|  |
| --- |
| IUT Paris Descartes |
| Exposé d’ACSI |
| Le Web Sémantique |
|  |
| Cyrille PEROIS – John BENCHETRIT – Nadia ZOUINE |
| 05/11/2012 |

|  |
| --- |
|  |

Sommaire

[I. Objectifs du Web sémantique 2](#_Toc340577602)

[1) Un web de données 2](#_Toc340577603)

[2) Défis du web sémantique 3](#_Toc340577604)

[II. Les moyens de mise en œuvre du Web sémantique 4](#_Toc340577605)

[1) L’ontologie, principe de baise du Web sémantique 4](#_Toc340577606)

[2) Les technologies existantes 5](#_Toc340577607)

[a) RDF 5](#_Toc340577608)

[b) RDFS 5](#_Toc340577609)

[c) OWL 5](#_Toc340577610)

[d) SPARQL 6](#_Toc340577611)

[III. Les potentiels du Web sémantique 7](#_Toc340577612)

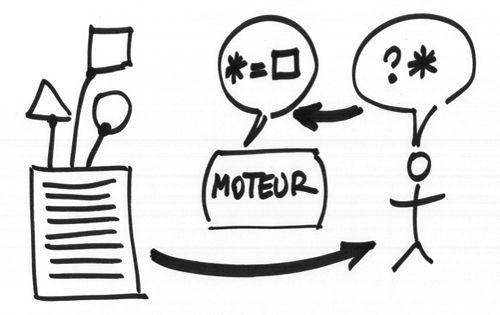
[1) Les avantages et les inconvénients 7](#_Toc340577613)

[2) Le rôle du Web sémantique dans le Web 3.0 ? 8](#_Toc340577614)

# I. Objectifs du Web sémantique

## 1) Un web de données

Le web sémantique est un terme inventé par Tim Berners Lee, co-directeur du web et directeur du W3C, on appelle aussi le web sémantique, web des données, car il s’agit de permettre aux machines de comprendre la sémantique, c’est-à-dire la signification de l’information sur le web. Le but étant de permettre aux machines d’accéder de manière la plus optimale à toutes les informations sur le web, afin que celles-ci puissent effectuer des recherches de manière plus précise que les utilisateurs eux-mêmes. Tim Berners Lee montre que l’hypertexte (Mise en relation de documents par des liens) doit évoluer et permettre aux machines de relier automatiquement des données sur le web aux éléments du monde réel.  
Autrement dit repose sur le fait d’ajouter des compléments d’information à un texte par exemple ce qui va permettre de classifier les informations et de le mettre en relation avec d’autres textes. Ces compléments d’informations s ‘apparentent à des mots clés et s’appellent des META-DONNEES



*38- Le web sémantique s’appuie sur un système de méta-données (données complémentaires)*

Il devient ainsi possible de mettre en relation un texte à un ou plusieurs domaines de connaissances, mais il est possible aussi de détecter l’intention de la pensée contenue dans une phrase. Google se sert déjà de ce mécanisme pour orienter au mieux les utilisateurs. L’avenir potentiel du web sémantique serait la sémantique comportementale, c’est-à-dire guider de manière encore plus forte l’internaute en détectant et analysant même ses arrières pensés, mais aussi en apportant des informations en fonction de la géolocalisation de l’internaute ce qui va permettre de cerner le contexte dans lequel une requête est formulée.

## 2) Défis du web sémantique

Le web sémantique doit répondre à de nombreux défis posés par le web actuel, dont l’immensité, l’imprécision, l’incertitude, l’incohérence et la tromperie.

