

# Spinning the semantic web: bringing the world wide web to its full potential par Dieter Fensel, James Hendler, Henry Lieberman and Wolfgang Wahlster

Jérôme Euzenat, Amedeo Napoli

▶ **To cite this version:**

Jérôme Euzenat, Amedeo Napoli. Spinning the semantic web: bringing the world wide web to its full potential par Dieter Fensel, James Hendler, Henry Lieberman and Wolfgang Wahlster. Bulletin de l'Association Française pour l'Intelligence Artificielle, 2003, 56-57, pp.18-21. hal-00906623

**HAL Id: hal-00906623**

**https://inria.hal.science/hal-00906623**

Submitted on 20 Nov 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Spinning the semantic web: bringing the world wide web to its full potential

***Dieter Fensel, James Hendler, Henry Lieberman and Wolfgang Wahlster (eds.)***

*The MIT press, Cambridge (MA US),*
*479 pages, 2003 (0-262-06232-1, 40 USD)*

*Recension par Jérôme Euzenat (INRIA Rhône-Alpes) et Amedeo Napoli (LORIA)*

**Introduction**

Le web sémantique est un sujet d'intérêt pour plusieurs communautés et pour les utilisateurs du web en général. Il est parfois considéré comme la prochaine génération du web, un web que les machines peuvent appréhender pour aider leurs utilisateurs. Le web sémantique est né à la fin du siècle dernier en tant que sujet de recherches et d'expérimentations.

Il n'y a eu que peu de livres, plutôt focalisés sur la recherche et la programmation du web sémantique. Ce dernier livre est né d'un séminaire Daghstul qui s'est tenu en 2000 au tout début du développement du web sémantique. Il rassemble les contributions de nombre des principaux acteurs du démarrage du web sémantique. Le volume de 479 pages est un bel objet comme on peut l'espérer d'un ouvrage des presses du MIT. C'est une sorte de « handbook » sur les premiers pas du web sémantique. Il comprend quelques chapitres généraux sur le sujet ainsi que des chapitres retraçant des expériences et choix de conception de systèmes. Les travaux décrits ont eu lieu juste avant l'avènement de OWL, si bien que cet ouvrage peut être considéré comme un point de départ et de référence pour comprendre les problèmes de conception du web sémantique à venir.

Du point de vue de la forme, il n'y a pas de formalisation dans cet ouvrage (il n'est pas très technique à l'exception de quelques exemples de code). L'ouvrage se concentre plutôt sur la description de systèmes, de langages et de leurs choix de conception.

Alors que le volume est organisé en parties que nous reprenons ci-dessous, nous avons trouvé qu'il contenait des chapitres plutôt centrés sur les problèmes posés et d'autres offrant des solutions. Bien entendu, cette séparation n'est pas aussi tranchée : présenter une solution requiert la présentation du problème résolu et les auteurs insistant sur un problème offrent souvent des pistes pour le résoudre. Mais les chapitres de ce livre se laissent facilement classer ainsi par leur ton et leur style. Bien entendu, cette classification sommaire n'est pas indépendante du plan de l'ouvrage. Ainsi, tous les chapitres de la partie « langage » proposent des solutions à certains problèmes. Ce qui compte ici est qu'à la fois les problèmes et les solutions contribuent à éclairer les développements du futur web sémantique.

Le livre commence par une motivation techno-philosophique du web sémantique par Tim Berners-Lee, suivie d'une introduction extensive des éditeurs : Dieter Fensel, James Hendler, Henry Lieberman et Wolfgang Wahlster. Elle présente les applications que le web sémantique est destiné à faciliter, les technologies nécessaires pour cela et une introduction à l'ouvrage lui-même. Berners-Lee se concentre sur l'information et les liens alors que les éditeurs se concentrent plus sur les applications et les ressources : ils présentent l'infrastructure du web sémantique comme un ensemble de langages, d'ontologies, de manières d'accéder à l'information et d'applications utilisant ces trois ingrédients. Ceci est globalement aligné avec la structure du livre que nous suivons aussi.

**I. Langages**

Afin d'être appréhendable par les machines, les pages ou les ressources du web sémantique doivent être exprimées dans un langage conçu à cet effet. Un tel langage doit être un hybride entre les langages de représentation de connaissance provenant de l'intelligence artificielle et les langages de marquage provenant de l'ingénierie documentaire. La première partie de l'ouvrage a pour but de présenter les langages développés dans une perspective voisine de celle du web sémantique. Il rassemble des exposés sur des langages (SHOE, OIL, XMLS, RDF, RDFS, UPML) qui ne sont pas toujours facilement accessibles. C'est donc une très bonne source pour comprendre ces langages et les principes de leur conception.

Le chapitre 2 décrit SHOE et les développements associés. SHOE est fondé sur HTML et peut être considéré comme un précurseur du web sémantique. Ce chapitre soulève les premières questions sur la particularité d'utiliser un langage de représentation de connaissance dans un environnement distribué : interopérabilité, évolution et passage à l'échelle.

