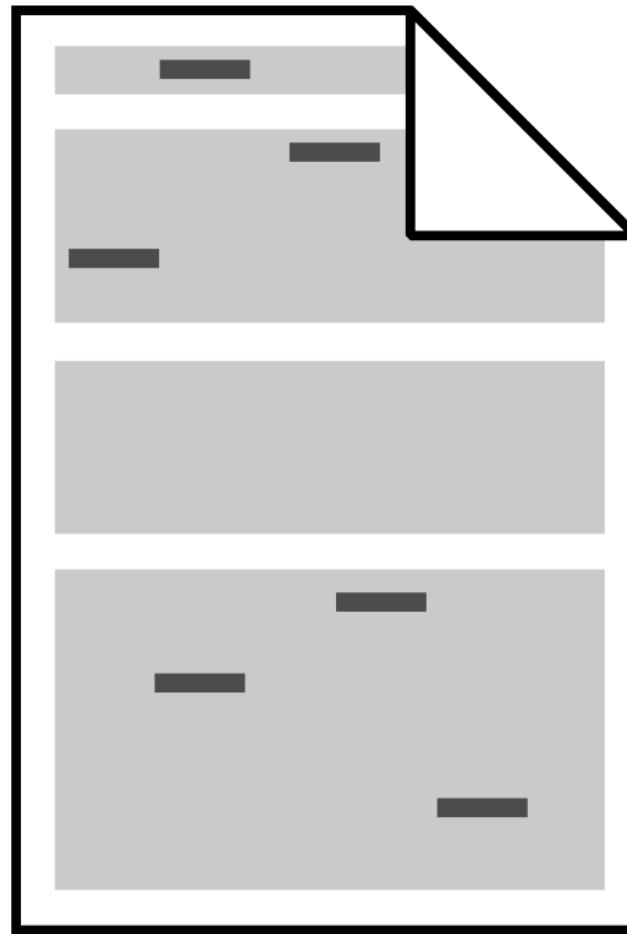






PERTINENCE, CALCUL DE SCORE

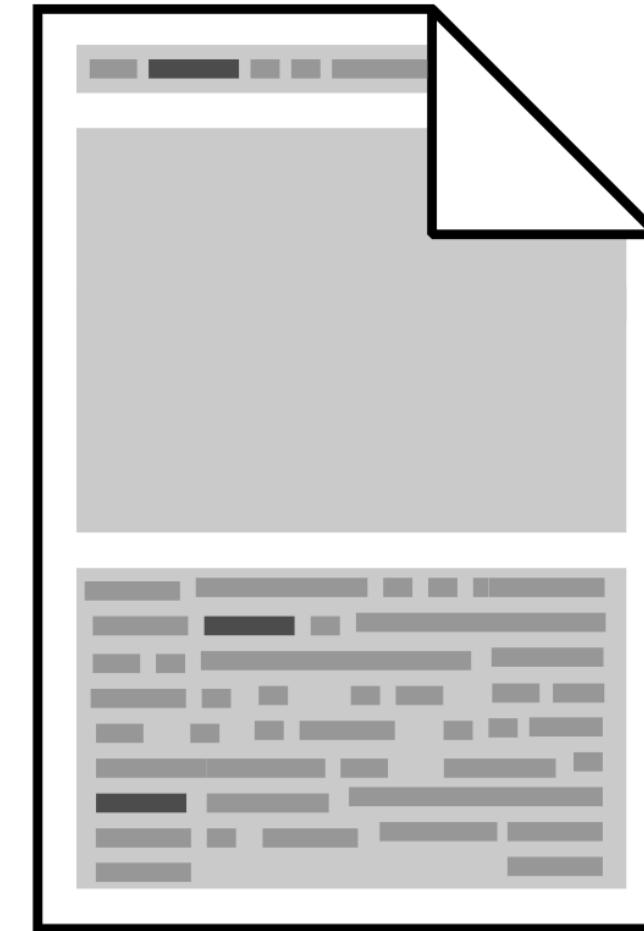
Choix techniques



term frequency



*inverse
term frequency*



field-length norm

1. Te Fenua
2. Elasticsearch
- 3. Aspect sémantique**
4. Aspect géospatial

Une couche de données

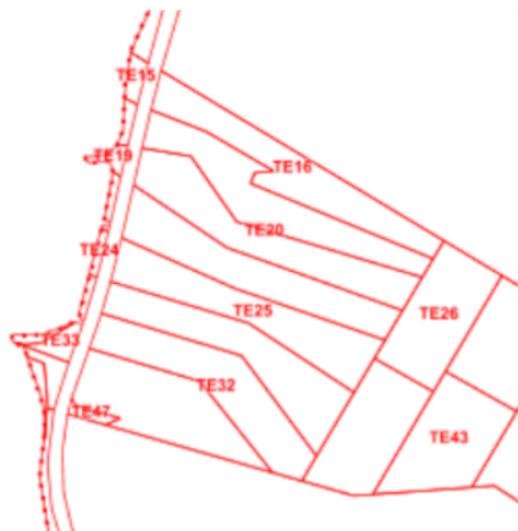


Un index Elasticsearch

Dans un index | une même thématique
Elasticsearch | un même type de données
| une même structure

EXPRESSIONS RÉGULIÈRES

Aspect sémantique

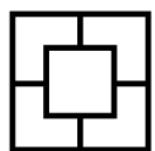


Parcelles cadastrales
