

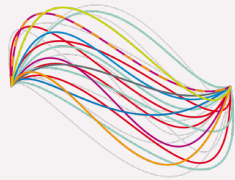
# SYSTÈMES INTELLIGENTS POUR LA TRANSMISSION DES HUMANITÉS NUMÉRIQUES ET POUR LA RECHERCHE EN SANTÉ

Directeur de thèse (HDR) : Thibaud HULIN  
Présentateur : Hao ZHANG

LECLA

25 juin 2021

UNIVERSITÉ de FRANCHE-COMTÉ **LECLA** ECOLE DOCTORALE I-SITE BFC



# Sommaire

- ① À propos de moi
- ② Humanités Numériques (HN)
- ③ À propos du projet
  - Environnement et partenaires
  - Objectifs de la recherche en IA
  - Démarche globale
  - Web sémantique et intelligence artificielle
  - Exemple d'ontologie

- ④ Conclusion

# À propos de moi

## Formation

- Étudiant de Mastère Spécialisé Smart System & IoT (Internet of Things) de CY Tech, une grande école administrée par CY Cergy Paris Université
- Diplôme d'ingénieur de l'ESLIV (École Supérieure d'Ingénieurs Léonard de Vinci) en IBO (Informatique, Big Data et Object Connecté) en option data science
- Formation de 5 ans de médecine ce qui est équivalente à une licence
- Participation à un programme d'échange franco-chinois sur la médecine d'urgences
- Note de 3e année du cycle ingénieur : 14,92 ; note de stage : 16,72 ; TOEIC 800 ; DALF Science C1 (français)



# À propos de moi

## Expérience

- Bonne maîtrise des langages de programmation (Python, R, JavaScript, etc.) et des bases de données (SQL, MongoDB, etc.)
- Bonne connaissance du domaine de Data Science, Machine Learning, TAL(Traitement automatique des Langues) et IoT
- Participation aux projets d'IA comme le traitement d'image et de langage
- Plus d'un an d'expérience dans une entreprise pharmaceutique en étudiant de nouveaux médicaments et leur résultat d'essai clinique
- Êtes curieux et ouvert d'esprit à propos des humanités en général

# Sommaire

- ① À propos de moi
- ② Humanités Numériques (HN)
- ③ À propos du projet
  - Environnement et partenaires
  - Objectifs de la recherche en IA
  - Démarche globale
  - Web sémantique et intelligence artificielle
  - Exemple d'ontologie

- ④ Conclusion

# Humanités Numériques (HN)

- Les Humanités Numériques (HN) forment un socle de partage, d'étude et de formation à des concepts ou à des compétences de haut niveau (dont dépendent les autres compétences), pour apprendre avec les technologies numériques.
- Avec les humanités numériques et les méthodes d'intelligence artificielle, la machine peut facilement reconnaître les textes manuscrits dans les archives, rechercher l'information et représenter des données via des graphes.



# Sommaire

## ① À propos de moi

## ② Humanités Numériques (HN)

## ③ À propos du projet

Environnement et partenaires

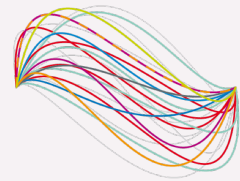
Objectifs de la recherche en IA

Démarche globale

Web sémantique et intelligence artificielle

Exemple d'ontologie

## ④ Conclusion







# Objectifs de la recherche en IA

## Les données sont :

- les scénarios de cours produits par les enseignants
- les fonds d'archives numérisés, à valoriser en pédagogie, et pour les chercheurs (notamment du pôle thématique LLC)

## L'IA :

- suggère des ressources pertinentes pour les enseignants
- permet de visualiser les données pour les chercheurs en HN, voire en pédagogie
- fournit d'autres services comme : reconnaissance du texte dans les manuscrits, système de recommandation, etc.

# Démarche globale

## Démarche globale 1 / 2

- Cartographier le champ grâce à la mobilisation d'experts du domaine. (Ce travail a déjà commencé.)
- Collecter les données d'expérience pédagogiques sur une plateforme dédiée.
- Concevoir des services en comparant des cas d'utilisation en santé et en humanités. (ex. : classification, croisement des données, regroupements, etc.)

