

# Outils, réseaux et données pour l'étude pluri et interdisciplinaire de la thématique Violences, Santé, Sociétés



**Jean-Luc Demonsant**

**Plateforme Universitaire de Données de Toulouse**

**TGIR PROGEDO CNRS-EHESS, USR 2006**

**MSHS-T CNRS, USR 3414**

**Université Fédérale Toulouse Midi-Pyrénées**

# Les Plateformes Universitaire de Données (PUD)

2

PUD: prolongement régional de la Très Grande Infrastructure de Recherche (TGIR) PROGEDO acronyme de PROduction et GEstion des DONnées qui a pour but de **développer la culture des données en SHS**

Implantées au sein des Maisons des Sciences de l'Homme (MSH), les PUD visent à apporter un soutien méthodologique aux utilisateurs de données (accompagnement à la recherche, formations)



# La Plateforme Universitaire de Données de Toulouse

3

Objectif => structurer des communautés de chercheurs autour des données en SHS favorisant:

- une complémentarité approches quali/quantitative
- **Interdisciplinaire:** intra et hors- SHS: épidémiologie, informatique, mathématique...
- **Coloration « santé »:** Philippe Terral, UT3, référent scientifique PUD-T et département Santé de PROGEDO (et collaboration avec IFERISS)

=> Le projet « Violence santé sociétés » porté par Thomas Lefèvre (UP13) illustre bien cette volonté

# Violences, Santé, Sociétés

4

Une problématique centrale en médecine légale:  
détermination de « **ITT** »: Incapacité totale de travail / **8 jours** (question **judiciaire** : indemnisation, qualification de l'infraction)

Une **iniquité** de traitement **médical** et **judiciaire** => **hétérogénéité majeure** inter centres et inter médecins

# Comment aborder ces inégalités ?

5

## Deux grandes options

- étudier les **déterminants**, et **agir** sur ceux qui sont **causaux**
- prendre en compte la **variabilité** et en rendre compte

## Moyens

- études multicentriques, **observationnelles** puis **interventionnelles**
- **formation** des acteurs
- mise à disposition des **outils adaptés**  
informatisation, capture de l'information

**Continuum Recherche – formation – clinique**

# Méthodes:

6

- i) utiliser le certificat médical
- ii) utiliser les outils de big data et d'IA pour extraire des données communes pseudonymisées
- iii) standardiser les données des différents centres
- iv) organiser l'accès simplifié aux données et aux traitements de ces données pour les centres
- v) création d'une cohorte ouverte permettant l'appariement à d'autres données (INSEE, INED, IAU) et des études ancillaires ciblées.

# Etat des lieux

7

- phase pilote projet ORFéAD validée (janvier 2019)
- participation de 11 centres (des 47 en France)
- fédération de plus de 12 000 certificats médicaux d'examens de victimes de violences volontaires, âgées de plus de 10 ans, à l'exclusion des violences sexuelles.
- En cours: tester les outils développés et exploiter les données extraites (étude des déterminants de l'ITT)
- Projet financé par la mission Droit et Justice, le GH Paris Seine Saint-Denis et la PME Tekliko, soutenu par la PUD-T

# Compétences développées

8

- Capacité à travailler au-delà des frontières des disciplines en **SHS, santé, informatique**
- En liant recherche – formation – clinique
- **Données** : utilisation, découpage et appariement de données (enquêtes INED, INSEE, IAU..) habituellement produites et utilisées en des lieux et des moments distincts (extraction de données de certificats de 11 centres), d'origines et de natures variées
- **Traitements** : analysées par des méthodes classiques ou non classiques pour la biomédecine (dont « l'intelligence artificielle »)



## Autre projet pertinent accompagné par la PUD-T

9

“Social Science Knowledge Accumulation Initiative” (SKY) initiative de Sylvain Chabé-Ferret (TSE-INRA) visant à repenser la diffusion de la recherche

Projet multi disciplinaire sur les preuves empiriques visant à compiler des méta analyses notamment sur les effets des interventions sociales et en particulier sur la santé:

<https://chabefer.github.io/SKY/Interventions.html>

# Contact accompagnement projets SHS santé

10

Jean-Luc Demonsant, PhD

[jean-luc.demonsant@univ-toulouse.fr](mailto:jean-luc.demonsant@univ-toulouse.fr)

Ingénieur Statisticien contractuel

Plateforme Universitaire de Données de Toulouse (PUD-T)

TGIR PROGEDO CNRS-EHESS, USR 2006

<http://www.progedo.fr/promouvoir/plates-formes-universitaires-de-donnees>

MSHS-T CNRS, USR 3414

Université Fédérale Toulouse Midi-Pyrénées



# Annexe technique

12

## ○ Traitements multiples :

### ✦ Fédération des données **non structurées**

Traitement Automatisé de la **Langue** (partiellement)

**pseudonymisation**

analyse syntaxique, **sémantique** et prédictive pour l'extraction de variables

### ✦ Élaboration des **modèles prédictifs**

analyse sémantique pour structuration des données textuelles

**réductions** non linéaires de dimension

**sélection supervisée** ou automatique des variables

recherche des meilleurs modèles selon plusieurs métriques

externalisation sous forme d'**API automatisée**

**prédiction perso ITT en temps réel (en cours)**

# Annexe technique

13

- ✦ Aide en consultation:
  - capture** à la volée des données de consultation
  - intégration** du modèle **prédictif** dans le logiciel métier
- ✦ Recherche:
  - réseaux bayésiens et calculs individualisées des probabilités conditionnelles
  - clustering** robuste et consensuel
  - visualisation 3-D** et exploration / segmentation temps réel des données

# Ce qui a été - sera nécessaire

14

Réunion de **compétences**...

...et surtout de **volontés**

...et de **circonstances** : accès SI

La recherche du **compromis** sous **contraintes** fortes  
**humaines**

**économiques**

(réglementaires)

Le **paravent** du POC

organisation « projet » et recherche clinique incompatible

question de l'**industrialisation**