



Aurélie Névéol, CR1 CNRS

Les textes cliniques Comme nouvelle source de données pour la recherche En épidémiologie et santé publique





Le Traitement Automatique de la Langue peut avoir un impact significatif en Santé Publique

→ Production de connaissances
épidémiologiques [Pham et al. 2014] :

- Indicateurs épidémiologiques: Prévalence des Incidentalomes?
- Évaluation de la qualité des soins: Apport du phléboscaner dans le diagnostic de la maladie thromboembolique
- Aide à la décision clinique

Je vois ce jour en consultation Monsieur Marc DURAND, né le 18/05/1954 avec les résultats d'un angioscanner thoracique et phleboscanner que j'avais demandés le 24 novembre 2004 pour suspicion d'embolie pulmonaire.

L'examen n'objective pas d'EP, ni de TVP des membres inférieurs. On observe cependant un processus ganglio-tumoral hilair gauche avec épanchement pleural bilatéral et colapsus pulmonaire associés.

Je vois ce jour en consultation Monsieur Jean DUPOND, né le 26/11/1952 avec les résultats d'un angioscanner thoracique et phleboscanner que j'avais demandés le 4 juin 2003 pour suspicion d'embolie pulmonaire. L'examen n'objective pas d'EP, ni de TVP des membres inférieurs. On observe cependant un processus ganglio-tumoral hilair gauche avec épanchement pleural bilatéral et colapsus pulmonaire associés.

Désidentification

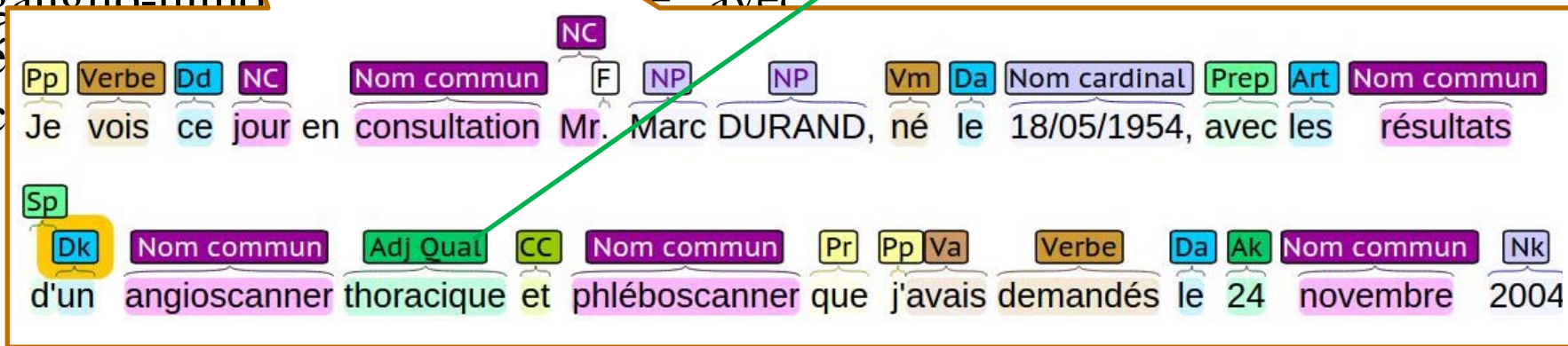
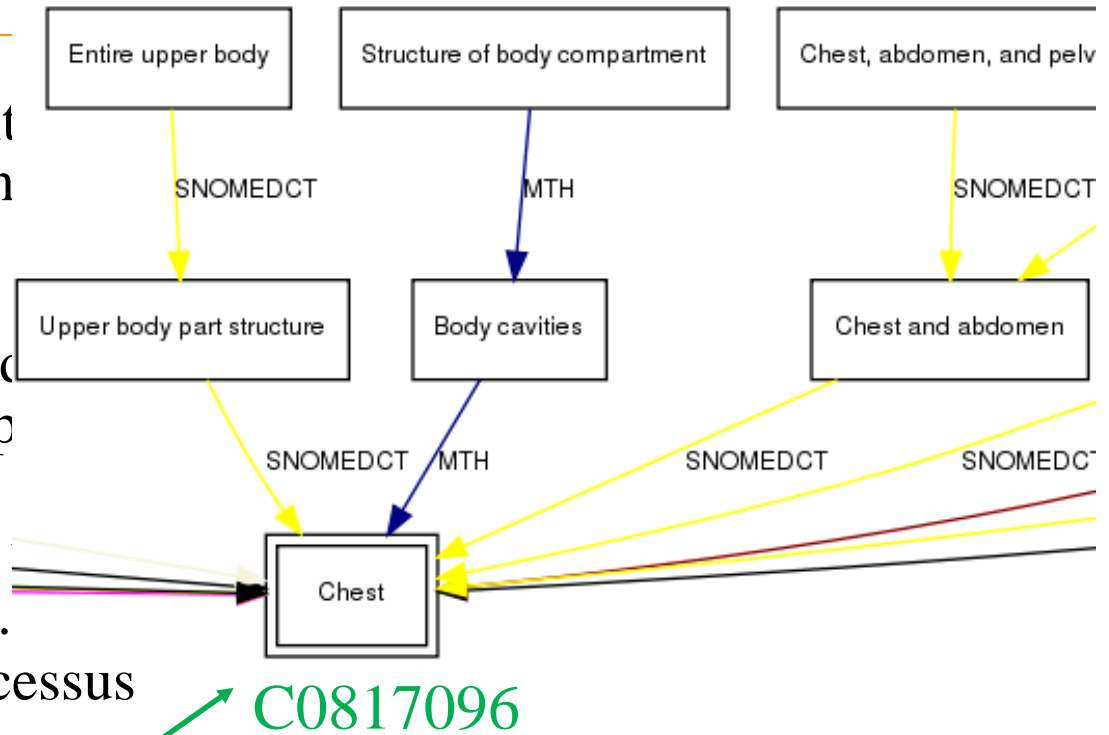
- Outillée
- Pas de ré-identification hors CHU





Traitement Automatique de la Langue Clinique

Je vois ce jour en consult
 Monsieur Marc DURAND, n
 18/05/1954 avec les résultats
 angioscanner thoracique
 phleboscanner que j'avais demandé
 24 novembre 2004 pour susp
 d'embolie pulmonaire.
 L'examen est négatif pas d'EP,
 TVP des membres inférieurs.
 observe ce processus
 ganglio-tumoral.





En pratique:

Accès à des textes cliniques en France

Données issues des dossiers patient:

- En l'absence de consentement, les données doivent être anonymisées
- Accord des comités d'éthique IRB/CERNI/CCTIRS

Limites:

- Utilisation restreinte dans le cadre d'un projet clinique (pas de « challenges » internationaux)
- Démarches mobilisant des collègues médecins:
 - Textes non pris en compte dans le cadre réglementaire
 - Quelle démarche effectuer?



Besoin: Corpus de textes cliniques désidentifiés disponibles pour l'expérimentation en TAL, partageable, avec contribution volontaire des hôpitaux

- Accès aux données grâce à un dispositif similaire à celui du SNIIRAM et des caisses d'assurance maladie
 - Ouvert mais contraint par un accord d'utilisation
 - Formation des signataires au travail avec des données humaines <https://phrp.nihtraining.com>
 - Interdiction de ré-identifier les patients
 - Mutualiser les corpus
 - Cout de développement; temps
- Comment le mettre en place dans la nouvelle réglementation?