

Blogs e gerenciamento da imagem das organizações: análise de ferramentas de busca de blogs¹

Prof. Dra. Cíntia Carvalho²

Prof. Dra. Helaine Rosa³

Prof. Ms. Rodrigo Goulart⁴

Prof Dra. Sandra Portella Montardo⁵

Resumo: Cada vez mais, as organizações percebem o poder dos blogs sobre sua imagem, que podem se configurar tanto como uma oportunidade de relacionamento com seus públicos como uma grande ameaça para o seu negócio. Logo, monitorar os blogs passa a ser tarefa imprescindível para favorecer a percepção e análise de risco para a imagem e reputação das organizações. No entanto, as ferramentas de busca por blogs disponíveis oferecem restrições quando utilizadas para este fim (Pesquisa Google de Blogs e Search do Technorati). Este artigo pretende identificar possibilidades e restrições dessas ferramentas. Em seguida, destaca-se a pertinência da proposição de novos métodos para o processamento automático ou semi-automático de blogs, com vista na análise semântica dos mesmos e seu relacionamento com a imagem das organizações, por meio de métodos e técnicas de Processamento da Linguagem Natural (PLN).

Palavras-chave: Blogs; organizações; ferramentas de busca; Processamento de Linguagem Natural; websemântica

Introdução

Os blogs se fazem presente na vida das empresas, estejam elas a par ou não do que se diz sobre elas neste canal. Na medida em que os blogs se popularizam no país, aumentam os riscos de que seus autores se utilizem dessa ferramenta para beneficiar ou comprometer a imagem e a reputação de uma corporação. Assim, percebe-se a importância de se usar essa nova forma de comunicação a favor das organizações no sentido de monitorar essas ferramentas e ampliar a possibilidade de diálogo com seus públicos.

Ainda que haja uma produção significativa de livros a respeito de diversas modalidades sobre comunicação digital em web marketing, poucas publicações se ativeram aos blogs especificamente. Da mesma forma, nos congressos da área da comunicação⁶, os blogs aparecem como objeto de pesquisa, principalmente, quando o tema tratado é redes sociais e suas implicações do jornalismo, não tendo havido até então um foco no tema proposto.

Além disso, até o momento, não se sabe da existência de softwares ou sistemas específicos para a análise automática de blogs que permitam uma avaliação qualitativa e quantitativa da relação de blogs com uma determinada organização, entidade ou pessoa.

¹ Trabalho a ser apresentado no NP Tecnologias da Informação e da Comunicação no XXVIII Encontro da Sociedade Brasileira de Estudos Interdisciplinares da Comunicação.

² Doutora pela PUCRS – Linha de Pesquisa Comunicação e Poder nas Organizações. Professora dos Cursos de Relações Públicas na ULBRA/RS e na Feevale e pesquisadora do Grupo de Pesquisa nesta instituição. E-mail: cicarp@terra.com.br

³ Doutora pela PUCRS – Linha de Pesquisa Comunicação e Poder nas Organizações. Professora do Curso de Relações Públicas da Feevale e pesquisadora do Grupo de Pesquisa Comunicação e Cultura desta mesma instituição. E-mail: helainerosa@feevale.br

⁴ Mestre em Ciência da Computação pela PUCRS – Linha de Pesquisa Inteligência Artificial. Professor do curso de Informática da Feevale e pesquisador do Grupo de Pesquisa em Tecnologias da Informação na mesma instituição. E-mail: rodrigo@feevale.br

⁵ Doutora em Comunicação pela PUCRS – Linha de Pesquisa Comunicação e Tecnologias do Imaginário. Professora do Curso de Publicidade e Propaganda da Feevale e pesquisadora do Grupo de Pesquisa Comunicação e Cultura na mesma instituição. E-mail: sandramontardo@feevale.br

⁶ Compós e Intercom, em nível nacional.

1. Imagem e risco

Quando uma crise atinge pessoas ou organizações despreparadas, ainda que vencedoras, põe imediatamente em risco a credibilidade corporativa. Não existe uma hora marcada para a instalação da crise, porém, há possibilidades de preveni-la. A sobrevivência de uma organização depende da habilidade que esta tiver para processá-la (ROSA, 2001).

A premissa básica destas condições encontra-se diretamente associada aos processos de comunicação instituídos entre uma organização e seus *stakeholders* - grupos de interesse que se relacionam com a organização afetando-a ou sendo por ela afetados -, sendo estes identificados como reforçadores das relações firmadas.

A comunicação pensada de forma estratégica⁷, no sentido de “*esboçar e implementar programas de mediação entre interesses sociais, políticos e econômicos, capazes de influenciar o crescimento e a sobrevivência da organização*”⁸ procura estabelecer, segundo Kunsch (1999, p.74), “*processos interativos e as mediações da organização com seus diferentes públicos, a opinião pública e a sociedade em geral*”, por seu papel decisivo neste cenário. Esse processo é vital para uma organização porque ninguém pode gerar sozinho informações para administrar, prevenir ou evitar uma crise.

Quando se estabelecem relações divergentes, as partes envolvidas tendem a obstacularizar e a contribuir para que fracasse toda a forma de vinculação. Emergem sentimentos de suspeita e desprezo caracterizando motivações agressivamente competitivas e despertando a desconfiança. Consolida-se, a partir destas sensações, uma relação movida pela falta de legitimidade outorgada ao outro.

Neste aspecto, a *web* – especificamente os blogs – apresenta-se como uma ferramenta que pode auxiliar na melhoria do desempenho organizacional, na implementação de estratégias de Comunicação e na administração do relacionamento da corporação com seus complexos segmentos de públicos. Pode, também, constituir-se como uma alternativa para promover a reaproximação de partes em disputa, reativando as negociações interrompidas ou aconselhando no sentido de propiciar o surgimento de soluções para o impasse. O que irá diferenciar uma organização da outra e os impactos causados por estas situações, são as resoluções políticas a partir das negociações.

Quando há uma exposição negativa de uma empresa causada por boatos, denúncias, informações distorcidas, intrigas ou rumores, a credibilidade organizacional é abalada e, conseqüentemente, a imagem desta organização também é atingida. Para uma grande parte de indivíduos, imagem implica em algo irreal, ilusório, passageiro. Ela pode ser considerada como uma representação simplificada que emerge na mente das pessoas como síntese de uma ou várias sensações e percepções. Segundo Vaz (1995), em termos mercadológicos, a imagem é um quadro de referências a que o consumidor/cliente recorre para avaliar se determinada idéia é merecedora ou não do seu interesse, da sua simpatia, do seu apoio; confrontando-a com outras idéias não necessariamente específicas do objeto, mas pertinentes à sua apreciação, que são referenciais próprios de cada pessoa, resultantes de nossas experiências individuais de vida.

