

Web Semântica e Cérebro Global juntos por uma boa causa

Semantic Web and Global Brain together for a worthy cause

Vitor Barrigão Gonçalves

Escola Superior de Educação do Instituto Politécnico de Bragança
vg@ipb.pt

Eurico Manuel Carrapatoso

Faculdade de Engenharia da Universidade do Porto / INESC Porto
emc@fe.up.pt

Resumo

Nestes últimos anos, a investigação rumo a uma Web mais inteligente estava dividida em duas correntes: por um lado, a Web Semântica e seus agentes inteligentes e, por outro, o Cérebro Global e seus agentes inteligentes. No primeiro caso, o objectivo reside em desenvolver linguagens para que as máquinas possam entender o conteúdo dos documentos que armazenam. No segundo, a ênfase vai para a auto-organização da Web de acordo com o comportamento dos utilizadores, recorrendo à analogia Internet/Cérebro. Embora muito haja ainda por fazer em cada uma das linhas de investigação, estamos convictos de que apostar na convergência entre as visões da Web Semântica e do Cérebro Global resultará numa Web mais inteligente. Por conseguinte, neste ensaio teórico são analisadas as características e implicações de cada uma destas visões de modo a perspectivar a articulação entre elas, em prol das organizações em geral e das instituições educativas em particular.

Palavras-chave: Web Semântica, Cérebro Global, Metadados, Ontologias, Agentes inteligentes.

Abstract

In recent years, research towards a more intelligent Web has been split into two currents: one, the Semantic Web and its intelligent agents, and the other, the Global Brain and its intelligent agents. In the first case, the aim is to develop languages for which the machines can understand the contents of documents that they store. In the second, the emphasis goes to the self-organization of the Web according to the behavior of users, using the analogy Internet / Brain. Although there is still much to be done in each line of research, we believe that convergence between the visions of the Semantic Web and the Global Brain will result in a more intelligent Web. Therefore, this paper analyzes the characteristics and implications of each of these visions in order to look for connections between them, on behalf of organizations in general and of educational institutions in particular.

KeyWords: Semantic Web, Global Brain, Metadata, Ontologies, Intelligent Agents.

Introdução

O sucesso da Internet deve-se essencialmente à crescente utilização da World Wide Web (WWW ou simplesmente Web) como principal serviço de comunicação e disseminação de informação. O seu crescimento envolve todas as áreas, desde a institucional à comercial e à educativa, podendo-se encontrar todo o tipo de informação na Web, o que obriga os utilizadores a despendem demasiado tempo para encontrar o que procuram. Mas, se a informatização dos documentos e a sua distribuição na Web visam facilitar a consulta e introduzir rapidez no processo de busca, porque é que isso não acontece?

A informação, outrora escassa, é hoje tão abundante que pode conduzir à desinformação. Esta avalanche de informação é cada vez mais preocupante (García, 2002).

Acreditar que dispor de toda a informação que queiramos é sinónimo de mais conhecimento seria uma quimera. Passar de uma sociedade da informação para uma sociedade do conhecimento que permita seleccionar e relacionar a informação imprescindível com vista à geração de conhecimento e de inteligência é talvez o desafio actual mais importante. Para tal, exige-se uma Web que “pense” connosco.

A Web actual é como que uma grande biblioteca desordenada porque não existem bibliotecários a estruturá-la e organizá-la ou, quando existem, nem sempre seguem as mesmas normas de estruturação e organização dos conteúdos. A Web Semântica será como um espaço repleto de recursos devidamente indexados e catalogados em que a estrutura de qualquer conteúdo terá significado tanto para os humanos, como para as máquinas.

Expressões, tais como: “inteligência conectiva” (Derrick de Kerckhove), “inteligência colectiva” (Pierre Lévy), “inteligência emergente” (Steven Johnson), “colectivos inteligentes” (Howard Rheingold), “redes inteligentes” (Albert Barabasi) e “cérebro global” (Francis Heylighen), têm sido cada vez mais usadas entre teóricos reconhecidos. Deliberadamente ou não, estes autores defendem que existem muitas semelhanças entre a Internet e o cérebro humano. Neste sentido podemos afirmar que a emergente Web Semântica contribuirá significativamente para aumentar essa semelhança e constitui o primeiro passo rumo a uma Web mais inteligente. Parece uma ideia para um filme? Talvez não, é uma visão optimista sobre o futuro da Internet e do impacto que a Web Semântica poderá representar.

Se lembrarmos a mitologia, o Deus Apolo resolveu construir um templo que servisse de oráculo para os homens. Falamos do Oráculo de Delfos que respondia a todo o tipo de perguntas dos humanos, como se de um extraordinário motor de busca inteligente se tratasse, proporcionando a resposta correcta à pergunta formulada (García, 2002). Analogamente, tem vindo a ser idealizado um novo oráculo no ciberespaço o qual, em vez de nos disponibilizar as centenas de possíveis respostas à questão submetida, tal como acontece nos actuais motores de busca, nos apresentará as respostas o mais possível filtradas e precisas, separando o trigo do joio.

Podemos afirmar que existem duas visões para converter a actual Internet numa Rede mais inteligente: uma via de investigação aposta num sistema que permita que a procura e recuperação de informação sejam mais intuitivas e precisas através da aplicação da semântica, de metadados, de ontologias e de agentes inteligentes (Tim Berners-Lee, Ora Lassila, Hans-Georg Stork e James Hendler, entre outros) e a outra baseia-se na possibilidade dos servidores poderem ser capazes de aprender o percurso que fazem os

cibernautas com vista a que a rede possa funcionar como um Cérebro Global (Francis Heylighen, Johan Bollen, Valentin Turchin e Cliff Joslyn, entre outros). Finalmente, com vista à simbiose desta duas visões no âmbito da estruturação e partilha do conhecimento, os Sistemas de Conhecimento Distribuído assumem particular importância. Embora diferentes, assim estas visões podem ser complementares.

