

Mise en place d'une base de données FAIR interconnectée avec VIP

Contexte

La plateforme d'imagerie multimodale et expérimentale LyonTech PILOT située sur le campus de La Doua à Villeurbanne regroupe les modalités d'imagerie par résonance magnétique (IRM), échographie (US) et optique. Le domaine d'application est le médical avec la production d'images et/ou de spectres dans ces 3 modalités pour extraire des biomarqueurs les plus pertinents pour diverses pathologies. Certaines données sont acquises sur des modèles d'animaux sains et pathologiques et sur l'homme généralement pathologique. Ces données sont exploitées pendant la période de financement pour produire des résultats sous forme d'un ou plusieurs articles scientifiques. Le devenir des données après publication est totalement inconnu car aucun suivi, référencement, ou encore une réutilisation n'est mise en place.

Girder (<https://girder.readthedocs.io/en/stable/>) est une plateforme de gestion de données gratuite, open-source et développée par Kitware. Girder permet une bonne organisation des données et une gestion facilitée des utilisateurs. Le laboratoire CREATIS a une bonne expérience avec cette plateforme, plusieurs instances étant déployées au laboratoire. Cette solution de base de données (BD) conviendra très bien pour le stockage et la gestion des données issues de la plateforme PILOT.

VIP (Virtual Imaging Platform) est un portail web (<https://vip.creatis.insa-lyon.fr>) pour l'imagerie médicale. VIP compte, en octobre 2019, plus de 1000 utilisateurs enregistrés et une vingtaine d'applications. En utilisant des ressources de calcul et stockage distribuées, VIP offre à ses utilisateurs académiques des services leur permettant d'exécuter simplement et de manière distribuée des applications d'imagerie médicale. Dans le cadre de précédents travaux, nous avons déjà développé un plugin Girder permettant de lancer des traitements à travers les applications disponibles dans VIP. Suite à ce stage, les données de la base de données pourront donc naturellement profiter de cette fonctionnalité.

Objectifs

Ce stage traitera de problématiques de stockage et de traitement FAIR (localisable, accessible, interopérable et retrouvable) des données de recherche issues de la plateforme PILOT.

Pour cela, nous proposons de mettre en place une base de données de type Girder, qui permettra d'accueillir les données issues de PILOT et de gérer leurs droits d'accès conformément au souhait de leur propriétaires. Nous proposons d'extraire et publier les métadonnées pertinentes permettant de rendre ces données localisables et retrouvables. Enfin, pour faire le lien avec les traitements qui peuvent être lancés sur ces données, la base de données sera interconnectée avec la plateforme VIP.

Plus précisément, l'étudiant recruté devra travailler sur :

- La prise en main d'une instance de BD Girder
- L'étude des données disponibles sur PILOT afin de proposer une structure adéquate pour leur intégration dans la BD
- Définition des métadonnées à extraire et publier
- Automatisation de l'extraction de ces métadonnées
- Mise en place et test d'un pipeline complet de l'import des données dans la BD, l'extraction des métadonnées et l'exécution de traitements avec VIP

Encadrement

Le stage sera principalement encadré par Laurent Mahieu-Williams, ingénieur de recherche au laboratoire Creatis et expert de la plateforme PILOT. Sorina Pop et Frédéric Cervenansky interviendront aussi pour les aspects liés à la mise en place de base de données Girder et le portail VIP.

Durée : 5 à 6 mois

Compétences requises :

- Bonnes compétences en informatique, programmation et systèmes Linux
- Connaissances en bases de données et/ou traitement de données
- Rigueur, esprit critique, intérêt pour la recherche

Profil recherché : Informatique, niveau Master ou 5ème année ingénieur

Contact : Sorina Pop, email : sorina.pop@creatis.insa-lyon.fr