

Jornalismo e audiência na Web: o uso de ferramentas computacionais para capturar leitor¹

Luana Teixeira de Souza Cruz*

Resumo

A medição de audiência da notícia em ambiente digital sofre consequências da mudança no uso da Web, pois os produtores de conteúdo começam a lidar com ferramentas não necessariamente jornalísticas para capturar o leitor em ambientes hipermediáticos. Esses instrumentos, diferentes dos critérios de noticiabilidade, estão ajudando os jornalistas a “ganhar” mais leitura em matérias online. A ascensão dos instrumentos computacionais da Web Semântica – como Google Analytics e Google Trends –, além da influência de ferramentas de busca na Web, faz parte desse contexto. Nessa teia de significações, as tags funcionam como nós “atratores”, conforme neste estudo se mostrar no exemplo do Caso “Incêndio em Santa Maria”. A Web Semântica está fornecendo estruturas e dando significado ao conteúdo das páginas da Web, criando um ambiente no qual agentes de softwares e leitores podem trabalhar de forma cooperativa. Essas tecnologias estão ajudando a entender o que se passa na cabeça do leitor e atraí-lo na rota hipertextual, pois trouxeram todo o potencial que tem origem, nada mais, nada menos, no pensamento humano.

Palavras-chave: *Jornalismo. Web Semântica. Tags. Google. Hipertexto.*

¹ Texto baseado na dissertação de Mestrado da autora. (Cf. CRUZ, 2014)

* Mestre em Estudos de Linguagens pelo Centro Federal de Educação Tecnológica de Minas Gerais (Cefet-MG).
E-mail: luanatsc@gmail.com.

Introdução

A mudança de uso da Web, desde o surgimento em 1992, modificou o cenário comunicacional e a forma de produção da informação. Os pilares da narrativa e a medição de audiência da notícia sofrem consequências dessas mudanças, pois o repórter começa a lidar com ferramentas não necessariamente jornalísticas para capturar o leitor em ambientes hipermediáticos. Neste artigo, são descritas algumas dessas ferramentas, diferentes dos critérios de noticiabilidade², que estão ajudando os jornalistas a “ganhar” mais leitura em suas matérias online.

Nesse contexto, é importante pensar sobre dois temas: a ascensão dos instrumentos computacionais da Web Semântica e a influência ferramentas de busca na Web. Esses dois aspectos vão estar presentes no caso *Incêndio em Santa Maria*, usado para exemplificar a captura de leitores durante uma cobertura jornalística. A discussão que perpassa os três assuntos é o gerenciamento de *tags* no processo de produção e distribuição das notícias, que também será abordado neste estudo.

A audiência na Web está relacionada à necessidade de informação que as pessoas têm. No processo de comunicação, a audiência é o conjunto de respostas dos leitores aos conteúdos. Segundo Ruótolo (1998), essas respostas podem ser internas (como uma mudança de opinião) ou externas (como a compra de um produto). As respostas que os indivíduos dão aos conteúdos da comunicação resultam de contextos sociais, considerando interesses culturais e as formas de conhecimento dos leitores de sites noticiosos.

Os estudos sobre audiência se desenvolveram a partir da década de 1940, baseados em análises de recepção focadas na programação televisiva. De acordo com Ruótolo (1998), não se encontra no escopo dos estudos de audiência aspectos muito importantes da comunicação social que não fazem parte das respostas do receptor. No entanto, o que se propõe neste estudo é justamente descrever os aspectos não relacionados à recepção, avaliando as estratégias do jornalista de Web para capturar leitores e a influência imediata da audiência no processo de produção. A natureza da leitura na Web altera resultados de audiência nas mídias porque permite ao leitor traçar caminhos imprevisíveis.

A exploração mercadológica da audiência pelas mídias é uma ideia já conhecida: mais leitores, mais audiência e mais possibilidades de anunciantes nas páginas online. No entanto, aqui não interessa discutir essa questão, pois o simbólico, a rota do leitor e a experiência do produtor são muito mais importantes.

2 Critérios usados para produção e hierarquização de notícias. Ao escrever uma matéria, os jornalistas pensam em termos valor de um fato – refletindo se aquele caso tem impacto, interesse humano, significância e singularidade, utilidade, repercussão, originalidade, negatividade, raridade, proximidade, oportunidade, dentre outros. Os mesmo critérios são usados para a hierarquização das matérias no processo de definição dos destaques mais importantes (manchetes). (SODRÉ, 2005)

O jornalista vive um desafio diário de equilíbrio entre critérios de noticiabilidade e interesses empresariais para capturar sua audiência. Os novos aspectos desse jogo são as estratégias computacionais – ferramentas de busca, uso de *tags* e otimização de sites, que agora compõem a rotina de trabalho.

Lógica da rede

De acordo com a ComScore (2013)³, o Brasil tem mais de 44 milhões de usuários conectados na internet, que gastam mais de 27 horas por mês nesse ambiente, sendo 36% do tempo médio no uso de redes sociais. Isso aponta para uma mudança de paradigma do produtor de conteúdo, que se vê obrigado a capturar, na rede, esse leitor de caminhos imprevisíveis.

A partir da década de 1990, alguns autores direcionaram seus estudos para a metáfora da rede, conceito que eles acreditam ser a nova metáfora da realidade contemporânea. Para França (2002), por exemplo, rede é a metáfora que surgiu no final do século XX e substituiu o modelo do telégrafo, no qual alguns polos de produção possuíam onipotência sobre os homens que foram reduzidos ao papel de receptores passivos. Nos ambientes em rede, em especial a Web, as diversas interações são capazes de quebrar a lógica da matriz de transmissão. A realidade do jornalismo nesse contexto é o surgimento de um novo sistema que promove distribuição global de textos, imagens e sons.

Pensando em produção jornalística na Web, mais importante que entender a lógica da rede é saber que, diferentemente dos meios tradicionais, a escrita e a leitura no ambiente hipermidiático apresentam características específicas. Trata-se de uma leitura não sequencial e não linear. Estamos falando de hipertexto. Popularmente, os hipertextos são concebidos como uma série de textos conectados por *links* que oferecem ao leitor diferentes caminhos, conforme definiu Ted Nelson (THEODOR, 2013)⁴ no início da década de 1960.

