

5.

A Web dos significados

“Esperanzas y utopías aparte, acaso lo más lúcido que sobre el lenguaje se ha escrito son estas palabras de Chesterton: ‘El hombre sabe que hay en el alma tintes más desconcertantes, más innumerables y más anónimos que los colores de una selva otoñal...cree, sin embargo, que esos tintes, en todas sus fusiones y conversiones, son representables con precisión por un mecanismo arbitrario de gruñidos y de chillidos. Cree que del interior de un bolsista salen realmente ruidos que significan todos los misterios de la memoria y todas las agonías del anhelo.’” (Jorge Luis Borges, El idioma analítico de John Wilkins. In: *Otras inquisiciones*, 1996)

Com a Revolução Digital, a discussão a respeito da busca por uma língua única de comunicação cedeu lugar ao debate sobre as inúmeras possibilidades de comunicação. A língua utilizada durante a CMC não parece ser tão relevante quanto a própria comunicação, principalmente durante o segundo e o terceiro momentos da Rede considerados momentos multilíngues e chamados, respectivamente, de Web 2.0 e Web 3.0 (ou Web Semântica). Para que estes momentos sejam mais bem compreendidos, é necessário retornar ao passado, isto é, ao primeiro momento da Internet, chamado de Web 1.0⁸¹.

5.1.

O primeiro momento da Rede

Os impactos sociais e pessoais produzidos pelo surgimento da Internet no mundo na década de 1990 são comparados às mudanças geradas pela Revolução Industrial a partir do século XVIII. Nicolaci-da-Costa (1998) explica que a máquina a vapor teve o poder de desencadear uma revolução que, além de alterar os meios de produção, alterou também as formas de viver, de sentir, de se relacionar com o mundo, com os outros e consigo mesmo. Ela prossegue afirmando que:

⁸¹ É importante ressaltar que esta investigação não tem como foco explorar os sistemas operacionais utilizados nos três momentos da Internet.

“... Com o aparecimento dos trens movidos a vapor..., ganhou-se uma mobilidade nunca antes sonhada, o que acabou resolvendo o problema da escassez de alimentos que, agora, podiam ser transportados com rapidez de terras férteis porém longínquas. Com o aparecimento das indústrias, a população do campo deslocou-se para as sedes das mesmas e isso resultou no surgimento das grandes metrópoles com seus parques industriais. Emergiu o capitalismo com suas linhas de montagens e longas horas de trabalho competitivo.” (Nicolaci-da-Costa, 1998, p. 5)

Certamente, a consequência de tais mudanças provocaram alterações radicais na vida das pessoas daquela época, principalmente daquelas oriundas das comunidades rurais que, ao fixarem residência nas grandes metrópoles, passaram a viver no meio de estranhos e a se sentir solitárias. A Revolução Industrial fez surgir novas formas de ver o mundo, de viver e de sentir. A Revolução Digital não foi diferente. Nicolaci-da-Costa (2002) nos mostra que a Internet gerou um novo espaço, equivalente ao que a metrópole representou para a Revolução Industrial, chamado ciberespaço. Este é descrito por Lévy (1999, p. 92) como: “... espaço de comunicação aberto pela interconexão mundial dos computadores e das memórias dos computadores.” O ciberespaço também gerou e continua gerando novas formas de organização social e psicológica (Nicolaci-da-Costa, 2005a).

O primeiro momento da Rede, conhecido como Web 1.0, caracterizou-se pela emergência de sites e grandes portais. Inúmeros jornais, revistas, estabelecimentos comerciais, empresas, bibliotecas, escolas, universidades, etc passaram a disponibilizar informações, produtos e serviços para os usuários da Rede. Estes sites e grandes portais eram confeccionados por pessoas especializadas em tecnologia e com conhecimento de programação. Além disso, durante a década de 1990, surgiram também os sites de busca como os americanos AltaVista (www.altavista.com – 1996) e Google⁸² (www.google.com – 1998) e os brasileiros Aonde (www.aonde.com.br – 1993) e Cadê (www.cade.com.br – 1995).

Devido às dificuldades na confecção dos sites e portais, os conteúdos disponibilizados na Web 1.0 eram controlados pelos programadores, designers e donos dos sites e portais. Também era possível encontrar páginas pessoais, nessa primeira fase da Internet. No entanto, para que estas fossem confeccionadas, era necessário o conhecimento das linguagens de programação

⁸² O Google é atualmente o site mais acessado de toda a Rede e não se restringe mais a uma ferramenta de busca. Vários outros serviços são disponibilizados como mapas (*Google Maps*), email (*Gmail*), ferramenta para criar *blogs* (*Blogger*), *chat* (*Google Talk*), ferramenta para criar e disponibilizar álbuns de fotos (*Picasa*), rede social (*Orkut*), site para publicação de vídeos (*YouTube*), tradutor automático (*Google Translator*), dentre outros. Para as estatísticas da Rede verificar o site <http://www.alexa.com/topsites>.

ou a contratação de profissionais especializados. Sendo assim, a Web 1.0 pode ser caracterizada como a Web das publicações eletrônicas feitas por alguns especialistas em tecnologia para muitos usuários.

Porém, de forma menos sofisticada do que em suas versões mais recentes, uma das principais características da Web - a interatividade – já se achava presente na Web 1.0. No início dos anos de 1990 era possível utilizar programas de bate-papo on-line e em tempo real como IRC (*Internet Relay Chat*) ou ICQ (*I Seek You*)⁸³. Ferramentas como estas revolucionaram as formas de relacionamento e comunicação entre as pessoas. Consequentemente, como já explorado no capítulo 4, surgiu um novo tipo de linguagem: o *netspeak*.

Os impactos gerados pela chegada da Web não foram pequenos. O susto, temor e preconceito iniciais deram lugar à curiosidade, atração e excitação por parte de seus usuários (Nicolaci-da-Costa, 2006). As possibilidades ilimitadas oferecidas pela Rede de conhecer tudo e interagir com todos atraíram milhões de usuários. Em pouco tempo, uma nova plataforma de vida havia sido criada, isto é, um espaço alternativo integrado ao mundo “real”. De acordo com Nicolaci-da-Costa (2006):

“... a vida que acontece nesse espaço “virtual”, à semelhança daquela que se desenrola no espaço “real”, está em constante transformação. Essa mesma vida “virtual” atua ainda como uma força transformadora da vida “real”. E, certamente, o mesmo acontece no sentido inverso, tendo em vista que a plataforma “virtual” é criada e mantida por homens concretos e máquinas concretas que habitam o mundo físico.” (Nicolaci-da-Costa, 2006, p. 35)

Começam a surgir os nativos digitais (Prensky, 2001), que, ao contrário dos imigrantes digitais, nasceram após a disseminação das tecnologias digitais e da Web e veem todas essas novidades como naturais⁸⁴. Nicolaci-da-Costa (2009b) descreve essas duas gerações da seguinte forma:

“Os membros [dessas duas gerações]... já foram batizados de “nativos digitais” e “imigrantes digitais”. Segundo Prensky (2001), que cunhou essa nomenclatura, “nativos digitais” são aqueles jovens que já cresceram (ou estão crescendo) em um mundo dominado por vídeo games, computadores, internet, celulares, etc. São aqueles para quem a linguagem digital e a concepção de mundo por ela difundida parecem tão naturais quanto sua língua materna. São aqueles que, entre suas características subjetivas marcantes, ostentam as seguintes: a constante disponibilidade para contatos online e via celulares; a auto-exposição online de gostos, sensibilidades e talentos; a fluidez de movimentação virtual e

⁸³ O IRC não é mais utilizado atualmente. O ICQ ainda está disponível em mais de 10 línguas, mas perdeu muitos usuários após a popularização do *MSN Messenger*. Para mais detalhes sobre o ICQ verificar em: <http://www.icq.com/>.

⁸⁴ Os nativos digitais são conhecidos também como geração Y e os imigrantes digitais como geração X.

física; a expectativa de livre acesso à informação de todos os tipos; a ausência de planejamento; o imediatismo e a pouca tolerância à espera ou ao adiamento de gratificação... Já os “imigrantes digitais”, ainda segundo Prensky, são os homens e mulheres que se assemelham aos imigrantes tradicionais que nunca chegam a dominar com perfeição a língua nativa dos países para os quais migraram (sempre mantendo algum sotaque) nem a se liberar completamente das crenças e costumes de suas culturas de origem. Eles podem ter recebido bem as inovações digitais, podem dominar e usar cotidianamente várias delas, mas nunca deixarão de sofrer a influência de seu passado analógico.” (Nicolaci-da-Costa, 2009b, p.237)

O caráter dinâmico e transformador da Rede é inegável, assim como seu constante e cada vez mais rápido crescimento e aperfeiçoamento. Parece difícil até para os nativos digitais acompanharem tantas mudanças e inovações. A Internet vem crescendo e se aperfeiçoando muito rapidamente. Novos sistemas operacionais e ferramentas fazem com que os usuários se tornem cada vez mais participativos na Rede e esta cada vez mais multilíngue, tendo em vista que inúmeras pessoas das mais diferentes partes do mundo a ela vêm se conectando. A evolução da Web 1.0, que já era tão revolucionária e transformadora, fez surgir a segunda fase da Rede: a Web 2.0.

5.2. A Web 2.0

Enquanto a primeira fase da Internet pode ser caracterizada pela centralização e controle da informação por aqueles que dominavam a tecnologia, a Web 2.0 é conhecida como a Web colaborativa e participativa, a Web social.

O termo Web 2.0 surgiu em 2004 quando Tim O'Reilly⁸⁵ e John Battelle⁸⁶ criaram a *Web 2.0 Conference*, chamada hoje de *Web 2.0 Summit*⁸⁷. Esta conferência foi criada para discutir e apresentar novidades tecnológicas e o termo Web 2.0 tentava marcar a emergência de uma nova concepção para a Web. A Web 2.0, contudo, não surgiu em 2004, mas no final da década de 1990 e início dos anos 2000, como poderá ser observado nas próximas seções.

⁸⁵ Tim O'Reilly é o fundador e CEO do O'Reilly Media, Inc. Esta empresa americana organiza grandes eventos que têm a tecnologia como tema principal. Dentre eles estão: *Web 2.0 Summit*, *Web 2.0 Expo*, *O'Reilly Open Source Convention*, *Gov 2.0 Summit* e *Gov 2.0 Expo*. A visão da O'Reilly Media, Inc. é disseminar conhecimento e apresentar inovações sobre tecnologia. Para mais detalhes verificar: <http://tim.oreilly.com/>.

⁸⁶ John Battelle é jornalista e fundador do *Federated Media Publishing*. Esta empresa americana ajuda sites independentes a se inserirem no mercado. Para mais detalhes verificar: <http://www.federatedmedia.net/about/index>.

⁸⁷ Para mais informações sobre a conferência de 2009 acessar: <http://www.web2summit.com/web2009>.

Tim O'Reilly alega que o segundo momento da Rede teve início com o *boom* da filosofia pontocom que, de acordo com ele, começou por volta de 2001 com a expansão de negócios on-line, dos sites de busca, dos sites de relacionamentos e, principalmente, das redes sociais. A filosofia básica da Web 2.0 é a Web como plataforma, ou seja, como um lugar, ambiente, e não somente como uma Rede de computadores interligados. A Internet não sofreu grandes alterações em seus sistemas operacionais a partir da Web 2.0, o que mudou foram as maneiras de encará-la e de usá-la, foi uma mudança mais social do que tecnológica. A Rede passou a ser colaborativa e coletiva. É importante descrever com alguns exemplos como isso passou a acontecer, para que seja possível ter uma visão geral da segunda geração da Rede e dos motivos pelos quais esta fase vem se tornando cada vez integrada, apesar de multilíngue.

5.2.1.

O início da Web colaborativa

A segunda geração da Internet se caracteriza principalmente pela possibilidade que os usuários têm de participar ativamente na produção de conteúdo na Rede. O que anteriormente era feito por empresas ou pessoas especializadas em tecnologia (primeira geração da Internet), passou a poder ser realizado pelo próprio usuário. Este passou a gerar, alterar e compartilhar conteúdo na Rede – textos, informações, imagens, arquivos de áudio e vídeo. As possibilidades, facilidades e liberdade que a Web 2.0 proporciona aos usuários fazem com que ela esteja aberta a diversas culturas e línguas diferentes, isto é, seja cada vez mais democrática. Tal abertura reforça o caráter multilíngue da Rede, citado no capítulo anterior.

É possível afirmar que um dos fatores responsáveis pelo início da Web colaborativa foi o surgimento dos *blogs*. Estes começaram a surgir em 1997 e podem ser caracterizados como espaços textuais que permitem que qualquer pessoa com acesso a Internet publique, leia, interaja com escritores e/ou leitores, bem como comente textos, fotos, arquivos de áudio e vídeo sem necessariamente dominar as linguagens de programação. Em uma pesquisa realizada em 2005, revelo que:

“... o termo *weblog*, posteriormente reduzido para *blog*, foi criado por Jorn Barger, editor do *site Robot Wisdom* (www.robotwisdom.com), em 1997. Jorn Barger explica em seu *site* pessoal que um *weblog*, às vezes chamado de *blog*, página de notícias ou filtro, é uma página da *web* onde um *weblogger*, também

chamado de *blogger*, blogueiro ou pré-surfista, *logs* (registra por escrito uma viagem ou evento, como em um diário de bordo) outras páginas da *web* e textos que considera interessantes. De acordo com Barger, o blogueiro adiciona a publicação mais recente, também chamada de *post*, no topo da página. Abaixo ou acima do *post*, podemos encontrar a data e a hora da publicação. Além disso, também é comum encontrar abaixo de cada texto publicado o nome ou o apelido do blogueiro. Desta forma, os leitores podem acompanhar o blog lendo as publicações de forma cronologicamente inversa, ou seja, sempre da publicação mais recente para a mais antiga.” (Di Luccio, 2005, p.42)

Das simples publicações de sugestões e *links* para outras páginas na Rede, como mencionado acima, os *blogs* evoluíram para os diários pessoais, devido às facilidades de confecção e manutenção destes, tendo como consequência a sua rápida popularização.

Além disso, as seções de comentários contribuíram e continuam contribuindo para o grande sucesso dos *blogs*. Abaixo de cada publicação, pode ser encontrado um *link* chamado *comments* ou comentários. Este *link*, ao ser clicado, abre uma janela na qual os leitores podem fazer comentários, críticas ou sugestões a respeito dos textos ou imagens publicados pelos blogueiros. Nesta mesma janela, os escritores também podem responder aos comentários dos leitores e interagir com eles. Os comentários vão sendo acumulados em ordem cronologicamente inversa e podem ser lidos não somente pelo escritor do *blog*, mas também por qualquer usuário que clicar no *link* e abrir a janela dos comentários. Ainda na página principal dos *blogs* é possível encontrar uma breve descrição do blogueiro e também vários *links* que, ao serem clicados, abrem novas páginas. Os *links* podem nos levar para arquivos de outros meses do próprio blogueiro, para seus *sites* favoritos e para outros *blogs* ou páginas mantidas pelo próprio blogueiro. Podemos encontrar também *links* que exibem a tradução do *blog* para outras línguas (grande parte das vezes utilizando a ferramenta de tradução automática *Google Translator*), o e-mail e o currículo do blogueiro, os *blogs* de seus amigos e seus *blogs* favoritos. Alguns blogueiros inserem também *links* que nos direcionam para suas galerias de fotos (*Fotolog*, *Flickr* ou *Picasa*)⁸⁸. Os *links* que nos levam a outros *blogs* contribuem para a construção de uma rede de blogueiros conhecida como blogosfera. As possibilidades oferecidas pelos *blogs* de relacionamentos entre blogueiros, leitores e inúmeros *blogs* viabilizam o estabelecimentos de laços sociais entre

⁸⁸ O *Fotolog* (<http://www.fotolog.com.br/>) e o *Flickr* (<http://www.flickr.com/>) são sites de hospedagem e compartilhamento de fotos, sendo que o primeiro é pouco utilizado nos dias de hoje em relação ao segundo. Ambos podem ser caracterizados como *blogs* de fotos. Já o *Picasa* (<http://picasa.google.com.br/>) é um programa do Google que pode ser baixado pelo usuário e permite que este organize e edite suas fotos.

blogueiros e leitores, fazendo dos *blogs* redes sociais importantes no dias de hoje (Recuero, 2009).

Já se passaram mais de 10 anos desde o aparecimento dos primeiros *blogs* na Rede. Estes evoluíram (tecnologias e usos) consideravelmente e se mantiveram populares até os dias de hoje. É possível encontrar *blogs* jornalísticos, políticos, literários, empresariais, tecnológicos, automobilísticos, esportivos, religiosos, financeiros, de guerra, de protestos, de moda, de entretenimento, de artes, de saúde, ou seja, existe uma infinidade de tipos de *blogs* na atualidade. A grande maioria dos jornais, revistas, canais de televisão, empresas, escolas, universidades, estabelecimentos comerciais, repartições públicas e bibliotecas do mundo todo já possuem seus *blogs* na Internet.

