

# galáxia

Galáxia

E-ISSN: 1982-2553

aidarprado@gmail.com

Pontifícia Universidade Católica de São  
Paulo  
Brasil

Oliveira Barbosa, Suzana; Torres, Vitor  
O paradigma 'Jornalismo Digital em Base de Dados': modos de narrar, formatos e  
visualização para conteúdos  
Galáxia, núm. 25, junio, 2013, pp. 152-164  
Pontifícia Universidade Católica de São Paulo  
São Paulo, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=399641251013>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

re<sup>o</sup>alyc.org

Sistema de Informação Científica  
Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal  
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# O paradigma 'Jornalismo Digital em Base de Dados': modos de narrar, formatos e visualização para conteúdos<sup>1</sup>

Suzana Oliveira Barbosa

Vitor Torres

**Resumo:** Neste artigo, retomamos formulações anteriores acerca do Paradigma Jornalismo Digital em Base de Dados (JDBD) para refletir sobre sua preponderância crescente, apontando extensões para o modelo no jornalismo contemporâneo. Uma das mais visíveis e sobre a qual nos centramos é o que se vem chamando Jornalismo Guiado por Dados, considerada aqui uma das vertentes do JDBD, pois está no escopo de abrangência do seu conceito. A partir do referencial teórico e do mapeamento de exemplos empíricos, inferimos que as novas extensões do JDBD materializam, principalmente, modos de narrar, de compor as peças informativas, de formatos variados para os conteúdos, bem como para apresentação das informações jornalísticas. Assim, é possível indicar a conformação de hipernarrativas ou narrativas interativas dinâmicas, de cariz complexo, também afinadas com o que se denomina estética base de dados.

**Palavras-chave:** jornalismo; Jornalismo Digital em Base de Dados; cibercultura

**Abstract:** *The Paradigm of Database Journalism: narrative modes, formats and content visualization*

– This article examines earlier formulations of the paradigm of Digital Database Journalism (DDBJ) to reflect about its increasing preponderance, pointing out expansions of the model in contemporary journalism. One of the most visible expansions upon which we focus is known as Data-Driven Journalism, which we consider one of the approaches of DDBJ, since it lies within the scope of the DDBJ concept. From the theoretical framework and mapping of empirical examples, we infer that the new expansions of DDBJ materialize mainly modes of narration, of composing informative essays, of varied content formats, and of presenting journalistic information. Thus, it is possible to indicate the shaping of dynamic interactive hypertexts or narratives, of a complex aspect, which are also in tune with so-called database aesthetics.

**Keywords:** journalism; digital database journalism; cyberculture.

<sup>1</sup> Este artigo, originalmente apresentado no GT Estudos de Jornalismo do XXI Encontro da Compós (2012), foi revisto e atualizado para envio a esta revista.

## Paradigma JDBD em expansão contínua

Como um modelo teórico formulado para o melhor entendimento sobre o papel das bases de dados no jornalismo contemporâneo, o Paradigma JDBD – Jornalismo Digital em Base de Dados – (BARBOSA, 2007, 2008, 2009, 2011; FIDALGO, 2004, 2007; MACHADO, 2006; RAMOS, 2011a, 2011b<sup>2</sup>) expande-se em sucessivas apropriações, cada vez mais demarcando distinções para os cibermeios operando segundo a lógica multiplataforma. No contexto atual, a consolidação das bases de dados como estruturantes da atividade jornalística e como agentes singulares no processo de convergência de meios nos permite afirmar a sua preponderância, ao tempo que verificamos extensões para o modelo.

Uma das mais visíveis é o que se vem chamando Jornalismo Guiado por Dados ou Data Journalism (BRIGGS, 2010; BRADSHAW, ROHUMAA, 2011; ROGERS, 2011; GRAY, BOUNEGRU, CHAMBERS, 2012). Segundo definição mais simples, Jornalismo Guiado por Dados é aquele produzido com dados, os quais podem ser gerados e disponibilizados por uma diversidade de fontes públicas e privadas – inclusive as próprias organizações jornalísticas do *mainstream* – e podem estar estruturados em sua forma mais bruta, comum, em planilhas Excel, ou mesmo publicados segundo padrões de *design* e formatos diversos para a narrativa jornalística que tiram partido de recursos variados para a melhor apresentação e compreensão do leitor/usuário, do público.

O termo enfatiza, especialmente, a existência de informações públicas de países, governos, entidades de diversos perfis tornadas cada vez mais acessíveis e diversificadas, ao tempo que permitem o uso, a exploração e a publicação por meio de formatos diversos e ferramentas de visualização específicas utilizadas para sua apresentação. Gera o que estudiosos classificam como o fenômeno do “*big data*” ou “a revolução industrial do dado”, conforme indicam Bradshaw e Rohumaa (2011, p. 48). Talvez, um dos exemplos emblemáticos da força dos dados tenha sido o episódio da divulgação dos 250 mil telegramas confidenciais de comunicações diplomáticas de governos, principalmente dos Estados Unidos, por parte da organização WikiLeaks, em novembro de 2010.