[a-z]{1,2}[0-9]{1,3}

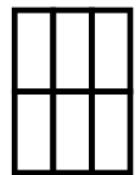


ESM

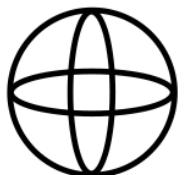
esm [0-9]{1,3}-{0,1}[0-9]{0,3}



DFCI [-0-9]+[:]+[0-9]+[:]+[0-9]+[,]+-[-0-9]+[:]+[0-9]+[:]+[0-9]



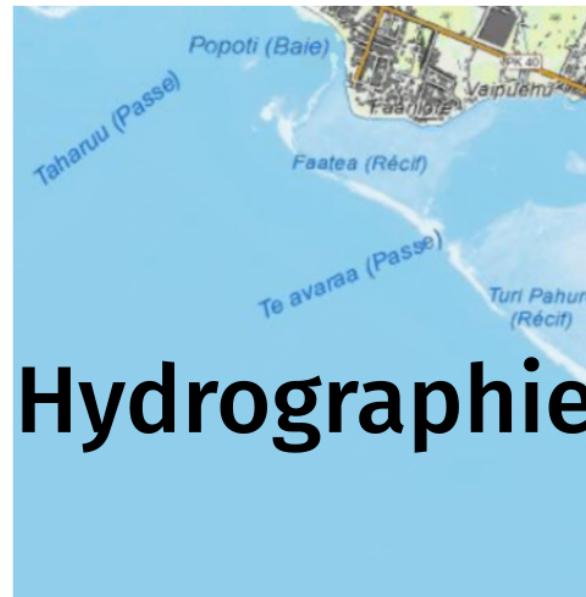
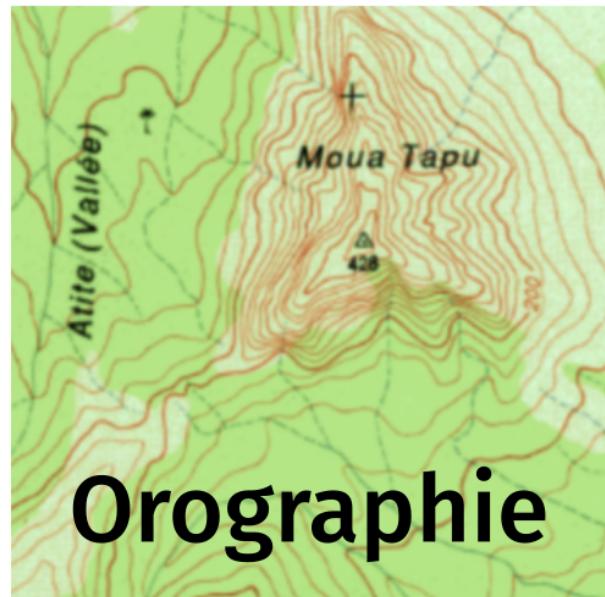
UTM [0-9][][0-9]{6}[][0-9]{7}



DMS [-0-9]+[:]+[0-9]+[:]+[0-9]+[,]+-[-0-9]+[:]+[0-9]+[:]+[0-9]

MOTS-CLÉS THÉMATIQUES

Aspect sémantique



• • •

mont,
montagne,
mt, moua, pic,
sommet, crete,
col, falaise,
plateau, motu,
ilot, vallee, ravin,
pointe, isthme,
rocher, vallee...

