



Les données, l'IA et la médecine

Pr Christian Lovis MD MPH

Service des sciences de l'information médicale @HUG

Doctorat Sciences de la Vie, "Génomique et santé numérique" @UNIGE

DALL-E, <https://labs.openai.com>

Sciences de l'information médicale



UNIVERSITÉ
DE GENÈVE

FACULTÉ DE MÉDECINE

L'équipe DataRaptor remerciements conflits d'intérêts





"The plural of
anecdote



is not
data."

attributed to
Erika Engelhaupt

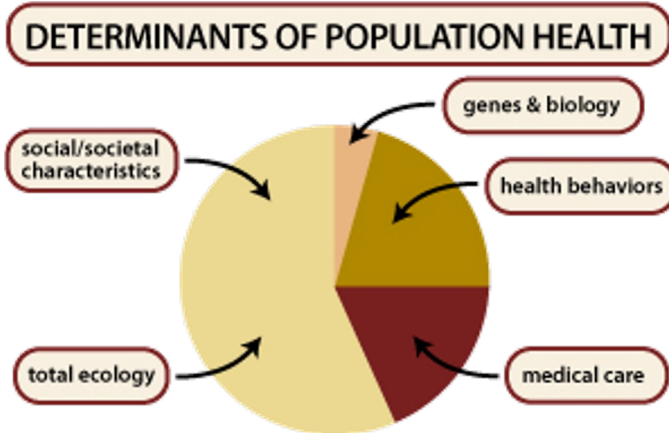


The Skeptics' Guide to the Universe - Episode 934

médecine != big data

médecine = wide data

Haute dimensionnalité
Peu de data / cas ...



Tarlov, A.R., Public Policy Frameworks for Improving Population Health. Annals of the New York Academy of Sciences, 1999. 896(SOCIOECONOMIC STATUS AND HEALTH IN INDUSTRIAL NATIONS: SOCIAL, PSYCHOLOGICAL, AND BIOLOGICAL PATHWAYS): p. 281-293.



Sciences de l'information médicale



UNIVERSITÉ DE GENÈVE
FACULTÉ DE MÉDECINE

La catégorisation altère l'information

Histoire
évoquée



Représentation
narrative

Blanche-Neige endormie,
entourée des 7 nains éplorés

Représentation de
la CIM-10

- R40.2 Coma
- T57.0 Effet toxique de l'arsenic et de ses composés
- Z63 Autres problèmes liés au groupe de soutien principal, y compris la situation familiale
- E34.3 Petite taille, non classée ailleurs

