



Universidade do Estado do Rio de Janeiro

Centro de Educação e Humanidades

Faculdade de Comunicação Social

Felipe Rodrigues Grandin

A contribuição do jornalismo guiado por dados para a criação de valor nas organizações jornalísticas

Rio de Janeiro

2014

Felipe Rodrigues Grandin

**A contribuição do jornalismo guiado por dados para a criação de valor nas organizações
jornalísticas**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-graduação em Comunicação, da Universidade do Estado do Rio de Janeiro. Área de concentração: Cultura de Massa, Cidade e Representação Social.

Orientadora: Prof^ª. Dr^ª. Sonia Virgínia Moreira

Rio de Janeiro

2014

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CEH/A

G753 Grandin, Felipe Rodrigues.
A contribuição do jornalismo guiado por dados para a criação de valor nas organizações jornalísticas / Felipe Rodrigues Grandin. – 2014.
121 f.

Orientadora: Sonia Virgínia Moreira.
Dissertação (Mestrado) – Universidade do Estado do Rio de Janeiro.
Faculdade de Comunicação Social.

1. Comunicação Social – Teses. 2. Jornalismo – Teses. 3. Jornais brasileiros – Teses. I. Moreira, Sonia Virgínia. II. Universidade do Estado do Rio de Janeiro. Faculdade de Comunicação Social. III. Título.

es

CDU 080(81)

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

Assinatura

Data

Felipe Rodrigues Grandin

**A contribuição do jornalismo guiado por dados para a criação de valor
nas organizações jornalísticas**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre ao Programa de Pós-graduação em Comunicação, da Universidade do Estado do Rio de Janeiro. Área de concentração: Cultura de Massa, Cidade e Representação Social.

Aprovada em 21 de março 2014.

Banca Examinadora:

Prof.^a. Dr.^a. Sonia Virgínia Moreira (Orientadora)

Faculdade de Comunicação Social da UERJ

Prof.^a. Dr.^a. Letícia Cantarela Matheus

Faculdade de Comunicação Social da UERJ

Prof. Dr. João Paulo de Jesus Faustino

Universidade do Porto

Rio de Janeiro

2014

RESUMO

GRANDIN, Felipe Rodrigues. *A contribuição do jornalismo guiado por dados para a criação de valor nas organizações jornalísticas*. 2014. 121 f. Dissertação (Mestrado em Comunicação) – Faculdade de Comunicação Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2014.

O jornalismo passa por um momento de transição como resultado do surgimento das novas tecnologias de informação e comunicação, que tem fortes reflexos sociais, políticos, culturais e econômicos, tornando a sociedade contemporânea mais global, horizontal e transparente (CASTELLS, 1999, 2003; JENKINS, 2008; MANOVICH, 2001). Um dos efeitos é o crescimento exponencial no volume de informações em circulação, bem como a transformação da maneira como são recebidas, manipuladas e compreendidas. As organizações jornalísticas – que historicamente tiveram forte controle sobre suas atividades e mercados – precisam lidar agora com a concorrência onipresente pela coleta, seleção, produção e distribuição de notícias e informações (ALBARRAN, 2010; DOYLE, 2013; PICARD, 2009, 2013; RICHERI, 2012). O resultado é a perda de audiência e de anunciantes, que ameaçam a sustentabilidade econômica da indústria. O Jornalismo Guiado por Dados (JGD) emerge nesse contexto de crise financeira e estrutural das organizações jornalísticas. O aumento exponencial das bases de dados de interesse público criou uma nova oportunidade para que o jornalismo assuma a frente no exercício dessas funções. O objetivo dessa dissertação é identificar como o JGD pode ser usado para resgatar a relevância e o valor do jornalismo ao inserir novas práticas da computação e das ciências sociais em seu processo produtivo e criando novos produtos e serviços. Tendo o estudo de caso como base para a pesquisa, centrado no núcleo Estadão Dados, do jornal *O Estado de S. Paulo*, busca-se um melhor entendimento sobre essa prática e seus possíveis benefícios para a sustentabilidade das organizações jornalísticas.

Palavras-chave: Jornalismo guiado por dados. Criação de valor. Economia de mídia. Produção de sentidos

ABSTRACT

GRANDIN, Felipe Rodrigues. *The contribution of data-driven journalism to value creation in media organizations*. 2014. 121 f. Dissertação (Mestrado em Comunicação) – Faculdade de Comunicação Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2014.

Journalism is going through a moment of transition as a result of a broader set of changes caused by the emergence of new communication and information technologies. These changes have disruptive social, political, cultural and economical impacts, which are redefining contemporary society as one more global, horizontal and transparent (CASTELLS, 1999, 2003; JENKINS, 2008; MANOVICH, 2001). Among other effects, this phenomena result in an exponential growth of production, storage and distribution of information, changing as well the way people receive, manipulate and understand information (GOLDHABER, 2006, 2012; SIMON, 1972; LANTHAM, 1994). Journalism organizations, which for decades long had strict control of their activities and markets, now have to deal with ubiquitous competition over production, gathering, selection and distribution of information and news (ALBARRAN, 2010; DOYLE, 2013; PICARD, 2009, 2013; RICHERI, 2012). The result is a general loss of audience and advertisement that threatens the industry's sustainability. Data-driven Journalism emerges in this context of structural and financial crises. Shaken by increased competition and lost of hegemony, journalists have to deal with the devaluation of functions that traditionally guaranteed its economical value: Accessing sources, determining the significance of information, and conveying it effectively (PICARD, 2009). The exponential rise of available public interest information created an opportunity for journalists to step up and lead the process of making sense of these data. The current dissertation aims to identify in what ways data-driven journalism can be used to rescue the value and relevance of journalism to the society by using new methods, skills and tools, and creating new products and services. The case study method is the baseline for this research, centered in the Estadão Dados core, from *O Estado de S. Paulo* newspaper, this work aims to better understand these new practices and their possible positive effects on the sustainability of journalism organizations.

Keywords: Data-driven journalism. Value creation. Media economics. Sensemaking.

AGRADECIMENTOS

Agradeço a família, pelo amor e o apoio incondicional. Aos meus pais, Ana Léa e Carlos, minha mulher, Mariana, minha filha, Júlia, minha irmã, Carla, meu cunhado, Eduardo e minha sobrinha, Laura. Amo vocês.

Agradeço especialmente a minha mãe, maior incentivadora da minha vida acadêmica, cujo apoio, carinho, persistência e paciência foram coautores deste trabalho.

A meu pai, pelo exemplo, que me inspirou a seguir a profissão jornalística – mesmo que a contragosto – e me ensinou a diferença entre o bom e o mau jornalismo.

A minha orientadora, Sonia, pela confiança, o apoio, a orientação e a amizade. Por combinar como poucas pessoas seriedade e bom humor. Por acreditar em mim e no meu projeto desde o início.

Ao meu padrinho, Alceste, que sempre me ajudou e incentivou na vida e nos estudos, por quem tenho enorme carinho.

Aos professores Paulo Faustino e Letícia Matheus, que me ajudaram bastante apontando caminhos para melhorar a pesquisa.

Aos professores do PPGCOM-UERJ, pelo apoio e por compartilharem seus conhecimentos. Aos colegas de mestrado que tornaram essa experiência mais alegre.

Aos colegas da Brasil Energia, por compreenderem a importância da vida acadêmica e terem me dado a liberdade necessária para fazer essa pesquisa.

Por fim, a minha mulher, que está sempre ao meu lado, agradeço o companheirismo, a compreensão e a paciência, sem as quais esse trabalho não seria possível.

E a Julinha, por nos fazer felizes todos os dias.

LISTA DE FIGURAS

Figura 1-	Diagrama das causas de mortalidade do exército do Oeste, por Florence Nightingale	24
Figura 2 -	Coletando dados: um fluxograma de ferramentas e técnicas	39
Figura 3 -	Processo de criação de valor no Jornalismo Guiado por Dados	59
Figura 4 -	A pirâmide invertida do jornalismo de dados	60
Figura 5 -	Reprodução da ferramenta Basômetro	85

LISTA DE QUADROS

Quadro 1 - Perfil das equipes de Jornalismo Guiado por Dados nas redações de cinco veículos.....	31
Quadro 2 - Perfil das equipes de projetos de JGD no Brasil	32
Quadro 3 - Tipologias do JGD	36

SUMÁRIO

	INTRODUÇÃO	10
1	JORNALISMO GUIADO POR DADOS COMO FORMA CONTEMPORÂNEA DE PRODUÇÃO DE SENTIDO	23
1.1	Contexto histórico	23
1.2	O JGD nas redações	30
1.3	Em busca de um conceito de Jornalismo Guiado por Dados	33
1.4	Multidisciplinaridade	38
1.5	A produção de sentido no JGD	40
1.5.1	<u>A ilusão da objetividade dos dados</u>	43
1.5.2	<u>Limitações e riscos do JGD</u>	45
2	A CRIAÇÃO DE VALOR PELO JORNALISMO GUIADO POR DADOS	47
2.1	A economia de mídia	47
2.1.1	<u>A economia de mídia no Brasil</u>	50
2.2	O conceito de valor aplicado ao jornalismo	51
2.3	Criação de valor no JGD	58
2.3.1	<u>Acesso a fontes: dados públicos, informação exclusiva</u>	61
2.3.2	<u>Seleção e hierarquização de informações: pólo de dados confiáveis</u>	63
2.3.3	<u>Distribuição: Personalização do conteúdo</u>	63
2.3.4	<u>Produção de sentido e contextualização de volumes de informações</u>	65
2.3.5	<u>Venda e licenciamento de aplicativos e ferramentas</u>	66
2.4	Estratégias de criação de valor	67
3	ESTUDO DE CASO – ESTADÃO DADOS	70
3.1	Perfil do Grupo Estado	70
3.2	O Estadão Dados	73
3.2.1	<u>Perfil e contexto histórico</u>	73
3.2.2	<u>Diferenças em relação às práticas tradicionais</u>	76
3.2.2.1	Pauta	78
3.2.2.2	Produção	80
3.2.2.3	Formatação	81
3.2.2.4	Distribuição	82
3.2.3	<u>Impacto no impresso</u>	83

3.2.4	<u>Basômetro</u>	84
	CONCLUSÃO	90
	REFERÊNCIAS	100
	ANEXO A - Entrevista José Roberto De Toledo	108
	ANEXO B - Entrevista Daniel Bramatti.....	111
	ANEXO C - Entrevista Gustavo Faleiros	117

INTRODUÇÃO

O homem contemporâneo está exposto a um volume de informações inédito na história. Hilbert e López (2011) apontam que em 2007 cada pessoa já recebia, em média, quatro vezes mais dados que em 1986 – o equivalente a 174 jornais de 85 páginas por dia, provenientes da internet, celulares, rádios, jornais e TVs. Estima-se que trafeguem mensalmente no globo 10,8 Exabytes, equivalente a 2,7 bilhões de DVDs – o suficiente para armazenar 540 vezes todo o conteúdo existente nos arquivos da maior biblioteca do mundo, a Biblioteca do Congresso dos Estados Unidos. Mais do que um fenômeno tecnológico, a popularização das mídias digitais tem importantes ramificações sociais e forte impacto na maneira como se formam as representações sociais, aqui compreendidas como conjunto de ideias, significados e valores socialmente compartilhados, e as implicações psicológicas e econômicas desse fenômeno, como a superabundância de informações que leva à incapacidade de o homem contemporâneo processar todos os estímulos que recebe.

Ainda que na maior parte das vezes estejam invisíveis, as bases de dados são praticamente onipresentes na vida contemporânea. Por regerem a comunicação, o registro e a construção de narrativas nos meios digitais levaram Manovich (2001) a considerá-las não só uma nova forma cultural, mas a essência da expressão cultural na era dos computadores. Não por acaso, em praticamente todas as áreas de interesse público – saúde, educação, política, justiça, governo, criminalística etc. – há uma profusão de informações sendo colocadas na internet por governos, empresas, instituições e pessoas. Essas bases de dados podem, e devem, ser analisadas e usadas para produzir conteúdo jornalístico. É o que tem sido feito com sucesso por veículos como *The Guardian* e *The New York Times* em coberturas de impacto, como o caso dos escândalos Wikileaks¹ e a eleição presidencial norte-americana de 2012².

¹ Trata-se do vazamento de documentos secretos do exército norte-americano nas invasões do Afeganistão e do Iraque. Durante o ano de 2010, foram divulgados pelos veículos *The Guardian*, *The New York Times*, *Der Spiegel*, Al Jazeera e Bureau of investigative Reporting 92 mil relatórios confidenciais sobre a ação no Afeganistão e outros 391.832 documentos secretos sobre a Guerra do Iraque fornecidos pela organização sem fins lucrativos Wikileaks, com enorme repercussão nacional. A forma com essa imensa quantidade de dados foi analisada e divulgada especialmente pelos jornais *The New York Times* e *The Guardian*, que usaram novas técnicas e ferramentas de visualização de bases de dados, foi considerada inovadora

² Ao longo da eleição para a presidência dos Estados Unidos em 2012, o blogueiro Nate Silver aplicou metodologias estatísticas para analisar as chances de cada candidato a partir da imensa base de pesquisas eleitoras realizadas no país por diversos institutos e conferir um caráter científico à análise daqueles dados, até então avaliados de forma anedótica e impressionista por analistas políticos e membros de campanhas e

Esse processo ocorre ao mesmo tempo em que as organizações jornalísticas perdem de maneira significativa sua estabilidade e posição históricas na sociedade e têm sua sustentabilidade ameaçada por novas formas de produzir e consumir notícias, que estão abalando práticas e modelos de negócios centenários, sem que uma resposta à altura tenha sido dada.

Se revermos as mudanças ocorridas nas organizações noticiosas nas últimas décadas, fica claro que a evolução tem sido restrita, limitada, superficial e incapaz de preservar uma idade de ouro do jornalismo. (...) Concomitante com os esforços limitados de inovação, a retenção dos lucros pelos proprietários corporativos, cujos interesses primordiais não correspondem ao aperfeiçoamento das organizações noticiosas e à sua habilidade em exercer seu papel social, tem sido constante e provoca um deletério desperdício de recursos (PICARD, 2013, p. 23).

Na sociedade contemporânea mediada pelas tecnologias digitais, em que as bases de dados surgem como forma simbólica dominante (MANOVICH, 2001), o papel dos jornalistas como mediadores da informação está passando por uma radical transformação. Até o surgimento e a popularização das mídias digitais os profissionais e organizações jornalísticas tinham papel predominante na esfera pública na leitura, seleção e interpretação dos acontecimentos de maior relevância social. Hoje vêm perdendo a preponderância como resultado do acesso direto dos leitores/espectadores às fontes primárias de informação e ao espaço público por meio da internet e das mídias digitais. É a partir das bases de dados que são construídas as narrativas, nas quais o usuário tem papel ativo. É o leitor quem escolhe os caminhos que irá seguir e quando irá parar – construindo, portanto, cada narrativa a partir do acesso aos links armazenados nas bases de dados.

Está sendo redefinida a relação entre os meios de comunicação e suas audiências e também seus próprios papéis. Hoje, cada indivíduo transmite em média o equivalente a seis jornais diários, por meio de e-mails, mensagens de telefone e redes sociais (HILBERT; LÓPEZ, 2011). O cidadão comum prescinde cada vez mais da intermediação da imprensa para ter acesso a informações. Dados antes monopolizados por grupos de mídia estão à disposição de qualquer um que tenha interesse em analisá-los, processá-los e retransmiti-los. Governos, empresas e instituições estão publicando suas bases de dados na rede, por vontade própria ou por pressão da opinião pública. Estão surgindo novas formas de reportagem assim

partidos. A iniciativa de Silver foi vista inicialmente com ceticismo pelos especialistas, que debochavam da precisão de suas previsões em um cenário historicamente considerado incerto. Mas suas análises se mostraram acertadas ao identificar com bastante antecedência que Barack Obama seria reeleito e ao acertar os resultados em todos os 50 estados norte-americanos, superando, de longe, todos os outros analistas.

como uma nova geração de serviços, com o público participando do processo de produção de notícias.

Como resultado desse processo e, principalmente, da ampliação e diversificação dos canais de comunicação, o modelo de negócio tradicional do jornalismo, baseado no duplo mercado – anunciantes e audiência – está em crise devido à concorrência com outras plataformas que atacam nas duas frentes: disputam a verba publicitária e oferecem maior volume de conteúdo a um custo mais baixo ou mesmo gratuitamente. É natural que o papel do jornalista mude para se adaptar ao novo cenário, assim como das organizações jornalísticas.

Como afirma Picard (2009, p. 3), o valor econômico do trabalho jornalístico sempre resultou da distribuição do conhecimento alheio, por meio de três funções e suas habilidades associadas: acessar fontes, determinar a relevância das informações e transmiti-las de forma eficiente. Enquanto os profissionais e organizações de mídia tinham o monopólio sobre essas funções, seu valor era alto. Agora, com o desenvolvimento tecnológico da mídia, qualquer pessoa pode assumir esses papéis, o que reduz esse valor a praticamente zero. Para reverter essa situação, segundo Picard, é necessário inovar para criar valor, fornecendo informações e conhecimento que não estão disponíveis em outros lugares, em formas diferentes, mais úteis e relevantes para as audiências.