Foi a partir desses dados coletados e disponibilizados pelo WikiLeaks – conhecido como *cablegate* – que *The New York Times*, *The Guardian*, *El País*, *Le Monde* e a revista *Der Spiegel* puderam dar sentido às informações contidas nos dados, relacionando-os, fazendo cruzamentos, análises e, com isso, puderam revelar assuntos que provavelmente jamais seriam conhecidos. Esse é também um exemplo referencial por ter levado as organizações jornalísticas a buscarem recursos e parâmetros que melhor pudessem estruturar as narrativas publicadas nos respectivos *websites*, com distintos formatos<sup>3</sup> para os

<sup>2</sup> Aqui, são indicados trabalhos de autores que têm investigado o emprego das bases de dados no jornalismo desde o início da década de 2000, a partir da perspectiva que as concebe como fundamentais para estruturação da atividade jornalística, e não apenas como ferramentas auxiliares.

<sup>3</sup> O termo formato é usado considerando o âmbito da narrativa jornalística, ou seja, a forma, a aparência que assume o conteúdo jornalístico apresentado e que incorpora os diversos recursos propiciados pelo suporte digital para modelizar as narrativas (MIELNICZUK, 2003, p. 72).

conteúdos: de textos a documentos em pdf, videogalerias, infográficos interativos, galerias de foto, mapas gerados a partir de *mashups*<sup>4</sup>, *slideshows*, base de dados navegável com os telegramas cuja interface para acesso se apresenta como mapa interativo, uso de técnicas de visualização para apresentação das informações contidas nos *cables*<sup>5</sup>, entre outros.

Aqui, entendemos o Jornalismo Guiado por Dados<sup>6</sup> como uma das vertentes, um dos aspectos compreendidos pelo Paradigma JDBD, uma vez que está no escopo de abrangência do seu conceito, sintetizado como sendo: o modelo que tem as bases de dados como definidoras da estrutura e da organização, bem como da composição e da apresentação dos conteúdos de natureza jornalística, de acordo com *funcionalidades* e *categorias* específicas, que também vão permitir a criação, a manutenção, a atualização, a disponibilização, a publicação e a circulação de cibermeios dinâmicos em multiplataformas. Sendo assim, o Jornalismo Guiado por Dados é compreendido como uma das extensões para o Paradigma JDBD no jornalismo contemporâneo, uma vez que demarca a ampliação das possibilidades de emprego das bases de dados no processo de produção de conteúdos jornalísticos, no seu consumo e circulação.

Entre as mais de vinte funcionalidades sistematizadas para o Paradigma JDBD (BARBOSA, 2007, 2008, 2009), citamos as mais pertinentes para a discussão que aqui se estabelece: 1) integrar os processos de apuração, composição, documentação e edição dos conteúdos; 2) orientar e apoiar o processo de apuração, coleta, e contextualização dos conteúdos; 3) regular o sistema de categorização e qualificação das distintas fontes jornalísticas, indicando a relevância delas; 4) habilitar o uso de metadados para análise de informações e extração de conhecimento, por meio de técnicas estatísticas ou métodos de visualização e exploração como o *data mining*<sup>7</sup>. Também assegurando a aplicação da técnica do *tagging*<sup>8</sup>; e 5) garantir a flexibilidade combinatória e o relacionamento entre os conteúdos.

Quanto às categorias expressivas do modelo, temos a da *dinamicidade* como aquela que permite às demais adquirir representatividade. É a partir dela que se demarca um novo padrão para os cibermeios, em contraposição aos anteriores *sites* estáticos. São elas:

- **Automatização.** Inerente ao uso de bases de dados nos processos de armazenamento, estruturação, organização e apresentação das informações. Permite agilidade nos processos de apuração, formatação de conteúdos a partir do que

<sup>4</sup> Consiste na operação de combinar conteúdos de duas ou mais fontes para gerar um novo.

<sup>5</sup> Ver em: <<http://migre.me/aTego>>, <<http://migre.me/aTeij>>, <<http://migre.me/aTej5>>, <<http://migre.me/aTek6>>, <<http://migre.me/aTelK>>. Mais sobre o cablegate e seus desdobramentos, em: <<http://migre.me/aTem1>>.

<sup>6</sup> De certa maneira, amplia a noção que se firmou na década de 1980 como Reportagem Assistida por Computador (RAC) e que teve no NICAR, ([www.ire.org/nicar/](http://www.ire.org/nicar/)), o principal irradiador de suas técnicas para dar suporte às reportagens investigativas, muitas delas vencedoras de vários prêmios Pulitzer. No Brasil, a Associação Brasileira de Jornalismo Investigativo ([www.abraji.org.br/](http://www.abraji.org.br/)), entidade criada em 2002, realiza cursos e oficinas sobre RAC.

<sup>7</sup> Mineração ou exploração de dados. Conjunto de operações de diversos tipos que se pode realizar para explorar relações, tendências e projeções em conjuntos de dados de grandes dimensões (COLLE, 2002, p. 255). Ferramentas de *data mining* se baseiam no conhecimento acumulado que está invisível nas bases de dados.

<sup>8</sup> Descritor usado como metadado e agregado ao conteúdo do texto de uma notícia, de uma reportagem, ou a qualquer outro formato, como vídeo, infografia interativa, *slideshow*, mapa, etc. Disponível em: <<http://migre.me/aTemN>>. Acesso em: 11 fev. 2012.

está armazenado no arquivo e também as chamadas estatísticas dinâmicas ou Sistemas de Recomendação de Notícias (SRN), entre outros. Há três tipos de automatização: parcial, procedimental (nível intermediário) e total.

