

Avis de Soutenance

Monsieur Nicolas HIOT

Informatique

Soutiendra publiquement ses travaux de thèse intitulés

Construction automatique de bases de données pour le domaine médical : Intégration de texte et maintien de la cohérence

dirigés par Madame MIRIAN HALFELD FERRARI ALVES

Ecole doctorale : Mathématiques, Informatique, Physique Théorique et Ingénierie des Systèmes - MIPTIS
Unité de recherche : LIFO - Laboratoire d'Informatique Fondamentale d'Orléans

Soutenance prévue le **lundi 01 juillet 2024** à 14h00
Lieu : 6 rue Léonard de Vinci, Bâtiment IIIA, 45067 Orléans
Salle : Amphithéâtre HERBRAND

Composition du jury proposé

Mme MIRIAN HALFELD FERRARI ALVES	Université d'Orléans	Directrice de thèse
M. Jacques CHABIN	Université d'Orléans	Co-encadrant de thèse
Mme Salima BENBERNOU	Université Paris Descartes	Examinatrice
M. Hubert NAACKE	Université Paris Sorbonne	Rapporteur
Mme Zoubida KEDAD	UVSQ Université Paris-Saclay	Rapporteuse
M. Joshua AMAVI	Ennov	Invité

Mots-clés : base de données graphes, base de connaissances, pharmacovigilance, structuration automatique de données, données intelligentes, contraintes

Résumé :

La construction automatique de bases de données dans le domaine médical représente un défi majeur pour garantir une gestion efficace de l'information et faciliter les prises de décision. Ce projet de recherche se concentre sur l'utilisation des bases de données graphes, une approche qui offre une représentation dynamique et une interrogation efficace des données et en particulier de leur topologie. Notre projet explore la convergence entre les bases de données et le traitement automatique du langage, avec deux objectifs centraux. Tout d'abord, notre attention se porte sur le maintien de la cohérence au sein des bases de données graphes lors des mises à jour, en particulier avec des données incomplètes et des règles métiers spécifiques. Maintenir la cohérence lors des mises à jour permet de garantir un niveau de qualité de données uniforme pour tous les utilisateurs et de faciliter l'analyse. Dans un monde en constante évolution, nous donnons la priorité aux mises à jour, qui peuvent impliquer des modifications de l'instance pour accueillir de nouvelles informations. Mais comment gérer efficacement ces mises à jour successives au sein d'un système de gestion de base de données graphes ? Dans un second temps, nous nous concentrons sur l'intégration des informations extraites de documents textuels, une source de données majeure dans le domaine médical. En particulier, nous examinons les cas cliniques et de pharmacovigilance, un domaine crucial pour identifier les risques et les effets indésirables associés à l'utilisation des médicaments. Comment détecter l'information dans les textes ? Comment intégrer ces données non structurées de manière efficace dans une base de données graphe ? Comment les structurer automatiquement ? Et enfin, qu'est-ce qu'une structure valide dans ce contexte ? On s'intéresse en particulier à favoriser la recherche reproductible en adoptant une démarche transparente et documentée pour permettre la vérification et la validation indépendante de nos résultats.