Relacionar as duas linhas de investigação mencionadas de modo a perspectivar uma Web mais inteligente constitui o principal objectivo deste artigo. Num futuro próximo, localizar informação via Web com vista a gerar novo conhecimento poderá não ser suficiente para obter vantagens competitivas. Sabendo que a inteligência é a habilidade de utilizar eficazmente o conhecimento, a vantagem sobre os concorrentes estará na inteligência, pelo que o modelo para a Web deverá a pouco e pouco permitir gerar e recuperar não só conhecimento, mas também inteligência.

Assim, este ensaio teórico foi estruturado de modo a: i) caracterizar a Web actual e, face a algumas das suas principais limitações, perspectivar a próxima geração da Web – a Web Semântica; ii) descrever as principais ideias das abordagens “Web semântica” e “Cérebro Global”; iii) evidenciar a importância que as metalinguagens assumem nesta evolução; iv) relacionar os modelos das duas abordagens referidas, com vista a perspectivar uma infraestrutura de inteligência distribuída para a Web.

Da partilha de informação à partilha de inteligência

A Internet é uma rede de redes de computadores à escala mundial. Constitui a infraestrutura de suporte das denominadas auto-estradas da informação, onde dados circulam de um lado para o outro, as 24 horas do dia, os 7 dias da semana, os 365 dias do ano.

Actualmente, a Web assume-se como a melhor infra-estrutura para a comunicação, publicação, partilha ou disseminação de informação (conteúdos). Estes factores favoreceram o aparecimento de vários sítios de informação (Websites), desde as páginas Web pessoais até aos mais complexos Sistemas de Comércio Electrónico e de Ensino Aberto a Distância (EAD).

A Web pode ser vista como uma rede de servidores que veio tornar a Internet mais divertida e amigável. Estes servidores proporcionam o acesso a um conjunto de documentos ou páginas ligadas (através de hiperligações) que podem conter texto, imagens, gráficos, som, animação ou vídeo (Gonçalves, 2002). Aquando da sua apresentação, não se previa que a Web tivesse o impacto que teve na sociedade em geral. Tornou-se numa mediateca de

documentos à escala mundial, constituindo actualmente o maior repositório de informação que disponibiliza conteúdos e oferece serviços para praticamente todos os sectores da sociedade.

A evolução da Internet, e mais concretamente da Web, pode resumir-se em três gerações consequentes e cumulativas. Enquanto que a primeira geração da rede permitia apenas a troca de dados entre máquinas distintas sendo a Web resultado da criação manual de páginas em HTML (páginas estáticas), a segunda provocou uma revolução ao disponibilizar uma vasta gama de aplicações e informação para as pessoas (preocupação na apresentação de informação recorrendo a editores e utilizando modelos e folhas de estilo, geração automática de páginas Web baseadas em base de dados e interacção com o utilizador através de formulários e de mecanismos de busca - páginas dinâmicas e interactivas), tornando também possível o ensino/aprendizagem e o comércio electrónicos. A próxima geração da rede tem vindo a ser designada Web Semântica ou Inteligente (*Semantic Web*) apostando na estruturação de informação, na descrição da mesma através de metadados e no recurso a agentes inteligentes para apoiar o utilizador na realização de determinadas tarefas. A tabela 1 caracteriza sucintamente a evolução da Web rumo a uma Web semântica.

Característica	Antes da Web Semântica	Rumo à Web Semântica
Destinatários	Seres humanos	Seres humanos e computadores (agentes)
Linguagem principal	HTML (<i>HyperText Markup Language</i>)	XML (<i>eXtensible Markup Language</i>)
Estrutura dos conteúdos	Documentos não estruturados	Documentos estruturados
Exibição dos conteúdos	Dados e sua apresentação no mesmo documento	Dados num documento XML e apresentação num documento XSL por cada interface
Edição dos conteúdos	Documentos essencialmente estáticos	Documentos dinâmicos
Interactividade	Difusão unidireccional	Difusão bidireccional (Web editável)
Semântica dos conteúdos	Semântica implícita	Semântica explícita (metadados)
Decomposição e recomposição	Websites independentes	Agregação num Website de conteúdos de diversos Websites
Publicação	Centralizada (servidor Web)	Descentralizada (P2P)

Tabela 1 - Caracterização da Web

Face à proliferação de conteúdos na Web, surgiram mecanismos de pesquisa, cuja principal função é facilitar a exploração e recuperação de informação. Os mecanismos de pesquisa podem classificar-se em: a) *Directories* (listas, catálogos ou directórios, tais como:

Yahoo! e Lycos), onde a informação se encontra catalogada de acordo com as descrições das páginas fornecidas por quem as submete; b) *Search engines* (motores de pesquisa ou motores de busca de informação, tais como: Google e Altavista), que baseiam a busca de informação em palavras-chave, em robots que realizam o *download* das páginas Web, em indexadores que criam índices extraíndo automaticamente os termos chave das páginas ou em interfaces de consulta que comparam os termos recebidos com a base de termos indexados (Ramalho 2002; Sullivan 2002). Contudo, o constante aumento do número de servidores e de páginas Web faz com que a teia de informação assuma dimensões gigantescas, tornando difícil encontrar aquilo que se deseja ou necessita, mesmo recorrendo aos agentes dos motores de busca, pois o número de resultados devolvidos é exageradamente grande e impreciso. As bases de dados de alguns motores de pesquisa são construídas automaticamente com um tipo de programas ou agentes de software denominados rastejadores (*crawlers*) que analisam periodicamente a informação disponibilizada através de vários serviços da Internet e criam registos sobre o seu conteúdo.