La quantité d'information diminue

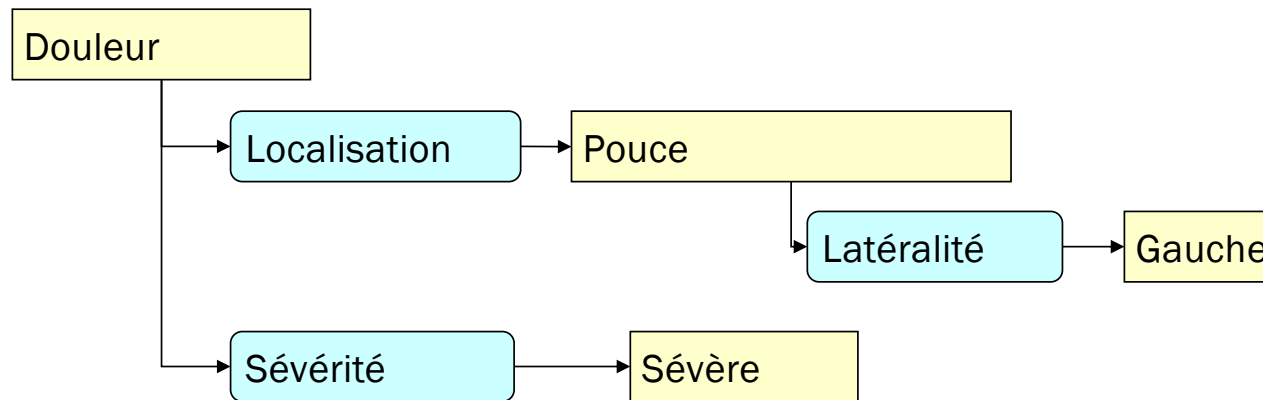
et les pertes augmentent

© Christophe Gaudet-Blavignac, adaptation

Sciences de l'information médicale



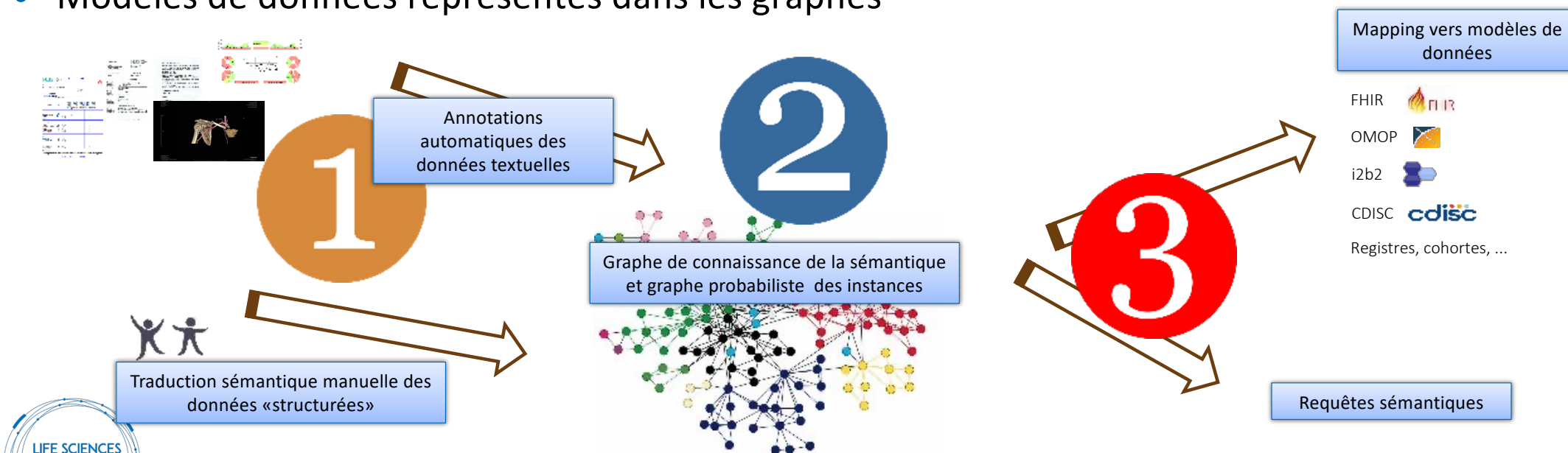
Grammaire formelle -> graphes



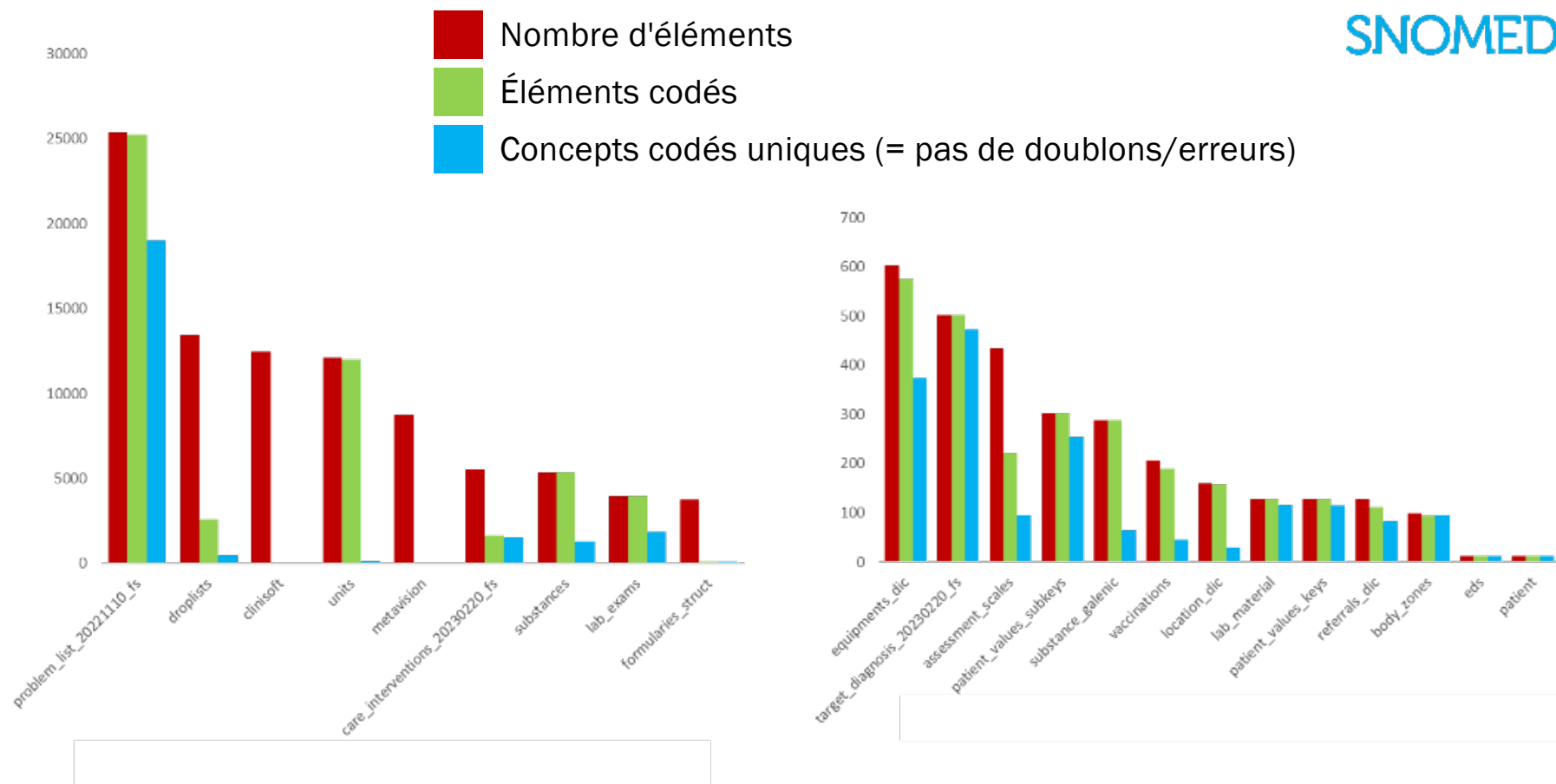
22253000|douleur| :
363698007|site de recherche|=
(76505004|structure du pouce| :
272741003|laterality|=7771000|left|),
246112005|severity|=24484000|severe|

Traduction détaillée // réutilisation généralisable

- Sémantique manuelle sur les données catégorielles
- Annotations automatiques sur les textes
- Bases de données graphes formels et probabilistes
- Modèles de données représentés dans les graphes

