

Stage Big Data au CHU Grenoble Alpes : Plateforme PREDIMED, imagerie

Le CHU Grenoble Alpes

CHU du bassin de santé de Grenoble (800 000 habitants) et établissement de référence pour le sillon alpin (Isère, Savoie et Haute-Savoie : 2 millions d'habitants), le CHU Grenoble Alpes (CHUGA) est particulièrement actif en soins de haute spécialité (activité de recours et d'expertise) et en clinique : il bénéficie par ailleurs de l'environnement scientifique exceptionnel de l'Université Grenoble ALPES (UGA), partenaire du projet, et plus globalement de l'agglomération grenobloise et du carrefour européen Grenoble-Lyon-Genève. Pour une présentation complète du CHUGA, nous vous invitons à vous référer au site internet institutionnel du CHUGA <https://www.chu-grenoble.fr/>.

Missions

Le CHUGA a mis en place un Entrepôt de Données de Santé, PREDIMED, une plateforme de Données Massives, destinée à exploiter pour la recherche, amélioration de la prise en charge des patients et le pilotage d'établissement des données issues de sources hétérogènes de données de soins du CHUGA (biologie, documents texte, prescription, données génomiques, autres). Les données d'imagerie, bien qu'essentielles pour la recherche médicale ne sont encore que peu exploitées dans PREDIMED.

Ce stage, à l'interface entre la recherche médicale et les dernières avancées en informatique, a comme objectif d'améliorer l'intégration et l'exploitation de cette source de données dans PREDIMED. En pratique, il s'agit :

1. De s'approprier l'utilisation de la plateforme PREDIMED, les contraintes de sécurité et de confidentialité liées aux données personnelles et les sources de données nécessaires au projet.
2. D'apprendre à rechercher les images dans le PACS (système d'archivage et de transmission d'images) en les (pré-)triant selon des critères présents dans les métadonnées des images, les comptes rendus texte (analyse de texte), et les bases de données relationnelles associées, via des scripts dédiés.
3. De développer des scripts génériques qui permettront le chargement de données identifiées vers la plateforme PREDIMED.
4. D'explorer en profondeur les données d'imagerie, les métadonnées et les comptes rendus textes associés issus des logiciels métiers utilisé pour le soin.
5. De constituer le jeu de données contenant les images nécessaires à un projet pilote, dans un espace dédié, et d'en évaluer la qualité et la cohérence des données extraites.
6. De traiter les différentes sources d'information non seulement de manière indépendante mais également de démontrer la faisabilité à exploiter toute la richesse du croisement de l'exploitation de toutes ces sources.
7. D'explorer la faisabilité du chargement des données machine à l'origine des images reconstruites dans la plateforme en étroite collaboration avec le service d'imagerie du CHUGA.
8. De documenter le travail effectué.
9. De travailler au sein de PREDIMED en collaboration entre le Pôle Santé Publique, la Direction des Services Numériques et le service de Radiologie du CHUGA ainsi qu'avec une jeune startup de la région et en collaboration avec le laboratoire TIMC.

Ce stage est l'opportunité d'approfondir ses compétences informatique dans un environnement innovant (Hadoop, Orthanc, SolR) et cela autour du domaine d'application passionnant de l'imagerie médicale contextualisée cliniquement. Un des enjeux de ce projet réside également dans une première formalisation des pipelines nécessaires à la réalisation du projet (des données sources à la base de données qualifiée). L'application de méthode d'IA pour la fouille dans les données pourra être envisagée en seconde partie du stage.

Le stage sera réalisé au sein du Département de Méthodologie de l'Information en Santé du Pôle Santé Publique du CHUGA.

Fonctions

- Les responsables de ce stage seront Dr K. Charrière et Dr C. Lenne.
- Un compte-rendu hebdomadaire sera réalisé auprès du groupe de travail « Données Massives ».

Livrables

- Les scripts automatisés d'alimentation de la plateforme PREDIMED avec des données d'imagerie
- Les scripts d'analyse de métadonnées et des textes associés pour trier des images avant le chargement sur la plateforme de big data.
- Les scripts de croisement de données issues de différentes sources afin, en particulier, de proposer de nouveaux indicateurs caractéristiques de la qualité associée au jeu de données.
- L'étude de faisabilité de la récupération des données « machine » à l'origine des images reconstruites.
- Le jeu de données pour le projet extrait avec sa qualification sur un espace dédié.
- Une documentation détaillée sur tous les travaux réalisés pendant le stage.

Savoir-faire et savoir être

- Expérience en développement informatique
- Connaissances dans le domaine de l'imagerie seraient un plus
- Rigueur, initiative et curiosité, autonomie et travail en équipe

Contacts (envoyer un CV et une lettre de motivation)

Katia Charrière (kcharriere@chu-grenoble.fr) et Christian Lenne (EXT-CLenne@chu-grenoble.fr)