

# Tim Berners Lee, Jim Hendler et Ora Lassila, « Le web sémantique », 2001

---

Traduction par Elisabeth Lacombe et Jo Link-Pezet de l'article de Tim Berners-Lee, James Hendler, Ora Lassila

<http://www.urfist.cict.fr/archive/lettres/lettre28/lettre28-22.html>

**Berners Lee, Tim, James Hendler, and Ora Lassila. « The Semantic Web », *The Scientific American* (2001). <http://www.scientificamerican.com/article.cfm?id=the-semantic-web> (accédé le 6 février 2013).**

La sono hurlait la chanson des Beatles "We can work it out" (Ensemble, on peut y arriver) quand le téléphone sonna. Quand Pete répondit, son téléphone baissa le son en envoyant un message aux autres appareils locaux possédant un contrôleur de volume. Sa sœur Lucie l'appelait depuis le cabinet du médecin : "Maman a besoin de voir un spécialiste et de suivre des séances de thérapie, deux ou trois fois par semaine. Je vais demander à mon agent de nous arranger les rendez-vous". Pete accepta tout de suite la responsabilité de la conduire en voiture à tour de rôle.

Depuis le cabinet du médecin, Lucie donna des instructions à son agent du Web sémantique par le biais de son navigateur. L'agent trouva rapidement l'information concernant le traitement prescrit à Maman à partir de l'agent du médecin, parcourut plusieurs listes de fournisseurs de services, vérifia la couverture d'assurance pour la mère pour trouver un thérapeute dans un rayon de 20 miles de son domicile et prit en considération les taux de satisfaction "excellent" ou "très bon" attribués par des services d'évaluation fiables. Il fallut ensuite essayer de faire coïncider les heures de rendez-vous possibles (fournies par les agents de fournisseurs individuels à travers leur site Web) avec les emplois du temps chargés de Pete et Lucy. (Les mots clés en italique indiquent les termes dont la sémantique ou la signification ont été définies pour l'agent à travers le Web sémantique).

L'agent leur fournit un plan en quelques minutes. Pete ne fut pas d'accord : l'hôpital de l'université se trouvait de l'autre côté de la ville par rapport au domicile de sa mère, et il lui faudrait la ramener à l'heure de pointe. Il fit refaire la recherche par son propre agent en ajoutant des critères de choix plus précis comme le lieu et l'horaire. L'agent de Lucy, qui a une confiance totale dans l'agent de Pete dans le contexte particulier de cette tâche lui a apporté automatiquement de l'aide en lui fournissant des codes d'accès et des raccourcis à partir des données qu'il avait déjà triées.

Un nouveau plan fut présenté instantanément : il y avait une clinique beaucoup plus proche avec des horaires plus matinaux, mais il y avait aussi deux avertissements. D'abord, Pete devrait reprogrammer deux de ses rendez-vous (parmi les moins

importants). Il vérifia de quoi il s'agissait : ce n'était pas un problème. L'autre remarque concernait la liste des compagnies d'assurance qui avait oublié d'inclure ce fournisseur (NDLR : la clinique) dans la liste des thérapeutes médicaux. "Le type de service ainsi que le statut du plan d'assurance ont été vérifiés de manière sûre par d'autres moyens "le rassura l'agent." Détails ?".

Lucy enregistra son accord au moment même où Pete grommelait. "Epargne-moi les détails" et tout fut réglé. (Bien sûr, Pete n'a pu s'empêcher de vouloir connaître ces détails : plus tard ce soir-là, son agent lui expliqua comment il avait découvert ce fournisseur bien qu'il ne fût pas sur la liste.)

## Expression du sens.

Pete et Lucy ont pu utiliser leurs agents pour mener à bien leur recherche non pas grâce à la toile telle qu'elle est aujourd'hui, mais plutôt grâce au Web sémantique de demain. Actuellement, presque tout le contenu du Web est destiné à être lu, il n'est pas fait pour être manipulé de façon intelligente par des programmes informatiques. Les ordinateurs peuvent parcourir les pages du Web pour un traitement de routine ou de mise en page - ici une en-tête, là un lien vers une autre page- mais en général les ordinateurs n'ont pas de méthode fiable pour s'occuper de la sémantique : ce lien mène à la page d'accueil de la clinique médicale Hartman et Strauss, cet autre lien conduit au CV du Dr Hartman.