Nesse contexto, emerge uma nova prática que pode contribuir para retomar, ao mesmo tempo, a relevância do papel jornalístico, assim como parte da atratividade econômica perdida nas últimas décadas: o Jornalismo Guiado por Dados (BARBOSA; TORRES, 2013; BRADSHAW, 2013; GRAY; CHAMBERS; BOUNEGRU, 2012; LORENZ, 2011; MACHADO, 2006). Apesar de estarem perdendo a exclusividade em acessar fontes de informação e mediar determinadas relações, os meios de comunicação de massa, mais especificamente os veículos jornalísticos, acabam, na prática, mantendo o seu papel de mediador, mas de outra maneira. A função de transmissor e a de *gatekeeper* perdem valor em relação à de produtor de sentido e as organizações de mídia passam a ser um centro de referência de informações com credibilidade, um *hub of trusted data* (LORENZ; KAYSERBRIL; MCGHEE, 2011), o que pode ser traduzido como “pólo de dados confiáveis”, daí a emergência do JGD, que está sendo incorporado nos processos produtivos dos grupos de mídia e já é, em muitos casos, incluído na hierarquia organizacional por meio da criação de editorias ou até mesmo de empresas voltadas exclusivamente para a produção, interpretação e visualização de dados.

Entre as organizações pioneiras, estão aquelas cujo principal produto é jornalístico, informação, notadamente impressos diários ou agências de notícias. Jornais como *The New York Times* (EUA), *The Guardian* (Reino Unido), *El País* (Espanha), *Clarín* (Argentina) e *O Estado de S. Paulo* (Brasil), por exemplo, formaram equipes específicas dentro de sua estrutura e seções de dados em seus websites, oferecendo acesso a bases de dados, ferramentas de visualização e interpretação, com novas formas narrativas e de interação com o público.

Ao mesmo tempo – por suas características e, dentro de uma perspectiva mais aberta e colaborativa, por sua origem e desenvolvimento histórico – o JGD abre espaço para a atuação de qualquer pessoa, independentemente de ser ligada ou não a uma empresa ou organização jornalística. Para Simon Rogers (2012), o JGD tem o caráter quase revolucionário ao nivelar o campo da mídia e permitir a prática igualitária por qualquer um. Isso porque introduz uma prática nova, que exige habilidades diferentes das usualmente adotadas pelo jornalismo convencional e que pode ser realizada com o uso de dados e ferramentas gratuitos e disponíveis na internet. É um processo semelhante ao ocorrido com a emergência dos blogs e redes sociais usados para a publicação de notícias ou a cobertura de grandes eventos. No caso do JGD, porém, os jornalistas partiram muitas vezes em desvantagem por não ter as competências necessárias para a atividade, como a programação e a análise estatística. O sucesso das previsões do blogueiro Nate Silver na cobertura das eleições presidenciais norte-americanas de 2008 e 2012 – superando, sozinho, todos os conglomerados de mídia do país no índice de acertos – é uma amostra do potencial da introdução de novas habilidades no jornalismo.

Por isso é importante, então, analisar a emergência do JGD como produtor de sentidos em um ambiente de crescente circulação de informações e de onipresença das bases de dados – especialmente a criação de áreas voltadas para análise, processamento e visualização de dados dentro de conglomerados de mídia e os novos papéis exercidos pelos jornalistas, protagonistas desse fenômeno.

Ao longo da experiência de oito anos deste autor na produção de reportagens para os veículos *Jornal da Tarde*, *O Estado de S. Paulo* e *Brasil Energia*, observou-se a crescente necessidade de usar bases de dados públicas como fonte e suporte para a elaboração de reportagens. O movimento de transparência e abertura de dados por órgãos públicos ocorrido na última década disponibilizou uma enorme quantidade de informações de interesse social e jornalístico que não apenas abre espaço para novas leituras e compreensões, como permite superar fontes oficiais ou diretamente envolvidas nos temas. Em boa parte das vezes, porém,

esses dados não fazem sentido em estado bruto. Sendo passíveis de análise e compreensão da forma como são divulgados.

Como ilustração pode-se usar o exemplo da cobertura da área de segurança. No Brasil, os dados sobre segurança em cada estado são inseridos pelas Polícias Civil, que registra as ocorrências e é responsável pelas investigações dos crimes, e Militar, que faz o policiamento ostensivo e reprime diretamente a criminalidade. Ambas são subordinadas às secretarias de segurança estaduais, que controlam os dados de criminalidade e são diretamente responsáveis pela execução das políticas públicas de segurança. Na experiência do autor cobrindo o tema em São Paulo, é possível afirmar que há interesse em divulgar informações positivas, como a queda da criminalidade, e esconder as negativas, como o aumento da criminalidade, por parte daqueles que têm o controle dos dados. Não por acaso, as estatísticas de segurança estaduais são divulgadas com menor frequência e detalhamento do que o possível ou desejável para a compreensão da real situação da segurança em cada estado – no caso de São Paulo, passou a ser publicada mensalmente e por município apenas em 2011. Importante notar que os boletins de ocorrência são registrados em tempo real em uma base de dados integrada – ou seja, não há impedimentos técnicos para o aumento da frequência de divulgação – e incluem o detalhamento completo dos crimes – como, por exemplo, o local em que ocorreu.

Tradicionalmente busca-se superar esses obstáculos com o trabalho de reportagem e apuração a fim de conseguir informações mais detalhadas, por meio do vazamento de informações por membros das corporações ou do governo que tenham acesso aos dados completos. Outra forma é buscar informações nas bases de dados de mortalidade mantidas pelo Ministério da Saúde e pelas secretarias estaduais e municipais de saúde. O sistema da Secretaria de Saúde da Prefeitura de São Paulo³, por exemplo, permite o detalhamento de agressões físicas e psicológicas por faixa etária, idade, sexo, local, horário, entre outros, com atualização mensal e possibilidade de comparar os dados com os seis anos anteriores. Com base nele, é possível verificar as áreas da cidade que têm maior incidência de crimes, qual o perfil das vítimas mais vulneráveis e, com isso, validar, questionar e até mesmo sugerir políticas públicas, o que não seria possível apenas com base nas informações da polícia ou da Secretaria de Segurança⁴. Esse é apenas um exemplo dentre vários semelhantes, de como o uso de bases de dados pode contribuir para melhorar a cobertura jornalística e garantir o

³ Ver em: <<http://www.prefeitura.sp.gov.br/cidade/secretarias/saude/tabnet/>>

⁴ O uso das bases de dados para reduzir a dependência das fontes oficiais é discutida por Paraise e Dagiral (2013)

acesso a informações de interesse público e exclusivas, alternativas às fontes oficiais ou diretamente envolvidas.

Ao longo de três anos e sete meses, entre maio de 2007 e dezembro de 2010, este pesquisador trabalhou como jornalista no Grupo Estado, produzindo reportagens publicadas nos periódicos *Jornal da Tarde* e *O Estado de S. Paulo* e no portal estadão.com.br. Previamente, de setembro a dezembro de 2005, fez o Curso Intensivo de Jornalismo Aplicado, no próprio Grupo Estado, durante o qual conheceu a história e o funcionamento do grupo e acompanhou o trabalho cotidiano das redações de seus veículos.

Durante esse período pôde observar e participar do funcionamento da redação e acompanhar as dinâmicas da produção jornalística. Testemunhou a integração do impresso com o site, a introdução da multimídia no site estadao.com.br, com a criação da TV Estadão e a produção de vídeos pelos repórteres, além da criação do site do *Jornal da Tarde*. Como repórter, o pesquisador começou a entrar em contato, informalmente, com o uso de bases de dados no jornalismo, tema desta pesquisa, durante a apuração de reportagens. Em 2007 fez curso ministrado pela Associação Brasileira de Jornalismo Investigativo (Abraji) sobre Reportagem com Auxílio de Computador (RAC) e a partir daí começou a usar mais ativamente alguns métodos e bases de dados para reportagens sobre temas variados, como saúde, educação, segurança, transporte, política e meio ambiente. Foi possível não só praticar o RAC, usar as ferramentas e as práticas e métodos, como vivenciar as dificuldades para obtenção de informações públicas, a necessidade de se considerar o viés e as particularidades de cada base de dados e problemas, como falta de transparência de governos, baixa fidelidade de dados, entre outros pontos. O trabalho prático de reportagem no veículo analisado, especialmente a aplicação da RAC, foi determinante para identificar as peculiaridades do Jornalismo Guiado por Dados e as características que o diferenciam de outras práticas jornalísticas. Essa divisão não está bastante clara mesmo para muitos que trabalham com ambas as práticas. Sem a vivência profissional, seria muito mais difícil para este pesquisador separar as atividades.

Desde janeiro de 2011, este autor trabalha como jornalista na Editora Brasil Energia, primeiro como repórter e depois como editor da revista *Brasil Energia*, publicação mensal especializada no setor de energia. Durante esse período também desenvolveu trabalhos com o uso de RAC, aplicando as bases de dados públicas cotidianamente. A experiência foi mais prolífica, no entanto, no que se refere ao entendimento do mercado de mídia brasileiro, a gestão estratégica de empresas de mídia e a aplicação do conceito de valor econômico no jornalismo. Internamente, foi possível ter visão do funcionamento de uma empresa que

oferece serviços jornalísticos altamente especializados a uma audiência de nicho, tanto do processo produtivo como das estratégias de negócio. Externamente, o pesquisador teve contato cotidiano com os leitores, buscando saber quais informações e serviços têm maior ou menor valor para o público. Já que a publicação tem como público-alvo prioritário os executivos de empresas do setor de energia (sua principal função é fornecer informações que auxiliem na tomada de decisões estratégicas), considera-se internamente que entender o que tem ou não valor para esse público é parte inerente do trabalho cotidiano dos jornalistas que atuam no veículo. Esse público é particularmente interessante para esse tipo de pesquisa porque, pelo que observou o pesquisador, esses leitores sabem melhor do que o público comum qual o valor econômico dos produtos informativos, de acordo com o impacto em seus negócios.

A partir dessas experiências, este autor buscou mais informações para entender o fenômeno do JGD e aplicá-lo no cotidiano profissional e observou que, apesar de uma crescente e dinâmica comunidade de jornalistas interessados no tema, a literatura é bastante escassa, especialmente no âmbito acadêmico. Prova disso é que boa parte das referências bibliográficas diretamente relacionadas ao tema foram produzidas concomitantemente a este trabalho, entre 2012 e 2014. Mesmo nas grandes organizações jornalísticas do Brasil o uso das novas habilidades e práticas trazidas pelo jornalismo de dados na última década ainda é relativamente incipiente. Coincidentemente, *O Estado de S. Paulo* é um dos pioneiros na formação de uma equipe de dados e tem o grupo mais bem estruturado, o que foi a principal motivação para a escolha dessa organização específica para o estudo de caso, além do fato de ser um dos maiores e mais tradicionais grupos jornalísticos do país, que hoje passa por um momento de crise e mudança de estratégia, na qual o meio online tem importante papel.

Ao longo da experiência deste pesquisador como repórter, ao mesmo tempo em que crescia o interesse pelo JGD aumentava a inquietação sobre o impacto dos fatores econômicos não apenas no cotidiano dos profissionais como na qualidade das coberturas, das reportagens, e dos produtos jornalísticos como um todo. Pode-se afirmar sem medo de exagero que, em nenhuma outra época, o negócio esteve tão presente nas redações. E nunca a gestão e as estratégias foram tão importantes para a sobrevivência dos veículos. A crise do jornalismo (financeira e de identidade) é um fato declarado pelos seus próprios atores e por teóricos do campo da comunicação e da economia.

Assim surgiu a hipótese de que o jornalismo de dados poderia ser uma das saídas para recuperar o valor e a relevância das organizações de mídia. Dentre o pessimismo e o mal-estar que se observava nas redações, não apenas do *Jornal da Tarde* onde o autor trabalhava – e

que não por acaso foi fechado em 2012 –, mas de maneira geral, o jornalismo de dados aparecia como uma novidade positiva, que poderia contribuir para contrabalançar a queda de qualidade resultante dos cortes de equipes, a menor receita, e a perda de relevância, resultante da maior concorrência com outros veículos e meios digitais.

Esta pesquisa, portanto, é resultado dessas inquietações e uma tentativa de contribuir, ainda que de forma modesta, para a compreensão desses fenômenos e na busca de soluções para os problemas que afetam a prática jornalística na contemporaneidade.

Objetivos

- a) Geral: Fazer um mapeamento da incorporação, por um grupo de mídia brasileiro (Grupo Estado), das práticas do JGD, e identificar se e de que maneira o jornalismo de dados pode contribuir para a criação de valor na organização a partir da análise de casos relevantes, visando identificar suas particularidades em relação a outras formas / métodos / práticas jornalísticas e verificar se essas características agregam valor ou não à organização. Busca-se identificar essas potencialidades e discutir de que forma podem ser usadas para combater a perda do valor do jornalismo para seus diferentes públicos.
- b) Específicos: Verificar quais as discontinuidades em relação ao processo tradicional de apuração, produção e veiculação jornalística que configuram o método conhecido como Jornalismo Guiado por Dados na organização de mídia O Estado de S. Paulo, mais especificamente por meio do núcleo de JGD Estadão Dados. Verificar se e de que forma a adoção das habilidades, práticas e ferramentas características do JGD pode contribuir para a criação de valor na organização jornalística analisada, tanto internamente quanto externamente.

Hipóteses

- a) O JGD se diferencia de outras formas jornalísticas ao usar novas fontes, habilidades e métodos de coleta, apuração e processamento de informações, criando novas formas de visualização, possibilidades de interação com a audiência e produtos de mídia.
- b) O uso do JGD pode colaborar para a criação de valor das organizações jornalísticas recuperarem o valor de três funções do jornalismo elencadas por Picard (2009): acesso exclusivo a fontes, seleção e hierarquização de informações; e distribuição eficiente de

informações. Essa contribuição se daria com a produção de sentido a partir de novas fontes de informação – as bases de dados de interesse público disponíveis online –, a introdução de novas ferramentas, competências e métodos de produção nas redações, maior precisão e credibilidade das notícias, ampliação do engajamento com públicos específicos e a criação de novos produtos e serviços jornalísticos.

Metodologia

A presente pesquisa é qualitativa e tem caráter exploratório, usando o recurso de estudo de caso (DUARTE e BARROS, 2012; EISENHARDT, 1989; POZZEBON e FREITAS, 1998; ROWLEY, 2002; STAKE, 1995; YIN, 1984) como método principal de coleta de informação. Segundo Eisenhardt, os estudos de caso se aplicam particularmente na pesquisa de fenômenos recentes e pouco estudados, em estágios iniciais de pesquisa ou em novas áreas de pesquisa (1989, pp. 548-549). Para Yin, o estudo de caso deve ser aplicado quando “investiga um fenômeno contemporâneo em seu contexto natural, em situações em que as fronteiras entre o contexto e o fenômeno não são claramente evidentes, utilizando múltiplas fontes de evidência. (1984, p.23, tradução do autor)”. O autor defende que o uso desta metodologia se justifica quando o tema é pouco investigado, exigindo estudo aprofundado de poucos casos, que leve à identificação de categorias de observação ou à geração de hipóteses para estudos posteriores. Esse método se distingue também de outras modalidades de pesquisa nas ciências sociais pelo tipo de questões adotadas. Segundo o autor, ela é geralmente usada quando as questões de interesse do estudo referem-se ao como e ao porquê.

Dentre as quatro tipologias de Yin (1984, p.22) – estudo de caso único incorporado, único holístico, múltiplo incorporado e múltiplo holístico –, o presente trabalho segue a última: estudo de caso único holístico, com unidade única de análise e único caso. A diferença entre os casos holísticos e incorporados se dá pela definição da unidade de análise e flexibilidade de condução do estudo. Os holísticos tendem a ter uma estrutura mais flexível em razão dos resultados obtidos, não sendo percebida diretamente nenhuma subanálise no estudo. Os incorporados, por sua vez, comportam mais de uma unidade de análise. Este último tem como distinção também a presença de dados quantitativos.

A opção pelo estudo de um único meio de comunicação se deve a um conjunto de fatores. O primeiro deles é a hodiernidade do fenômeno que se pretende investigar. É exatamente o que ocorre com o Jornalismo Guiado por Dados, tema pouco explorado na

academia e mesmo nas organizações jornalísticas nacionais, como apontam as informações coletadas nesta pesquisa. Mesmo o conceito de jornalismo de dados não é consensual entre os praticantes. O mesmo se aplica, ainda que em menor escala, ao campo escolhido, a economia de mídia. Relativamente novo, especialmente no Brasil, ainda carece de maior volume de literatura, o que demanda a adoção de abordagem mais exploratória. Se aplica também às questões levantadas por este estudo, preocupadas em entender como o JGD se difere de outras práticas jornalísticas e pode contribuir para a criação de valor nas organizações de mídia.

