

## WEB SEMÂNTICA: DIFICULDADES E DESAFIOS

Daniel Santos de Andrade<sup>1</sup>

### RESUMO

A Web Semântica, projeto idealizado e conduzido pelo consórcio W3C, promete ser uma evolução da Web atual, sendo que sua principal meta é fazer com que os computadores sejam capazes de entender os elementos presentes nas páginas da Internet, utilizando-se de padrões e linguagens como RDF e XML. Este artigo visa esclarecer o que é este projeto, mostrando seu histórico, desde sua idealização até os dias atuais, e os esforços e estudos que vêm sendo feitos para que ele se torne realidade. São enfocadas as dificuldades, que a Web Semântica vem enfrentando em sua implantação, com ontologias e desestruturação da Web - razões pelas quais ainda não é um sucesso, mesmo sendo esse um sistema útil e que poderá ajudar nos processos de diversas áreas de atuação.

Palavras-chave: Web. Semântica. Web Semântica. Ontologias. Metadados.

### 1 INTRODUÇÃO

A Web Semântica é ao mesmo tempo uma espécie de evolução e de extensão da web atual, que pode, entre outras vantagens, tornar ainda mais colaborativa a relação entre o homem e o computador.

Quando se acessa uma página na Internet, a mente humana é capaz de reconhecer seus elementos, distinguir seus textos, interpretar seus significados e compreender suas informações, mas um computador não é capaz de nada disso. Tudo o que ele pode entender é uma complexa seqüência de caracteres binários que indicam apenas como exibir o conteúdo de maneira correta. Dessa forma, um dos principais objetivos da web semântica é tornar documentos dispostos na internet legíveis para sistemas informatizados, assim como o são para os seres humanos, e fazer com que os computadores sejam capazes de distinguir e entender os elementos presentes nas páginas da web. Para que isso ocorra, devem ser utilizadas, além de tecnologias avançadas (como os agentes), as ontologias, com o fim de

relacionar termos, fazendo com que os sistemas de buscas se tornem ainda mais poderosos e eficientes e trazendo melhores resultados ao usuário final. A padronização das tecnologias, metadados descritivos e linguagens utilizadas no desenvolvimento de sistemas na web também são fatores imprescindíveis para a concretização do sistema.

A web semântica foi idealizada em 2001, pelo mesmo homem que foi um dos maiores responsáveis pela criação da internet, Tim Berners-Lee, e seu projeto vem sendo conduzido pelo consórcio W3C (World Wide Web Consortium), sucedendo o antigo projeto Metadados e herdando dele seus princípios e metas. A diferença entre os dois projetos é que o Metadados trabalha apenas com a linguagem HTML, uma linguagem estática, que não permite a utilização de semântica e o relacionamento entre termos, e, serve somente para definir a maneira como a página será exibida. Já a web semântica utiliza-se da linguagem XML (Extensible Markup Language), muito mais dinâmica, e que torna possível que os dados sejam semanticamente descritos e categorizados, de acordo

<sup>1</sup> Tecnólogo em Informática para a Gestão de Negócios pela Faculdade de Tecnologia de Mauá - FATEC Mauá. E-mail: daniel.andrade@msn.com.

com as definições do próprio usuário, através do uso de tags.

A Web Semântica não deve servir para substituir a Web atual, mas virá a ser uma extensão dela. As duas deverão, portanto, coexistir.

Os padrões que podem fazer a Web Semântica se tornar realidade já existem há algum tempo, porém o número de sites que se utilizam deles ainda é ínfimo.

Apesar de todos os estudos a respeito do assunto, a implantação e o uso da semântica na Web não é um trabalho simples, e, tem se deparado com diversos obstáculos e dificuldades ao longo do tempo, entre eles o fato da Web atual estar totalmente desestruturada e com seus dados desorganizados. Sanar esse problema pode exigir que toda a informação atualmente presente na Internet seja reestruturada.

A Web semântica pode ser uma das maiores e mais importantes das evoluções já vistas em relação à internet, e, seu futuro uso pode mudar de uma vez por todas a maneira de busca de conteúdo na web, e, também, a forma como os resultados sobre este conteúdo será disponibilizado. Filtrar resultados de buscas não será mais tarefa exclusiva dos seres humanos, mas ficará também a cargo de agentes, que são aplicativos capazes de executar tarefas complexas, simplificando cada vez mais a navegação pela internet.