Sobre esta escrita/leitura não linear, Landow (1997) diz que o hipertexto provê um sistema infinitamente recentrável cujo ponto provisional de focalização depende do leitor, que assim se torna um leitor verdadeiramente ativo em mais um sentido do que o comum. Uma das características fundamentais do hipertexto é que ele é composto de corpos de textos ligados que não têm um eixo primário de organização.

Lévy (2003) define o hipertexto sob a terminologia de uma metáfora, à luz de dois paradigmas: o técnico e o funcional. Tecnicamente, hiper-

3 Empresa especializada em medição, análise de audiência, conteúdo e origem de tráfego na Web.

4 Theodor Holm Nelson, ou simplesmente Ted Nelson, é um filósofo e sociólogo estadunidense nascido em 1937. Pioneiro da Tecnologia da Informação, inventou os termos “hipertexto” e “hipermídia”, em 1963, e publicou-os em livro, em 1965. (Cf. THEODOR, 2013)

texto é um conjunto de nós ligados por conexões. Os itens de informação não são ligados linearmente, como uma corda com nós, mas cada um deles, ou a maioria, estende suas conexões em estrela, de modo reticular⁵. Funcionalmente, hipertexto é um tipo de programa para a organização de conhecimentos ou de dados, a aquisição de informações e a comunicação. Assim como a metáfora da rede, a metáfora do hipertexto é usada para pensar a comunicação que rompe como o esquema transmissionista.

No jornalismo, no contexto de comunicação mediada por computador, emergiram termos como “publicação e recepção para incorporar, anotar, comentar, responder, agregar, cortar, compartilhar, além dos termos em inglês *download*, *upload*, *input* e *output*” (MANOVICH, 2008 *apud* MOHERDAUI, 2010, p. 214). É particularmente relevante para a área acadêmica resultados que apontem especificidades das novas mídias e as mudanças nas mídias tradicionais, sob a perspectiva de como o jornalismo se desenvolve e se adapta às realidades hipermediáticas.

O jornalismo tem a obrigação de arquitetar os acontecimentos e expor os fatos à experiência coletiva social. Essa arquitetura está vulnerável aos novos modos de fazer comunicação – influenciados pelas hiper mídias, em especial a Web. Não cabe aqui dizer se os jornais vão alterar radicalmente o que Sodré (2005) considera captura de acontecimentos sob critérios de noticiabilidade definidos pela prática jornalística. Eles, porém, vão incorporar outras estratégias, como já se vê nas redações. O caminho até a realidade atual foi longo, podendo ser dividido em três fases: Jornalismo na Web 1.0; Jornalismo na Web 2.0 ou Web Colaborativa; Jornalismo baseado em *tags*: a fase da Web Semântica;

Jornalismo na Web 1.0

A Web 1.0, originária da década de 1990, era uma espécie de canal de transmissão de informações, com aplicativos fechados e a adoção de um fluxo de comunicação unidirecional. A primeira fase da Web é conhecida também como informativa ou etapa transpositiva para o jornalismo (MIELNICZUK, 2001). Com características do jornal impresso, inicialmente, as versões online invadiam a Web. No Brasil, o primeiro jornal online, lançado em 1995, foi o *Jornal do Brasil* (JB). Logo depois, o jornalismo se associou ao comércio eletrônico e tornou-se comum o termo “portal”⁶, materializando a ideia de uma Web em miniatura, simplificada e acessível. Nessa primeira fase, os portais cumpriam papel

5 Outra referência nesse sentido é o conceito desenvolvido por Deleuze e Guattari (1995), o rizoma (raízes que se elevam para fora da terra), termo inspirado na botânica e adaptado ao ambiente digital. Rizomas são como linhas suplementares, que não se ligam a um ponto, emergem de vários pontos, sem exatidão nem destino. A metáfora do rizoma apresenta alguns princípios de funcionamento que se aplicam, também, à metáfora da rede e a todo o contexto de comunicação hipermediática.

6 Um portal é um site na Web que funciona como centro aglomerador e distribuidor de conteúdo para uma série de outros sites ou subsites dentro e também fora do domínio ou subdomínio da empresa gestora do portal. (Cf. PORTAL..., 2013)

importante de mídias de massa com os chamados “portais horizontais”, de conteúdos variados.

Em 1997, o Yahoo! lançou o modelo de página de entrada – *stick application* – um portal por onde o usuário acessava tudo que sua conta pessoal permitia. É nesse ponto que os portais começaram a fazer uma proposta de conteúdos e serviços mais personalizados, baseados no tripé interesse, interação e envolvimento. Segundo Ferrari (2004), esses modelos são os portais verticais. Em 1998, a Época Online inovou uma de suas edições, ao permitir que o usuário assistisse a um vídeo relacionado à matéria durante a leitura do conteúdo. No mesmo ano, iniciou-se o investimento das TVs na Web surgiram versões online para a ESPN Americana e, em 2000, a Globo lançou o globo.com com proposta de ser o maior portal brasileiro. Ainda segundo Ferrari (2004), em 1999, surgiu a Google, com uma versão inicial do buscador e que viria a se tornar uma das empresas mais influentes do mundo moderno.

A essência dos portais residia não na produção de conteúdo exclusivamente feito para Web, mas na distribuição da informação. Segundo Borges (2010), essa fase foi marcada pela baixa intervenção do leitor no conteúdo da comunicação, baixa capacidade de personalização do conteúdo, predomínio do emissor sobre o controle da informação e de suas relações com o usuário. Muito rapidamente, a Web deixou de ser uma estrada de mão única. Grandes portais como Yahoo!, UOL e ZAZ criaram ferramentas, com as quais o usuário poderia interagir com as páginas do site, mandar recados, opinar, votar em enquetes, teclar em *chats* e criar páginas pessoais. Esses recursos serviram de experimentação para o surgimento de uma etapa posterior, a Web 2.0.