Optou-se, neste trabalho, por utilizar instrumentos múltiplos de coleta de dados, que incluem: análise documental e observação participante como apoio metodológico, entrevistas semi-estruturadas e análise de conteúdo. Foi realizada pesquisa e mapeamento de empresas jornalísticas e seus cibermeios, identificando aquelas que adotam o Jornalismo Guiado por Dados em sua estrutura formal e as que despertam maior interesse levando em conta os critérios de originalidade, representatividade e diversidade. Fez-se uso de pesquisas pela internet, análise documental, participação em fóruns e listas de discussão ligadas ao tema – das quais se destacam TransparênciaHacker⁵, Data Driven Journalism and Open Data in Journalism⁶, além de dois Cursos Online Abertos e Massivos (MOOC, na sigla em inglês) voltados para o tema, oferecidos pela Knight Center for Journalism in the Americas at The University of Texas at Austin: *Data-Driven Journalism: The Basics* e *Introduction to Infographics and Data Visualization*.

De forma concomitante, foram elaboradas as hipóteses que norteariam a partir daí o estudo de caso, levando em conta as similaridades e descontinuidades verificadas durante a revisão bibliográfica e a análise das organizações. O caso escolhido para análise foi o Estádio Dados (<http://estadaodados.com>), núcleo da organização de mídia S.A. O Estado de S. Paulo especializado em reportagens baseadas em estatísticas e no desenvolvimento de projetos especiais de visualização de dados. É apropriado ao método, pois, de acordo com Pozzebon e Freitas (1998), a unidade de análise, nos estudos de caso, pode ser composta por indivíduos, grupos ou organizações, ou ainda por projetos, sistemas ou processos decisórios específicos.

A seleção de um e não múltiplos casos se deve à características peculiares do objeto em questão. A opção pelo Estádio Dados se deve ao fato de não apenas pertencer a uma organização tradicional e de grande porte, com presença em diversos setores, mas

⁵ Grupo com 1.511 membros hospedado pelo Google, formado após o 1º Transparência Hackday, maratona de desenvolvimento de aplicativos e extração de dados públicos, realizado na Casa de Cultura Digital, em São Paulo, nos dias 04 e 05 de Outubro de 2009.

⁶ Lista de discussão com 1571 membros inscritos criada pelo European Journalism Centre (ejc.net) e hospedada pela Open Knowledge Foundation (okfn.org)

principalmente por ser a única com uma equipe multidisciplinar dedicada ao JGD no Brasil e por ter desenvolvido uma ferramenta sofisticada de análise de dados, o Basômetro, o que abre espaço para um estudo de caso mais rico e completo. Não foi encontrada nenhuma outra grande organização de mídia brasileira com essas características, o que por si só torna esse um caso único. Outros dois projetos brasileiros relacionados ao tema também foram pesquisados, mas de forma menos aprofundada, o FolhaSPDados, do grupo de mídia Folha de S. Paulo, e o InfoAmazônia, da organização de mídia sem fins lucrativos O Eco. Inicialmente, havia-se planejado a realização de um estudo de caso múltiplo, mas verificou-se que apenas um exemplo atendia as características buscadas.

A seleção das fontes foi feita de maneira intencional, ou seja, os informantes foram escolhidos com base em seu conhecimento, envolvimento com o objeto e capacidade de contribuir para responder as questões da pesquisa. Segue-se a orientação de Jorge Duarte, que afirma: “Nos estudos qualitativos, são preferíveis poucas fontes, mas de qualidade, a muitas, sem relevo” (DUARTE In DUARTE; BARROS, 2012, p. 68). Foram ouvidos o editor responsável pelo núcleo Estadão Dados, José Roberto Toledo e o repórter Daniel Bramatti, que atua desde o criação do núcleo, além de repórteres do jornal *O Estado de S. Paulo* de maneira anônima para verificar a relação do núcleo com o restante da redação. Além disso, analisou-se o conteúdo e as ferramentas produzidas pelo núcleo, documentos e depoimentos a ele referidos, buscando-se também práticas de mercado que pudessem ser comparadas. Foi realizada ainda entrevista semi-estruturada com Gustavo Faleiros, responsável pelos projetos de JGD InfoAmazônia e FolhaSPDados, com o objetivo de ter um maior entendimento das práticas correntes de outras organizações de mídia no âmbito nacional, de maneira a fazer uma relação com aquelas observadas no estudo de caso, enriquecendo dessa maneira a pesquisa.

Referencial teórico

Esta pesquisa está fundamentada nas teorias da comunicação (HOHLFELDT; MARTINO; FRANÇA, 2001; LOPES, 2004; , especialmente as relacionadas à produção de sentido e à corrente do *newsmaking* (LAGE, 2000; MCCOMBS, 1972, SCHUDSON, 1982, TRAQUINA, 1999, TUCHMAN, 1978; VIZEU, 2002; WILLIANS; CARPINI, 2004), e nas teorias da economia de mídia (ALBARRAN, 1996; ALEXANDER et al., 1993, DOYLE, 2013; LINS, 2005; MOREIRA, 2011, 2014; PICARD, 1989, 2013; RICHERI, 2012; SONNAC; GABSZEWICZ, 2013; TOUSSAINT-DESMOULINS, 2011). Na busca pelas

origens e pela definição do Jornalismo Guiado por Dados foram usados diversos autores tanto do meio acadêmico quanto profissionais da área do jornalismo (BARBOSA; TORRES, 2013, BRADSHAW, 2013, GRAY; CHAMBERS; BOUNEGRU, 2012; GORDON, 2013; HOLOVATY, 2006; HUME, 1998; LORENZ, 2011; MACHADO, 2006; MEYER, 1973, 1999; ROGERS, 2010, 2011, 2012; TOLEDO, 2013; TRÄSEL, 2013), buscando-se dessa forma obter um entendimento holístico do fenômeno e delimitar da melhor forma possível sua abrangência. A análise da criação de valor por meio do JGD foi baseada em conceitos e teorias especialmente do campo da gestão das organizações de mídia (ALBARRAN, 2004, 2010; DOYLE, 2013; PICARD, 2009, 2013; RICHERI, 2012) e do Marketing (CARNEIRO; CAVALCANTI; SILVA, 1997; FAUSTINO, 2010; MAGRETTA, 2013; PORTER, 1985, 1990, 1996). Para a contextualização das transformações tecnológicas e seus efeitos sobre a comunicação e o jornalismo foram relevantes tanto dos campos da Teoria da Informação (HILBERT; LÓPEZ, 2011), Economia da atenção (GOLDHABER, 2006, 2012; SIMON, 1972; LANHAM, 1994), quanto do campo das ciências sociais e da comunicação (CASTELLS, 1999, 2003; HARPER, 1998; HUNTER; WASSENHOVE, 2010; JENKINS, 2008; JENKINS; FORD; GREEN, 2013; MCADAMS, 2012; MIELNICZUK, 2003; NIELSEN, 2012; SILVA, 2011).

A estrutura da dissertação

Esta dissertação está dividida em três capítulos, além da Introdução e da Conclusão. No primeiro capítulo, intitulado *Jornalismo Guiado por Dados como forma contemporânea de produção de sentido*, busca-se localizar o JGD como um fenômeno contemporâneo e enquadrá-lo como uma das práticas jornalísticas usadas com o objetivo de se produzir sentido. Para isso, detalha-se o contexto histórico que levou ao surgimento do JGD, ressaltando suas raízes na Reportagem com o Auxílio de Computador e o Jornalismo de Precisão e os desenvolvimentos que o levaram a se tornar uma prática com características individuais, além do estágio de sua inserção nas redações nacionais e estrangeiras. São mostradas as diferentes tentativas de delimitar um conceito que defina o JGD, assim como as dificuldades para se alcançar esse objetivo, e definidos os limites escolhidos pelo autor e usados como baliza para esta pesquisa. Por fim, o JGD é analisado como prática produtora de sentido, destacando-se os principais elementos que podem influenciar seus praticantes.

No segundo capítulo, intitulado *A criação de valor pelo Jornalismo Guiado por Dados*, analisa-se o JGD sob a perspectiva econômica, mais especificamente a criação de

valor nas organizações. Inicialmente, é feita uma introdução sobre os estudos da Economia de Mídia e sua relação com o campo da Comunicação, bem como um panorama das indústrias de mídia no Brasil. Discute-se em seguida o conceito de valor e sua aplicação no jornalismo e no JGD. A partir da análise de exemplos brasileiros e estrangeiros e de textos de diferentes pesquisadores e de adeptos do JGD, foram identificadas formas por meio das quais o JGD pode contribuir para a criação de valor nas organizações.

No terceiro capítulo, *Estudo de Caso – Estadão Dados*, é feita a análise mais detalhada de um caso específico de inserção do JGD em uma organização brasileira, o mais completo encontrado no Brasil. O objetivo é ressaltar as diferenças em relação ao processo tradicional de produção jornalística e verificar se as possíveis formas de contribuição de valor identificadas anteriormente se apresentam no exemplo em questão. Para isso é traçado um perfil da organização, o Grupo Estado, e do núcleo de JGD, o Estadão Dados, além de seu histórico de criação. Detalha-se o processo produtivo, dividido entre as etapas de criação de valor, destacando-se as particularidades dos métodos e dos conteúdos produzidos pelo núcleo em relação ao do restante da redação. Para apoiar a diferenciação foi feita uma análise mais aprofundada da ferramenta Basômetro, desenvolvida pelo Estadão Dados, comparando as reportagens produzidas a partir do seu uso com as matérias tradicionais sobre o mesmo tema: a base aliada do governo federal no Congresso.

1 JORNALISMO GUIADO POR DADOS COMO FORMA CONTEMPORÂNEA DE PRODUÇÃO DE SENTIDO

1.1 Contexto histórico

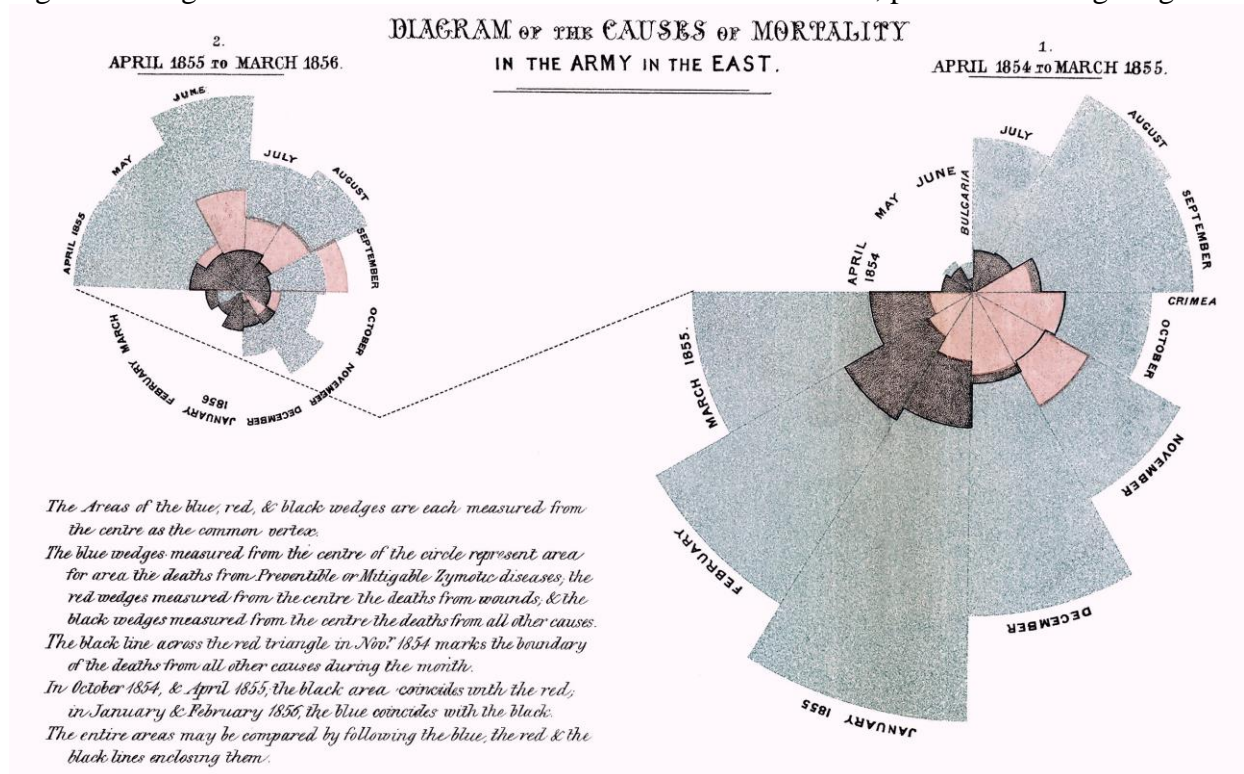
As bases de dados estão presentes no jornalismo há quase tanto tempo quanto a própria atividade. O jornal britânico *The Guardian* já exibia em sua primeira edição, publicada em 5 de Maio de 1821, uma tabela mostrando a lista das escolas das cidades de Manchester e Salford, na Inglaterra, com o número de alunos e o gasto médio anual do governo com cada uma, revelando pela primeira vez quantos recebiam educação gratuita⁷ do estado. Um exemplo de reportagem que revela a notícia a partir da investigação e análise de bases de dados. Mesmo quando se trata das sofisticadas representações gráficas de informações de interesse público há registros que se assemelham de forma assombrosa aos atuais e datam do século XIX, como a visualização das mortes do exército britânico na Guerra da Criméia (1853-1856), elaborada pela enfermeira e estatística britânica Florence Nightingale em 1858, que poderia facilmente estar em um site de notícias contemporâneo (Figura 1). Ou o gráfico criado pelo engenheiro e estatístico francês Charles Joseph Minard exibindo as perdas do exército napoleônico na invasão da Rússia, entre 1812 e 1813⁸.

Até recentemente, no entanto, ter acesso a bases de dados de interesse público exigia um esforço de reportagem que envolvia o convencimento de autoridades e instituições a liberar documentos, a coleta física desses documentos, normalmente impressos, e sua análise pelos jornalistas, o que limitava sobremaneira a quantidade de arquivos que poderiam ser analisados e o tamanho de cada arquivo. Isso porque, quanto maior o documento, mais pessoas ou mais tempo era necessário para analisá-lo, o que implicava em mais gastos e, muitas vezes, tornava a atividade inviável. O mesmo vale para a visualização desses dados, que exigia um trabalho praticamente artesanal e também demorado.

⁷ Ver em: <<http://www.theguardian.com/news/datablog/2011/sep/26/data-journalism-guardian#zoomed-picture/>>

⁸ Ver em: <<http://upload.wikimedia.org/wikipedia/commons/2/29/Minard.png/>>

Figura 1- Diagrama das causas de mortalidade do exército do Oeste, por Florence Nightingale



Fonte: British Monarchy/Wikimedia Commons. Disponível em: <<http://commons.wikimedia.org/wiki/File:Nightingale-mortality.jpg>>

Esse cenário começou a mudar na década de 1950 quando jornalistas usaram pela primeira vez um computador para processar dados que seriam usados em uma reportagem. A rede de TV norte-americana CBS lançou mão de um computador Remington Rand UNIVAC para projetar o resultado da disputa eleitoral pela presidência dos Estados Unidos em 1952, entre Dwight Eisenhower e Adlai Stevenson, dando origem à Reportagem com Auxílio de Computador (RAC), um conceito abrangente que abarca qualquer prática que envolva o uso de computadores para auxiliar o processo de obtenção de informações jornalísticas (COX, 2000). Na ocasião, toda a mídia esperava uma disputa acirrada entre os dois candidatos, mas as projeções do computador – feitas com base nos primeiros dados de apuração de votos e em fórmulas desenvolvidas especialmente para extrapolar esses resultados parciais – indicavam uma vitória esmagadora de Eisenhower. A emissora não acreditou nas previsões e adiou por horas sua divulgação, até que as urnas confirmaram o que o computador já havia adiantado. Desde então, as projeções por computador se tornaram prática obrigatória na mídia em todas as eleições americanas.

Um dos pioneiros do RAC no mundo foi jornalista norte-americano Philip Meyer, que, a partir do trabalho com computadores, desenvolveu novas técnicas de reportagem que aplicavam os métodos científicos na produção jornalística, em busca de maior veracidade.

Suas idéias foram compiladas no livro *Precision Journalism: a reporter's introduction to social science methods*, lançado em 1973, e influenciam até hoje os praticantes do RAC, que tem como expoentes o National Institute for Computer-Assisted Reporting⁹ (NICAR), criado em 1989, e, no Brasil, a Associação Brasileira de Jornalismo Investigativo¹⁰ (Abraji), criada em 2002, e também os adeptos do JGD. Logo na primeira página de sua clássica obra, Meyer afirma, com uma clarividência impressionante, décadas antes da popularização da internet e das mídias digitais:

The world has become so complicated, the growth of available information so explosive, that the journalist needs to be a filter as well as a transmitter, an organizer and interpreter as well as one who gathers and delivers facts. In addition to knowing how to get information into print or on the air, he or she also must know how to get it into the receiver's head. In short, a journalist has to be a database manager, a data processor, and a data analyst (MEYER, 1973, p.1).