Le Web sémantique va utiliser la structure pour donner du sens au contenu des pages Web, en créant un environnement où les agents logiciels en parcourant les pages pourront réaliser rapidement des tâches compliquées pour les utilisateurs. Par exemple, un agent arrivant sur la page d'accueil de la clinique saura non seulement que la page a des mots-clés comme "traitement, médecine, physique, thérapie" (de la façon dont cela peut être codé aujourd'hui), mais aussi que le Dr Hartman travaille dans sa clinique les lundi, mercredi et vendredi. Le script fait un classement par date en format aaaa-mm-jj et renvoie des heures de rendez-vous. Il "saura" tout cela sans avoir recours à l'intelligence artificielle comme avec Hal (l'ordinateur de 2001 Odyssée de l'espace) ou avec C-3PO de la Guerre des étoiles. Au contraire, cette sémantique a été encodée dans la page Web au moment où le directeur de la clinique l'a mise en forme en utilisant un logiciel d'écriture de pages Web sémantique comprenant les ressources du site de l'Association de kinésithérapeutes.

Le Web sémantique n'est pas un Web à part, mais une extension du Web courant, dans lequel on donne à une information un sens bien défini pour permettre aux ordinateurs et aux gens de travailler en coopération. Les premiers tissages du Web sémantique dans la structure du Web existant sont déjà faits. Dans un futur proche, ces développements atteindront des fonctionnalités significatives à mesure que les machines vont devenir de plus en plus capables de traiter et "comprendre"les données qu'elles ne font que visualiser pour le moment.

La propriété essentielle de la toile est son universalité. La puissance d'un lien hypertexte est due au fait que "tout peut être lié à tout". La technologie Web cependant ne sait pas faire de différence entre une ébauche gribouillée et un produit achevé, entre une information commerciale et une information universitaire, ou entre les cultures, les langages, les médias, etc.... L'information varie selon de nombreux axes. L'un de ceux-ci est la différence entre l'information produite principalement pour la consommation humaine et celle produite essentiellement pour des machines. A un bout de l'échelle, nous trouvons de tout : des publicités pour la télévision commerciale qui durent cinq secondes à de la poésie. A l'autre bout, nous avons des bases de données, des programmes et des indicateurs. A ce jour, le Web s'est développé plus rapidement comme médium de documents destinés à des gens plutôt qu'en tant que données et information pouvant être traitées de manière automatique.

Comme l'Internet, le Web sémantique sera aussi décentralisé que possible. De tels systèmes engendrent beaucoup d'enthousiasme à tous les niveaux, de l'entreprise à l'utilisateur individuel et procureront des bénéfices difficiles ou impossibles à prédire à l'heure actuelle. La décentralisation requiert des compromis : le Web a dû se débarrasser de l'idéal de compatibilité absolue de toutes ses interconnexions, qui débouchait souvent sur l'insupportable message : "Error 404 :Not Found", mais qui permettait une croissance exponentielle et incontrôlable.

## **La représentation de la connaissance.**

Pour que le Web sémantique fonctionne, les ordinateurs doivent avoir accès à des collections structurées d'informations et d'ensembles de règles d'inférence qu'ils peuvent utiliser pour parvenir à un raisonnement automatisé. Les chercheurs en intelligence artificielle ont étudié ces systèmes longtemps avant que le Web ne se développe. Cette technologie souvent appelée représentation de la connaissance en est actuellement à un stade comparable à celui de l'hypertexte avant le Web. Il est évident que l'idée est bonne, que de très beaux exemples existent, mais cela n'a pas encore changé la face du monde. Cette technologie contient le germe d'importantes applications, elle doit être liée à un système global et unique pour réaliser tout son potentiel.

Les systèmes traditionnels de représentation de la connaissance sont centralisés et nécessitent que chacun partage exactement une définition identique des concepts communs comme "parent" ou "véhicule". Ce contrôle central est étouffant et l'augmentation de la taille et de la portée d'un tel système devient rapidement ingérable.

De surcroît, ces systèmes limitent les questions pouvant être posées (pour que l'ordinateur puisse répondre de manière fiable et à toutes les questions). Le problème rappelle le théorème de Gödel en mathématiques : tout système assez complexe pour être utile renferme des questions auxquelles il est impossible de répondre, ce qui est tout à fait comparable aux versions élaborées du paradoxe de base : "Cette phrase est

fausse". Pour éviter de tels problèmes, les systèmes traditionnels de représentation de la connaissance ont en général un ensemble de règles qui leur sont propres : étroites et idiosyncratiques afin de faire des inférences sur leurs données. Par exemple, un système de généalogie reposant sur une base de données d'arbres généalogiques devrait inclure la règle suivante : "la femme d'un oncle est une tante". Même si les données peuvent être transférées d'un système à une autre, on ne peut généralement pas le faire pour des règles qui existent sous une forme complètement différente.

