

Le W3C donne un coup d'accélérateur au Web sémantique avec Sparql

Peter Judge, publié le 21 janvier 2008, ZDNet UK

Le langage Sparql, pilier du Web sémantique, a été validé par le World Wide Web Consortium. Selon ses auteurs, il est capable d'interroger des sources de données multiples, même du Web 2.0, grâce à des requêtes basées sur des expressions sophistiquées.

Le concept de « Web sémantique » progresse avec la [validation](#) des spécifications techniques du langage de requête Sparql par le W3C (World Wide Web Consortium), qui l'a créé. Le consortium vient de publier une recommandation officielle, invitant ainsi l'industrie informatique à l'implémenter dans ses produits.

Sparql est un des piliers du [web sémantique](#), qui veut que toute information est universelle et peut être facilement récupérée et exploitée, quel que soit le format du fichier où elle se trouve et sa langue de publication. Comment ? Grâce à des métadonnées spécifiques qui utilisent plusieurs langages, dont Sparql, et non pas le langage naturel - les différentes langues du monde.

Interrogé plusieurs sources avec une même requête

Concrètement Sparql exprime des requêtes en utilisant des expressions très sophistiquées, qui l'autorisent à interroger aussi les nouvelles sources de données du Web 2.0. « Sparql permet de chercher des informations stockées dans des bases de données et autres sources, où qu'elles se trouvent », selon Tim Berners-Lee, directeur du W3C et créateur du World Wide Web.

Les langages de requête traditionnels, tels que SQL, ont été conçus pour accéder à une seule source de données. Mais ils donnent de mauvais résultats lorsqu'il s'agit de chercher des informations à partir de sources multiples. Sparql, lui, est capable d'interroger plusieurs sources en même temps pour une seule et même requête, et ensuite de combiner les résultats pour fournir une réponse plus riche et précise.

« Parce que Sparql est indépendant de tout format de base de données, il peut tirer parti des données de la déferlante Web 2.0 et la mêler à d'autres ressources sémantiques du web », explique le W3C dans un communiqué.

« Vouloir exploiter le web sémantique sans Sparql revient à tenter d'utiliser une base de données relationnelle sans SQL », indique Tim Berners Lee, à titre de comparaison.

Sparql fonctionne de pair avec les technologies RDF (Resource Description Framework) et WSDL (Web Services Description Language), déjà validées par le W3C.

Déjà des projets utilisant Sparql

Il a déjà reçu le soutien de nombreux industriels influents, notamment Hewlett-Packard qui l'a intégré à ses produits. HP annonce ainsi la sortie de SDB, un système de base de données basé sur Sparql pour son produit de *framework* open source Jena. « Les clients de HP vont bénéficier d'une meilleure utilisation des informations grâce aux technologies de web sémantique », a affirmé Jean-Luc Chatelain, directeur des technologies de l'activité Software Information Management de HP.

SDB est l'une des quatorze implémentations connues à ce jour de Sparql. Parmi les autres, on peut citer celles de l'italien Asemanics pour trois gros chantiers.

Le premier est la création d'un moteur de recherche d'images satellite pour le compte de l'Agence spatiale européenne (ESA). Le deuxième est un moteur, développé pour les autorités maritimes néerlandaises, qui permet d'exploiter les données des marées recensées depuis 400 ans. Le dernier est un agrégateur de données pour les archives de la BBC.