⁷ Significa pensar numa ação de comunicação excelente e eficaz, a partir de uma análise ambiental e de uma auditoria social regida pela flexibilidade, percepção e por uma avaliação mensurável dos resultados que devem beneficiar não só a organização, mas também seus públicos (KUNSCH, 1999).

⁸ Citação extraída do artigo “Relações com a mídia”, de Heloiza Matos. In: KUNSCH, Margarida. Obtendo Resultados com Relações Públicas. São Paulo: Pioneira, 1999, p. 199.

Rosa (2002) salienta que a formação de uma imagem organizacional desenvolvida a partir de mentiras e falsidades, torna-a passível de destruição. A imagem, nestas circunstâncias, até pode proporcionar uma certa vantagem momentânea, mas isso não deve perdurar, a não ser que os fatos reais sejam modificados para concordar com uma condição favorável desta.

O referido autor apresenta cinco pilares de sustentação da imagem de uma organização. O primeiro condiciona a verdadeira percepção pública de uma imagem não se atendo apenas aos fatos, significando dizer que a imagem criada das coisas ou pessoas não são baseadas apenas nos fatos, sendo condicionadas, também, pelos valores inseridos neste processo.

O segundo pilar da imagem consiste que os fatos podem assumir novos significados, dependendo dos valores vigentes. Entende-se, neste momento, o motivo pelo qual a imagem organizacional ou pessoal se modifica ao longo do tempo.

A construção da imagem real, faz com que o ser humano aceite incontestáveis “verdades” transmitidas pelo conhecimento. Caracterizando-se este como o terceiro pilar da imagem: o ser humano precisa viver uma realidade para se convencer da imagem a ela associada. Prosseguindo, o mesmo autor (2003) indica que “*O conhecimento, o saber, o estudo, a tradição, a cultura permitem que o homem construa uma imagem do mundo, aceitando como verdadeiras premissas que ele não pode checar pessoalmente*” (ROSA, 2003, p. 85).

Imagem também se traduz na cultura, a qual produz premissas e lógicas próprias que condicionam a forma como as imagens públicas são percebidas individualmente. Os valores do mundo como um todo, não são absolutos. Neste sentido, tem-se o quarto pilar, onde uma imagem depende do meio, da cultura a que ela está exposta para poder saber como será vista.

O quinto pilar está voltado para a idéia de que as imagens são abstrações, desconsiderando-se a idéia de que não são reais e nem falsas. O que importa, neste caso, não é saber se a imagem é verdadeira, mas sim, investigar se possui seus atributos essenciais: se é estável, consistente e coerente, e se tem capacidade de sobrevivência e de organização.

Neste contexto, tem-se que o histórico de uma organização e seu comportamento ético, ao longo de sua existência, são fatores importantes para a formação de uma marca ou para o reconhecimento de uma companhia pela opinião pública, especialmente para os formadores de opinião.

Discutir opinião pública, segundo Poyares (1998), é entendê-la como uma manifestação de juízos em tal volume, intensidade e continuidade que resulta na formação de uma corrente capaz de identificar-se no espaço e no tempo. Ele também indica que inúmeros elementos agem continuamente numa sociedade democrática para a formação das correntes de opinião, como os valores culturais e formação espiritual e profissional; os meios de comunicação social (rádio, tv, jornais, livros e revistas), estatutos jurídicos e decisões de direito; a atuação de líderes e a pressão de grupos de interesse.

Nesta direção, desvendar a dimensão das informações que permeiam os processos comunicativos, o conhecimento das necessidades, expectativas sobre as organizações, do contexto sócio-econômico-político e cultural, do cenário competitivo, das formas mais adequadas de mediação destas relações, são apenas alguns elementos que poderão auxiliar no desenvolvimento de habilidades para solucionar impasses e construir uma imagem organizacional fortalecida.

Adequação aos novos valores, novos estilos de vida, novos modos de comunicação, novas tecnologias – no caso específico deste estudo, os blogs - e novas relações geopolíticas, são variáveis a serem consideradas para a sobrevivência organizacional, assim como o investimento na construção de um sistema comunicacional que propicie condições para que mudanças se efetivem de forma menos traumática possível.

Os blogs, neste caso, possuem uma linguagem mais descontraída, exalam independência e há quem considere como um canal alternativo de notícias e informação. A força dos blogs está em possibilitar que qualquer pessoa, sem nenhum conhecimento técnico, publique suas idéias e opiniões na *web* e que milhões de outras pessoas publiquem comentários sobre o que foi escrito, criando um grande debate aberto a todos.

As conversas, idéias e opiniões que se dão através dos blogs podem facilmente se tornar virais e atingir um enorme número de pessoas em um curto espaço de tempo. São milhões de “blogueiros” opinando sobre notícias, acontecimentos, produtos e serviços, entretenimento, etc. Eles podem estar falando sobre sua empresa, interagindo com seus funcionários, discutindo sobre seus produtos e avaliando seu atendimento, relatando experiências pessoais boas ou ruins, contando suas frustrações ou elogiando sua área de atendimento ao cliente. Estas opiniões certamente atingem a imagem que os leitores formam sobre sua empresa e seus produtos e influencia decisões de compra, no melhor estilo “boca-a-boca” virtual – com força e um alto nível de credibilidade. E este é apenas um aspecto da importância dos blogs e seu impacto no mundo corporativo.

Justamente, para auxiliar na criação de fóruns de discussão dos conflitos e crises organizacionais, na interação e intercâmbio de informações relativas a ameaças para as organizações é que os blogs podem ser inseridos, apresentando-se como uma ferramenta interativa para monitorar e encontrar soluções para impasses (possíveis e/ou já deflagrados). Estes canais favorecem a percepção e análise de riscos⁹, pois se constituem como ambientes onde sinais de advertência sobre fragilidades organizacionais podem ser identificados e favorecem tomadas de decisões pró-ativas, antes que uma crise de fato possa ser deflagrada.

Todas estas vozes podem ser encaradas como um problema ou como oportunidades. Desde que se esteja atento a elas, não as ignore, porque elas, certamente, não irão embora de repente. É uma nova realidade que veio para ficar e, com ela, a necessidade das empresas se adaptarem.

2. Blogs e percepção de risco à imagem das organizações

Web logs ou blogs¹⁰, na sua versão abreviada, são ferramentas de fácil atualização em web sites em HTML¹¹, que se baseiam no sistema de microconteúdos e na atualização quase que diária dos mesmos.

Em abril de 2006, o Technorati¹² rastreava 37,3 milhões de blogs, além de sinalizar que a blogosfera¹³ dobra de tamanho a cada seis meses, conforme Dave Sifry (2006)¹⁴. Ele ainda comenta que, em média, um novo blog é criado a cada segundo por dia, enquanto que 19,4 milhões de blogueiros (55%) continuam

⁹ Risco é a probabilidade de ocorrer um dano à organização.