Indubitavelmente, os *blogs* se tornaram um dos maiores canais de publicação e comunicação on-line na contemporaneidade e parece muito difícil imaginar a Rede sem a presença deles. De acordo com o maior diretório de *blogs* na Rede, o *Technorati*⁸⁹, não é possível saber com precisão, mas possivelmente existem mais de 100 milhões de *blogs* na Internet nos dias de hoje. O *blog* se tornou uma rede social forte e poderosa, passando a ser utilizado muitas vezes como fonte de renda.

Uma pesquisa⁹⁰ feita pelo site *Technorati* e apresentada em outubro de 2009 mostrou que 28% dos 3 mil blogueiros de várias partes do mundo entrevistados e cadastrados no diretório de *blogs* se consideram profissionais. Dentre esses 28%, 17% fazem de seus *blogs* suas principais fontes de renda. São números altos para o que até pouco tempo atrás era considerado um “diário virtual”. Certamente os *blogs* vêm deixando de lado o caráter estritamente pessoal do início e se profissionalizando. Muitos blogueiros populares na blogosfera estão lucrando com campanhas publicitárias feitas em seus *blogs* ou escrevendo *posts* encomendados sobre produtos ou serviços de uma ou outra empresa. Um exemplo interessante do uso dos *blogs* nos dias de hoje no Brasil é o *Blog do Tas*, no qual é possível observar no canto direito da tela inicial os chamados *links* patrocinados, isto é, *links* que, a cada clique do usuário, dão ao blogueiro um percentual em dinheiro dos anunciantes:

⁸⁹ www.technorati.com

⁹⁰ Para mais detalhes da pesquisa realizada pela *Technorati* verificar: <http://technorati.com/blogging/article/richard-jalichandra-keynote-blogworld-2009/>.

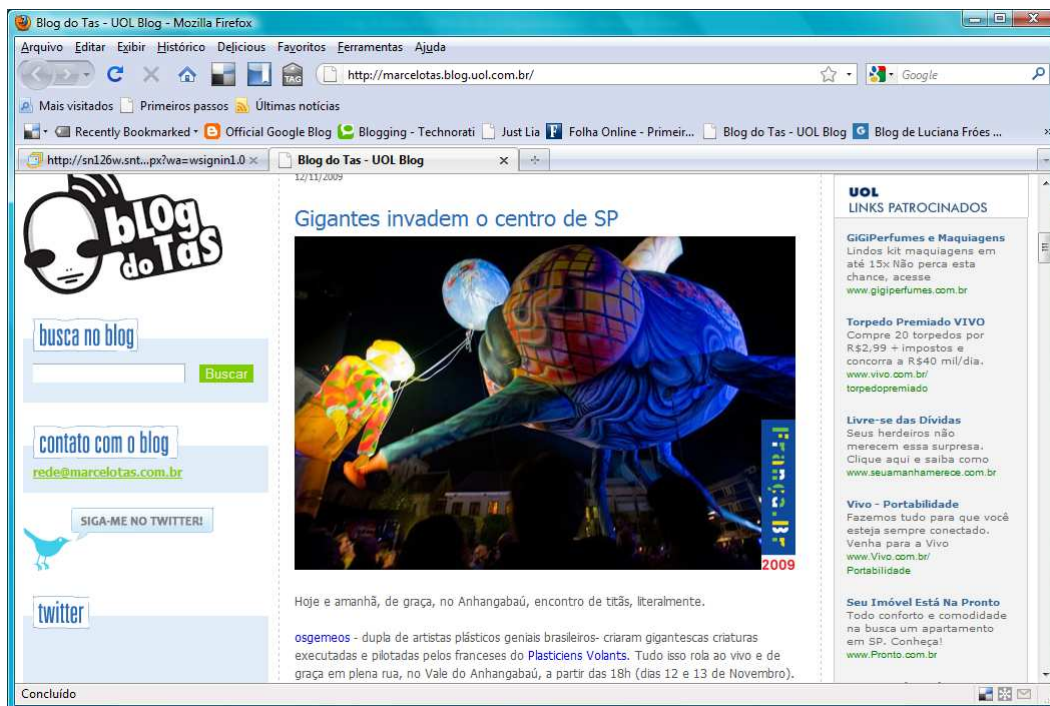


Figura 1. Reprodução da tela inicial do *Blog do Tas*, disponível em: <http://marcelotas.blog.uol.com.br/> - Acesso em 11 de novembro de 2009.

Ainda no *Blog do Tas* é possível perceber um *banner* no topo da página. O blogueiro certamente foi pago para disponibilizar o espaço para a publicidade do site www.meucarronovo.com.br:

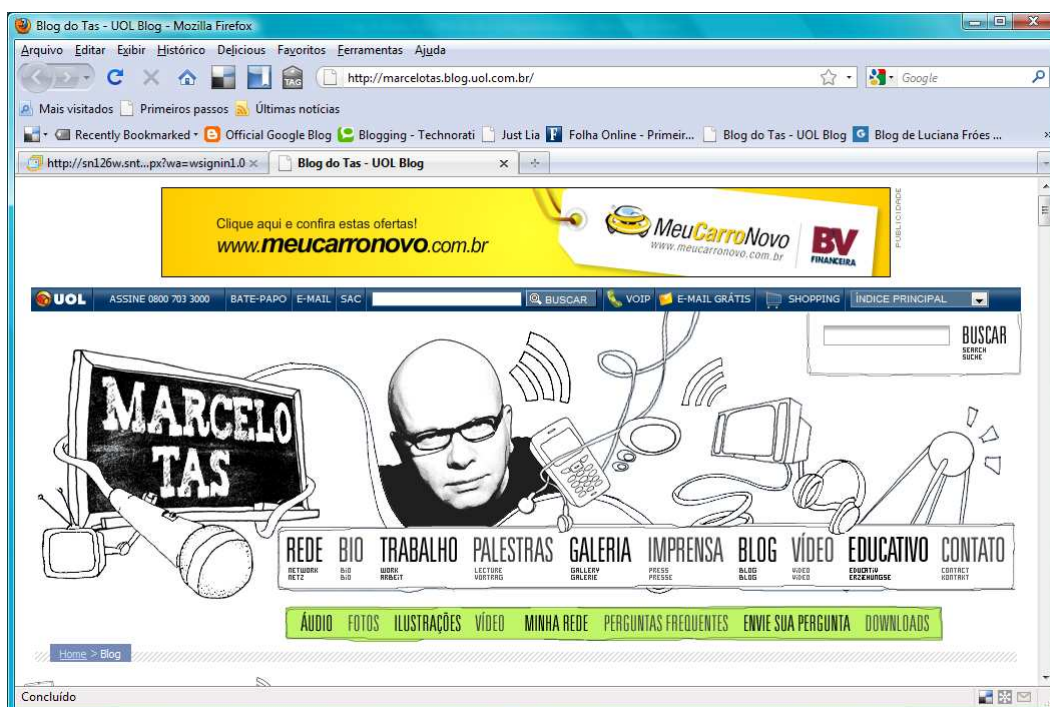


Figura 2. Reprodução da tela inicial do *Blog do Tas*, disponível em: <http://marcelotas.blog.uol.com.br/> - Acesso em 11 de novembro de 2009.

Em outro *blog* brasileiro, o *Brainstorm9*, é possível notar uma publicação patrocinada, também conhecida como *post* patrocinado. Logo abaixo do título da publicação está a especificação “post patrocinado”, indicando que uma determinada empresa comprou aquele espaço do *blog*, bem como pagou o blogueiro (certamente um formador de opinião) pela publicação assinada. A netiqueta indica que o blogueiro deve especificar para o leitor sempre que uma publicação for patrocinada, embora nem todos sigam essa regra. É possível observar abaixo a indicação do blogueiro de que se trata de um *post* patrocinado:



Figura 3. Reprodução da tela inicial do *blog Brainstorm9*, disponível em: <http://www.brainstorm9.com.br/2009/10/27/ibm-apresenta-maltos-e-associados-em-arquivo-morto/> - Acesso em 11 de novembro de 2009.

Ainda no *blog Brainstorm9* é possível encontrar um *link* que leva o leitor à seção “anuncie”, na qual o blogueiro explica as possibilidades de anúncios no *blog* (*banners* ou *posts* patrocinados), apresenta percentuais sobre os perfis de seus leitores (idade, classe social e profissão) e uma lista com as empresas que já anunciaram em seu *blog*. Parece que, além de disseminarem os conteúdos produzidos livremente pelos blogueiros, os *blogs* foram inseridos no mundo dos veículos de comunicação para campanhas publicitárias. Vejamos a seguir:

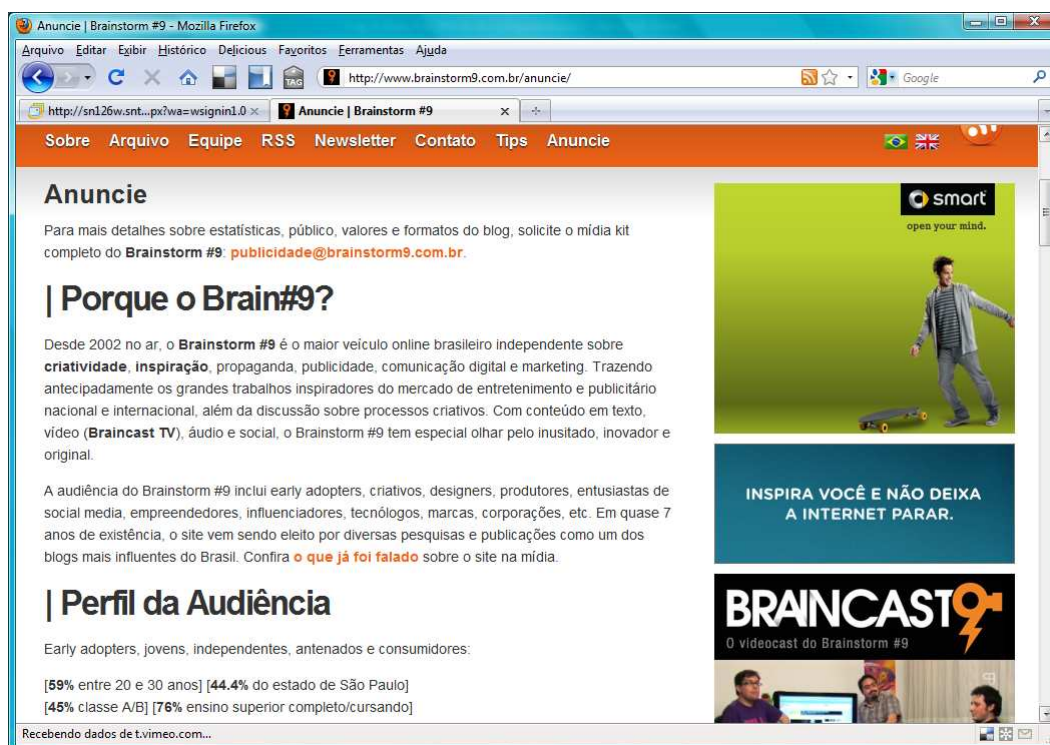


Figura 4. Reprodução da tela da seção “anuncie” do *blog Barinstorm9*, disponível em: <http://www.brainstorm9.com.br/anuncie/> - Acesso em 11 de novembro de 2009.

Outro exemplo relevante do poder dos *blogs* como produtores de conteúdo e formadores de opinião foi o uso desta ferramenta na campanha presidencial de Barack Obama, nos Estados Unidos, em 2008. Certamente seu *blog* foi uma maneira inovadora de fazer campanha política e representou um poderoso canal de comunicação, com os apoiadores e possíveis eleitores. Após o sucesso estrondoso e os resultados positivos no uso desta rede social, políticos de várias partes do mundo passaram a utilizar as redes sociais como canais de comunicação com o público durante suas campanhas. O *blog* de Barack Obama fez tanto sucesso que foi mantido após sua vitória em 2008.



Figura 5. Reprodução da tela inicial do *blog Organizing for America*, disponível em: <http://my.barackobama.com/page/content/hqblog> - Acesso em 11 de novembro de 2009.

A importância, o poder dos *blogs* e as maneiras de utilizá-los são fatos já consolidados na Rede e continuam crescendo vertiginosamente. Certamente, os *blogs* marcaram o início da segunda geração da Internet.

5.2.2. A Web das wikis

A Web 2.0 também é conhecida como a Web das *wikis*. A primeira *wiki*, que no idioma havaiano significa “super rápido”, foi criada pelo programador americano Howard G. "Ward" Cunningham em 1994. Ele desenvolveu o software colaborativo *WikiWikiWeb* e o instalou em seu website de consultoria em softwares. Uma *wiki* é um site que possibilita a criação e edição coletiva e colaborativa de um número indeterminado de páginas na Rede interligadas entre si. Assim como nos *blogs*, não é necessário dominar linguagens de programação e os usuários produzem todo o conteúdo que é publicado na página com liberdade para alterá-lo, complementá-lo e/ou apagá-lo a qualquer momento.

Em uma entrevista publicada em 2003 no site especializado em assuntos relacionados ao desenvolvimento de softwares, *Artima Developer*⁹¹, Howard G. "Ward" Cunningham explica que criou as *wikis* pensando na tendência natural do ser humano de contar histórias e na possibilidade e facilidade que estes contadores de histórias teriam de exercer suas autorias com liberdade. O diferencial está na coletividade da autoria, nas *wikis* todos podem ser autores, editores e/ou leitores dos textos em um espaço para troca de ideias e compartilhamento de conhecimento livre.

A maior e mais importante *wiki* da Web 2.0 nos dias de hoje é a enciclopédia colaborativa Wikipédia⁹², criada em janeiro de 2001 pelos americanos Jimmy Wales e Larry Sanger. Nesta enciclopédia, os usuários têm a possibilidade de criar, corrigir e expandir os verbetes que, no presente (novembro de 2009), já ultrapassam os 13 milhões. Inicialmente, a Wikipédia começou a ser desenvolvida em língua inglesa e seu sucesso foi tão grande e rápido que, no fim de 2001 e início de 2002, já existiam versões dessa enciclopédia em espanhol, português, francês, alemão, chinês, russo, japonês, entre outras línguas. O acesso de pessoas de culturas, regiões e línguas diferentes, que queriam contribuir e editar a enciclopédia, foi fazendo com que a Wikipédia se tornasse cada vez mais multilíngue. O próprio site atualiza seus dados e mostra que existem hoje publicações em 271 línguas, das quais 240 se mantêm ativas, quer dizer, com novas publicações ou correções.

A própria Wikipédia se autodefine, em sua página de abertura, com a seguinte frase: "Wikipédia: a enciclopédia livre". Apesar desta definição, o conceito de enciclopédia livre deve ser questionado. A Wikipédia utiliza programas específicos para combater vandalismos, *spams*, e outros tipos de ataques. Além desta proteção, a própria comunidade de editores voluntários, também conhecidos como wikipedistas, se une com a intenção de evitar publicações de informações falsas. De acordo com o artigo do site americano *Howstuffworks*⁹³, alterações de publicações já existentes ou publicações novas não vão para a página sem antes passar pelo crivo de um corpo editorial voluntário, que poderá aceitar, alterar ou rejeitar o texto. Uma matéria da *Boston Review*⁹⁴ de novembro de 2009 afirma que existe uma comunidade bem homogênea de editores da Wikipédia nos dias de hoje. A reportagem revelou também que 80% de seus editores são do sexo masculino, mais de 65% são

⁹¹ <http://www.artima.com/intv/wiki.html>

⁹² www.wikipedia.com

⁹³ Para a matéria completa checar: <http://computer.howstuffworks.com/wiki2.htm>.

⁹⁴ <http://bostonreview.net/BR34.6/morozov.php>

solteiros, mais de 85% não têm filhos e 70% têm menos de 30 anos. Estas estatísticas revelam que a homogeneidade dos perfis dos editores possivelmente coloca em questionamento o caráter tão diversificado e libertador da Wikipédia. Além dessas questões, a veracidade das informações contidas nesta enciclopédia é e continua sendo colocada em questão.