Jornalismo na Web 2.0 ou Web Colaborativa

Essa fase é baseada na colaboração e produção de conteúdo, o marco inicial para a etapa é no surgimento do *e-commerce* como novo canal de vendas e o consequente aumento das ações das empresas pontocom. A Web 2.0 apresenta aspectos que não eram observados na Web 1.0, ou pelo menos não eram muito explorados, como a produção coletiva de conteúdo, maior interatividade⁷ e o leitor como produtor. Utilizado para descrever a segunda geração da *World Wide Web*, o termo “Web 2.0” foi criado por O’Reilly (2005) e é caracterizado pelo autor como a mudança para uma Web como plataforma. Para ele, a regra mais importante da Web 2.0 é aproveitar a participação das pessoas, com anseios de encontrar a inteligência coletiva.

7 Conforme Primo (2003, 2007), a interatividade ocorre quando dois ou mais participantes desenvolvem uma atividade mútua e simultânea, geralmente buscando o mesmo objetivo. A interatividade vai além do argumento comercial dos sites ditos dinâmicos por causa das possibilidades de clicks: ela une pessoas num processo de comunicação mediada por computador.

Segundo Kelly (2009), os quatro pilares da Web 2.0 são o compartilhamento, a cooperação, a colaboração e o coletivismo. O jornalismo sentiu os impactos da popularização da Web 2.0 quando os leitores, considerados até então receptores, passam a ter papel ativo na publicação e edição de conteúdos, por meio do “Jornalismo *Open Source*” ou “Jornalismo Participativo”, que surgiu em 2004. Nesse mesmo ano, o jornalismo passou a contar com o potencial das redes sociais⁸. Primeiro nasceram o Orkut, Myspace, o Facebook e anos depois o Twitter. O alcance da notícia, por meio dessas redes, é algo incalculável, pois o número de conexões estabelecidas entre os leitores nessas ferramentas potencializa muito a distribuição da informação. Sem falar na tecnologia RSS *feed*⁹, que contribui para a personalização e mudança na distribuição das matérias online.

Segundo Borges (2010), a tônica dessa geração pode ser resumida nos seguintes atributos: ruptura do predomínio do polo de emissão, criação de canais de informação e conversação independentes das fontes formais, alto grau de envolvimento e personalização por parte dos usuários e alto grau de articulação coletiva.

A revolução da Web 2.0 na produção jornalística não cessou os recursos para inovação. Pelo contrário, abriu portas para o que O’Reilly (2005) chama de “potencializar a plataforma Web”. Na segunda fase, muito se evoluiu nas ferramentas de busca online otimizada e principalmente no uso da *folksonomy*¹⁰, a produção de conteúdo baseada na ação *tagging*. Ambos os termos ajudam a pensar o jornalismo baseado em nuvens de palavras-chave atualmente espalhadas pela Web.

A metáfora do jornal impresso ainda permanece na interface, conforme alerta Moherdau (2010), pois os produtores seguem a ordem de hierarquização da informação com manchetes, setas que indicam direções para frente e para trás – iguais às páginas do impresso. No entanto, os agenciamentos começam a se alterar. Por causa do processo de *tagging*, a produção e a leitura enfrentam novo paradigma hipertextual. O que o jornalismo começa a fazer, como parte da potencialização de recursos da Web, é organizar a nuvem de dados do ambiente hipermediático para estruturar a distribuição de notícias ancorada nas palavras-chave.

8 Uma rede social é definida como um conjunto de dois elementos: atores (pessoas, instituições ou grupos, os nós da rede) e suas conexões (interações ou laços sociais). Uma rede é uma metáfora para observar os padrões de conexão de um grupo social, partindo das conexões estabelecidas entre os diversos atores. (RECUERO, 2009, p. 23)

9 O RSS é a interface que pode ser usada para criar sites ancorados em base de dados. Permite que alguém não apenas acesse uma página, mas faça uma assinatura, sendo notificado cada vez que houver atualização na página. (O’REILLY, 2005)

10 A *folksonomia* é uma maneira de indexar informações, uma analogia à taxonomia, mas inclui o prefixo *folks*, palavra da língua inglesa que significa pessoas. O ponto forte da *folksonomia* é sua construção com base no linguajar natural da comunidade que a utiliza. A *folksonomia* permite a cada usuário da informação classificar com uma ou mais palavras-chave, conhecidas como *tags*. (FOLKSONOMIA..., 2013)

Jornalismo baseado em *tags*¹¹: a fase da Web Semântica¹²

“A Web Semântica não é uma Web separada, mas uma extensão da atual. Nela a informação é dada com um significado bem definido, permitindo melhor interação entre os computadores e as pessoas” (BERNERS-LEE; LASSILA; HENDLER, 2001 *apud* ALVARENGA; SOUZA, 2004, p. 214). A Web Semântica é a organização e o uso de maneira mais inteligente de todo o conhecimento já disponível nessa plataforma. As informações são compartilhadas na rede digital e uma possível potencialização de recursos pode reduzir o tempo gasto pelo usuário em buscas frustradas.

Somente o uso de tecnologias Web Semântica e a participação ampla de usuários (como na Web 2.0) levará ao resultado evolutivo, com novas oportunidades de negócio para a comunicação mediada por computador. Isso significa dizer que as características da Web 2.0, como ubiquidade, interconectividade, não linearidade, hibridização, interatividade, colaboratividade, hipercomplexidade e imprevisibilidade, levantadas por Cruz *et al.* (2009) estão englobadas e potencializadas. Bernes-Lee, Lassila e Hendler (2001, sem paginação) descrevem a Web Semântica como “uma nova forma de conteúdo que [...] desencadeará uma revolução de novas possibilidades.” O uso do metadados¹³ é a aposta para aprimorar a Web; eles possibilitam a classificação do conteúdo e tornam as buscas mais eficazes. Alvarenga e Souza (2004, p. 133) explicam o termo Web Semântica afirmando que a expressão só se justifica se observarmos as aumentadas possibilidades que o uso de metadados traz:

Embora ‘semântica’ signifique ‘estudo do sentido das palavras’, Guiraud (1975) reconhece três ordens principais de problemas semânticos: 1) a ordem dos problemas psicológicos, que relaciona os estados fisiológicos e psíquicos dos interlocutores nos processos de comunicação de signos; 2) a ordem dos problemas lógicos, que estabelece as relações dos signos com a realidade no processo de significação; 3) a ordem dos problemas linguísticos, que estabelece a natureza e as funções dos vários sistemas de signos. [...] O uso da conotação “semântica” para a Web ampliada está ancorado na segunda definição, e se justifica se observarmos as aumentadas possibilidades de associações dos documentos a seus significados por meio dos metadados descritivos.

