

## **REUNION DOING - 24 OCTOBRE 2019**

**LIFO, salle SR3:**

### **Participants (à cocher selon présence dans la réunion)**

- Jacques Chabin: oui
- Cédric Eichler:oui
- Mirian Halfeld Ferrari (redaction ODJ et CR):oui
- Carmem Hara: oui
- Nicolas Hiot:oui
- Anne Lyse Minard:oui
- Ciro Morais Medeiros:oui
- Martin Musicante-oui
- Christophe Rodrigues :oui
- Agata Savary: oui
- Genoveva Vargas-Solar: oui (VISIO)
- Bich Dao :oui

### **Planning**

- **13h45min: Accueil**
- **14h: Introduction**
  - Rappel: Idée générale de DOING [DOING](#)
  - Le points principaux de la dernière réunion - La liste de travaux envisageables dans le cadre médical/pharmaco
- **14h 15min: Présentations (20 minutes chaque + questions)**
  - Carmem Hara : voir Slides
  - Martin Musicante: voir Slides
- **15h 15min: Discussions (comment concrétiser nos collaborations? Demandes de stage, prochaines actions)**

### **Liste de travaux envisageables et stages à demander (DIAMS/ICVL/LIFO et autres?)**

#### **A) Extraction d'une instance (et ensuite d'un schéma) à partir d'un texte.**

Description: En prenant un domaine, par exemple pharmacovigilance (Ennov/Joshua), et un certain format-type, comment construire une base de connaissance (graphe) à partir d'un texte?

- Instances: les instances d'une classe et leurs relations...
- Schéma: à dériver à partir de l'instance
- Contraintes: quelles contraintes peuvent être dérivées à partir de l'instance
- Anonymisation: à prendre en compte déjà dans cette étape, des anonymisations à faire?
- LIEN POSSIBLE avec l'analyse d'opinion (cf Damien Nouvel) ? résumé de textes (cf Christophe Rodrigues)?

#### **Décision: 1 demande de stage dans cette thématique (extraction d'information pour peupler une BD)**

- Personnes impliquées/intéressées: Anne-Lyse, Jean-Yves, Agata, Nicolas, Mirian
- Demande de stage : Anne-Lyse fera la rédaction pour le 25 Octobre
- Références données par Genoveva concernant la question des sources de données: [KAGGLE](#); [Azure AI Gallery](#); [FHIR](#)

## B) Fusion et évolution de schéma

Description: Cet aspect serait une partie importante pour la thèse de Nicolas.

- Comment faire évoluer le schéma d'une bd graphe comme conséquence de l'évolution d'une instance ou comme conséquence d'une fusion entre deux ou plusieurs schéma ?
- Liens avec travaux précédents: Carmem Hara, Martin Musicante, Pierre Réty  
Anonymisation: à prendre en compte lors d'une fusion?

Personnes impliquées/intéressés: Nicolas, Joshua, Cedric, Jacques, Mirian...

Demande de stage? **Non, pour l'instant**

## C) Taxonomie de types de requêtes pour cadrer et illustrer *(proposition rédigée par Genoveva)*

Description:

- comment les requêtes classiques (qui attendent une réponse) peuvent intégrer des opérations d'agrégation, de calcul, et des opérations plus style data mining et
- comment articuler également des requêtes style data science (exploration) avec des techniques d'apprentissage supervisé et non supervisé. Prendre des cas d'usage grand débat / médicale
- A terme, énoncer les nouveaux aspects concernant l'expression une langue naturelle, ré-écriture, et calcul de résultats ("at speed of thought"?, optimisation, analyse dynamique)

### **Décision: 1 demande de stage dans cette thématique (data science query)**

- Une demande DIAMS est faite. La même demande sera aussi présentée à CASCIMODOT.
- Encadrants: Genoveva, Bich, Christel et Mirian (voir proposition [ICI](#)) avec participation de Nicolas

Christophe RODRIGUES nous à présenté l'article [A Neural Autoencoder Approach for Document Ranking and Query Refinement in Pharmacogenomic Information Retrieval](#) comme source d'inspiration pour des travaux liés à DOING. L'idée semble intéressante mais pas encore connectable avec les stages proposés - il faut la garder en tête pour évaluer les possibilités futures.