- **Flexibilidade.** Relacionada à facilidade de recuperação dos conteúdos para a contextualização e a própria flexibilidade combinatória entre distintos tipos de conteúdos, o que é fundamental no contexto da produção jornalística em redações integradas e no âmbito do jornalismo móvel.
- **Inter-relacionamento/hiperlinkagem.** Permite incorporar materiais distintos que estejam disponíveis na base de dados do cibermeio, bem como nas de outras fontes com informações abertas para oferecer análise e contextualização. Por outro lado, agrega elementos da hipernarrativa ou da narrativa interativa às peças jornalísticas, assegurando, em paralelo, tematizações novas para os conteúdos.
- **Densidade informativa.** Delimitada, assim como a próxima categoria, a partir da concepção de Resolução Semântica (FIDALGO, 2004, 2007b), considerada imanente ao jornalismo digital (MIELNICZUK et al, 2010). A densidade informativa é assegurada por uma oferta abrangente e diversificada de conteúdos (tanto quanto aos gêneros jornalísticos como aos modos de narrar e aos formatos) e de recursos associados à cobertura jornalística dos acontecimentos, de acordo com o tratamento, o processamento e a publicação deles no âmbito de cada cibermeio.
- **Diversidade temática.** Novas tematizações podem ser trabalhadas para assegurar também maior densidade informativa e vice-versa. Com a incorporação crescente das bases de dados para a estruturação do material jornalístico, para a construção das peças informativas e para a apresentação dos conteúdos, também maiores serão as possibilidades para novas tematizações, ou seja, que ultrapassem o eixo temático mais comum utilizado para organizar os conteúdos, como política, economia, esportes, cultura, além de ciência, saúde, tecnologia. Assim, as novas editorias/seções denominadas “Data” ou “Dados” presentes em alguns cibermeios (*Los Angeles Times*, *The Texas Tribune*, *Guardian*, *The New York Times*, *Gazeta do Povo*) são exemplos atuais ilustrativos de novas tematizações que estão diretamente relacionadas com a força dos dados no jornalismo.
- **Visualização.** Diz respeito aos modos diferenciados para se representar informações jornalísticas, a partir da sua estruturação em base de dados. Aqui são intrínsecas as noções de metadados ou metainformações, as técnicas de *data mining*, do *tagging*, entre outras. As ferramentas de visualização mais empregadas são a *ManyEyes*, desenvolvida pela IBM, a *Your.FlowingData*, para coletar dados através do Twitter, a *Visual.ly* e a *Treemaps*, aplicada para analisar

extensa quantidade de dados, que gera o padrão de visualização conhecido como *Squarified* (retângulos alongados<sup>9</sup>).

- **Convergência.** Esta categoria foi agregada às demais a partir do entendimento de que as bases de dados são agentes centrais no processo de convergência jornalística (BARBOSA, 2008, 2009), conformando plataformas de gestão de conteúdos que dão suporte às redações integradas, segundo a lógica do jornalismo convergente, multiplataforma, que também requer habilidades múltiplas dos profissionais sobre os modos de narrar, o uso da hipertextualidade, da multimídia, da memória, da interatividade nas peças jornalísticas, sobre os formatos e as maneiras de apresentar os conteúdos.

## Modos de narrar: a hipernarrativa em base de dados

Ao compreender o Jornalismo Guiado por Dados como um dos aspectos do Paradigma JDBD, consideramos os produtos e os formatos dele derivados como configuradores de modos novos de narrar, assegurados pela hipernarrativa em base de dados. A referência principal aqui é o trabalho de Lev Manovich (2001), o qual atribui *status* de nova forma cultural do contemporâneo para as bases de dados, concepção adotada por alguns pesquisadores do campo do jornalismo, do cinema e das artes na nova mídia.

Segundo Manovich, as bases de dados possibilitam a redefinição da concepção de narrativa, gerando a hipernarrativa (numa analogia ao hipertexto), resultado da soma de múltiplas trajetórias efetuadas através das bases de dados. Ou seja, num nível material, como sustenta o autor, a narrativa é apenas um conjunto de *links*, pois os elementos em si permanecem armazenados em uma base de dados (MANOVICH, 2001, p. 231). Para demonstrar a pertinência da sua formulação, ele aplica a teoria semiológica que prevê as dimensões sintagmática (relações se dão de maneira explícita) e paradigmática (relações ocorrem de maneira implícita) para compreender os relacionamentos entre os elementos de um sistema.

