

Modélisation et enrichissement de données PFAS

Fabien Duchateau (prénom.nom@liris.cnrs.fr)

Contexte. Les PFAS (pour [Per- and polyFluoroAlkyl Substances](#), aussi appelés polluants éternels) sont des composés chimiques employés dans une multitude de produits et d'usages [1]. Le [projet PDH](#) recense les contaminations et les utilisateurs de PFAS dans une douzaines de pays européens. Les données sont actuellement stockées dans des fichiers CSV, et une carte permet de les visualiser.

En enrichissant les données existantes, le projet PDH pourrait être analysé par des toxicologues, des sociologues, etc [2]. Par exemple, les PFAS sont détaillées dans des bases de données comme [CompTox](#) ou [EcoTox](#), ce qui permettrait de connaître la composition chimique ou la structure 3D du PFAS (e.g., la molécule 375-01-9 correspond à [Comptox DTXSID4059914](#)). Quant à l'INSEE, il produit des données socio-démographiques et socio-économiques par quartier, ce qui permettrait d'étudier les populations les plus affectées par ces PFAS.

Objectifs. Ce projet cherche à atteindre les objectifs suivants :

- Proposer une représentation des données selon le modèle document (MongoDB et [format GeoJSON](#) pour les données spatiales);
- Intégrer les données des PFAS stockées en CSV dans le nouveau modèle MongoDB, ainsi que les données INSEE sur les territoires et les populations;
- Développer un algorithme d'enrichissement des données en interrogeant les bases [CompTox](#) et [EcoTox](#);
- Développer un prototype (python + web) qui permet visualiser sur une carte l'ensemble des données.

Bibliographie.

- [1] Emma L Schymanski, Jian Zhang, Paul A Thiessen, Parviel Chirsir, Todor Kondic, and Evan E Bolton. Per- and polyfluoroalkyl substances (pfas) in pubchem: 7 million and growing. *Environmental Science & Technology*, 57(44):16918–16928, 2023. [Link EVT2023](#).
- [2] Ramin Moghadasi, Tabea Mumberg, and Philipp Wanner. Spatial prediction of concentrations of per-and polyfluoroalkyl substances (pfas) in european soils. *Environmental Science & Technology Letters*, 10(11):1125–1129, 2023. [Link ESTL2023](#).