Independentemente do grau de confiabilidade das publicações e de liberdade oferecida aos inúmeros editores, o fato é que a Wikipédia continua sendo a *wiki* mais multilíngue existente na Rede na atualidade. A tela abaixo ilustra, parcialmente, a variedade linguística desta enciclopédia. À esquerda é possível visualizar parte da lista com as línguas disponibilizadas pela Wikipédia:

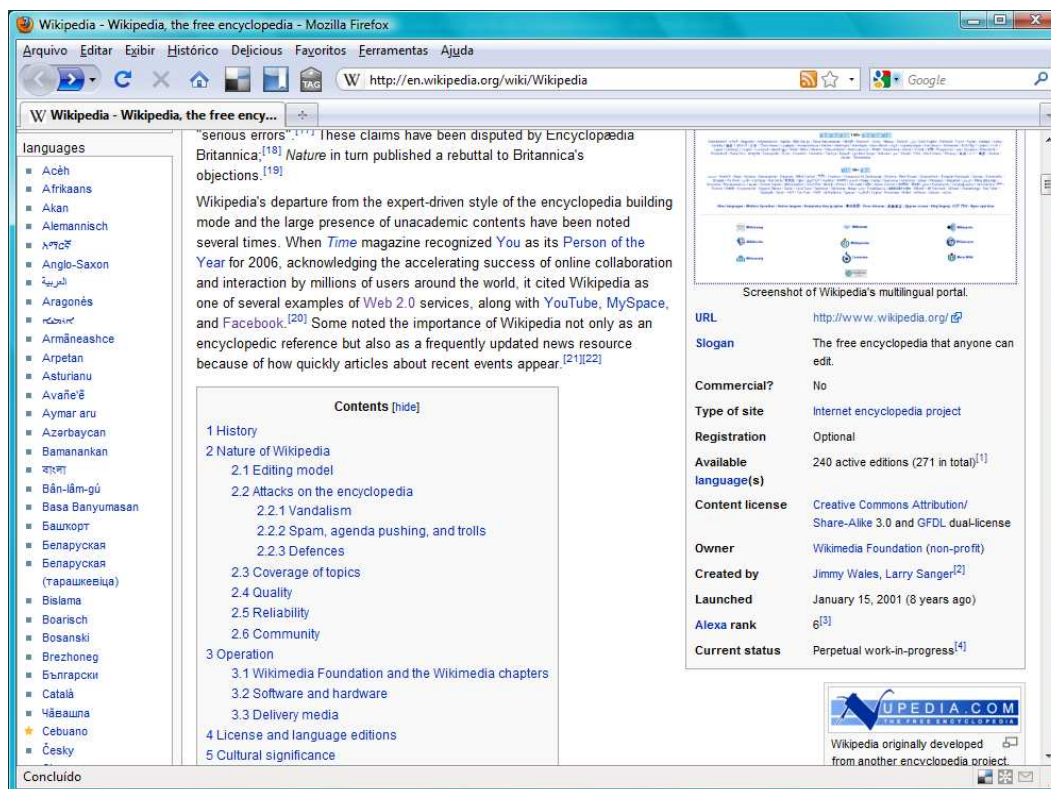


Figura 6. Reprodução da tela da enciclopédia Wikipédia. Esta tela contém a definição do termo Wikipédia em língua inglesa. Disponível em: <http://en.wikipedia.org/wiki/Wikipedia> - Acesso em 11 de novembro de 2009.

O exemplo seguinte mostra a tela de um verbete pronto para ser editado, na qual é possível observar algumas regras para os editores. A tela é a do verbete “*wiki*” em língua portuguesa. É importante observar que alguns verbetes são fechados permanentemente para edição por já serem considerados completos ou por serem alterados somente pelos administradores da Wikipédia.

As proteções⁹⁵ existentes na Wikipédia também têm o intuito de evitar vandalismo, atualizações feitas por editores que não se identifiquem previamente ou que não sejam cadastrados na comunidade de editores, e que uma página sobre um determinado assunto seja recriada.

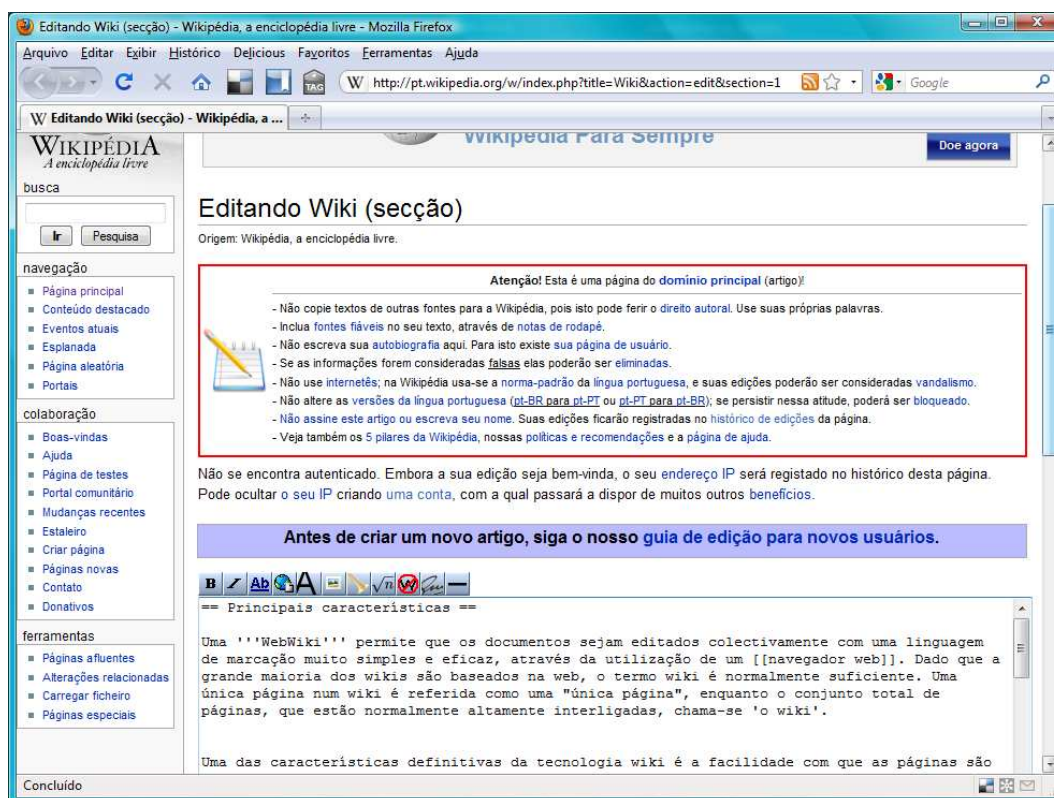


Figura 7. Reprodução da tela de edição do verbete “wiki” em português da Wikipédia. Disponível em: <http://pt.wikipedia.org/w/index.php?title=Wiki&action=edit§ion=1> - Acesso em 11 de novembro de 2009.

Além da Wikipédia, outro exemplo de *wiki* interessante na Rede é o *Wiktionary*⁹⁶. Trata-se de um dicionário colaborativo, coletivo e multilíngue com uma proposta similar à da Wikipédia em relação às publicações e edições. Este dicionário foi criado em 2002 e é possível encontrar sua versão em quase 200 línguas. Além da definição de palavras e expressões, é possível encontrar a tradução destas para várias línguas, pronúncia com áudio, imagens, etimologia e classe gramatical das palavras. Na tela abaixo podemos observar a definição do termo “on-line” oferecida pela versão em língua inglesa do *Wiktionary*:

⁹⁵ Para mais informações sobre as políticas de proteção na Wikipédia, acessar o *link*: http://en.wikipedia.org/wiki/Wikipedia:Protection_policy#semi.

⁹⁶ A versão em português deste dicionário colaborativo está disponível em: http://pt.wiktionary.org/wiki/Wikcion%C3%A1rio:P%C3%A1gina_principal.

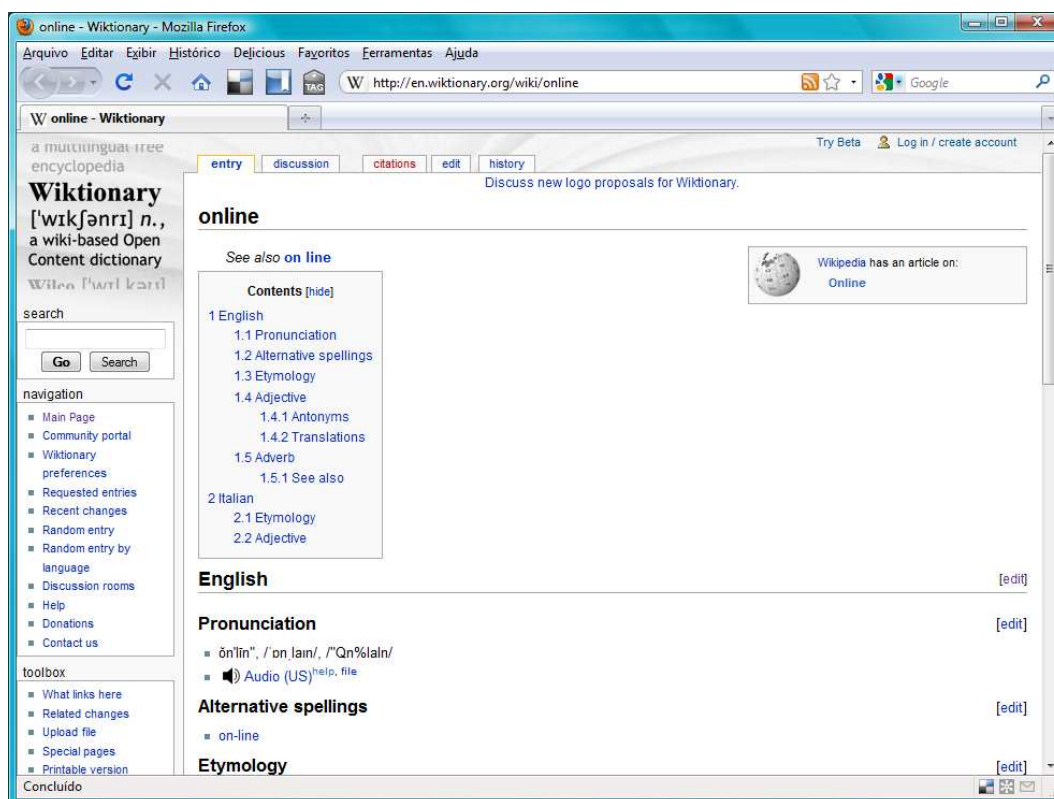


Figura 8. Reprodução da tela do verbete “on-line” da versão em língua inglesa do Disponível em: <http://en.wiktionary.org/wiki/online> - Acesso em 12 de novembro de 2009.

É possível observar a seguir a tela de edição do mesmo verbete. As políticas de edição, publicação e privacidade seguem as mesmas regras descritas anteriormente para a Wikipédia.

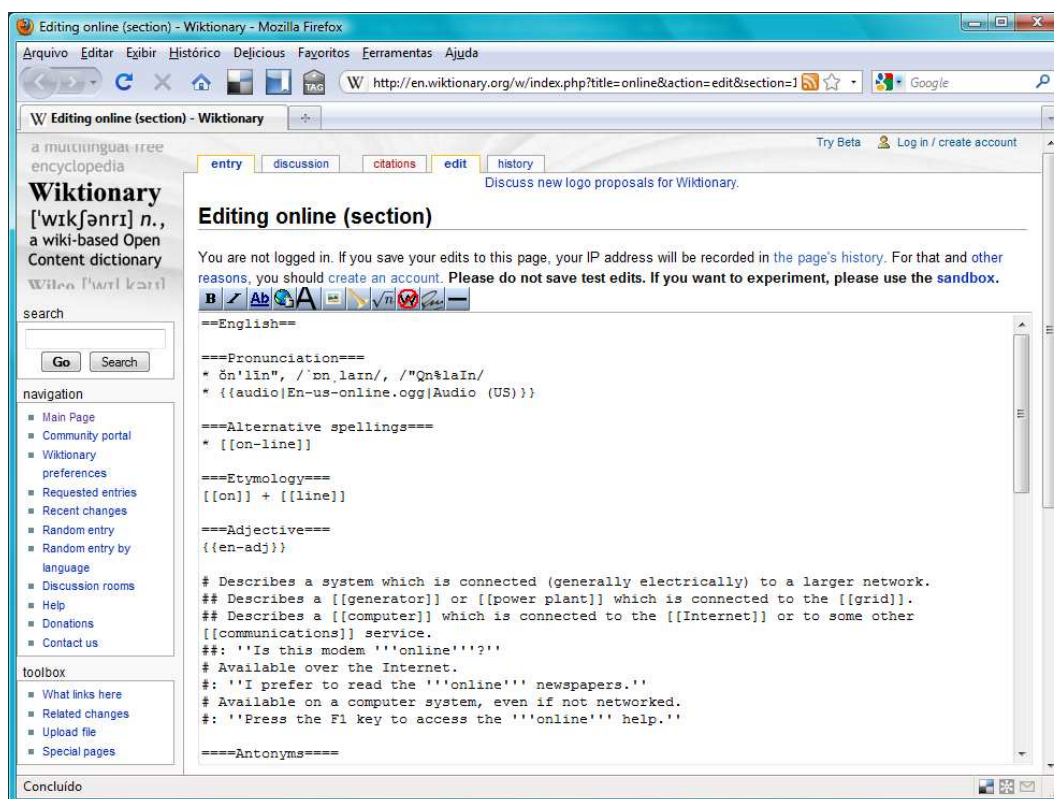


Figura 9. Reprodução da tela de edição verbete “on-line” da versão em língua inglesa do Disponível em: <http://en.wiktionary.org/w/index.php?title=online&action=edit§ion=1> - Acesso em 12 de novembro de 2009.

Existem várias outras *wikis* na Rede. A *Wikimedia Foundation* é uma organização sem fins lucrativos, criada em 2003 nos Estados Unidos por Jimmy Wales. Seu intuito foi criar diversos tipos de *wikis* colaborativas e multilíngues. Além da Wikipédia e do Wiktionary, esta organização criou algumas outras *wikis* que são menores e menos utilizadas do que a Wikipédia e o Wiktionary, mas possuem características muito similares no que concerne à edição e publicação nas páginas:

- (a) *Wikibooks* (Wikilivros) - *wiki* criada em 2003 que visa à criação coletiva de livros, apostilas e textos didáticos sobre os mais variados assuntos em 121 idiomas diferentes.
- (b) *Wikijunior* - também visa à criação coletiva de livros para crianças entre zero e 12 anos. Os editores são professores, pedagogos e jovens. O conteúdo é muito variado, com presença de muitas imagens. Este site, que também data de 2003, existe em somente em 17 idiomas.

- (c) *Wikisource* – é um acervo virtual e coletivo de textos originais (livros, documentos, cartas, leis, discursos, comunicados de imprensa) que estejam em domínio público. Atualmente em 56 idiomas, foi criado em 2003.
- (d) *Wikiquote* – criada em 2003 e em mais de 40 línguas, esta *wiki* tem como objetivo construir, colaborativamente, uma página de citações de pessoas proeminentes, de livros e provérbios dos mais variados assuntos e períodos.
- (e) *Wikimedia Commons* – criada em 2004 e disponível em mais de 80 línguas, esta *wiki* tem como objetivo criar, coletivamente, uma compilação de arquivos de mídia.
- (f) *Wikinews* (Wikinotícias) – é uma *wiki* que propõe o jornalismo coletivo e colaborativo. Foi criada em 2003 e existe em quase 30 línguas.
- (g) *Wikiversity* (Wikiversidade) – criada em 2006 e disponível somente em 10 línguas, esta *wiki* visa à criação de projetos educacionais por alunos, professores e pesquisadores.
- (h) *Wikispecies* – *wiki* que visa criar um catálogo com todas as espécies existentes no mundo. Esta *wiki* foi criada em 2004 e é encontrada em 60 línguas.

Dentre as inúmeras línguas disponíveis nas *wikis* criadas pela *Wikimedia Foundation*, uma merece destaque especial: o *simple English* (inglês simples). Na própria página da Wikipédia é possível encontrar a definição desta língua. Esta é descrita como uma simplificação da língua inglesa, visando facilitar a leitura das *wikis* por crianças, estudantes, adultos com dificuldades de aprendizado e pessoas que estão aprendendo a língua inglesa. De acordo com a explicação da Wikipédia, o *simple English* somente é utilizado para simplificar *wikis* já existentes em língua inglesa, sendo assim, nenhuma publicação nova é feita em *simple English*. A sugestão é que os artigos nesta língua sejam escritos utilizando as 1.000 palavras mais comuns da língua inglesa e gramática simplificada, no entanto, não são oferecidos detalhes de quais seriam essas palavras ou de quais construções gramaticais seriam consideradas simples. No

caso de artigos mais complexos sobre ciências, política e religião, por exemplo, que necessitem entre 2.000 e 3.000 palavras, é sugerido utilizar mais palavras, além das 1.000 sugeridas inicialmente, mas estas devem ser simples também. É proposto também que os editores tenham como base o BASIC. Como descrito no capítulo 3 desta investigação, o BASIC foi um projeto de língua inglesa única e simplificada com base no inglês, criado em 1930 por Charles Kay Ogden.

É curioso observar que, apesar do caráter indubitavelmente multilíngue da Internet e da proposta da *Wikimedia Foundation* de criar os mais diversos tipos de *wikis* colaborativas e multilíngues, o desejo de se criar uma língua única de comunicação, assim como Crystal, Seidlhofer e Rajagopalan, permanece, pelo menos para a *Wikimedia Foundation*, independentemente da filosofia da Web 2.0 e de todas as possibilidades, oferecidas por esta, de liberdade na criação, por parte dos usuários, de conteúdos em qualquer língua. Parece que, neste caso, a *Wikimedia Foundation* está na contramão em relação ao caminho que a Rede vem percorrendo na atualidade.

