

# Web sémantique

Autor(en): **Sokhn, Maria / Mugellini, Elena**

Objektyp: **Article**

Zeitschrift: **Bulletin.ch : Fachzeitschrift und Verbandsinformationen von Electrosuisse, VSE = revue spécialisée et informations des associations Electrosuisse, AES**

Band (Jahr): **101 (2010)**

Heft 9: **100 Jahre Diskurs zur schweizerischen Energiepolitik = 100 ans de discussion sur la politique énergétique suisse**

PDF erstellt am: **12.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-856123>

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

## **Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

# Web sémantique

## Evolution vers un web intelligent

Après le web de documents, une nouvelle extension se profile, celle du web intelligent caractérisé par les aspects sémantiques, mais aussi d'interopérabilité des données. Le web sémantique permettra aux machines de prendre « conscience » du sens des données qu'elles regroupent. L'exploitation de cette richesse enfouie dans le web transmettra aux machines un semblant de l'intelligence humaine facilitant ainsi la communication homme-machine.

Maria Sokhn, Elena Mugellini

Dans les années 90, après s'être implantés dans les bureaux, les ordinateurs personnels s'installent dans les maisons et se relient entre eux grâce au réseau informatique mondial : l'internet. Les utilisateurs peuvent désormais communiquer et s'informer par le biais des sites sur la toile. Le web de documents est né.

Vers le début du siècle, le nombre d'utilisateurs ainsi que la quantité de données disponibles explosent. Le web se transforme ! Les interfaces s'améliorent, faisant place à des interfaces plus ergonomiques et plus interactives. Le web permet à chacun de construire sa page, de mettre en ligne ses textes, ses images, ses vidéos, ses données. L'utilisateur prend le pouvoir. Le web devient participatif. Le web de données voit le jour.

### Le web sémantique : un web de données

Les différentes évolutions technologiques autour du web ont permis de passer d'un réseau de machines au web, un ensemble de documents accessibles à tous. De nos jours, on tend vers une toile de données reliées entre elles. Relations dont le sens est pour l'instant principalement réservé à l'intelligence humaine (figure 1a). Tout l'enjeu du web de données est d'assigner du sens, compréhensible pour les machines, aux liens entre les contenus, de prendre en compte le contexte de ces derniers et d'interpréter et de donner du sens au contenu des données (figure 1b). On parle ainsi de web sémantique ou web de données.

### Qu'est ce que ça apporte ?

« Le web sémantique n'est pas un web séparé, mais une extension du web actuel dans laquelle l'information est munie d'une signification bien définie permettant aux ordinateurs et aux personnes de mieux travailler en coopération » [1].

En effet, le web de données a pour but d'offrir une interopérabilité entre l'ensemble des données enfouies dans nos ordinateurs. L'ajout de sens et la structuration de l'information augmentent la pertinence des contenus. Leur mise en relation nous permettra d'exploiter au mieux cette mine et par conséquent d'enrichir nos connaissances. Le but ultime étant de pouvoir communiquer avec les ordinateurs comme avec les personnes.

En leur transmettant un semblant de l'intelligence humaine, les ordinateurs comprendront le sens des relations et pourront les utiliser pour offrir de

meilleurs services aux utilisateurs. Le web sémantique apportera par conséquent des améliorations à différents niveaux : amélioration des techniques de recherche, développement des réseaux sociaux, personnalisation des interfaces web, veille automatique par l'intermédiaire de systèmes d'agrégation de contenu, simplification des possibilités de publications personnelles, partage de données multimédia, etc. (figure 2).

### Un des piliers du web sémantique : l'ontologie

Une ontologie est une spécification formelle d'une conceptualisation partagée. Dite appliquée (applied ontology), elle explicite des concepts relatifs à un niveau de connaissance particulier et les règles de leur utilisation dans une application définie. Une ontologie cherche à représenter le sens des concepts et des relations qui les lient. Les concepts sont organisés dans un graphe dont les relations peuvent être des relations sémantiques ou des relations taxonomiques (figure 3). Les ontologies décrivent principalement : les individus (les objets de base), les classes ou concepts (ensembles, collections), les attributs (propriétés, fonctionnalités), les relations (les liens entre les objets), et les événements (changements des attributs ou des relations).