Les chercheurs du Web sémantique, au contraire, acceptent les paradoxes. Les questions sans réponse sont le prix à payer pour acquérir de la souplesse. Nous construisons un langage de règles aussi significatif que nécessaire pour permettre au Web de raisonner autant qu'on le veut. Cette philosophie ressemble à celle du Web classique : dès le développement du Web, ses détracteurs ont souligné qu'il ne pourrait jamais être une bibliothèque bien organisée, que sans base de données centrale et sans structure arborescente, on ne pourrait jamais être sûr de tout trouver. Ils avaient raison. Mais la puissance d'expression du système a mis à la disposition du public des quantités importantes d'information et les moteurs de recherche (qui auraient paru tout à fait irréalisables il y a 10 ans) permettent de trouver des ressources. Le défi du Web sémantique, cependant, est de fournir un langage qui exprime à la fois des données et des règles pour raisonner sur les données et pour que les règles de n'importe quel système de représentation de la connaissance puisse être exportées sur le Web.

Ajouter de la logique au Web (c'est-à-dire lui donner la possibilité d'utiliser les règles pour faire des inférences, choisir des cours d'action et répondre aux questions), c'est le travail qui attend la communauté du Web sémantique maintenant. Un mélange de décisions mathématiques et techniques complique cette tâche. La logique doit être assez puissante pour décrire les propriétés complexes des objets, mais assez puissante pour ne pas empêcher les agents de résoudre un paradoxe. Heureusement, une bonne partie de l'information que nous voulons donner est entre les lignes et du type " un boulon cruciforme est une sorte de boulon décolleté " difficulté qui se résout en langage réel par un peu plus de mots.

Deux importantes technologies de développement du Web sémantique existent déjà : eXtensible Markup Language (XML) et le Resource Description Framework (RDF). XML permet à chacun de créer ses propres balises -marques cachées comme <code postal> ou <université> qui mettent des annotations sur les pages Web ou les sections de texte d'une page. Les scripts ou les programmes peuvent utiliser ces balises de manière compliquée, mais le programmeur doit savoir dans quel but l'auteur utilise chaque balise.

En bref, XML permet aux utilisateurs d'ajouter une structure arbitraire à leurs documents sans rien dire de la signification des structures [voir XML and the Second-Generation Web, par Jon Bosak et Tim Bray, Scientific American, May 1999].

**LE WEB SEMANTIQUE PERMETTRA AUX MACHINES DE COMPRENDRE LES DOCUMENTS ET LES DONNEES SEMANTIQUES**, mais **PAS** la parole **ET LES ECRITS HUMAINS**. Le sens est donné par RDF, qui le code dans des ensembles de triples, chaque triple jouant le rôle du sujet, du verbe et de l'objet dans une phrase élémentaire. On peut écrire ces triples en

utilisant les balises XML. Avec RDF, dans un document, on part du principe que des choses particulières (des gens, des pages Web, ou quoi que ce soit) ont des propriétés (comme est la sœur de, est l'auteur de) avec certaines valeurs (une autre personne, une autre page Web). Cette structure se révèle être une façon naturelle de décrire la plupart des données traitées par les machines. Le sujet et l'objet sont chacun identifiés par un Identifiant de Ressource Universelle (URI), exactement comme dans les liens sur une page Web (les URLs, Uniform Resource Locators, sont les types d'URI les plus courants). Les verbes sont également identifiés par des URIs, ce qui permet à chacun de définir un nouveau concept, un nouveau verbe, simplement en définissant une URI quelque part sur le Web.

Le langage humain s'accommode de l'utilisation d'un terme identique pour désigner des choses différentes, mais le traitement automatique des données, non. Imaginez que je loue un clown pour apporter des ballons à mes clients le jour de leur anniversaire. Malheureusement, le service transfère les adresses de ma base de données à sa base de données, en ignorant que les adresses de ma base sont celles où sont envoyées les factures et que la plupart sont des boîtes postales. Mes clowns viendront divertir les employés de la poste - ce qui n'est certainement pas une mauvaise chose en soi, mais pas ce que j'attends. L'utilisation d'une URI différente pour chaque concept spécifique résout ce problème. Une adresse de courrier peut être distinguée d'une adresse de rue, et les deux peuvent se différencier d'une adresse qui est un discours.

