

Projet étudiant en traitement du langage naturel

Niveau L3 · LISN · Université Paris-Saclay

Contexte. Dans le cadre de recherches en traitement automatique des langues (TAL) portant sur les *propositions atomiques* (des unités sémantiques minimales extraites automatiquement de textes), je propose un mini-projet étudiant d'annotation multilingue et d'approfondissement thématique.

Description du projet.

Volet annotation. Après une présentation du sujet de recherche et de l'interface d'annotation, les participant(e)s contribueront à l'annotation de données en **anglais, français, espagnol, allemand, italien** et/ou **portugais** (selon les compétences linguistiques).

Volet projet. Un approfondissement personnel au choix :

- ▷ **Apprentissage automatique** : fine-tuning de modèles de langue (PyTorch, Transformers)
- ▷ **Logique formelle** : sémantique computationnelle, théorie de l'information (Carnap-Bar-Hillel)

Modalités. Durée minimum **2 semaines**, démarrage **flexible**, sur site au **bâtiment 507 du LISN** (Orsay).

Ce que vous y gagnez

- Immersion dans un **laboratoire de recherche public** (CNRS / Université Paris-Saclay)
- Expérience **valorisable sur CV** : participation à un projet de recherche actif
- **Encadrement personnalisé** et initiation aux pratiques de la recherche en TAL

Profil recherché : étudiant.e de **L3**, motivé.e, avec un intérêt pour le TAL, la linguistique computationnelle ou la logique. *Aucun prérequis technique avancé n'est nécessaire.*

Contact : Luc POMMERET
pommeret@lisn.fr