

# eOnsight

## Stage 2022 – Data engineering

### Contexte

Dans le cadre du programme Copernicus, l'Agence Spatiale Européenne (ESA) a conçu et a partiellement déployé la série de satellites d'observation Sentinel. Ayant chacun leurs spécificités, ils sont équipés d'un ensemble d'instruments et de capteurs.

L'objectif d'eOnsight est de croiser les données pouvant être captées par ces capteurs avec les données disponibles sur Internet.

### Objectif

Dans le cadre de cet exercice, nous nous intéresserons à agréger des données sur Wikipedia.

### Résultat attendu

En vous appuyant sur la page wikipedia [https://fr.wikipedia.org/wiki/Liste\\_de\\_ponts\\_d%27Italie](https://fr.wikipedia.org/wiki/Liste_de_ponts_d%27Italie) écrire un code source permettant **d'agréger les ponts (noms, coordonnées géographiques...) qui se trouvent à Gênes en Italie et les mettre dans une base de données**. Le tout doit être **exécuté depuis un serveur distant** au choix (Heroku, serveur perso, AWS...) puis **exposer la base de données**.

### Suggestion

Si vous voulez coder sous Python, vous pouvez utiliser Beautiful Soup 4. Vous pouvez aussi stocker les données sur une base SQLite ou PostgreSQL.

### Pour aller plus loin

Si le sujet vous inspire, vous pouvez aussi essayer d'agréger les ponts de Gênes via OpenStreetMap en utilisant la balise « bridge » <https://wiki.openstreetmap.org/wiki/FR:Key:bridge>