¹⁰ Web se refere à rede, no sentido de entrelaçamento, navegação na Internet, e log, a registro e conexão. O programador Peter Merholz foi o primeiro a abreviar web log para blog, segundo Thompson (2006, internet).

¹¹ HTML – Hypertext Markup Language, linguagem em que se baseia boa parte da programação de web sites para a Internet.

¹² Disponível em <http://www.technorati.com>.

¹³ Blogosfera se refere a tudo que circula no mundo dos blogs.

¹⁴ State of blogosphere. Part II, April 2006. On Language and Tagging. Disponível em <http://www.technorati.com>

atualizando-os durante dos três meses após a sua criação. Ainda em seu relatório sobre o Estado da Blogosfera, Sifry (2006) aponta que 50.000 postagens¹⁵ são feitas a cada hora. Em maio de 2006, o Technorati já rastreava 40,7 milhões de blogs e 2,4 bilhões de *links*.

De acordo com Pinto (2002), essas páginas da Web podem ser criadas automaticamente por um sistema pré-definido em algum servidor. O mais popular servidor é o Blogger¹⁶, que foi adquirido pelo Google em 2004. O primeiro passo é criar sua conta no sistema e, a partir da formatação escolhida, basta começar a escrever os textos e selecionar imagens desejadas para a publicação (postagens).

São partes constituintes de um blog: as postagens e os comentários. “Cada mensagem enviada é apresentada na página como a entrada de um diário, com a data e a hora em que foi postada. Cada uma dessas entradas chama-se postagens. Em geral, as postagens são apresentados na ordem inversa à que foram enviadas. ou seja, a primeira postagem da página é geralmente a mais recente (isto pode ser mudado pelo dono do blog)” (PINTO, 2002, p. 23). Nestas, é possível disponibilizar texto, fotos e *links* para qualquer outra página.

Logo abaixo da postagem, há um *link* a partir do qual os visitantes do site podem deixar o seu comentário a respeito do que foi postado. O dono do blog também pode comentar sobre os comentários deixados nesse *link*. Os blogs apresentam um sistema de arquivamento das postagens mais antigos, que são facilmente acessíveis, disponibilizados por meio de *links* nos títulos dos postagens. Outro mecanismo é o *trackback*, usado nos blogs para apresentar uma lista de postagens de outros *blogs* que tenham alguma relação com o assunto.

Desde meados do ano passado, o interesse suscitado por blogs no mundo corporativo vem tomando as capas e as páginas principais de publicações especializadas em negócios no Brasil. De acordo com matéria de Terzian (2005), no *web site* da GM¹⁷, há uma área chamada GM Blogs em que Bob Lutz, vice-presidente de desenvolvimento global da companhia, fala sobre os produtos da empresa. No entanto, matéria da revista Exame¹⁸ afirma que todos os comentários colocados no blog de Lutz são submetidos a uma aprovação prévia.

A mesma matéria traz que Randy Baseler, vice-presidente de marketing da Boeing, mantém um blog no site da empresa, no qual tece críticas à concorrente Airbus. Porém, não há espaço para comentários. Outro caso, ainda de acordo com esta matéria, é o da Microsoft. Robert Scoble, dono do blog Scobleizer, foi contratado por Bill Gates e tem contribuído para a mudança da imagem da empresa. Segundo ele, trata-se de um blog franco que inclui várias críticas a Microsoft.

Já James Finn, vice-presidente de comunicação corporativa da IBM, classifica o blog corporativo como uma ferramenta de negócios de alto potencial. Para Carlos Eduardo Franco, consultor associado da consultoria TerraForum, os blogs corporativos têm dois objetivos: aproximar a empresa e o cliente, funcionando como uma ferramenta de marketing e de relacionamento, e criar uma rede de conhecimentos entre funcionários. No primeiro caso, trata-se de um blog externo, permitindo que os funcionários falem de sua rotina de trabalho, ainda que não possa revelar dados sigilosos. O segundo caso mostra referência a um tipo de blog restrito ao

¹⁵ Chama-se postagens todas as inserções de texto e imagens, contendo ou não *links*, feitas pelos autores dos blogs.

¹⁶ Disponível em <http://www.blogger.com> ou <http://www.blogspot.com>

¹⁷ Disponível em <http://www.gmblogs.com>.

¹⁸ Os blogs vão mudar os seus negócios in: **Revista Exame**. São Paulo, Abril, 1º. de fevereiro de 2006, p. 21.

ambiente corporativo, já que se destina a disponibilizar informações sobre projetos, cronogramas, relatórios, etc.

Em setembro de 2005, o Technorati contabilizou cerca de 5 mil blogs corporativos. Entre as empresas que mantêm blogs, aparecem muitas empresas de informática, como a Microsoft, IBM, Sun, SAP, Macromedia e Oracle.

Há outras formas de exploração comercial dos blogs que se aproximam aos moldes da mídia tradicional: o anúncio em blogs de muita audiência. Em matéria divulgada no jornal O Sul, o jornalista Ricardo Noblat, que mantém um blog sobre política, afirma que nos Estados Unidos deve haver umas 20 pessoas que vivem apenas da veiculação de anúncios em seus blogs. Em 2005, foram comercializados 100 milhões de dólares nesse tipo de anúncio no mundo todo. O serviço AdSense do Google permite que os donos de blogs ganhem alguns centavos por clique em anúncios nos seus blogs. Ainda assim, os donos do segundo blog mais popular do mundo, o Boing Boing¹⁹, mantém atividades paralelas.

A matéria ainda indica que, em outubro do ano passado, a Time Warner pagou 25 milhões de dólares pela Weblogs Inc., um conjunto de blogs. No entanto, o objetivo dos blogueiros, em um primeiro momento, não é de ganhar dinheiro, nem mesmo os dos blogs da mídia tradicional. A força do blog está em ser mais um canal de comunicação entre as empresas e seus públicos, como foi mostrado anteriormente.

E é exatamente por isso que os blogs merecem atenção: nunca os consumidores puderam se expressar com tamanha liberdade e alcance, o que pode representar oportunidades e ameaças para as empresas, pois sua reclamação ou elogio pode ser facilmente localizado nos mecanismos de busca específicos para blogs por consumidores interessados em determinadas marcas.

Em julho de 2005, o Relações Públicas Steve Rubel²⁰ revelou a importância de as empresas prestarem atenção ao que está sendo dito nos blogs a respeito delas. Seu trabalho é planejar estratégias para minimizar os prejuízos dos clientes diante de comentários negativos encontrados nos blogs.

Segundo ele, os blogs são publicações espontâneas e imediatas que podem beneficiar ou comprometer uma companhia. Além disso, o executivo ressalta que 51% dos jornalistas norte-americanos lêem blogs. As gigantes IBM, GM e Microsoft contam com equipes e ferramentas para monitorar o que se está falando sobre elas nos blogs.