Assim, a base de dados, que até então cumpria um papel implícito, adquire existência material, enquanto a narrativa, de explícita, passa à condição de desmaterialidade. O paradigma, portanto, é privilegiado sobre o sintagma e, desse modo, a narrativa se torna virtualizada e a base de dados passa a existir materialmente (ibidem). Essa é a hipernarrativa presente em objetos e produtos da nova mídia, já que são produzidos a partir de base de dados e, neles, traços da narrativa tradicional também estão presentes, resultando em muitos híbridos. Por outra parte, na leitura, na interação ou na exploração desses produtos (entre eles um cibermeio), o usuário atua como um "performador" de ações para levar a narrativa adiante. Pois, se ele não se movimenta, clicando nos *links* e escolhendo o que lerá ou verá em seguida, a narrativa para. Complementando, Ellen Lupton (2010, p. 93)

<sup>9</sup> Ver aplicação no *Newsmap.jp* e em: <<http://news.bbc.co.uk/2/hi/8562801.stm>>. Sobre a técnica: <[www.cs.umd.edu/hcil/treemap/](http://www.cs.umd.edu/hcil/treemap/)>.

afirma que as bases de dados constituem a estrutura definidora da informação, que dá sustentação a jornais, revistas, jogos eletrônicos, catálogos etc., gêneros que criam um espaço de informação.

Outros investigadores também têm acolhido a concepção de Manovich como partida para propor novos entendimentos acerca da narrativa. Entre eles, alguns são da área do cinema (KINDER, 2002: "narrativa interativa em base de dados"; CAMERON, 2008: "narrativa modular"; SIMONS, 2008: "narrativa complexa"), enquanto outros são pesquisadores das artes na nova mídia, a exemplo de Christiane Paul (2007) e Victoria Vesna (2007), que vão trabalhar na perspectiva de uma estética base de dados: termo usado para descrever os princípios estéticos aplicados na imponente lógica das bases de dados para qualquer tipo de informação, filtrando coleções e visualizando os dados (PAUL, 2007, p. 95).

No campo do jornalismo, encontramos investigadores que empregam a concepção de narrativa interativa ou hipernarrativa desenvolvida por Manovich (2001) para pensá-la em relação à composição de narrativas jornalísticas. Para Machado (2006), por exemplo, a narrativa, em vez de uma sucessão de ações, configura-se cada vez mais como uma viagem através do espaço constituído pelos conjuntos estruturados de itens organizados na forma base de dados e torna-se um conjunto contínuo de ações narrativas e explorações (2006, p. 50). Ele argumenta que, ao contrário da narrativa moderna, na qual o ouvinte, o leitor ou o telespectador acompanha a narração sem interferir na lógica interna das ações, o fluxo dos novos modelos de narrativa, mais do que incorporar, vai depender diretamente da intervenção do usuário.

Mielniczuk (2008), por sua vez, destaca que numa plataforma de gestão de conteúdos jornalísticos estruturada em base de dados, o hipertexto funciona conforme a lógica de um sistema dinâmico, o que implica diretamente mudanças para as narrativas jornalísticas. Principalmente, quanto ao uso dos *links*, que "ganham movimento/animação fazendo com que as lexias<sup>10</sup> hipertextuais tornem-se muito mais 'fluidas' (a troca de uma lexia para a outra já não depende mais de uma ação, de um clique)". Como destaca a autora, é no transcorrer da narrativa que as lexias podem ir fluindo de uma para a seguinte, e assim por diante. "Poderíamos pensar que, nesses casos, os *links* seriam invisíveis" (ibidem, p. 173).

Já Ramos (2011a, 2011b), tomando como parâmetro o Paradigma JDBD e considerando-o um texto da cultura – segundo o referencial da escola de semiótica russa –, compreende que as bases de dados dominam hierarquicamente a criação e a manipulação de conteúdos diversos, estando no centro da criação jornalística e gerando linguagens, modelizadas pelos formatos<sup>11</sup>. O resultado desse processo, segundo ela, são experiências narrativas, pois em estado puro não há narrativas nas bases de dados (2011b, p. 11). Para a autora, uma base de dados admite a narração, mas não há nada na sua lógica que

<sup>10</sup> Na teoria do hipertexto, é a unidade mínima da narrativa, ao passo que, na narratologia, o texto é a unidade mínima da narrativa.

<sup>11</sup> Para Ramos, formatos próprios ao meio digital levam em conta o signo informático e as características estruturais do texto JDBD.

fomente a criação de uma história. Conforme afirma, uma narrativa como uma navegação em bases de dados seria o que Jenkins (2008, p. 158) chama de narrativa transmidiática, “justamente como uma possibilidade de múltiplas combinações em diversas bases de dados, ou, como diz o autor, em diversas mídias” (2011b, p. 12).

Neste artigo, em consonância com a conceituação do Paradigma JDBD e a constatação de extensibilidades para tal modelo (principalmente com o Jornalismo Guiado por Dados), avalia-se que a noção de hipernarrativas ou narrativas interativas dinâmicas produzidas a partir de conteúdos jornalísticos estruturados em bases de dados é adequada para caracterizar o que se encontrou no mapeamento de exemplos – descritos a seguir – que materializam modos novos de narrar, novos formatos para os conteúdos, assim como as distintas possibilidades para a visualização das informações, as quais estão afinadas também com uma estética base de dados. Infere-se que são padrões que estão se estabilizando porque são mais afinados com os aprimoramentos técnicos e com a melhor compreensão acerca das propriedades específicas do jornalismo digital nessas duas décadas de desenvolvimento.

