

Approche d'annotation automatique des événements

Rim Faiz*, Aymen Elklifi**

* LARODEC, IHEC de Carthage, 2016 Carthage Présidence, Tunisie

Rim.Faiz@ihed.rnu.tn

**LARODEC, ISG de Tunis, 2000 Le Bardo, Tunisie

Aymen_elklifi@yahoo.fr

Résumé. Quotidiennement, plusieurs agences de presse publient des milliers d'articles contenant plusieurs événements de toutes sortes (*politiques, économiques, culturels, etc.*). Les preneurs de décision, se trouvent face à ce grand nombre d'événements dont seulement quelques uns les concernent. Le traitement automatique de tels événements devient de plus en plus nécessaires. Pour cela, nous proposons une approche, qui se base sur l'apprentissage automatique, et qui permet d'annoter les articles de presse pour générer un résumé automatique contenant les principaux événements. Nous avons validé notre approche par le développement du système "AnnotEv".

1 Introduction

Acquérir de la connaissance à partir de textes est une nécessité qui s'est accrue ces vingt dernières années, avec l'essor considérable de la masse de documents disponibles en format électronique, qu'il faut gérer afin d'extraire ou de filtrer les informations pertinentes parmi toutes celles contenues dans ces documents (Faiz, 2006). A titre d'exemple; les événements boursiers sont nombreux et diversifiés. Les experts de la bourse doivent analyser ces événements en un temps relativement raisonnable pour prendre des décisions importantes. Il s'agit, donc, d'annoter les documents présentant des événements pour pouvoir extraire ceux qui sont pertinents. C'est dans ce cadre que s'inscrit notre travail dont l'objectif est de développer une approche qui annote automatiquement ces articles de Presse.

La suite du document est organisée comme suit : nous commençons, dans la section 2, par décrire les principaux systèmes d'annotations existants. Au cours de la section 3, nous présentons notre approche d'annotation, qui a été validée par le système AnnotEv lequel sera présenté et évalué dans la section 4. Enfin, dans la section 5, nous présentons quelques perspectives de notre travail.

2 Présentation de quelques systèmes d'annotation

Plusieurs méthodes et techniques sont utilisées par les systèmes d'annotations dédiés au Web sémantique telles que l'Exploration Contextuelle (Desclés, 1997), les graphes conceptuels (Roussey et al, 2002), les méta-thésaurus (Khelif et al., 2004) et les indicateurs linguistiques (Muller et al., 2004). Nous pouvons citer :

Le système EXCOM (Djaoua et al., 2006) utilise un ensemble d'outils linguistiques qui visent à annoter un document par un ensemble de connaissances aussi bien internes