* **Immensité :** tous les systèmes automatisés qui constituent le web sémantique doivent prendre en compte et gérer de très nombreuses informations.
* **Incertitude :** Un utilisateur peut faire une recherche avec un ensemble de termes précis mais qui va correspondre à plusieurs résultats (par exemple une liste de symptômes avec plusieurs maladies possible). Dans ce cas un raisonnement probabiliste va être appliqué.
* **Tromperie :** Tromperie volontaire de la part du développeur afin de tromper les utilisateurs et de les mener vers des pages n’ayant pas de rapport avec la recherche des utilisateurs

# II. Les moyens de mise en œuvre du Web sémantique

## L’ontologie, principe de baise du Web sémantique

Une ontologie définit les termes employés pour décrire et représenter un domaine de connaissances. En informatique, une ontologie a pour but de représenter les connaissances de notre monde, de façons à ce qu'un ordinateur soit capable de les comprendre et des les manipuler (effectuer des opérations sur celles-ci). Elle permet donc de modéliser un ensemble de connaissances.

Pour représenter ces connaissances, une ontologie met en place différents concepts :

* les individus : les objets de base
* les classes : ensembles d'objets
* les attributs : caractéristiques que les objets possèdent
* les relations : liens entre les objets
* les évènements : changements subis par des attributs ou des relations

[Exemple d’ontologie]

Les "vraies" ontologies sont évidemment plus complexes que celle-ci. Différentes personnes et organismes sont déjà en train d'en mettre en place.

## Les technologies existantes

1. RDF

Acronyme de Resource Description Framework, c'est la base du web sémantique. C'est un modèle de graphe servant à décrire les ressources et les métadonnées associées à celles-ci, permettant leur traitement automatique. Plusieurs syntaxes existent, dont le RDF/XML reposant comme son nom l'indique sur le langage de description XML. Cette syntaxe est considérée comme lourde, c'est pourquoi des syntaxes plus lisibles sont apparues par la suite (exemple : Notation3). La structure fondamentale d'une expression en RDF est une collection de triplets composés d'un sujet, d'un prédicat et d'un objet.

[Schéma sujet ->prédicat -> objet]

1. RDFS

RDF Schéma a pour but de créer des vocabulaires de métadonnées, c'est à dire de donner un sens aux informations stockées sous forme de triplets RDF.

[Un petit exemple]

1. OWL

Web Ontology Language est fonda sur la syntaxe RDF/XML et offre un moyen d'écrire des ontologies pour le web. Il offre aux machines une plus grande capacité d'interprétation du contenu web que RDF et RDFS car il intègre, en plus de la possibilité de décrire des classes et des propriétés, des outils de comparaison des propriétés et des classes. Afin d'éviter qu'OWL soit un outil très complexe, le W3C l'a doté de 3 sous langages :

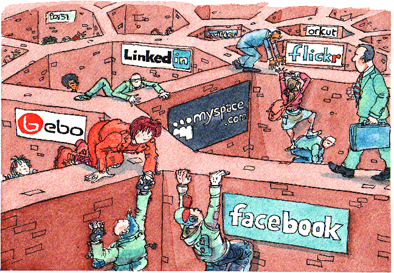
* OWL Lite, qui est le plus simple, destiné aux utilisateurs ayant besoin d'exprimer une hiérarchie de concepts simple
* OWL DL (pour Description Logics), plus complexe que OWL Lite, permet une expressivité plus importante. Il est, comme son nom l'indique fondé sur la logique descriptive, ce qui lui confère une complétude des raisonnements (c'est à dire que toutes les inférences sont calculables), ainsi que leur décidabilité (c'est à dire que leur calcul se fait en une durée limitée)
* OWL Full, qui est la description la plus complexe d'OWL. Elle permet aussi le plus haut niveau d'expressivité. Il convient donc aux situations nécessitant un haut niveau de capacité de description, au gré de la complétude et de la décidabilité des calculs liés à l'ontologie. Il permet aussi d'étendre le vocabulaire par défaut de OWL.

1. SPARQL

SPARQL (pour SPARQL Protocol And Query Language) est l'équivalent du SQL car il permet d'accéder aux données. C'est à dire que grâce à ce standard, il serait possible (en théorie) d'accéder à n'importe quelle donnée du web. Tout comme SQL, il permet d'exprimer des requêtes interrogatives grâce à SELECT et la clause WHERE afin d'extraire d'un graphe RDF un sous-graphe représentant les données recherchées. Il permet aussi d'ajouter des parties à un graphe existant grâce à CONSTRUCT.

