

Stage de Master 2 en développement bioinformatique

“Evolution de l’interface web d’Omicrobe afin d’ordonner les données selon leur qualité”

Durée

Six mois de stage.

Lieu

Centre de recherche INRAE de Jouy-en-Josas (78).

Contexte

L’unité Mathématiques et Informatique Appliquées du Génome à l’Environnement (MaIAGE)¹ est située sur le centre INRAE² de Jouy-en-Josas. Cette unité de recherche regroupe des mathématiciens, des informaticiens, des bioinformaticiens et des biologistes qui développent des méthodes pour répondre à des questions de biologie et agro-écologie, allant de l’échelle moléculaire à l’échelle du paysage en passant par l’étude d’individus, de populations ou d’écosystèmes. MaIAGE est structurée en cinq équipes dont l’équipe Acquisition et formalisation des connaissances à partir de textes (Bibliome)³, l’équipe Bioinformatique et statistique des données “omiques” (StatInfOmics)⁴ et la plateforme bioinformatique Migale⁵. Bibliome développe des méthodes de traitement automatique des langues (TAL) et d’apprentissage automatique (ML) pour extraire des informations de textes par des ontologies dans le domaine de la biologie. StatInfOmics développe et met en œuvre des méthodes statistiques et bioinformatiques dédiées à l’analyse de données “omiques”. Migale fournit des services à la communauté des sciences de la vie. Cette proposition de stage s’inscrit sur un projet commun aux trois équipes.

Missions

L’unité MaIAGE développe l’application Omicrobe⁶ qui rassemble des informations sur les habitats, les phénotypes et les usages des micro-organismes, extraites automatiquement de sources textuelles (PubMed, GenBank, DSMZ, CIRM). Les données sont accessibles via une interface web⁷ et une interface programmatique⁸ (API). Afin de faciliter la lecture par l’utilisateur, le ou la stagiaire aura pour mission de modifier l’interface web afin d’ordonner les informations de l’interface en fonction de leur qualité. Le ou la stagiaire utilisera des métriques basées sur la combinaison de l’évaluation de la qualité des relations agrégées et des informations prédites par les outils de *text-mining*.

Profil du candidat

Master 2 / dernière année d’école d’ingénieur en bio-informatique.

Compétences souhaitées :

- Javascript / Python
- PostgreSQL

Contacts

- Louise Deléger, louise.deleger@inrae.fr
- Sandra Dérozier, sandra.derozier@inrae.fr

Références

- ¹ Unité MaIAGE, <https://maiage.inrae.fr/>
- ² INRAE, <https://www.inrae.fr/>
- ³ Equipe Bibliome, <https://maiage.inrae.fr/fr/bibliome>
- ⁴ Equipe StatInfOmics, <https://maiage.inrae.fr/fr/statinfomics>
- ⁵ Plateforme Migale, <https://migale.inrae.fr>
- ⁶ Dérozier S *et al.* Omnicrobe, an open-access database of microbial habitats and phenotypes using a comprehensive text mining and data fusion approach. *bioRxiv.* 2022. DOI: <https://doi.org/10.1101/2022.07.21.500958>
- ⁷ Interface web Omnicrobe, <https://omnicrobe.migale.inrae.fr/>
- ⁸ Interface programmatique Omnicrobe, <https://omnicrobe.migale.inrae.fr/api>