### Aspects techniques

Peu à peu, le web sémantique se dote de nouveaux langages et d'outils de plus

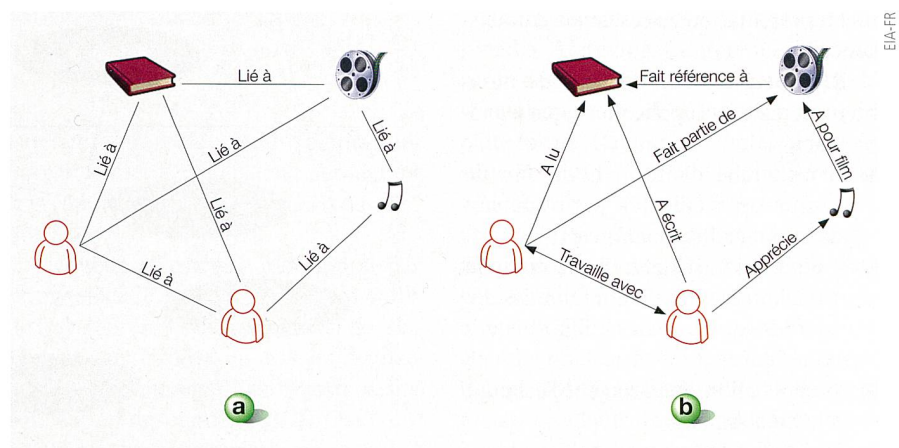
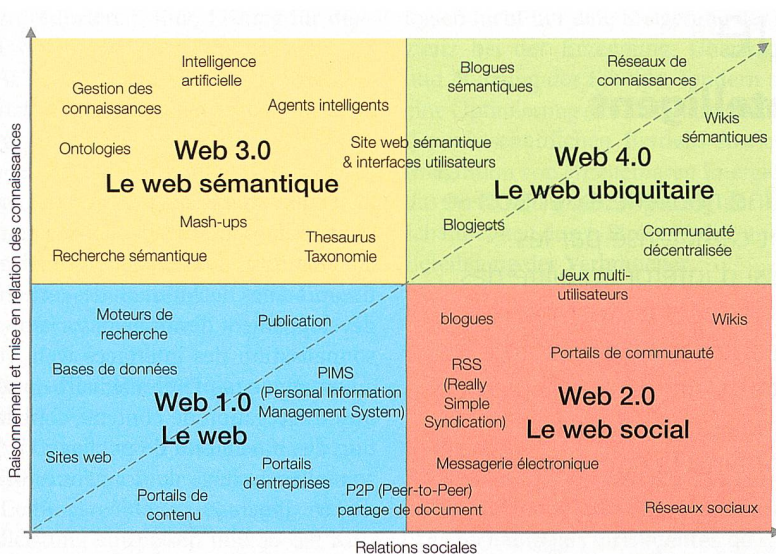


Figure 1 (a) Liens dépourvus de tout sens. (b) Liens dotés d'un sens donnant une structure ainsi qu'une intelligence aux données.





**Figure 2** Evolution du web et de ses applications.

en plus performants. Cette section présente le langage de description, le langage de requête et des outils de travail.

**Langage de description**

Afin de décrire une ontologie, il existe plusieurs formalismes parmi lesquels RDF (Resource Description Framework). RDF est un langage pour la conceptualisation et le traitement des connaissances d'un domaine précis. Il s'appuie essentiellement sur la représentation des objets et des entités qui composent ce domaine en tant que ressources caractérisées par des propriétés et des valeurs attribuées à ces propriétés. Cette association, dite de triplet (Sujet-Prédicat-Objet), permet de décrire les caractéristiques des classes avec le modèle attribut/valeur et lier ainsi les différentes classes (puisque les valeurs des attributs peuvent être à leur tour des liens sémantiques vers d'autres classes). On retrouve ainsi une représentation en réseau de connaissances.

RDF peut être utilisé dans de nombreux domaines d'applications, par exemple dans :

- la recherche d'informations afin de fournir de meilleures performances aux moteurs informatiques.
- la description structurelle du contenu et de l'interaction du contenu des documents web ou des bibliothèques numériques.
- les applications de partage et d'échange de connaissances.

Toutefois, RDF n'est pas suffisamment riche pour décrire les liaisons logiques entre les triplets. C'est là qu'intervient

OWL (Web Ontology Language), un dialecte XML (Extensible Markup Language) basé sur la syntaxe du RDF.

OWL fournit des règles et des contraintes logiques qui seront associées aux triplets définis en RDF. OWL permet de définir des relations telles que la transitivité, l'inverse etc. RDF et OWL, utilisés ensemble, permettent de définir des relations et de décrire des logiques qui transmettront aux ordinateurs une part du raisonnement et de l'intelligence humaine.