Juntamente com os blogs, as redes de relacionamento e listas de discussão indicam que as instituições perderam o privilégio da informação, como aponta Ronald Mincheff, presidente da filial brasileira da Empresa de Relações Públicas Edelman²¹. O executivo ainda aponta que as opiniões das pessoas comuns ganham cada vez mais credibilidade do que as veiculadas por um meio de comunicação.

Barbosa Lima²², indica que, em junho de 2005, ocorreu nos Estados Unidos a *Weblog Business*

¹⁹ Boing Boing: a Directory of Wonderful Things é mantido por cinco jornalistas que trabalharam na Revista Wired. É um blog sobre tecnologia e está disponível em <http://www.boingboing.com>

²⁰ Na época desta entrevista, concedida à Revista Exame (20 de julho de 2005, página 116), Steve Rubel trabalhava no escritório de Relações Públicas CooperKatz. Em fevereiro deste ano, Rubel é contratado como Vice-Presidente Senior da Edelman, que atende empresas como a Unilever, AmBev e produtos como o Xbox, da Microsoft.

²¹ Os blogs vão mudar os seus negócios in: **Revista Exame**. São Paulo, Abril, 1º. de fevereiro de 2006, p. 21.

²² Disponível em <http://www.tecnologia.terra.com.br>. Acesso em 16 de abril de 2006.

Strategies Conferences and Expo, enfocando os blogs apenas como potenciais ferramentas de marketing e promoção. Entre suas idéias para o que denomina *blogmarketing* estão:

- a) **Blogs como testemunhais de produtos ou serviços:** empresas enviariam produtos ou serviços para autores de blogs populares que comentariam sua experiência no blog;
- b) **Blogs como objeto de conversação:** na medida em que gerem *links* sobre comentários positivos sobre uma marca.
- c) **Blogs como parte de uma grande ação de marketing específica:** blogs patrocinados fariam sobre o dia-a-dia de um consumidor em fase de experimentação de um novo produto ou serviço. O autor destaca que os textos não podem parecer publicitários, mas espontâneos, como os de um blog autêntico.
- d) **Relato de experiência do consumidor pós-compra:** o exemplo dado é o caso de uma construtora propor que os novos moradores de um prédio ou de um condomínio relatem seu dia-a-dia no endereço novo.
- e) **Blogs fechados para pesquisa qualitativa:** para que as empresas captem experiências diárias de consumidores que estejam testando produtos e serviços.

No Brasil, a empresa de calçados Kildare se valeu do blog para testar a aceitação do novo conceito da campanha publicitária em meados do ano passado²³. Segundo Fernando Alano, diretor comercial da Kildare, futuramente o blog²⁴ poderá ser utilizado para verificar a recepção de produtos novos por parte de quem utiliza o calçado.

Uma pesquisa veiculada na mesma matéria indica que 5% dos blogueiros cria um blog com o interesse de gerar negócios. Este estudo foi realizado pela empresa de Relações Públicas Edelman, tendo sido aplicada em 53 países, inclusive no Brasil. Ainda sobre essa pesquisa:

O levantamento da Edelman indicou ainda que 74% dos blogueiros comentam sobre empresas e produtos pelo menos uma vez por semana em sua página. Por causa disso, existem serviços especializados – até no Brasil – em monitorar o que dizem de seus clientes na rede. Daí o papel do blog corporativo também como resposta aos diários de consumidores. (NUNES, 2006, p. 24)

De acordo com o Ibope Net/Ratings, 37% dos internautas lêem blogs regularmente e passam, em média, 20% do seu tempo de navegação visitando blogs ou comunidades virtuais²⁵, um dado que só vem a enfatizar a importância de se monitorar o que está sendo dito a fim de que se obtenha condições de fazer dos blogs ferramentas de gerenciamento da imagem e da reputação das organizações. O primeiro passo para isso, é investigar de que forma as ferramentas de busca por blogs operam e em que medida podem atender a essa finalidade.

3. Análise das ferramentas de busca de blogs

Nesta parte do trabalho, serão descritas as ferramentas de busca de blog do Google e do site Technorati. Em seguida, serão efetuados testes com buscas por determinados nomes de organizações, a fim de que se possa analisar sua pertinência em virtude da necessidade de localização de ameaças à imagem e à reputação das mesmas em blogs.

Vale ressaltar que esta pesquisa não pretende ser exaustiva quanto aos conteúdos localizados em seus resultados. Nesse sentido, pretende-se selecionar situações gerais que sejam suficientes para demonstrar seu

²³ Diários virtuais no mundo dos negócios in: **Zero Hora** do dia 19 de fevereiro de 2006, página 24.

²⁴ Disponível em <http://www.vocenasceuhomem.com.br>

²⁵ Os blogs vão mudar os seus negócios in: **Revista Exame**. São Paulo, Abril, 1º. de fevereiro de 2006, p. 21.

alcance e apontar limitações destas ferramentas e sugerir um método de busca mais apropriado aos fins pretendidos.

3.1. Pesquisa Google de Blogs

De acordo com o site da Google²⁶, na Pesquisa Google de Blogs²⁷, os resultados incluem todos os blogs que publique um *site feed* (RSS ou ATOM), não apenas aqueles publicados pelo Blogger. O índice de blogs do buscador do Google é atualizado continuamente. Esta pesquisa faz a indexação pelos *site feeds*, que são verificados com frequência quanto ao novo conteúdo. Isso significa que os resultados da pesquisa de um determinado blog serão atualizados com conteúdo novo mais rápido que as pesquisas padrão da web. Além disso, devido aos dados estruturados dentro dos *site feeds*, é possível encontrar postagens e períodos específicos com muito mais precisão.

Na opção pesquisa avançada, pode-se refinar a busca por blogs das seguintes formas: 1) por postagem, em que é possível efetuar a busca por palavras no corpo do texto das postagens e nos títulos; 2) em blogs determinados, seja por palavras no título ou pelos seus endereços; 3) pelo autor, através da busca de blogs e postagens de sua autoria; 4) pela data de postagem, por meio de duas opções: qualquer hora²⁸ ou períodos compreendidos entre datas com dia, mês e ano; 5) idioma, oferecendo 35 opções de idiomas e 6) possibilidade de usar ou não o filtro *Safe Search*, para restringir blogs impróprios para menores.

A pesquisa de blogs faz a indexação pelos *site feeds*, que são verificados com frequência quanto ao novo conteúdo. Isso significa que os resultados da pesquisa de um determinado blog são atualizados com conteúdo novo mais rápido que as pesquisas padrão da web. Além disso, devido aos dados estruturados dentro dos *site feeds*, é possível encontrar postagens e períodos específicos com muito mais precisão.

