

Acquisition de relations sémantiques pour un web biomédical

Laurent Alamarguy, Rose Dieng-Kuntz

Projet Acacia, INRIA Sophia Antipolis

{Laurent.Alamarguy, Rose.Dieng}@sophia.inria.fr

Plan

- **Introduction: problématique**
 - Recherche Information & domaine bio-médical: WSM
 - Extraction Information & domaine bio-médical: TALN
- **Méthode**
 - Acquisition de connaissances
 - Chaîne de traitement
- **Conclusion**

Problématique 1: RI & bio-médical

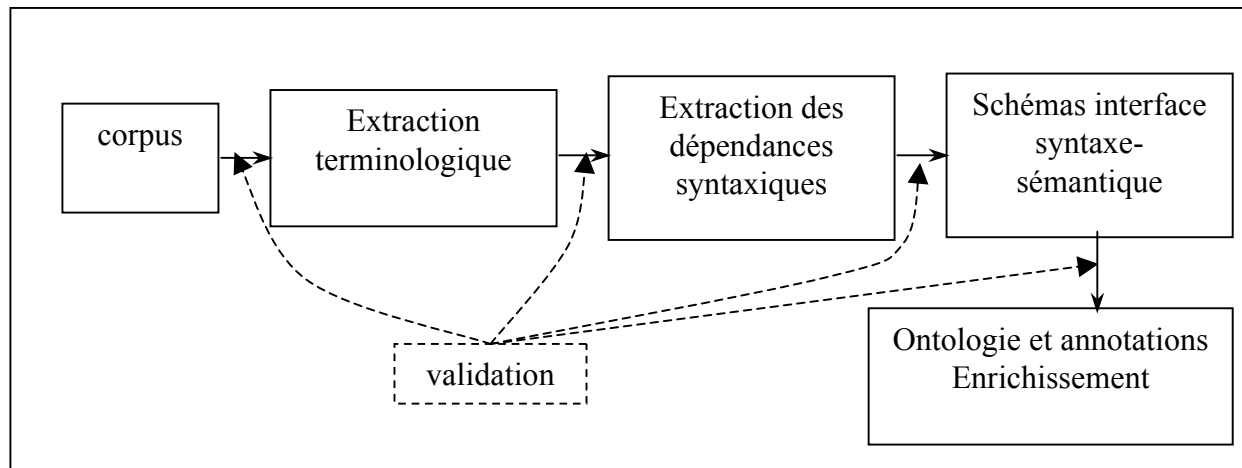
- Recherche de fonctions géniques
 - ➔ structurer le web bio-médical en fonction des problématiques de la génomique
- Moteur de recherche reposant sur des ontologies et annotations reflétant les problématiques de la génomique

Problématique 2: EI & bio-médical

- Acquisition de relations sémantiques traduisant les fonctions géniques
e.g., hypoxaemia triggers cardiovascular events in dialysis patients
- Importance des contraintes entre les relations et les concepts d'une ontologie

Acquisition de connaissances: méthodologie

- Construction et enrichissement d'ontologies et d'annotations
 - Sur corpus
 - Méthode coopérative



Méthodologie 1

- Constitution du corpus
- Extraction terminologique (Nomino, Fastr, Syntex)
e.g., [dialysis patient], [cardiovascular event]
- Extraction des dépendances syntaxiques
e.g., (1) [_{subj} *hypoxaemia*] trigger [_{obj} (*cardiovascular*) *events*] [_{prep-in} (*dialysis*) *patient*]
(2) [_{subj} (*cardiovascular*) *events*] [_{prep-in} (*dialysis*) *patient*] are triggered by [_{obj} *hypoxaemia*]

Méthodologie 2

- Interface syntaxe – sémantique

→ plusieurs schémas syntaxiques = une structure sémantique

- Grammaires d'unification

e.g., (1) *hypoxaemia triggers cardiovascular events in dialysis patients* et (2) *cardiovascular events in dialysis patients are triggered by hypoxaemia* =

Triggering-

(déclencheur)-->hypoxaemia

(déclenché)-->cardiovascular events

(domaine d'action)-->dialysis patient

Méthodologie 3

- Enrichissement d'ontologies et d'annotations
 - UMLS via MetaMap
 - Correspondance avec concepts et types conceptuels déjà existants; e.g., HYPOXAEMIA
 - Ajout de nouveaux concepts; e.g., [triggering], [cardiovascular event][dialysis patient]
- Formalisation en RDFS et RDF
- RI par Corese

Conclusion

- Structuration de relations sémantiques pour un web bio-médical
- Acquisition de relations sémantiques via les dépendances syntaxiques transcrivant les fonctions géniques