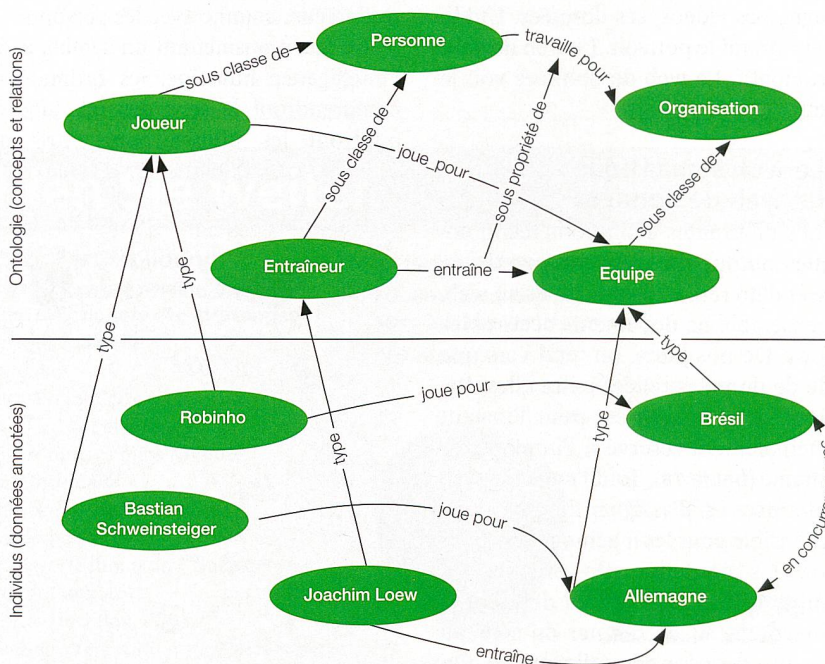
**Langage de requête**

Afin d'exploiter le web sémantique, un langage de requête normalisé par le W3C (World Wide Web Consortium) a été mis en place : SPARQL (SPARQL Protocol and RDF Query Language). SPARQL est parfaitement compatible avec les autres technologies du web sémantique. Il permet d'exploiter le web sémantique à travers :

- son langage de requête, dont la syntaxe est basée sur les triplets.
- le protocole d'accès, qui comme un service web met en œuvre son interface, ses opérations, etc. par des liaisons telles que HTTP et SOAP (Simple Object Access Protocol).
- son langage de présentation des résultats (XML).

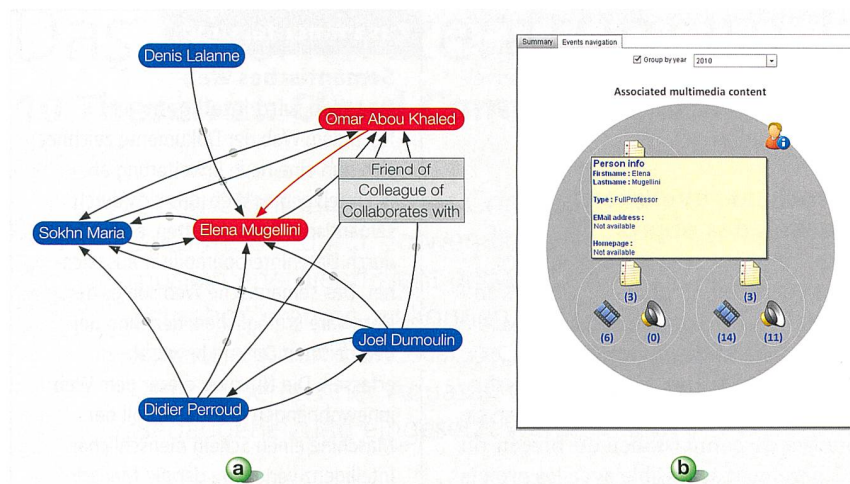
SPARQL emprunte une partie de sa syntaxe au SQL (Structured Query Language), un langage de requête pour les bases de données relationnelles. Il interroge n'importe quel composant d'un triplet qui a la forme « Sujet-Prédicat-Objet » et permet d'exprimer des requêtes interrogatives ou constructives.

Les requêtes interrogatives se distinguent par le mot clé « select » qui permet d'extraire un sous-graphe composé des ressources qui vérifient les conditions des contraintes de la requête. Les requêtes constructives se définissent avec le mot clé « construct », qui permet de construire un nouveau graphe sur la base du graphe



**Figure 3** Exemple d'une ontologie.





**Figure 4** Recherche et visualisation de l'information. (a) Exemple de relations entre les personnes et leurs données. (b) Interface développée dans le cadre du projet d'approfondissement de Joël Dumoulin [5].

tion, de la gestion ou de la manipulation d'une ontologie (classes, propriétés, etc.). Il génère des descriptions en XML qui peuvent être parsées à l'aide des parseurs (analyseurs syntaxiques) spécialisés comme par exemple Jena, un autre exemple d'outil de travail.