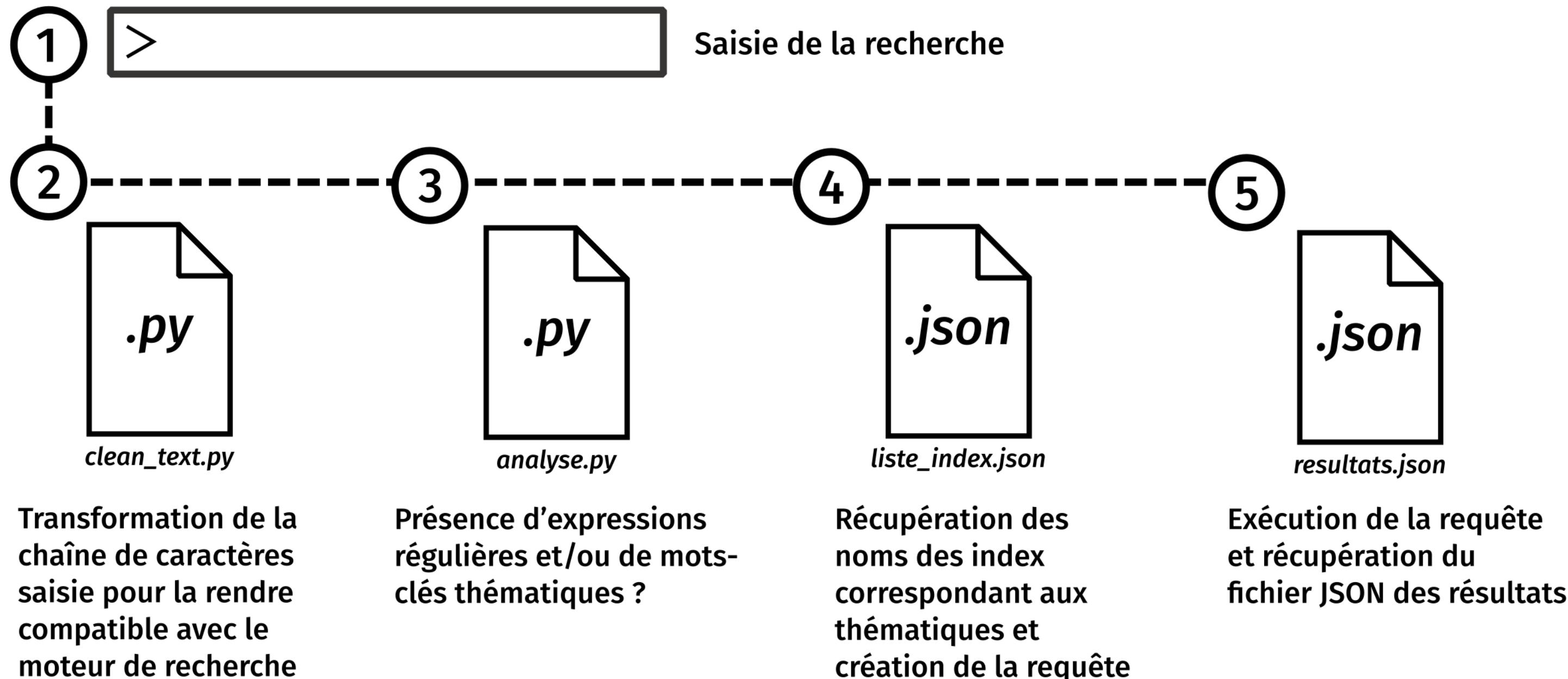
amer, baie
banc, lago
passe, recif,
riviere, fleuve,
ruisseau, torrent,
source, cascade,
lac, barrage,
atoll, ile,
archipel, baie...

rue, route,
avenue, av
boulevard,
servitude,
impasse, place,
cours, rond-
point, tunnel,
carrefour,
sentier, chemin...

hopital,
urgence,
clinique,
pharmacie,
incendie,
pompier,
caserne, police,
commissariat

PROCESSUS

Aspect sémantique



Transformation de la chaîne de caractères saisie pour la rendre compatible avec le moteur de recherche

Présence d'expressions régulières et/ou de mots-clés thématiques ?