A importância da utilização da web semântica poderá ser sentida não apenas pelos usuários domésticos, mas também pelas empresas, pois ela fará com que sistemas diferentes se comuniquem entre si, realizando a chamada interoperabilidade. Podem-se esperar também mudanças na atuação de todos os profissionais de informática, com a criação de novos motores de busca e melhoria dos motores já existentes, o surgimento de novas e melhores interfaces com o usuário e a inserção de significados nos documentos. Até mesmo a educação à distância pode ser beneficiada pela Web semântica. Mesmo assim, essa nova forma de internet ainda não é amplamente utilizada.

O desenvolvimento deste artigo pautou-se em análise bibliográfica e em estudo de sites que já se utilizam do sistema.

## 2 HISTÓRICO E EVOLUÇÃO DA WEB SEMÂNTICA

Em um artigo chamado "Web Semântica: um novo formato de conteúdo para a Web que tem significado para computadores vai iniciar uma revolução de novas possibilidades.", publicado em 2001 na revista Scientific American, Tim Berners-Lee, James Hendler e Ora Lassila definiram a Web Semântica como sendo "uma extensão da Web atual, na qual é dado à informação um significado bem definido, permitindo que computadores e pessoas trabalhem em cooperação." (BERNERS-LEE; HENDLER; LASSILA, 2001, p. 2).

Nesse artigo, Berners-Lee, que anos antes fora um dos principais responsáveis pela criação do sistema WWW (World Wide Web), descreveu um cenário que, na época, pareceria oriundo de algum filme de ficção científica, no qual agentes seriam capazes de marcar consultas médicas e sessões de fisioterapia em horários baseados na agenda do usuário e levando em consideração seu histórico de saúde. Da mesma forma, seria possível que ao toque do telefone os demais aparelhos de uma residência fossem automaticamente desligados ou tivessem o som reduzido, desde que essa função fosse previamente programada pelo usuário. Inicialmente, o W3C desejava que os agentes da web semântica estivessem presentes também em dispositivos móveis, aumentando exponencialmente sua abrangência, deixando claro que a Web Semântica seria um fator importante para facilitar a vida de seus usuários.

A ideia desse novo formato de web é baseada nos princípios e metas do antigo projeto Metadados, que em 1994 já apontava a necessidade da utilização

de semântica para descrição de recursos dispostos na internet. (BREITMAN, 2005). Esse antigo projeto e seu sucessor são conduzidos pelo consórcio W3C (World Wide Web Consortium), um consórcio fundado em 1994 por Tim Berners-Lee, no qual empresas filiadas, uma equipe própria e até mesmo o público trabalham em conjunto para desenvolver padrões para a web, com a finalidade de levá-la ao seu máximo potencial.

### 3 O PROJETO METADADOS

A definição mais comumente dada aos metadados é que eles são "dados sobre dados. O termo se refere a qualquer informação utilizada para a identificação, descrição e localização de recursos." (BREITMAN, 2005, p. 16). O consórcio W3C apresenta uma visão menos abrangente do termo, porém mais voltada à web semântica: "informações para a web que podem ser compreendidas por máquinas." Esta definição dada pelo W3C limita o termo metadados a sistemas eletrônicos, o que não está completamente correto, uma vez que até mesmo os arcaicos cartões de biblioteca podem ser considerados metadados, pois apresentam as informações referentes à identificação, descrição e localização de livros. Os metadados vêm sendo utilizados há séculos pelas mais diversas áreas, não se limitando, portanto, a meios digitais.

Sabendo da importância que dados sobre dados poderiam ter para o futuro da internet, o W3C lançou nos anos 1990 o projeto Metadados, cujas maiores preocupações eram as formas de modelar e codificar metadados, e, visava inserir significados nos documentos da internet, o que mais tarde viria a ser um dos principais objetivos da web semântica. A forte preocupação do W3C com os metadados teve como consequência a criação dos padrões RDF (Resource Description Framework) e PICS (Platform for Internet

Content Selection). Os trabalhos com o sistema PICS foram concluídos, mas os estudos com a linguagem RDF continuam.