Les triples de RDF constituent des Webs d'information sur des choses qui sont en relation avec le document. Parce que RDF utilise des URI pour coder l'information dans un document, les URI vérifient que les concepts ne sont pas simplement des mots dans un document mais sont corrélés à une définition unique que tout le monde peut trouver sur le Web. Par exemple, imaginons que nous ayons accès à plusieurs bases de données donnant des informations sur des gens, avec leurs adresses. Si nous voulons retrouver des gens avec un code postal spécifique, nous avons besoin de savoir quels sont les champs dans chaque base de données qui représentent les noms et quels sont les champs qui représentent les codes postaux. RDF peut spécifier que "(le champ 5 dans la base de données A) (est un champ de type) (code postal)" en utilisant des URIs plutôt que des phrases pour chaque terme.

## **Ontologies.**

Bien sûr, ceci n'est pas la fin de l'histoire, car deux bases de données peuvent utiliser des identifiants différents pour ce qui est en fait le même concept, code postal par exemple. Un programme qui veut comparer ou combiner l'information provenant de deux bases de données doit savoir que ces deux termes sont utilisés pour désigner la même chose. Idéalement, le programme doit pouvoir découvrir des sens identiques pour n'importe quelle base de données rencontrée.

La solution à ce problème se trouve dans le troisième composant basique du Web sémantique : l'ensemble d'informations appelé ontologies. En philosophie, une

ontologie est une théorie à propos de la nature de l'existence, des types de choses qui existent ; l'ontologie en tant que discipline étudie ces aspects théoriques. L'intelligence artificielle et les chercheurs du Web ont adopté ce terme dans leur propre jargon, et pour eux une ontologie est un document ou fichier qui définit de façon formelle les relations entre les termes. Un type d'ontologie caractéristique du Web possédera une taxonomie et un ensemble de règles d'inférence. La taxonomie définit des classes d'objets et les relations entre eux. Par exemple, une adresse peut être définie comme type de lieu, et les codes postaux de ville peuvent être définis pour s'appliquer seulement à des lieux, et ainsi de suite. Les classes, les sous-classes et les relations entre les entités sont un très puissant outil pour utiliser le Web. Nous pouvons exprimer un grand nombre de relations entre les entités en attribuant des propriétés aux classes et en permettant à des sous-classes d'hériter de leurs propriétés. Si les codes postaux doivent être de type ville et sachant que les villes ont généralement des sites Web, nous pouvons associer un site Web au code postal d'une ville même si aucune base de données ne relie directement un code à un site.

Les règles d'inférence dans les ontologies sont encore plus puissantes. Une ontologie peut exprimer la règle suivante : " si un code postal de ville est associé à un code d'état et qu'une adresse utilise ce code de ville, alors cette adresse est associée au code de l'état. " Un programme pourrait en déduire, par exemple, que l'adresse de l'université Cornell, située dans l'Ithaca, doit se trouver dans l'état de New-York, qui est aux Etats - Unis, et devrait donc être formatée selon les standards américains. L'ordinateur ne comprend pas vraiment cette information, mais il peut maintenant manipuler les termes de manière beaucoup plus efficace et significative pour un usager humain.

Avec les pages d'ontologie sur le Web, les solutions aux problèmes de terminologie (et autres) commencent à voir le jour. La signification des termes ou codes XML utilisés sur une page Web peut être définie par des pointeurs de la page vers une ontologie. Il est clair que les vieux problèmes sont prêts à ressurgir . Par exemple, si je pointe vers une ontologie qui définit des adresses comme contenant un zip code et que vous pointez vers une adresse qui utilise un postal code. On peut résoudre cette confusion si les ontologies (ou d'autres services du Web) fournissent des relations d'équivalence : une ou deux de nos ontologies peuvent contenir l'information selon laquelle zip code équivaut à postal code.

