

A BIG DATA E O JORNALISMO: A BASE DE DADOS FAVORECENDO A INFORMAÇÃO

LUIZ RICARDO GOULART HÜTTNER¹; GABRIELA DA SILVA ZAGO²

¹Universidade Federal de Pelotas – luiz.ricardo@ufpel.edu.br

²Universidade Federal de Pelotas – gabrielaz@gmail.com

1. INTRODUÇÃO

O jornalismo sempre sofreu com as transformações tecnológicas e informacionais que ocorreram durante séculos por todo o mundo. Desde o surgimento da prensa móvel de Gutenberg, onde os escritos puderam ser distribuídos em grandes quantidades até a chegada da internet, o jornalismo se viu modificado. Assim, as formas de apuração, produção, circulação e consumo de informações, também mudaram consideradamente.

Na internet, o jornalismo passou por diversas fases. Segundo Mielniczuk (2001) desde o seu surgimento, a internet mudou (quase que por completo) o modo de fazer jornalismo. A primeira apropriação que o jornalismo faz da internet é a transposição literal do jornal para a web. Em segundo, os sites noticiosos começam a usar algum recurso típico da internet, como o uso de imagens. A terceira fase já explora os recursos de vídeos, imagens, hiperlinks, infográficos e outros recursos que são disponibilizados na web 2.0 (O'Really, 2005, online), que diferentemente de sua versão antiga, tem como característica também, o gerenciamento de bancos de dados das companhias Web 2.0, que podemos definir não como softwares, mas sim como “infoware” (O'Really, 2005, online).

Porém, antes mesmo da invenção da Web, a informação era repassada através da internet, mas eram direcionados para públicos específicos.

Antes da invenção do World Wide Web (WWW ou Web), a rede já era utilizada para a divulgação de informações jornalísticas, porém os serviços oferecidos eram direcionados para públicos muito específicos e funcionavam através da distribuição de e-mails, de boletins disponibilizados através do Gopher¹ ou de recursos semelhantes (PALACIOS, MIELNICZUK, BARBOSA, RIBAS, NARITA, 2002. online).

Somente com o início da comercialização da Web, no começo dos anos 1990, a internet começa a atender, de uma forma mais expressiva, as atividades jornalísticas.

Com todo o avanço da tecnologia, muitos pesquisadores e estudiosos questionam qual o futuro do jornal impresso e para onde caminha o jornalismo digital, ou como preferimos chamar neste artigo, de ciberjornalismo. Quanto ao desaparecimento do jornal impresso, não nós preocuparemos em responder essa questão, mas afirmamos, conforme já pesquisado, o jornal impresso mudará seu formato, mudará o jeito que entregará a informação ao seu leitor (TORRES, 2011, online), passando de notícias exclusivas, para notícias mais contextualizadas, pois a notícia de última hora já é postada por veículos de comunicação que estão na internet, graças as funções e a agilidade do ciberjornalismo. Dentre as

¹ O Gopher é um sistema que possibilita o acesso a informações mantidas em diversos computadores da rede. O acesso é feito através de menus e este sistema comporta apenas textos.

principais características do ciberjornalismo, destacamos a hipertextualidade, a atualização contínua, a multimídia, a interatividade, a personalização e a memória (PALACIOS, 2003).

Na internet, o futuro do ciberjornalismo também é discutido. Visto que ele já passou por diversas mudanças e o avanço técnico/tecnológico cresce também a cada dia, a pergunta que se faz é qual o futuro do ciberjornalismo?

Suzana Barbosa propõe um jornalismo de base de dados (JDBD). Segundo a autora, este seria o jornalismo em que

Pressupõe base tecnológica ampliada, acesso expandido por meio de conexões banda larga, proliferação de plataformas móveis, redação descentralizada e adoção de sistemas que permitam a participação do usuário, produtos criados originalmente para veiculação no ciberespaço, conteúdos dinâmicos formatados em narrativas multimídia, experimentação de novos elementos conceituais para organização da informação, assim como de novos gêneros (BARBOSA, 2007, p. 128)

Segundo a linha da autora, vemos crescer um novo momento nesta base de dados na internet. Surge também o termo Jornalismo Guiado por Dados (JBD),

que compreende diversas práticas profissionais, cujo ponto em comum é o uso de dados como principal fonte de informação para a produção de notícias. Barbosa e Torres (2012) consideram o fenômeno como pertencente ao paradigma Jornalismo Digital em Bases de Dados (JDBD), caracterizando-se como uma extensão ou ampliação das práticas de JDBD (TRÄSEL, 2013, online).