O projeto Metadados utilizava-se da linguagem HTML, o que tornava maior a dificuldade em atingir seus objetivos, uma vez que essa linguagem é estática e serve apenas para a renderização da página, ou seja, o modo de exibi-la. Dessa forma, o HTML não permite a inclusão de significados ou descrição dos itens da página exibida. Já a web semântica utiliza-se dos padrões XML e RDF, ambos muito mais dinâmicos e que permitem a inclusão de semântica no conteúdo, sendo que RDF são modelos de dados utilizados para representar objetos e seus relacionamentos, e, a semântica necessária para tal é representada pela linguagem XML.

As principais funções do RDF são fornecer modelos de dados e sintaxe para a codificação de metadados, de modo que esses possam ser entendidos por máquinas, e, também, fornecer a interoperabilidade entre agentes que tenham a necessidade de troca de informações pela internet. (BREITMAN, 2005)

A expectativa com todas as facilidades e funcionalidades que a Web Semântica pode trazer, em conjunto com as linguagens já desenvolvidas e que poderiam ser aplicadas neste novo projeto, permitiu o W3C entrar em uma grande euforia em relação ao futuro da internet. Assim, no ano de 2001, o W3C abandona de uma vez por todas o antigo Projeto Metadados e substituí-lo pelo projeto Web Semântica. Desde então, pesquisadores de inteligência artificial vêm propondo uma série de modelos a serem utilizados no novo formato de Web. (BREITMAN, 2005, p. 5). Mas, embora a Web Semântica tenha sido idealizada há quase dez anos, sua utilização ainda não está em patamares avançados de desenvolvimento.

As linguagens, padrões e sistemas a serem utilizados na chamada "Web 3.0" já estão plenamente

desenvolvidos. O próprio James Hendler, um de seus idealizadores, afirma que há mais de dois anos já existem programas na Web que aplicam essas tecnologias em larga escala, e que tantos outros continuam a surgir (MORAES, 2009, p. 60). Ainda assim, buscadores que utilizam os métodos convencionais e antigos de busca pela internet, como Google e Yahoo, ainda são indispensáveis.

"Apesar de a Web Semântica trazer benefícios, sua adoção ampla continua sendo uma promessa." (MORAES, 2009, p. 60)

Já existem sites construídos sobre os padrões da nova Web, até mesmo no Brasil, mas o número deles é ainda é ínfimo em comparação ao que se esperava que houvesse. A popularização da Web Semântica parece ter se restringido a sistemas pessoais ou restritos de busca, como as Intranets de algumas empresas.

#### **4 AS BARREIRAS ENFRENTADAS PELA WEBSEMÂNTICA**

Diversos fatores são apontados como as possíveis causas que impedem o crescimento da Web Semântica, dentre eles a falta de infraestrutura necessária para fazê-la acontecer, a dificuldade com a criação de ontologias, e a desestruturação e desorganização da Web atual.

Wilson (2006, p.1) aponta mais alguns fatores, afirmando que é quase impossível que empresas e sites existentes dediquem o tempo e os recursos necessários para adicionar todos os metadados, e que os softwares utilizados para este fim podem ainda não ser suficientes para tornar o projeto viável.

#### **5 PROBLEMAS: FALTA DE INFRAESTRUTURA**

Muito se fala que a falta de infraestrutura é um grande empecilho para a concretização da Web Semântica. Sobre esta afirmação, Hendler (2001, p. 1)

diz:

Em uma palestra que assisti recentemente, o palestrante descreveu como 'ficção científica' a visão compreendendo agentes percorrendo a Web e realizando tarefas complexas para seus usuários. Esse palestrante defendeu que estamos longe do dia em que esta visão se tornará realidade, pois ainda não temos a infra-estrutura necessária para fazê-la acontecer.

Hendler fecha a citação, dizendo que concorda com o palestrante no que se refere à falta de infraestrutura, mas deixa claro que acha a expressão "ficção científica" pessimista demais, pois um aspecto crucial dessa infraestrutura, uma linguagem padronizada de ontologias, está surgindo.

Da mesma forma, Souza e Alvarenga (2004, p.1) afirmam que a infraestrutura comum da internet pode ser utilizada também na Web semântica. Assim sendo, a infraestrutura não pode ser considerada como um problema tão grande.

#### **6 PROBLEMAS: DIFICULDADES COM ONTOLOGIAS**

A palavra ontologia vem do grego "ontos" (ser) + "logos" (palavra). (BREITMAN, 2005, p. 30)

Segundo a Wikipédia ([S.d.]), no que se refere à filosofia, ontologias tratam da natureza do ser, da realidade, da existência dos entes e das questões metafísicas em geral. A ontologia trata do ser enquanto ser, isto é, do ser concebido como tendo uma natureza comum que é inerente a todos e a cada um dos seres.

