



## Stage Master 1 (ou équivalent) en informatique / Machine learning Apprentissage de langues étrangères assisté par ordinateur

### Contexte

**ALAI A (Apprentissage des Langues assisté par Intelligence Artificielle)**<sup>1</sup> est un laboratoire commun entre la PME Archean Technologies (Montauban, France) et l'IRIT (Institut de Recherche en Informatique de Toulouse). Ce laboratoire est soutenu par le programme LabCom de l'Agence Nationale de la Recherche (2019-2023). Il réunit les compétences complémentaires pour le développement d'applications d'apprentissage des langues étrangères selon une approche multidisciplinaire combinant pédagogie des langues étrangères, linguistique et science des données.

Un travail important est mené sur la **catégorisation des erreurs de prononciation des apprenants**. A partir d'un corpus d'enregistrements audio, contenant les énoncés produits par des étudiants japonophones apprenant le français vous appliquerez des méthodes pour catégoriser ces erreurs suivant les critères établis par des spécialistes d'enseignement du français langue étrangère (FLE).

### Description de la mission

La mission consistera plus précisément à développer, par des techniques de **machine learning et d'augmentation de données**, un système prenant en entrée **un enregistrement** (énoncé prononcé par un apprenant) et déterminant si le **son ciblé** (un phonème particulier) est correctement prononcé ou bien correspond à une des catégories d'erreurs identifiées.

Les performances de ce système devront être évaluées et comparées à celles obtenues avec une autre approche déjà développée et reposant sur l'utilisation d'un système de reconnaissance automatique de la parole. Plusieurs phonèmes cibles pourront être traités au cours du stage. Le code développé sera intégré au démonstrateur développé par l'entreprise.

### **Compétences techniques**

- Programmation (Python, Shell), traitement du signal (si possible de parole) et apprentissage automatique (développement et évaluation).
- Boîtes à outils d'apprentissage automatique (Sklearn, PyTorch, Tensorflow, ...).

### **Autres compétences**

- Aptitude pour le travail en équipe, bonne capacité d'organisation et de restitution de son travail.
- Bonnes pratiques de développement informatique (commentaires, documentation, versionnage et archivage).

### **Conditions de travail**

- Lieu de travail : IRIT – TOULOUSE
- Rémunération : gratification selon les tarifs en vigueur.
- Durée de la mission : ~4 mois, à partir d'avril **2023**

### **Candidature**

CV et lettre de motivation à envoyer à : [isabelle.ferrane@irit.fr](mailto:isabelle.ferrane@irit.fr) et [lfontan@archean.tech](mailto:lfontan@archean.tech)

---

<sup>1</sup> <https://www.irit.fr/SAMOVA/site/projects/current/labcom-alaia/>