5.2.3. As Redes Sociais

Como já foi dito, a Web 2.0 é a Web social, participativa e colaborativa, na qual o conteúdo é produzido pelos próprios usuários. Também é a Web das redes sociais. Além dos *blogs*, que foram os precursores da segunda geração da Internet, é importante citar outras redes sociais que se destacam muito nessa fase da Rede.

O site de relacionamentos *Orkut*⁹⁷, mencionado no capítulo 4, é uma rede social afiliada ao site de busca Google. Foi criado em janeiro de 2004 nos Estados Unidos com o objetivo de ajudar seus membros a criar novas amizades e manter relacionamentos on-line. Seu nome é inspirado no projetista chefe de origem turca, Orkut Büyükkökten, analista de sistemas do Google. Trata-se da rede social mais acessada do Brasil. É importante lembrar que o *Orkut*, originalmente americano, está disponível em mais de 40 línguas nos dias de hoje e conta com mais de 60 milhões de usuários, dos quais mais de 50% são brasileiros, já são mais de 60 milhões. O conteúdo do *Orkut* é composto de textos, imagens e vídeos publicados pelos próprios usuários. O próprio site

⁹⁷ www.orkut.com.br

oferece dados demográficos mostrando a penetração do *Orkut* no Brasil. Como é possível observar abaixo:

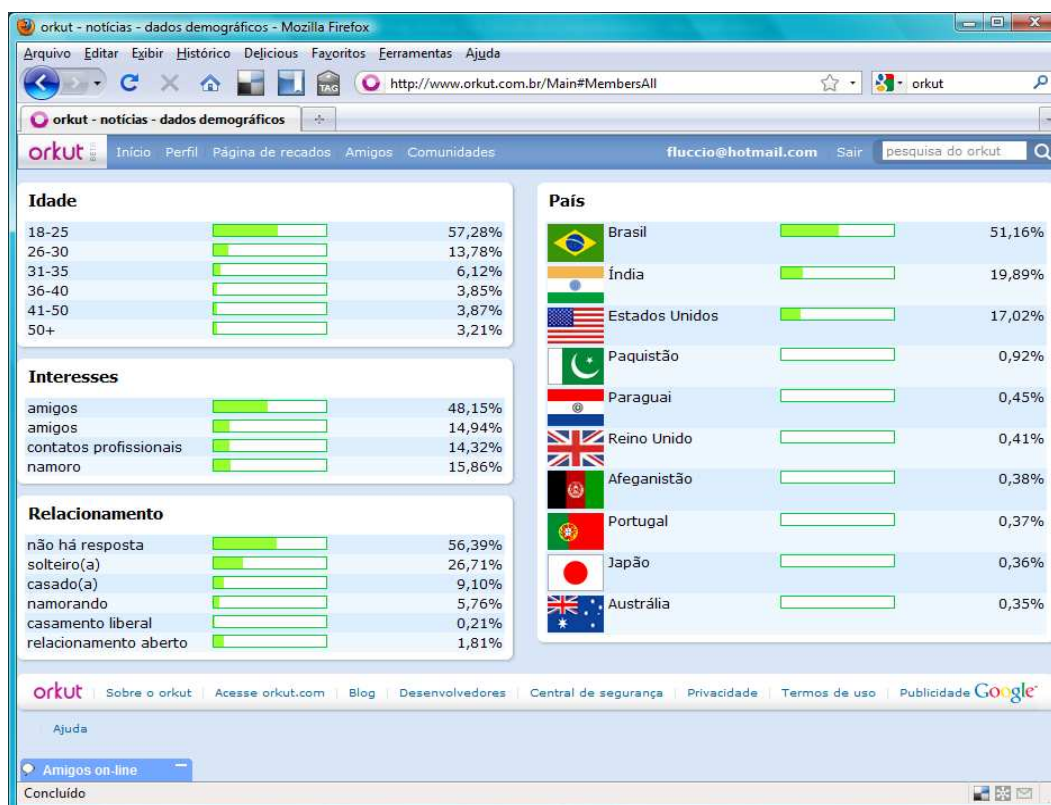


Figura 10. Reprodução da tela dos dados demográficos do *Orkut*. Disponível em: <http://www.orkut.com.br/Main#MembersAll> somente para membros cadastrados - Acesso em 15 de novembro de 2009.

Na página de configurações do usuário, a seguir, é possível observar a lista de algumas línguas nas quais o site pode ser acessado. No primeiro item – idioma de exibição – o próprio site convida seu usuário a utilizar o *Orkut* no idioma que escolher, dentre os que estão na lista oferecida:



Figura 11. Reprodução da tela das configurações do meu perfil no *Orkut*. Disponível em: <http://www.orkut.com.br/Main#GeneralSettings?rl=ls&xid=1270709720122515107> - Acesso em 15 de novembro de 2009.

Outra rede social multilíngue que vem se destacando no mundo todo é o *Facebook*⁹⁸. Esta foi criada em fevereiro de 2004 pelo estudante de Harvard e programador Mark Elliot Zuckerberg. Trata-se de uma rede social, similar ao *Orkut*, na qual as pessoas podem se encontrar, conversar e se informar sobre amigos e familiares, publicar fotos, vídeos e música. No *Facebook* o usuário também é responsável pela produção do seu conteúdo. Como mencionado no capítulo 4, esta rede social já está disponível em mais de 70 línguas. Atualmente é a rede social mais popular na Internet, com mais de 300 milhões de usuários, de acordo com a revista Info Exame de novembro de 2009. É possível observar, na tela a seguir, uma janela com a lista das línguas disponibilizadas pelo *Facebook* atualmente:

⁹⁸ www.facebook.com

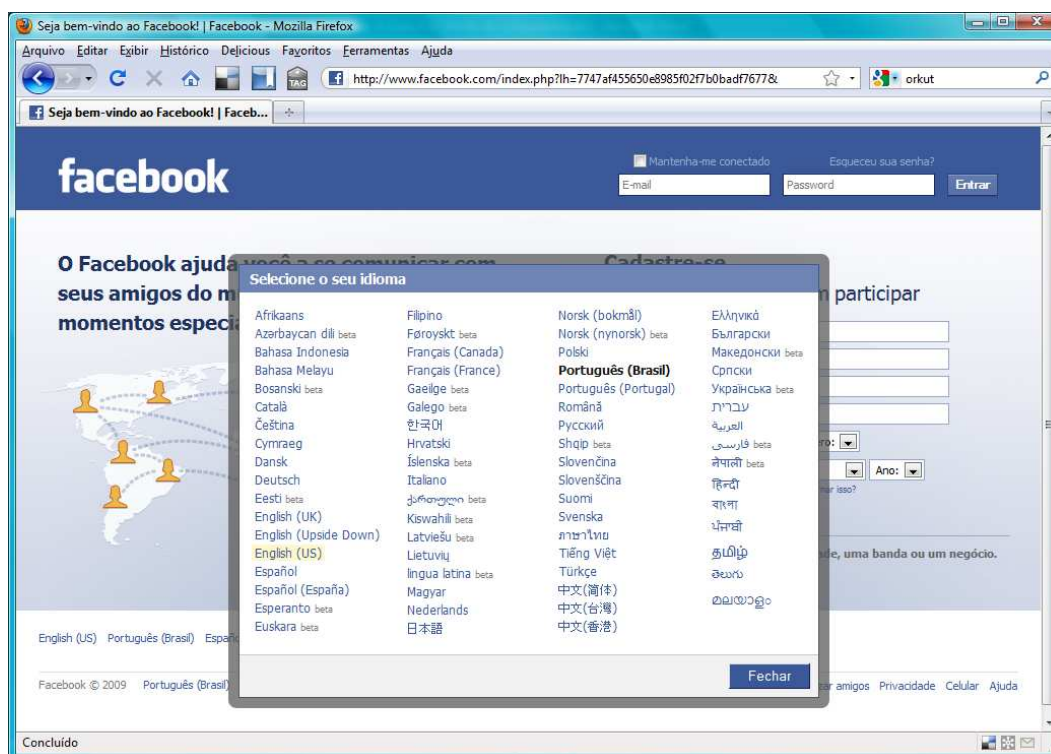


Figura 12. Reprodução da tela inicial do Facebook com a janela das línguas disponibilizadas. Disponível em: <http://www.facebook.com/index.php?lh=7747af455650e8985f02f7b0badf7677&> - Acesso em 15 de novembro de 2009.

A rede social *MySpace*⁹⁹, também americana, foi criada em 2003 e se tornou a rede social de músicos profissionais e amadores, devido a sua capacidade distintiva de hospedar MP3. Esta fez com que muitas bandas e músicos de diversas nacionalidades criassem seus perfis no *MySpace* e fizessem destes suas páginas oficiais de trabalho. Nesta rede social, os músicos têm a possibilidade de publicar suas discografias completas ou parciais, vendê-las e interagir com o público sobre seus trabalhos. Os usuários também publicam muitos vídeos e fotos. Já é possível acessar o *MySpace* em quase 20 línguas. A seguir observa-se a tela inicial de um perfil no *MySpace* com a janela do *MySpace International* aberta no canto direito. Esta indica os locais e as línguas nos quais esta rede social pode ser utilizada.

⁹⁹ www.myspace.com

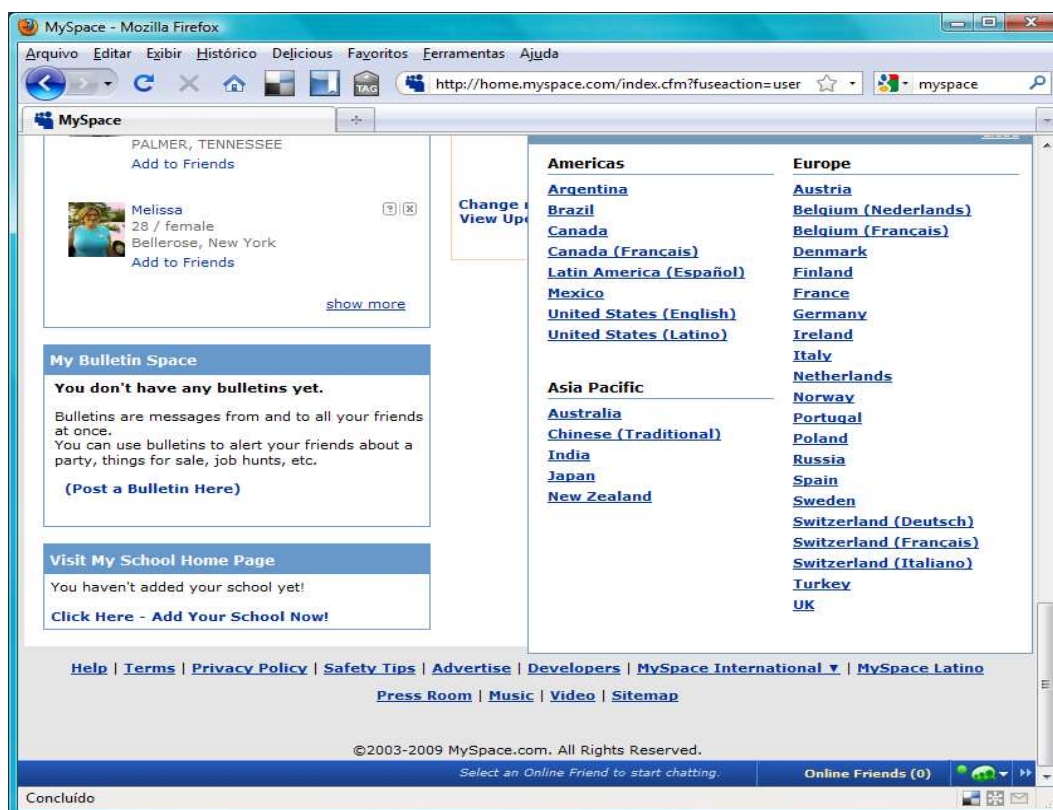


Figura 13. Reprodução da tela inicial do meu perfil no *MySpace* com a janela dos locais e línguas disponíveis. Disponível em: <http://home.myspace.com/index.cfm?fuseaction=user> - Acesso em 15 de novembro de 2009.

É relevante destacar outras duas redes sociais muito utilizadas na Rede por usuários de várias partes do mundo, mas ainda não tão multilíngues quanto as citadas anteriormente. A rede social *LinkedIn*, criada nos Estados Unidos em 2003, tem como foco ser uma rede social profissional. No *LinkedIn* os usuários podem criar seus perfis com suas experiências profissionais, fazer novos contatos, recomendar profissionais através de depoimentos, anunciar eventos ou encontrar empregos na seção “jobs”. Muitas empresas vêm utilizando o *LinkedIn* para recrutar novos funcionários e anunciar vagas. Apesar de seu diferencial, o *LinkedIn* somente está disponível em inglês, espanhol, alemão e francês, como é possível observar no canto superior direito da tela a seguir:

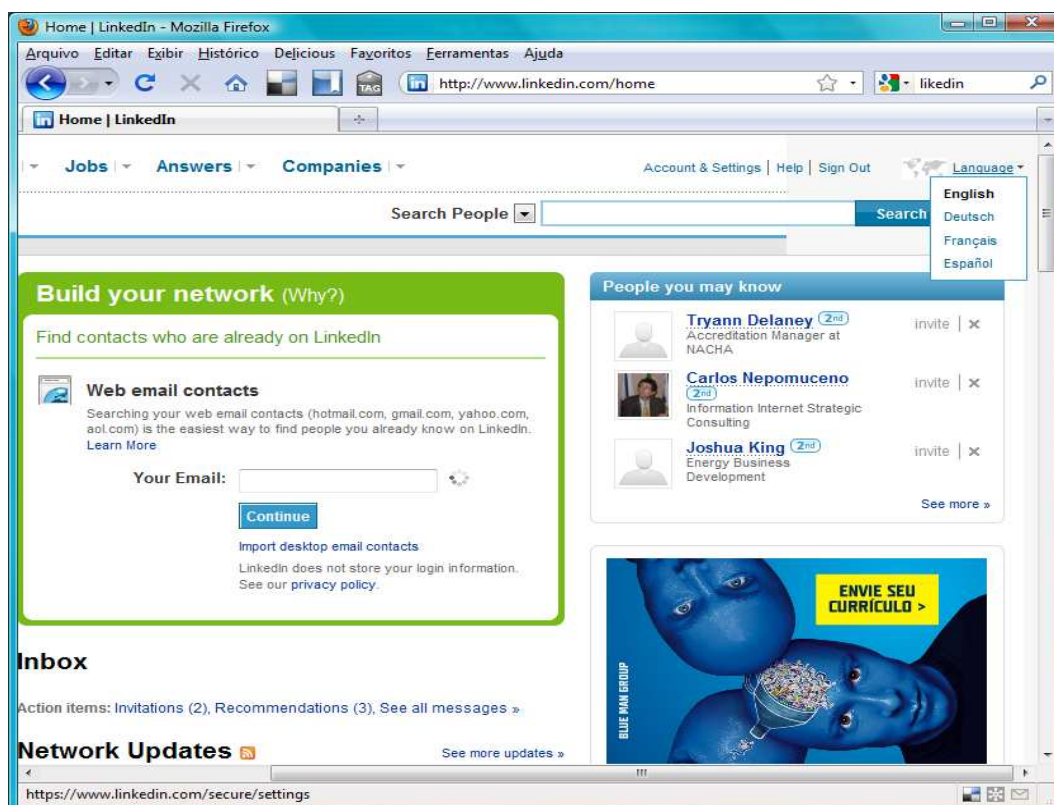


Figura 14. Reprodução da tela inicial do meu perfil no *LinkedIn*. Disponível em: http://www.linkedin.com/home?trk=hb_logo - Acesso em 15 de novembro de 2009.

A rede social mais recente da Internet é o *Twitter*¹⁰⁰. Criada em 2006, somente se popularizou no Brasil em 2009. O *Twitter* também é conhecido como *microblogging*, pois é comparado a um *blog* compacto. Os usuários, também conhecidos como *tuiteiros*, publicam mensagens de até 140 caracteres com base em uma pergunta que se localiza no topo da página: “*What’s happening?*” (“O que está acontecendo?”). Estas mensagens curtas podem ser publicadas via Web, SMS e dispositivos portáteis. Neste espaço para somente 140 caracteres, os usuários têm a possibilidade de publicar *links* que podem levar seus leitores, também chamados de *followers* (seguidores) para outras páginas da Rede. Estes *links* têm suas URLs automaticamente abreviadas como por exemplo, o *link* <http://g1.globo.com/Noticias/Brasil/0,,MUL1379810-5598,00.html> da página G1 do portal Globo.com foi abreviado para <http://short.ie/w9rfoe>. Esta abreviação de URLs pode ocorrer automaticamente pelo *Twitter* ou através de ferramentas existentes na Rede que têm como objetivo encurtar os *links* da Web, visando facilitar sua publicação ou memorização pelos usuários. Dentre essas

¹⁰⁰ www.twitter.com

ferramentas estão o brasileiro Migre.me¹⁰¹ e o americano Bit.ly¹⁰², este é um dos mais utilizados na Rede¹⁰³.