# III. Les potentiels du Web sémantique

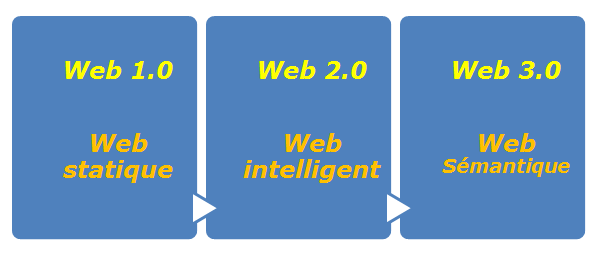
## Les avantages et les inconvénients

Un des avantages du concept de Web sémantique est de pourvoir lier des informations qui restent d’habitude isolées comme par exemple le fait de rapprocher des peintres de la même école et de recenser ses œuvres. De plus, les informations ne seraient plus seulement stockée mais aussi être comprises par les ordinateurs, afin d'apporter à l'utilisateur ce qu'il cherche vraiment, ce qui n’est pas le cas avec le Web actuel qui est considéré comme étant un Web syntaxique.

Le Web sémantique apporte une grande richesse, mais amène des problématiques comme la gestion de la sécurité et de la confidentialité. En effet, les machines pourront elles aussi consulter et interpréter des informations, notamment les données personnelles qui seraient exploitées par plusieurs vocabulaires tels que FOAF ou SIOC pour les sites « sociaux », et tout ceci sans aucune action humaine. De plus, comment allons-nous aider les utilisateurs non spécialistes à créer et à exploiter des ressources dans le cadre des infrastructures et des outils proposés par le Web sémantique ? En d’autre terme, saurons-nous mettre en place des ressources formelles, des ontologies, pour le Web sémantique ou devrons-nous en passer par des ressources moins formelles mais appréhendables par un utilisateur lambda ?

## Le rôle du Web sémantique dans le Web 3.0 ?

Historiquement, le Web 3.0 est une expression qui désigne la troisième étape en cours des transformations majeures dont le [Web](http://fr.wikipedia.org/wiki/Web) fait l'objet depuis son lancement :

* Le Web 1.0 est l'[internet](http://fr.wikipedia.org/wiki/Internet) qui se réfère aux [pages web](http://fr.wikipedia.org/wiki/Pages_web) liées entre elles par des [hyperliens](http://fr.wikipedia.org/wiki/Hyperliens) qui a été créé au début des [années 1990](http://fr.wikipedia.org/wiki/Ann%C3%A9es_1990).
* Le [Web 2.0](http://fr.wikipedia.org/wiki/Web_2.0) est l'[internet liant les personnes](http://fr.wikipedia.org/wiki/Web_social) entre elles, au départ par le [courrier électronique](http://fr.wikipedia.org/wiki/Courrier_%C3%A9lectronique), puis qui s'est généralisé avec le phénomène des [blogs](http://fr.wikipedia.org/wiki/Blog), des [forums](http://fr.wikipedia.org/wiki/Forum_internet) de discussion agrégeant des [communautés](http://fr.wikipedia.org/wiki/Communaut%C3%A9s) autour de [sites internet](http://fr.wikipedia.org/wiki/Sites_internet) et enfin avec les [réseaux sociaux](http://fr.wikipedia.org/wiki/R%C3%A9seaux_sociaux).

Le web 3.0, lui, n'est pas vraiment défini. En fait, l'expression est employée par tous les spécialistes pour expliquer ce que sera selon eux la prochaine étape de développement du web. Les deux thèses dominantes sont de considérer le web 3.0 comme l'[Internet des objets](http://fr.wikipedia.org/wiki/Internet_des_objets)[1](http://fr.wikipedia.org/wiki/Web_3.0#cite_note-0), qui émerge depuis 2008, l'autre thèse dominante est d'en faire le [web sémantique](http://fr.wikipedia.org/wiki/Web_s%C3%A9mantique).