

ÉDITORIAL

Dans un monde de plus en plus numérique, les données sont au cœur de nos systèmes et de nos environnements quotidiens. Elles ne sont plus confinées dans les « applications de gestion », et se retrouvent à l'échelle des applications du web, des réseaux sociaux, des applications scientifiques ou des réseaux de capteurs, pour n'en mentionner que quelques-unes. Ce domaine de recherche, communément regroupé sous le terme de *bases de données*, est foisonnant et touche de nombreux secteurs d'application, avec un potentiel sociétal qui le rend unique au regard des autres domaines de recherche en informatique.

Dans ce contexte, *Gestion de données incertaines et distribuées* propose des travaux de recherche récents posés par la prolifération des données. Nous avons choisi de focaliser ce numéro sur deux aspects de la gestion de données, à savoir l'incertitude et la distribution.

L'incertitude est associée à de nombreuses données, que ce soit pour représenter le niveau de confiance que l'on peut avoir sur la donnée elle-même ou sur toute autre suspicion liée à sa qualité. Deux articles sont associés à cet aspect.

– L'article de Mouhamadou Lamine Ba, Talel Abdesslem et Pierre Senellart porte sur la gestion de versions incertaines de documents XML. Ce problème est important à l'échelle du Web lors de l'édition collaborative de documents par des contributeurs divers et variés (e.g. fiabilité, réputation sur le domaine). La gestion des versions est donc cruciale pour tracer l'évolution du contenu partagé et la provenance des contributions.

– L'article de Haizhou Li, François Pinet et Farouk Toumani porte sur les processus métiers manipulant des données imprécises ou incertaines. Dans ce contexte, il étudie précisément le test de simulation pour les processus métier centrés « données probabilistes » en donnant des résultats de complexité.

La distribution est un autre pilier de la gestion de données, inhérente au mode même de production des données : elles peuvent être générées par tout type de dispositif, des plus sophistiqués aux plus simples (du télescope au capteur de température), à l'échelle d'une entreprise ou du web. Trois articles sont associés à cet aspect.

– L'article de Muhammad Faheem et Pierre Senellart porte sur le « Crawl » intelligent et adaptatif d'applications web pour l'archivage du web en intégrant une dimension sémantique (par exemple en prenant en compte le type de CMS utilisé).

Une comparaison avec les robots web traditionnels montre très bien l'intérêt de cette approche.

– L'article de Dario Colazzo, François Goasdoué, Ioana Manolescu et Alexandra Roatis porte sur l'analyse de données issues du web sémantique (RDF). Il présente un cadre formel pour l'analyse multidimensionnelle de données RDF, à l'instar des approches OLAP dans le cadre relationnel.

– L'article de Quoc Cuong To, Benjamin Nguyen et Philippe Pucheral porte sur l'exécution sécurisée de requêtes avec agrégats sur des données distribuées. Il apporte une contribution effective à la gestion de données personnelles dans des architectures décentralisées en visant à rendre le contrôle des données à leur propriétaire.

Ce numéro regroupe des contributions étendues d'articles présentés au 29^e congrès BDA qui s'est déroulé à Nantes du 25 au 28 octobre 2013.

Nous remercions les auteurs pour leur contribution, ainsi que les membres du comité de lecture pour leur participation active dans la relecture des articles de ce numéro.

JEAN-MARC PETIT
INSA Lyon
LIRIS (UMR 5205 CNRS)

COMITÉ DE LECTURE

Christophe Bobineau – LIG, INP Grenoble
Dario Colazzo – LRI, Univ Paris Sud, Orsay
Christine Collet – LIG, INP Grenoble
Jerome Darmont – ERIC, Univ Lyon 2
Bruno Defude – SAMOVAR, TSP Evry
Anne Doucet – LIP6, UPMC, Paris
Walid Gaaloul – SAMOVAR, TSP Evry
Daniela Gregory – LAMSADE, Paris-Dauphine
David Gross-Amblard – IRISA, Univ Rennes 1
Melanie Herschel – LRI, Univ Paris Sud, Orsay
Hélène Jaudoin – IRISA, Univ Rennes 1
Frédérique Laforest – LT2C, Saint-Etienne
Florence Sedes – IRIT, Toulouse
Genoveva Vargas Solar – LIG, CNRS