11 Expressão criada por Moherdau (2010).

12 A expressão Web Semântica foi cunhada pela primeira vez por John Markoff, jornalista do *The New York Times*. (Cf. MARKOFF, 2006)

13 “‘Meta’ é um prefixo de autorreferência, de forma que 21 ‘metadados’ sejam ‘dados sobre dados’. [...] Em documentos na *Web* têm a função de especificar características dos dados que descrevem a forma com que serão utilizados, exibidos, ou mesmo seu significado em um contexto.” (ALVARENGA; SOUZA, 2004, p. 134)

Esse “tratamento de dados” está diretamente ligado às *tags*. As palavras-chave são a representação da metalinguagem *XML* (*eXtra Markup Language*), uma ampliação da linguagem HTML (utilizada para incluir hipertextos na Web)¹⁴. A nova fase da Web engloba a característica de criação de redes e comunidades, por isso Moherdau (2010) afirma que a noção de *collaborative tagging*, com a qual os leitores podem incluir metadados sob forma de palavras-chave, ajuda na publicação de informações categorizadas, um sinal da organização proposta pela Web Semântica.

Se as *tags* são, literalmente, a chave para a que a Web deixe de ser um mar de documentos para se tornar um mar de dados, é possível aplicar o mesmo raciocínio à composição de notícias. Nesse contexto, Moherdau (2010, p. 222) coloca um grande desafio “pensar o jornalismo a partir de uma estética do banco de dados, baseada em *tags* e remixada em nuvens de comunicação”. Uma tarefa que foi inserida na realidade de jornalista nessa etapa é o gerenciamento de *tags* e a análise de audiência em tempo real. É papel do produtor de conteúdo pensar em termos certos que devem ser escolhidos no texto online para potencializar a captura do leitor.

Ascensão de ferramentas da Web Semântica

A Web Semântica tem exercido papel relevante no estabelecimento de padrões tecnológicos para a recuperação da informação. Bernes Lee comanda o projeto World Wide Web Consortium (W3C)¹⁵, que pretende alcançar elevado nível dos resultados de buscas na rede por meio do desenvolvimento de regras semânticas visando à comunicação entre homem-máquina, de forma que, mesmo sem deter maiores conhecimentos acerca de estratégias de busca, qualquer usuário seja capaz de recuperar a informação de forma mais precisa. Isso porque tal tecnologia pretende permitir que a máquina “compreenda” a necessidade de informação de quem a busca (CUNHA, 2006).

Alguns trabalhos que já são desenvolvidos chegam muito perto da previsão feita pela equipe W3C no início do projeto, porque usam essencialmente padrões tecnológicos e semânticos para a recuperação e a indexação da informação. Todos esses sites jogam com as *tags*, jogo que pode ser controlado por outras ferramentas, que serão descritas a seguir. São instrumentos que o Google disponibiliza para que produtores de conteúdo possam medir o valor de suas informações, a origem de tráfego na Web e, principalmente, saber se as *tags* escolhidas são “competentes”.

14 Para saber mais sobre HTML e XML, cf. XHTML2 Working Group Home Page. Disponível em: <<http://www.w3.org/MarkUp/>>. Acesso em: 1º abr. 2013.

15 Cf. WORLD Wide Web Consortium (W3C). Disponível em: <<http://www.w3.org/Consortium/>>. Acesso em: 1º abr. 2013.

Google Analytics

O Google Analytics¹⁶ é uma ferramenta usada para medir audiência na Web. É necessário cadastrar o site no sistema e definir usuários que podem visualizar as informações disponibilizadas. A medição direta de acessos é feita por meio da contagem de vezes que as páginas dos sites são carregadas pelos leitores. Coloca-se um código em todas as páginas do site e, quando alguém acessa a página, o código é lido pelo programa, que fica “anotando” a quantidade de leitores, o tempo que eles gastam na página, de onde eles “vieram” (origem de tráfego). Há outras ferramentas parecidas que fazem a Web Analytics, como o Omniture¹⁷.

Para o jornalismo online, o Analytics tornou-se essencial na rotina dos repórteres que acompanham em tempo real as matérias mais escolhidas pelos leitores, podendo controlar quais são as notícias preferidas, quais devem permanecer em destaques nas capas dos portais e os que assuntos rendem continuidade. Pela URL (endereço da página) da matéria, a ferramenta do Google faz um ranking de notícias quase em tempo real.

O acompanhamento é essencialmente uma função do repórter e se torna também parâmetro de medição de produtividade para os editores. O monitoramento em tempo real é usado de minuto em minuto por jornalistas na redação, por trazer essa noção instantânea. É possível também fazer levantamentos de audiência mensal, semestral ou de qualquer período determinado, porém esses dados não ajudam tanto na produção diária. Pode-se, também, analisar dados gerais ou divididos por editoriais.

Google Trends

O Google Trends¹⁸ é uma ferramenta que aponta o volume de busca de uma palavra-chave no Google. Com esse sistema, é possível ver padrões de busca ao longo do tempo. Digitando uma *tag* ou expressão, o programa traça um gráfico sobre o comportamento de busca daquela palavra desde 2004 até o dia da solicitação, sendo o eixo horizontal a representação do tempo e o vertical, a frequência que o termo é procurado globalmente. É possível saber em quais localidades a palavra é mais “googlada” ou virou tendência, em quais períodos e as combinações que geralmente os internautas usam para buscar a palavra.