No *Twitter* os usuários podem também publicar *links* que levam seus seguidores ao serviço de armazenamento de imagens chamado *TwitPic*. É importante lembrar que no *Twitter*, um usuário pode seguir (*follow*) ou ser seguido por outros usuários, denominados *followers* (seguidores). Na tela que se segue é possível observar, no canto superior direito, que a usuária está seguindo (*following*) 146 pessoas e está sendo seguida por 92 pessoas (*followers*). Desta forma, todas as publicações dos usuários são vistas pelos seus *followers* (seguidores). O *Twitter* se popularizou no mundo todo como uma forma eficiente e rápida de divulgação de notícias, imagens, vídeos, ou seja, qualquer conteúdo hospedado na Rede. Muitas empresas já estão usando esta rede social para divulgar notícias, produtos e serviços. Nos dias de hoje, 30% da audiência dos sites de notícias têm origem no *Twitter*, de acordo com a revista Info Exame de novembro de 2009.

A seguir observa-se a tela inicial de um perfil no *Twitter* com a pergunta inicial no topo da página, os *followings* e os *followers* no canto direito.

¹⁰¹ <http://migre.me/sobre/>

¹⁰² <http://bit.ly/pages/about/>

¹⁰³ Dados da revista Info Exame de novembro de 2009.

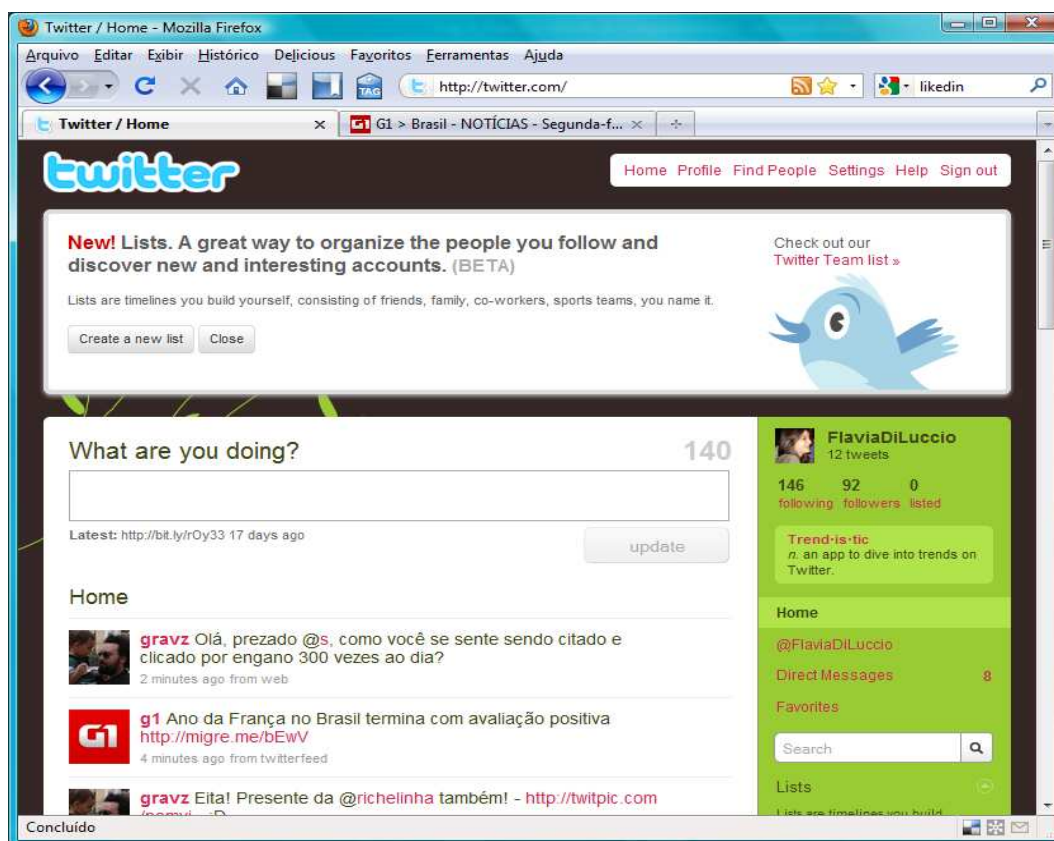


Figura 15. Reprodução da tela inicial do meu perfil no *Twitter*. Disponível em: <http://twitter.com/FlaviaDiLuccio> - Acesso em 15 de novembro de 2009.

No momento, o *Twitter* só está disponível em inglês, espanhol e japonês, mas devido à sua rápida popularização no mundo, possivelmente estará disponível, em breve, em várias outras línguas.

5.2.3.1. Sites de compartilhamento

Os chamados sites de compartilhamento são redes sociais com foco no compartilhamento de imagens, arquivos de áudio e vídeo. Três destes serão destacados nesta investigação.

O primeiro deles é o *YouTube*¹⁰⁴. Criado em 2005, este site tem como lema a frase “*broadcast yourself*”, permite que os usuários publiquem, assistam e compartilhem vídeos em formato digital. Além disso, como nos *blogs*, os usuários podem fazer comentários sobre vídeos assistidos, criar uma lista de favoritos e se comunicar com outros usuários através da seção de comentários ou por email. Os vídeos são organizados por categorias: vídeos assistidos no

¹⁰⁴ www.youtube.com

momento, vídeos destacados pela equipe do *YouTube* e vídeos recomendados para usuários cadastrados no site com base em seus históricos de acesso. Todo o conteúdo do *YouTube* depende dos usuários e já está disponível em 16 línguas. O acesso ao *YouTube* é gratuito¹⁰⁵ e a partir de um cadastro, qualquer usuário, de qualquer lugar, pode publicar um vídeo utilizando qualquer língua, mesmo que não seja uma das 16 línguas para as quais o site já foi traduzido.

O segundo exemplo desta categoria é o *Flickr*¹⁰⁶. Sua proposta é muito similar a do *YouTube*, mas ao invés de vídeos, os usuários publicam fotos, criam álbuns, fazem contatos com outros usuários, comentários sobre as imagens e contatos profissionais. O *Flickr* está disponível em 8 línguas somente.

O último exemplo de site de compartilhamento é o *Blip.fm*¹⁰⁷. Trata-se de um site de compartilhamento de músicas, no entanto este ainda não se difundiu como os outros e somente está disponível em língua inglesa. Sua página é muito similar à do *Twitter* e os usuários podem publicar comentários sobre as músicas postadas por outros usuários e criar listas com músicas e DJs favoritos.

O movimento Web 2.0 vem se fortalecendo há algum tempo, tanto que a revista *Time* de dezembro de 2006 tinha como chamada da capa: “Você é eleito a pessoa do ano pela revista *Time*.” A edição destacou e elogiou o público que contribuiu muito para a produção do conteúdo da Internet. Além da produção coletiva de conteúdo, a revista destacou a possibilidade de interação que as ferramentas da Web 2.0 proporcionam aos usuários. Estes têm a possibilidade de interagir com outros usuários, com empresas, com escritores, com leitores, etc, em inúmeras línguas. É a geração da produção coletiva e colaborativa de conteúdo. Por outro lado, é importante destacar também que a possibilidade, a partir da Web 2.0, de qualquer usuário produzir e publicar conteúdo sem intermediários faz com que o próprio usuário tenha a tarefa extra de filtrar a quantidade de informações disponibilizadas na Rede, tendo, eventualmente certa dificuldade em encontrar o que procura. É possível afirmar que a liberdade trouxe um controle de qualidade relativamente baixo na Rede.

A tela a seguir mostra quem controla as informações na Rede, de acordo com a revista *Time* de 2006:

¹⁰⁵ O acesso a todas as redes sociais citadas neste capítulo também é gratuito.

¹⁰⁶ www.flickr.com

¹⁰⁷ <http://blip.fm/home>



Figura 16. Reprodução da capa da revista Time de 25 de dezembro de 2006. Disponível em: <http://www.time.com/time/covers/0,16641,20061225,00.html>. - Acesso em 13 de outubro de 2007.

Se a Web 2.0 traz uma revolução em relação à produção de conteúdo na Internet, propicia também uma revolução linguística, como pudemos observar nos exemplos de sites confeccionados durante a Web 2.0 e citados anteriormente. Durante a sua primeira geração, a Internet parecia linguisticamente homogênea, com a língua inglesa no domínio. No entanto, parece que a Internet e seu movimento Web 2.0 também estão contribuindo para o enfraquecimento do inglês como língua única da Rede¹⁰⁸. Como foi possível

¹⁰⁸ De acordo com o estudo feito pelo *International Telecommunication Union* (http://www.itu.int/ITU-D/ict/material/Telecom09_flyer.pdf) apresentado em outubro de 2009, mais de ¼ da população mundial está usando a Internet. O site de pesquisas estatísticas sobre a Rede *Internet World Stats* (<http://www.internetworldstats.com/stats7.htm>) apresentou os dados de 2009. Estes revelam que até setembro a língua inglesa era utilizada por 27,6% dos usuários da Rede, seguida do chinês com 22,1% e do espanhol com 7,9% dos usuários. Os números são baseados nos falantes nativos das línguas, no entanto, o site não oferece detalhes sobre a região desses falantes e nem sobre os critérios utilizados para o recrutamento dos falantes considerados nativos. A única informação disponibilizada é que muitas pessoas são bilíngues ou multilíngues e que para a pesquisa foi designada somente uma língua por pessoa. No entanto, não fica claro como foi feita a escolha de uma língua nativa no caso dos falantes bilíngues ou multilíngues. As informações completas com as 10 línguas mais acessadas da Rede estão em: <http://www.internetworldstats.com/stats7.htm>. Sobre os 20 países com maior número de usuários

observar, este movimento representa a abundância de conteúdo, de informações e de comunicação na Rede. A Web 2.0 parece ter revolucionado o que já era revolucionário e surpreendente, a Revolução Digital. Seria possível também revolucionar a segunda geração da Rede? Parece já ser possível afirmar que a resposta tende a ser positiva e que a Web está se encaminhando para a sua terceira geração, a chamada Web 3.0 ou Web Semântica. Vejamos.

5.3.

Web 3.0: o presente

Após os vários exemplos oferecidos nas seções anteriores, chega-se à conclusão que a Web se tornou um espaço infinito de informações multilíngues que podem ser armazenadas e acessadas, a partir de *links*, por qualquer usuário de qualquer parte do mundo. Além disso, oferece aos usuários inúmeras possibilidades de comunicação e relacionamento. Nesse contexto, parece que a partir da Web 2.0 a Rede chegou ao grau mais alto de possibilidades. No entanto, esta segunda geração da Rede trouxe consigo algumas consequências nem sempre positivas.

A Revolução Digital foi extremamente importante, pois fez com que os computadores pudessem se integrar em uma Rede. Para que isso se concretizasse, a linguagem de programação HTML (*HyperText Markup Language*/Linguagem de Marcação de Hipertexto) passou a ser utilizada para descrever as estruturas das páginas da Web. O HTML é a linguagem da Web 1.0 e da Web 2.0. Esta foi criada por Tim Berners-Lee¹⁰⁹, considerado o pai da *World Wide Web*, no início dos anos de 1990 e permite:

- (a) o armazenamento e publicação on-line de qualquer tipo de documento: textos, tabelas, listas, fotos, etc;
- (b) a recuperação de qualquer informação on-line através de *links* hipertextuais, isto é, a partir de um clique;

na Rede, acessar: <http://www.internetworldstats.com/top20.htm>. Até junho de 2009 a China estava em primeiro lugar, seguida dos EUA, Japão, Índia e Brasil.

¹⁰⁹ Criador do *World Wide Web Consortium* (W3C), comunidade internacional que visa à criação de padrões com o objetivo de garantir o crescimento da Rede. No W3C é possível encontrar informações sobre Web Semântica, utilizadas durante a confecção deste capítulo. Para detalhes verificar: <http://www.w3.org/>.

- (c) o desenvolvimento de ferramentas que possam ser utilizadas para efetuar transações com serviços remotos, como busca de informações, reservas de hotéis, compras de produtos, pagamentos de contas, etc;
- (d) a inclusão de planilhas, arquivos de áudio e vídeo diretamente em um documento.

É importante ressaltar que a linguagem HTML possui uma sintaxe que é compreendida pelo computador. Trata-se de uma linguagem de marcação estrutural e não semântica. Como mencionado acima, qualquer documento pode ser armazenado e recuperado pelo usuário com base em palavras-chave, também conhecidas como *tags* ou etiquetas que descrevem e classificam as informações existentes nas páginas da Web para que estas sejam recuperadas e acessadas pelos usuários. As *tags* são escolhas pessoais e informais dos criadores dos conteúdos on-line e não existe uma padronização para tais escolhas. Podem ser consideradas representantes importantes da segunda geração da Web, pois para que os *blogs*, *wikis* ou redes sociais sejam encontrados dentre as inúmeras páginas da Rede é necessário que seus criadores escolham *tags* que melhor representem os conteúdos de suas páginas. Apesar de as *tags* facilitarem a recuperação dos documentos armazenados na Rede, o que existe hoje é um grande amontoado de documentos que podem ser acessados por *links* que são listados em sites de busca, como o Google. Nesse contexto, é relevante tentar encontrar soluções para as questões a seguir: como é possível que o usuário encontre o que procura nesse mundo de informações? Como é possível fazer com que o sistema organize essas informações de acordo com os interesses do usuário?

A Internet permite que os computadores estejam ligados uns aos outros e formem uma grande Rede na qual é possível trocar informações. No entanto, estas informações não são compreendidas pelos próprios computadores. Sendo assim, estes não têm a possibilidade de organizar as inúmeras informações contidas na Web, fazendo com que esta pareça uma grande biblioteca, com uma infinidade de livros não catalogados em inúmeras línguas. A semelhança é grande com a Biblioteca de Babel descrita por Borges (1996), composta por um número indefinido e infinito de galerias hexagonais, com vastos poços de ventilação e pequenas varandas, de onde se pode ver, interminavelmente, qualquer um dos pisos inferiores e superiores. A Biblioteca de Babel é infinita e possui todas as informações do mundo, incluindo biografias de qualquer pessoa, previsões para o futuro e traduções de qualquer livro para qualquer língua,

trazendo, assim, a esperança de que todos os mistérios básicos da humanidade se resolvam. Por outro lado, essa esperança é logo transformada em depressão excessiva diante da impossibilidade de encontrar os livros mais preciosos no meio daquela infinidade de volumes. Parece que a Rede, com seus intermináveis *links*, se assemelha a tal biblioteca, com seu acervo em desordem.

É possível imaginar que o usuário da Web tem certeza que existem documentos preciosos na infinidade de *links* disponíveis, mas sua esperança de encontrar exatamente o que busca dá lugar à frustração ao constatar a dificuldade em acessar os documentos desejados. A partir desta comparação, outra questão que deve ser discutida é se seria possível catalogar e organizar toda essa “biblioteca” composta por uma infinidade de *links* existente na Rede.

5.3.1.

Uma proposta repleta de desafios

De acordo com Berners-Lee (2001), é possível organizar os inúmeros *links* existentes na Rede e fazer com que os computadores compreendam as informações que armazenam. Isto quer dizer que os questionamentos apresentados na seção anterior podem ser solucionados. Tais soluções parecem estar na concepção de Web Semântica, também conhecida como Web 3.0 ou terceira geração da Rede. Esta foi inicialmente descrita por Berners-Lee no texto intitulado *The Semantic Web* (“A Web Semântica”) em 2001. De forma resumida, a Web Semântica é a Web do significado, ou seja, é a geração da Web na qual os computadores entendem os significados que estão por trás de cada página, isto é, de cada documento armazenado na Rede. Enquanto a Web 2.0 foi a passagem da Web das publicações eletrônicas (Web 1.0) para a Web social (Web 2.0), a Web 3.0 é a passagem desta para a Web dos significados, baseada no armazenamento, integração, combinação, interpretação e interação das informações contidas em diferentes partes da Rede. Diferentemente da mudança social ocorrida com o início da Web 2.0, a Web 3.0 acarretou uma mudança tecnológica.