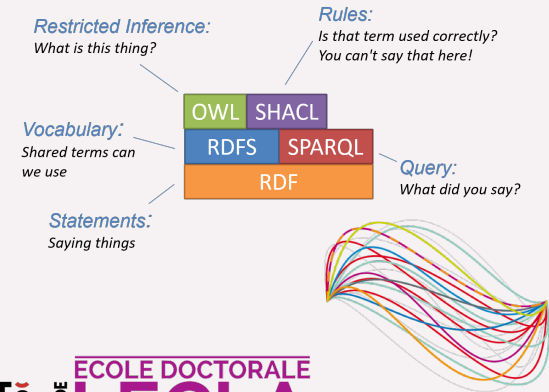
# Démarche globale

## Démarche globale 2 / 2

- Développer des services intelligents pour la recherche d'information, la visualisation et l'analyse des données.
- Évaluer les usages (analyse des traces, mouvements oculaires, enquêtes...). La recherche de eye-tracking peut être en collaboration avec l'équipe ERCOS de ELLIADD.
- Observer les pratiques d'enseignement et communiquer sur le projet.

# Utiliser les technologies du web sémantique pour construire des services intelligents

- Le W3C est un organisme de standardisation à but non lucratif pour faciliter l'échange de données sur le Web. La co-directrice de thèse, experte en informatique et en sciences cognitives, Mme McGuiness en fait partie et y contribue.
- Le Web sémantique est une extension du Web standardisé par le W3C. Ses standards encouragent l'utilisation de formats de données et de protocoles d'échange normés sur le Web, en s'appuyant sur le modèle Resource Description Framework (RDF) qui est la technologie centrale du Web sémantique.



# Web sémantique et intelligence artificielle

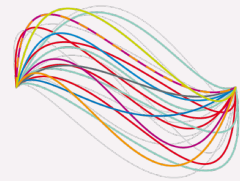
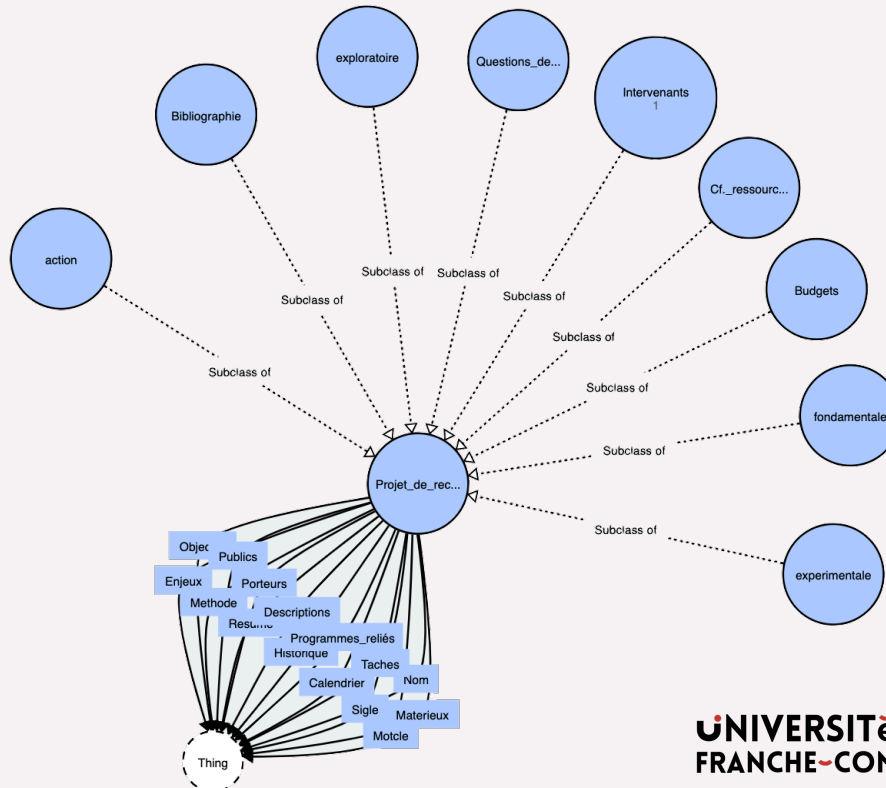
## Rôle du Web Sémantique

- modéliser le domaine d'un discours
- identifier les éléments (terminologie), les relations, notamment entre classes et individus
- construire une ontologie (base de connaissance)