A empresa escolhida para que se fizesse o teste da busca é a Brasil Telecom. A Brasil Telecom integra operações de telefonia fixa e móvel, dados e voz, longa distância nacional e internacional, data center, Internet grátis, banda larga e acesso discado para uso residencial e empresarial. De acordo com o site da empresa²⁹, em julho de 1999 a mesma passou a oferecer o código CSP 14 para ligações interurbanas e, em janeiro de 2004, o CSP 14 passou a operar para todo o país e em ligações internacionais. Além disso, provê conexão à internet em alta velocidade com o Turbo, que usa a tecnologia ADSL para acesso rápido à rede sem ocupação da linha telefônica, e oferece conteúdo 100% banda larga pelo provedor BrTurbo e BrTurbo Empresas. Quanto à mobilidade para acessar a rede sem fio, há o BrTurbo ASAS. Para acesso discado à internet, a Brasil Telecom oferece o iBest, o segundo maior provedor gratuito do mercado brasileiro. O site de relacionamentos Orkut traz 24 comunidades do tipo “Eu odeio a Brasil Telecom”, enquanto que para sua concorrente GVT existem apenas duas comunidades desse tipo, motivando pela busca de opinião sobre a empresa.

Ao buscar por “Brasil Telecom” na pesquisa geral por blogs do Google, em todas as línguas, aparecem

²⁶ Disponível em <http://www.google.com.br>.

²⁷ Disponível em <http://www.blogsearch.google.com>, <http://www.search.blogger.com>, no painel do Blogger, disponível em <http://www.blogger.com> e em Navbar, disponível em qualquer blog do Blog* Spot em <http://www.help.blogger.com>. A pesquisa realizada em todos os links é a mesma. A única diferença é que na Navbar há a opção para se pesquisar no blog que estiver visualizado.

²⁸ As opções para qualquer hora são: última hora, últimas 12 horas, último dia, semana passada e mês passado.

²⁹ Disponível em <http://www.brasiltelecom.com.br>.

2853 resultados³⁰. Destas postagens, 1753 foram feitas em português. Os resultados são mostrados com o link para a postagem pesquisada, bem como com *links* para o blog que contém a postagem. No blog No place like home.com³¹, uma postagem traz:

Canção do ódio

Eu odeio a Brasil Telecom.

Eu odeio a Brasil Telecom.

Eu odeio a Brasil Telecom.

Faz três dias que não consigo acessar minha conta de mail principal. E ninguém sabe me dizer qual o problema.

Eu odeio a Brasil Telecom.

Eu odeio a Brasil Telecom.

Eu odeio a Brasil Telecom.

Por outro lado, o blog Garota sem fio³², faz referência à marca de uma maneira positiva, indicando o pioneirismo da Brasil Telecom em associar serviço com celular, wi-fi e VoIP, conforme a citação que segue:

Brasil Telecom lançará serviço com celular + wi-fi + VoIP

A **Brasil Telecom** não tem medo do VoIP, ao contrário: quer aproveitar a onda para faturar algum. E, visionária, é a primeira operadora brasileira a anunciar um serviço híbrido de telefonia celular e VoIP. A idéia é excelente para os usuários de telefonia em geral, principalmente móvel. O serviço significará mais economia no fim do mês. No segundo semestre deste ano, a **Brasil Telecom** comercializará um aparelho híbrido, o GF200, da empresa parceira UTStarcom. O GF200 se conecta à rede wi-fi e faz chamadas usando a tecnologia VoIP. Onde não houver sinal wi-fi, ele funcionará como um celular normal. O preço do aparelho ainda não está definido. A UTStarcom produz o modelo na China e pretende ter uma segunda opção de terminal até o final do ano, de acordo com a companhia. A **Brasil Telecom**, por sua vez, acredita que será a primeira operadora do mundo a lançar comercialmente a opção em GSM. [...].

Em postagem publicada no Blog do Mello³³, há uma menção a uma dimensão que pode ferir a reputação da marca Brasil Telecom ao relacioná-la com o valerioduto, esquema de desvio de verbas públicas deflagrado no Governo Lula no ano passado, que desencadeou o escândalo político sem precedentes:

Serraglio versus Daniel Dantas

A **Brasil Telecom** será citada como um dos possíveis caminhos para alimentação do valerioduto - afirmou o relator da CPI, deputado Osmar Serraglio (PMDB-PR), segundo o Jornal do Brasil. A **Brasil Telecom**, à época, era controlada pelo banco Opportunity, do banqueiro Daniel Dantas. Serraglio será enfim o homem capaz de enfrentar aquele que faz tremer a República e o empresariado? Ou Serraglio terá o mesmo destino da juíza Márcia Cunha, da 2ª Vara Empresarial da Justiça do Rio, que afastou o Opportunity da direção da **Brasil Telecom** e que, segundo suas palavras, teve sua vida transformada num inferno, a partir daquele instante? Recentemente, a juíza declarou-se sob suspeição para julgar processos do banco e de companhias controladas por ele. Segundo Márcia Cunha, ela tomou essa decisão "por não ter força para enfrentar o poder econômico" do Opportunity. A partir da sentença, boatos teriam sido espalhados contra a juíza, ela disse que teria sofrido intimidação por parte de estranhos, que seu gabinete teria sido invadido e que ela e o filho sofreram ameaças.

Essa é uma boa *Opportunity* para testar o tamanho do poder do senhor Daniel Dantas.

Esta postagem revela que se torna necessário mencionar outras palavras associadas ao momento por que

³⁰ Pesquisa feita dia 19 de maio de 2006.

³¹ Disponível em <http://www.homepontocom.blogspot.com>.

³² Disponível em <http://www.odontopalm.com.br/gsf>.

³³ Disponível em <http://www.blogdomello.blogspot.com>.

passa a empresa para localizar o alcance total que ela tem na blogosfera, citando-se eventos os quais ela pode ser associada.

Já na pesquisa avançada do Google Blogs, especificando as opções *Encontrar postagens com todas as palavras* “Brasil Telecom” e *com a expressão* “eu odeio a Brasil Telecom”, trouxe dois resultados: o do blog já citado anteriormente (No place like home) e a do blog Cerveja Blog³⁴, que trouxe a expressão no título da postagem:

EuODEIOaBrasilTelecom

Habilitei meu ADSL em casa em 29/03/2006. Às 23:03h. E concordei em pagar uma absurda taxa de habilitação de R\$ 64,95. E concordei com um plano de fidelidade que me prendia à **BT** por um ano, pra merecer míseros R\$ 10,00 de desconto ao mês, comprometendo-me a pagar mensalmente o ainda ultrajante valor de R\$ 54,90 por 256kbps de velocidade, visto que na concorrência paga-se até R\$ 39,90 por 2Mbps.