Actualmente, os Websites podem disponibilizar mecanismos de pesquisa personalizados que, após a indexação dos conteúdos internos, facilitam a procura de informação nesse sítio Web, através de palavras-chave. Muitos Websites recorrem a bases de dados para armazenar grandes quantidades de informação de forma estruturada, de modo a que a localização de informação armazenada possa ser feita de forma rápida e eficiente. No entanto, os motores de busca externos ao Website não conseguem ter acesso aos conteúdos gerados a partir de uma base de dados, uma vez que não têm forma de perceber como os dados se encontram organizados, sendo difícil para um agente externo ao Website realizar inferências sobre as bases de dados.

Na perspectiva desse tipo de motores de busca externos ao Website, a grande maioria das bases de dados, embora armazenem muita informação, é em grande parte improfícua, porque está inacessível, profunda ou invisível, daí que essa parte da Web se designe de Web profunda ou invisível (*Deep Web*), para além de que muitas vezes a informação é vaga, ambígua ou simplesmente desactualizada.

A evolução da tipografia simples para SGML (*Standard Generalized Markup Language*), da HTML para a XML e, finalmente, para a RDF (*Resource Description Framework*) e OWL (*Ontology Web Language*) no âmbito da Web Semântica, tem sido um processo evolutivo com vista à interpretação dos dados armazenados, por parte das máquinas. As aplicações para a Web Semântica permitirão processar, reutilizar e partilhar a informação inteligentemente, com ou sem ajuda humana. A Web transformar-se-á assim numa enorme base de

conhecimento compartilhado, constantemente lida e escrita não só por humanos, mas também por aplicações semânticas.

Diversos autores defendem que existe uma grande similaridade entre o cérebro humano e a Internet. Se todos podemos estar ligados à Internet, poderá esta assumir-se como o cérebro do nosso planeta? E o seu comportamento colectivo será inteligente, podendo constituir uma mente global? Poderá esta forma colectiva de inteligência comparar-se à de uma mente humana individual, ou de um grupo de mentes humanas?

A Internet pode não estar organizada exactamente da mesma forma que o cérebro humano, mas podemos dizer que é uma extensão dele. McLuhan afirmou que as tecnologias podiam ser vistas como extensões dos nossos próprios membros e sentidos. Enquanto extensões de nós próprios, qualquer tecnologia exterioriza, amplifica e modifica diversas funções cognitivas. Consequentemente, podemos afirmar que a Web poderá vir a transformar-se numa extensão dos nossos sentidos e das nossas mentes, um cérebro global da sociedade em que vivemos. Tal como o cérebro inclui um sistema de memória onde pode armazenar, localizar e recordar informações, também a Internet disponibiliza espaços e protocolos para armazenar, partilhar e recuperar informações, favorecendo a cognição colaborativa de forma descentralizada. Este espaço de partilha de informação é como uma memória colectiva da mente global.

Podemos, então, afirmar que a Internet (tecnologia/hardware) tem vindo a evoluir no sentido de um cérebro global distribuído, e a sua actividade (agentes de software, seres humanos e dados) representa o processo cognitivo de uma mente global cada vez mais inteligente.

Cérebro Global

Em primeira instância, a ideia do “Cérebro Global” ou “Mente Global”, incide na comparação da sociedade com um organismo. Embora não seja um termo ou ideia recente, ganhou novo fôlego com o aparecimento da Internet e, em particular, da Web. Francis Heylighen e Johan Bollen (2003) têm à sua disposição a Web, um sistema de inteligência distribuída de carácter global, o que lhes confere uma grande vantagem sobre os seus antecessores, no sentido de desenvolver uma filosofia sistémica evolutiva da mente global. O projecto “Principia Cibernetica” (<http://pespmc1.vub.ac.be/>) materializa esta linha de investigação.

Podemos afirmar que um dos desafios actuais é converter a Web numa mente global, isto é, num sistema nervoso central composto de milhares de milhões de neurónios que se abastecem da inteligência distribuída da Rede. Cada ser humano só por si é uma fonte de

conhecimento para outro. Por conseguinte, a inteligência colectiva de todos os membros da Rede Internet constituirá o Cérebro Global.

Neste sentido, a aplicação da ideia do Cérebro Global à Web implicará a reconstrução das ligações de acordo com as necessidades do utilizador. O servidor Web do projecto Principia Cibernética, desenvolvido por Heylighen e Bollen, tenta materializar este modelo de cérebro global ao permitir a actualização e recomposição constante das suas ligações de acordo com as procuras do utilizador, eliminando as que vão perdendo actualidade pelo escasso ou nulo grau de consultas que recebem.

O servidor Web adjudica um agente inteligente a cada visitante cujo algoritmo permite detectar qual o percurso que este segue dentro da Web e memorizar o histórico do seu comportamento. Quando voltar a aceder ao servidor, o agente inteligente mostrará as páginas que supostamente interessam ao utilizador e ajustará a estrutura de ligações à esfera dos seus interesses. Se os agentes comunicarem entre si, poderão intercambiar informações com base no grau de utilização ou satisfação dos utilizadores sobre determinadas ligações.

A inteligência do cérebro global da Web não corresponderá apenas ao somatório das inteligências individuais da Rede, mas sim a uma inteligência colectiva em plena evolução.

Podemos afirmar que a Web é como uma criança que foi adquirindo informação, mas que continua ávida de conhecimento. Para que os agentes Web possam tomar decisões inteligentes, a Web precisa, primeiramente, de se ajustar aos interesses do utilizador para, posteriormente, favorecer a interpretação da informação disponível e a consequente geração de conhecimento.

O cérebro global parte do princípio de que a sociedade pode ser vista como um sistema baseado na inteligência colectiva de todos os seus membros, ou seja, uma estrutura que une cada nó (computador ligado à rede) como se fossem os neurónios de um cérebro (figura 1).

