

Stage « TAL pour les corpus en littérature »

Équipe Littératures, Arts et Numérique (ELAN)
au sein du laboratoire Litt&Arts (UMR 5316), Grenoble
Année 2020

En appui à nos projets de recherche, nous proposons 1 stage de 4 à 6 mois en enrichissement de corpus numériques TEI.

Spécialité recherchée : traitement automatique des langues et corpus textuels

Profil du candidat : étudiant·e en TAL ou linguistique outillée ayant un intérêt pour les corpus littéraires et la TEI ou étudiant·e dans l'une des disciplines des SHS dans une filière à très forte coloration numérique et/ou informatique

Contexte

Au sein de l'UMR Litt&Arts, ELAN est une équipe d'ingénieur·e·s accompagnant les projets de recherche de l'unité. De nombreux projets, qu'ils traitent de manuscrits, de correspondances, de bibliographies ou d'autres types de données, nous amènent à manipuler des corpus en TEI (de leur modélisation à leur visualisation en ligne en passant par des outils de recherche). Le stage propose de travailler sur les données d'un ensemble de projets choisis et de concevoir - selon les cas et les objectifs du projet - une ou plusieurs approches utilisant des outils et méthodes issues du TAL afin d'enrichir les corpus, leur visualisation ou encore les exploitations faites de ces données.

Objectifs du travail

Le stagiaire devra, sous la responsabilité de deux ingénieures d'ELAN :

- étudier les spécificités des corpus littéraires vs les corpus linguistiques et identifier les besoins de nos projets ;
- faire un état de l'art d'outils utilisés en linguistique et en TAL pour répondre à ces besoins ;
- compiler les corpus issus de différents projets partageant une même problématique afin de constituer un ensemble conséquent ;
- tester différents outils et méthodes pour analyser le corpus ainsi constitué ;
- documenter ces tests et rédiger des méthodologies à destination d'utilisateurs néophytes ou peu formés.

Compétences recherchées

- bonne connaissance du domaine du TAL ;
- bonnes capacités rédactionnelles ;
- lecture courante de l'anglais ;
- la connaissance du format XML, voire de la TEI sera appréciée ;
- capacité d'organiser son propre travail avec rigueur.

Nous attendons du ou de la candidate un goût certain pour la transmission du savoir et des capacités didactiques. En effet, les documentations devront viser une autonomie maximale des membres des projets, tant sur la mise en œuvre de l'outil ou de la méthode que sur l'utilisation, l'exploitation et l'interprétation des résultats.

Cadre du stage

La·e stagiaire sera accueilli au sein du Laboratoire Arts et pratique du texte, de l'image, de l'écran et de la scène (Litt&Arts, UMR 5316, UGA/CNRS). Il ou elle sera accompagné·e dans son travail par deux encadrantes et travaillera en collaboration avec plusieurs chercheuse·eur·s de l'unité.

Durée et date de début du contrat

Le stage est prévu au printemps-été 2020.

Répondre à l'offre

Les candidats doivent envoyer :

(1) un CV, (2) une lettre de motivation et (3) une lettre de recommandation d'un enseignant

par mail, ayant comme objet :

[Stage ELAN] Candidature de M/Mme Prénom NOM

à :

Anne Garcia-Fernandez : annegf@univ-grenoble-alpes.fr

Elisabeth Greslou : gresloue@univ-grenoble-alpes.fr