Récupération des noms des index correspondant aux thématiques et création de la requête

Exécution de la requête et récupération du fichier JSON des résultats

1. Te Fenua
2. Elasticsearch
3. Aspect sémantique
- 4. Aspect géospatial**

COMPORTEMENTS

Aspect géospatial



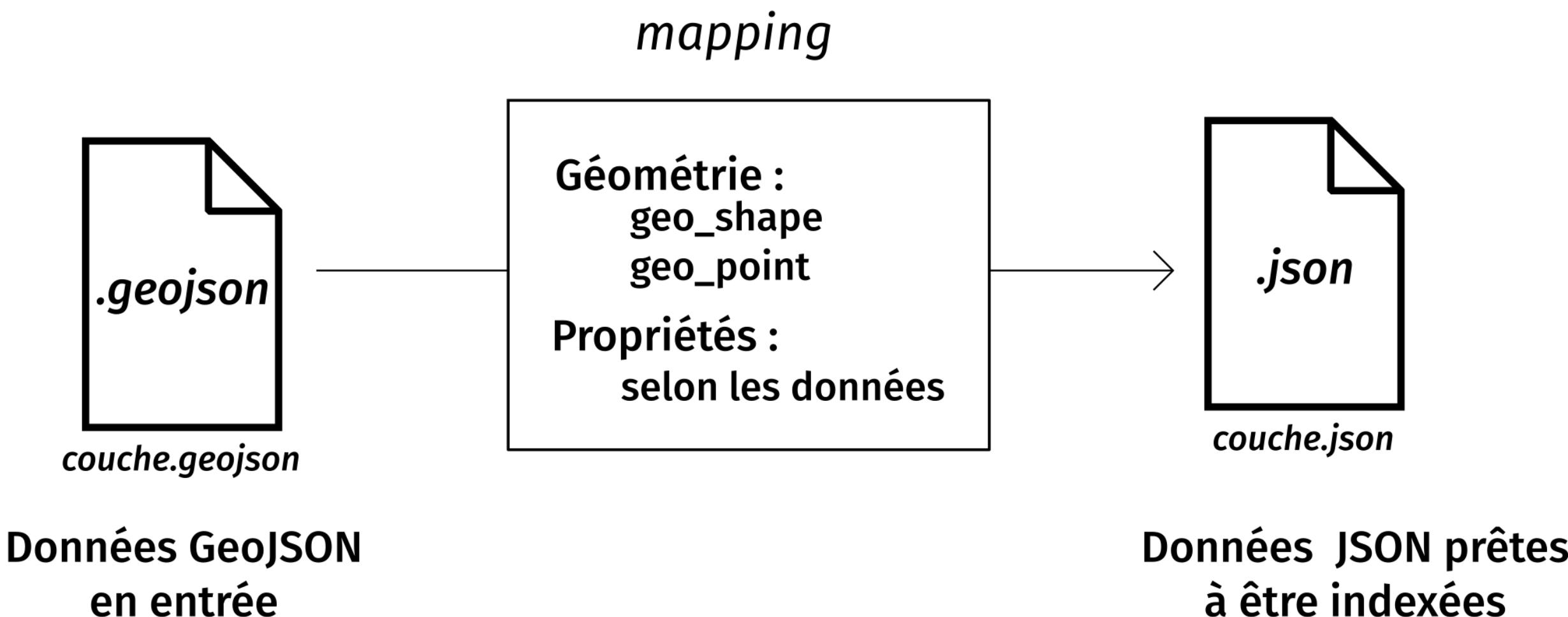
Adapter l'emprise puis rechercher

Rechercher

D'après l'enquête *Tests utilisateurs Te Fenua, juin-juillet 2018*, Patricia Gallot-Lavallée

