

Referências

- ALBÓ, X. *Cultura, interculturalidade, inculturação*. São Paulo: Edições Loyola, 2005.
- ASSIS-PETERSON, A. A.; COX, M. I. P. O inglês em tempos de globalização: para além do bem e mal. *Calidoscópico*, v. 5, n. 1, p. 5-14, jan./abr 2007.
- CANDAU, V. M. F. Direitos humanos, educação e interculturalidade: as tensões entre igualdade e diferença. *Revista Brasileira de Educação*, v. 13, p. 45-56, 2008.
- CANEVACCI, M. *Sincretismos: uma exploração das hibridizações culturais*. São Paulo: Studio Nobel, 1996.
- COX, M. I. P.; ASSIS-PETERSON, A. A. Transculturalidade e transglossia: para compreender o fenômeno das fricções lingüístico-culturais em sociedades contemporâneas sem nostalgia. In: CAVALCANTI, M. C.; BORTONI-RICARDO S. M. (Orgs.). *Transculturalidade, linguagem e educação*. Campinas: Mercado de Letras, 2007. p. 23-43.
- FAIRCLOUGH, N. Teoria social do discurso. In: FAIRCLOUGH, N. *Discurso e mudança social*. Tradução de Izabel Magalhães. Brasília: Editora Universidade de Brasília, 2001 [1992]. p. 89-131.
- FREIRE, P. *Pedagogia do oprimido*. Rio de Janeiro: Paz e Terra, 2005.
- HALL, S. *A identidade cultural na pós-modernidade*. Rio de Janeiro: DP&A, 2006 [1992].
- HOOKS, b. *Teaching to transgress: education as the practice of freedom*. New York: Routledge, 1994.
- KUMARAVADIVELU, B. A lingüística aplicada na era da globalização. In: MOITA LOPES, L. P. (Org.). *Por uma lingüística aplicada indisciplinar*. São Paulo: Parábola, 2006. p. 129-148.
- MOITA LOPES, L. P. Lingüística aplicada e vida contemporânea: problematização dos construtos que tem orientado a pesquisa. In: MOITA LOPES, L. P. (Org.). *Por uma lingüística aplicada indisciplinar*. São Paulo: Parábola, 2006. p. 85-107.
- ORTIZ, R. *Mundialização: saberes e crenças*. São Paulo: Brasiliense, 2006.
- PENNYCOOK, A. A lingüística aplicada dos anos 90: em defesa de uma abordagem crítica. In: SIGNORINI, I.; CAVALCANTI, M. C. (Orgs.). *Lingüística aplicada e transdisciplinaridade*. Campinas: Mercado de Letras, 1998. p. 23-49.
- PIMENTEL DA SILVA, M. S. As línguas indígenas na escola: da desvalorização à revitalização. *Signótica*, v. 18, n. 2, p. 381-395, jul./dez. 2006.
- RAJAGOPALAN, K. Postcolonial world and postmodern identity: some implications for language teaching. *DELTA*, v. 21, Fac. Especial, p. 11-20, 2005. Disponível em: <http://www.scielo.br/scielo.php?pid=S0102-44502005000300003&script=sci_arttext>. Acesso em: 11 jun. 2009.



A SOCIEDADE DA INFORMAÇÃO E DO CONHECIMENTO: PRESENTE E FUTURO

João Carlos da Silva¹ e Cedric Luiz de Carvalho²

Introdução

Nos dias atuais, a quantidade de dados disponíveis no mundo, tanto para os indivíduos quanto para as organizações, é gigantesca e cresce vertiginosamente a cada dia. Individualmente, as pessoas armazenam em suas máquinas dados nos mais diversos formatos. Por exemplo, uma pessoa pode ter guardados em seu computador arquivos de vídeo, codificados no formato MPEG4, músicas no formato MP3, planilhas, documentos em formato texto, dados armazenados em um banco de dados, etc. As instituições, da mesma forma, também produzem e armazenam dados nos mais diversos formatos, gerados por diversos indivíduos, departamentos e processos.