Jena est un environnement de travail écrit en Java, qui vise à faciliter le développement des applications dédiées au web sémantique. En plus de manipuler des ressources de type RDF, RDFS (RDF Schema) et OWL, Jena propose un moteur d'inférence qui permet de faire des raisonnements sur les ontologies.

On retrouve également une autre catégorie d'outils, des validateurs comme le validateur RDF du W3C ou le WonderWeb OWL Ontology, qui permettent de s'assurer de la cohérence des concepts et de la syntaxe d'une ontologie, favorisant ainsi l'interopérabilité et l'intégrité des documents.

EIA-FR



**Figure 5** Visualisation de l'information : agrégation de différentes ressources et de divers types d'information.

interrogé. Par exemple, sur un graphe RDF contenant des informations sur une conférence scientifique, une requête « select » permet de trouver les articles d'une personne donnée, alors qu'une requête « construct » ajoutera des relations de « travaille avec », « est collègue de » ou autre, qui ne seraient pas explicitement définies dans le graphe de départ.

**Outils de travail**

La création d'une ontologie est une tâche fastidieuse qui peut être facilitée

par l'utilisation d'un outil d'édition qui offre une interface utilisateur permettant la création des classes et de leurs descriptions.

L'éditeur « Protégé », développé par l'Université de Stanford, est un des outils de travail présents dans le monde du web sémantique. Il offre un environnement complet et puissant au niveau de la conception des modèles RDF/OWL, de plus il est doté d'une interface simple et ergonomique regroupant l'ensemble des notions qui entrent en jeu lors de la créa-

**Potentiel du web sémantique**

Les nouvelles technologies du web sémantique ont permis un essor dans plusieurs domaines, comme par exemple dans le domaine de la recherche et de la visualisation de l'information ainsi que dans l'interopérabilité et l'échange de données.

**Recherche et visualisation de l'information**

Face à l'explosion informationnelle sur le web, le premier défi reste celui de la gestion et de l'exploitation de cette mine d'information. Plus particulièrement, la gestion et l'exploitation des ressources de type multimédia s'avèrent une tâche complexe. En effet, le contenu de ces données, leur contexte ainsi que les relations entre les différentes ressources ne sont pas suffisamment exploités.

Dans ce contexte, SMAC (Smart Multimedia Archiving Conference [2]) tire profit des nouvelles technologies du web sémantique mises à disposition dans la plate-forme Calimera (Conference Advanced Level Information Management and Retrieval [3]) développée par le groupe MISG (Multimedia and Information Systems Group [4]) de l'Ecole d'ingénieurs de Fribourg, en collaboration avec le CERN (Centre européen de recherche nucléaire) et l'Université de Fribourg. SMAC est une plate-forme permettant l'acquisition, l'archivage, l'analyse, la recherche et la consultation d'archives de conférences multimédia. Cet ensemble de modules permet une gestion



complète du cycle de vie d'un événement, depuis sa création (logistique, thématique, participants, etc.), en passant par la recherche sémantique offrant à l'utilisateur une vue sur l'ensemble des ressources et leurs relations (figure 4), jusqu'à sa consultation où, sur une seule interface, sont agrégées des données multimédia parfaitement synchronisées (figure 5).

### Interopérabilité et échange de données

Le modèle actuel de l'administration publique se base sur une collecte et un usage de données segmentées limitant ainsi l'échange d'informations. Le défi imposé par le développement de la cyber-administration est de rendre les différentes entités interopérables afin de garantir une meilleure qualité de service. L'utilisation des technologies du web sémantique dans ce contexte présente un grand potentiel. En effet, l'intégration de ces technologies permettra de faciliter l'échange et la gestion des données en préservant leur sens exact et en les rendant compréhensibles à toutes les personnes, applications et entités concernées.

L'exploitation des technologies sémantiques au service des administrations publiques est l'un des axes stratégiques du pôle de compétence « eGov Center CH » [6], une initiative de collaboration entre partenaires académiques, cantons, communes et secteurs privés actifs dans le domaine IT (technologies de l'information). Cette initiative est coordonnée par l'Ecole d'ingénieurs et d'architectes de Fribourg et parrainée par le Secrétariat d'Etat à l'économie (SECO). Le pôle de compétence « eGov Center CH » vise principalement à appuyer et accompagner les communes et administrations

publiques dans leurs efforts d'harmonisation et d'optimisation des procédures administratives, en exploitant les nouvelles technologies de l'information et de la communication.

