

BILAN ANNUEL

PROJET CNEPRU

Année : 2015

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
FODIL	Leila	Doctorante et Attaché de recherche	Membre
BOURIB	Samira	Attachée de recherche	Membre

Laboratoire : LCPTS

Département : Telecom

Faculté : Electronique et Informatique

Code du projet : S*00220130032

Intitulé :

Classification et Exploration Linguistique de Données Textuelles

Bilan scientifique

Le projet a démarré en janvier 2014, il est impliqué aux domaines de recherche suivants : fouille textuelle de données, technologie de l'information et catégorisation automatique de données textuelles.

Auparavant, nous avons conçu deux corpus textuels stylométriques, que nous avons appelé:

- *A4P database* ou Authorship Attribution of Ancient Arabic Philosophers (avec tentative de bruitage);

- et *StyloSpeech database* : Stylometric based and Speech based Identification (projet non achevé).

Durant cette année, nous avons implanté des méthodes de reconnaissance automatique de thème (Topic Recognition) que l'on a proposé et soumis à des conférences internationales. Ce projet de recherche a permis d'avoir 2 articles acceptés et présentés à des conférences :

- un article présenté à ICTRC'2015, Mai 2015 et un article de conference accepté et présenté à la conference TIR'2015, September 2015.
- De plus, nous avons suivi et encadré plusieurs thésards dans ce thème de recherche : 4 thésards en Doctorat et 1 thésards en Master.

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

- Bruitage du corpus textuel : - A4P database : Authorship Attribution of Ancient Arabic Philosophers et étude de robustesse.
- Mise au point d'un corpus textuel/parole : StyloSpeech : projet non achevé.
- ICTRC-2015 - Abu-Dhabi, UAE, May 17-19, 2015.
Title: Topic Identification of Arabic Noisy Texts Based on KNN. International Conference on Information and Communication Technology Research (ICTRC'2015), May 17-

19, 2015, pp 89-92. <http://events.kustar.ac.ae/ictrc2015/>

- TIR-2015- Valencia, Spain (with DEXA 2015), September 1 (1-4), 2015.

Title: Topic Identification of Noisy Arabic Texts Using Graph Approaches. International Workshop on Text-based Information Retrieval TIR/DEXA2015. pp 254-258.

<http://www.uni-weimar.de/medien/webis/events/tir-15/tir15-web/>

- Nous avons suivi et encadré plusieurs thésards dans ce thème de recherche :

- 4 thésards en Doctorat, Université USTHB.
- 1 thésards en Master, Université de Msila (soutenance Juin 2015)

Le chef de projet