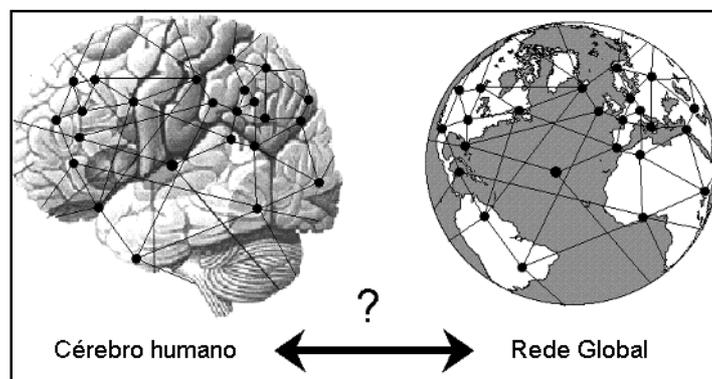


Figura 6 - Analogia entre Cérebro humano e Rede Internet (<http://pespmc1.vub.ac.be/Images/Brain-Earth.GIF>, acedido em 24/5/2005)

Esta ideia baseia-se no conceito de “inteligência partilhada”, segundo o qual a inteligência do grupo é maior que a soma das inteligências dos seus membros. A Internet viabiliza este conceito com uma dimensão muito maior do que qualquer outra infra-estrutura de comunicação anterior.

No âmbito do Projecto Principia Cibernética, o cérebro global está organizado como uma rede de associações, tal como se se tratasse de neurónios interligados que se desenvolvem mediante aprendizagem (Bollen e Heylighen, 2003). Sempre que se encontram as mesmas ligações repetidamente, o cérebro global aprende. Se pensar é a activação sucessiva de neurónios quando nos concentramos em determinados assuntos, e se essa activação provoca ligações com outros neurónios que, por sua vez, activam outros conceitos, então na Web também se podem estabelecer novas ligações entre páginas consoante o percurso efectuado pelo utilizador.

O funcionamento do cérebro global aplicado à estrutura da Internet pode traduzir-se na criação de novas hiperligações que se activam quando se detecta que o utilizador altera a sua intenção de pesquisa, de modo a que as ligações se actualizem ou eliminem automaticamente consoante o seu uso. Assim, as ligações da Internet na perspectiva do cérebro global assemelhar-se-ão à estrutura de um cérebro humano, porque se criam novas ligações em função do uso, tal como os neurónios. Segundo Francis Heylighen (2002), qualquer questão que surja por parte do utilizador deixará de ser uma tarefa árdua porque a rede organizar-se-á e adaptar-se-á ao que o utilizador espera dela.

O Cérebro Global pretende disponibilizar uma Web que “pensa e aprende”. Uma Web que quanto melhor nos conhecer, melhor saberá os nossos requisitos, mesmo que nós próprios não saibamos o que procuramos ou que palavras-chave necessitamos para o encontrar.

Observando o funcionamento do cérebro humano, sabemos que a recuperação de dados é flexível e eficiente, que a nossa massa encefálica é capaz de recolher da nossa memória a informação ou a resposta que necessitamos. Os nossos neurónios ligam-se entre si mediante sinapses (*synapses*) e a sua união reforça-se pelo uso. Este exemplo de aprendizagem baseia-se na chamada regra de Hebb: se dois neurónios se activam sucessivamente, o seu poder de interligação aumenta (figura 2).

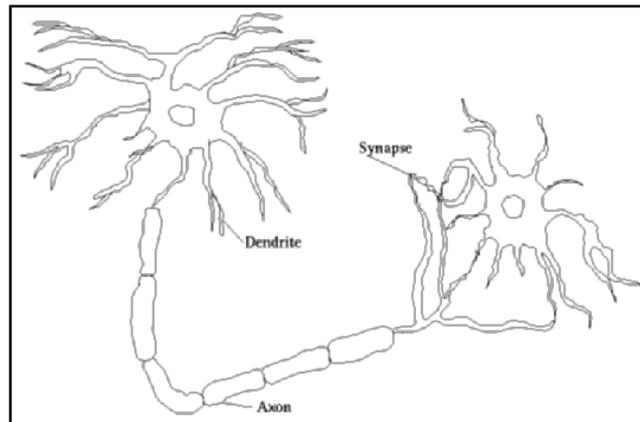


Figura 7 – Regra de Hebb (<http://pespmc1.vub.ac.be>, acessado em 24/5/2005)

Da mesma forma, se duas páginas Web são consultadas pelo mesmo utilizador num breve intervalo de tempo, a hiperligação existente entre elas assume um peso mais importante ou dá origem a uma nova ligação. Pelo contrário, ligações que não são consultadas assiduamente, podem desaparecer. Inconscientemente, estamos a seleccionar de entre os milhares de milhões de páginas Web aquelas que mais nos interessam.

Em suma, os agentes inteligentes recordarão as páginas Web mais usadas e, através da activação sucessiva de ligações, encontrarão outras páginas Web associadas ou relacionadas que nos recomendarão caso sejam do nosso interesse. Um exemplo básico disto pode ser já visualizado no Website *Amazon*: ao procurar um livro, obtemos também recomendações relacionadas com esse livro feitas por outros utilizadores, para além de críticas ou outras páginas relacionadas.

Em resumo, a Web transformar-se-ia num sistema inteligente que aprende com os seus utilizadores, reage aos seus desejos, reorganiza-se, interpreta, deduz e associa, enfim, pensa pelos seus utilizadores. Esta Web constitui assim um Cérebro Global, uma extensão do nosso próprio cérebro.

Cada acção do utilizador implica a reestruturação e reorganização da Rede, criando novas ligações. A Rede altera-se constantemente, tornando-a mais eficiente e inteligente. Desta forma, cada um de nós terá uma Internet personalizada.