Existem outras dezenas de ferramentas que usam a *tag* e a lógica da busca como principal estratégia de captura. Os exemplos são úteis à publicidade como o Google AdWords¹⁹, e para o jornalismo, como o

16 Cf. GOOGLE Analytics. Disponível em: <<http://www.google.com/analytics/http://www.w3.org/MarkUp/>>. Acesso em: 1º abr. 2013.

17 Cf. OMNITURE. Disponível em: <<http://www.omniture.comhttp://www.w3.org/MarkUp/>>. Acesso em: 1º abr. 2013.

18 Cf. GOOGLE Trends. Disponível em: <<http://www.google.com/trends>>. Acesso em: 30 mar. 2013.

19 Cf. GOOGLE AdWords. Disponível em: <<https://adwords.google.com/>>. Acesso em: 30 mar. 2013.

Google News²⁰. Não é possível cercar neste estudo todas as possibilidades; o mais importante é pensar nos efeitos do uso desses instrumentos.

O próximo passo é entender a engenharia que está por trás disso, que faz um site ser escolhido na rota do leitor quando ele resolve usar uma ferramenta de busca como Google, Bing, Avg, Conduit, Babylon, dentre outros.

Influência das ferramentas de busca na Web: o que é SEO?

De acordo com dados do Instituto Verificador de Circulação (IVC) (2011-2012), entidade sem fins lucrativos que certifica as métricas de desempenho de veículos impressos e digitais, a origem de acesso em sites no Brasil pode ser dividida em três categorias: buscadores, direto e redes sociais. A primeira categoria detém 40% do tráfego da rede, a segunda 54% e a terceira, 6%. No tráfego desses buscadores, a participação do Google é 94%, seguido pelo Bing (2%) e outras ferramentas da mesma categoria (4%). Outro dado importante é que as pessoas se prendem a um grupo específico de palavras-chave na hora das buscas orgânicas²¹. O IVC identificou 100 expressões mais usadas de cada site analisado no estudo e traçou um “top 100” com essas expressões. Como resultado, constatou-se que quase metade do tráfego gerado pelos buscadores veio dessa *tags* específicas.

Isso significa que os buscadores têm grande relevância quando um leitor quer encontrar um site, mesmo que a categoria não domine o *ranking* de origem. Significa, também, que as *tags* são verdadeiras “atratoras” de internautas, portanto, os produtores que souberem usar bem essa inteligência das palavras-chave estarão em vantagem. Os portais de notícia estão fazendo de tudo para serem encontrados na Web. Eles querem estar na rota do leitor, e essa tentativa de aparecer no topo dos sites de busca é resultado de estratégias de Search Engine Optimization (SEO), traduzido para ao português como “otimização de sites”. Os mecanismos de busca usam um algoritmo – embasado em um cálculo matemático – que interpreta a relevância de uma página da Web. As empresas de comunicação não pensam diretamente nesse algoritmo, mas elas querem entender o comportamento do leitor para manipular o algoritmo. Muitas delas, já estão ensinando SEO para os jornalistas nas redações.

Segundo Formaggio (2008), os passos que um site de notícias deve seguir na elaboração de um plano para que seja encontrado em mecanismos

20 Cf. GOOGLE News. Disponível em: <<https://news.google.com/>>. Acesso em: 30 mar. 2013.

21 Uma busca orgânica é o processo que permite aos usuários da World Wide Web realizar pesquisas na rede por meio das listagens dos motores de busca, ao contrário do que ocorre com listagens de resultados onde pode constar publicidade paga e anúncios (*pay per click*) exibidos entre os resultados de procura. Listas de resultados de busca orgânicas são baseadas na relevância de uma palavra-chave em uma página de um Website. (Cf. BÚSCA..., 2015)

de busca são: 1) identificar e elaborar as *tags* importantes para o site, 2) identificar mudanças ou adaptações na página são necessários para que o site seja encontrado, 3) identificar as páginas de entrada desejadas para essas *tags*, 4) saber qual o posicionamento dos concorrentes por aquelas *tags* importantes para seu site. Muitos desses aspectos dizem respeito ao trabalho do desenvolvedor do site, e não ao jornalista, mas o repórter pode, sim, ajudar e aumentar a relevância do site escolhendo títulos assertivos, inserindo *tags* certas em suas matérias ou relacionando conteúdos de forma mais contextualizada possível. O exemplo da cobertura do Incêndio em Santa Maria será a prova dessa função do repórter.

Caso Incêndio em Santa Maria

No dia 27 de janeiro de 2013 o Brasil assistiu a uma das maiores tragédias de todos os tempos. Mais de 240 jovens morreram no incêndio da Boate Kiss, na cidade de Santa Maria, no Rio Grande do Sul. Era um domingo, e todas as atenções da imprensa nacional, assim como dos brasileiros, voltaram-se para o fato. Os jovens, na sua maioria, universitários, participavam de uma festa na casa de shows, quando um integrante da banda – que comanda a noite – usou pirotecnia na apresentação e acabou provocando o incêndio.

A notícia, provavelmente, tornou-se o topo da audiência nos sites de notícias, e um dos fatores é a própria essência jornalística que atribui valor ao fato – o critério de noticialidade (FORMAGGIO, 2008, p. 3). É um caso que englobou os seguintes fatores: 1) interesse social – porque suscitou questões de segurança em casas noturnas que se aplicam a qualquer cidade; 2) imprevisibilidade – surpreendeu a população brasileira; 3) significância e singularidade – foi um fato intenso, com centenas de envolvidos que sofreram consequências, além de ser inédito; 4) atualidade – naquele domingo o que havia de recente nos noticiários eram as informações de cada momento sobre a tragédia; 5) continuidade – as notícias sobre o caso perpetuaram naquele domingo, durante dias, semanas e meses; 6) emoção – o fato foi tratado como imagem da comoção nacional porque sensibilizou a população brasileira; 7) proximidade afetiva e geográfica – familiares e amigos das vítimas se tornaram leitores potenciais pelo interesse em informações; 8) negatividade – no senso comum é tratado como interesse pela tragédia do outro; esse desvio que as pessoas têm para notícias trágicas é psicanaliticamente um sistema emocional de autodefesa que dá uma sensação de alívio ou tensão ao saber que outra pessoa passou pela situação violenta.