A internet, especialmente a sua porção multimídia, a *World Wide Web*, ou simplesmente *web*, vem se consolidando mundialmente como um campo extremamente fértil para o compartilhamento e disseminação de dados dos mais variados tipos, sejam de natureza acadêmica, científica, comercial ou até mesmo pessoal. Através dela, indivíduos e organizações podem colocar seus dados à disposição.

Essa facilidade oferecida pela internet tem provocado um crescimento astronômico da quantidade de dados disponíveis na *web*. Esse crescimento gera

1. Doutor em Ciência da Computação (University of East Anglia) e professor do Instituto de Informática da UFG.

2. Doutor em Ciência da Computação (UFMG) e professor do Instituto de Informática da UFG.

uma dificuldade cada vez maior na obtenção de informação específica de uma forma rápida. Além das deficiências de comunicação, os próprios métodos de recuperação estão muito aquém dos desejos e necessidades daqueles que deles fazem uso.

O estrondoso crescimento do número de páginas e documentos disponíveis na *web* tem dificultado, cada vez mais, a localização, o acesso e a manutenção de dados na rede. Esses problemas vêm do fato que usuários, com diferentes níveis de conhecimento, oferecem uma vasta quantidade de recursos, sem nenhuma padronização.

Surge daí a insatisfação dos próprios usuários, pois quando estes realizam uma pesquisa na *web*, geralmente devem esperar um tempo prolongado para obter um resultado que lhes seja satisfatório. Normalmente, os mecanismos de busca hoje disponíveis retornam, como resultado das consultas, um grande volume de informações, referenciadas por endereços de páginas da internet. Não são raras as vezes em que recebemos milhares de endereços (algumas vezes milhões) como resultado de uma consulta. Fica então a cargo do usuário a filtragem destes endereços para obter respostas às suas consultas.

Tome-se como exemplo o seguinte contexto: um docente da UFG deve participar de um congresso em São Paulo, no período de 14/9/2010 a 18/9/2010. Para marcar a data da viagem e providenciar reserva de hotel, ele está interessado em obter informações sobre disponibilidade de voos, ida e volta, com tarifas econômicas. Além disso, ele também está interessado em obter informações sobre hotéis próximos ao local do evento, dentro de uma faixa específica de preços, para o período desejado.

No cenário atual, o interessado deverá executar uma sequência, possivelmente longa, de consultas a páginas da internet. Primeiramente, podem-se consultar sítios de companhias aéreas em busca de horários de voos convenientes

para as viagens de ida e de volta, com tarifas atraentes. Essa consulta será executada pelo menos uma vez para cada companhia aérea disponível e, a cada repetição, o usuário deve fornecer datas de ida e de volta, cidades de origem e de destino, e quantidade de passageiros. O resultado recebido inclui todos os voos que satisfazem os dados fornecidos pelo usuário, cabendo a ele escolher o que melhor lhe convém. Como o usuário deve consultar diferentes companhias aéreas, somente depois de várias tentativas, analisando melhores horários e tarifas, é que ele chegará a alguma conclusão.

Um trabalho semelhante terá de ser executado para obter informações sobre hotéis. Neste processo, o usuário analisará conceitos como disponibilidade no período desejado, localização e tarifas.

Depois de tantas idas e vindas, de cansativas consultas a sítios de companhias aéreas, de pesquisas aparentemente intermináveis a guias turísticos em busca de hotéis, a tarefa inicial parece ter chegado ao fim. Infelizmente não. Ainda resta ao interessado debruçar-se sobre uma pilha de anotações tomadas, de forma a combinar os resultados de “melhor opção de viagem” com “melhor opção de hotel”.

Como pode ser constatado através do exemplo anterior, a *web* foi concebida para ser utilizada por seres humanos. O inconveniente desta forma em que a *web* está estruturada é que o desenvolvimento de ferramentas automáticas, voltadas para a recuperação dos dados nela armazenados, é uma tarefa bastante complicada para ser realizada com os recursos tecnológicos disponíveis atualmente. A questão central é: como extrair, automaticamente, conhecimentos a partir dos dados disponíveis, utilizando computadores. Ou ainda, como identificar o conhecimento embutido em dados.