Nos anos 1980, os computadores já estavam presentes na maior parte das redações norte-americanas, o que só veio a ocorrer na década seguinte no Brasil. Inicialmente, os computadores apenas substituíam as máquinas de escrever como editores de texto. Posteriormente, começaram a ser usadas para integrar os textos em sistemas online. Somente na virada do século, com a popularização da internet em escala global, os computadores se tornaram também fonte de informação regular para os jornalistas.

Há pouca informação sobre a introdução e desenvolvimento do RAC no país, mas Träsel (in: GRAY; CHAMBERS; BOUNEGRU, 2014) faz um esclarecedor resumo histórico identificando suas origens na cobertura política e a influência de associações norte-americanas na formação dos profissionais nacionais. A prática do RAC começa com a série de reportagens assinadas por Mário Rosa que denunciou o escândalo do desvio de verbas da Legião Brasileira de Assistência (LBA) pela então primeira-dama Rosane Collor, mulher do presidente Fernando Collor de Melo. Foi iniciada com a reportagem *Rosane sai sob suspeita da LBA*, publicada em 26 de Agosto de 1991 no *Jornal do Brasil*¹¹. Era a primeira vez que um repórter usava dados do Sistema Integrado de Administração Financeira do Governo Federal (Siafi), criado em 1987, que registra todos os pagamentos feitos pelo governo federal. O acesso foi feito com senha emprestada pelo então senador Eduardo Suplicy (PT), pois apenas os parlamentares tinham direito a entrar no sistema. “O jornalismo ganhava uma nova e

⁹ Ver em: <www.ire.org/nicar/>

¹⁰ Ver em: <www.abraji.org.br/>

¹¹ A edição digitalizada da edição está disponível em: <<http://goo.gl/syjecD>>

importante fonte de informação, mais técnica, quase científica. Estavam superados os métodos mais arcaicos de apuração, que envolviam, eventualmente, o enfrentamento com jagunços”, afirmou o jornalista Lúcio Vaz, que descreve o caso no livro *A ética da malandragem* (2005, p.171).

Nos anos seguintes, conta Träsel, a prática ganhou corpo especialmente no jornal Folha de S. Paulo, introduzida por tutores da National Institute for Computer-Assisted Reporting dos Estados Unidos, uma subdivisão da associação Investigative Reporters and Editors (IRE/NICAR). Capitaneado por jornalistas como Fernando Rodrigues e José Roberto Toledo, o RAC foi disseminado na redação e passou a integrar o currículo do programa de *trainees* do jornal. Em 1998, Fernando Rodrigues começou a construir o banco de dados Políticos do Brasil¹² e foi reconhecido com o Prêmio Esso de Melhor Contribuição à Imprensa pelo arquivo de declarações de bens de políticos brasileiros Controle Público, em 2002, e pelo livro *Políticos do Brasil*, em 2006.

Em 2002, José Roberto de Toledo se tornou um dos sócios-fundadores e vice-presidente da Associação Brasileira de Jornalismo Investigativo (Abraji). A influência norte-americana aparece mais uma vez por meio do Centro Knight para o Jornalismo nas Américas, que promoveu um seminário em dezembro daquele ano que serviu de inspiração para a formação da Abraji. Os principais palestrantes foram Brant Houston, autor de um manual de RAC e então diretor do IRE, e Pedro Armendares, da organização mexicana Periodistas de Investigación, um dos tutores dos cursos de RAC organizados pela Folha de S. Paulo. A Abraji se tornou a maior divulgadora do RAC no país e forte defensora da liberdade de acesso à informação, como uma das entidades integrantes do Fórum de Direito de Acesso a Informações Públicas, criado em 2003, e através de cursos e palestras – dois fatores fundamentais para a emergência do jornalismo guiado por dados ao longo da década de 2000. Desde então, já treinou mais de 4 mil jornalistas.

Outro pioneiro no uso do Siafi é o jornalista Ascânio Seleme, atual diretor de redação do jornal O Globo, que trabalhou no início dos anos 1990 em colaboração com o analista econômico Gil Castello Branco, que criaria em 2005 a Organização Não-Governamental Contas Abertas. A ONG, dedicada à transparência dos gastos públicos, acompanha a execução orçamentária e financeira da União, através de monitoramento do Siafi, e promove o treinamento de jornalistas para fiscalizar gastos públicos. Junto com a ONG Transparência Brasil, foi um dos principais vetores de divulgação de bases de dados de interesse público no

¹² Ver em: <<http://noticias.uol.com.br/politica/politicos-brasil/>>

país, permitindo a repórteres realizar reportagens quando o acesso às informações do Estado dependia de gestão caso-a-caso junto a órgãos do governo e às redações não investiam neste tipo de recurso. A Transparência Brasil foi criada em 2000 com o objetivo de construir e manter bases de dados sobre financiamento eleitoral, histórico de vida pública e processos sofridos por parlamentares em nível municipal, estadual e federal, notícias sobre corrupção publicadas nos principais jornais brasileiros e sobre o desempenho dos juízes membros do Supremo Tribunal Federal. Ambas receberiam também o Prêmio Esso de Melhor Contribuição à Imprensa. Em 2006, a Transparência Brasil, e no ano seguinte, a Contas Abertas.

Na esteira do avanço da web no Brasil e no mundo, surgiram também outras mobilizações em defesa da divulgação de informações de interesse público por governos e instituições privadas, alinhadas com as características da internet – aberta, não-hierárquica, descentralizada. Como resultado, governos, empresas e instituições passaram a produzir e lançar uma torrente de bases de dados na internet, num volume que cresce exponencialmente a cada ano. Ainda que algumas das bases estivessem lançadas, a difusão desmesurada de dados só começou a ocorrer a partir dos anos 1990 e, de forma mais global e sistemática, na década seguinte, com a consolidação dos movimentos de dados abertos, transparência governamental e da pressão social pela liberação de informações de maneira geral. Esse processo avançou mais rapidamente nas democracias ocidentais economicamente desenvolvidas, que já tinham uma cultura de prestação de contas e transparência, como os Estados Unidos, e se espalhou com maior atraso para o Brasil – a Lei de Acesso à Informação (LAI) brasileira, ou Lei nº 12.527, foi promulgada em 18 de novembro de 2011, já o “Freedom of Information Act” (FOI), legislação similar norte-americana, para efeito de comparação, entrou em vigor em 4 de Julho de 1966.

A profusão de bases de dados nas redes, muitas vezes ininteligíveis para o cidadão comum, mesmo os ávidos por informação, criou uma demanda por pela análise dessas informações de maneira que fizessem sentido e pudessem ser usadas para efetivamente cobrar governos, empresas e outras instituições, embasar ações políticas ou, simplesmente, permitir a tomada de decisões, sejam de cunho social ou cotidiano, com maior conhecimento. Isso acontece de forma concomitante com a disponibilização de softwares e ferramentas de análise e visualização de dados, muitas vezes gratuitas e colaborativas, dentro do movimento de software livre. Com isso, se tornou mais fácil produzir novos conteúdos e novos sentidos a partir das bases de dados, tanto fora como nas redações. Segundo o inventor da world wide web, Tim Berners-Lee, o jornalismo não só pode como deve assumir o papel de interpretar os

dados públicos: “The responsibility needs to be with the press (...) Journalists need to be data-savvy”¹³.

O Jornalismo Guiado por Dados (JGD) surge então como a prática que pretende transformar essa massa de dados em informação que seja relevante e faça sentido para a sociedade. Os exemplos jornalísticos pioneiros datam de meados dos anos 2000, como o *Chicagocrime.org*, lançado em 2005 pelo jornal *Chicago Tribune* para localizar os crimes cometidos na cidade a partir de dados da polícia municipal. A ferramenta combinava dados de criminalidade divulgados pela Polícia de Chicago com mapas online disponíveis gratuitamente na internet, permitindo pela primeira vez que qualquer pessoa verificasse quais áreas da cidade tinham maior incidência de quais crimes e, com isso, pudessem cobrar maior presença da polícia na repressão aos criminosos. Seus efeitos sobre a produção jornalística foram estudados e são discutidos por Parasie e Dagiral (2013).

Outras iniciativas foram criadas nos anos seguintes tanto por indivíduos como por organizações de mídia vários países, com a introdução de programadores nas redações e nas quais jornalistas buscavam maior conhecimento de programação e de análise e visualização de dados. Começavam a surgir os termos *data-driven journalism* (jornalismo guiado por dados) e *data journalism* (jornalismo de dados) para definir a nova prática que usava conhecimento de computação e linguagens de programação, como JavaScript, Python e Ruby, para destrinchar bases de dados online e produzir histórias com auxílio de ferramentas de visualização, como *OpenStreetMap*, *GoogleMaps* e *ManyEyes*, criando narrativas envolventes e interativas. Em 2006, Adrian Holovaty (2006) publicou um artigo que é considerado o primeiro a identificar as mudanças pelas quais o jornalismo estava passando e recomendando que os sites dos jornais deixassem de ser totalmente concentrados em histórias para também oferecer os dados por trás das histórias de maneira estruturada. Em 24 de agosto de 2010, foi realizado o primeiro congresso dedicado ao JGD, em Amsterdã, Holanda, pelo Centro Europeu de Jornalismo. Os 50 jornalistas ali reunidos começavam a oficializar o conceito de jornalismo guiado por dados.

A notoriedade, no entanto, veio com a publicação de 92.000 relatórios do exército norte-americano sobre a Guerra do Afeganistão, naquele mesmo ano. O escândalo – que reuniu uma organização independente, a *Wikileaks*, e três conglomerados de mídia tradicionais: o norte-americano *The New York Times*, o britânico *The Guardian* e o alemão

¹³ “A responsabilidade deve ser da imprensa (...) Jornalistas precisam ser aficionados por dados.” (tradução do autor), disse Berners-Lee em evento de divulgação de novas bases de dados de gastos públicos do governo britânico em 2010, segundo registrou o jornal *The Guardian*. Ver em: <<http://www.theguardian.com/media/2010/nov/22/data-analysis-tim-berners-lee/>>

Der Spiegel – chamou a atenção da opinião pública para uma prática até então marginal nas redações: a produção de reportagens com o uso de imensos bancos de dados e formas inovadoras de visualização e acesso por parte dos leitores. A enorme repercussão da cobertura chamou a atenção para o jornalismo guiado por dados, incentivando a prática em vários países, além de investimentos por partes das empresas de comunicação na contratação de profissionais especializados e montagem de equipes específicas para a produção do JGD. Foi o que fizeram não só as três empresas, mas também outras publicações, como *Chicago Tribune*, *Financial Times*, *El País*, *O Estado de S. Paulo* e *Gazeta do Povo*, para citar alguns exemplos. O primeiro manual dedicado ao tema foi publicado em 2012, o *Data Journalism Handbook*. Apenas nos últimos três anos a prática ganhou força e notoriedade na classe jornalística e na opinião pública, por meio de grandes coberturas baseadas em dados, como os escândalos Wikileaks. No Brasil, a consagração viria com a série de reportagens “Diários Secretos”, publicada na *Gazeta do Povo*, do Paraná, que recebeu o Prêmio Esso de Jornalismo, principal premiação do jornalismo brasileiro. Os repórteres criaram uma base de dados com as nomeações realizadas pela Assembléia Legislativa paranaense entre 2006 e 2010 a partir dos diários oficiais e, por meio do cruzamento de informações usando o software Excel descobriram casos de funcionários fantasmas e nepotismo.

O JGD está, portanto, intimamente relacionado a um contexto de massificação da internet e das novas mídias digitais e de movimentos de transparência e divulgação de informações. Pressupõe a existência de uma massa de bases de dados públicas disponíveis na rede e de ferramentas acessíveis de visualização, análise, interpretação e divulgação desses dados. Pode ser considerada, portanto, uma prática jornalística característica da contemporaneidade, pois não existiria em outra época, ou em outras condições. É possivelmente uma das práticas, dentro do jornalismo, que pode aproveitar de maneira plena o potencial das mídias digitais.

Lorenz, Kaiser-Bril e McGhee (2011) vão além e acreditam que a prática pode transformar completamente o modelo de negócios das organizações de mídia, transformando-as em grandes “pólos de dados confiáveis” e permitindo o fornecimento de novos produtos e serviços de informação, que sirvam como fonte de receita alternativa aos anúncios, como aplicativos pagos baseados nas bases de dados de interesse público coletada, processadas e arquivadas pelas empresas. Um exemplo de serviço cujo valor pode ser facilmente identificado e quantificado é o aplicativo *Is It Better to Buy or Rent*¹⁴, uma sofisticada

¹⁴ Ver em: <http://www.nytimes.com/interactive/business/buy-rent-calculator.html?_r=0>

ferramenta gratuita desenvolvida pelo The New York Times que informa se é mais vantajoso alugar ou comprar um imóvel com base em uma série de cálculos que consideram não só os custos e impostos associados à operação como o custo de oportunidade do investimento.

Ao se voltar a para a interpretação daquilo que forma o suporte da comunicação e da cultura na atualidade, as bases de dados, se apresenta como uma ferramenta de produção de sentido por excelência da era dos computadores, servindo como contraponto à obsolescência dos meios de comunicação de massa tradicionais e se apresentando como uma das saídas para a crise existencial do jornalismo no século XXI.

1.2 O JGD nas redações

Em estudo pioneiro sobre a inserção do Jornalismo Guiado por Dados nas redações, Zanchelli e Crucianelli (2013) identificaram quatro fatores determinantes para a criação de equipes de JGD bem sucedidas, a partir da análise de casos em seis organizações de mídia: quatro nos Estados Unidos (*Washington Post*, *New York Times*, *USA Today* e *Chicago Tribune*), uma no Reino Unido (*BBC*) e uma no Brasil (*Estadão*). Os fatores são: posicionar a equipe de dados próxima de quem faz as notícias; encorajar repórteres e desenvolvedores a trabalhar em conjunto para ter idéias de reportagens guiadas por dados; recrutar repórteres e desenvolvedores que façam a ponte entre as habilidades existentes e as necessárias; produzir reportagens que mostrem o que os dados significam e porque a audiência deve se importar com eles (tradução do autor)¹⁵.

Os pesquisadores observaram que a inserção do JGD nas redações acontece geralmente por iniciativa de repórteres e se desenvolve – com o aumento das equipes e do investimento da empresa na área – na medida em que obtém resultados concretos, notadamente o aumento da audiência e de prestígio por meio de reportagens de grande apelo. Deixam claro também a necessidade de novas habilidades e o caráter multidisciplinar – jornalismo e programação – e coletivo – jornalistas, programadores, designers – da prática. Esses são fatores intrínsecos ao JGD que o diferenciam de outras práticas jornalísticas que fazem uso de bases de dados como o RAC e o Jornalismo de Precisão.

¹⁵ “Locating the data-journalism team close to the news desk; encouraging reporters and developers to work together to come up with ideas for data-driven stories; recruiting reporters and developers who bridge the skills gap; producing stories that show what data mean and why the audience should care” (ZANCHELLI; CRUCIANELLI, 2013).

É natural que qualquer nova prática seja inserida de forma paulatina nas redações especialmente se envolver uma mudança de cultura, a adoção de novas ferramentas, a transformação dos produtos jornalístico e a incorporação de profissionais de outras áreas no cotidiano de produção de notícias, como é o caso do JGD. A partir dos achados de Zanchelli e Crucianelli, foi elaborado o quadro abaixo, condensando o perfil das seis equipes de dados.

Nota-se que em todos os casos, as equipes de JGD estão integradas à redação, mesmo quando o trabalho não é feito de forma conjunta com outros repórteres, como no caso do jornal *The Guardian*, na qual os jornalistas guiados por dados controlam todo o processo produtivo e trabalham apenas com desenvolvedores e designers, mas não com repórteres de outras áreas. A dinâmica de produção é bem diferente em cada caso e envolve desde a produção isolada à integração total, como é o caso da BBC, em que não há uma divisão formal da equipe de dados e todos produzem conteúdo JGD. A formação das equipes também é bem diversa e envolve desde grupos só de jornalistas especializados, como no jornal *The New York Times*, como a integração total a ponto de se tornar um serviço, como é o caso do *Washington Post*, em que os desenvolvedores que atuam na produção do JGD estão ligados ao setor de Tecnologia da Informação.

Quadro 1 - Perfil das equipes de Jornalismo Guiado por Dados nas redações de cinco veículos

Veículos	Distribuição	Estrutura da equipe	Dinâmica de trabalho na redação	Área específica no site	Formato longo ou curto ¹⁶	Dentro da redação
<i>BBC</i>	Online	8 jornalistas, 5 desenvolvedores, 5 designers	Isolado e colaborativo	Sim	Ambos	Sim
<i>New York Times</i>	Online e impresso	20 a 40 jornalistas com domínio de programação e design	Isolado	Sim	Ambos	Sim
<i>The Guardian</i>	Online e impresso	3 jornalistas com domínio de programação	Colaborativo	Sim	Ambos	Sim
<i>Chicago Tribune</i>	Online e impresso	n/d	Colaborativo	Sim	Ambos	Sim
<i>USA Today</i>	Online e impresso	5 jornalistas e 1 editor com domínio de programação, design e estatística	Colaborativo	Não	Ambos	Sim
<i>Washington Post</i>	Online e impresso	12 a 13 desenvolvedores na redação	Colaborativo	Não	Ambos	Sim

Fonte: ZANCHELLI e CRUCIANELLI, 2013.

