
Editorial

Les données scientifiques et techniques, qu'elles soient textuelles ou factuelles, formelles ou informelles, constituent des mines d'informations stratégiques aussi bien pour les décideurs (intelligence économique, veille concurrentielle) que pour les chercheurs et les ingénieurs (veille scientifique et technologique).

Cependant, devant la masse croissante d'informations, les organismes ont besoin de systèmes d'aide à l'analyse de plus en plus performants. Ces systèmes doivent offrir des possibilités d'exploration très fine et de représentation synthétique de l'information recueillie et des nouvelles connaissances déduites. En amont, ils doivent assurer la recherche, la collecte, la sélection et le filtrage de l'information électronique disponible dans des bases spécialisées internes, externes et sur internet, ainsi que la prise en compte des données informelles collectées, sur le terrain, par les différents acteurs de la veille. En aval, pour la communication et la restitution des résultats, ils doivent privilégier l'ergonomie dans les fonctions de présentation, d'exploration, de navigation et de synthèse.

L'objectif de ce numéro spécial est de présenter des travaux de recherche et de développement industriel innovants dans le domaine des systèmes de Veille Stratégique Scientifique et Technologique.

Les deux premières contributions de ce numéro spécial s'intéressent à la notion de réseaux interentreprises.

La première contribution correspond à une revue de la littérature relative aux stratégies de coopération en intelligence économique. Elle en étudie les diverses formes en se focalisant sur les réseaux interentreprises. L'auteur considère que les réseaux peuvent s'approprier le processus d'intelligence économique par l'apprentissage relationnel. Cet article décrit ainsi les liens, dans les deux sens, entre l'intelligence économique et les réseaux d'entreprises en mettant l'accent sur le mécanisme de coopération et sur l'apprentissage.

La deuxième contribution étudie les relations entre la structure d'un réseau d'alliances interentreprises et l'innovation en se focalisant sur les mécanismes de formation et d'évolution de réseaux. L'analyse s'appuie sur l'outil Tétralogie intégrant des fonctions d'analyse fondées sur la théorie des graphes. Le réseau représente les entreprises d'un secteur d'activité et les lie selon les alliances existant entre elles. En utilisant certaines propriétés de la théorie des graphes, on peut remarquer certaines particularités des réseaux ainsi constitués et extraire une vision macroscopique d'un secteur d'activité. L'article applique cet outil à la recherche d'un lien entre la structure du réseau et l'innovation. A partir de ces

connaissances sur le réseau, l'entreprise peut alors en tirer profit pour mieux s'adapter et s'organiser en fonction de ses alliés.

Les contributions suivantes se focalisent sur des tâches complémentaires inhérentes au traitement de l'information lorsqu'elle est volumineuse : construction de vocabulaires et annotation sémantique, résumé de textes et classification de documents.

La troisième contribution décrit une plate-forme propriétaire pour la construction semi-automatique de vocabulaires/terminologies, l'enrichissement semi-automatique de bases de connaissance et l'annotation sémantique des documents. Les solutions mises en place se basent sur des techniques et des méthodes de traitement automatique du langage naturel dans le contexte du web sémantique et d'extraction d'information. La plate-forme (ITM) comprend différentes composantes ainsi qu'un lien avec un outil propriétaire d'extraction d'information. Le cadre de comptes rendus de décision de cour de cassation illustre l'article.

La quatrième contribution propose une nouvelle méthode de construction automatique de résumés de textes basée sur l'utilisation d'algorithmes d'ordonnancement et qui utilisent une combinaison de caractéristiques. Ces caractéristiques sont issues de marqueurs linguistiques, d'acronymes et les mots-clés du titre du document, etc. Ces caractéristiques donnent un score à chaque phrase d'un document.

Enfin, la cinquième contribution traite de la classification de courriers électroniques. Après avoir introduit le problème de classification de courriers électroniques, l'auteur présente une méthode basée sur une lemmatisation et un dictionnaire de raccourcis phonétiques des mots. Les auteurs comparent une méthode non supervisée avec une méthode supervisée basée sur les SVM. Ils montrent qu'une approche intégrant un étiquetage par K-means donne de bonnes performances, évitant ainsi l'écueil de l'expertise d'un nombre volumineux de données d'apprentissage.

Josiane Mothe
IRIT, IUFM – Toulouse

COMITÉ DE LECTURE

Henri Briand – Polytech LINA, Université de Nantes
Isabelle Comyn-Wattiau – CNAM et ESSEC, Cergy Pontoise
Taoufiq Dkaki – GRIMM-ISYCOM et IRIT, Toulouse
Henri Dou – CRRM, Université Paul Cézanne, Aix-Marseille
Bernard Dousset – IRIT, Toulouse
Jacques Ducloy – Institut de l'Information Scientifique et Technique, Nancy
Mohand-Saïd Hacid – LIRIS, Université Claude Bernard, Lyon
Olivier Jouve – SPSS, Paris
José Martinez – Atlas-GRIM – LINA et INRIA, Ecole polytechnique de Nantes
Josiane Mothe – IRIT, IUFM, Toulouse
Jean-Marc Petit – LIRIS, INSA, Lyon
Gérard Ramstein – Université de Nantes
Maryse Salles – IRIT, Université Toulouse 1
Florence Sèdes – IRIT, Toulouse

