

Domaine	Mathématiques et Informatiques	
Chef de projet	Dardouri Lakhdar	Dardouri.Lakhdar@univ-oeb.dz
Membres	Guerram Taha	
	Tennachi Abd-ed-Daim	
	Boussaha Karima	
	Bendriss Mohamed	

Intitulé
Text Mining et traitement automatique des langues naturelles : application à la langue Arabe

Problématique
<p>Le Traitement Automatique des Langues Naturelles (TALN) est le traitement automatique des données linguistiques (textuelles ou orales) exprimées dans une langue naturelle. C'est une discipline à la frontière de la linguistique, de l'informatique qui consiste à l'application de programmes et techniques informatiques à tous les aspects du langage humain, c'est pour cette raison que le TALN est aussi connu sous l'appellation linguistique informatique. Parmi ses applications, on peut citer à titre indicatif, l'analyse automatique des textes, la compréhension automatique des textes, résumé automatiques des textes et la traduction automatique.</p> <p>Le Text Mining (TM) ou « Fouille de textes » est l'extraction automatique de connaissances potentiellement utiles et cachées dans des sources de données textuelles et volumineuses.</p> <p>Le thème de ce projet de recherche est le concours des deux branches de l'IA, TALN et TM appliqués à la langue Arabe. Relativement aux autres langues naturelles, la langue Arabe présente bien des particularités telles la richesse morphologique et sémantique, la voyellation (TASHKIL), absence de la notion de lettre minuscule et majuscule, la dualité, etc.... Ces particularités doivent être prises en compte et traduites sur le plan algorithmique. L'objectif visé par le présent projet consiste à faire interférer les deux disciplines DM et TALN de manière à pouvoir extraire à partir des textes arabes des connaissances cachées et utiles. Dans le cadre de ce projet, les applications visées sont la classification des documents arabes, la génération automatique d'ontologies et l'analyse des opinions électroniques (blogs en Arabe), relatives à un évènement ou à un produit ou équipement quelconque.</p>