

16|17 NOVEMBRE
2016

CITÉ DES SCIENCES ET
DE L'INDUSTRIE - PARIS

Les rencontres du
NUMÉRIQUE
de l'ANR

ANR
10
ANS

Les rencontres du

NUMÉRIQUE

de l'ANR

CABeRneT:

Compréhension Automatique

de textes Biomédicaux

pour la Recherche Translationnelle

Travail avec des données cliniques

Partage de corpus, d'annotations et d'outils

16|17 NOVEMBRE
2016

Accès à des textes cliniques en France


- Accès au contenu de 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?

Problèmes rencontrés

- Comment développer un corpus de textes cliniques désidentifiés, disponibles pour l'expérimentation en TAL, partageable, avec contribution volontaire des hôpitaux?


Éléments de Solution Proposés

- Recherche
 - Désidentification outillée [[Grouin et Névéol 2014](#), [Grouin et al. 2014](#)]
 - Étude des risques de réidentification [[Grouin et al. 2015](#)]
- Réglementation
 - Dialogue avec la CNIL et les partenaires hospitaliers



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 hilare 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 hilare gauche avec épanchement pleural bilatéral et colapsus pulmonaire associés.

Données utilisées dans CABeRneT

- Cliniques
 - Corpus HEGP (accès restreint dans le cadre d'un projet sur les incidentalomes)
 - Corpus LERUDI (accès restreint dans le cadre d'un projet sur la RI dans le DEP)
- Biomédicaux
 - Corpus QUAERO (libre, utilisé dans la campagne CLEF eHealth)
 - Corpus CéPiDC (utilisé dans la campagne CLEF eHealth avec accord d'utilisation)

Données en préparation dans CABeRneT

- Corpus de textes cliniques désidentifiés, disponibles pour l'expérimentation en TAL, partageable, avec contribution d'au moins trois partenaires hospitaliers
 - Recueil du consentement des patients
 - Pathologies non rares
 - Désidentification outillée
 - Accord d'utilisation