

RAPPORT DE SYNTHÈSE D'ACTIVITÉ DE RECHERCHE

PROJET CNEPRU

Semestre : 1er **Année :** 2016

1. Informations sur l'équipe de recherche

NOM	Prénom	Grade de chercheur	Qualité
SAYOUD	Halim	Directeur de recherche	Chef de projet
OUAMOUR	Siham	Maître de recherche	Membre
KHENNOUF	Salah	Attaché de recherche	Membre
BOURIB	Samira	Attachée de recherche	Membre
FODIL	Leila	Doctorante	Membre

Laboratoire : LCPTS

Département : Telecom

Faculté : Electronique et Informatique

Code du projet : S*00220130032

Intitulé : Classification et Exploration Linguistique de Données Textuelles

2. Bilan scientifique

Le projet de recherche a démarré en 2014, il s'intéresse à la fouille textuelle de données, la technologie de l'information et la catégorisation automatique de données textuelles.

Des expériences de classification de documents textuels et de documents audiovisuels, desquels on a extrait la parole transcrite, ont été entreprises. Les critères de catégorisation sont : par thème, par auteur et par langue.

Les résultats obtenus sur la parole pure étaient satisfaisants, tandis que les résultats de la parole transcrite étaient non convaincants.

Durant cette année, on a pu assurer 3 conférences internationales, plusieurs PFE de Master et une publication dans un journal international.

Nous avons aussi conçu un corpus textuel stylo-métrique, que nous avons appelé: OCR5P.

Par ailleurs, nous avons implanté des algorithmes d'identification de langue, d'auteur et de thème que l'on a proposé et soumis à des conférences internationales. Ceci a permis d'avoir plusieurs articles de conférences acceptés : soit 3 conférences -A-.

De plus, nous avons fait soutenir un thésard en Doctorat dans ce thème de recherche : Mr Abainia Kheireddine, avec la mention très honorable.

Finalement, un stage de courte durée a été effectué au sein de l'université de Wolverhampton où on a pu échanger un certain nombre d'idées et d'informations importantes concernant ce domaine de recherche.

3. Travaux réalisés (Publications, Communications, PFE,...)

-Bases de Données conçues :

-Mise au point d'un corpus textuel : Base de données des 5 philosophes brutes que nous avons appelée OCR5P.

-SOUTENANCE DE DOCTORAT :

Etudiant : Abainia Kheireddine, Doctorat TTI, FEI, USTHB.

Date de soutenance : 08.06.2016, à l'USTHB, Mention Très honorable.

-ENCADREMENT D'UN MASTER :

CONTRIBUTION A LA DISCRIMINATION DES LOCUTEURS DANS DES MILIEUX DEFAVORABLES BASEE SUR LA FUSION DES CARACTERISTIQUES. Etudiant: Meftah Ibrahim MEBKHOUT, Soutenu le 06 juin 2016.

-ENCADREMENT D'UN MASTER :

ETUDE ET ANALYSE DE L'EFFET D'ACQUISITION OPTIQUE DES DOCUMENTS TEXTUELS SUR L'ATTRIBUTION D'AUTEURS. Présenté par BENZERROUG Hocine, Soutenu le 06 juin 2016.

-PUBLICATION DANS UNE REVUE INTERNATIONALE :

Journal: Information Processing & Management. Elsevier.

Title: Effective Language Identification of Forum Texts Based on Statistical Approaches.

References: Elsevier, Information Processing & Management, vol.52:4, 2016. pp: 491-512, July 2016. doi:10.1016/j.ipm.2015.12.003.

-ARTICLES ACCEPTEES ET PRESENTES DANS DES CONFERENCES :

-ICISP-2016- Quebec, Canada, May 30- June 01, 2016.

Title: Speaker Discrimination Using Several Classifiers and a Relativistic Speaker Characterization.

International Conference on Image and Signal Processing, ICISP'2016, Quebec Canada.

Book:Image and Signal Processing. Chapter No: 21. Chapter DOI:10.1007/978-3-319-33618-3_21. pp 203-212.

-ICISP-2016- Quebec, Canada, May 30- June 01, 2016.

Title: Speaker Discrimination Based on a Fusion between Neural and Statistical Classifiers.

International Conference on Image and Signal Processing, ICISP'2016, Quebec Canada.

Book:Image and Signal Processing. Chapter No: 21. Chapter DOI:10.1007/978-3-319-33618-3_22. pp 213-221

-ICCSAMA-2016- Vienna, Austria, May 02-03, 2016.

Title: Effect of the Text Size on Stylometry -Application on Arabic Religious Texts. International Conference on Computer Science Applied Mathematics and Applications, ICCSAMA'2016. pp 215-228.

Date et signature du chef de projet