DAML-ONT est décrit dans le chapitre 3. Il s'agit du langage promu par le programme DAML de la DARPA et peut être vu comme un langage de frames au-dessus du langage RDF-Schema.

Le chapitre 4 introduit successivement les langages XML-Schema, RDF-Schema et OIL. C'est un exposé très intéressant de ces langages et de leurs mérites respectifs. Le chapitre donne quelques détails sur des points spécifiques comme la manière d'imbriquer OIL (et par conséquent OWL) dans RDF-Schema.

Le chapitre 5 évoque les services web au travers d'UPML (Unified Problem solving Method development Language) et de iBrow. Le premier permet de décrire des méthodes de résolution de problèmes alors que le second permet de résoudre un problème en trouvant les méthodes adéquates distribuées sur le web et en les enchaînant de la façon appropriée. Ainsi, UPML peut être vu comme un précurseur des langages de description de services web actuels.

Le chapitre 6 propose différentes définitions de ce qui est appelé une ontologie et quels types d'applications peuvent en tirer parti. C'est plutôt un chapitre introductif, explorant la jungle des définitions, qui peut être lu en préalable aux chapitres de cette partie.

Ces chapitres seraient plus faciles à lire s'ils avaient partagé un exemple commun. Mais c'est une tâche difficile dans les ouvrages à auteurs multiples. Ils mettent très bien en évidence le besoin d'un pouvoir d'expression accru d'un langage à un autre (spécialement le chapitre 4 consacré à OIL). Ils concernent des langages antérieurs à DAML+OIL et OWL ce qui permet de comprendre la genèse de ces derniers.

**II. Support**

La seconde partie est intitulée « support par la connaissance (Knowledge support) » et présente des outils et architectures qui tirent parti des ressources que l'on devrait trouver sur un web sémantique. Il contient deux chapitres présentant des solutions : le chapitre 7 présente Sesame, une architecture pour le stockage et l'interrogation de RDF, et le chapitre 11 présente SEAL, une architecture, méthodologie et ensemble d'outils pour construire des « sites web sémantiques ». Les autres chapitres, bien qu'ils puissent proposer des solutions, sont plutôt consacrés à la mise en évidence d'un problème auquel le développement d'un web sémantique devra faire face : approcher le sens des ressources en fonction des tâches à accomplir (chapitre 8), construire des ontologies avec leur contexte de sorte de pouvoir les réinterpréter (chapitre 9) et modéliser les relations entre ontologies de manière à assurer l'interopérabilité (chapitre 10).

Sesame, présenté au chapitre 7, est une architecture pour le stockage efficace et la recherche d'une importante quantité de métadonnées. Le chapitre inclus une rapide présentation des langages RDF et RDFS ainsi

que du langage de requêtes RQL (une adaptation de OQL pour RDF). Il mène une discussion intéressante sur les différents niveaux (syntaxique, structurel et sémantique) de réponse à une requête. Les auteurs proposent d'utiliser un langage de requête du type de RQL pour interpréter les requêtes au niveau sémantique. Une expérience concrète avec le SGBD PostGres est présentée.

Le chapitre 8 est plus orienté vers l'industrie. Il plaide en faveur de l'interprétation des ressources du web en fonction des tâches à accomplir et propose une analyse systématique des questions auxquelles ces ressources devraient répondre en fonction de leur fréquence et de la possibilité de les anticiper ou non. Finalement, le chapitre présente une application d'aide en ligne qui tire parti du web et de ressources indexées par des connaissances, mais qui ne mets pas en évidence l'aspect tâche. Malheureusement, aucune piste n'est donnée pour généraliser une telle application à l'échelle du web.

Le chapitre 9 discute des risques pour l'interopérabilité à utiliser des ontologies à cause de différences de granularité, d'inconsistance ou de modélisations non compatibles. La solution promue consiste à lier les sources de connaissances à leurs sources documentaires. Ceci devrait assurer la mobilité des ontologies. Cette perspective serait mieux adaptée au web sémantique si elle utilisait les techniques du web qui reste encore le contexte idéal pour lier des ressources à leur contexte.

Le chapitre 10 présente les relations comme l'élément fondamental pour aider à l'utilisation de sources de connaissance. En cas de requêtes impliquant plusieurs domaines à la fois, l'interopérabilité entre ontologies peut être un problème. Pour le résoudre, il est proposé de rendre explicite les relations entre ontologies. Ce chapitre propose plusieurs exemples très intéressants dont on peut s'inspirer pour concevoir un système sur le web sémantique prenant en compte des relations complexes entre sources de connaissance.