Notre projet d' envoyer des clowns divertir mes clients trouve une solution partielle dès lors que les deux bases de données pointent vers des définitions différentes d'adresse. Le programme qui utilise des URIs distinctes pour différents concepts de l'adresse, ne les confondra pas et de fait aura besoin de découvrir que les concepts sont liés entre eux. Le programme pourrait ensuite utiliser un service qui travaille sur une liste d'adresses postales (première ontologie) et la convertit en une liste d'adresses physiques (seconde ontologie) en reconnaissant et en enlevant les boîtes postales et autres adresses inutiles. La structure et la sémantique fournies par les ontologies rend plus facile la fourniture d'un service et son utilisation complètement transparente.

Les ontologies peuvent valoriser le fonctionnement du Web de plusieurs façons. On peut les utiliser de façon simple pour améliorer la pertinence des recherches - le programme

de recherche peut ne rechercher que les pages faisant référence à un concept précis au lieu de celles qui utilisent des mots-clés ambigus.

Des applications plus avancées utiliseront des ontologies pour associer l'information d'une page à des structures de connaissance et à des règles d'inférence. Un exemple de page marquée à cet usage est consultable à l'adresse <http://www.cs.umd.edu/~hendler>. Si vous allez consulter cette page, vous verrez la page du Web normal intitulée " Dr James A. Hendler ". En tant qu'humain, vous trouverez rapidement le lien vers une courte note biographique et lirez que Hendler a reçu un titre de docteur de l'université Brown. Un programme d'ordinateur qui essaie de trouver ce genre d'information devrait être très sophistiqué pour deviner que l'information est peut-être une biographie et comprendre l'anglais utilisé dans ce cas précis.

Pour les ordinateurs, la page est liée à une page d'ontologie qui définit l'information concernant les départements d'informatique. Par exemple, les professeurs travaillent dans des universités et ils ont généralement des doctorats. Le marquage d'une page (qui n'est pas visualisé par les navigateurs du Web) utilise les concepts ontologiques pour signifier que Hendler a reçu son doctorat de l'entité décrite à l'URI <http://www.brown.edu> - la page Web de Brown. Les ordinateurs peuvent également trouver qu' Hendler est membre d'un projet de recherche particulier, a une adresse électronique particulière, et ainsi de suite. Toute ces informations sont traitées instantanément par un ordinateur et peuvent être utilisées pour répondre à des questions (où le Dr Hendler a-t-il reçu son diplôme ?) qui exigeraient d'un homme de passer au crible le contenu des pages ramenées par un moteur de recherche.

De plus, ce marquage facilite le développement de programmes qui peuvent s'atteler à des questions compliquées dont les réponses ne se trouvent pas sur une seule page Web. Supposons que vous vouliez trouver la Ms Cook que vous avez rencontrée à une conférence professionnelle l'année précédente. Vous ne connaissez pas son prénom, mais vous vous rappelez qu'elle a travaillé pour un de vos clients et que son fils était étudiant dans votre université. Un programme de recherche intelligent peut filtrer les pages des personnes dont le nom est Cook (en sautant toutes les pages relatives à cooks, cooking, Cook islands...), trouver ceux qui travaillent pour une entreprise qui est sur votre liste de clients et suivre les liens vers les pages Web de leurs enfants pour découvrir s'il y en a qui sont scolarisés à la bonne école.

## **Les agents.**

Le pouvoir véritable du Web sémantique sera atteint quand les gens créeront de nombreux programmes qui collecteront les contenus du Web à partir de sources diverses, qui traiteront l'information et échangeront les résultats avec d'autres programmes. L'efficacité de ces agents logiciels croîtra de manière exponentielle au fur et à mesure que seront disponibles des contenus du Web lisibles par des machines et des services automatisés (comprenant d'autres agents). Le Web sémantique promeut cette synergie : même les agents qui ne sont pas expressément fabriqués pour travailler

ensemble peuvent transférer des données entre eux si les données sont accompagnées de sémantique.

Un aspect important du fonctionnement des agents sera l'échange de " preuves " écrites dans le langage unifié du Web sémantique (langage qui exprime les inférences logiques en utilisant des règles et des informations comme celles qui sont spécifiées par les ontologies). Par exemple, supposons que l'information pour contacter Ms Cook ait été localisée par un service en ligne, et à votre grande surprise il la situe à Johannesburg. Vous voulez naturellement vérifier cela, donc votre ordinateur demande au service de prouver sa réponse, ce qu'il fait rapidement en traduisant son raisonnement interne en langage unifié du Web sémantique. Un moteur d'inférence dans votre ordinateur vérifie que la Ms Cook en question est bien celle que vous cherchiez, et il peut vous montrer les pages Web pertinentes si vous avez encore des doutes. Bien qu'ils soient encore loin d'atteindre les profondeurs du potentiel du Web sémantique, quelques programmes sont déjà capables d'échanger des preuves de cette façon, en utilisant les premières versions courantes du langage unifié.