Além dos fatores tradicionais do jornalismo, as ferramentas computacionais foram essenciais para os sites noticiosos. As *tags* e as ferramentas de busca mudaram o cenário da cobertura. Um exemplo é o que ocorreu com os jornalistas do pelo site do Estado de Minas (em.com.br), portal focado no noticiário local em Minas Gerais, naquela manhã de domingo. Veja o relato do jornalista Emerson Campos, que trabalhou na cobertura:

A gente já tava com uma audiência que era bem atípica pro dia de domingo porque o fato em si pedia, gerava essa audiência grande então a gente já tinha notado que obviamente a audiência tinha subido muito, só que a partir de certo momento a gente viu que ela se multiplicou quatro, cinco, seis, sete vezes foi aumentando e a gente ficou meio sem entender. Quando a gente entrou nessa notícia em si que tava dando mais audiência do que todas as outras e não estava com destaque tão grande na capa quanto as outras, a gente foi ver o que tinha de diferente nela e a gente percebeu que era a *tag* no final, as palavras-chave de Santa Maria e a gente até deu uma olhada nos outros sites e ninguém tava usando, tava todo mundo puxando como tragédia sul, boate, incêndio, mas não usava Santa Maria. A gente pegou e usou o mesmo artifício para colocar em outros conteúdos que estavam no ar a mesma *tag*. A partir do momento que a gente colocou essa mesma *tag* nos outros conteúdos, os outros conteúdos também começaram a responder e audiência a subir. Não subiu tanto como essa primeira acho que foi uma coisa inicial, veio antes das outras ela teve/gerou mais acesso no Google do que as demais, mas as outras também responderam em algum nível e começaram a gerar acesso pelo Google e não mais pela capa do nosso portal. (CAMPOS, 2013).

O conteúdo citado pelo jornalista e que fez disparar os acessos do site no dia 27 foi a galeria de fotos intitulada “Incêndio em Santa Maria”. Conforme dados do Google Analytics²², naquele domingo, 93,19% do tráfego na página da galeria teve origem em ferramentas de busca orgânica, ficando o restante para o tráfego direto (pessoa que clicou no *link* da galeria em uma matéria ou em outro site que havia usado a galeria), e ao chamado tráfego de referência²³ (pessoas que chegaram até a galeria pelas capas do Portal Uai, em.com.br, Google News, estaminas.com.br ou Facebook). De todo esse tráfego, 98,19% veio do buscador Google. As combinações procuradas nos buscadores que renderam maior audiência à galeria foram “incêndio em santa maria” (18,81%), “fotos do incêndio em santa maria” (12,92%) e “fotos incêndio santa maria” (5,42%).

22 A autora teve acesso aos dados porque era repórter do em.com.br à época do incêndio e da pesquisa.

23 Esses sites do tráfego de referência são definidos pelo proprietário do site e passam a ser monitorados como origens importantes de “porta de entrada” dos leitores.

A diferença entre as combinações de *tag* parecem sutis, podem ser por apenas uma preposição, mas mudam o cenário da audiência conforme relata o repórter:

A galeria que começou a dar muito acesso, eram fotos iniciais, o conteúdo em si não justificava tanto a quantidade de acesso. A gente sabia que não era por causa destaque na capa, porque não tinha uma foto tão boa ainda para chamar acesso pela capa e a gente foi notar que era mais a semântica da coisa mesmo, de ser a palavra-chave que estava buscando acesso e não o conteúdo em si. (CAMPOS, 2013)

Para se ter uma ideia da importância desse exemplo do uso de *tag*, o número absoluto de acessos dessa galeria de fotos a colocou em primeiro lugar como a mais clicada de todos os tempos desde o surgimento do portal em.com.br, em setembro de 2010. A audiência, medida desde o dia da tragédia até dois meses depois do fato (27 de março), é o dobro da galeria, que vem em segundo lugar – fotos de um grave acidente na Avenida Nossa Senhora do Carmo, em Belo Horizonte, que está no ar desde 6 de junho de 2012. Nem mesmo o tempo de exposição online, o fator regional do site, a proximidade e tantos outros critérios de noticiabilidade venceram a competência da otimização mecânica.

O uso de *tags* específica alterou o processo de produção dos repórteres – que passaram a incluir a expressão em todo o conteúdo sobre a tragédia – e capturou leitores automaticamente. Possivelmente, antes da tragédia, a *tag* “Santa Maria” tinha outros valores semânticos quando procurada em buscadores na Web. Agora não é mais possível fazer esse teste, porque a palavra-chave assumiu o significado atrelado ao incêndio, mas é provável que antes os resultados remetessem à santa – mãe de Jesus – ou a alguma cidade que não fosse aquela do Rio Grande do Sul, ou alguma música, avenida, peça teatral, igreja, algum bairro, enfim, uma centena de significados que a expressão pode lembrar. No Google Trends, a *tag* alcançou o topo máximo de buscas em janeiro de 2013 e ganhou uma combinação improvável antes do fato: “tragédia santa maria”. Talvez o grande erro dos outros portais de notícia no processo de *tagging* de seus conteúdos pode ter sido o uso de “boate” ou “Kiss”, que são palavras com uma infinidade semântica bem maior que Santa Maria. O em.com.br teve mais assertividade usando “incêndio em santa maria”.

