

# Détection automatique de phrases parallèles dans un corpus biomédical comparable technique/simplifié

Rémi Cardon, Natalia Grabar

CNRS, UdLille, UMR 8163 - STL - Savoirs Textes Langage, F-59000 Lille, France  
{prénom.nom}@univ-lille.fr

Symposium MADICS 2019



# Simplification automatique de textes médicaux

- Les méthodes actuelles nécessitent des corpus parallèles
- Pas de ressource disponible en français...
- ...mais nous avons des corpus comparables
  - Revues de littérature médicale
  - Information sur les médicaments
  - Articles encyclopédiques
- Objectif : alignement de phrases pour constituer un corpus parallèle
- Approche : classification binaire (paire de phrases alignées / non alignées)