Segundo a linha do armazenamento e processamento de dados, surge também o que hoje chamamos de Big Data. Antes, tecnologia da informação era disponível (em princípio) para órgãos governamentais, grandes empresas e de instituições de educação, até o ponto que surge a computação em nuvem, que reduz o custo para o armazenamento de informação, através de aplicativos e softwares, que em princípio, geram poucos (ou nenhum) custo para o usuário. Quando falamos em nuvem, nos referimos aos dados que ficam armazenados em softwares de código aberto, que podem ser acessados remotamente através de dispositivos com acesso à internet.

Por definição

"Big data" refere-se ao conjunto de dados (dataset) cujo tamanho está além da habilidade de ferramentas típicas de banco de capturar, gerenciar e analisar. A definição é intencionalmente subjetiva e incorpora uma definição móvel de como um grande conjunto de dados necessita a fim de ser considerada Big data. Não definimos Big data em termos de ser maior do que certo número de Terabytes (milhares de Gigabytes). Assumimos que, como a tecnologia avança sobre o tempo, o tamanho de datasets que quantificado como Big data também aumentará² (JUNIOR, 2011, online).

Para melhor entendimento deste conceito, alguns especialistas classificam 5 Vs: velocidade, variedade, volume, veracidade e valor. O volume de informação

² Big data: the next frontier for innovation, competition, and productivity. McKinsey Global Institute, p.1, maio de 2011. Disponível em www.mckinsey.com/mgi/.../big_data/pdfs/MGI_big_data_full_report.pdf. Acessado em 20 de maio de 2011.

que é gerado, não somente por empresas jornalísticas, mas também por usuários da internet, através, principalmente, dos sites de rede social. Cada dia centenas de milhares de informações são geradas pelos interagentes nessas redes. Assim, interagentes e grandes empresas navegam em um labirinto cada vez maior de dados e informações.

Por exemplo, segundo dados do site inglês³ há uma média de 2,5 mil milhões de gigabytes de dados criados diariamente, composto por 200 milhões de tweets e 30 bilhões de peças de conteúdo compartilhado no Facebook a cada mês. Ainda segundo a publicação, se seguirmos nesta perspectiva, no ano de 2020, poderemos chegar a um escalonamento de 43000000000000 gigabytes, com seis bilhões de pessoas em posse de telefones celulares inteligentes, com acesso à internet.

2. METODOLOGIA

Por ainda se tratar de um tema relativamente novo e no jornalismo, mais novo ainda, utilizaremos estudos recentes que dissertem sobre esse tema. Iremos analisar como esses dados podem ser usados para a produção de informação. O quanto importante é o jornalista saber lidar com esses dados.

Várias empresas já trabalham com o processamento e análise de dados em todo mundo. Mas empresas jornalísticas já separam em editorias especiais, as matérias jornalísticas que e utilizam de dados para repassar e sustentar uma informação. Observamos que sites noticiosos, como o inglês The Guardian⁴ e o Estadão no Brasil⁵, que já criam o seu próprio banco de dados⁶, onde o leitor pode acessar dados que o jornal teve acesso, através de matérias jornalísticas.

3. RESULTADOS E DISCUSSÃO

As vantagens do Big Data perpassam o limite de uma boa reportagem. Através do direcionamento que uma reportagem pode dar ao seu leitor, este mesmo leitor pode cruzar dados e informações “verdadeiras”, pois estarão analisando a informação diretamente da fonte e não somente dados através de um texto, que sofre a modificação de jornalistas (e sua crença pessoal) e editores (que adaptam o texto para as linhas editorias da empresa).

