

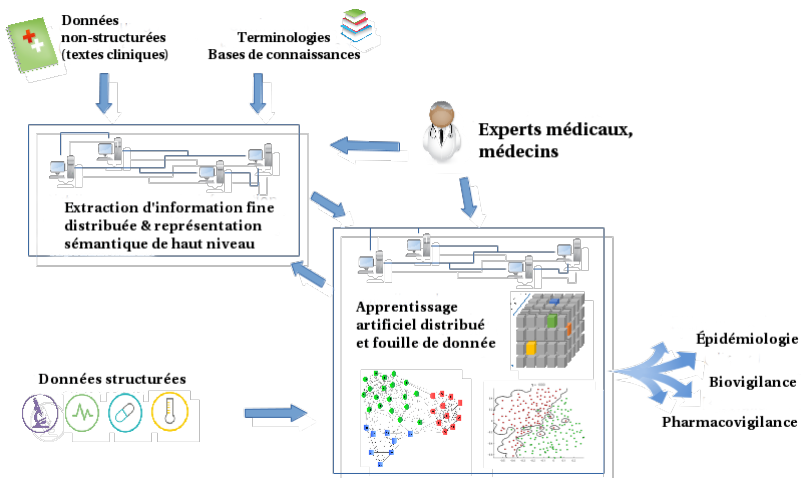
Fouille de texte et extraction d'informations dans les données cliniques

Clément Dalloux**, Natalia Grabar* & Vincent Claveau**

*UMR CNRS 8163 – STL, F-59000 Lille, France
natalia.grabar@univ-lille.fr

**CNRS, IRISA, Rennes, France,
vincent.claveau@irisa.fr
clement.dalloux@irisa.fr

BigClin : Big Data Analytics for Unstructured Clinical Data



Données médicales & apprentissage

Problèmes

- données anonymisées/désidentifiées \neq données librement accessibles
- anglais : données médicales/cliniques annotées disponible (MIMIC, I2B2, N2C2, etc.)
- français : Peu de données médicales annotées, **pas de données cliniques librement accessibles**
- résultats d'études difficilement reproductibles pour le français

Contributions

- corpus de données médicales annotées librement accessibles,
- outils d'annotation automatique (machine learning, deep learning)