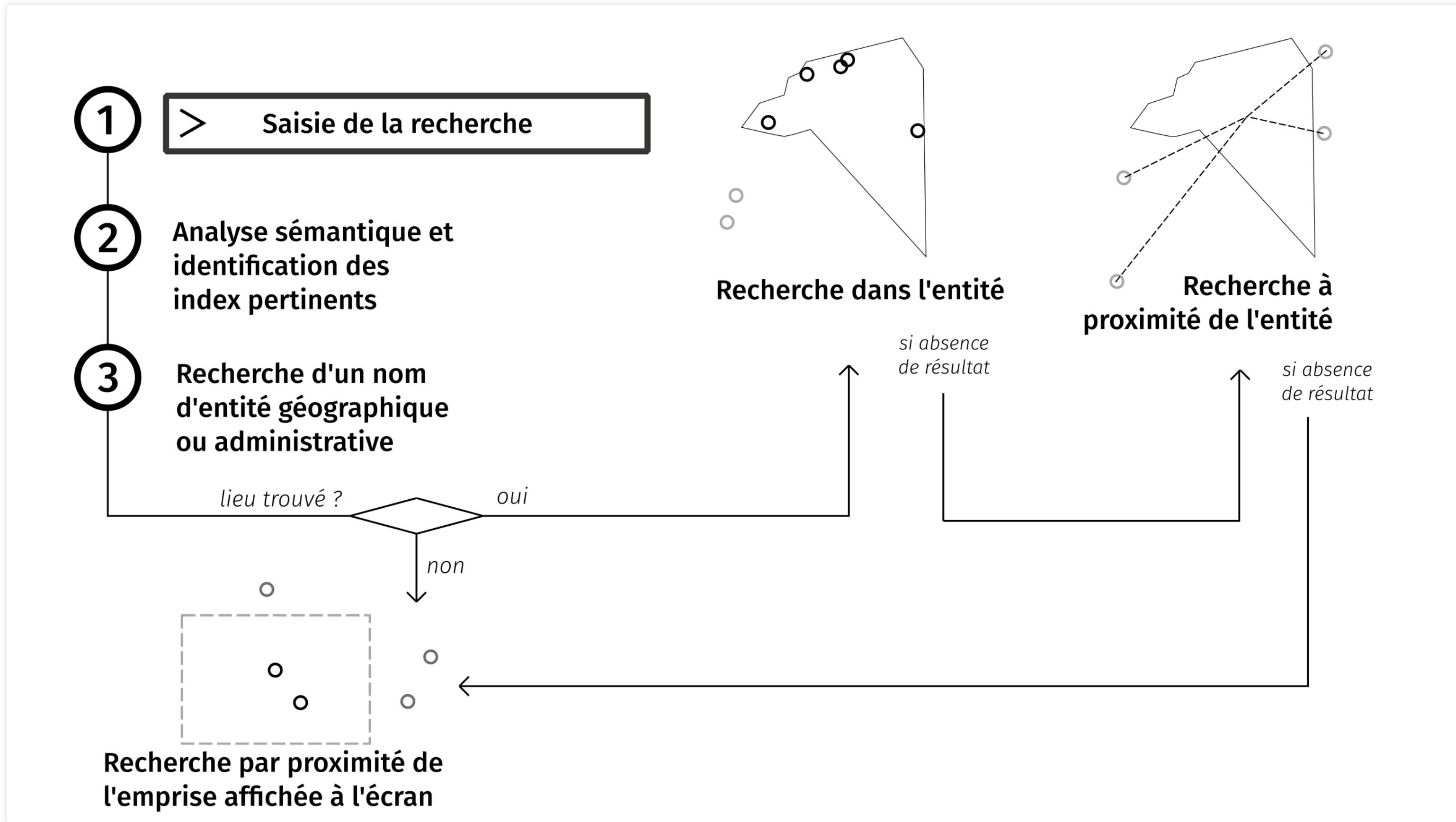
Utilisation de deux fonctions de recherche d'Elasticsearch :

- fonction de **score liée à la distance** ;
- fonction de recherche de **documents inclus dans une emprise**.



PROCESSUS

Aspect géospatial



Prolongement

SUITE, RÉUTILISATION ET PERFECTIONNEMENT

Tous mes remerciements à:

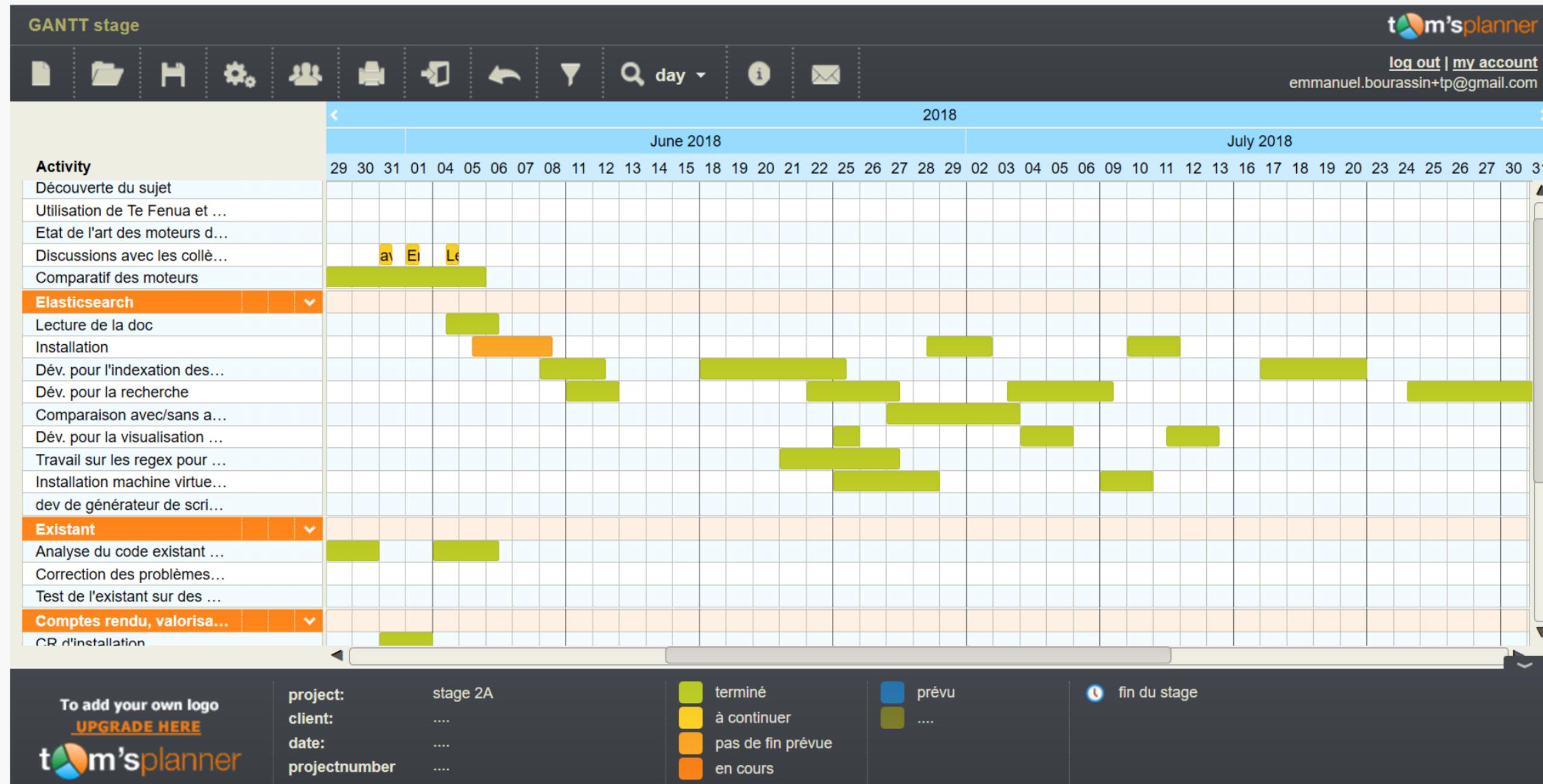
**Emmanuel BOUINOT (SIPF/Cellule SIG),
Tehei TAIORE (SIPF/Cellule SIG),
Benoît COSTES (IGN/ENSG/DE/DEI)**

pour leurs judicieux conseils et bienveillant accompagnement.

Merci également à Hubert Leviel, Mathieu Thouement, Bryan Lot, Léonard Tavae, Fabrice Fauchon, Philippe Siquin, et l'ensemble du personnel du SIPF.

DIAGRAMME DE GANTT

Bonus



Création et mise à jour d'un diagramme de GANTT via le service *Tom'splanner*

The screenshot shows a GitLab project page for 'tefenua-elasticsearch'. The left sidebar contains navigation links: Project (selected), Details, Activity, Cycle Analytics, Repository, Issues (0), Merge Requests (0), CI / CD, Operations, Wiki, Snippets, and a 'Collapse sidebar' button. The main content area has a title 'Expérimentation sur Elasticsearch'. Below it, a text block states: 'Cette expérimentation a été conduite du 2018-05-22 au 2018-08-10 par Emmanuel Bourasson (ENSG, emmanuel.bourasson@gmail.com). La version d'Elasticsearch utilisée est elasticsearch-6.3.1.' A section titled 'Import des données' explains the process of extracting data from PostgreSQL and converting it to GeoJSON using Python scripts. Another section, 'Logstash', discusses the potential use of Logstash for automating data transformation. A third section, 'Indexation des données', mentions a script for indexing data. The footer of the page includes a search bar and various icons.

Versement du code développé et commenté sur la plateforme GitLab.gov.pf

Bonus



Bonus



Bonus



Bonus



Bonus