Visando uma melhor compreensão da Web Semântica, Berners-Lee (2001) inicia sua descrição com um exemplo fictício interessante. Em seu exemplo, Pete¹¹⁰ estava ouvindo a música dos Beatles “*We can work it out*” no

¹¹⁰ Pete e Lucy são personagens fictícios presentes no exemplo de Berners-Lee (2001).

seu sistema de entretenimento quando o telefone tocou. Ao atender a ligação, seu telefone enviou uma mensagem aos aparelhos locais que fizeram com que o volume da música diminuísse. Do outro lado da linha, Lucy, sua irmã que ligava do consultório do médico, dizia que a mãe deles precisava ver um especialista e posteriormente fazer uma série de sessões de fisioterapia duas vezes por semana. Pete concordou imediatamente com Lucy em adotar um sistema de revezamento para levar a mãe ao fisioterapeuta. Do consultório médico, Lucy instruiu seu agente de Web Semântica de um navegador portátil. Este armazenou informações sobre o tratamento da mãe e fez uma rápida pesquisa sobre os melhores e mais econômicos seguros que pudessem dar cobertura à sua mãe em um raio de até 32 quilômetros de casa durante o tratamento. Além disso, seu agente fez um cruzamento de sua agenda com a de seu irmão e do fisioterapeuta e apresentou um planejamento para o tratamento. No entanto, Pete não concordou com o planejamento e alegou que a clínica era longe da casa de sua mãe e que haveria muito trânsito nos horários sugeridos. Em poucos minutos, um novo planejamento foi apresentado pelo agente, considerando a localização e horários. O sistema também remarcou os compromissos considerados menos importantes na agenda de Pete que estavam coincidindo com o novo planejamento oferecido. Tudo isso só seria realmente possível por intermédio da Web Semântica.

Conforme afirmam Berners-Lee, Shadbolt & Hall (2006), a Web Semântica representa uma extensão tanto da Web 1.0 quanto da Web 2.0 que consistiam basicamente em documentos que poderiam ser acessados e lidos pelos usuários, para a Rede que inclui informações e documentos que são manipulados pelo computador. É a chamada era das informações ativas geradas pelo próprio documentos a partir da teoria semântica que possibilita a interpretação de símbolos. Desta forma, os documentos existentes na Rede podem ser interpretados e integrados pelo próprio sistema, possibilitando o acesso dos usuários às informações corretas para um contexto específico, de forma mais rápida e direta. A Web 3.0 permitiria, por exemplo, que o usuário abrisse seu calendário no computador e visse uma viagem de trabalho agendada. Com base nessa informação, o sistema disponibilizaria imediatamente todas as possibilidades de voo para o seu destino naquele dia, os hotéis disponíveis e próximos do local de trabalho, os restaurantes e tudo relacionado àquela viagem específica. A chave para que tudo isso se torne realidade parece estar na integração dos mais diversos dados e informações contidos na Rede, para que as máquinas e os homens trabalhem cada vez mais

em cooperação. É possível afirmar que a Web Semântica visa à universalização da Rede, a qual qualquer coisa/pessoa pode se ligar a qualquer coisa/pessoa. Em outras palavras, é a Rede descentralizada ao máximo.

A seguir, será feita uma descrição simplificada e breve de como a Web Semântica funciona na prática.

5.3.1.1. A Web Semântica na prática

O exemplo da seção anterior mostra que a filosofia principal da Web Semântica é a integração dos dados existentes na Rede. É possível afirmar que se trata de uma extensão da Web que já existe nos dias de hoje, diferenciando-se por possibilitar que os usuários e os computadores trabalhem cooperativamente. Para que isso se concretize é necessária a utilização de linguagens específicas que sejam capazes de expressar os dados e regras de representação de significado. Segue outro exemplo fictício de como a Web 3.0 funcionaria.

João coleciona selos há muitos anos. Para cada item de sua coleção ele fez um documento com informações detalhadas, criando, assim, inúmeros documentos independentes. No momento, ele quer encontrar um selo vermelho do Reino Unido utilizado entre 1981 e 1995 e criado por John Bryan Dunmore. No entanto, parece bastante difícil para ele encontrar o selo que procura dentre os inúmeros documentos criados. Este exemplo nos remete às ferramentas de busca existentes na Rede nos dias de hoje. O selo vermelho procurado por João representa um item específico buscado pelo usuário na Web e os inúmeros documentos criados por João representam a Web atual, isto é, um enorme conjunto de documentos independentes. Sendo assim, é possível que João encontre o selo que procura, mas certamente encontrará dificuldades e levará tempo para isso. Os usuários da Rede também têm a possibilidade de encontrar o que procuram, mas, como João, precisam selecionar nas listas oferecidas pelos sites de busca, o que, muitas vezes, não é uma tarefa simples ou rápida, podendo gerar frustrações. A partir do momento em que os dados dos inúmeros documentos da coleção de selos de João forem descritos de forma mais estruturada, detalhada e se integrarem, a busca dele será muito mais bem sucedida. Desta forma, João precisa ter um banco de dados com os seus selos, outro com os países, outro com as cores, outro com as datas e integrá-los. Esta

é a filosofia da Web Semântica: integrar as informações da Rede, independentemente de onde estas estejam.

Como já descrito, a linguagem HTML se limita a oferecer informações sobre a estrutura de um documento da Web, indicando os parágrafos, fontes (estilos, cores e tamanhos) utilizadas, títulos, *links*, imagens, etc. As indicações aparecem entre os sinais gráficos < >. Como observado abaixo nos trechos de uma reportagem retirada do jornal O Globo on-line¹¹¹:

(a) indicação do título do texto: **<title>**MP denuncia PMs do caso AfroReggae - O Globo**</title>**

(b) indicação de parágrafo com o símbolo **<p>**: **<p>**RIO - O Ministério Público do Rio ofereceu nesta terça-feira denúncia contra os policiais militares Dennys Leonard Nogueira Bizzarro e Marcos de Oliveira Sales pelos crimes de prevaricação e furto qualificado, no caso da...

(c) indicação de imagem e de suas dimensões com o símbolo ****: ****

Por outro lado, a linguagem XML (*Extensible Markup Language*/Linguagem de Marcação Extensiva) se propõe a incluir informações semânticas em um documento (*tags* semânticas), descrevendo assim seus dados; é uma das linguagens utilizadas pela Web 3.0. Abaixo é possível observar um exemplo¹¹² de uma compra feita on-line em linguagem XML. Nota-se que as informações que estão entre os símbolos < > não são somente estruturais, mas metadados, isto é, dados sobre outros dados do texto como data da compra, país, nome do destinatário, nome da rua, número da residência, cidade, estado, CEP, dados de cobrança, comentários, nome do produto, etc. A seguir o exemplo com as *tags* semânticas entre os símbolos < >:

¹¹¹ Para o texto completo acessar: <http://oglobo.globo.com/rio/mat/2009/11/17/mp-denuncia-pms-do-caso-afroreggae-914803641.asp>

```

<?xml version="1.0"?>
<purchaseOrder orderDate="1999-10-20">
  <shipTo country="US">
    <name>Alice Smith</name>
    <street>123 Maple Street</street>
    <city>Mill Valley</city>
    <state>CA</state>
    <zip>90952</zip>
  </shipTo>
  <billTo country="US">
    <name>Robert Smith</name>
    <street>8 Oak Avenue</street>
    <city>Old Town</city>
    <state>PA</state>
    <zip>95819</zip>
  </billTo>
  <comment>Hurry, my lawn is going wild</comment>
  <items>
    <item partNum="872-AA">
      <productName>Lawnmower</productName>
      <quantity>1</quantity>
      <USPrice>148.95</USPrice>
      <comment>Confirm this is electric</comment>
    </item>
    <item partNum="926-AA">
      <productName>Baby Monitor</productName>
      <quantity>1</quantity>
      <USPrice>39.98</USPrice>
      <shipDate>1999-05-21</shipDate>
    </item>
  </items>
</purchaseOrder>

```

Endereço do destinatário.

Como afirma Berners-Lee (2001), a linguagem XML permite que qualquer um crie suas *tags* livre e arbitrariamente. É muito mais expressiva do que a linguagem HTML e mais fácil de ser entendida por ter muitas informações extras sobre os dados.

Além da linguagem XML, a Web Semântica utiliza a linguagem RDF¹¹³ (*Resource Description Framework*). Esta é uma linguagem de representação para identificação universal de recursos, mais especificamente, estabelece as relações entre os elementos de um documento em triplos (*triples*): sujeito – predicado – objeto. Estes triplos expressam significados dos documentos que são construídos em XML, como por exemplo, explicam que uma pessoa X tem propriedades (é irmão de Y ou autor do livro Z) e adere a certos valores (X e Y pertencem à família W vinda para o Brasil da Itália durante a Segunda Guerra

¹¹² Exemplo retirado do W3C em: <http://www.w3.org/TR/2004/REC-xmlschema-0-20041028/>.

¹¹³ Para mais detalhes acessar: <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>.

Mundial ou que o livro Z publicado por Y recebeu o prêmio Jabuti em 1979). A estrutura RDF parece ser uma forma de descrever grande parte dos dados existentes na Rede e criar as relações entre os elementos de um documento.

5.3.1.2.

A Web das ontologias sob a ótica de Wittgenstein

A terceira linguagem utilizada para a concretização da Web Semântica é a OWL (*Web Ontology Language*), linguagem das ontologias definida por Berners-Lee (2001) como um documento ou arquivo que define, de modo formal, as relações entre os termos. Inicialmente, é importante compreender o conceito de ontologia com base na filosofia.

Japiassú e Marcondes (2006) explicam que o termo ontologia se refere ao estudo da questão mais geral da metafísica, a do “ser enquanto ser”, ou seja, o ser considerado independentemente de suas determinações particulares e no que constitui sua inteligibilidade própria. Os autores prosseguem afirmando que, na filosofia, o termo “ontologia” é muitas vezes sinônimo de “metafísica”. De acordo com a definição de Diderot e D’Alembert na *Encyclopédia*, a metafísica se refere ao ramo da filosofia que examina as propriedades gerais dos seres humanos como a existência, a possibilidade e a duração; é a ciência do ser. Como complementa Marcondes (2007), a ontologia consiste na análise do que realmente existe, isto é, do real.

A Inteligência Artificial (IA) e a Web se apropriam do termo “ontologia”, como define Berners-Lee (2001), para designar um documento que define, formalmente, as relações entre os termos. Como mostra Freitas (2007), as ontologias devem ser gerais e independentes de língua, pois estas descrevem a realidade, considerada a mesma para todos. No que se refere à concepção da IA, as fontes computacionais devem ser as mais gerais possíveis, reutilizáveis e compartilháveis entre a própria comunidade de IA. Por outro lado, esta noção de ontologias gerais na IA parece ser problemática. Como alternativa, passou-se a investigar e criar ontologias específicas, com objetivo de aprimorar e facilitar as buscas semânticas.

A linguagem OWL tem como base a taxonomia, isto é, a classificação de palavras e hierarquia de termos, suas propriedades e relações. Não há dúvidas de que a OWL pode aprimorar muito os sistemas de busca da Rede, fazendo com que estes apresentem ao usuário somente as páginas relevantes aos seus

interesses e não todas as páginas relacionadas aos termos utilizados na busca, independentemente do contexto e significado.

É possível observar alguns exemplos já existentes de ontologias na Rede. O primeiro selecionado foi a Cyc¹¹⁴, base de conhecimento multicontextual desenvolvida pela empresa americana Cycorp. Esta empresa é provedora de tecnologia semântica fundada em 1984. Tem como objetivo representar formalmente o vasto conhecimento humano em inúmeras áreas e possui no momento mais de 200 mil termos com explicações. Estas explicações vão sendo desenvolvidas constantemente por pessoas especializadas, visando reduzir as ambiguidades que possam não ser reconhecidas pelos sistemas. Um dos exemplos oferecidos pelo Cyc é:

- (a) *Fred saw the plane flying over Zurich.* (Fred viu o avião sobrevoando Zurique.)
- (b) *Fred saw the mountains flying over Zurich.* (Fred viu as montanhas quando sobrevoava Zurique.)

Mesmo os exemplos sendo bem similares, um falante de língua inglesa não teria dificuldades em entender que no exemplo (a) o verbo *flying* se refere ao avião, já no exemplo (b), o mesmo verbo se refere a Fred. No entanto, sem o auxílio da Web Semântica, parece praticamente impossível que um sistema tradicional da Rede consiga fazer essa diferenciação. Por outro lado, como o sistema de ontologias Cyc sabe que aviões voam, ao contrário das montanhas, será possível analisar os exemplos e descartar interpretações sem sentido, como dizer que as montanhas estavam voando.

Com uma base de conhecimento muito grande, a Cyc apresenta uma pirâmide para que seja possível visualizar todas as categorias ontológicas existentes. A pirâmide é dividida em áreas e cada uma destas representa uma categoria. O texto com explicações logo abaixo da pirâmide se altera conforme as categorias da pirâmide são clicadas. A seguir é possível visualizar a pirâmide completa:

¹¹⁴ <http://www.cyc.com/>

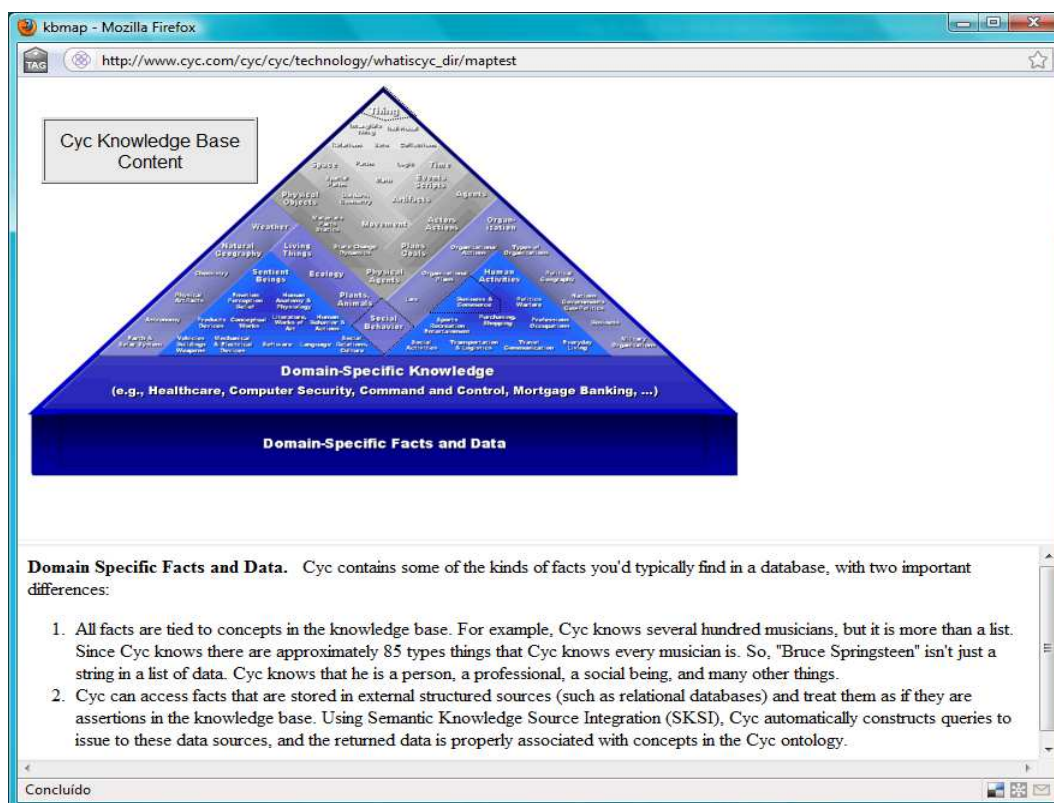


Figura 17. Reprodução da tela da base de conhecimento da Cyc. Disponível em: http://www.cyc.com/cyc/cyc/technology/whatis_cyc_dir/maptest - Acesso em 19 de novembro de 2009.