Un autre trait vital sera les signatures digitales, qui sont des blocs encryptés de données que les ordinateurs et les agents peuvent utiliser pour vérifier que l'information attachée a été fournie par une source fiable. Vous voulez être tout à fait sûr qu'un état envoyé à votre programme comptable qui dit que vous devez de l'argent à un détaillant en ligne n'est pas une contrefaçon produite par l'ordinateur de l'adolescent d'à côté. Les agents devraient se méfier des affirmations qu'ils lisent sur le Web sémantique jusqu'à ce qu'ils aient vérifié les sources d'information. (Nous souhaitons que davantage de personnes apprennent à le faire sur le Web tel qu'il est !)

Beaucoup de services Web automatisés existent déjà sans sémantique, mais d'autres programmes tels que les agents n'ont pas les moyens de localiser celui qui remplit une fonction spécifique. Ce traitement, appelé découverte de service, ne peut se faire que s'il y a un langage commun pour décrire un service afin que les autres agents comprennent à la fois la fonction offerte et la façon de s'en servir. Les services et les agents peuvent signaler leur fonction par exemple en mettant des descriptions dans des répertoires analogues aux Pages Jaunes.

Quelques systèmes de bas niveau de recherche de services sont actuellement disponibles, tels que ceux fournis par Microsoft (Universal Plug and Play), qui tente de connecter différents services de nature différente ou encore Sun Microsystems (Jini) qui essaie aussi de connecter des services. Ces initiatives prennent le problème au niveau structurel et syntaxique et s'appuient lourdement sur la standardisation d'un ensemble descriptif de fonctionnalités pré-déterminées. La standardisation atteint ici ses limites, car il est impossible d'anticiper tous les besoins futurs.

### **PROPREMENT FAIT, LE WEB SEMANTIQUE POURRA FAVORISER L'EVOLUTION DE LA CONNAISSANCE HUMAINE DANS SON ENSEMBLE.**

En effet, le Web sémantique est plus flexible. Les agents utilisateurs et producteurs peuvent parvenir à un état de compréhension mutuelle en échangeant des ontologies qui fournissent le vocabulaire nécessaire à la discussion. Les agents peuvent même



développer de nouvelles capacités de raisonnement lorsqu'ils découvrent des ontologies nouvelles. La sémantique facilite l'utilisation d'un service qui ne répond que partiellement à une demande.

Un processus typique inclura la création d'une chaîne de valeur dans laquelle des sous-ensembles d'information sont transmis d'un agent à l'autre, chacun ajoutant de la valeur pour construire le produit final répondant à la requête de l'utilisateur final. Ne nous méprenons pas : pour créer une chaîne de valeur automatiquement et à la demande, quelques agents associeront des technologies d'intelligence artificielle au Web sémantique. Mais c'est le Web sémantique qui donne les fondements et le cadre de développement de ces technologies.

La prise en considération simultanée de diverses caractéristiques ressemble aux compétences mises en œuvre par les agents de Pete et de Lucy dans le scénario qui introduit cet article. Leurs agents auraient délégué des morceaux de tâches à réaliser à d'autres agents et à des services découverts au fur à mesure. Par exemple, ils auraient utilisé un service certifié pour trouver une liste de fournisseurs de service et pour déterminer les plus appropriés d'entre eux en termes de services et de qualité de traitement. La liste des fournisseurs aurait pu être fournie par un autre fournisseur de service, etc.... Ces activités forment une chaîne dans laquelle une grande quantité de données distribuées dans le Web (pratiquement insignifiantes dans leur format présent) se trouve réduite à un petit ensemble de données significatives pour Lucy et Pete - constituer un calendrier de rendez-vous en rapport avec leurs emplois du temps respectifs et leurs contraintes respectives.