¹⁶ Do original “long-form” e “short”, que se referem tanto à extensão do textos das reportagens (reportagens de fôlego e notas curtas), quanto à complexidade das ferramentas e aplicativos desenvolvidos pelas equipes de dados

Tomando como base o modelo de Zanchelli e Crucianelli, foi elaborado o seguinte quadro com o perfil das equipes de dados identificadas hoje no Brasil, além do Estadão Dados, que já foi contemplado no estudo, mas teve os dados atualizados. Importante ressaltar que, com exceção da *Gazeta do Povo*, não foram encontrados exemplos de estruturas formais de JGD nas redações brasileiras, mas sim projetos temporário (FolhaSPDados) ou independente (InfoAmazônia).

Observa-se também nos exemplos nacionais encontrados que os jornalistas de dados ficam integrados à redação. Em todos os casos foram aproveitados profissionais que já atuam nas organizações e buscaram conhecimento específico por iniciativa própria ou durante o uso diário de bases de dados com o restante da equipe. O número de jornalistas é significativamente menor que nos casos internacionais, o que demonstra menor capacidade financeira das empresas e um estágio menos avançado de inserção do JGD.

Quadro 2 - Perfil das equipes de projetos de JGD no Brasil

Veículos/ Projeto	Distribuição	Estrutura da equipe	Relação com outros jornalistas	Área no site	Formato longo ou curto¹⁷	Status	Dentro da redação
Folha de S. Paulo/ FolhaSPDados	Online	2 jornalistas	Isolado	Sim	Ambos	Temporário	Sim
Estadão Dados	Online e impresso	3 jornalistas com domínio de RAC e Jornalismo de Precisão, 1 jornalista especialista em estatísticas e 1 desenvolvedor	Isolado e colaborativo	Sim	Ambos	Fixo	Sim
Gazeta do Povo	Online e impresso	1 designer, 1 desenvolvedor	Colaborativo	Sim	Ambos	Fixo	Sim
O Eco/ EcoLab / InfoAmazônia	Online e impresso	1 jornalista (coordenador), 2 desenvolvedores, 2 designers	Colaborativo	Sim	Ambos	Fixo	Sim

Fonte: elaboração do autor

Pode-se destacar novamente a influência de entidades estrangeiras na disseminação do uso das bases de dados nas redações brasileiras. A instituição norte-americana International Center for Journalists (ICFJ) foi uma das financiadoras do projeto InfoAmazônia, ao conceder uma bolsa ao jornalista Gustavo Faleiros, de O Eco, dentro de seu programa do programa Knight International Journalism Fellowships. A iniciativa também foi custeada pela Climate

¹⁷ Do original “long-form” e “short”, que se referem tanto à extensão do textos das reportagens (reportagens de fôlego e notas curtas), quanto à complexidade das ferramentas e aplicativos desenvolvidos pelas equipes de dados

and Development Knowledge Network (CDKN), sediada no Reino Unido, e por um doador anônimo dos Estados Unidos. Posteriormente, Faleiros foi convidado a implantar o projeto FolhaSPDados, uma iniciativa acertada anteriormente entre o ICFJ e a Folha de S. Paulo.

O InfoAmazônia recebeu investimento de aproximadamente US\$ 150 mil, proveniente das três fontes mencionadas, dos quais cerca de 70% foi aplicado no desenvolvimento de uma plataforma de produção, visualização e distribuição de mapas e notícias (programação e design) e o restante, na coordenação e mão-de-obra jornalística (FALEIROS, 2013). A maior parte do conteúdo noticioso foi herdada de um projeto anterior de O Eco, com custo de US\$ 200 mil, voltado para a produção de reportagens sobre a Amazônia. Aproveitando o conhecimento, a equipe e os produtos desenvolvidos, Faleiros criou posteriormente o EcoLab, laboratório de desenvolvimento de plataformas e aplicativos jornalísticos, com foco em temas relacionados ao meio ambiente, mapeamento de dados e jornalismo colaborativo.

O FolhaSPDados foi criado em 2012, com o objetivo de divulgar informações relevantes sobre a cidade de São Paulo contidas em bases de dados públicas às vésperas da eleição municipal daquele ano, usando ferramentas de JGD. Era parte das atribuições de Faleiros atreladas a sua bolsa do ICFJ. A ideia inicial era que o projeto servisse como ponto de partida para a disseminação das ferramentas de análise e visualização de dados para a redação e fosse mantido futuramente pelos próprios jornalistas. No entanto, devido às limitações de pessoal e a falta de tempo dos jornalistas, envolvidos com a cobertura tradicional das eleições, esse objetivo não foi alcançado (FALEIROS, 2013).

1.3 Em busca de um conceito de Jornalismo Guiado por Dados

Por ser uma prática relativamente nova, ainda não há um conceito consensual e bem delimitado para o Jornalismo Guiado por Dados (BARBOSA e TORRES, 2012; LORENZ, 2011; TRÄSEL, 2013a, 2013b), que por vezes é chamado também de jornalismo de dados (GRAY, 2012a; GRAY et al., 2013, TOLEDO, 2013) ou jornalismo computacional (COHEN; TURNER; HAMILTON, 2011; DIAKOPOULOS, 2012; LIMA JR., 2011; PEARSON), o que evidencia essa indefinição. Desde o surgimento do termo no fim da década passada, vários autores encontraram diferentes formas de definir este fenômeno, mas o conceito ainda carece de limites, dada a abrangência de seu uso nos processos de produção jornalística, assim como

os pontos de contato e similaridades com a RAC, além da incipiente literatura acadêmica e da carência de instituições que a legitimem.

Apesar das similaridades, diferenciá-lo do Jornalismo de Precisão é tarefa mais fácil. Ao desenvolver a prática, Phillippe Meyer (1973) estava preocupado em introduzir os métodos científicos no jornalismo como forma de ampliar sua legitimidade na busca pela verdade factual, aproximando-o das ciências sociais. O computador e as bases de dados eram apenas instrumentos que podem ou não ser usados para se alcançar esse objetivo. A invenção de Meyer surge mais de sua busca por exatidão do que de um novo contexto de fluxos informativos que exija a criação de práticas e ferramentas que sejam capazes de dar conta de seu volume e complexidade, como é o caso do JGD.

O contexto histórico também pode fornecer indícios importantes para sua definição. A prática surge em um cenário de superabundância de informações de interesse público online, em volume e formato que estão além da compreensão dos cidadãos comuns, e atende a uma demanda pela seleção e produção de sentido a partir desses dados. Jonathan Gray (2012) tem uma definição bastante abrangente que ajuda a entender quão ampla é a aplicação do jornalismo de dados e começa a delinear suas particularidades, ainda que não de forma suficiente para delimitar o conceito:

Broadly speaking, “data journalism” is a fairly recent term that is used to describe a set of practices that use data to improve the news. These range from using databases and analytical tools to write better stories and do better investigations, to publishing relevant datasets alongside stories, and using datasets to deliver interactive data visualizations or news apps. (GRAY, 2012a)¹⁸

Mirko Lorenz (2010) define o jornalismo de dados como um novo processo produtivo jornalístico, que tem as bases de dados como fio condutor: “Data-driven journalism is a workflow where data is the basis for analysis, visualization and storytelling¹⁹”. Já Bradshaw (in: GRAY; CHAMBERS; BOUNEGRU, 2012²⁰) indica que entre os diferenciais do jornalismo de dados está sua capacidade de contar histórias envolventes que abarquem a escala e a abrangência da informação digital, transformar grandes números em infográficos

¹⁸ “De maneira geral, ‘jornalismo de dados’ é um termo relativamente recente usado para descrever um conjunto de práticas que usam dados para incrementar as notícias. Elas vão desde usar bases de dados e ferramentas de análise para escrever melhores matérias e investigar melhor a publicar conjuntos de dados relevantes junto com as matérias e usar dados para produzir visualizações e aplicativos (tradução do autor)”

¹⁹ “Jornalismo guiado por dados é um fluxo de trabalho em que os dados são a base para a análise, visualização e a narração de histórias” (tradução do autor)

²⁰ Livros eletrônicos e e-books, como este, não permitem referenciar a página da qual foram retiradas as citações, pois a paginação varia de acordo com a formatação das letras.

atraentes e ajudar a relacionar a notícias com os indivíduos, com o uso de ferramentas interativas. José Roberto de Toledo (2013) pensa de maneira similar ao afirmar que “O jornalismo de dados é uma adaptação do RAC aos petabytes de dados que o mundo produz diariamente” (TOLEDO, 2013) e que são exatamente essas características – volume e método (ferramentas e processos) – que o diferenciam do RAC.

Rogers (2012) compara o impacto do fenômeno do jornalismo de dados na comunicação com o do fenômeno punk na música, pois “qualquer um pode fazer”. Sua afirmação tem um tom de provocação e claramente busca “evangelizar” pessoas para atrair novos integrantes para a prática do jornalismo de dados, mas tem fundamento ao ressaltar a democratização dos dados e do conhecimento e ferramentas para processá-los. Esta visão é desafiada por Gray (2012a), para quem o otimismo deve ser substituído pela análise crítica dos dados, o estudo das ferramentas de análise e o aprofundamento nos assuntos que são abordados, da mesma forma que ocorre em outras práticas jornalísticas. A visão de Rogers também contrasta com a de Toledo, para quem o JGD é um trabalho coletivo, por exigir uma equipe multidisciplinar em sua execução.

Barbosa e Torres (2013) dão um importante passo na conceituação do Jornalismo Guiado por Dados ao considerá-lo como uma nova forma narrativa, que seria uma extensão do Jornalismo Digital em Bases de Dados (JDBD)²¹. Dessa forma, o colocam como parte de uma transformação mais ampla na atividade jornalística, operada a partir da popularização dos meios digitais e da web e da ascensão das bases de dados como forma cultural simbólica característica da era dos computadores, segundo o postulado por Lev Manovich (2001). Diferentemente da narrativa dos meios de comunicação de massa tradicionais (impressos, televisão, rádio e cinema), linear e pré-determinada, na qual o ouvinte, o leitor ou o telespectador acompanha a narração sem interferir na lógica interna das ações, nos meios digitais o fluxo narrativo, mais do que incorporar, vai depender diretamente da intervenção do usuário (MACHADO, 2006). É de se esperar, portanto, que a emergência de uma prática jornalística inteiramente ligada às bases de dados, como o JGD, tenha impacto sobre as convenções narrativas existentes. Afirmam Barbosa e Torres:

Ao compreender o Jornalismo Guiado por Dados como um dos aspectos do Paradigma JDBD, consideramos os produtos e os formatos dele derivados como

²¹ Definido como “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” (BARBOSA, 2007)

configuradores de modos novos de narrar, assegurados pela hipernarrativa em base de dados. (BARBOSA; TORRES, 2013, p.5)

Após análise de exemplos empíricos, Barbosa e Torres (2013, p.8) identificaram quatro tipologias para ordenar e distinguir as iniciativas do Jornalismo Guiado por Dados (Quadro 3). Dessas tipologias, apenas uma, os SRN, não foram identificadas pelo autor desta pesquisa como características do Jornalismo Guiado por Dados, uma vez que, sua produção não exige a aplicação de trabalho e conhecimento jornalístico em sua produção, seja multidisciplinar ou individual, ou a adoção de técnicas ou ferramentas de análise de dados, sendo apenas uma forma de visualização de bases de dados.

Quadro 3 - Tipologias do JGD

Tipos	Definição	Exemplos
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.	“SOPA Opera – Where Do Your Members of Congress Stand on SOPA and PIPA?”, <i>ProPublica</i> ; “Seven Billion People”, <i>BBC</i>
Dados	Nova denominação para editoriais/seções em alguns cibermeios nas quais são apresentados conteúdos produzidos a partir de dados disponibilizados pelos governos de países, estados, cidades, entre outros.	Data, <i>The Guardian</i> ; Data Desk, <i>Los Angeles Times</i> ; e Dados, <i>Gazeta do Povo</i>
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 13), disponibilizam dados coletados e estruturados por suas equipes.	Data.gov.br; Wikileaks
Sistemas de Recomendação de Notícias (SRN)	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.	“Nube de Noticias”, <i>La Nación</i>

Fonte: Barbosa e Torres, 2013

Talvez uma das maiores particularidades do JGD seja o uso da programação para coleta, análise ou visualização dos dados, habilidade que se mostra presente seja por meio dos próprios jornalistas ou no envolvimento de programadores/desenvolvedores de software, o que resulta muitas vezes em novos produtos, como ferramentas, aplicativos. Todas são formas novas de se contar histórias. Em comum, têm o fato de exigirem o conhecimento de linguagem de programação em certo grau, seja indiretamente, com a ajuda de softwares feitos

especificamente para a criação de produtos como esses, ou diretamente, com o uso desenvolvimento próprio de novas plataformas.

A inserção da programação e dos algoritmos no jornalismo é central para os autores que trabalham com o conceito de “jornalismo computacional”, termo surgido em meados da década passada que ganhou força com uma série de artigos e estudos elaborados por Sarah Cohen, Fred Turner e James Hamilton entre 2009 e 2011. Para eles, jornalismo computacional “Broadly defined, it can involve changing how stories are discovered, presented, aggregated, monetized, and archived. Computation can advance journalism by drawing on innovations in topic detection, video analysis, personalization, aggregation, visualization, and sensemaking”²² É realmente uma definição abrangente, que, porém, não colabora muito para especificar a prática. Em outro artigo, os dois últimos definem o conceito de forma mais clara, como “(...) the combination of algorithms, data, and knowledge from the social sciences to supplement the accountability function of journalism (TURNER; HAMILTON, 2011, p.2)”²³. A termo tem maior aceitação no campo das ciências da computação e tem como desvantagem a ênfase na tecnologia.

Lembrando o alerta de Meyer (1973) sobre o risco do foco excessivo nas tecnologias em detrimento do conhecimento gerado por elas, Träsel (2013b) defende o termo Jornalismo Guiado por Dados pela vantagem semântica de direcionar a atenção para os resultados em vez dos instrumentos utilizados. Além disso, pode-se afirmar, está sendo consagrada pelo uso. Ele propõe a seguinte definição, que é até agora a melhor encontrada por este pesquisador: “O jornalismo guiado por dados é a aplicação da computação e dos saberes das ciências sociais na interpretação de dados, com o objetivo é ampliar a função da imprensa como defensora do interesse público” (TRÄSEL, 2013b, p.3).

O Jornalismo Guiado por Dados, então, é um novo fluxo de produção jornalística, eminentemente multidisciplinar, que cria novos modos narrativos a partir do uso das bases de dados não apenas como fonte de informação, mas como suporte e definidoras do processo criativo e narrativo. Na presente pesquisa, dada a discussão sobre o conceito ainda estar em andamento, optou-se por um recorte temporal – a partir da primeira década deste século – e por apenas exemplos e casos que se apresentam ou são reconhecidos como jornalismo de dados ou JGD.

²² “Definido de forma abrangente, pode envolver mudar como as histórias são descobertas, apresentadas, agregadas, monetizadas e arquivadas. A computação pode avançar o jornalismo ao criar inovações em detecção de temas, análise de vídeo, personalização, agregação, visualização e produção de sentido”

²³ “(...) combinação de algoritmos, dados e conhecimentos das ciências sociais para suplementar a função de fiscalização do jornalismo”(tradução do autor)

1.4 Multidisciplinaridade

Além das técnicas tradicionais de reportagem e edição, é necessário dominar um conjunto de novas ferramentas e habilidades e, em última instância, uma nova linguagem para a produção do jornalismo de dados. Novas formas de analisar, organizar e visualizar as informações, além de aplicativos e softwares, podem ser criados para se adequar ao conteúdo e ao formato exigidos para cada reportagem. Não que os jornalistas devam se tornar programadores, designers ou estatísticos, mas é imprescindível entender o funcionamento, a dinâmica e, principalmente, os potenciais desse novo suporte, para saber como usá-lo.

Tradicionalmente, foi possível atuar nos meios de comunicação sem domínio dos números, mas a sobrevivência desses profissionais está a cada dia mais ameaçada pelo avanço das bases de dados no dia a dia dos cidadãos. Há fontes de informação públicas e disponíveis em praticamente todas as áreas da cobertura jornalística que podem tornar mais interessantes ou dar mais contexto às matérias mais simples, seja sobre economia, criminalidade, política, cidades ou esportes. Não que todos devam se tornar jornalistas guiados por dados, mas um mínimo de conhecimento dos conceitos básicos de estatística é essencial para usar os números adequadamente.