O segundo exemplo de ontologia é a *WordNet*¹¹⁵. Este é um léxico semântico desenvolvido para língua inglesa na universidade de Princeton nos Estados Unidos e disponível para uso on-line. Na *WordNet*, substantivos, verbos, adjetivos e advérbio são separados em grupos de sinônimos cognitivos, cada grupo é chamado de *synset*. Cada um destes grupos expressa um conceito distinto. De acordo com Freitas (2007) o objetivo desta ontologia é oferecer uma combinação de dicionário e tesouro e dar suporte a tarefas que envolvem análise automática de textos. A maior parte dos *synsets* se conecta com outros *synsets* por meio de relações semânticas. Até o início deste ano, já havia mais de 150.000 palavras organizadas em mais de 115.000 *synsets*. A seguir é possível observar a tela de busca do *WordNet* com os resultados para a busca da palavra *book*:

¹¹⁵ <http://wordnet.princeton.edu/>

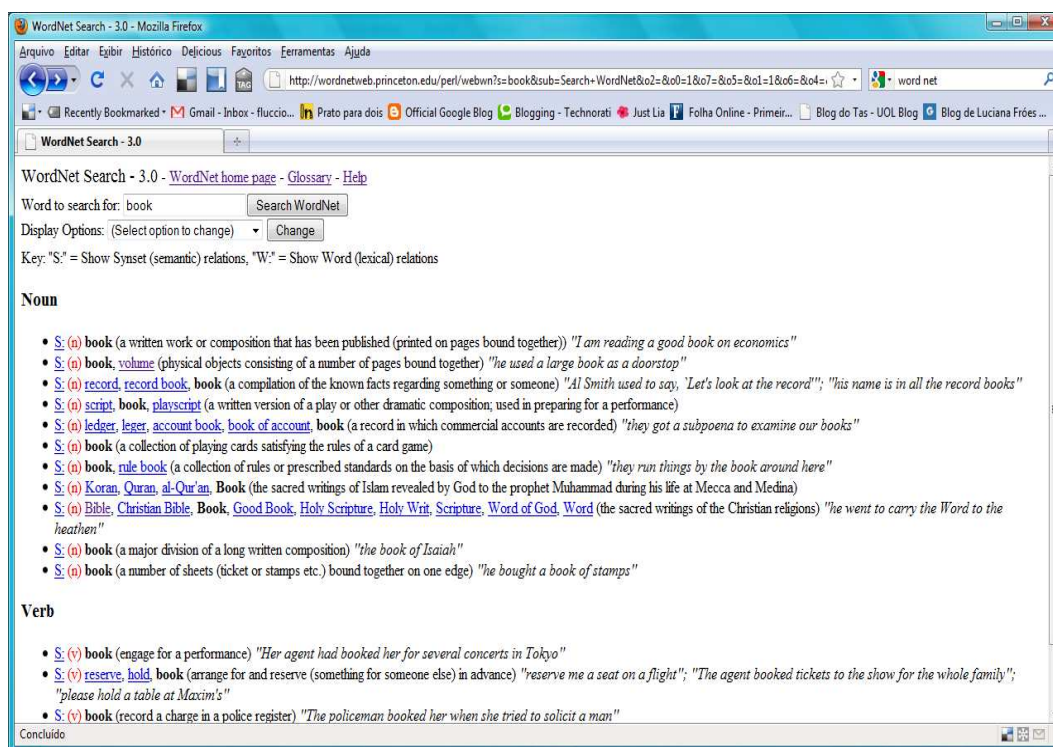


Figura 18. Reprodução da tela de busca do WordNet com o resultado para *book*. Disponível em: [b.princeton.edu/perl/webwn?s=book&sub=Search+WordNet&o2=&o0=1&o7=&o5=&o1=1&o6=&o4=&o3=&h=0000000](http://wordnetweb.princeton.edu/perl/webwn?s=book&sub=Search+WordNet&o2=&o0=1&o7=&o5=&o1=1&o6=&o4=&o3=&h=0000000) - Acesso em 19 de novembro de 2009.

É possível encontrar também ontologias multilíngues. A *EuroWordNet*¹¹⁶ é descrita no próprio site como uma base de dados com termos em 7 línguas (holandês, italiano, espanhol, alemão, francês, checo e estoniano) e segue a mesma estrutura da americana *WordNet*. O diferencial é a possibilidade de conectar as línguas, ou seja, ser possível ir de uma palavra em uma das 7 línguas para outras similares nas outras línguas disponíveis. Além disso, é possível encontrar as distinções semânticas entre as línguas. Este projeto foi criado em 1999 e está inativo no momento. No entanto, várias instituições e grupos de pesquisa estão desenvolvendo ontologias similares. Existe também a *WordNet.Br*¹¹⁷ que é a versão brasileira da *WordNet*. De acordo com Freitas (2007) a maior parte da elaboração deste trabalho é feita manualmente e reaproveita materiais disponíveis em outras fontes como dicionários de sinônimos, antônimos e verbos. A *WordNet.Br* faz parte dos projetos existentes no NILC (Núcleo Interinstitucional de Linguística Computacional)¹¹⁸ na USP. É importante destacar também a *WordNet.Pt*¹¹⁹,

¹¹⁶ <http://www.ilc.uva.nl/EuroWordNet/>.

¹¹⁷ <http://www.nilc.icmc.usp.br/~arianidf/WordNet-BR.html>. Este projeto ainda está em andamento na UNESP com apoio do CNPQ.

¹¹⁸ <http://www.nilc.icmc.usp.br/nilc/>

base de dados de conhecimento linguístico do português de Portugal e o HAREM¹²⁰, que tem como objetivo avaliar o sucesso na identificação e classificação dos nomes próprios da língua portuguesa. Este projeto faz parte da Linguateca¹²¹ que é um centro de recursos para o processamento computacional da língua portuguesa, com pólos em Portugal, na Dinamarca e na Noruega.

Na prática, as páginas da Rede estão ligadas às ontologias existentes na Web que definem informações relacionadas. Berners-Lee (2001) oferece como exemplo os professores universitários que geralmente têm doutorado. Os conceitos da OWL são utilizados para especificar, por exemplo, que o professor Hendler concluiu seu doutorado na instituição <http://brown.edu> (o *link* já é oferecido no momento em que o nome da universidade é apresentado). Também é possível saber que o professor Hendler é membro de um determinado grupo de pesquisadores (URL é disponibilizada neste caso também) e tem o email hendler@brown.edu; seu *blog*, perfil no *LinkedIn*, *Facebook*, *Flickr* e *Twitter* também são disponibilizados. Sendo assim, os dados relacionados ao professor Hendler estão interligados.

Apesar das ontologias já existentes e de outras que estão sendo criadas, é difícil imaginar a elaboração de conceitos únicos, universais e sem ligação com as línguas. Parece tratar-se de uma tentativa de estabelecer uma verdade absoluta, desconsiderando a linguagem. Esta concepção se assemelha à visão Agostiniana de linguagem. O filósofo Wittgenstein, em sua segunda fase (*Investigações Filosóficas* 1958), apresenta uma concepção revolucionária da linguagem, criticando e questionando o mito existente que via a linguagem como um conjunto de regras. Sua filosofia é uma crítica à visão Agostiniana de linguagem, que pressupõe que o papel da linguagem é o de descrever objetos, dar nome às coisas, ou seja, representar a realidade. Esta é a visão da aquisição da linguagem de acordo com a tradição na qual a unidade básica do significado é a palavra. Wittgenstein (2004/1958) ressalta que, na visão de linguagem de Santo Agostinho, toda palavra tem um significado e o significado é o objeto que a palavra designa.

De acordo com Wittgenstein (2004/1958), a linguagem não deve ser concebida como um sistema complexo baseado em correlações corretas entre as palavras e as coisas. Correlacionar palavras e objetos é, para o filósofo, insuficiente para dar conta do significado dos objetos designados, pois uma

¹¹⁹ <http://cvc.instituto-camoes.pt/wordnet/index.html>

¹²⁰ <http://www.linguateca.pt/HAREM/>

¹²¹ <http://www.linguateca.pt/>

mesma palavra pode designar vários objetos que não têm nada em comum entre si e criar assim uma confusão de significados. Wittgenstein (2004/1958) deixa claro que considera a posição de Santo Agostinho muito limitada, pois não fala de uma diferença de espécies de palavras:

“Poderíamos dizer que Santo Agostinho descreve um sistema de comunicação; só que nem tudo que chamamos de linguagem é este sistema. E isto precisa ser dito em certos casos onde se levanta a questão: ‘Esta exposição é útil ou inútil?’ A resposta: é útil; mas somente para este domínio estritamente circunscrito, não para a totalidade que você pretendia expor.” (Wittgenstein, 2004/1958, p.16, §3)

Para Faustino (1995), a imagem Agostiniana da linguagem é apresentada por Wittgenstein em *Investigações Filosóficas* como a que identifica o ato de “significar” com o ato de “apontar” o significado de um signo. Wittgenstein nos diz que a “gramática do significar” não é semelhante à da expressão “representar-se algo”. Ele nos chama a atenção para o fato de que usar a linguagem como representação de objetos, como defendido por Santo Agostinho com sua visão cartesiana de linguagem, é completamente diferente de usar a linguagem como instrumento de ação. Wittgenstein contrapõe a visão Agostiniana de linguagem afirmando que nem sempre usamos a linguagem para representar e, quando a usamos, tal representação não é de um objeto e sim do social. O significado não é determinado pela referência, ele nos é dado pelo contexto de prática de uma língua, pois as palavras fazem parte de uma prática, “O significado de uma palavra é seu uso na linguagem.” (Wittgenstein, 1958/2004, p.38, §43). Desta forma: “Todo signo, sozinho, parece morto. O que lhe confere vida? – Ele está vivo no uso.” (Wittgenstein, 1958/2004, p.173, §432).

Sendo assim, a visão Agostiniana de linguagem parece criar dificuldades quando se trata da elaboração de ontologias na Web. Deve-se levar sempre em consideração que as línguas são dinâmicas e estão constantemente sofrendo mudanças, de acordo com o grupo que a utiliza, o local e a época. Os significados não são estáticos ou fixos, eles variam de acordo com contexto, ou seja, com os diferentes usos. A totalidade formada pela linguagem e pelas atividades com as quais ela vem entrelaçada é denominada por Wittgenstein de jogos de linguagem (*language-games*). Uma mesma palavra pode participar de vários contextos, ser usada de diversas formas e ter significados variados. Os jogos de Wittgenstein procuram mostrar que a linguagem é parte de uma atividade, uma ação ou uma forma de vida. Como pode ser observado no trecho: “A expressão ‘jogo de linguagem’ deve salientar aqui que falar uma língua é parte de uma atividade ou de uma forma de vida.” (Wittgenstein, 1958/2004,

p.26, §23). Os jogos de linguagem são muito diversificados e podem ser usados de diferentes formas como para: ordenar, pedir, agradecer, inventar, representar, descrever, traduzir, relatar, cantar, praguejar, etc. Existem diferentes jogos de linguagem, com objetivos distintos. No entanto, todos esses jogos possuem uma essência ou característica básica definidora que é compartilhada. Tais características básicas semelhantes foram denominadas por Wittgenstein de “semelhança de família”, ou seja, os jogos são jogados de formas diferentes, com objetivos diferentes, mas com propriedades comuns, como ocorre com os membros de uma mesma família.

“...vemos uma complicada rede de semelhanças que se sobrepõem uma às outras e se entrecruzam. Semelhanças em grande e em pequena escala. Não posso caracterizar melhor essas semelhanças do que por meio das palavras ‘semelhanças familiares’; pois assim se sobrepõem e se entrecruzam as várias semelhanças que existem entre os membros de uma família... E eu direi: os ‘jogos’ formam uma família.” (Wittgenstein, 1958/2004, p.52, §66/67).

A partir dos jogos de linguagem, o foco principal são os objetivos de quem usa a linguagem e o contexto de uso. Como afirma Marcondes (2000):

“Assim sendo, podemos afirmar que ao investigarmos a linguagem estamos investigando igualmente a sociedade da qual ela é linguagem, o contexto social e cultural na qual é usada, as práticas sociais, os paradigmas e valores, a ‘racionalidade’ desta comunidade; e desta forma pode-se dizer que não há uma separação radical entre ‘linguagem’ e ‘mundo’ já que a ‘realidade’ é constituída exatamente pelo modo como aprendemos a linguagem e a usamos.” (Marcondes, 2000, p. 17)

A partir da visão wittgensteiniana de linguagem e da noção de jogos é possível questionar a criação de ontologias na Rede. Parece que a Rede está se encaminhando para uma tentativa de uniformização das ontologias para fazer com que os documentos existentes on-line se comuniquem e se integrem. No entanto, seguindo o raciocínio de Wittgenstein, parece que a OWL é um dos grandes desafios da Web Semântica, pois para se criar uma ontologia universal on-line, seria necessário uma padronização dos conceitos, culturas, valores, assim como desconsiderar as diferenças linguísticas existentes no mundo. Este caminho é pouco provável, se levarmos em consideração a concepção da linguagem de Hacker a seguir (2000, p.16): “Nós somos essencialmente seres que utilizam linguagem. Nossa linguagem e as formas de nossa linguagem moldam nossa natureza, dão forma ao nosso pensamento, preenchem nossa vida.” Como complementa Marcondes (2000):

“Ao usar a linguagem, estamos agindo em um contexto social, e nossos atos são significativos e eficazes apenas na medida em que correspondem às determinações destas ‘formas de vida’, destas práticas e instituições sociais. Em seu uso da linguagem os falantes seguem regras não apenas lingüísticas *stricto sensu* (isto é, gramaticais, fonéticas e semânticas), mas sobretudo *pragmáticas*... o significado das expressões não depende essencialmente das intenções dos falantes, mas, na verdade, as intenções são formadas e tornadas possíveis elas próprias a partir dos hábitos, práticas e instituições de uma comunidade lingüística.” (Marcondes, 2000, p. 38)

Wittgenstein (1958) mostra que nossas formas de representar o mundo não podem ser consideradas independentes das práticas sociais. A linguagem é uma prática social e faz parte desta. Como diz Marcondes (2000, p.38) sobre Wittgenstein e os jogos de linguagem: “Na realidade, o que o indivíduo aprende não é pura e simplesmente uma palavra, ou expressão, mas um jogo de linguagem completo... como usar uma determinada expressão lingüística em um contexto determinado para obter certos fins.” Existem em nossa prática social, on-line ou off-line, inúmeros jogos de linguagem, ao usá-los, estamos agindo em um contexto social onde nossos atos serão significativos. Coutinho (1985) complementa:

“A linguagem é vista como uma prática social. Aprender uma linguagem requer mais do que o aprender de uma linguagem no sentido sintático e semântico usual. Aprender uma linguagem é introduzir-se no conjunto de regras explícitas e implícitas que constituem as convenções sociais – as instituições, as atitudes, as crenças, os valores, as intenções, os comportamentos, etc. – de uma determinada “forma de vida” (cultura). Seguir uma regra, nessa concepção, significa então aprender os usos polimorfos dos termos, isto é, os “jogos de linguagem”. A noção de “gramática” é redefinida para incluir a totalidade das regras implicadas pela prática lingüística... uma visão fluida e abrangente da linguagem sendo esta constituída por uma multiplicidade de “jogos de linguagem” que se entrecruzam e cujos significados não podem ser abstraídos do contexto social de uso.” (Coutinho, 1985, p.76)

O grande desafio de Wittgenstein foi justamente redefinir a concepção de significado sem utilizar somente as regras lingüísticas, que relacionam as palavras com os objetos, mas levando em consideração como as pessoas usam as palavras, em quais condições e contextos e, principalmente, o que elas fazem (significam) ao usarem determinadas palavras e expressões. Como afirma Marcondes (1999), Wittgenstein nos mostra em seus primeiros parágrafos de *Investigações Filosóficas* um “jogo de linguagem” como uma combinação de palavras, atos, atitudes e formas de comportamento; ou seja, é como compreendemos o processo de uso da linguagem em sua totalidade. Sendo assim, não aprendemos pura e simplesmente uma palavra ou expressão, mas como usá-las em um contexto determinado para obter certos fins. A OWL como

uma proposta de linguagem de representação universal na Rede parece estar fora da realidade humana. A questão que permanece é: será que a terceira geração da Rede conseguirá concretizar o projeto de uma ontologia universal para que a Web Semântica seja realmente possível e eficaz?

5.3.1.3.

A concretização da Web 3.0

Ainda é cedo para tentar encontrar uma resposta para o questionamento feito acima. O fato é que, apesar de todas as dificuldades encontradas na criação de uma OWL universal e na real integração dos dados na Rede, já existe uma ferramenta de busca em funcionamento e se chama *Wolfram Alpha*¹²². Esta é descrita, na página inicial de seu site, como ferramenta computacional de conhecimento e foi lançada em maio de 2009 pelo matemático Stephen Wolfram. O *Wolfram Alpha* é um projeto muito novo e ambicioso que se propõe a responder perguntas feitas pelos usuários sobre os mais variados temas, ou seja, pretende fazer todo conhecimento sistemático processável por qualquer pessoa. Em um vídeo¹²³, no qual seu criador explica com vários exemplos sua ferramenta de busca, Stephen Wolfram afirma que este é um projeto longo que tenta reunir a maior parte possível de conhecimento existente no mundo e torná-lo computável para que qualquer pessoa em qualquer parte do mundo possa acessá-lo. A seguir é possível observar a página inicial desta ferramenta de busca:

¹²² <http://www.wolframalpha.com/>

¹²³ <http://www.wolframalpha.com/screencast/introducingwolframalpha.html>

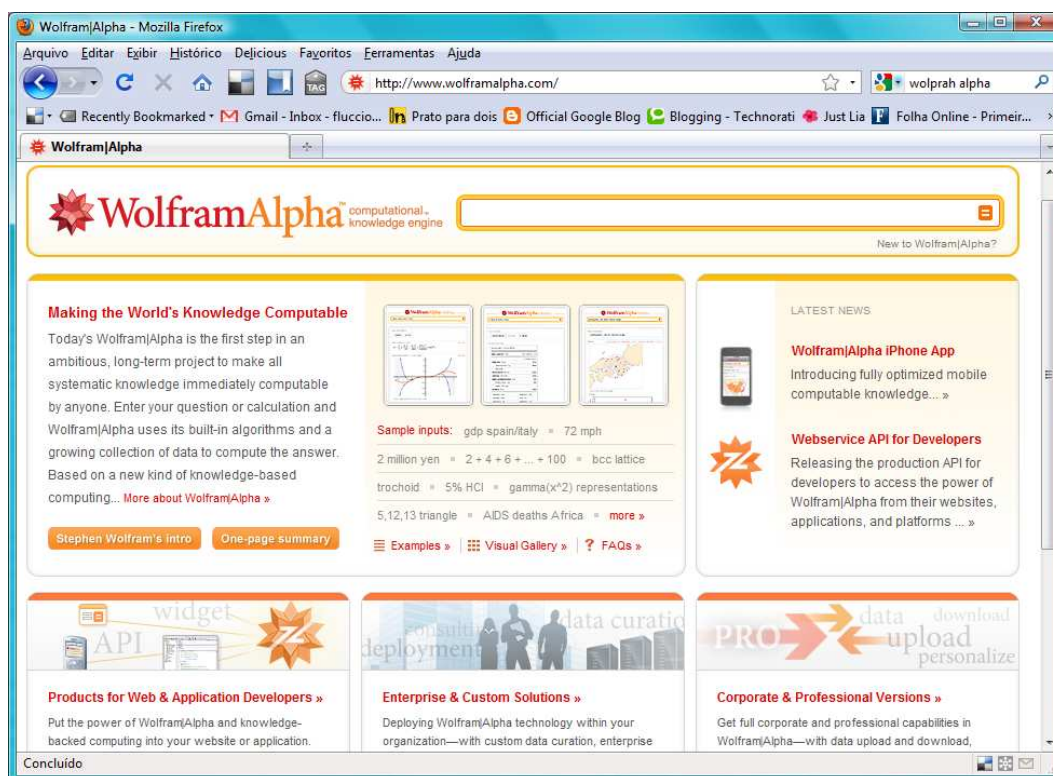


Figura 19. Reprodução da tela inicial do *Wolfram Alpha*. Disponível em: <http://www.wolframalpha.com/> - Acesso em 19 de novembro de 2009.

