



Proposition de stage de Master 2 : premier semestre 2024

Nom de l'entreprise ou du laboratoire : URGI, INRAE

Adresse où se déroulera le stage :

INRAE, URGI Bat 18, RD10 Route de Saint-Cyr, 78000 Versailles

Responsable du stage (personne qui sera contactée par les candidats):

- Nom, Prénom : POMMIER Cyril, <https://cv.hal.science/cyril-pommier>
- Statut : Ingénieur de recherche
- Coordonnées : cyril.pommier@inrae.fr

Titre du stage : Développement de workflow ETL et d'interfaces web pour le portail de données plantes FAIDARE.

Mots clés résumant les méthodes et techniques à utiliser au cours du stage :

- Langages et outils de développement : Python, conteneur Docker, GNU/Linux, Shell/Bash, Java, Spring, Angular, RDF, JSON, web services REST.

Résumé du projet de stage :

L'URGI est une unité de recherche INRAE dédiée aux plantes. Elle héberge une plateforme de bio-informatique dont l'une des activités majeures est d'assurer, en suivant les principes FAIR, le partage, la standardisation et la structuration des données de recherche de projets nationaux et internationaux. Le partage des données passe, entre autres, par l'utilisation de portails de données, dont en particulier le portail FAIDARE (<https://urgi.versailles.inrae.fr/aidare/>) qui est développé à l'URGI. Il s'agit d'un portail de recherche de données scientifiques indexant plus de 30 bases de données internationales à ce jour. Divers supports de formation^{1,2} donnent un aperçu de ses fonctionnalités. Ce logiciel sera l'objet de ce stage, son architecture repose sur : une base de données NoSQL Elasticsearch, alimentée avec des JSON compatibles avec la BreedingAPI (www.brapi.org) via un ETL en python³ et des scripts de chargements en bash. L'interface de requêtage⁴ repose sur un *backend* Java + Spring Boot et un *frontend* en Angular, communiquant par web services REST.

La personne recrutée pour ce stage participera à la mise à niveau vers la BrAPIV2 de

¹ https://moodle.france-bioinformatique.fr/pluginfile.php/638/course/section/118/Module6.2_FAIDARE-overview.pptx?time=1686776684906

² <https://hal.inrae.fr/hal-03851454>

³ <https://github.com/elixir-europe/plant-brapi-etl-faidare>

⁴ <https://github.com/elixir-europe/plant-faidare>

l'application, la mise en place de fonctionnalités d'export de données, l'amélioration du workflow d'ETL, et éventuellement l'ajout de fonctionnalités de recherche utilisant des alignements d'ontologies⁵. Elle sera intégrée à une équipe de développement de cinq personnes travaillant sur des projets parallèles. Elle sera encadrée en suivant les méthodologies agiles afin de lui permettre de prendre en main les technologies et l'architecture dans de bonnes conditions.

Montant des indemnités de stage

Approximativement 500€.

Modalités de candidature

Les candidatures (CV + lettre de motivation) doivent être adressées au plus tard le 15/11/2023 par courriel à cyril.pommier@inrae.fr avec l'objet suivant : [stage 2024 – développement - FAIDARE]. Les candidatures seront évaluées au fur et à mesure.

⁵ <https://doi.org/10.57745/ZLJYQQ>