Conclusão

A tecnologia, aplicada aos processos comunicativos na Web, chega a um ponto hipercomplexo. As ferramentas da Web Semântica nos ajudam a entender o que se passa na cabeça do leitor, a atraí-lo na rota

hipertextual, tentando adiantar a lógica de pensamento dele quando navega pelo ambiente digital. Essas tecnologias trouxeram o potencial, cuja origem, nada mais nada menos, está no pensamento humano. Para Lévy (2004), as ações dos jornalistas e comunicadores farão com que o interpretante entenda as mensagens de várias formas, pois na comunicação a informação se precisa por meio do contexto e do sentido. Segundo o autor, as palavras, sons, sinais e imagens acabam criando redes de significação na mente do leitor. Dar sentido a texto é o mesmo que ligá-lo, conectá-lo a outros textos, portanto, é o mesmo que construir um hipertexto. Nessa teia de significações, as *tags* funcionam como nós “atratores”, conforme vimos no caso *Incêndio em Santa Maria*.

A informação ganhou *status* de insumo básico para a construção do conhecimento e, ainda, para tomada de decisão. A Web Semântica está fornecendo estruturas e dando significado ao conteúdo das páginas da Web, criando um ambiente no qual agentes de *softwares* e leitores podem trabalhar de forma cooperativa. Por fim, vale refletir sobre o caso específico do jornalismo, em que o repórter está sujeito à mistura de sua subjetividade aos instrumentos computacionais. O profissional vai precisar agregar gerenciamento de *tags* na rotina de trabalho.

Journalism and audience on the Web: the use of computer tools to attract readers

Abstract

News audience measurement in the digital environment suffers the consequences from change in Web use as content producers begin to deal with not necessarily journalistic tools to lure the reader into hypermedia environments. These instruments, different from the criteria of newsworthiness, are helping journalists to “earn” more reading in online materials. The rise of computational tools on the Semantic Web – such as Google Analytics and Google Trends – besides the influence of search engines on the Web, is part of that context. In this web of meanings, tags work like “attractors”, as this study shows in the “Fire in Santa Maria” Case. The Semantic Web is providing structures and giving meaning to the content of Web pages, creating an environment in which software agents and readers can work cooperatively. These technologies are helping to understand what goes on in the reader’s mind and draw him into the hypertextual route because they brought all the potential that originates nothing less than in human thought.

Key-words: Journalism, Semantic Web, tags, Google, hypertext.

Referências

ALVARENGA, Lídia; SOUZA, Renato Rocha. A web semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 132-141, 2004. Disponível em: <<http://dx.doi.org/10.1590/S0100-19652004000100016>>. Acesso em: 20 mar. 2013.

BERNERS-LEE, Tim *et al.* The semantic Eeb. *Scientific American*, Canadá, v. 284, n. 5, p. 28-37, 2001.

BORGES, Rosane da Silva. Espaço e design do jornalismo impresso: confluências com o jornalismo digital na Web 3.0. In: CONGRESSO INTERNACIONAL COMUNICACIÓN 3.0, 2., Salamanca, 2010 (Libro nuevos medios, nueva comunicación). Disponível em: <<http://campus.usal.es/~comunicacion3punto0/comunicaciones/089.pdf>>. Acesso em: 7 jun. 2013.

BUSCA orgânica. In: WIKIPEDIA: a enciclopédia livre. 2015. Disponível em <http://pt.wikipedia.org/wiki/Busca_org%C3%A2nica> Acesso em: 30 mar. 2013.

CAMPOS, Emerson. *Cobertura do incêndio em Santa Maria*. Belo Horizonte, 12 ago. 2013. Entrevista concedida a Luana Teixeira de Souza Cruz. 1 arquivo .mp3 (7 min.).

COMSCORE. *Brazil digital future in focus*. São Paulo, 2013. Disponível em: <<http://www.comscore.com/por/Imprensa-e-eventos/Apresentacoes-e-documentos/2013/2013-Brazil-Digital-Future-in-Focus>>. Acesso em: 7 jun. 2013.

CRUZ, Luana *et al.* *Discussão dos modelos teóricos de comunicação em rede e suas aplicações práticas em produtos para hiperídia*: um estudo de caso do iGoogle. 2009. Disponível em: <<http://www.bocc.ubi.pt/pag/gomes-peixoto-cruz-pires-ferreira-discussao-dos-modelos-teoricos.pdf>>. Acesso em: 3 mar. 2013.

CRUZ, Luana Teixeira de Souza. *Rota hipertextual baseada em tags*: discussão de processos de produção e leitura como sistemas complexos no contexto da Web Semântica. 2014. 126 f. Dissertação (Mestrado em Estudos de Linguagens) – Programa de Pós-Graduação *Stricto Sensu* em Estudos de Linguagens (Posling), Centro Federal de Educação Tecnológica de Minas Gerais (Cefet-MG). Belo Horizonte, 2014. Disponível em: <www.academia.edu/.../Rota_hipertextual_baseada_em_tags_Discussão_de_processos_...>. Acesso em: 7 jun. 2013.

CUNHA, Jacqueline de Araújo. *Web semântica: o estado da arte*. 2006. 74 f. Monografia (Graduação em Biblioteconomia) – Universidade Federal do Rio Grande do Norte, Natal, 2006. Disponível em: <https://monografias.ufrn.br/jspui/bitstream/1/86/1/JacquelineAC_Monografia.pdf>. Acesso em: 7 jun. 2013.

DELEUZE, Gilles; GUATTARI, Félix. Introdução: rizoma. In: _____; _____. *Mil platôs: capitalismo e esquizofrenia*. São Paulo: Editora 34, 1995. v. 1, p. 10-36. Disponível em: <www.escolanomade.org/wp-content/downloads/deleuze-guattari-mil-platos-vol1.pdf>. Acesso em: 3 mar. 2013.

FERRARI, Poliana. *Jornalismo digital*. São Paulo: Contexto, 2004.

FOLKSONOMIA. In: WIKIPEDIA: a enciclopédia livre. 2013. Disponível em: <<http://pt.wikipedia.org/wiki/Folksonomia>>. Acesso em: 20 mar. 2013.

FORMAGGIO, Erick Beltrami. *SEO: otimização de sites: aplicando técnicas de otimização de sites com uma abordagem prática*. Rio de Janeiro: Brasport, 2008.