Nesta tela inicial é possível perceber um espaço, no topo da página, reservado para o usuário escrever sua pergunta, expressão, termo, equação, ou seja, qualquer tipo de informação que busca. É possível encontrar também exemplos e uma lista com vários assuntos já disponíveis para consulta. As informações são divididas por categorias (como por exemplo, “mapas e geografia”) e cada categoria possui uma lista com subcategorias (como mapas, distâncias, países, cidades, lagos, etc). A busca pode ser feita por esses tópicos já existentes, ou no espaço reservado para busca no topo da página. Se o usuário fizer uma busca, por exemplo, pelas palavras “apple” e “orange”, a ferramenta irá automaticamente entender que “apple” se refere a uma fruta e que o usuário talvez esteja procurando informações comparativas sobre essas duas frutas. O *Wolfram Alpha* disponibilizará todas as informações nutricionais com tabelas e gráficos. É importante lembrar que esta ferramenta inclui, no canto direito da tela, um *link* para que esta mesma busca seja feita no Google. Como é possível observar na tela que se segue:

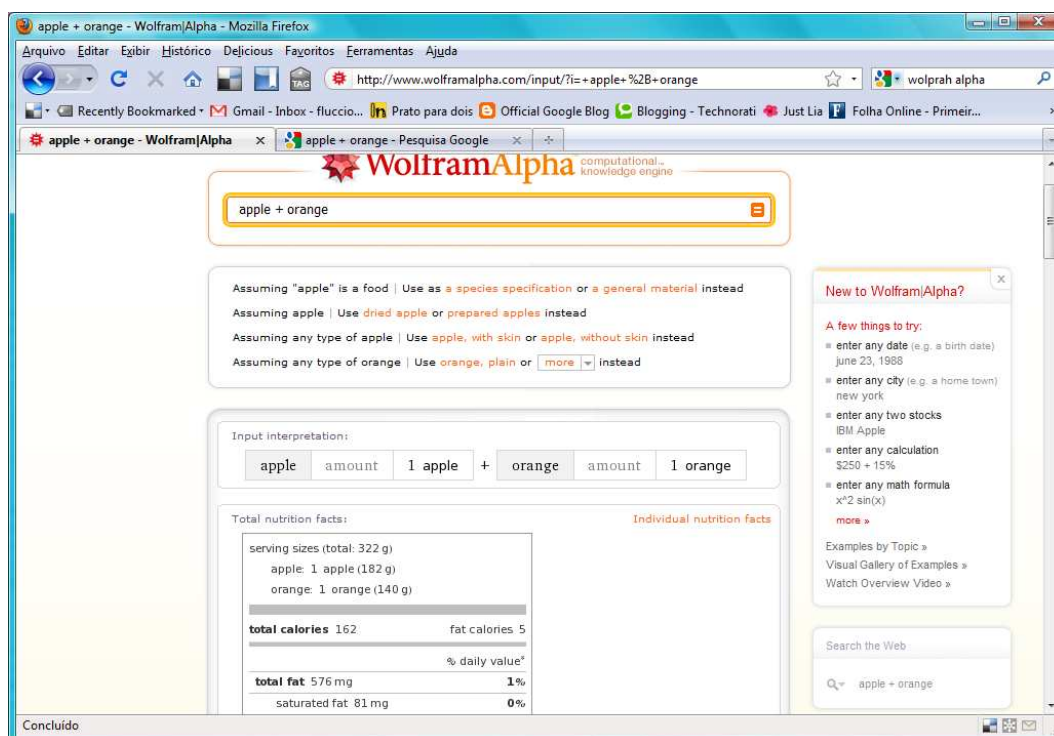


Figura 20. Reprodução do resultado da pesquisa “apple+orange” no Wolfram Alpha. Disponível em: <http://www.wolframalpha.com/input/?i=apple+%2B+orange> - Acesso em 19 de novembro de 2009.

Se o usuário alterar a sua busca para “apple” e “ibm” a ferramenta entenderá, automaticamente, que a busca é pela comparação das duas empresas. O resultado mostrará, portanto, tabelas, gráficos e informações financeiras:

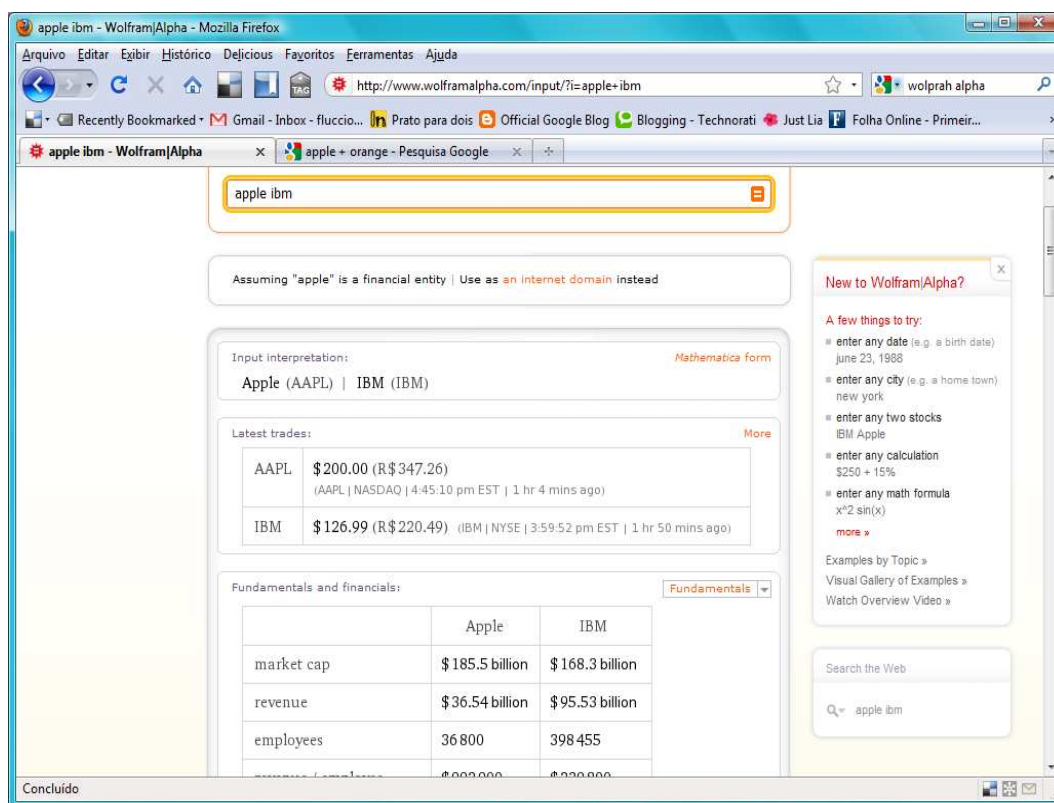


Figura 21. Reprodução do resultado da pesquisa “apple+ibm” no *Wolfram Alpha*. Disponível em: <http://www.wolframalpha.com/input/?i=apple+ibm> - Acesso em 19 de novembro de 2009.

Se o usuário digitar somente “apple” o sistema escolherá, aleatoriamente, mostrar como resultado informações sobre a empresa ou sobre a fruta e oferecerá ao usuário a possibilidade de clicar em *links* que o levarão para os resultados sobre a fruta ou uma espécie de planta ou sobre um domínio da Internet ou sobre uma nave espacial. É possível observar essas possibilidades no segundo retângulo da tela a seguir:

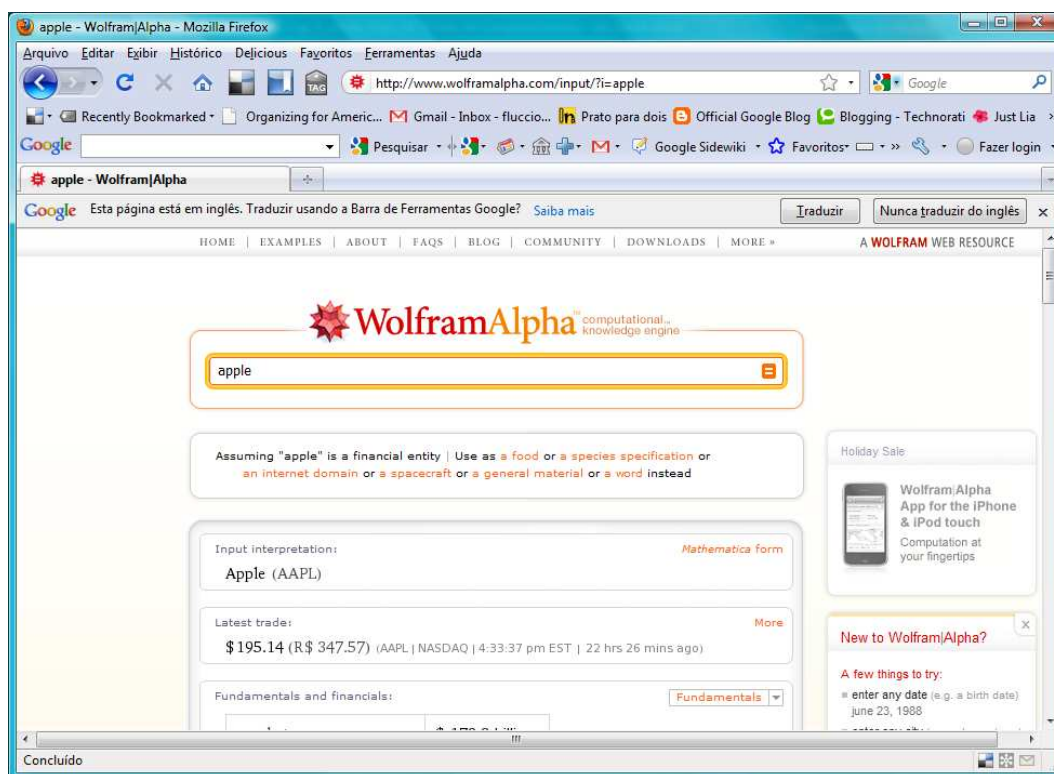


Figura 22. Reprodução do resultado da pesquisa “apple” no Wolfram Alpha. Disponível em: <http://www.wolframalpha.com/input/?i=apple> - Acesso em 19 de novembro de 2009.

A ferramenta de busca fará o mesmo, caso o usuário busque somente por “orange”. O sistema escolherá automaticamente um resultado e disponibilizará *links* que levará o usuário a outras possibilidades, como informações sobre a fruta, espécie de planta, cidade, nome de pessoa, etc.

No caso dessas mesmas pesquisas serem feitas no buscador Google, o usuário vai receber uma lista interminável de *links* desorganizados e não conseguirá entender no primeiro caso (Figura 20) que “apple” é uma fruta e não a empresa. Desta forma, incluirá, em sua lista, informações sobre as frutas e as empresas e caberá ao usuário selecionar os *links* de seu interesse e fazer as comparações. Como se observa a seguir, a pesquisa referente à “apple” e “orange” resultou em uma mistura de resultados entre as frutas e a empresa Apple:

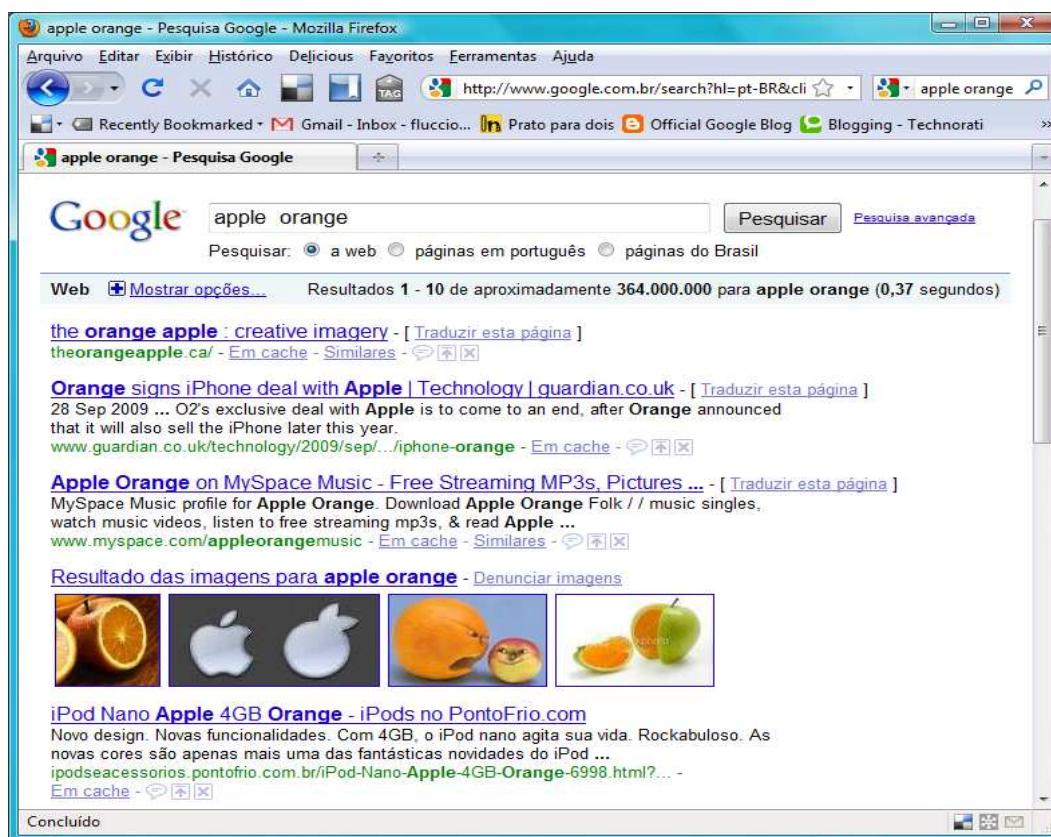


Figura 23. Reprodução do resultado da pesquisa “apple+orange no Google”. Disponível em: <http://www.google.com.br/search?hl=pt-BR&client=firefox-a&rls=org.mozilla%3Apt-BR%3Aofficial&hs=OEx&q=apple++orange&btnG=Pesquisar&meta=&aq=f&oq=> - Acesso em 19 de novembro de 2009.

Por outro lado, ao se pesquisar no Google “apple” e “ibm”, a ferramenta entende que o usuário busca informações sobre as empresas. No entanto, não as ordena comparativamente. Isso mostra que o Google já está aprimorando sua busca, mas ainda há um longo caminho a ser percorrido. Em março de 2009, o Google anunciou aprimoramentos em sua ferramenta de busca, visando expandir a lista dos itens relacionados a uma busca e apresentando descrições de resultados mais longas (palavras e expressões em negrito em cada *link* resultante de uma busca). Para que essas melhorias sejam possíveis, a Web Semântica já está sendo utilizada¹²⁴ pelo Google. Como afirma Berners-Lee (2001), se desenvolvida de forma apropriada, a Web Semântica poderá contribuir para a evolução do conhecimento humano como um todo.

¹²⁴ Mais detalhes no *blog* oficial do Google: <http://googleblog.blogspot.com/2009/03/two-new-improvements-to-google-results.html>.