Também pode-se considerar como novas convenções de mídia que geram padrões culturais compartilhados e nos permitem dar sentido a informações complexas, de acordo com Janet Murray (2012, Kindle location 264-5405). Para ela, a criação dessas convenções, assim como de *templates* é necessária para organizar o universo expandido da cobertura jornalística: “Templates de notícias online potencialmente podem conter muito mais informação como imagens em movimento, documentos originais, visualizações interativas, comentários dos leitores, sistemas de recomendação” (ibidem, 279-5405). Mas, para isso, afirma: é preciso estabelecer formas estáveis de mídia comparáveis à metáfora do jornal impresso. E isso demanda esforço coletivo tanto de jornalistas como de *designers* da informação.

## Mapeamento de exemplos empíricos

O conjunto de exemplos pesquisados, totalizando 48 *sites* analisados, está intrinsecamente ligado a produtos e formatos trabalhados segundo a noção do Jornalismo Guiado por Dados e, portanto, concernentes com o que se define como Paradigma JDBD.

Partimos do mapeamento de diferentes exemplos que abrangem desde fontes públicas disseminadoras de dados, modelizados segundo diversos padrões de design, até formatos específicos de conteúdos jornalísticos que empregam recursos diferenciados para estruturar as hipernarrativas. Esses exemplos estão publicados e distribuídos em *sites* de cibermeios do *mainstream*, organizações independentes, governos, cibermeios originalmente nascidos na *web*, além de *sites* de universidades ou grupos de pesquisa.

Por meio da análise dos exemplos chegou-se a uma tipologia que tem por objetivo melhor ordenar e distinguir as diversas iniciativas encontradas. Essa tipologia considera: *Mashups*, *Data/Dados*, *Sistemas de Recomendação de Notícias (SRN)*, *Fontes*, e são assim caracterizados:



- **Mashups.** Consiste na produção de conteúdos, de formatos distintos, a partir da combinação de variadas fontes públicas de dados. Com isso, os novos conteúdos gerados podem ganhar apresentações também diversificadas, afinadas assim com a estética base de dados. Exemplos de formatos criados através de *mashups* são infografias interativas e *newsgames*.
- **Data/Dados.** Nova denominação para editorias/seções em alguns cibermeios nas quais são apresentados conteúdos produzidos a partir de dados disponibilizados pelos governos de países, estados, cidades, entre outros. Exemplos são as editorias *Data*, no *Guardian*; *Data Desk*, no *Los Angeles Times*; e *Dados*, na *Gazeta do Povo* (Paraná)<sup>12</sup>.
- **Sistemas de Recomendação de Notícias (SRN)** ou estatísticas dinâmicas. Permitem visualizar – a partir do emprego de variadas técnicas de visualização da informação – o que é mais popular na cobertura informativa de um site jornalístico ou cibermeio no *real time* em que acessamos. Aparecem principalmente como seções sob as seguintes denominações: “*Most popular*”, “Notícias mais lidas”, “Mais comentadas”, “Mais compartilhadas”, “Mais enviadas”, “Mais vistas”.
- **Fontes.** Sites mantidos por governos de países, estados, cidades, além daqueles de organizações independentes que distribuem documentos referentes à administração pública, entre outros, sob os mais diversos formatos. Também como “Fontes” encaixam-se as iniciativas de sites de organizações jornalísticas que, através de APIs (Application Programming Interface<sup>13</sup>), disponibilizam dados coletados e estruturados por suas equipes. É o que faz o *Guardian*, por exemplo.

Feito isso, cruzamos nossas descobertas quanto às iniciativas de aplicação do Jornalismo Guiado por Dados às categorias expressivas do Paradigma JDBD, o que nos permite corroborar as proposições trabalhadas neste artigo quanto à existência de extensibilidades concretizadas em novos modos de narrar, de uma diversidade de novos formatos para conteúdos jornalísticos e de novas maneiras para apresentá-los.

## Análise aplicada

### Mashups

Entre os exemplos encontrados em cibermeios brasileiros e estrangeiros, detivemos em dois que foram selecionados por conta da composição do conteúdo ofertado e das diferenças de visualização. O “*SOPA Opera – Where Do Your Members of Congress Stand on SOPA and PIPA?*”, desenvolvido pela ProPublica<sup>14</sup>, é uma infografia interativa,

<sup>12</sup> <[www.guardian.co.uk/data](http://www.guardian.co.uk/data)>; <<http://datadesk.latimes.com/>> e <[www.gazetadopovo.com.br/dados/](http://www.gazetadopovo.com.br/dados/)>.

<sup>13</sup> Conjunto de instruções de programação para acesso e comunicação a determinado aplicativo de *software* baseado na plataforma *web* por outro aplicativo.

<sup>14</sup> Organização independente dos Estados Unidos, especializada em jornalismo investigativo: Ver em: <<http://projects.propublica.org/sopa/>>. Acesso em: 13 fev. 2012.

permanentemente atualizada. Funciona como uma base de dados sobre os membros do congresso dos Estados Unidos e permite conhecer seus posicionamentos políticos diante do projeto de lei que visa alterar regras sobre compartilhamento de conteúdos na *web*, o *Stop Online Piracy Act (SOPA)*. Nela, o leitor/usuário pode filtrar as possibilidades de acesso para navegar, interagir, envolver-se com o conteúdo. Os dados empregados são de fontes originárias de iniciativas *open-data* independentes e do próprio governo estadunidense: *OpenSecrets* (Center for Responsive Politics<sup>15</sup>); *OpenCongress*<sup>16</sup>; *GovTrack.us*<sup>17</sup>; *Library of Congress THOMAS*<sup>18</sup>, e do jornal *The New York Times* (*New York Times Congress API*<sup>19</sup>).