Em Ciência da Computação e Ciência da Informação, uma ontologia é um modelo de dados que representa um conjunto de conceitos dentro de um domínio e os relacionamentos entre estes. Uma ontologia é utilizada para realizar inferência sobre os objetos do domínio. (WIKIPEDIA, [S.d.])

O Consórcio W3C conceitua uma ontologia

como "a definição dos termos utilizados na descrição e na representação de uma área do conhecimento."

Para Wilson (2006, p. 5), uma ontologia é um "vocabulário que descreve objetos e como eles se relacionam."

As ontologias vêm sendo utilizadas nas mais diversas áreas do conhecimento, como inteligência artificial, processamento de linguagem natural, engenharia de software e na própria web semântica, onde, de acordo com Breitman (2005, p. 43), seu papel é "explicitar o vocabulário utilizado e fornecer um padrão para o compartilhamento de informação." Dessa forma, elas seriam capazes de fornecer um modelo comum que tornaria possível a troca de informações entre agentes de software e aplicações.

"Para a Web Semântica ser eficaz, as ontologias precisam ser detalhadas e compreensivas" (STRICKLAND, 2008, p. 1). Porém, a criação de ontologias não é tarefa fácil para a maioria das pessoas, e por isso elas podem ser as principais responsáveis pelo atual insucesso da Web Semântica, como será abordado a seguir.

"O potencial do uso de ontologias para lidar com o problema da semântica de recursos de informação, sobretudo quando há grandes volumes de informação, tem sido largamente explorado pelas áreas de pesquisa da Web Semântica." (FALBO et al, 2004, p. 3) O objetivo dessas equipes é explorar o potencial da utilização de ontologias, principalmente levando em consideração que essas podem ser de extrema utilidade na organização dos grandes volumes de informação que hoje são encontrados na Internet.

Mas, as ontologias são hoje apontadas como uma das principais dificuldades encontradas no processo de implantação da web semântica.

Strickland (2008, p. 1) afirma: "A construção de ontologias requer muito trabalho. De fato, esse é um dos grandes obstáculos que a Web Semântica enfrenta." Ainda, cita: "Críticas sugerem que a tarefa de criar e manter tais arquivos complexos é

trabalho demais para a maioria das pessoas". Diante das dificuldades, Strickland (2008, p. 1) ainda alerta para o fato de que as pessoas podem não estar dispostas a fazer ontologias compreensivas em seus sites, e caso estas páginas necessitem de alguma mudança posterior, os usuários podem não manter as ontologias da maneira como estavam antes das alterações realizadas.

O problema com ontologias é que são muito difíceis de criar, implementar e manter. Dependendo de seu escopo, podem ser enormes, definindo uma grande variedade de conceitos e relações. Alguns desenvolvedores preferem se concentrar em lógica e regras ao invés de ontologias por causa de suas dificuldades. Desacordos a respeito dos papéis que estas regras deveriam desempenhar podem ser uma armadilha para a Web Semântica. (WILSON, 2006, p. 5)

James Hendler (2001, p. 2) afirma que os usuários não irão inserir ontologias nas suas páginas se não perceberem o quão importante é fazê-lo, e ferramentas para demonstrar essa importância não serão desenvolvidas se as ontologias não forem utilizadas. Dessa forma, cria-se um verdadeiro impasse, e para resolvê-lo estão sendo desenvolvidas ferramentas gratuitas de manipulação de conteúdo, com o intuito de demonstrar ao governo e a outras partes da sociedade que a Web Semântica pode ser uma realidade.