Lors de l'étape suivante, le Web sémantique brisera les limites du monde virtuel et s'étendra jusqu'au monde physique. Les URIs peuvent tout décrire et tout pointer, y compris des entités physiques, ce qui signifie que nous pouvons utiliser le langage RDF pour décrire des dispositifs comme des téléphones cellulaires ou des télévisions. De tels dispositifs peuvent annoncer leurs fonctionnalités : ce qu'ils peuvent faire et la manière dont ils sont contrôlés, tout comme des agents logiciels. Comme ils sont beaucoup plus flexibles que des systèmes de bas niveau comme Universal Plug and Play, une approche sémantique ouvre une multitude de possibilités intéressantes. Ce que nous appelons aujourd'hui domotique exige des configurations extrêmement précises. Des descriptions sémantiques des capacités et des fonctionnalités nous permettront d'atteindre de bons niveaux d'automatisation tout en réduisant l'intervention humaine. Nous pouvons donner un exemple simple avec Pete qui répond au téléphone et dont la stéréo va automatiquement baisser le niveau sonore. Au lieu de devoir programmer chaque élément de façon spécifique, il programmera une telle fonction une fois pour toutes pour prendre en charge tous les outils qui ont un niveau sonore, qu'il s'agisse du téléviseur, du lecteur de DVD, de l'ordinateur portable apporté pour travailler à la maison. Les premiers résultats concrets qui existent dans ce domaine à partir des travaux réalisés pour développer un standard de description de capacités fonctionnelles de services (tels que la dimension des écrans) et les préférences des utilisateurs. Réalisés en RDF, ce standard s'appelle CC (Capacité Composite)/ PP (Profil de Préférence). A l'origine, il est fait pour permettre à des téléphones cellulaires et à d'autres clients Web

non-standard de décrire leurs caractéristiques de sorte que le contenu du Web puisse être taillé sur mesure pour eux. Plus tard, lorsque nous pourrons utiliser la flexibilité de la langue pour manipuler les ontologies et les règles logiques, les différents ustensiles pourront eux-mêmes chercher et utiliser ces services pour ajouter de l'information et des fonctionnalités. Il n'est pas difficile d'imaginer qu'un four à micro-ondes géré par le Web pourra consulter le site Web du fabricant de congélateurs pour optimiser les temps de cuisson.

## **L'évolution de la connaissance.**

Le Web sémantique n'est pas simplement un outil pour réaliser les tâches individuelles dont nous avons déjà parlé. De plus, s'il est bien fait, le Web sémantique pourra favoriser l'évolution de la connaissance humaine dans son ensemble. Les efforts humains sont en permanence pris dans une tension entre l'efficacité de petits groupes d'individus agissant indépendamment et les besoins de tissage d'une vaste communauté. Un petit groupe peut innover de façon rapide et cohérente, mais cette situation crée une sous-culture dont les concepts ne sont pas compris par les autres. A l'opposé, des actions de coordination portant sur une large communauté sont lentes et douloureuses et nécessitent beaucoup d'efforts de communication. Le monde travaille entre ces deux extrêmes, avec une tendance à commencer avec de petits groupes et des idées personnelles et à se développer vers une compréhension plus large avec le temps.

Il est essentiel de rassembler ces deux sous-cultures lorsqu'un langage commun plus large est nécessaire. Souvent les deux groupes développent indépendamment des concepts très voisins et la description de la relation qui existe entre eux peut apporter de larges bénéfices. Comme le dictionnaire franco-danois, ou une table de conversion poids et mesures, les relations permettent de communiquer et collaborer même lorsque les points communs entre les concepts n'ont pas encore entraîné de terminologie commune. Le Web sémantique en nommant chaque concept simplement dans une URI permet à chacun d'exprimer de nouveaux concepts inventés avec un minimum d'efforts. Ce langage logique unificateur permettra à ces nouveaux concepts d'être progressivement reliés au Web universel. Cette structure ouvrira la connaissance et les travaux humains à l'analyse signifiante d'agents logiciels donnant accès à une nouvelle classe d'outils grâce auxquels nous pourrons vivre, travailler, apprendre ensemble.

On peut trouver plus d'informations dans :

- Weaving the Web : The original design and ultimate destiny of the World Wide Web by its Inventors, Tim Berners-Lee, with Mark Fischetti. Harper, San Francisco, 1999
- [World Wide Web](#)
- [An introduction to ontologies](#)
- [W3C Semantic Web Activity](#)

- ["Is there an intelligent agent in your future \(Nature Webmatters, 3/99\)"](#) which was the recipient of the "2000 AAAI Effective Expository Writing Award."
- [Agents on the Semantic Web](#) (to appear in IEEE Intelligent Systems) which describes a model of how agents and ontologies play together in the web of the future.