Dado, informação e conhecimento

Num mundo cada vez mais competitivo, mais do que dados, as pessoas e as organizações necessitam das informações a eles associadas e, conseqüentemente, do conhecimento que estas informações podem gerar. O conhecimento é um valor essencial para pessoas que são responsáveis pelo processo de tomada de decisão. Nos países mais avançados, o acúmulo de informações para estruturar esquemas de conhecimentos cada vez maiores e mais sofisticados permitiu mudanças qualitativas no setor econômico, por exemplo. Hoje em dia, o poderio econômico internacional de um país está diretamente relacionado à sua capacidade de gerar novos conhecimentos.

A *web* é um caminho de acesso à informação espalhada pela internet. É um modelo de domínio público que foi construído para a busca de informações. Se por um lado a tecnologia e o desenvolvimento das redes de comunicações possibilitam a difusão do conhecimento, por outro lado impulsionam a publicação direta da fonte ao consumidor. Este contexto propaga falta de padrões para oferta de documentos ou informação na internet e, conseqüentemente, dificulta a busca e a recuperação da informação.

O processamento da informação é uma atividade complexa, uma vez que, dependendo do contexto e do domínio do conhecimento, pode ter significados diversos, ou seja, um termo ora representa um conceito, ora pode representar outro e constituir relações conceituais diferentes.

O cenário exemplificado na seção anterior poderia dar origem a uma consulta formulada da seguinte maneira: “Quais são os horários de voos de Goiânia para São Paulo, com ida em 14/9/2010 e volta em 18/9/2010,

para um passageiro, com tarifa econômica, e quais hotéis próximos ao Masp, com diária de solteiro entre 100 e 200 reais, têm vaga para entrada em 14/9/2010 e saída em 18/9/2010?”

Este exemplo de consulta contém dados que fazem referência a cidades, datas, valores monetários, etc. Porém, para dar uma resposta que satisfaça às expectativas do usuário, computadores devem ser dotados de capacidade para analisar semanticamente os dados disponíveis. Além disso, devem ser capazes de executar mecanismos que acrescentem habilidades de interpretação de conceitos, como, por exemplo: viagem, tarifa econômica, próximo, diária, etc. Somente a capacidade de associar dados com conceitos, isto é, transformar dados em informação e conhecimento, pode permitir que respostas adequadas para a consulta sejam obtidas.

A Web Semântica

Embora a web atual apresente muitas deficiências, ela exerce um papel muito importante para a sociedade como um todo. Várias iniciativas, como as desenvolvidas pelo Consórcio W3C (*World Wide Web Consortium*), buscam, por intermédio da criação de padrões, arquitetura de metadados, serviços de inferências e ontologias, uma melhor forma de tornar as informações processáveis pelas máquinas.

Neste caso, metadados são descrições sobre os dados. Por exemplo, para um arquivo que armazena uma fotografia, o autor, o formato e a data em que foi tirada podem ser informações relevantes para sua identificação.

Inferência, em computação, é um processo de identificar informações implícitas associadas a dados armazenados. Por exemplo, não é necessário armazenar, em um banco de dados, informações sobre a relação avô e neto. Esta relação pode ser inferida a partir da relação entre pai e filho, caso esta relação esteja representada em um banco de dados, pois o avô pode ser considerado como o pai do pai de um indivíduo.

Uma ontologia, de uma forma bem simplificada, pode ser considerada como a especificação das coisas que existem e as relações entre estas coisas.

O uso de metadados proporciona a descrição abrangente de recursos na web, e é uma solução adequada para promover a recuperação de serviços com mais eficiência.

Há vários projetos direcionados para produção de arquiteturas de metadados com o objetivo de oferecer sustentação à codificação, ao transporte e ao intercâmbio de uma variedade de metadados desenvolvidos de forma independente. Porém, o intercâmbio destes metadados deve ser efetuado levando-se em consideração três aspectos de interoperabilidade: semântico, sintático e estrutural.