Le chapitre 11 place les ontologies comme l'élément central pour tous les types de communication sur le web. Il est fondé sur l'expérience d'Ontobroker et repose sur son application à site web sémantique présentant un laboratoire. On appelle ici site web sémantique, un site web engendré à partir de données systématiquement liées à une ontologie. Une discussion sur les ontologies, la manière de les développer et la comparaison d'ontologies est produite. L'ensemble d'outils intégrés SEAL, développé autour d'Ontobroker, est présenté en liaison avec d'importants sujets pour les sites web sémantiques comme l'évaluation des connaissances et la personnalisation de l'accès des utilisateurs. Le chapitre aborde aussi l'intégration de tels sites au sein du web sémantique par le biais de l'exportation de la connaissance en RDF.

Plusieurs chapitres de cette partie étant plutôt orientés vers des problèmes et non des solutions, l'aspect support de cette section n'est pas très adapté. Mais les discussions sur les problèmes que devra affronter le web sémantique et quelques principes généraux pour les résoudre restent stimulantes.

**III. Dynamique**

La dernière partie est consacrée aux aspects dynamiques et prolonge la partie précédente qui insistait déjà sur l'évolution et l'usage des ontologies. Cette dernière partie est plus centrée sur l'usage du web sémantique que sur l'évolution de ses ressources. Elle contient de nouveaux deux chapitres présentant des problèmes pour le web sémantique : la communication opportuniste entre des sources de connaissance hétérogènes et non préparées à communiquer (chapitre 12) et l'intégration de procédures au sein du web sémantique (chapitre 13). Les solutions concernent l'adaptation des documents à des dispositifs particuliers à l'aide de méthodes plutôt syntaxiques (chapitre 14) et un langage permettant l'intégration de description de l'activité de planification dans le web sémantique (chapitre 15).

Dans le chapitre 12, l'informatique ubiquitaire est décrite par de petits appareils connectés sans fil pour répondre aux besoins de la vie de tous les jours. Son but est la découverte et l'utilisation de services sans aide ni intervention humaine. Cependant, l'histoire présentée — la visite d'un musée — peut être interprétée de deux manières antagonistes: d'une part, le web sémantique doit nous aider dans une variété de situations de notre vie de tous les jours ; d'autre part, est-il souhaitable que les machines prennent de telles décisions de manière autonome ? À ce problème connu, aucune réponse n'est apportée.

Le chapitre 13 aborde l'aspect dynamique du web lors de son interaction avec les utilisateurs. Il défend l'inclusion d'un aspect procédural à l'intérieur du web sémantique à l'aide d'un langage à prototype dédié intégré à XML.

Le chapitre 14 décrit un moteur d'adaptation de contenu dont le but est d'adapter des documents en accord avec les possibilités de l'application cliente. Ce chapitre est strictement concerné par les transformations syntaxiques (le mot « sémantique » est utilisé pour dénoter le contexte de la transformation plutôt que son action), mais il montre la voie pour étendre ces techniques aux descriptions sémantiques.

Le chapitre 15 propose le langage INOVA pour représenter les problèmes, les tâches et les plans et le langage OP3 pour représenter les processus de planification. Ce chapitre propose ainsi une méthodologie pour intégrer les planificateurs dans le web sémantique. On peut regretter que ce chapitre ne soit pas plus lié à la présentation d'UPML (chapitre 5).

Cette partie illustre les diverses situations dans lesquelles le web sémantique devra faire preuve d'un comportement dynamique. Cet aspect n'est pas véritablement considéré dans les développements actuels du web sémantique, mais devra jouer un rôle important dès lors que les ressources du web sémantique seront utilisées en contexte.

**Synthèse**

*Spinning the semantic web* est une somme sur les origines et la conception du web sémantique. Du côté négatif, on aurait pu espérer une meilleure coordination entre les chapitres : des références croisées auraient sans doute facilité la lecture et la comparaison des approches. Les nombreuses définitions contradictoires d'« ontologie » et présentations de RDF font quelque peu désordre. Une construction plus intégrée, expliquant au besoin les désaccords, aurait contribué à la lisibilité. Sur un plan historique, une présentation plus approfondie des objectifs et du contexte sous-tendant Ontobroker et le programme DAML auraient permis de compléter le panorama.

À l'exception de ces remarques formelles, il faut remercier les éditeurs pour avoir rassemblé en un seul volume ces contributions témoignant de la genèse du web sémantique. La plupart d'entre-elles n'auraient pas été disponibles autrement.

*Spinning the semantic web* est hautement recommandable pour quiconque veut travailler sur le web sémantique. Il sera certainement la référence sur le web sémantique en train de se faire, à la fois sur les aspects techniques et historiques. Il répond à des questions auxquelles les futurs langages et outils pour le web sémantique répondront « par construction ». Il aidera donc beaucoup d'étudiants à saisir l'esprit derrière le web sémantique.

Ceci appelle une suite dans laquelle les approches présentées ici seront intégrées dans leur présentation, mais surtout dans leur implémentation.