O segundo é o especial da *BBC.co.uk*, "Seven Billion People"<sup>20</sup>, cujo tema é a população da Terra. Neste especial, que usa mecânica de jogos como uma das técnicas para compor e apresentar o diversificado conteúdo, dados de diferentes instituições foram cruzados, combinados. Todos aqueles referentes à população, por exemplo, estão baseados em estimativas da UN Population Division<sup>21</sup> e todos os cálculos são fornecidos pelo UN Population Fund<sup>22</sup>. Os dados restantes são de outras seções da ONU, como Global Footprint Network<sup>23</sup> e International Telecommunication Union<sup>24</sup>.

Os dois exemplos estão em conformidade com as categorias do Paradigma JDBD, pois: refletem a dinamicidade; possuem flexibilidade, porque facilitam a recuperação dos conteúdos para contextualização; também possuem a característica da densidade informativa, pois, à medida que novas informações são apuradas, os conteúdos são atualizados, assegurando o uso de fontes diversas; assim como estão marcados pela diversidade temática. Ambos também apresentam modos diferenciados para visualizar informações jornalísticas. O especial da *BBC.co.uk* empregou a técnica *Visual.ly*, gratuita, enquanto a *ProPublica* usou ferramenta própria para a interface em mosaico, formada por fotos de cada congressista. A variedade de formatos (texto, foto, infografia, slideshows, mapas etc.) para os conteúdos revela, ainda, a convergência. Outra propriedade presente é a de inter-relacionamento/hiperlinkagem. Ela aparece com maior intensidade no exemplo da *ProPublica*: a interface de imagens em mosaico é toda clicável, guiando o leitor/usuário para novos conteúdos e possibilitando maior aprofundamento e contextualização no tema.

## Data/Dados

Pelo que se observa, a incorporação das novas editorias/seções denominadas *Data/Dados* compondo o *menu* informativo dos cibermeios é crescente. Nelas, se disponibiliza

<sup>15</sup> <[www.opensecrets.org/](http://www.opensecrets.org/)>.

<sup>16</sup> <[www.opencongress.org/](http://www.opencongress.org/)>.

<sup>17</sup> <[www.govtrack.us/](http://www.govtrack.us/)>.

<sup>18</sup> <<http://thomas.loc.gov/home/thomas.php>>.

<sup>19</sup> <[http://developer.nytimes.com/docs/congress\\_api](http://developer.nytimes.com/docs/congress_api)>.

<sup>20</sup> <<http://www.bbc.co.uk/news/world-15391515>>. Acesso em: 13 fev. 2012.

<sup>21</sup> <[www.un.org/esa/population/](http://www.un.org/esa/population/)>.

<sup>22</sup> <[www.unfpa.org/public/](http://www.unfpa.org/public/)>.

<sup>23</sup> <[www.footprintnetwork.org/en/index.php/GFN/](http://www.footprintnetwork.org/en/index.php/GFN/)>.

<sup>24</sup> <[www.itu.int/en/Pages/default.aspx](http://www.itu.int/en/Pages/default.aspx)>.

uma diversidade de dados, presentes em reportagens, geralmente relacionados à região onde a organização jornalística a qual pertence o cibermeio está localizada, assim como modelos diferenciados de visualização das informações são empregados para envolver o usuário, principalmente ativando recursos multimídia e de interatividade. Esse tipo de ação reforça o Jornalismo Digital em Base de Dados como paradigma e possibilita o surgimento de novos modos de narrar e de formatos para conteúdos noticiosos.

Em tais editorias, nas quais se encontrou os elementos pertinentes às categorias descritas na introdução deste artigo para o Paradigma JDBD, há também dados disponíveis para qualquer usuário criar e desenvolver novas aplicações a partir deles. Outra vantagem é que, com a proliferação dessas aplicações – as *APIs* – o público tem diversificadas orientações para ler e contextualizar determinados assuntos<sup>25</sup>.

### Sistemas de Recomendação de Notícias

Os SRN são formatados por algoritmos e linguagens de programação que selecionam itens, dados, de uma determinada base de dados e personalizam o modo como uma informação específica será visualizada no momento em que se acessa o *site* (horário e região, por exemplo, são características definidoras de personalização de alguns SRN). Esse tipo de conteúdo – consolidado em seções nas *homes* e nos menus informativos – confere dinamicidade aos cibermeios, ao tempo que também é caracterizado por um alto nível de automatização. A indicação para as notícias “Mais lidas”, “Mais comentadas”, “Mais compartilhadas” etc. está amplamente presente em cibermeios nacionais e estrangeiros. As possibilidades de hierarquização da informação nos SRN variam de acordo com as técnicas utilizadas. Contudo, a presença de uma base de dados é indispensável para o seu funcionamento, o que faz com que elas sejam parte definidora dos SRN – outro formato para os conteúdos jornalísticos materializados pelo Paradigma JDBD.