Web Semântica

A Web é uma biblioteca global desordenada. Segundo Hans-Georg Stork (2002), a Web Semântica tem como objectivo ordenar o caos da Internet. Para tal, as máquinas têm que entender o conteúdo de uma página Web, interpretá-lo, analisá-lo e classificá-lo de acordo com critérios ou ontologias para que o utilizador, ao realizar uma pesquisa, obtenha informação o mais pertinente possível para a sua procura. Refira-se que uma ontologia é

uma forma de representar o conhecimento através de hierarquias elementares (Hendler, 2001).

A visão da Web Semântica baseia-se na possibilidade de estruturar e organizar a informação Web para que possa ser usada pelas máquinas, não apenas para a sua apresentação, mas para a sua compreensão, integração e reutilização entre várias aplicações (W3C, 2001).

Esta Web Inteligente é uma Web que saberá responder às nossas questões através de metadados, ontologias e agentes inteligentes. Tal como afirma James Hendler (2001), a Web está a ser transformada num conjunto de recursos que poderão ser entendidos e activados por agentes inteligentes e não apenas por humanos.

Para tornar possível esta visão, têm vindo a ser desenvolvidas linguagens e ferramentas, das quais se destacam a XML e a RDF (Daconta, 2003). A XML é usada para etiquetar páginas atribuindo uma estrutura aos dados e a RDF para descrever esses dados, atribuindo significado aos conteúdos.

Por exemplo, usamos a RDF para dizer que o termo “homem” e o termo “indivíduo” se referem ao mesmo conceito. Mas estas duas linguagens poderão não ser suficientes dadas as limitações inerentes às bases de dados, aplicações e outros recursos não-textuais. Para estabelecer relações entre recursos, necessitamos de recorrer a ontologias, tais como *RDF Schema* (RDFS) e *OWL* (Davies, 2003).

Uma vez que existem palavras que têm vários significados, as ontologias têm como função esclarecer o significado universalmente aceite para cada uma delas. Mas como explicar a uma máquina o significado de democracia se esta pode ser entendida de várias maneiras? Representar e relacionar todos e cada um dos conceitos existentes não é tarefa fácil, uma vez que podem variar de país para país e de contexto para contexto. Por exemplo, na nossa cultura, a cor preta é sinónimo de tristeza, enquanto que noutras culturas simboliza alegria.

Mas, não basta que as máquinas entendam os conteúdos textuais, também devem perceber o significado das imagens, sons, filmes ou outros conteúdos multimédia. Um quadro que seja considerado arte (por exemplo, mostrando mulheres nuas), não pode ser confundido com uma imagem pornográfica. Caso contrário, o material poderia vir a ser censurado sem ter em conta a sua qualidade artística.

Parafraseando Hans-Georg Stork (2002), a Web Semântica abarcará muito conhecimento, mas não todo. Portanto, embora as ontologias se limitem a áreas de interesse, serão numerosas e competirão entre si. Stork (2002) defende a multiplicidade de ontologias e assegura que é desejável a confrontação das mesmas.

A Web Semântica continuará a ser um espaço de comunicação tradicional, uma vez que, pelo menos no início, continuará a ser composta de “ilhas de confiança” (bibliotecas digitais, por exemplo) às quais se recorre para recuperar informação.

É aos agentes inteligentes que nos dirigiremos para perguntar. E a eles cabe a tarefa de procurar até encontrarem o que necessitamos, para além de poderem mostrar informação adicional verdadeiramente relacionada com o que procuramos.

A Web Semântica é a Web da linguagem, uma vez que se serve dela para se estruturar. É a Web das ontologias e do significado das palavras. É a Web que nos compreende ou, pelo menos aparentemente, é isso que nos parecerá. A figura 3 ilustra genericamente a visão da Web Semântica.

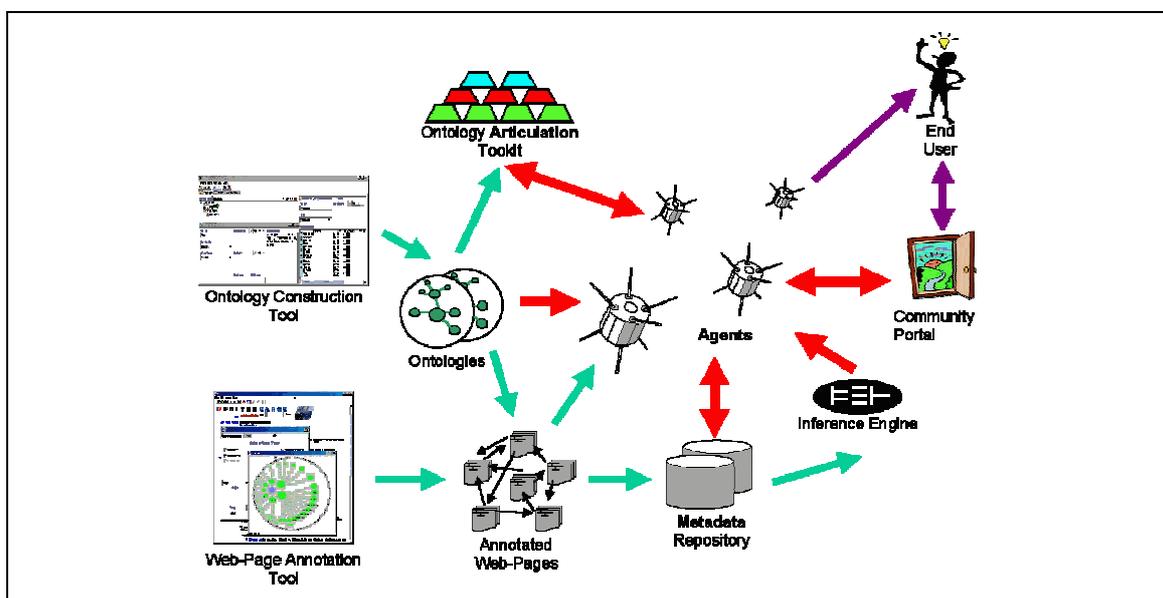


Figura 8 – Visão global da Web Semântica (<http://www.w3.org>, acessado em 24/5/2005)

Metalinguagens e Metadados

A Web Semântica favorece a comunicação formal a um nível mais elevado através da utilização de metalinguagens. As metalinguagens baseiam-se na utilização de metadados que permitem descrever o conhecimento indicando como interpretar e usar a informação. As metalinguagens, em particular as metalinguagens semânticas que descrevem as relações entre a informação e os sistemas de conceitos, estimulam a comunicação e o processamento da informação a um nível próximo do humano. A combinação das redes computacionais com as metalinguagens semânticas representa um passo importante na história da comunicação e da inteligência (Spivack, 2004).

