
Editorial

L'augmentation des volumes d'information disponible sous format électronique et l'importance de l'information dans la société actuelle conduit au besoin d'une nouvelle génération d'outils et de méthodes. Ces outils doivent permettre à un utilisateur d'accéder à l'information dont il a réellement besoin. L'accès à l'information brute telle que le permettent les systèmes de recherche d'information ne correspond pas au seul type de besoin en information que peut avoir un utilisateur. En effet, la réponse à un besoin d'information peut ne pas se trouver dans un document mais nécessiter une variété de ressources. L'utilisateur peut en effet être intéressé par une information plus élaborée : connaissance, liens entre informations, signaux faibles, visualisations avancées, etc. qui nécessite l'analyse d'ensembles de documents. La masse d'informations, leur caractère volumineux, distribué et hétérogène, le profil de l'utilisateur sont autant de problèmes à résoudre pour offrir des mécanismes permettant aux utilisateurs d'accéder à l'information qui leur est vraiment utile.

L'objectif de ce numéro spécial est de publier des travaux dans le domaine de la recherche, l'extraction et l'exploration d'information répondant à ces problématiques.

La première contribution montre comment les ontologies peuvent aider à la recherche et l'exploration d'information. Les ontologies semblent en effet un des moyens pertinents pour rendre les systèmes de recherche et d'exploration d'information plus performants grâce à la manipulation de connaissances sémantiques. Le web sémantique en est un exemple. Dans cet article, l'auteur présente une application basée sur deux types d'ontologie : une ontologie propre à la tâche à réaliser (ici l'exploration d'information ou la veille scientifique) et une ontologie propre au domaine étudié (ici l'astronomie). L'association de ces ontologies au corpus ainsi qu'une interface permettant de les mettre à profit sont présentées dans cet article.

La deuxième contribution propose une méthodologie d'analyse de résultats obtenus par différents systèmes dans une tâche de recherche d'information. Ainsi, des outils d'analyse de données sont utilisés pour analyser finement les résultats obtenus et pour dégager des groupes de requêtes ou de systèmes ayant les mêmes comportements. Cette analyse est un premier pas vers la fusion intelligente de systèmes de recherche d'information, c'est-à-dire l'utilisation conjointe de systèmes complémentaires, choisis en fonction des contextes.

La troisième contribution s'intéresse aux données médicales cliniques. Généralement les données de ce type proviennent de différentes sources hétérogènes et distribuées. Cet article présente le système HealthObs qui a pour objectif de découvrir des associations intéressantes à partir de données cliniques. Ce système offre différents services comme l'homogénéisation sémantique de sources hétérogènes au travers de modèles de données et d'ontologies, l'aide à la formulation de requêtes, l'exploration d'information et la visualisation pour l'interprétation des résultats. Un cas d'étude illustre l'utilisation et l'intérêt d'un tel système. Il s'agit de découvrir les relations potentiellement intéressantes entre des résultats de tests biochimiques anormaux et les affections du système circulatoire.

La quatrième contribution s'intéresse aux données vidéo qui occupent une part de plus en plus importante des sources d'information. Les sources sont diverses (caméras de surveillance, domaine médical, etc.). La recherche d'éléments dans des vidéos fait l'objet de nombreuses études. Par exemple, le programme d'évaluation TREC (Text Retrieval Conference, trec.nist.gov) propose cette tâche pour la cinquième année consécutive. L'article présenté ici s'intéresse à une des étapes préliminaires à la recherche par le contenu des vidéos qui est leur segmentation en unités sémantiques. Une méthode originale de détection des scènes est ainsi proposée. Elle repose sur l'utilisation des algorithmes génétiques pour explorer les données et apprendre automatiquement des règles de détection en fonction du type de vidéos traitées.

La cinquième contribution s'intéresse au paradigme P2P comme base au développement de services variés pour la recherche et l'accès à l'information. Plus spécifiquement, cet article s'intéresse au cas des images volumineuses qui doivent être archivées à long terme et sur lesquelles doivent être ensuite appliquées des méthodes d'exploration de données. L'astronomie, la médecine, les arts sont producteurs de ce type de sources. Dans ce type de cas, le niveau de compression ou d'agrégation et de résolution associé aux images est crucial. L'approche proposée dans cet article a été validée au travers d'un prototype de service web développé dans l'environnement JXTA.

Enfin, la sixième contribution propose une couche interface qui vise à intégrer des sources hétérogènes de par leur contenu au travers d'un espace d'information unifié. La prise en compte du contexte est ici importante et les techniques de minage d'information jouent un rôle clef.

Josiane Mothe
IRIT, Toulouse

COMITE DE LECTURE

Maha Abdallah, Laboratoire d'Informatique de Paris VI

Marie-France Bruandet, CLIPS-IMAG, Grenoble

Max Chevalier, LGC/Institut de Recherche en Informatique de Toulouse

Jérôme Darmont, Laboratoire E.R.I.C, Lyon

Jean-Pierre Guedon, Nantes

Mohand-Said Hacid, LIRIS, Université Lyon 1

Frédérique Laforest, INSA, Lyon

Mounia Lalmas, Queen Mary, University of London, Grande Bretagne

Jacques Lemaître, Université de Toulon et du Var

Hervé Martin, LSR-IMAG, Grenoble

Josiane Mothe, Institut de Recherche en Informatique de Toulouse

Olivier Nerot, Cluster & GRID groupe LRI,

Alan Smeaton, Dublin City University, Irlande