Mas tudo bem. Lá na minha rua não chegou a concorrência, nem Virtua nem GVT. Passei cinco (5) dias conectado. Acessando a internet, olhando meu banco, baixando músicas, feliz da vida.

Ontem (04/04/06) liguei o micro depois que cheguei em casa. E não tinha internet! Liguei para o suporte, para ver o que havia acontecido... uma retardada chamada Adriana só soube ler na tela dela que a minha conta havia sido CANCELADA, sem se oferecer para ajudar em nada o meu problema! Pedi o telefone da ouvidoria, que ela prontamente me forneceu e desligou o telefone. Pra eu ligar pra lá e descobrir que só funciona até às 18h. Resumindo a papagaiada: ALGUÉM lá dentro, provavelmente por psicografia, recebeu um pedido de cancelar a minha conta, e cancelou. Depois de muito berrar nos ouvidos de pelo menos 3 atendentes, conseguimos no mínimo que não nos cobrassem a habilitação e muito menos a multa rescisória (por ter "cancelado" antes de completar 1 ano). Mas aguardem, gafanhotos. Essa conversa já chegou à Anatel.

Por acaso, a postagem inserida imediatamente antes desta, falava do mesmo assunto sem, portanto, mencionar a expressão “eu odeio” antes de Brasil Telecom, ainda que o sentido da expressão fosse o mesmo. Falou-se também nos concorrentes (GVT e Netfone) no título da postagem: “Netfone e GVT, corram! Quem chegar primeiro à minha rua leva!”. Além disso, os comentários fazem referência ao tema, sem mencionar a marca Brasil Telecom, propriamente dita, mas, apenas de maneira indireta, como mostra o seguinte comentário:

Já tive problemas com ADSL e a Anatel resolveu, tiveram que me devolver quase tudo o que eu gastei. Agora eu não tenho mais ADSL tenho acesso discado, ADSL é muito caro. Mas continuo tendo problemas, eles não me mandam a conta telefônica para pagar e já me cortaram o telefone duas vezes. Vão cortar a terceira provavelmente pois hoje já é dia 05 e eu teria que ter pago dia 04/05/06, porcaria ... parece que eles odeiam a gente. Edna

Paullus, nem me fale em provedor. Os caras vieram aqui e arrumaram minha conexão. Em compensação, desconfiguraram toda minha rede local (como se ela tivesse alguma coisa haver com a internet) e desativaram meu firewall. Aliás, a culpa maior era do meu firewall. Eu só não sei como é que ele virou o vilão depois de 2 anos funcionando ok. Nem o pessoal de lá sabe, pois também não souberam me explicar. Take.

Frente a esta situação, fez-se a busca avançada nesta mesma ferramenta por Brasil Telecom, com a expressão "eu odeio a Brasil Telecom" e contendo qualquer uma das palavras GVT ou Netfone. No entanto, o resultado para a busca foi de que não havia resultados para esta busca, apesar de sabermos que havia.

Quanto a isso, é válido ressaltar que os comentários fazem referência à postagem (NetFone e GVT, corram! Quem chegar primeiro leva!) e citado abaixo do detectado pela ferramenta de busca que solicitou a

³⁴ Disponível em <http://www.cervejablog.com.br>.

expressão “eu odeio a Brasil Telecom”. Certamente, esta postagem foi localizada e estava entre os 1753 resultados obtidos na busca pela ferramenta do Google, na primeira busca, já que trazia a marca Brasil Telecom. No entanto, o título desta postagem não mencionava a marca, mas destacava a concorrência, o que mostra que a busca deveria se dar de modo a contemplar a concorrência também, o que traria mais e mais resultados dos que os que forma listados.

Cabe ainda enfatizar que os comentários desta postagem mencionavam apenas produtos associados a Brasil Telecom, como ADSL, e outros termos, como Anatel, alertando para que a busca por Brasil Telecom deve contemplar também esses produtos e outros termos associados. Ou seja, para a pesquisa de blogs que trazem ameaças à imagem e à reputação das empresas, é necessário ter acesso a esse tipo de texto que refere uma marca sem mencioná-la de uma maneira mais precisa.

3.2. Technorati

Além da ferramenta de busca por blogs, o site Technorati disponibiliza tags sobre temas específicos comentados em blogs, um diretório de blogs (Blog Finder) e o blog do próprio site, com pesquisas e análises sobre a Blogosfera. A *homepage* traz ainda uma lista dos dez termos mais buscados na última hora em blogs, notícias, livros e filmes mais populares do dia, assim como os 100 blogs mais populares.

A ferramenta de busca de blogs do Technorati proporciona uma busca por blogs com alguns recursos diferentes da Pesquisa Google de blogs, pois permite selecionar o grau de autoridade sobre o assunto postados/comentados nos blogs pesquisados. Da mesma forma que a ferramenta anteriormente analisada, o sistema interpreta como autoridade o número de *links* que apontam para o blog. No entanto, pode-se selecionar a busca por blogs com: 1) qualquer autoridade, em que são mostrados todos os resultados; 2) um pouco de autoridade, em que são listados blogs com pelo menos um link; 3) autoridade, em que são mostrados blogs com alguns *links* e 4) muita autoridade, aqueles que tem centenas de *links* apontando para eles.

A ferramenta de busca do Technorati permite que se pesquise através das opções: 1)postagens contendo este texto (com todas estas palavras; com esta exata frase; com ao menos uma destas palavras; não contendo nenhuma destas palavras); 2) com *links* para esta URL; 3) por postagens categorizadas. Além disso, há a possibilidade de se optar por pesquisa em: a) todos os blogs; b) blogs sobre determinados assuntos; c) em determinados blogs.

Na busca por “Brasil Telecom”, os resultados trouxeram 1.912 postagens no total. Em seguida, apareceu a opção de se pesquisar por idiomas. Optando-se pelo idioma português, a busca por blogs com qualquer autoridade trouxe 1297 resultados; com pouca autoridade, 398 postagens; com alguma autoridade, 109 postagens e, com muita autoridade, foram localizadas, primeiramente, 3 postagens. É válido ressaltar que na repetição da busca de blogs, surgiram resultados diferentes em todas as opções (às vezes, passando a números menores).

A forma como são apresentados os resultados sinalizam, ainda, há quanto tempo foi inserida a postagem. Da mesma forma que a ferramenta analisada anteriormente, surge como opções de navegação clicar na postagem propriamente dita e no blog, havendo também a indicação do título do blog e de seu autor. A ferramenta ainda disponibilizou um gráfico sobre as postagens contendo “Brasil Telecom” nos últimos 90 dias.

Nesta busca, foram encontradas postagens com teor semelhante aos transcritos anteriormente. Entre eles, foram listadas algumas postagens que citavam a empresa apenas como uma referência sem valor para os propósitos pretendidos nesta pesquisa. Seguem dois exemplos de blogs diferentes nesse sentido.

