

***Réalisé par :***

* ***Waël KOUJUK***
* ***Lucien KERISIT***
* ***Patty RANDRIAMBOLOLONA***

***Encadrant :***

* ***Line JAKUBIEC-JAMET***

**Sémantique et langue naturelle**

# Introduction

Le Traitement Automatique des Langues (TAL) est un domaine pluridisciplinaire ou travaillent en collaboration étroite des linguistes et des informaticiens dans le but de développer des applications capables de traiter des données linguistiques.

Ce projet s'inscrit dans le cadre du TER de première année de Master Informatique à l'Université Aix-Marseille. Le sujet proposé par Madame Line Jakubiec-Jamet s'insère dans le domaine de la formalisation et du traitement des langues naturelles.

Il nous est demandé de développer une interface Web qui permettra d'attribuer à chaque mot d'un texte une étiquette sémantique, en fonction d'une description déjà définie au préalable, dans le but d'évaluer ensuite la sémantique de phrases entières voire du texte complet.

L'application, évaluant la sémantique de mots rencontrés dans un texte, vise un public qui par exemple veut apprendre le français.

Dans ce rapport nous allons vous présenter notre projet et détailler sa conception de manière explicite de la phase d'analyse au produit final.

Pour cela, nous ferons dans un premier temps un examen attentif du sujet et nous évoquerons les méthodes et outils pouvant être utilisés pour mener à terme le projet.

Dans une deuxième partie, il s'agira de détailler le travail réalisé.

Enfin, une troisième partie fera office de conclusion dans laquelle on s'efforcera de faire le point sur la réalisation du projet par rapport aux objectifs fixés par le cahier des charges.

# Sujet détaillé

## 1) Objectifs

L’application doit récupérer en entrée un texte pour ensuite l’analyser et afficher la classification de chaque mot en fonction d’une description déjà définie dans un domaine choisi. L’analyse des mots permettra ensuite de vérifier la sémantique des phrases.

L’interface Web de l’application doit être conviviale, ergonomique et facile d’utilisation.

## 2) Domaine de classification

Le choix du domaine pour la classification des mots