Para isso, a profissão cultura jornalística deve ser repensada. Pois ao que tudo indica, num futuro próximo, o jornalismo de base de dados irá se consolidar cada vez mais e é tarefa do jornalista adquirir habilidades que antes eram dispensáveis, já que a “Era do Big Data possibilita que os dados digitalizados possam ser cruzados e relacionados por jornalistas com habilidades multidisciplinares ou por pessoas que possuem conhecimento sobre as tecnologias digitais conectadas” (JUNIOR, 2011, online).

4. CONCLUSÕES

Desta forma, caminhamos cada vez mais para um ciberjornalismo dinâmico, em que ele “é continuamente renovado, inter-relacionado, contextualizado, trabalhado segundo distintos formatos multimídia e com parâmetros novos para a sua apresentação de modo a propiciar maior

³ <http://thoughtsoncloud.com/2014/02/cloud-computing-and-big-data-an-ideal-combination/>

⁴ <http://www.theguardian.com/uk>

⁵ <http://blog.estadaodados.com/>

⁶ <http://www.theguardian.com/news/datablog>

envolvimento na maneira como os usuários vão consumir e interagir com ele e a partir dele” (BARBOSA, 2012).

Por fim, vemos o jornalismo guiado por dados pode devolver, aliando os recursos da sociologia e a rapidez do cálculo dos computadores, o retorno da objetividade para as redações e para o jornalismo (TRASEL, 2013).

5. REFERÊNCIAS BIBLIOGRÁFICAS

BARBOSA, S. Sistematizando conceitos e características sobre o jornalismo digital em base de dados. In: BARBOSA, S. *Jornalismo Digital de Terceira Geração*. Labcom: Covilhã, 2007.

BARBOSA, Suzana. Agentes de inovação, renovação e de reconfiguração para o jornalismo em tempos de convergência. In: *Revista de Cibercomunicación*. 2012. Disponível em: <http://revistadecibercomunicacion.com/index.php/cibercomunicacion/article/download/5/16> Acessado em: 30 jul. 2014.

JUNIOR, W. T. L. Big Data, Jornalismo Computacional e Data Journalism: estrutura, pensamento e prática profissional na Web de dados. In: **SBPJOR**, 2011. Disponível em: <http://www.ec.ubi.pt/ec/12/pdf/EC12-2012Dez-11.pdf> Acessado em: 20 jul. 2014.

MIELNICZUK, L. Características e implicações do jornalismo na Web. Online, 2001. Disponível em: http://www.facom.ufba.br/jol/pdf/2001_mielniczuk_caracteristicasimplicacoes.pdf Acessado em: 30 jul. 2014.

O'Reilly, T. *What Is Web 2.0*. Online, 2005. Disponível em: <http://oreilly.com/web2/archive/what-is-web-20.html> Acessado em: 30 jul. 2014.

PALACIOS, M; MIELNICZUK, L; BARBOSA, S; RIBAS, B; NARITA, S. Um mapeamento de características e tendências no jornalismo online brasileiro. In: Universidade Federal da Bahia, Salvador, 2002. Repositório UFBA. Disponível em: http://www.facom.ufba.br/jol/pdf/2002_palacios_mapeamentojol.pdf Acessado em: 30 jul. 2014.

PALACIOS, M. Jornalismo Online, informação e memória: apontamentos para debate. In: FIDALGO, A; SERRA, P. *Jornalismo Online*. Universidade da Beira Interior: Covilhã, Portugal, 2003.

TORRES, C. O que as mídias sociais podem ensinar às mídias tradicionais? Online. Disponível em: <http://blogmidia8.com/2011/06/o-que-as-midias-sociais-podem-ensinar.html> Acessado em: 30 jul. 2014.

TRASEL, M. Jornalismo Guiado por Dados: relações da cultura hacker com a cultura jornalística. In: **XXII ENCONTRO ANUAL DA COMPÓS**, UFBA, Bahia, 2013. Disponível em: http://compos.org.br/data/biblioteca_2065.pdf Acessado em: 30 jul. 2014.