E enquanto me contento com uns dias assim sem sal nem açúcar, espero o telefone tocar. Aquela espera bem cinematográfica, ridícula até... celular a tira-colo até no banho... ainda que saiba que a cobertura nacional **Brasil Telecom** Celular não funciona no trecho São Paulo - Mato Grosso do Sul...³⁵

fui no strike com a gaby, o ygor e o henrique... sem coments... so q o strike estava ultra lotado!!! confraternização **brasil telecom!!!** (ah proposito a reforma ficou muito boa)...³⁶

No blog *Questão de Opinião*³⁷, havia um comentário bem específico sobre a Brasil Telecom, sobre comparação da relação custo-benefício do serviço de Internet banda larga desta empresa com uma concorrente.

Nas cidades atendidas pela Net, já é possível navegar na internet com um link de dois megas por menos de 50 reais. Uma lentidão perto da velocidade desfrutada em alguns países, com mais de 20 mbps, mas um raio perto de uma linha discada ou de uma plano básico de ADSL. Na **Brasil Telecom**, um link de 800mbps custa mais de R\$ 120,00. Menos da metade da velocidade, por mais que o dobro do preço...

Já a pesquisa por Tags contendo “Brasil Telecom”, trazia dois blogs como resultado³⁸. No entanto, nenhuma das postagens localizadas traziam em seu texto o nome da empresa procurada.

Com base em todas essas observações, destaca-se o fato de que os mecanismos de busca em blogs analisados podem ser considerados sites de busca que agregam alguns mecanismos especiais para pesquisa exclusiva blogs, por eles relacionados. Em muitos casos, blogs não indexados pelos buscadores de blogs são encontrados nos buscadores convencionais, comprovando que a blogosfera pode ser muito maior do que a sugeria pelos especialistas. Outro fato é de nenhum dos buscadores estende ou refina suas respostas com base em informações lingüísticas ou semânticas (relacionadas a um determinado assunto) com o intuito de desambiguar termos ou classificar blogs, o que poderia melhorar os resultados obtidos.

Nesse contexto, torna-se pertinente a utilização de técnicas automáticas ou semi-automáticas para o tratamento das consultas e dos resultados em buscas por blogs relacionados a determinados temas, produtos e organizações. A área da computação que trata desse tema é a do Processamento da Linguagem Natural (PLN).

4. Processamento de Linguagem Natural na busca por blogs

A gramática da língua portuguesa, analisada sob o ponto de vista normativo (Gramática Normativa), divide o seu estudo em fonologia, morfologia e sintaxe. Por meio do estudo de textos anotados manualmente³⁹, as regras de gramática presentes nos textos são identificadas com o objetivo de gerar regras formais (passíveis

³⁵ Blog *Avassaladoras*. Disponível em: <http://spaces.msn.com/>

³⁶ Blog " <http://spaces.msn.com/brunogok/Blog/>

³⁷ Blog *Questão de opinião*. Disponível em <http://www.nataldosanjos.com.br>.

³⁸ <http://blogeditor.ibest.com.br/wpmbhack> e <http://people.telecorp.com.br/lft>.

³⁹ Símbolos determinados por especialistas humanos que representam informações lingüísticas e que podem ser interpretados por computadores.

de serem implementadas em computadores) que descrevam o raciocínio humano sobre a língua culta de uma determinada prática social.

A área da computação que trata das questões relacionadas ao estudo da língua, a Linguística Computacional (LC), tem como objetivo estudar e propor soluções computacionais para os fenômenos da língua na modalidade escrita e falada. Neste contexto, a LC divide-se em subáreas, com a finalidade de investigar diferentes aspectos e aplicações do estudo da língua em computação, como por exemplo os trabalhos de Bick (2000) com análise morfosintática automática, Pardo (2003) com a sumarização e geração automática de textos e Uryupina (2003) que investigam o estudo dos fenômenos de correferência e suas aplicações. Cada subárea tem como objetivo propor soluções para questões específicas do processamento da linguagem natural, sejam elas teóricas ou na forma de recursos computacionais, e que geralmente envolvem temas complexos do estudo da língua.

Além das questões linguísticas de um texto, outro tema relacionado é o da representação e, posteriormente, a utilização computacional do conhecimento adquirido. Nesse contexto se insere os conceitos de ontologia e websemântica.

O termo "Ontologia" vem sendo utilizado desde o início da década de 90 na área de Inteligência Artificial para representação computacional de conhecimento em áreas como Engenharia de Conhecimento e Processamento de Linguagem Natural. Segundo Gruber (1993), para sistemas de Inteligência Artificial, o que "existe" é o que pode ser representado.

O termo ontologia é concebido freqüentemente como um conjunto de conceitos, definições e relacionamentos entre eles (Gruber, 1993; Uschold and Gruninger 96, Loh et al, 2002). Os conceitos e as relações de uma ontologia podem ser definidos em diferentes níveis de formalidade, variando desde definições altamente informais, em linguagem natural, até definições rigorosas em linguagens formais, como a lógica de primeira ordem. Isso depende do grau de "especificação" que se necessita.

Dentre as várias linguagens de programação existentes, destaca-se a OWL – *Ontology Web Language*. A OWL⁴⁰ é uma linguagem para especificação de ontologias e o desenvolvimento de aplicações que processam informações semânticas. Ela é baseada em XML o que simplifica a sua disponibilização (Web) e manipulação. A utilização de ontologias e de mecanismos inteligentes com foco no significado das informações disponíveis na Web define a Websemântica. Na Internet, uma série de ferramentas para a edição de ontologias estão disponíveis para download. Delas destaca-se o Protégé. Protégé⁴¹ é um ambiente de plataforma independente e extensível para criação e edição de ontologias e bases de conhecimento. Permite a construção de ontologias de domínio, formulários de entrada de dados customizados e a entrada de dados (Boff, 2005).

A proposta deste trabalho é investigar o desenvolvimento de mecanismos inteligentes para descoberta de conhecimento em blogs que estejam relacionados à imagem de uma determinada organização. Por meio de técnicas de PLN, ontologias e um Robô Virtual, o conhecimento sobre uma determinada organização será reunido, organizado e utilizado de forma automática e semi-automática na busca de informações dispersas nos blogs. O conhecimento de especialistas será utilizado na construção da ontologia que, por sua vez, será utilizada

⁴⁰ A linguagem OWL é recomendada pelo consórcio W3C (<http://www.w3c.org>)

⁴¹ Protégé. Disponível em <http://protege.stanford.edu/>. Acesso em 21 de maio de 2006.

pelo Robô. Técnicas de PLN serão empregadas na construção da base de conhecimento e o mecanismo de inferência do Robô, para que então as utilize em buscas na internet.