A língua falada tinha como limitação principal a distribuição e o acesso à informação, uma vez que emissores e receptores teriam de estar fisicamente próximos para que

houvesse comunicação. Com a invenção da escrita passou a ser possível registar, representar, armazenar e comunicar ou partilhar o conhecimento independentemente da mente humana, derrubando, a pouco e pouco, barreiras temporais e espaciais. Com o aparecimento das redes de comunicações a comunicação tornou-se mais rápida, mais rica, mais alargada, mais acessível, tanto de forma síncrona como assíncrona.

A Web assume-se então como uma forma de distribuir e recuperar a informação, mas parte do processamento e interpretação da mesma mantém-se ainda nos seres humanos. Sem seres humanos a Web não seria nada mais do que uma colecção de informação relativamente estática e de aplicações isoladas. Mas, a Web Semântica poderá mudar este panorama, fazendo com que a inteligência funcione fora do cérebro humano, ao fornecer metalinguagens que dotem a Web de uma inteligência aparente que permita interpretar a informação existente e gerar nova informação e conhecimento.

Por conseguinte, o uso de metalinguagens semânticas possibilitará não só a distribuição da informação, mas também do conhecimento e da inteligência (entenda-se por inteligência o acto de processar e de usar a informação com vista a interpretá-la e compreendê-la tendo por objecto o conhecimento). As metalinguagens semânticas permitem expressar, distribuir e partilhar formalmente o conhecimento necessário à interpretação e utilização inteligente da informação, independentemente dos receptores da informação serem humanos ou máquinas. Ou seja, as linguagens escritas permitem apenas a portabilidade da informação, mas as linguagens semânticas possibilitam a geração de conhecimento (sistemas conceptuais) e mesmo de “inteligência” (isto é, procedimentos para processar o conhecimento) com base nessa informação portátil.

Desde que o software use estas metalinguagens para expressar o conhecimento sobre um qualquer domínio específico em metadados, passa a ser possível importar e usar esse mesmo conhecimento sempre e quando for requerido.

O aparecimento de normas para partilhar metadados semânticos que descrevem o significado da informação tem vindo a impulsionar uma nova era do conhecimento (e da inteligência) distribuído na Internet.

Num futuro não tão distante quanto possa parecer, a Internet tornar-se-á mais “inteligente”. Hoje a maior percentagem de inteligência reside nos cérebros humanos, dentro de alguns anos, essa percentagem será igualada, ou mesmo superada, pela que existirá fora dos cérebros humanos, ou seja, na Web semântica.

Rumo a uma inteligência colectiva distribuída

A Internet é o cérebro e a Web é a sua mente (Spivack, 2004). A ideia do cérebro global assume que cada um de nós faz parte dele. A novidade está em adicionar as máquinas, ou melhor, os agentes de software, a esta realidade, dotando-os de capacidades de raciocínio próximas das dos humanos.

Por conseguinte, cada um de nós, inconscientemente ou não, faz parte desta mente global. A mente global está emergindo em torno de nós, e através de nós, é a nossa criação, mas pode ganhar alguma independência, desde que os agentes de software interpretem metadados e consigam dialogar e colaborar na resolução de tarefas sem intervenção humana directa (Spivack, 2004).

A evolução da inteligência humana individual e colectiva tem sido um processo natural que dura já há alguns milhares de anos. A Web Semântica constitui apenas mais uma etapa neste processo evolutivo, uma etapa que visa habilitar as tecnologias da capacidade de interagir sem intervenção humana, de forma a assistir os humanos, libertando-os de tarefas morosas, repetitivas e entediantes, contribuindo significativamente para a evolução da inteligência planetária. Ou seja, com a Web Semântica é possível que os processos cognitivos comecem a ocorrer fora do cérebro humano. Numa visão mais futurista, a Web evoluirá para uma estrutura que excederá a nossa compreensão, ou seja, uma mente colectiva, da qual cada um de nós será apenas mais uma das suas partes.

Tal como as colónias de formigas e abelhas, que têm comportamentos colectivos “inteligentes” sem controlo centralizado, também os seres humanos e os agentes de software, que vagueiam aleatória e independentemente na Web Semântica, lendo, escrevendo, anotando, analisando, avaliando e agregando informação, deverão ter comportamentos inteligentes colectivos sem que os processos tenham que ser coordenados por alguns deles.

O valor da inteligência colectiva distribuída pode ser demonstrado pelas comunidades de desenvolvimento de software *open-source*, uma vez que desenvolvem melhor código, mais rápido, com menor custo e com uma adopção mais ampla do que aquela que é possível em entidades privadas (Spivack, 2004).

A infra-estrutura da inteligência distribuída

As arquitecturas de sistemas de conhecimento distribuído são desenvolvidas com base em normas e tecnologias Web para permitir a integração de (e acesso uniforme a)

componentes de sistemas inteligentes, possibilitando que os agentes interajam com os recursos de informação existentes na rede.