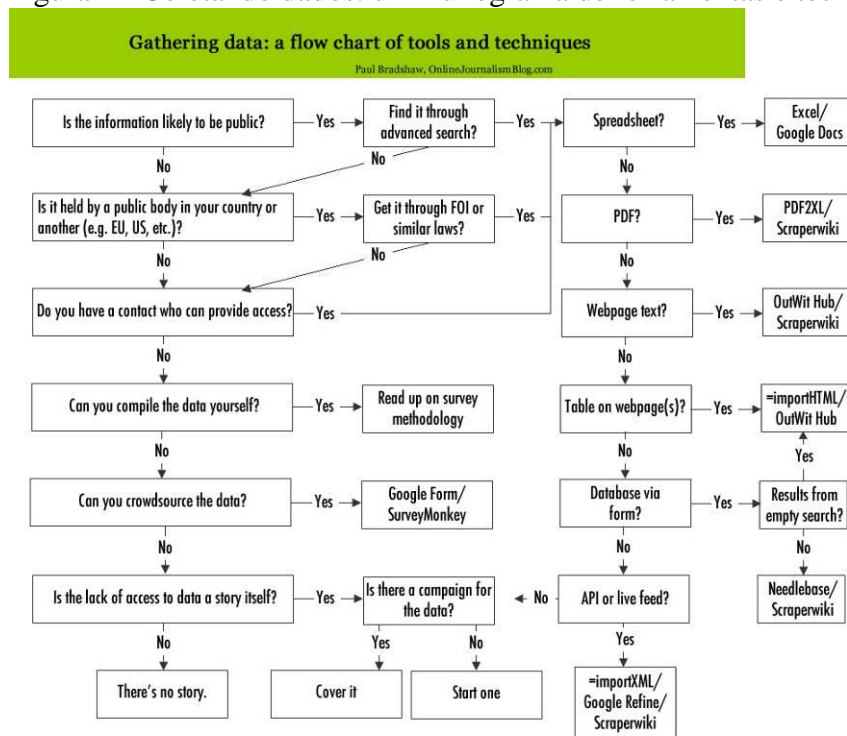
O sucesso das previsões do blogueiro e estatístico Nate Silver na cobertura das eleições presidenciais norte-americanas de 2012 mostra que o espaço para palpiteiros está diminuindo, como alerta Bradshaw (2012). A exemplo do computador da CBS em 1952, o estatístico previu acertadamente a vitória de Barack Obama muito antes de todos os grandes veículos de mídia, que apontavam uma disputa acirrada com Mitt Romney. Inicialmente, Silver divulgava suas previsões por meio de um site independente (<http://www.fivethirtyeight.com/>), criado em 2008. A visibilidade alcançada o levou a ser contratado pelo New York Times, que hospedou seu conteúdo entre agosto de 2010 e julho de 2013, quando o contrato foi encerrado. Se os jornalistas não melhorarem sua formação, outros profissionais – como Silver – ocuparão esse espaço.

O fluxograma criado por Bradshaw (Figura 2) para facilitar aquela que seria a primeira etapa do jornalismo de dados, a coleta, ilustra bem o volume de ferramentas e habilidades envolvidas nessas etapas. Nele, fica patente a multidisciplinaridade necessária para dar conta da atividade. Pode ser destacada a exigência de conhecimento prévio de ao menos três campos, como o Direito (leis de acesso à informação) Computação (linguagens de programação, *scrapping*) e Estatística (compilação de dados por conta própria).

A partir da análise de artigos, textos e opiniões de figuras destacadas do jornalismo de dados, foram identificadas cinco áreas de conhecimento considerados mais relevantes para a prática do jornalismo de dados. São elas: Programação, Legislação de acesso à informação, Estatística, Visualização/Design e Narrativas para novas mídias.

Essas cinco competências não necessariamente são obrigatórias para o jornalista de dados, que pode e deve atuar em conjunto com outros profissionais de maneira multidisciplinar, mas ao menos um conhecimento básico seria desejável, até mesmo para que possa conhecer as potencialidades de cada ferramenta. Dessas, provavelmente as mais características dos jornalistas seriam o conhecimento da legislação de acesso à informação (dos meios de conseguir as informações) e a narrativa, pois de resto, pode-se trabalhar em conjunto com programadores, designers e estatísticos para produzir as reportagens. Gordon (2013) elenca três profissionais como essenciais para a montagem de uma equipe de jornalismo de dados: repórter assistido por computador, desenvolvedor de aplicativos de notícias e especialista em visualização de dados. A esses pode-se adicionar outros integrantes, não essenciais, mas desejáveis, como um especialista em estatística, como é o caso do Estádio Dados.

Figura 2 - Coletando dados: um fluxograma de ferramentas e técnicas



Fonte: Bradshaw (2011b).

1.5 A produção de sentido no JGD

A ideia de notícia como reprodução da realidade, proposta por algumas teorias tradicionais do jornalismo, vem perdendo espaço há décadas no meio acadêmico à medida que novas pesquisas e teorias passaram a considerar a dimensão simbólica da atividade e demonstraram a influência de fatores subjetivos em todo o processo jornalístico, desde a seleção das fontes e das informações até a construção das narrativas e mesmo na formação dos profissionais e na distribuição do conteúdo (BORELLI, 2005; MCCOMBS, 1972; SCHUDSON, 1982; TRAQUINA, 1999; TUCHMAN, 1978; VIZEU, 2002; LAGE, 1999). A visão foi gradualmente substituída pela ideia de que o jornalismo promove uma construção social da realidade, de acordo com o contexto em que está inserido (TUCHMAN, 1978). Dentro do campo da mídia o jornalismo é um dos lugares privilegiados para a construção da atualidade e realidade social, pois é o espaço onde se constrói o sentido a partir do que de fato ocorreu ao contrário do entretenimento que trabalha com a ficção (BORELLI, 2005, p.6). A notícia tem caráter referencial (LAGE, 2009), pois não reflete, mas se refere a uma realidade observada, usando um arcabouço simbólico influenciado por fatores internos (valores, cultura, etc) e externos (dinâmica de produção e organização do veículo, relação com outros atores, etc). O jornalista é portanto um produtor de sentidos²⁴.

Naquele que é um dos principais estudos recentes sobre a função do jornalismo e dos jornalistas na sociedade, apresentado no livro *The Elements of Journalism*, Bill Kovach and Tom Rosenstiel afirmam que o jornalista na atualidade tem o papel de ajudar o público a produzir sentido a partir de um conjunto de informações cada vez maior e mais complexo:

The new journalist is no longer deciding what the public should know. She is helping audiences make order out of it. This does not mean simply adding interpretation or analysis to news reporting. The first task of the new journalist/sense maker, rather, is to verify what information is reliable and then order it so people can grasp it efficiently. (KOVACH; ROSENSTIEL, 2001)²⁵

²⁴ Como afirma BORELLI (2005, p. 6): “*sentido* é muito mais que o significado (dado como universal pelo uso e convenção social), pois excede os limites da frase posta, é muito mais que a aparência inscrita, é definido apenas na relação entre o eu e o tu, isto é, o leitor de um jornal precisa mobilizar sua cognição, seus conhecimentos, valores, culturas, sua biografia para estabelecer novos sentidos, seus vínculos com o que foi ofertado na relação”.

²⁵ “O novo jornalista não decide mais o que o público deve saber. Ele ajuda as audiências a organizar as informações. Isso não significa simplesmente adicionar interpretação ou análises às reportagens. A primeira tarefa do novo jornalista/ produtor de sentidos, em vez disso, é verificar quais informações são confiáveis e organizá-las de maneira que as pessoas possam compreendê-las eficientemente. (KOVACH e ROSENSTIEL, 2001)

De que maneira o jornalismo guiado por dados contribui para a produção de sentidos e a construção da realidade no mundo contemporâneo? Seu maior mérito parece ser o de trazer a público, de maneira inteligível e acessível, informações de interesse geral antes restritas a determinados grupos sociais, que detinham ou o controle sobre esses dados – e foram, por vontade própria ou por força da lei, levados a divulgá-la –, ou os meios para entendê-los e usá-los em benefício próprio. Grandes bases de dados, ainda que de interesse público, na maior parte das vezes não são compreensíveis pelos cidadãos comuns, que não têm o conhecimento, as habilidades ou o tempo e interesse para interpretá-las. Ao fazer a ponte entre esses dados e a população, o jornalismo fornece ferramentas que auxiliam estes cidadãos, e também empresas, governos e instituições, a tomar decisões.

Considere-se, por exemplo, o projeto de JGD *Dollar for Docs* (<<http://projects.propublica.org/docdollars>>), lançado em outubro de 2010 pela organização sem fins lucrativos ProPublica, em parceria com dois grupos de mídia (*Boston Globe* e *Chicago Tribune*), uma publicação de nicho (*Consumer Reports*) e duas emissoras públicas, de rádio (*NPR*) e televisão (*PBS*). A iniciativa é centrada em uma base de dados com os pagamentos feitos por 15 empresas farmacêuticas a médicos e pesquisadores norte-americanos, no valor de US\$ 2 bilhões, desde 2009, com palestras, consultorias, pesquisas, viagens e refeições. Uma equipe multidisciplinar, formada por jornalistas e programadores, coletou os dados e usou diferentes programas para fazer a conversão dos arquivos, padronizá-los e inseri-los em uma mesma base de dados que pudesse ser consultada na internet. Deu origem a mais de uma centena de reportagens em diversos veículos de mídia e revelou, partindo da análise dos dados, a existência de médicos que recebiam valores superiores a seus rendimentos profissionais para prescrever remédios para seus pacientes, incluindo para usos não aprovados pelas autoridades do país. Foi criado um aplicativo online que permite a qualquer pessoa buscar o seu médico ou hospital e saber quanto recebeu de quais companhias, além de fazer comparativos e análises relacionando empresas, profissionais de saúde e clínicas.

Individualmente, as informações tinham um sentido. Ao serem combinadas e analisadas, ganharam um novo sentido. Em resumo, dados de extrema importância para a sociedade só puderam ser divulgados e compreendidos pela população por meio do Jornalismo Guiado por Dados. É difícil acreditar que individualmente uma das pessoas afetadas pelo esquema de pagamentos teria tempo ou capacidade para realizar essa empreitada, ou que empresas ou organizações que não sejam de mídia se dedicassem a esse tipo de trabalho.

Outro exemplo é o projeto *kameraspotter* (<[http:// kameraspotter.information.dk/](http://kameraspotter.information.dk/)>), lançado em março de 2012 pelo jornal dinamarquês *Dagbladet Information* com o objetivo de mapear todas as câmeras de segurança da Dinamarca. Desde então cerca 2,2 mil câmeras foram identificadas por leitores e jornalistas através de um aplicativo para celular criado pelo jornal. Em conjunto foi elaborada uma série de reportagens sobre a vigilância e armazenamento de dados de telefones celulares intitulada “Watching the watchers”, publicada ao longo de dois meses no jornal impresso diário e na internet. Segundo os autores Sebastian Gjerding e Peter Anderson, o projeto foi baseado na premissa do jornalismo de dados de produzir sentido a partir de bases de dados.

Além de coletar os dados das câmeras, os jornalistas foram às ruas investigar as razões para a onipresença dos dispositivos de monitoramento no país e descobriram que o crescimento era motivado pelo medo de ataques terroristas – o país se tornou alvo de ameaças de islâmicos radicais desde que o jornal *Jyllands-Posten* publicou caricaturas ilustrando Maomé, o profeta muçulmano, mas as representações visuais são proibidas pela religião islâmica – e pelo barateamento da tecnologia. Interessante notar que, ao olhar para o cotidiano dos dinamarqueses, tentando identificar o porquê da existência das câmeras, os autores revelaram questões bem mais amplas de cunho político e econômico.

Ficou evidente, segundo eles, que há uma busca dos cidadãos dinamarqueses por mais controle e “segurança” como resultado do conflito político, econômico e ideológico entre Ocidente/Oriente e Cristãos/Muçulmanos, que se tornou dominante na esfera pública desde os atentados de 11 de setembro de 2001, em Nova York, nos Estados Unidos. O próprio discurso da “ameaça terrorista”, propagado incessantemente na opinião pública desde então, especialmente nos países desenvolvidos, pode ser considerado como uma forma de os poderes estabelecidos induzirem as pessoas a se submeter voluntariamente ao controle e considerá-lo um bem, por garantir a “segurança”, e não um instrumento de dominação. Ao ser colocado em contexto pelo JGD, o uso indiscriminado de câmeras ganha um novo sentido, criando um distanciamento e um espaço para reflexão sobre seu uso e funcionamento.

Não é á toa que o JGD tem uma forte relação com o jornalismo investigativo nos diversos países em que começa a se consolidar – no caso do Brasil, a Abraji tem sido a principal evangelizadora da prática nas redações. O jornalismo investigativo surgiu entre os anos 60 e 70, como resultado da percepção de que os jornalistas não são meros repassadores de informação, mas também atores e protagonistas, interferindo diretamente em fatos de interesse público (BORELLI, 2005). Na origem do JGD percebe-se também a disposição de revelar as informações que, ainda que públicas, restam escondidas em bases de dados muitas

vezes incompreensíveis para a maior parte da população. É quase paradoxal, pois a prática é marcada pela objetividade dos números e pelo uso de ferramentas, métodos e análises científicas. Mas, como afirma Borelli, “para a maioria dos autores contemporâneos, a objetividade tão difundida ao longo da história do jornalismo serviu apenas para propagar a superficialidade” (op. Cit). É preciso entender, portanto, as limitações e as forças que influenciam essa produção de sentido para melhorar o produto jornalístico. A ética profissional evolui, então, da objetividade para a transparência (HUNTER e WASSENHOVE, 2010). Ou, na maneira de Spinoza, a liberdade do jornalista é tanto maior quanto o conhecimento dos cordéis que o manipulam. E a credibilidade equivalente à publicidade que dá a esses cordéis.

1.5.1 A ilusão da objetividade dos dados

As bases de dados podem criar a ilusão de serem fontes imparciais e objetivas, mais próximas da realidade factual, mas o fato é que estão sujeitas às mesmas influências que as fontes tradicionais, uma vez que são criadas por pessoas, que têm interesses, sofrem pressões e estão inseridas em um contexto social específico. Efetivamente, em determinados casos podem ser menos sujeitas a influências e serem alternativas mais confiáveis, mas da forma ao que já acontece com outras fontes de informação, que podem ter diferentes graus de confiabilidade e credibilidade. Essa questão é central em Parasie e Dagiral (2013), que questionam se o uso de bases de dados realmente liberta os jornalistas das fontes.

Os jornalistas guiados por dados, dessa forma, não estão isentos de seguir os mesmo critérios de apuração e checagem adotados nas reportagens tradicionais. Como afirma Codrea-Rado (2012): “Data can be political. Numbers, facts, and figures are just as much a source as quotes from politicians, community leaders and academics. The graphics editor and data reporter have just as much responsibility to verify their numbers as the crime reporter does to accurately collect the facts of a police case.”²⁶

A quais influências estão submetidos e em que contexto estão inseridos aqueles que trabalham com Jornalismo de Dados? Essa é uma questão de resposta difícil, que exigiria uma

²⁶ “Dados podem ser políticos. Números, fatos e valores são fontes equivalentes a citações de políticos, líderes comunitários e acadêmicos. O editor e o repórter de dados têm tanta responsabilidade de verificar seus números quanto o repórter policial de apurar os fatos de um crime” (tradução do autor).

análise complexa das forças que se impõem neste processo e que não são objeto da presente pesquisa. No entanto, é possível levantar, de maneira introdutória, alguns possíveis fatores que podem influenciar a construção da notícia a partir de bases de dados e que seriam característicos dessa prática.

Um deles é a própria obtenção dos dados, cuja coleta primária é feita na maior parte das vezes por terceiros. Por depender da disponibilidade desse material, o processo de seleção pode se feito não de acordo com a relevância do tema, mas a viabilidade de uso. Dados de interesse público que estejam em formatos pouco manipuláveis, como Java ou PDF, por exemplo, podem ser deixados de lado por falta de tempo ou recursos para torná-los mais acessíveis, ou desconhecimento das ferramentas. Organizações e instituições que não tenham capacidade ou meios de criar bases de dados úteis também tendem a ser deixadas de lado.

Outra possibilidade é o repórter usar apenas bases de dados como fonte para a reportagem, sem ouvir outras fontes e checar informações, apresentando o material com o viés de quem o criou, sem análise crítica e contraposição. Essa é uma tentação real, pois o profissional pode se deixar levar pelas inovações técnicas – o volume e a complexidade dos dados, as diversas formas de manipulação e visualização – e não atentar para a maneira como foram coletados ou não ter mais tempo para realizar o trabalho de reportagem propriamente dentro dos prazos exigidos pelo veículo.

Uma forma alternativa de analisar essa questão é considerar que, em um contexto no qual boa parte das relações sociais – e da experiência humana como um todo – é registrada em bases de dados digitais, aqueles que tiverem a capacidade e as ferramentas para entender esses dados terão uma vantagem sobre o restante das pessoas e poderão usar esse conhecimento em benefício próprio. Como afirmou o chefe de notícias interativas do jornal inglês *Financial Times*, Martin Stabe:

When every industry now has access to the computing power to analyze data that could give them a leading edge in business, journalists need to be able to analyze the same data to find the truth, Stabe said, for the public interest(...) "Journalists are starting to be intellectually outgunned by other industries," he said, "and if we don't keep up we will be at the mercy of the PR machine" (in REID, 2014)²⁷.