FRANÇA, Vera R. Do telégrafo à rede: o trabalho dos modelos e a apreensão da comunicação. In: PRADO, José L. (Org.). *Crítica das práticas midiáticas: da sociedade de massa às Ciberculturas*. São Paulo: Hacker, 2002. p. 57-76.

GOOGLE AdWords. Disponível em: <<https://adwords.google.com/>>. Acesso em: 30 mar. 2013.

GOOGLE Analytics. Disponível em: <<http://www.google.com/analytics/http://www.w3.org/MarkUp/>>. Acesso em: 1º abr. 2013.

GOOGLE News. Disponível em: <<https://news.google.com/>>. Acesso em: 30 mar. 2013.

GOOGLE Trends. Disponível em: <<http://www.google.com/trends>>. Acesso em: 30 mar. 2013.

GUIRAUD, Pierre. *A semântica*. 2. ed. Rio de Janeiro : Difel, 1975.

INSTITUTO VERIFICADOR DE CIRCULAÇÃO (IVC). *Estudo sobre audiência de websites*: Base IVC: jan. 2011/dez. 2012. Disponível em: <http://www.ivcbrasil.org.br/conteudos/pesquisas_estudos/AudienciaWeb2012.pdf/>. Acesso em: 30 mar. 2013.

KELLY, Kevin. The new socialism: global collectivist society is coming online. *Revista Wired*, 2009. Disponível em: <http://www.wired.com/culture/culturereviews/magazine/17-06/nep_newsocialism?currentPage=all>. Acesso em: 19 mar. 2013.

LANDOW, George P. *Hypertext 3.0: critical theory and new media in an era of globalization*. Baltimore: Parallax, 1997.

LÉVY, Pierre. *As tecnologias da inteligência: o futuro do pensamento na era da informática*. Rio de Janeiro, Editora 34, 2004.

LÉVY, Pierre. Pela ciberdemocracia. In: MORAES, Denis de (Org.). *Por uma outra comunicação: mídia mundialização cultural e poder*. Rio de Janeiro: Record, 2003. p. 367-384.

MANOVICH, Lev. *Software takes command*. 2008. Disponível em <<http://www.maximsurin.info/wp-content/uploads/pdf/software-takes-command.pdf>>. Acesso em: 7 jun. 2013.

MARKOFF, John. Entrepreneurs see a web guided by common sense. *New York Times*, New York, v. 12, 2006. Disponível em: <http://www.nytimes.com/2006/11/12/business/12web.html?ei=&c_r=0>. Acesso em: 7 jun. 2013.

MIELNICZUK, Luciana. Características e implicações do jornalismo na Web. In: CONGRESSO DA ASSOCIAÇÃO PORTUGUESA DE CIÊNCIAS DA COMUNICAÇÃO (SOPCOM), 2., Lisboa. 2001. *Anais...*, Lisboa: Associação Portuguesa de Ciências da Comunicação, 2001. Disponível em: <http://www.facom.ufba.br/jol/pdf/2001_mielniczuk_caracteristicasimplicacoes.pdf>. Acesso em: 7 jun. 2013.

MOHERDAUI, Luciana. Jornalismo baseado em tags. In: RIBEIRO, Ana Elisa et al. (Org.). *Leitura e escrita em movimento*. São Paulo: Peirópolis, 2010. p. 214-228.

MORIN, Edgar. *Introdução a pensamento complexo*. Lisboa: Instituto Piaget, 2003.

OMNITURE. Disponível em: <<http://www.omniture.comhttp://www.w3.org/MarkUp/>>. Acesso em: 1º abr. 2013.

O'REILLY, Tim. *O que é Web 2.0?* 2005. Disponível em: <<http://pressdelete.files.wordpress.com/2006/12/o-que-e-Web-20.pdf>>. Acesso em: 15 mar. 2013.

O'REILLY, Tim. *Web 2.0: compact definition: trying again*. 2006. Disponível em: <<http://radar.oreilly.com/2006/12/Web-20-compact-definition-tryi.html>>. Acesso em: 15 mar. 2013.

PORTAL. In: WIKIPEDIA: a enciclopédia livre. 2013. Disponível em: <[http://pt.wikipedia.org/wiki/Portal_\(Web\)](http://pt.wikipedia.org/wiki/Portal_(Web))>. Acesso em: 20 mar. 2013

PRIMO, Alex. *Interação mediada por computador: comunicação, cibercultura, cognição*. Porto Alegre: Sulina, 2007.

PRIMO, Alex. Quão interativo é o hipertexto? Da interface potencial à escrita coletiva. *Fronteiras: estudos midiáticos*, São Leopoldo, v. 5, n. 2, p. 125-142, 2003.

RECUERO, Raquel. *Redes sociais na Web*. Porto Alegre: Sulina. 2009.

RIBEIRO, Ana Elisa et al. *Folheando de mentira: leituras de jornais impressos na web*. Contemporanea: revista de comunicação e cultura, Salvador, v. 7, n. 1, p. 1-32, 2009.

RUÓTOLO, A. C. Audiência e recepção: perspectivas. *Comunicação & Sociedade*, São Bernardo do Campo, n. 30, p.159-170, 1998.

SODRÉ, Muniz. O que é mesmo uma notícia? In: ENCONTRO ANUAL DA ASSOCIAÇÃO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO (COM-PÓS), 14., 2005, Niterói. *Anais...*, Niterói, RJ: Associação Nacional dos Programas de Pós-Graduação em Comunicação, 2005.

THEODOR Nelson. In: WIKIPEDIA: a enciclopédia livre. 2013. Disponível em: <http://pt.wikipedia.org/wiki/Theodor_Nelson>. Acesso: 7 jun. 2016.

WORLD Wide Web Consortium (W3C). Disponível em: <<http://www.w3.org/Consortium/>>. Acesso em: 1º abr. 2013.

XHTML2 Working Group Home Page. Disponível em: <<http://www.w3.org/MarkUp/>>. Acesso em: 1º abr. 2013.

Enviado em 30 de novembro de 2015.
Aceito em 25 de abril de 2016.