Tais arquitecturas permitem que um “computador A” possa recuperar o conhecimento de um “computador B”, ou seja, se usar agentes inteligentes, o “computador A” não recupera apenas informação, mas sim conhecimento gerado a partir da associação das informações existentes no “computador B”. Atendendo a que a acção de pensar e de compreender implica a activação de neurónios e a associação entre conceitos, estes sistemas de conceito distribuído devem ser capazes de entender tais conceitos. Ou seja, devem ser capazes de entender o conteúdo das páginas Web para poderem proporcionar informação precisa ao utilizador. É aqui que entra a Web Semântica e os agentes inteligentes.

Se a Web é literalmente o repositório da informação mundial, então ela constitui a memória da mente global. O desenvolvimento desta mente global dependerá da evolução dos sistemas distribuídos. A Web constitui, portanto, o catalizador chave para esta evolução. A inteligência distribuída requer a descentralização da informação e do seu processamento. A Web disponibiliza normas universalmente aceites para publicar e recuperar informação. Mas, sem integração explícita, as aplicações têm grande dificuldade em interoperar ou compreender os dados. A interoperabilidade e a descentralização facultadas pela HTML e HTTP (*HyperText Transfer Protocol*) não são suficientes para atingir um nível de eficiência e eficácia aceitáveis. Para que a inteligência distribuída possa ser uma realidade, é necessário ir mais longe e tornar possível que os agentes entre si possam publicar, trocar e usar a informação de e em qualquer lugar do sistema.

Podemos destacar 2 passos importantes rumo à formação da mente global suportada por uma infra-estrutura distribuída:

1) Se a Web é o repositório de informação do mundo, a XML permite encará-la como uma base de dados mundial. A XML permite definir, armazenar, recuperar, interagir e interpretar estruturas de dados com precisão. Com a XML qualquer sintaxe e esquema de dados pode ser definido e partilhado. A XML fornece uma estrutura à informação Web (à memória da mente global) que permite um acesso e manuseamento dos conteúdos mais eficiente e eficaz pelos agentes.

2) A emergente Web Semântica adiciona uma outra camada para além da XML. Permite que os agentes de software comecem a compreender e raciocinar sobre o significado da informação. Os agentes passam a trabalhar não apenas com dados, mas também com conceitos. Entenda-se conceitos como estruturas de informação ligadas aos sistemas

formais de conhecimento ou de ideias, isto é, informação com significado. A Web Semântica fornece normas para transformar estruturas ordinárias de informação em conceitos que podem ser compreendidos por aplicações ou agentes de software. O uso de metalinguagens para definir a semântica (RDFS e OWL) possibilita ligar elementos de dados aos conceitos formalmente definidos em sistemas de conhecimento expressos em ontologias. A etiquetação dos dados com base em ontologias (a ligação de dados a ontologias) permitirá que os agentes possam raciocinar inteligentemente sobre a informação.

Por exemplo, os conteúdos de uma disciplina de informática incluída num sistema de e-Learning poderiam estar ligados a uma ontologia de informática que definisse conceitos de hardware e software e suas relações. Com base nesta ontologia seria possível realizar pesquisas semânticas nos materiais educativos bem mais inteligentes do que as procuras através de palavras-chave, usadas actualmente na grande maioria dos motores de busca. Uma pesquisa semântica para “informação sobre a memória interna de um sistema informático” retornaria páginas ou registros de dados sobre as memórias ROM e RAM, mesmo que esses termos não fossem explicitamente submetidos.

Contudo, a realização de pesquisas mais inteligentes, apontando eficiente e eficazmente para os requisitos do utilizador, são apenas um dos muitos benefícios mais visíveis da Web Semântica.

A Web Semântica pode proporcionar também que as aplicações (ou os agentes de software que as compõem) aprendam, raciocinem, sugiram e controlem automaticamente tarefas e processos mais inteligentemente. Assim, passa a ser exequível a inteligência artificial distribuída na Internet, não só porque a computação dentro de um programa é mais inteligente, mas também porque isso ocorre entre os programas, sendo possível a partilha de conceitos e a colaboração estreita em tarefas do raciocínio.

A utilização de metalinguagens (ou de metadados) adiciona camadas de conhecimento adicional à informação. Estas camadas de informação podem fornecer definições, ligações a outros recursos, informação sobre a organização da informação dentro do original, relações lógicas entre conceitos, anotações de outros leitores, bem como regras para interpretar, raciocinar ou usar a informação. Em suma, os metadados fornecem pistas para que as máquinas possam compreender a informação. Os metadados semânticos permitem que os computadores tenham acesso a camadas de conhecimento que anteriormente só poderiam existir ou ser utilizadas no cérebro humano. A Web Semântica favorece não

somente a comunicação eficaz entre seres humanos, mas também entre humanos e agentes de software, e entre agentes de software.

Conclusão

A primeira geração da Rede preocupou-se com a implementação da infra-estrutura tecnológica da Internet e dos respectivos serviços, a segunda incidiu na construção de aplicações Web e, actualmente, a terceira avança rumo a uma Web mais inteligente.

Conhecimento não é informação. Conhecimento é o valor construído diariamente quando o focalizamos, é a análise e a acção da informação. Mas, deter informação e conhecimento poderá já não ser suficiente, a vantagem está na habilidade de utilizar eficazmente esse conhecimento (inteligência).

A Web não pode apenas armazenar a inteligência distribuída dos seus utilizadores, deverá conseguir aprender com o que os utilizadores fazem e comportar-se de acordo com essa aprendizagem.

Nesta perspectiva, a visão da Web Inteligente percebe-se em duas visões diferentes, mas complementares. Uma refere-se à analogia da Internet com o cérebro humano, a outra à capacidade dos servidores Web compreenderem o significado de todos os conteúdos que armazenam e distribuem pela Rede, independentemente da sua forma (texto, imagem, gráficos, áudio ou vídeo). Para que se possa compreendê-los, será necessário estruturá-los e organizá-los semanticamente. Portanto, as metalinguagens semânticas assumem particular importância neste contexto. Mas, para que os servidores possam negociar entre eles os conteúdos Web, tornar-se-á necessário desenvolver sistemas de agentes inteligentes. Referimo-nos pois ao projecto da Web Semântica.