²⁷ “Quando todas as indústrias tem acesso à capacidade computacional para analisar dados que podem dar a ela vantagens nos negócios, é de interesse público, disse Stabe, que os jornalistas sejam capazes de analisar esses mesmo dados para buscar a verdade (...) “Jornalistas estão começando a ser intelectualmente superados por outras indústrias e se nós não acompanharmos esse ritmo estaremos à mercê da máquina de relações públicas.” (tradução do autor)

Fica clara, então, a necessidade de dominar não apenas as ferramentas e habilidades necessárias para se analisar os dados, mas o conhecimento dos temas de maneira a se aproximar ao máximo da precisão factual, permitindo que as audiências possam formar seu entendimento da realidade da melhor forma possível. Como afirmam Kovach e Rosenstiel (2001), “o jornalismo é, em essência, a disciplina da verificação”.

1.5.2 Limitações e riscos do JGD

Como qualquer outro instrumento ou método usado no jornalismo, o JGD não é o ideal para ser usado em todos, mas sim em determinados casos e situações. As ferramentas e aplicativos criados a partir da sua prática também não são a melhor forma de esclarecer o público e favorecer sua compreensão sobre os diferentes assuntos, mas uma das várias possibilidades que podem e devem ser usadas inclusive em conjunto. A internet está recheada de exemplos do mau uso da tecnologia de visualização de dados, um efeito colateral de sua popularização e aplicação indiscriminada.²⁸ Um infográfico malfeito pode ser incompreensível, desinformar sua audiência ou até mesmo mentir, como afirma Alberto Cairo²⁹: “Can information graphics (infographics) and visualizations lie? Most designers and journalists I know would yell a rotund “yes” and rush to present us with examples of outrageously misleading charts and maps”³⁰.

O mesmo pode ser dito do uso de bases de dados para a elaboração de reportagens. Há sempre o risco de que o deslumbramento com o volume de dados ou, principalmente, a capacidade de compreendê-los, leve o jornalista a centrar sua reportagem somente nos números, esquecendo das pessoas a quem se referem e afetam. São duas as prováveis conseqüências: produzir um texto árido e pouco atraente, que não vai ser lido por ninguém, falhando assim em seu objetivo de informar o público, ou criar uma narrativa falha ou

²⁸ Alguns exemplos podem ser encontrados em: <http://www.theguardian.com/news/datablog/gallery/2013/aug/01/16-useless-infographics>, <http://terribleinfographics.tumblr.com/>, http://www.slate.com/blogs/future_tense/2013/11/08/infographic_top_6_reasons_infographics_are_terrible_infographic.html

²⁹ Ver em: <http://data-visualization-software.com/lying-with-infographics-and-visualization/>

³⁰ “Os gráficos (infográficos) e visualizações podem mentir? A maioria dos designers e jornalistas que conheço diriam um rotundo “sim” e correriam para mostrar exemplos de gráficos e mapas escandalosamente enganosos” (tradução do autor).

incompreensível, o que é ainda pior e pode ter como resultado desinformar o leitor. Ambos os erros já foram cometidos por este pesquisador.

Este é outro ponto importante: o JGD é complementar à atividade jornalística, que tem seus princípios e objetivos, e não um fim em si mesmo. Como costuma ser ressaltado por praticamente todos os principais jornalistas envolvidos com o JGD, nada substitui a reportagem tradicional, a curiosidade do repórter, a checagem dos fatos e, principalmente, uma boa história. O maior dos perigos é a confiança excessiva na tecnologia.

2 A CRIAÇÃO DE VALOR PELO JORNALISMO GUIADO POR DADOS

2.1 Economia de mídia

A Comunicação é um campo interdisciplinar e, como tal, a análise de seus problemas exige a integração de teorias e conceitos de diversas procedências (LOPES, 2004). Tradicionalmente, o estudo dos fenômenos comunicacionais se apropriou do ferramental teórico e conceitual e atraiu o interesse de pesquisadores de diversas áreas das ciências sociais, como História, Sociologia, Psicologia, Linguística e Antropologia. No caso da Economia, porém, sua análise foi incipiente ou limitada, especialmente à Economia Política, de orientação marxista.

Segundo Parkin, Powell e Matthews, a economia pode ser definida como o “estudo de como as pessoas tomam decisões para lidar com a escassez” (2008, p. 4 apud DOYLE, 2013, p.8). Já a economia de mídia é definida por Albarran como o “uso das teorias conceitos e princípios da economia para estudar os aspectos macroeconômicos e microeconômicos de empresas e indústrias de mídia de massa” (2004, p.13, tradução do autor). Para Picard, um dos pioneiros da disciplina, a economia de mídia se preocupa em estudar “como os operadores de mídia atendem os desejos e necessidades de informação e entretenimento das audiências, anunciantes e da sociedade com os recursos disponíveis” (1989, p.7, tradução do autor). Doyle afirma que o campo desafia os conceitos tradicionais da economia por, entre outros motivos, trabalhar com uma matéria-prima que não é escassa e distribuir produtos que muitas vezes não se esgotam para um duplo mercado, o que torna ainda mais interessante o seu estudo (DOYLE, 2013).

A análise econômica dos meios de comunicação, porém, começou tardiamente, a partir do fim do século XIX, e se voltou para seus fins políticos e sociais do que para a sua avaliação como uma atividade econômica em si, com poucas exceções, como afirma Lins: “Usualmente, porém, a análise econômica levada a termo no contexto das teorias da comunicação limita-se à constatação de dados anedóticos, procurando interpretá-los à luz do instrumental sociológico” (LINS, 2005).

Toussaint-Desmoulin (2011) credita a demora para se tratar da mídia como atividade econômica a três fatores. O primeiro é uma visão romântica do público em relação aos meios de comunicação, que confere uma aura a seus profissionais, tanto jornalistas, como apresentadores e estrelas de cinema, separando-os da “realidade” das práticas de produção e

distribuição e dos custos envolvidos. Em segundo lugar, a tendência a se considerar a mídia primordialmente como uma atividade de interesse público, o que cria uma resistência a enxergá-la como voltada para o lucro, dentro de um contexto de oferta e demanda. Por último, a falta de relevância econômica em si, tanto pelo porte reduzido das empresas em relação a outros setores, como pela falta de dados estatísticos, necessários para a sua análise.

Esse cenário começou a ser alterado a partir dos anos 1980 por uma série de fatores. A crescente industrialização; a desregulamentação da mídia em diversos países, que eliminou monopólios artificiais e estimulou a competição; a emergência de novas tecnologias e canais de distribuição e suporte de informação; a convergência dos meios de comunicação; a maior necessidade de capital, com a conseqüente diversificação das fontes de financiamento; e a concentração de mercado (ALBARRAN, 2010; DOYLE, 2013; TOUSSAINT-DESMOULINS, 2011).

As mudanças levaram a uma maior dependência tecnológica e econômica das empresas de mídia e impulsionaram um processo de concentração de mercado a partir da década de 1990, quando houve uma onda de fusões e aquisições que resultou na formação de grandes conglomerados, como Time-Warner, NBC Universal, Comcast, Viacom e News Corp. O porte das grandes corporações garante maior competitividade e permite a adoção de estratégias de internacionalização, já que têm capital e poder econômico para produzir e competir em todo o mundo, e de diversificação, pois reúnem diversas plataformas, aproveitando o mesmo conteúdo de várias maneiras e usando uma para impulsionar as outras. O peso dessas corporações de mídia nas economias nacionais e mundial aumentou, se tornando equivalente ao de outros importantes setores, como vestuário e combustíveis (ALBARRAN, 2010, p.16), no caso dos Estados Unidos.

O surgimento de novas plataformas de produção e distribuição de informação provocou uma ruptura não apenas nos formatos de produção e distribuição de conteúdo, como do perfil e comportamento dos consumidores e dos modelos de negócio da comunicação. Especialmente desde a virada do milênio, passa-se por um processo de ampliação e diversificação dos veículos e formatos informativos e noticiosos, que leva ao aumento da concorrência e da competição por audiência e receitas publicitárias (ALBARRAN, 2010; PICARD, 2013), especialmente por força da convergência midiática (JENKINS, 2008), que reúne conteúdos de mídia antes isolados em uma mesma plataforma, a internet.

É emblemático das transformações sofridas pela mídia enquanto atividade econômica que a empresa mais valiosa do mundo³¹ atualmente, a Apple, deva seu recente sucesso a uma mudança de estratégia que a transformou de uma fabricante de computadores em uma companhia de mídia³², que não só fornece o suporte para a distribuição de conteúdo³³, mas os distribui e negocia, por meio de suas lojas iTunes e App Store. O domínio tecnológico foi aliado à prestação de serviços ao consumidor e aos produtores de conteúdo, como forma de impulsionar as vendas de seus próprios equipamentos e reter para si receitas de outros setores. As transformações, como se pode observar, ocorreram não só do lado da oferta, mas também da demanda, pois consumo se tornou móvel, personalizado, multiplataforma.

As estratégias adotadas pelas empresas para lidar com esse novo modo de consumo também passaram a atrair a atenção de economistas, já que as mudanças não se restringem à mídia, mas se espalham por outros setores. Como resultado, o estudo da economia de mídia tem ganhado espaço nos últimos anos, o que pode ser observado pela publicação de livros voltados para o tema (ALBARRAN, 1996; DOYLE, 2013; PICARD, 1989, 2013; TOUSSAINT-DESMOULINS, 2011), assim como periódicos especializados, como *The Journal of Media Economics*, fundado em 1988, *The International Journal of Media Management*, em 1999 e o *Journal of Media Business Studies*, em 2004. Não por acaso, um dos campos em crescimento é o estudo dos aspectos de gestão empresarial da mídia, como ressalta Faustino (2010). Pode-se afirmar também que técnicas e estratégias de gestão empresarial têm-se tornado cada vez mais importantes no âmbito interno das organizações de mídia, especialmente por conta das fortes transformações no cenário tecnológico e econômico decorrentes da popularização da internet e das mídias digitais.

A mídia, portanto, passou por uma série de transformações nas últimas três décadas, provocadas por fatores econômicos, tecnológicos e políticos, que estão levando a mudanças significativas nos processos de produção, distribuição e consumo de seus produtos. Entre as principais causas estão a crescente industrialização dos meios de comunicação, o surgimento de novas tecnologias de produção, distribuição e recepção de conteúdo, a globalização e a digitalização. Entre as consequências, o aumento da dependência tecnológica e de

³¹ Kell, John. 10 Abr. 2012. *Dow Jones Newswires*. Apple Market Value Jumps Above \$600 Billion Mark. Disponível em: <http://online.wsj.com/article/BT-CO-20120410-710983.html>. Acesso em: 3 Jun 2012.

³² Vogelstein, Fred. 25 Out. 2010. *Wired*. Behold, the Next Media Titans: Apple, Google, Facebook, Amazon <http://www.wired.com/epicenter/2010/10/behold-the-next-media-titans/all/1>. Acesso em 3 Jun 2012.

³³ MP3 players (iPod), telefones celulares (iPhone), tablets (iPad), além de computadores (iMac), notebooks (MacBook) e até televisões (Apple TV)

financiamento por parte dos meios de comunicação, a crescente personalização, mobilidade e diversidade dos modos de consumo, e a concentração de empresas e de mercados.

Esses fenômenos podem ser observados e compreendidos sob a ótica econômica e os conceitos e teorias deste campo científico oferecem uma poderosa ferramenta de análise e pesquisa dos meios de comunicação e da mídia de maneira geral, contribuindo para o seu entendimento. Além disso, as teorias e conceitos da Economia, sejam de abordagem macroeconômica ou microeconômica, são úteis não só por enriquecer o arcabouço teórico usado no estudo dos fenômenos da comunicação, mas também porque são eles próprios usados por instituições (empresas, governos, organizações) para basear seu comportamento e suas ações. Deixá-los de lado é abrir mão de compreender uma parte importante dessa dinâmica social.

2.1.1 A economia de mídia no Brasil

A pesquisa acadêmica das organizações de mídia no Brasil sob a ótica econômica tem sido feita majoritariamente por autores identificados com a Economia Política, cujo foco é voltado para a análise do discurso e das relações de poder. Há na literatura estudos que abordam temas como concentração de mídia, monopólio e políticas públicas de comunicação, mas sempre sob essa perspectiva. Como afirma Moreira (2014), praticamente não existem pesquisas estruturantes sobre a economia das comunicações no país, “que teriam como objeto as forças e limitações que regulam a estrutura do mercado das indústrias de mídia e telecomunicações internamente e externamente e o desdobramento dos interesses e da capacidade de interação do usuário”.

O trabalho inaugural neste campo, com as características mencionadas, foi realizada pela própria autora como parte de esforço de pesquisa que reuniu autores de 30 países para realizar um levantamento abrangente dos dados econômicos relativos à propriedade e concentração de mídia no mundo. Coordenado por Eli Noam, do Columbia Institute for Tele-Information, é denominado *Media Concentration and Ownership Around the World: Who Owns the World's Media?* e será divulgado este ano.

No capítulo sobre o Brasil, Moreira aponta que houve uma série de transformações no cenário das indústrias de mídia e telecomunicações desde o fim da década de 1990, iniciadas com a privatização das companhias telefônicas em 1999, que levou a um processo de fusões e

aquisições, resultando na formação de grandes conglomerados de mídia com atuação nacional em diversos setores. Nesse período houve a popularização da internet, da TV a cabo e dos telefones celulares e o surgimento da banda larga. Em dez anos, os 25 milhões de telefones fixos aumentaram para 41,8 milhões e os 15 milhões de celulares passaram para 191,5 milhões; a banda larga que era inexistente passou a ser usada por 12,8 milhões de pessoas e a TV a cabo, por 9,1 milhões (MOREIRA, 2014, p.6); e 31% dos domicílios passaram a ter acesso à internet³⁴ – hoje são 40%³⁵. Ao fim da primeira década do século XXI, as empresas de telecomunicações somavam valor de mercado de US\$ 79,3 bilhões (MOREIRA, 2014, p.5). Houve, portanto, enormes mudanças tanto na dinâmica das indústrias de mídia quanto no perfil dos consumidores.

É importante destacar da abrangente pesquisa feita por Moreira, os principais dados dos setores em que atuam organizações de mídia eminentemente voltadas para o jornalismo – especialmente jornais impressos e sites noticiosos, que são o foco desta pesquisa – pois ajudarão a contextualização do caso em estudo. O mercado de jornais diários, aponta a autora, é dominado por quatro grupos – Folhapar, Infoglobo, S.A., O Estado de S. Paulo e RBS, que representam 30% da circulação total auditada. O setor de notícias online, diferentemente do que se poderia esperar, é muito mais concentrado. Os quatro grupos dominantes respondem por 78% da audiência: a Folhapar, controladora do portal *UOL*, responde por 32% do mercado; a espanhola Telefonica, com o portal, *Terra*, tem 21%; a Globopar, com o *Globo.com*, tem 13,3%; e a Telemar, com o *iG*, responde por 12% (MOREIRA, 2014, p.13-14). A audiência deriva principalmente do fato de os portais serem ligados aos provedores de internet, servindo como página inicial de boa parte dos usuários. Como destaca Moreira, as versões online dos jornais impressos, como *Folha Online* e *O Globo Online*, têm audiência inferior aos portais. Nota-se, portanto, que os meios jornalísticos tradicionais não são a fonte mais popular de notícias dos brasileiros na internet.

2.2 O conceito de valor aplicado ao jornalismo

O conceito de valor se origina da filosofia e tem definições particulares em diferentes áreas de conhecimento, mas todos remetem à ideia de juízo ou significação atribuído pelos

³⁴ In: “Pesquisa TIC domicílios e usuários 2010”. Rio de Janeiro: CETIC.br. Disponível em: <http://www.cetic.br/usuarios/tic/2010/analises.htm>

³⁵ In: “Pesquisa TIC domicílios e usuários 2012”. Rio de Janeiro: CETIC.br. Disponível em: <http://www.cetic.br/usuarios/tic/2012/>

homens àquilo que se relacionam. Está ligado à estima e ao apreço por determinadas pessoas ou objetos. Os valores podem ser classificados como intrínsecos ou instrumentais. Os primeiros são aqueles que têm valor por si mesmos, como a justiça, a liberdade, a verdade, a felicidade. Os outros são aqueles que têm valor na medida em que contribuem para alcançar os valores intrínsecos, como o entendimento, a participação democrática, compromisso. Podem ser divididos também entre valores individuais, como conhecimento e segurança, e sociais, como verdade, honestidade e interação social.