## Rôle de l'IA grâce au WS

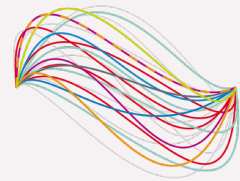
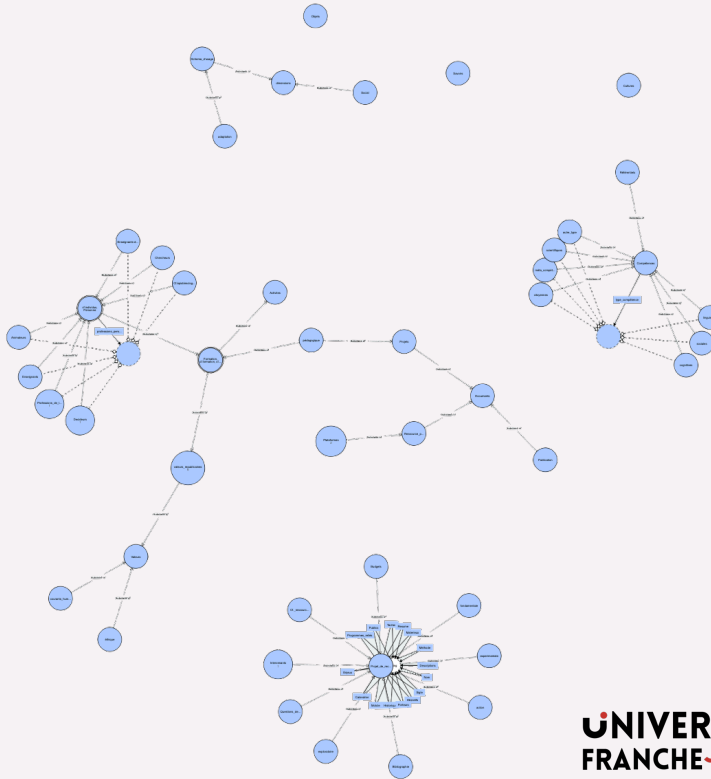
- trouver des régularités
- calculer des proximités entre ressources
- construire automatiquement une ontologie à partir d'un corpus de textes (via une intelligence artificielle de type connexionniste)

# Exemple d'ontologie



# Exemple d'ontologie

WebVOWL  
1.1.7



# Sommaire

- ① À propos de moi
- ② Humanités Numériques (HN)
- ③ À propos du projet
  - Environnement et partenaires
  - Objectifs de la recherche en IA
  - Démarche globale
  - Web sémantique et intelligence artificielle
  - Exemple d'ontologie

- ④ Conclusion



# Conclusion

Les apports de la thèse pour l'école doctorale sont :

- Valoriser des corpus issus de la **littérature** (ex. : le fond Queneau de l'UB), dans le cadre de séances de médiation des savoirs ou de médiation culturelles (fonds photographiques) ;
- Renouveler l'analyse **linguistique** de ces corpus grâce à la sémantisation des données issues des textes (via une ontologie comme Wordnet).
- L'ontologie permettra aussi de travailler sur les traductions des textes en différentes langues, ex. : les **traductions** des textes de Queneau.
- Ainsi, toutes les grandes disciplines de l'école doctorale seront conviées pour participer au projet sur **l'Émancipation du Pôle Thématique Lettres, Langues et Communication**.

# Conclusion

## Pour conclure...

- Une occasion unique de collaboration internationale avec celles et ceux qui font le web
- Cette thèse s'inscrira dans un environnement très stimulant, impliquant de nombreuses disciplines : toutes celles enseignées de l'école au lycée, en SHS (linguistique, Information-Communication, littérature, langues...), en informatique, en science cognitive et en santé.
- Des retours sont attendus pour tous les participants de laboratoires différents, via le programme Émancipation du Pothem LLC.
- Mon dossier a été sélectionné par l'Institut Polytechnique de Rensselaer (état de NY) et par M. HULIN grâce à mes compétences informatiques, en langues et en santé, mon expérience en IA et mon niveau d'ingénieur.



Merci !

## Quelques références

- Ageeva, N. A., Shapoval, G. N., Vlasova, V. N., Kartashova, E. A., Safronenko, A. V., & Sidorenko, Y. A. (2019). High level of legal awareness formation in medical students. Way from competencies to competence. *Espacios*, 40(9), 11.
- Berners-Lee, Tim, James Hendler, et Ora Lassila. « The semantic web ». *Scientific american* 284, no 5 (2001) : 34-43.
- Hulin, T. (2016). Vers une plateforme sémantique pour l'enseignement des sciences et de la culture numérique. *Tréma*, 42, 37-49.
- McCusker, J., Rashid, S. M., Agu, N., Bennett, K. P., & McGuinness, D. L. (2018). The Whyis Knowledge Graph Framework in Action. *International Semantic Web Conference (P&D/Industry/BlueSky)*.
- Roxin, I., & Tajariol, F. (2019). Information, communication et humanités numériques. Enjeux et défis pour un enrichissement épistémologique. *Actes du colloque « HumaNum » – I Roxin et all.*
- Hulin, T., Collet, L. & Rémond, É. (2021). Les enjeux de la formation de formateurs aux humanités numériques. In F. Paquienséguy & N. Pélissier, *Questionner les humanités numériques : Positions et propositions des SIC (SFSIC et CPdirsic, p. 257-273). CPDirSIC.*