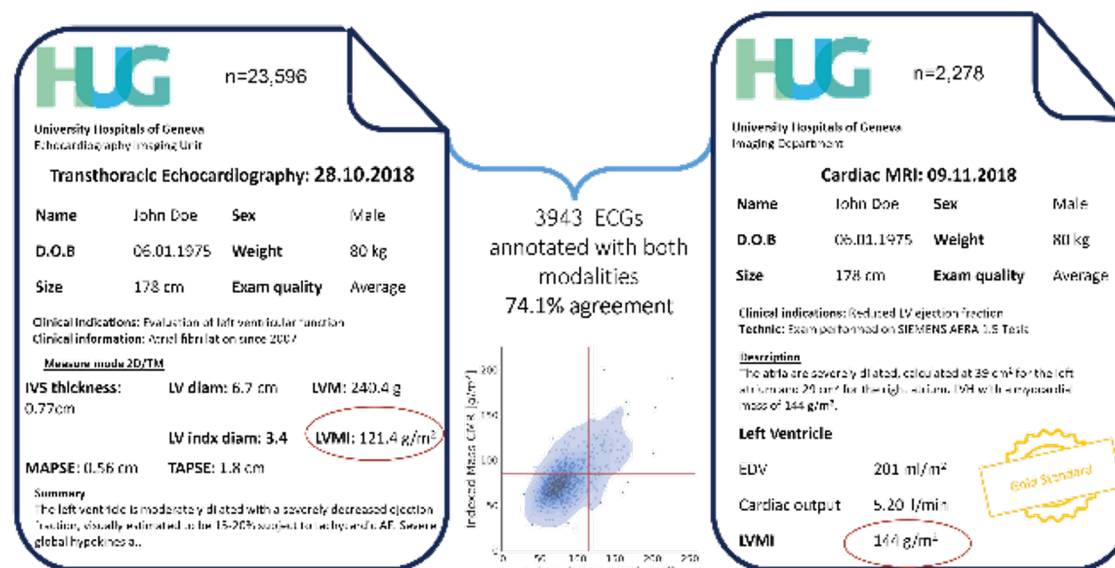


Exemple – impact sur la dimensionnalité



Exemple avec les électrocardiogrammes

- 450'000 et s'accroît chaque jour
- 200'000 ID uniques
- données brutes & DICOM d'un même dispositif d'ECG
- Annotation automatique à partir de sources cliniques

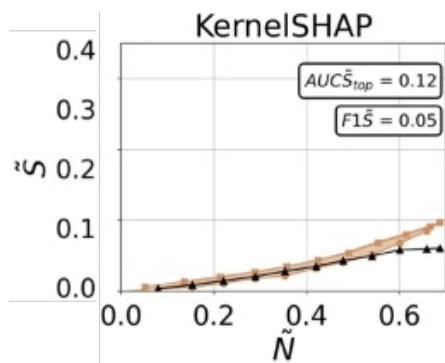
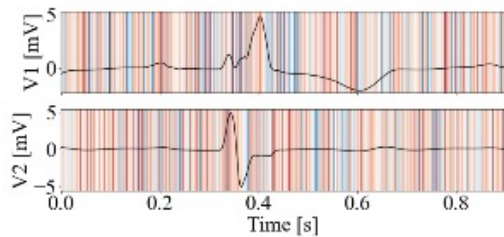


Turbé H et al. ACC 2023

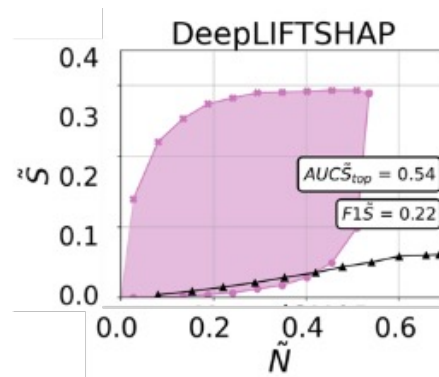
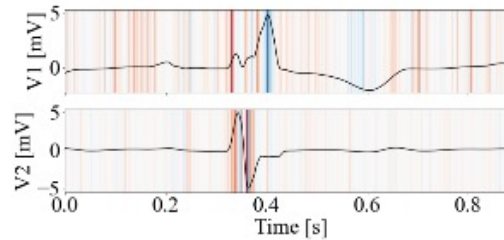
Interprétabilité = créer la confiance



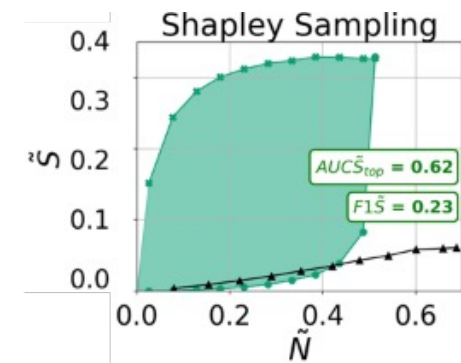
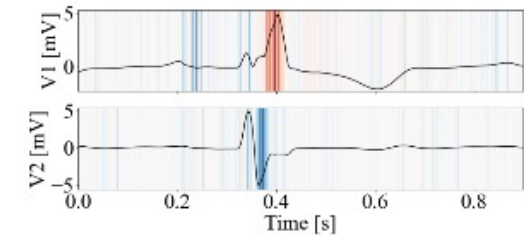
Carte bruyante



Carte approximative



carte correcte



Turbé et al, Nature Machine Intelligence 2023