### Prochaine évolution: le web des objets

En l'espace de quelques années, le web a fortement évolué. Le web « traditionnel » cède progressivement la place à un web participatif, un web d'échange de données et de connaissance. On assiste de nos jours à la création d'une base de données de connaissance au niveau du web librement accessible avec les projets DBpedia [7] et Linked Data [8]. La prochaine évolution sera certainement celle du web des objets, un web étendu à des objets et à des lieux de notre monde réel. Un web où les utilisateurs ne se connecteront plus via leurs mobiles ou ordinateurs, mais ce sera l'environnement qui sera en permanence connecté.

### Références

- [1] Traduction d'un extrait de T. Berners-Lee, J. Hendler and O. Lassila: The Semantic Web. Scientific American 284 (5), pp. 34-43, 2001.
- [2] Site web de SMAC, <http://smac.hefr.ch/>.
- [3] Site web de Calimera, <http://calimera.project.eia-fr.ch/>.
- [4] Site web du groupe MISG, [www.misg.ch/](http://www.misg.ch/).
- [5] Joël Dumoulin: NAVIR, Navigation et Visualisation de l'Information Recherchée. Master MSE, 2010. <http://joel.dumoulin.home.hefr.ch/NAVIR/Prototype/NAVIR.html>.
- [6] Site web du pôle eGov, [www.itvalley.ch/egovcenter/](http://www.itvalley.ch/egovcenter/).
- [7] Projet DBpedia, <http://dbpedia.org/>.
- [8] Projet Linked data, <http://linkeddata.org/>.
- [9] <http://maria.sokhn.home.hefr.ch/>.
- [10] <http://elena.mugellini.home.hefr.ch/>.

### Informations sur les auteurs

**Maria Sokhn** est ingénieure en informatique et réseaux de la Faculté d'Ingénierie de l'Université Saint-Joseph (ESIB). Elle est titulaire d'un diplôme de master

### Zusammenfassung

#### Semantisches Web

#### Das Web wird intelligent

Nach dem Web der Dokumente zeichnet sich nun eine neue Erweiterung ab: Ein intelligentes Web, das sich durch semantische Eigenschaften, aber auch durch Dateninteroperabilität auszeichnet. Das semantische Web soll es der Hardware ermöglichen, den Sinn der bearbeiteten Daten «bewusst» zu erfassen. Die Nutzung dieser dem Web innewohnenden Ressource soll der Maschine einen Schein menschlicher Intelligenz verleihen, der die Mensch-Maschinen-Kommunikation erleichtern könnte. Fernziel ist die Möglichkeit, mit Computern genauso kommunizieren zu können wie mit Menschen. CHE

spécialisé en production et création multimédia de Telecom ParisTech. Actuellement, elle mène sa thèse de doctorat qui se déroule en collaboration entre Telecom ParisTech et l'Ecole d'ingénieurs et d'architectes de Fribourg (EIA-FR). Elle dirige le centre technologique de cyberadministration « eGov Center CH » à l'EIA-FR [6]. Elle est membre du groupe de compétence sur les systèmes d'information multimédia MISG [4]. Ses principaux axes de recherche se concentrent sur la recherche d'information multimédia, la gestion, la modélisation et la visualisation des connaissances [9]. HES-SO Fribourg, Ecole d'ingénieurs et d'architectes, 1700 Fribourg, maria.sokhn@hefr.ch

**Elena Mugellini** est professeure à l'Ecole d'ingénieurs et d'architectes de Fribourg. Elle est titulaire d'un diplôme de master spécialisé en ingénierie de télécommunications et d'un doctorat en informatique. Elle est responsable du groupe de compétence sur les systèmes d'information multimédia MISG [4]. Ses principaux axes de recherche se concentrent sur les environnements intelligents, l'interaction homme-machine, la recherche d'information et les technologies du web sémantique [10].

HES-SO Fribourg, Ecole d'ingénieurs et d'architectes, 1700 Fribourg, elena.mugellini@hefr.ch

Anzeige

**ECG – PARTNER DER  
ENERGIEWIRTSCHAFT**



**THE ENERGY CONSULTING GROUP**

[www.the-ecgroup.com](http://www.the-ecgroup.com)