Mesmo com as ferramentas de manipulação de ontologias tomando o processo mais simples, não se pode esperar que usuários que não são experientes ou não tenham facilidade com lógica, e especialmente com ontologias, sejam capazes de criar conteúdos legíveis por computadores, mas Hendler (2001, p. 2) afirma que este é um aspecto crucial para a concretização da Web Semântica. Essa dificuldade, por parte dos usuários, confronta-se com as seguintes afirmações:

A concepção de ontologias deve ser conduzida como qualquer outro projeto de software, no sentido de serem tomadas decisões de projeto que determinam sua qualidade baseada em critérios como eficiência, legibilidade, portabilidade, extensibilidade, interoperabilidade e reuso. (FREITAS [S.d.], p. 27)

Em nome de uma compreensão mútua por parte dos intervenientes no processo de construção de ontologias, é desejável que, se não for possível encontrar uma terminologia comum, pelo menos a terminologia usada pelas várias partes seja compreendida reciprocamente. (FRANCA [S.d.], p. 1)

Nem todos os usuários são capazes de conceber ontologias de acordo com os critérios definidos por Freitas e França, principalmente se levarmos em consideração o fato de que a maioria deles até mesmo desconhece as ferramentas utilizadas para tal. Por isso, Hendler (2001, p. 2) afirma que o ideal seria que, exceto um pequeno número de desenvolvedores, os usuários comuns nem soubessem da existência de ontologias.

A diferença de vocabulário entre a linguagem natural humana e a linguagem de máquina é um grande problema que também precisa contornado. Sobre essa diferença, Wilson (2006, p. 4) diz:

Outro obstáculo para a Web Semântica é que computadores não têm o mesmo tipo de vocabulário que as pessoas. Você usa linguagem desde sempre, então é provável que seja fácil para você ver conexões entre palavras e conceitos diferentes e inferir significados com base em contextos. [...] Para entender o que as palavras significam e qual é a relação entre elas, o computador precisa ter documentos que descrevem todas as palavras e a lógica para fazer as conexões necessárias.

Moraes (2009, p. 60) acredita que para

contornar o problema com a criação de ontologias, ferramentas automatizadas baseadas em inteligência artificial poderiam cuidar de todo o trabalho sozinhas. Entretanto, este é outro impasse, pois criar estas ferramentas pode ser um trabalho ainda mais complexo do que criar ontologias.

Segundo Breitman (2005, p.68):

Cientistas do mundo todo estão buscando a metodologia mais adequada para o desenvolvimento de ontologias para a Web Semântica. Vários grupos de pesquisa estão empenhados nesta tarefa, entre eles pesquisadores das universidades de Manchester, Stanford, Karlsruhe, Politécnico de Milão, PUC-Rio, Toronto, Maryland, Vrije e MIT, além de organizações tais como DARPA, Bell Labs e W3C.

O objetivo desses pesquisadores é facilitar a criação de ontologias, para que talvez assim a Web semântica, enfim, torne-se realidade.

## 7 PROBLEMAS: A DESESTRUTURAÇÃO DA WEB ATUAL

Apesar de crescer exponencialmente a cada dia, a Web atual, também conhecida como Web Sintática, apresenta-se, cada dia, mais desestruturada. O volume de informações na Internet aumenta, juntamente com o número de páginas nela existentes, porém sem controle algum sobre seu conteúdo. Para se ter uma idéia, segundo o jornal O Globo (2007), no ano de 2006, o mundo gerou 161 bilhões de gigabytes de informação, volume maior em três milhões de vezes do que as informações contidas em todos os livros já escritos; este volume cresce a uma taxa anual de 57%.

Enquanto, os computadores conseguem apenas saber a maneira correta de exibir uma página, a função de identificar os seus conteúdos fica a cargo dos seres humanos. Ou seja, os motores de busca, como Google e Yahoo, trazem os resultados, mas o usuário

fica responsável por filtrá-los e por separar o que é relevante daquilo que não é. Na maioria das vezes, essa é uma tarefa árdua e que pode levar bastante tempo para ser realizada, dependendo do termo pesquisado e do número de resultados trazidos.

Por exemplo, uma pesquisa simples com a frase "apartamentos no centro de São Paulo" no Google trouxe 5.350.000 resultados, e no Yahoo 3.530.000, dessa forma, um usuário que realmente esteja interessado na compra de um apartamento no centro da cidade de São Paulo terá que dispor de um tempo razoável para filtrar essa busca e encontrar aquilo que deseja. É quase impossível saber quantos desses milhões de resultados realmente possuem relevância em meio a um volume enorme de conteúdo inútil.