Conclusão

Ao verificar-se o funcionamento das ferramentas de busca por blogs, percebemos que o monitoramento preciso destes com a finalidade de se descobrir o que seus autores pensam, escrevem e lêem sobre determinadas organizações exige métodos diferentes dos do mecanismos analisados. Em um aspecto geral, identificou-se, tanto na Pesquisa Google de Blogs quanto no site Technorati, que os sistemas levam em conta a autoridade dos blogs, ou seja, o número de *links* que apontam para ele, assim como sites de busca em geral. No entanto, é válido ressaltar que a imagem das empresas corre risco independentemente da popularidade do blog nesse sentido. Como mostra o Relações Públicas Steve Rubel, citado anteriormente, qualquer blog pode desencadear uma crise na imagem das organizações, de um dia para outro, dos menos aos mais populares. A dimensão viral que pode atingir uma mensagem independe desse tipo de autoridade. Com isso, percebe-se que essas ferramentas oferecem limitações à identificação de ameaças para a reputação das organizações, tornando-se pertinente que se proponha o desenvolvimento de um software baseado em técnicas de PLN, ontologias e um Robô Virtual que possibilite uma busca por blogs mais adequada aos fins descritos neste estudo.

Referências

A cada segundo, surge um novo blog. **Jornal O Sul**. Porto Alegre, p. 15, 6 de março de 2006.

BOFF, Rogério Boff; VIEIRA, Renata e GOULART, Rodrigo. Ensinando história da música com auxílio de ontologias. Anais do X Simpósio de Informática e V Mostra de Software Acadêmico da PUCRS - Uruguaiana (SIMS 2005). 2005

BICK, E. The Parsing System "Palavras". Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework. Aarhus University Press, 2000.

Brasil Telecom. Disponível em <http://www.brasiltelecom.com.br>. Acesso em 21 de maio de 2006.

Diários virtuais no mundo dos negócios in: **Zero Hora** do dia 19 de fevereiro de 2006, página 24.

GOOGLE. Disponível em <http://www.blogsearch.google.com> . Acesso em 19 de maio de 2006

GRUBER, T. (1993) Towards Principles for the Design of Ontologies used for Knowledge Sharing. Formal Ontology in Conceptual Analysis and Knowledge Representation, Kluwer Academic Publishers.

GUIMARÃES, Carlos. Os blogs vão mudar os seus negócios in: **Revista Exame** . São Paulo, Abril, 1º. de fevereiro de 2006, p. 21

KUNSCH, Margarida. Gestão integrada da comunicação organizacional e os desafios da sociedade contemporânea. **Comunicação e Sociedade** . São Bernardo do Campo: Programa de Pós-Graduação em Comunicação Social/Universidade Metodista de São Paulo/UMESP, nº 32, 2º semestre de 1999.

LIMA, Alessandro Barbosa. Blogmarketing: o que é e como usar. Disponível em: <http://www.tecnologia.terra.com.br>. Acesso em 16 de abril de 2006.

LOH, S., CASTRO, F., et al. (2002) Investigação sobre a Construção de Ontologias a Partir de Textos com Suporte de Ferramentas Automatizadas. Simpósio Brasileiro de Informática na Educação (SBIE 2002), Workshop de Ontologias.

NUNES, Vanessa. Diários virtuais no mundo dos negócios in: **Jornal Zero Hora**. Porto Alegre, p. 24, 19 de fevereiro de 2006.

- PARDO, T.A.S.; RINO, L.H.M.; NUNES, M.G.V. GistSumm: A summarization tool based on a new extractive method. Proceedings of 6th Workshop on Computational Processing of the Portuguese Language - Written and Spoken, pp. 210-218 (Lecture Notes in Artificial Intelligence 2721). Springer-Verlag, Germany. (Proc. of the VI PROPOR'2003. Universidade do Algarve-FCHS, Faro, Portugal. 26-27 Junho)
- PIMENTA, Ângela. Faça do blog um aliado. Entrevista com Steve Rubel in: **Revista Exame**. São Paulo, Abril, p. 116, 20 de julho de 2005.
- PINTO, Marcos José. **Blogs! Seja um editor na era digital**. São Paulo: Érica, 2002.
- POYARES, Walter. **Imagem Pública – Glória para uns, ruína para outros**. São Paulo: Globo, 1998.
- PROTÉGÉ. Disponível em <http://protege.stanford.edu/>. Acesso em 23 de maio de 2006.
- ROSA, Mario. **A síndrome de Aquiles**. São Paulo: Gente, 2001.
- _____. **A era do escândalo**. São Paulo: Gente, 2003.
- SIFRY, Dave. State of blogosphere. Part II, April 2006. On Language and Tagging. Disponível em <http://www.technorati.com>
- Technorati. Disponível em <http://www.technorati.com>. Acesso em 21 de maio de 2006.
- TERZIAN, Françoise. Blog de terno e gravata. In: **Coleção 2005 Info Exame: Podcasts, Blogs e RSS**. São Paulo, Abril, setembro de 2005.
- URYUPINA, O. High-precision identification of discourse new and unique noun phrases. Proceedings of the ACL Student Workshop, Sapporo, 2003.
- USCHOLD, M. and GRUNINGER, M. (1996) Ontologies: principles, methods, and applications. Knowledge Engineering Review, 11(2), pp. 93-155.
- VAZ, Gil Nuno. **Marketing Institucional**. São Paulo: Pioneira, 1995.
- Blogs pesquisados:**
- Blog No place like home. Disponível em <http://www.homepontocom.blogspot.com>. Acesso em 19 de maio de 2006.
- Blog Garota sem fio. Disponível em <http://www.odontopalm.com.br/gsf>. Acesso em 19 de maio de 2006.
- Blog do Mello. Disponível em <http://www.blogdomello.blogspot.com>. Acesso em 19 de maio de 2006.
- Blog Avassaladoras. Disponível em: <http://spaces.msn.com/>. Acesso em 21 de maio de 2006.
- Blog ". <http://spaces.msn.com/brunogok/Blog/>. Acesso em 21 de maio de 2006.
- Blog Questão de opinião. Disponível em <http://www.nataldosanjos.com.br>. Acesso em 21 de maio de 2006.
- Blog <http://blogeditor.ibest.com.br/wpmbhack>. Acesso em 21 de maio de 2006.
- Blog LFT Clipping. <http://people.telecorp.com.br/lft>. Acesso em 21 de maio de 2006.
- Blog Cerveja Blog. Disponível em <http://www.cervejablog.com.br>. Acesso em 19 de maio de 2006.
- Blog da Kildare. Disponível em <http://www.vocenasceuhomem.com.br>. Acesso em 11 de maio de 2006.
- Boing Boing. Disponível em <http://www.boingboing.com>. Acesso em 11 de maio de 2006.
- GM Blogs. Disponível em <http://www.gmblogs.com>. Acesso em 11 de maio de 2006.