Em “*Nube de Noticias*”, SRN do argentino *LaNación.com.ar*, emprega-se a ferramenta de visualização *Treemaps*<sup>26</sup>, uma das que estão relacionadas com a estética base de dados. Por meio dela, pode-se: ver como a informação noticiosa é hierarquizada; ver os conteúdos mais lidos ou aqueles que receberam mais comentários do público; consultar o arquivo dos conteúdos publicados no dia em que se está pesquisando ou de conteúdos postados há sete dias ou há um mês. Esses SRN permitem, ainda, ler material específico de determinada editoria ou combinar uma hierarquização entre todas as seções do cibermeio, refletindo, assim, o quão dinâmico, flexível, hiperlinkado e automatizado é. Também possui vasta densidade informativa, diversidade temática e conteúdos em multimídia.

<sup>25</sup> Ver tabela com exemplos em: <<http://migre.me/aTeo9>>.

<sup>26</sup> No *Guardian.co.uk*, em “*Zeitgeist*” (“espírito do tempo”), a ferramenta também é empregada: <[www.guardian.co.uk/zeitgeist](http://www.guardian.co.uk/zeitgeist)>. Acesso em: 13 fev. 2012.

## Fontes

No Jornalismo Guiado por Dados (BRIGGS, 2010; BRADSHAW, ROHUMAA, 2011; ROGERS, 2011; GRAY, BOUNEGRU, CHAMBERS, 2012), o jornalista deve saber pesquisar em vastas coleções de dados, interpretá-los e reorganizá-los para produzir um conteúdo noticioso mais contextualizado e com base documental mais ampla. Os materiais presentes em *sites* disponibilizados por governos de países, estados, cidades, dentre outros, mantidos por entidades e organizações independentes (como WikiLeaks) empregam recursos diversos para apresentar tais informações em distintos formatos, constituindo-se como importantes ferramentas para o trabalho jornalístico. A habilidade para aplicar e lidar com técnicas como o *data mining* ou mineração de dados (indicada nos processos do Jornalismo Guiado por Dados) **é uma das recomendadas para ser incorporada** pelo jornalista e é cada vez mais essencial, tendo em vista, ainda, a realidade das rotinas produtivas em redações integradas.

Com a denominação “Fontes” como uma tipologia acentuamos tanto a ampliação crescente no fluxo de informações disponíveis como a variedade criada para a apresentação dessas informações segundo diversos formatos. Por isso, entendemos que está igualmente relacionada com as categorias de dinamicidade, flexibilidade combinatória, diversidade temática, automatização, hiperlinkagem, densidade informativa, visualização e convergência.

A tendência mundial de governos em todo o mundo de aumentar a transparência, disponibilizando dados referentes à administração pública sob licenças abertas, assim como a expansão de projetos não governamentais que compartilham do mesmo ideal de criar iniciativas *open-data*<sup>27</sup>, certamente favorecerão a expansão do Jornalismo Guiado por Dados. Há, ainda, exemplos de *sites* de organizações jornalísticas que possuem seções<sup>28</sup> para disponibilizar dados produzidos, organizados e estruturados por seus jornalistas através de APIs.

O Brasil está seguindo essa tendência, embora num ritmo mais lento. São exemplos: iniciativas governamentais como o “Data.gov.br” e a Lei de Acesso à Informação – em vigor desde 16/5/2012 – bem como as que estão sendo levadas a cabo por organizações jornalísticas, como *O Estado de S. Paulo*, *Gazeta do Povo*, *Folha de S. Paulo*, e a agência Pública<sup>29</sup>, que vêm desenvolvendo projetos concernentes com a noção do Jornalismo Guiado por Dados. Para nós, tal cenário confirma a proposição defendida neste artigo quanto à expansão do Paradigma Jornalismo Digital em Base de Dados no jornalismo contemporâneo.

<sup>27</sup> Ver exemplos mapeados em: <<http://migre.me/aTeo9>>.

<sup>28</sup> Seções: Open Platform do *Guardian* <<http://www.guardian.co.uk/open-platform>>; e Developer Network do *The New York Times* <<http://developer.nytimes.com/>>.

<sup>29</sup> <<http://apublica.org/>> .

Suzana Oliveira Barbosa é professora do Depto de Comunicação e PósCOM da UFBA e doutora em Comunicação e Cultura Contemporâneas pela Universidade Federal da Bahia - UFBA. É pós-doutora pela Universidade de Santiago de Compostela. É autora do livro *Jornalismo digital de terceira geração*. Coordena o Projeto Laboratório de Jornalismo Convergente .

suzanaobarbosa@uol.com.br

Vitor Torres é mestrando no Programa de Pós-Graduação em Comunicação e Cultura Contemporâneas, da UFBA.

vitortorres.mid@gmail.com

## Referências

BARBOSA, S. Jornalismo em ambientes dinâmicos: perspectivas, tendências e desafios para a criação de conteúdos em tempos de convergência. In: Actas III Congreso Internacional de Ciberperiodismo y Web 2.0. Bilbao, Espanha: Universidad del País Vasco, 2011.