Schons (2007, p. 5) acredita que "um usuário, ao efetuar uma pesquisa na Internet, tende a ficar mais frustrado pelo excesso de 'respostas' que pela falta delas. Isto porque o excesso de informações não representa uma solução, mas um problema: a desinformação." Ainda: "A explosão informacional trouxe à tona graves problemas quanto à falta de organização e estruturação de informações na Internet, dificultando a recuperação de informações relevantes aos usuários." (SCHONS, 2007, p. 5)

Para resolver o caos informacional em que a internet atualmente está mergulhada, as áreas de pesquisa e desenvolvimento tornaram-se indispensáveis para tentar tornar possível uma nova forma de estruturação da Internet.

Breitman (2005, p. 5) afirma que pesquisadores de inteligência artificial vêm propondo uma série de modelos, com o objetivo de organizar as informações presentes na Internet. Segundo ela, a idéia central destes modelos é categorizar a informação de maneira padronizada, para facilitar seu acesso. Para isso, seriam usadas taxonomias semelhantes às utilizadas por biólogos na classificação de seres vivos. "Existem vários esforços no sentido de se criar um

modelo estruturado para a organização da informação da Internet." (BREITMAN, 2005, p. 5)

Segundo Schons (2007, p. 7), o desenvolvimento de novas tecnologias deve atuar nas seguintes vertentes:

- Aperfeiçoar os recursos já existentes, de modo que novos padrões possam aprimorar o processo de representar e armazenar informações e documentos já disponíveis;

- Prover novas metodologias e mecanismos para o registro, seleção, busca e representação da informação, de modo que o fluxo nessas atividades seja eficiente.

Como um dos principais objetivos da Web Semântica é fazer com que as informações sejam recuperadas mais rapidamente e de forma mais eficiente, a idéia de implantá-la, bem como seus padrões e ontologias podem ganhar campo como soluções para a atual desestruturação da Web. Mas, Schons (2007, p. 11) afirma:

[...] para que tal interação ocorra, é fundamental que as tecnologias sejam padronizadas, ou seja, que se estabeleçam padrões para descrição dos dados, além de uma linguagem comum que permita a construção e codificação de significados compartilhados.

Enquanto, os sistemas que serão capazes de resolver, ou de ao menos minimizar o problema com a desestruturação da Web sintática ainda não existem, a mesma continuará a padecer de crescimento excessivo e sem controle. Isto ocorre pelo fato da Internet ser um território que oferece extrema liberdade aos usuários, para estes postarem seus conteúdos, por mais triviais que esses possam ser.

## 8 PROBLEMAS: FALTA DE INTERESSE

A Web Semântica dificilmente irá ser um projeto bem sucedido, se não houver um interesse maior por parte dos usuários envolvidos. Porém, a maioria dos usuários e das pessoas que publicam conteúdo na Internet sequer sabe da existência de algo chamado Web Semântica, e aqueles que o sabem, chegam a tratar o assunto com certo desdém.

Em entrevista dada à revista Info Exame (2009, p. 60), Fileto afirmou que as empresas, muitas vezes, recusam-se a adotar padrões por simples interesses comerciais. Desta forma cria-se uma barreira no mundo corporativo, como um todo, embora seja justamente nas Intranets de empresas que a web Semântica tem se tornado mais popular. A maioria das empresas pode não ter enxergado os benefícios que a Web Semântica pode trazer, inclusive para elas próprias.

Já foi dito que para a Web Semântica funcionar, talvez toda a informação existente na Web atual deverá passar por uma grande reestruturação. "[...] Algumas pessoas realmente gostam de rotular ou identificar objetos e informações da Web." (STRICKLAND, 2008, p. 5). Apesar desta afirmação, acredita-se que a marcação de páginas exige um esforço que a grande maioria dos usuários não quer fazer.

Wilson (2006, p.1) afirma que "enquanto alguns sites já estão usando conceitos da Web Semântica, muitas das ferramentas necessárias ainda estão em desenvolvimento." Contrariando a afirmação de Wilson, Moraes (2009, p. 60) aponta que os padrões e linguagens da Web Semântica já existem há algum tempo. Dentre estes padrões destacam-se o RDF e o XML, ambos já apresentados neste estudo.

Quanto aos softwares utilizados, muito esforço por parte de desenvolvedores vem sendo feito, com o objetivo de criar ferramentas de software que possam facilitar o processo de criação de ontologias e marcação de metadados. A maioria, se não todas,

dessas ferramentas são públicas e de código aberto, e são fáceis de serem utilizadas, não devendo ser, portanto, uma barreira no processo de crescimento da Web Semântica.