A interoperabilidade semântica possibilita compreender o significado de cada elemento do recurso descrito, juntamente com as associações nele encontradas. Para assegurar este tipo de interoperabilidade, o uso de vocabulários específicos, ontologias e padrões de metadados é essencial. A interoperabilidade sintática determina como os metadados devem ser codificados para a transferência de informações. A interoperabilidade estrutural especifica como os recursos estão organizados, juntamente com os tipos envolvidos e os possíveis valores para cada tipo.

Recentemente, Tim Berners-Lee (cientista britânico, atualmente diretor do W3C) vislumbrou o desenvolvimento da Web Semântica, que, a partir do uso intensivo de metadados, visa oferecer acesso automatizado às informações através de máquinas, com base no processamento semântico dos dados e de heurísticas. Segundo Berners-Lee explica, a Web Semântica visa construir uma rede que, além de ligar documentos entre si, seja capaz de reconhecer o conteúdo desses documentos e, através de um processamento computadorizado, inferir novos conhecimentos.



Com o aperfeiçoamento da internet, procura-se fazer com que os computadores ligados à rede cumpram tarefas mais complexas na busca de informações. Um dos desafios é desenvolver mecanismos que permitam que computadores compreendam e associem dados disponíveis na rede, isto é, os computadores, através do uso de ontologias e regras de inferências, poderão encontrar respostas mais precisas para as consultas, descartando o que é irrelevante para o usuário.

Considerando o exemplo discutido no início deste texto, o objetivo central que impulsiona o desenvolvimento da Web Semântica é que, no futuro, a web ofereça ferramentas computacionais de busca automática que permitam que uma pessoa submeta uma consulta como: “Quais são os horários de voos de Goiânia para São Paulo, com ida em 14/9/2010 e volta em 18/9/2010, para um passageiro, com tarifa econômica, e quais hotéis próximos ao Masp, com diária de solteiro entre 100 e 200 reais, têm vaga para entrada em 14/9/2010 e saída em 18/9/2010?”, e receba como resposta as companhias aéreas e seus respectivos voos no período solicitado e uma relação de hotéis que satisfazem os requisitos da consulta.

Referências

BERNERS-LEE, T. H.; J. LASSILA, O. The Semantic Web. *Scientific American*, May 2001.

WORLD WIDE WEB CONSORTIUM. [Http://www.w3.org](http://www.w3.org).

SOBRE A ESTRUTURA FORMAL DAS PALAVRAS EM WILHELM VON HUMBOLDT

Sebastião Elias Milani¹

Introdução

A Morfologia tal qual é conhecida na atualidade, como a ciência que tem como objeto de estudo as formas mínimas do significado, foi desenvolvida no século XIX. August Schleicher na introdução da obra “As línguas da Europa moderna” (1897 [1857]) escreveu que “cada língua pode ser decomposta em dois elementos: noções de um lado e relações do outro” (p. 7); “que a essência de uma língua se baseia sobre a maneira como ela expressa acusticamente, quer dizer por uma palavra”, as significações e as relações (p. 7). Essas noções explicitam a formação básica das palavras nas línguas flexionais, ou nas línguas indo-europeias.

Essa visão terminava por conjugar séculos de estudo sobre a palavra, desde os gregos clássicos, fala-se aqui dos *Diálogos* de Platão, que tinham na palavra o objeto de estudo. Em Platão as noções de *convenção* e *sugestão* para as palavras estão muito bem explicadas e, durante os séculos que se sucederam, teorizou-se muito sobre a construção dos sentidos através das palavras. Isso de tal maneira que, no início do século XIX, a relação entre o pensamento e as palavras estava suficientemente compreendida para que se pudesse procurar na manifestação da língua a estrutura da representação do pensamento. De fato, é a descoberta do sânscrito que impulsionaria o desenvolvimento das pesquisas estruturais, porque expusera com clareza a condição de historicidade das línguas.

¹. Professor da Faculdade de Letras da UFG. sebas@letras.ufg.br