A Web Semântica possibilita que os servidores compreendam o conteúdo que armazenam e disponibilizam na rede. Os agentes inteligentes tratam da recuperação da informação relevante aos utilizadores que dela necessitam e assistem-nos na tomada de decisão.

De referir que, no âmbito da estruturação do conhecimento, os Sistemas de Conhecimento Distribuído em articulação com a Web Semântica constituem uma peça essencial para o desenvolvimento da arquitectura para uma Web Inteligente.

Mas, se for possível perceber o comportamento do utilizador, a Web pode personalizar-se de forma a colocar em primeiro plano aquilo que supostamente o utilizador necessita. Assim, todos estes esforços apontarão para o funcionamento da Rede segundo um modelo “Cérebro Global”. Ou seja, a Web Semântica viabiliza a ideia do Cérebro Global.

O Cérebro Global não é mais do que uma metáfora para caracterizar uma rede inteligente formada por pessoas que detêm conhecimento e tecnologias de comunicação que permitem a interligação de todos (Heylighen 2002).

Se os problemas se resolvem pensando, isto é, activando neurónios, então os resultados na Web Inteligente (segundo a perspectiva do Cérebro Global) obtêm-se activando conceitos relacionados entre si.

Através de um agente inteligente, é possível analisar o percurso de um utilizador através da Rede, e em função desse percurso, deduzir o que lhe interessa. A ideia é observar as ligações que um utilizador efectua entre as páginas Web. Consequentemente, a Web estará a aprender à medida que o utilizador navega, transformando-se numa rede associativa. O algoritmo responsável por essa aprendizagem reforça as ligações que mais se utilizam em detrimento das que menos se utilizam.

Mas a ideia do Cérebro Global não se deverá esgotar aqui. Esta é apenas uma forma dos agentes inteligentes se aperceberem do perfil dos utilizadores (os seus gostos e preferências). O cérebro global deverá ser fruto da inteligência de todos os utilizadores da rede, beneficiando não só cada um individualmente, mas também os grupos, tais como empresas ou escolas.

Porém, tal como questiona Pierre Lèvy (1999): Seremos capazes de nos mobilizar a favor de um projecto de sociedade centrado na inteligência colectiva que favoreça a partilha livre de conhecimentos? Teremos que ser, já que é um requisito para que as empresas e as escolas possam gerar vantagens competitivas para lidar com as exigências da sociedade actual. Os alunos e os professores necessitam que a Web se transforme não só num ambiente de conhecimento distribuído em prol da inteligência colectiva, mas também num parceiro da aprendizagem, acompanhando e facilitando os processos de ensino/aprendizagem. Se cada instituição educativa descrever e indexar os seus conteúdos através de metadados e ontologias e os disponibilizar internamente ou externamente estará a contribuir para acelerar a migração da Web actual para a Web Semântica e, consequentemente, criar as condições para uma Web mais inteligente baseada na abordagem do “Cérebro Global”.

Referências bibliográficas

Berners-Lee, T., Hendler, J., & Lassila, O. (2001, May). The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. New York: Scientific American, 284(5):34-43.

- Bollen, J. e Heylighen, F. (2003). Learning Web. In: Heylighen, Joslyn and Turchin (editors): *Principia Cybernetica Web*, Brussels: Principia Cybernetica.
- Daconta, Michael; OBRST, Leo e SMITH, Kevin T. (2003). *The Semantic Web - A guide to the future of XML, Web Services and Knowledge Management*, Wiley Publishing, Inc..
- Davies, J.; Fensel, D.; Van Harmelen, F. (2003). *Towards The Semantic Web – Ontology-Driven Knowledge Management*, John Wiley & Sons, Ltd.
- García, R. (2002). Búsqueda de información en Internet. In *Educar en Red - Internet como recurso para la educación*, pág. 83-100, Ediciones ALJIBE, SL.
- Gonçalves, Vitor (2002). *Desenvolvimento de Sistemas de Informação para a Web*, Dissertação de Mestrado em Tecnologia Multimédia, Universidade do Porto.
- Heylighen, Francis (2000). The Social Superorganism and its Global Brain. In: Heylighen, Joslyn and Turchin (editors): *Principia Cybernetica Web*, Brussels: Principia Cybernetica.
- Heylighen, Francis (2002). Conceptions of a Global Brain: an historical review. In: Heylighen, Joslyn and Turchin (editors): *Principia Cybernetica Web*, Brussels: Principia Cybernetica.
- Hendler, James (2001). *Agents and the Semantic Web*, IEEE Intelligent Systems Journal.
- Johnson, Steven (2001). *Emergence: the connected lives of ants, brains, cities, and software*. Nova Iorque: Scribner.
- Kerckhove, Derrick (1997). *Connected intelligence*. Toronto: Somerville House.
- Lévy, Pierre (1998). *A Inteligência Coletiva*. São Paulo: Loyola.
- Mcluhan, Marshall (1964-1969). *La comprensión de los medios como las extensiones del hombre*, Editorial Diana.
- Ramalho, Franklin de Souza (2002), *Web Semântica*, I Workshop de Dissertações da COPIN, UFCG / UFCG.
- Rheingold, Howard (2002). *Smart mobs: the next social revolution*, Perseus.
- Spivack, Nova (2004). *Minding the Planet: From Semantic Web to Global Mind*, Radar Networks.
- Stork, Hans-Georg (2002). The Semantic Web – A Web of knowledge and services, In *Workshop on DB/IS Research for Semantic Web and Enterprises*, 3 - 5 Abril de 2002.
- Sullivan, Danny (2002). How Search Engines Work, *Search Engine Watch*, Out 2002.
- W3c, WWW Consortium (2001), *Semantic Web*. Disponível em:
<http://www.w3.org/2001/sw/>