Porém, a barreira apresentada em relação a software, segundo Wilson (2006, p. 1) é: "No futuro, software produzido em série pode incluir opções para adicionar metadados ao criar novos documentos, mas esta ferramenta pode ainda não ser suficiente para tornar o projeto viável em grande escala."

Os principais software utilizados para a criação de ontologias e marcação de metadados são o OilEd, o Protege 2000 e o Ontoedit.

## 9 ONDE A WEB SEMÂNTICA JÁ EXISTE

Embora, a Web Semântica ainda não seja amplamente utilizada, ou mesmo conhecida, alguns sites e ferramentas já estão construídos sobre seus padrões. A maioria deles, em segundo plano.

Segundo Strickland (2008) vários blogs já incluem uma opção de identificação, facilitando a classificação de postagens sob tópicos específicos.

Algumas outras ferramentas aproximam-se dos objetivos da Web Semântica, muitas delas desenvolvidas pelo Google, como o Google Analytics e o Google Webmasters Tools.

Um dos sites onde as características da Web Semântica são mais claramente visíveis é o portal de finanças da Yahoo (Yahoo! Finance).

O Yahoo! Finance possui diversas informações financeiras relevantes ao mundo dos negócios, como cotação de ações de empresas e as variações nessas cotações, ocorridas recentemente. As mesmas são atualizadas automaticamente, em tempo real. Em uma barra na parte superior da página, permanecem sendo exibidas informações sobre as últimas empresas pesquisadas pelo usuário. A página é capaz de identificar dois elementos presentes: a última cotação e a variação sofrida por essa cotação. E abaixo das

informações financeiras são exibidas notícias relacionadas à empresa pesquisada, previamente identificada pelos agentes de busca.

## 10 CONCLUSÃO

A Web Semântica promete revolucionar as buscas pela internet, utilizando-se de relacionamento entre termos, descrição de conteúdo e ontologias. Linguagens como XML e padrões como o RDF e RDF-Schema também serão utilizados e estão prontos há algum tempo. A infraestrutura necessária também já existe. Tudo com o objetivo de aperfeiçoar os motores de busca e tornar as pesquisas mais simples e precisas, automatizando a filtragem de resultados, ao invés de deixá-la exclusivamente a cargo dos seres humanos.

Mesmo tendo sido idealizada há quase dez anos, e apesar dos visíveis esforços do consórcio W3C, até hoje a Web Semântica não é amplamente utilizada, e pode-se dizer até que é desconhecida pela grande maioria das pessoas.

A esperada Internet do futuro esbarra em obstáculos que nem mesmo todo o esforço já feito tem conseguido superar. As dificuldades com a utilização e criação de ontologias e a desestruturação da web atual são as maiores barreiras no caminho da Web Semântica, rumo ao seu máximo potencial. Esses fatores, aliados ao desinteresse por parte dos usuários e dos desenvolvedores, podem fazer com que, pelo menos, em curto prazo, a Web Semântica não passe de um projeto em fase experimental.

Ferramentas úteis e de fácil utilização podem

ser usadas para a criação e manipulação de ontologias, facilitando esse processo, mas dificilmente os usuários se interessarão por elas, e sem um interesse maior por parte destes e dos desenvolvedores, dificilmente o projeto da Web Semântica irá ter um futuro, e pode ser abandonado como alguns de seus antecessores, sendo o maior exemplo o projeto Metadados.

A desestruturação da Web atual é outro fator importante. O volume de dados atualmente presente na Internet e a desorganização em que eles estão dispostos faz com que a Web se tome um complexo emaranhado de informações, muitas delas inúteis. E a tendência futura é que esse problema só piore, pois a cada dia, milhares de novas informações e páginas surgem na Internet.

Alguns sites já conseguiram superar com relativo sucesso as dificuldades que a Web semântica enfrenta. Tomando esses sucessos como exemplo é possível supor que, afinal, essas barreiras não sejam intransponíveis e que outros sites podem seguir esse exemplo e superá-las também.

Um esforço enorme vem sendo feito pelas mais diversas pessoas e instituições, para que a Web Semântica deixe de ser um projeto e enfim se tome realidade, mas nem mesmo esse esforço trará resultados em curto prazo.

A menos que sejam feitos investimentos maciços em pesquisa e desenvolvimento de novas ferramentas, especialmente de manipulação de ontologias, e que a Internet sofra uma organização do seu conteúdo, a grande utilização da Web Semântica só poderá ser esperada para um futuro não muito próximo.