\_\_\_\_\_. Modelo JDBD e o ciberjornalismo de quarta geração. In: FLORES VIVAR, J. M.; RAMÍREZ, F. E. (Ed.). **Periodismo Web 2.0**. Madrid: Editorial Fragua, 2009. p. 271-283.

\_\_\_\_\_. Modelo Jornalismo Digital em Base de Dados (JDBD) em Interação com a Convergência Jornalística. In: **Textual & Visual Media. Revista de la Sociedad Española de Periodística**. vol. 1, Madrid, 2008, p. 87-106.

\_\_\_\_\_. Jornalismo Digital em Base de Dados (JDBD) – um paradigma para produtos jornalísticos digitais dinâmicos. (Tese de Doutorado). PósCOM/UFBA, 2007. Disponível em: <<http://migre.me/aTuYN>>. Acesso em: 4 fev. 2012.

BRADSHAW, P.; ROHUMAA, L. **The Online Journalism Handbook**. Skills to survive in the digital age. Harlow, England: Pearson, 2011.

BRIGGS, M. **Journalism Next: A Practical Guide to Digital Reporting and Publishing**. Washington, DC: CQPress, 2010.

CAMERON, A. **Modular narratives in contemporary cinema**. Houndmills, Basingstoke: Palgrave Macmillan, 2008.

COLLE, R. **Explotar la información noticiosa**. Data mining aplicado a la documentación periodística. Madrid: Universidad Complutense de Madrid, 2002.

FIDALGO, A. A resolução semântica no jornalismo online. In: BARBOSA, S. (Org.). **Jornalismo digital de terceira geração**. Covilhã, PT: LivrosLabCOM, 2007. p. 93-102.

\_\_\_\_\_. Do poliedro à esfera: os campos de classificação. A resolução semântica no jornalismo online. In: Anais II Encontro Nacional da SBPjor. Salvador-BA/Brasil, 2004.

GRAY, J.; BOUNEGRU, L.; CHAMBERS, L. (Ed.). **The Data Journalism Handbook**. How Journalists Can Use Data to Improve the News. Sebastopol: O'Reilly Media, 2012.

JENKINS, H. **Cultura da convergência**. São Paulo: Aleph, 2008.

KINDER, M. Hot spots, avatars, and narrative fields forever: Bunuel's legacy for new digital media and interactive database narrative. In: *Film Quarterly* 55, n. 4, 2002, p. 2-15. Disponível em: <<http://migre.me/aTvtT>>. Acesso em: 4 fev. 2012.

LUPTON, E. **Thinking with type**: a critical guide for designers, writers, editors, & students. New York: Princeton Architectural Press, 2010.

MACHADO, E. **O Jornalismo Digital em Base de Dados**. Florianópolis: Calandra, 2006.

MANOVICH, L. **The language of new media**. Cambridge, MA: MIT Press, 2001.

MIELNICZUK, L. et al. Estudos iniciais sobre a concepção de Resolução Semântica no jornalismo digital. In: Anais 19º Encontro Compós. n. 19, v.1, junho de 2010. Rio de Janeiro: PUC-Rio.

\_\_\_\_\_. O estudo da narratividade no ciberjornalismo. In: DÍAZ NOCI, J.; PALACIOS, M. **Metodologia para o estudo dos cibermeios**. Estudo da arte & perspectivas. Salvador: EDUFBA, 2008. p. 161-175.

\_\_\_\_\_. Jornalismo na web: uma contribuição para o estudo do formato da notícia na escrita hipertextual. (Tese de Doutorado). PósCOM-UFBA, 2003. Disponível em: <<http://migre.me/aTvx1>>. Acesso em: 27 ago. 2012.

MURRAY, J. *Inventing the Medium: Principles of Interaction Design as a Cultural Practice*. Cambridge, MA: MIT Press. Kindle Edition, 2012.

PAUL, C. The Database as System and Cultural Form: Anatomies of Cultural Narratives. In: VESNA, V. (Ed.). **Database Aesthetics**. Art in the Age of Information Overflow. Minneapolis: University of Minnesota Press, 2007. p. 95-109.

RAMOS, D. O. Formato: condição para a escrita do Jornalismo Digital em Bases de Dados. Uma contribuição da semiótica da cultura. (Tese de Doutorado). ECA/USP, 2011. Disponível em: <<http://migre.me/aTvzX>>. Acesso em: 27 ago. 2012.

\_\_\_\_\_. Jornalismo Digital em Base de Dados (JDBD) como um texto da cultura. In: Anais 20º Encontro Compós. n. 20, vol. 1, junho de 2011. Porto Alegre: UFRGS.

ROGERS, S. **Facts are Sacred**. The Power of Data. London: Guardian Books, 2011. [E-Book].

SIMON, J. Complex narratives. In: **New Review of Film and Television Studies**, 6: 2, 2008. p. 111-126. Disponível em: <<http://migre.me/aTvef>>. Acesso em: 4 fev. 2012.

VESNA, V. Seeing the World in a Grain of Sand: The Database Aesthetics of Everything. In: VESNA, V. (Ed.). **Database Aesthetics**. Art in the Age of Information Overflow. Minneapolis: University of Minnesota Press, 2007. p. 3-38.

*Artigo recebido em julho de 2012  
e aprovado em janeiro de 2013.*