## SEMANTIC WEB: DIFFICULTIES AND CHALLENGES

### ABSTRACT

The Semantic Web, that is a Project idealized and conducted by the W3C Consortium, promises to be an evolution of the nowadays Web, and its main goal is to make computers able to understand the elements present in the web pages using standards and languages like RDF and XML. This article aims to clarify what this project is, showing its history, since its conception till the present days, and the efforts and studies that have been made to make it become real. It focuses on the difficulties that the Semantic Web has faced in its implementation, such as: difficulties with ontologies and the Web non-structure, which are the reasons why It, even as a useful system that may help in several areas processes, is not being a success yet.

Keywords: Web. Semantic. Semantic Web. Ontologies. Metadata.

## REFERÊNCIAS

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. **The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities**. Publicação: 2001. Disponível em: < <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>>. Acesso em: 18 mar. 2010.

BREITMAN, Karin Koogan. **Web Semântica: A Internet do futuro**. Rio de Janeiro: LTC, 2005.

FALBO, Ricardo A.; RUY, Fabiano B.; PEZZIN, Juliana; MORO, Rodrigo Dal. **Ontologias e Ambientes de Desenvolvimento de Software Semânticos**. Publicação: 2004. Disponível em: <<http://www.inf.ufes.br/~falbo/download/pub/2004-JMSIC-1.pdf>>. Acesso em: 22 abr. 2010.

FREITAS, Frederico Luiz Gonçalves de. **Ontologias e a Web Semântica**. Tese apresentada à Universidade Católica de Santos (UniSantos) no Programa de Pós-Graduação em Informática. Disponível em: <[http://pages.udesc.br/~r4sm/material/Ontologia\\_Web\\_semantica%20Freitas.pdf](http://pages.udesc.br/~r4sm/material/Ontologia_Web_semantica%20Freitas.pdf)>. Acesso em: 3 mai. 2010.

FRANCA, Patricia Cunha. **Conceitos, classes e/ou universais: com o que é que se constrói uma ontologia?** ([S.d.]). Disponível em: <<http://linguamatica.com/index.php/linguamatica/article/view/10/13>>. Acesso em: 3. mai. 2010.

HENDLER, James. **Agents and the Semantic Web**. Publicação: 2001. Disponível em: <[http://classweb.gmu.edu/kersQh/infs770/Semantic\\_Web\\_16\\_2/Hendler.pdf](http://classweb.gmu.edu/kersQh/infs770/Semantic_Web_16_2/Hendler.pdf)>. Acesso em: 16 fev. 2010.

MORAES, Mauricio. A Web do Futuro é Semântica? **Revista Info Exame**. São Paulo: Abril, n. 280, p. 58-60, jun. 2009.

SCHONS, Cláudio Henrique. **O volume de informações na Internet e sua desorganização: Reflexões e perspectivas**. Publicação: 2007. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/1748/1497>>. Acesso em: 29 abr. 2010.

SCIAM. **Scientific American Magazine**. Publicação: 2001. Disponível em: <<http://www.scientificamerican.com/article.cfm?id=the-semantic-web>>. Acesso em: 13 mar. 2010.

SOUZA, Renato Rocha; ALVARENGA, Lidia. **A Web Semântica e suas contribuições para a ciência da informação**. Publicação: 2004. Disponível em: <[http://www.scielo.br/scielo.php?pid=S0100-19652004000100016&script=sci\\_arttext&tlng=pt](http://www.scielo.br/scielo.php?pid=S0100-19652004000100016&script=sci_arttext&tlng=pt)>. Acesso em: 07 mar. 2010.

STRICKLAND, Jonathan. **HowStuffWorks - Como funcionará a Web 3.0**. Publicação: 2008. Disponível em: <<http://informatica.hsw.uol.com.br/web-304.htm>>. Acesso em: 17 abr. 2010.

WIKIPÉDIA. **Ontologia**. Disponível em: < <http://pt.wikipedia.org/wiki/Ontologia>>. Acesso em: 22 abr. 2010.

WILSON, Tracy V. **HowStuffWorks - Como funciona a Web Semântica**. Publicação: 2006. Disponível em: <<http://informatica.hsw.uol.com.br/web-semantica6.htm>>. Acesso em: 15 abr. 2010.