Para a economia, o valor tem base material e está relacionado à capacidade de troca. O estudo econômico do valor se divide principalmente em duas correntes teóricas: a que considera o valor objetivo e a que o considera subjetivo. A primeira, sustentada por autores clássicos, como Adam Smith, David Ricardo, Jean-Baptiste Say, John Stuart Mill e Marx, enxerga o valor como um bem concreto e majoritariamente determinado pela quantidade de trabalho médio aplicado em sua produção, somados aos gastos com matéria-prima, equipamentos, aluguéis e outros custos. Em termos de ótica de criação de valor, é possível afirmar que se coaduna com a “mentalidade de produção” característica do modelo científico de gerenciamento de empresas. Conhecido como taylorismo, por ter sido desenvolvida pelo engenheiro norte-americano Frederic Winston Taylor (1856-1915), o modelo de administração de empresas consiste em gerenciar de forma racional e científica, usando metodologias e práticas emprestadas das ciências exatas. Como as companhias de então não conseguiam dar conta da demanda existente, a prioridade era ampliar a produção, sem se preocupar com a diferenciação do produto. Hegemônico até a primeira metade do século passado, é voltado exclusivamente para os processos internos e tem como meta aumentar a eficiência da produção, ou seja, produzir mais com o mesmo ou menos recursos. Se inseria em um contexto no qual a oferta de produtos e serviços era, de maneira geral, menor que a demanda, pois, se não havia capacidade suficiente para produzir carros para todos que o desejavam, por exemplo, a prioridade era fazer mais veículos com os recursos disponíveis (MAGRETTA, 2013).

A escola neoclássica, fundada no fim do século XIX e representada por teóricos como Stanley Jevons, Carl Menger e Knut Wicksell, considera o valor de troca subjetivo, definido pela importância que um bem tem para um indivíduo, o que depende da utilidade e escassez desse bem. Baseada na teoria da utilidade marginal, a corrente evoluiu para se tornar predominante na contemporaneidade, integrando o que se convencionou chamar de Economia Ortodoxa. Na administração de empresas, pode-se afirmar que se coaduna com a “mentalidade de marketing”, modelo de gestão inaugurado com a publicação em 1954 do

livro *Practice of Management*, de Peter Drucker (MAGRETTA, 2013) e que tem o consumidor como ponto de partida para todo o processo produtivo de bens e serviços. A partir de então, as empresas que estavam concentradas em seus processos operacionais passaram a se voltar para fora, para descobrir quais eram as necessidades dos consumidores e o quanto estavam dispostos a pagar por isso. A desvantagem das teorias neoclássicas é não diferenciar valor e preço, uma vez que igualam os dois conceitos considerando que o valor de um produto é aquilo que o mercado está disposto a pagar.

Na análise do jornalismo essa diferenciação se faz necessária, pois seu valor não é determinado pelo custo e trabalho aplicado e nem ao preço pago por seus produtos, mas por fatores sociais mais complexos e atemporais. Essa diferenciação é colocada de forma bem clara pelo megainvestidor Warren Buffet, cuja visão de valor é resumida na seguinte frase: “Preço é o que você paga. Valor é o que você leva”³⁶. No caso da mídia e do jornalismo, o valor é definido pela utilidade dos produtos e serviços percebida pelos diferentes grupos de pessoas por eles afetados: jornalistas, anunciantes, audiência, investidores e sociedade.

O jornalismo, segundo Picard (2009), não tem valores intrínsecos, mas apenas instrumentais, pois não é bom por si mesmo, e sim, na medida em que contribui para a democracia, ilumina a sociedade e facilita as relações sociais. O autor ressalta que o próprio jornalista não tem uma base de conhecimento exclusiva, como médicos ou professores, que possa negociar. O valor de seu trabalho e de suas habilidades se dá pela distribuição de conhecimento alheio. Há três tipos de conhecimento: sobre a existência das coisas, sobre o que fazer com as coisas e sobre o motivo pelo qual as coisas acontecem e suas implicações. O jornalismo historicamente está centrado no primeiro tipo de conhecimento, que é aquele de menor valor.

Sob o ponto de vista econômico, o jornalismo cria três tipos de valor para os consumidores diferentes dos criados por outras atividades e produtos:

Primeiro, o conteúdo jornalístico possui benefícios funcionais que auxiliam as audiências a entender o seu lugar no mundo e os acontecimentos ao redor dele, fornecendo também informações e conselhos que as ajudam nas suas próprias vidas e atividades. Em segundo lugar, o consumo de notícias e informações possui benefícios emocionais ao fornecer escape, compaixão, sentido de pertença e comunidade, prazer, segurança, reafirmação e liderança. Em terceiro, o uso dos meios de comunicação possui benefícios de auto expressão nos quais as audiências podem identificar-se com as perspectivas ideais, vozes e opiniões de uma fonte noticiosa em particular. O valor também é fornecido quando é dada ao público a

³⁶ Paráfrase das idéias contidas em carta enviada por Warren Buffet aos acionistas da Berkshire Hathaway Inc em 1996. Ver em: <http://www.berkshirehathaway.com/letters/1996.html>

capacidade de dialogar com as organizações noticiosas adicionando informação, comentando e selecionando o conteúdo que deseja receber (PICARD, 2013, p. 96).

Esses valores, por sua vez, segundo o autor, derivam de três funções, e suas habilidades associadas: exclusividade no acesso a fontes de conhecimento, na seleção e hierarquização das informações e distribuição eficiente dessas informações. Até a ascensão da internet e das mídias digitais, as organizações de mídia detinham a hegemonia dessas funções, pois operavam em mercados que tendiam sempre a se tornar monopolistas ou oligopolistas (ALBARRAN, 2010, DOYLE, 2013, TOUSSAINT-DESMOULINS, 2011), ou seja, com pequeno número de concorrentes. Seja no impresso, rádio ou televisão, apenas poucas organizações públicas e privadas ofereciam produtos e serviços que cumpriam essas funções, o que garantia com que tivessem um alto valor. Com a popularização da web e de todas as tecnologias associadas, essa exclusividade acabou. O acesso a fontes de informação e conhecimento, antes restrito, tem se espalhado por meio de ferramentas como emails, sites e redes sociais. Autoridades e personalidades antes praticamente inatingíveis pelo cidadão comum hoje podem ser alcançadas diretamente e responder perguntas de forma individual. O maior exemplo é o presidente dos Estados Unidos, Barack Obama, considerado o homem mais poderoso do mundo por conta de seu cargo, que responde a perguntas e interage com as pessoas de todo o mundo por meio de sua conta na rede social Twitter <<http://twitter.com/BarackObama>>.

Na era “analógica”, as organizações de mídia, em conjunto com elites políticas e econômicas, tinham forte influência sobre quais temas seriam considerados relevantes na esfera pública (teoria do agendamento), como eram apresentados e de que forma seriam debatidos (teoria do enquadramento) e, em última instância, sobre a construção da realidade social (teoria do *newsmaking*) (BORELLI, 2005; MCCOMBS; SHAW, 1972; SCHUDSON, 1982; TRAQUINA, 1999; TUCHMAN, 1978; VIZEU, 2002). Este poder sobre o debate público está diminuindo consideravelmente com a introdução das tecnologias digitais, que reúnem de forma inédita descentralização, flexibilidade, expressão individualizada e comunicação global e horizontal (CASTELLS, 2003). A difusão irrefreável de informações e de modos para produzi-las e distribuí-las torna impossível a qualquer grupo social exercer individualmente o controle sobre as informações (WILLIAMS; CAPRINI, 2004, p.1.217). O público passou a ter papel ativo na produção e distribuição dos conteúdos, o que já leva as

empresas a projetá-los de acordo com sua capacidade de espalhamento³⁷ (JENKINS; FORD; GREEN, 2013).

Os meios de comunicação de massa tradicionais são atualmente influenciados por blogs e redes sociais, onde as próprias pessoas escolhem e atribuem importância a cada informação. Um dos maiores exemplos dessa nova ordem é o site reddit (<<http://www.pt.reddit.com>>). Criado em 2005 nos Estados Unidos por dois universitários, é uma rede social na qual os usuários podem publicar conteúdo multimídia e votar no que está publicado, atribuindo nota positiva (*upvote*) ou negativa (*downvote*), influenciando o destaque que terão no site. Controlado pelo conglomerado Advance Publications,³⁸ dono da editora Condé Nast, o site registrou 56 bilhões de visualizações de página por 731 milhões pessoas em 2013³⁹. Não há qualquer hierarquia senão aquela do interesse da audiência⁴⁰. Entrevistas concedidas por autoridades e personalidades públicas, como Bill Gates, Barack Obama, Edward Snowden e Arnold Schwarzenegger, são colocadas ao lado de entrevistas de anônimos com histórias curiosas, memes, trechos da Wikipédia⁴¹, fotos pessoais, entre outros conteúdos.

Quebrou-se então a exclusividade dos jornalistas e organizações de mídia no acesso às fontes e às ferramentas para selecionar, produzir e distribuir informações. Considerando o valor econômico de troca como subjetivo e definido pelo consumidor de acordo com a utilidade e escassez, o valor econômico das funções tradicionais do jornalismo foi reduzido a quase zero. Segundo Picard (2009), além dos fatores externos, que afetaram o valor dos produtos e serviços jornalístico, outro fator, desta vez interno desvalorizou o trabalho jornalístico, a profissionalização. Ao padronizar as habilidades e técnicas dos profissionais, transformou-os em uma “commodity”⁴², aumentando sua oferta e permitindo que fossem facilmente substituídos.

³⁷ Tradução do autor do termo original “spreadability”.

³⁸ Ver em: <<http://www1.folha.uol.com.br/fsp/tec/65467-caotico-reddit-atrai-20-mi-de-usuarios.shtml/>>

³⁹ Ver em: <<http://blog.reddit.com/2013/12/top-posts-of-2013-stats-and-snoo-years.html/>>

⁴⁰ Ver em: <<http://www.theatlantic.com/technology/archive/2014/01/ama-how-a-weird-internet-thing-became-a-mainstream-delight/282860/>>

⁴¹ Enciclopédia colaborativa multilíngue baseada na web e criada em 2001.

⁴² O termo commodity no sentido denotativo se refere a bens padronizados e negociados em grandes quantidades para entrega imediata ou futura, como petróleo, grãos e minerais. O sentido usado pelo autor é conotativo, ou seja, significa que um profissional tem o mesmo valor de outro, pois, do ponto de vista da empresa, não há diferença entre eles, assim como um barril de petróleo é igual a outro.

A profissionalização dos jornalistas se insere na “mentalidade de produção”. A indústria jornalística incorporou ao longo dos anos, assim como as outras, a “mentalidade de marketing”, mas o avanço desse modelo de gestão pode ser acelerado ou retardado de acordo com as características do mercado em que as companhias estão inseridas e das idiossincrasias da atividade. Há setores em que a mentalidade de marketing predomina, e há uma forte influência do comportamento das audiências no conteúdo, como a televisão ou revistas especializadas. Em outros, como o jornal impresso, essa correlação não é tão direta.

É questionável que a prática jornalística adote integralmente a visão do marketing, uma vez que seu papel vai além das necessidades imediatas daqueles que consomem diretamente notícias e envolve funções sociais como esclarecimento da população, fiscalização do poder público, formação de cidadãos, fortalecimento das liberdades e da democracia, entre outras (KOVACH; ROSENSTIEL, 2001). As redes de televisão, por exemplo, já possuem forte mentalidade de marketing, com grandes departamentos de pesquisa e desenvolvimento de produtos, direcionados para auscultar a audiência. No caso dos programas jornalísticos, porém, a resposta do público nem sempre é determinante, sendo colocada de lado em favor da relevância dos temas, segundo os critérios de noticiabilidade. De qualquer maneira, a forte concorrência promovida pelos meios digitais torna premente a preocupação com as necessidades e os critérios valorativos do público. Mais do que nunca, é uma questão de sobrevivência que as empresas jornalísticas se voltem para fora para identificar como podem se tornar mais relevantes para os consumidores de informação.

Picard (2013) afirma que as estratégias tradicionais adotadas pelas organizações jornalísticas não tem sido bem sucedidas na criação de valor para seus públicos e que a saída é buscar novas formas de produzir e distribuir conteúdos para se diferenciar e criar valor. O que se observa nas redações, porém, é um conflito entre a pressão para fazer o que todos os veículos concorrentes fazem, e dessa forma não correr o risco de “tomar furos e ficar para trás” e a disposição para assumir riscos e buscar outras fontes e enquadramentos.

Esta opção se torna mais rara com a recorrente substituição de profissionais experientes, com salários mais altos, por inexperientes, com vencimentos mais baixos, decorrente dos processos de *downsizing*, comuns nos últimos anos nas organizações jornalísticas brasileiras, que, sem encontrar alternativas mais criativas, buscam aumentar a produtividade “na marra”, reduzindo custos de pessoal e exigindo maior volume de produção em detrimento da qualidade de ambos. Pesquisa com 538 jornalistas do estado de São Paulo, mercado que reúne quase um terço dos profissionais do setor, mostra que as mudanças nos processos produtivos das organizações de mídia ocorridas desde a década de 1990 levaram à

precarização do trabalho jornalístico⁴³, ou seja, a redução dos empregos fixos com carteira assinada e o aumento da rotatividade. O estudo capitaneado por Roseli Fígaro, coordenadora do Centro de Pesquisa em Comunicação e Trabalho da Escola de Comunicações e Artes da Universidade de São Paulo (ECA/USP), aponta que esse fenômeno afetou negativamente a qualidade de vida dos profissionais, o que tem conseqüências sobre a qualidade do conteúdo jornalístico. Outro fenômeno identificado é a transferência de profissionais experientes das redações para consultorias e empresas de comunicação, que oferecem melhores salários e condições de trabalho. O principal problema é que, de acordo com o estudo, essa saída se dá sem a transferência de conhecimento para os mais novos, pois haveria um conflito entre as diferentes gerações como resultado não apenas da diferença de idade, mas principalmente da forma como as empresas promoveram a renovação de seus quadros, sem planejamento e de maneira apressada.

Jornalistas menos experientes tendem a ter menos conhecimento e autoconfiança para tomar decisões que contrariem o senso comum na cobertura jornalística, o que cria um ciclo vicioso de repetição de notícias e coberturas por diferentes veículos. O fenômeno favorece a popularização do “jornalismo declaratório”, aquele que é baseado apenas nas declarações de terceiros (RUBIM; COLLING, 2007), sejam de fontes ou *press releases*, exigindo o mínimo de processamento das informações por parte dos jornalistas, e resultando em notícias pobres em conteúdo e padronizadas. A tendência é tão forte que provocou uma reação dos jornalistas e levou à disseminação do conceito de “jornalismo investigativo” para definir o bom jornalismo, aquele que se volta para a checagem e o cruzamento de informações, a contextualização, a revelação de fatos e informações inéditas e socialmente relevantes e que se preocupa em contar boas histórias⁴⁴.

Pode-se afirmar, portanto, que as transformações econômicas e tecnológicas não foram bem assimiladas pelas redações tradicionais levando a uma queda de qualidade do conteúdo, que se torna padronizado, declaratório e superficial. Esse não é um fenômeno exclusivo do Brasil, pois, como observa Picard (2013), há um “mal-estar” generalizado entre os jornalistas como resultado da piora das condições de trabalho, da percepção da perda da relevância e da falta de perspectivas, pois não foi encontrado um caminho ou modelo bem sucedido que possa ser replicado para a transição das organizações de mídia a este novo cenário. E é provável que não haja apenas um.

⁴³ Mangini, Jussara. Pesquisa investiga mudanças no jornalismo e no perfil do jornalista. Agência FAPESP, 2014. Disponível em: <http://agencia.fapesp.br/18409>

⁴⁴ A própria Abraji, que carrega o termo em seu nome, deixa a definição em aberto.

2.3 Criação de valor no JGD

Uma empresa cria valor quando o valor de seus produtos e serviços é maior do que o dos insumos usados para produzi-los. Michael Porter tornou mais claro o entendimento deste fenômeno ao criar em 1985 o conceito de cadeia de valor, que desagrega os processos mais importantes dentro de uma organização buscando identificar a contribuição de cada etapa para o produto final e os custos e as possibilidades de diferenciação dos concorrentes. A empresa é vista como um sistema, formada por subsistemas, cada um consistindo de entradas (inputs), processos e saídas (output), que envolvem a aquisição e consumo de recursos, como mão-de-obra, capital, equipamentos, gerenciamento. Uma companhia tem vantagem competitiva sobre seus concorrentes se consegue desempenhar essas atividades melhor ou a um custo mais barato. A soma da cadeia de valor de uma empresa com a de seus fornecedores, canais de distribuição e clientes forma um sistema de valor.

As indústrias de mídia se dividem tradicionalmente em quatro cadeias de valor, quatro conjuntos de atividades a partir dos quais se cria valor, segundo Richeri (2012): produção, embalagem, distribuição e venda (tradução do autor)⁴⁵. Historicamente, as organizações jornalísticas tinham domínio sobre todo esse processo, desde a produção de textos, imagens e sons, passando pela impressão de periódicos e gravação de programas de rádio e televisão, sua distribuição e, no caso dos jornais e revistas, a venda aos consumidores finais (fogem a esse controle as bancas de jornal, mas se incluem as assinaturas). Esse valor era mantido por meio de um rigoroso controle sobre a distribuição do conteúdo, que garantia a monetização desse conteúdo, vendido às audiências, e da exclusividade de acesso a essa audiência, vendida aos anunciantes.

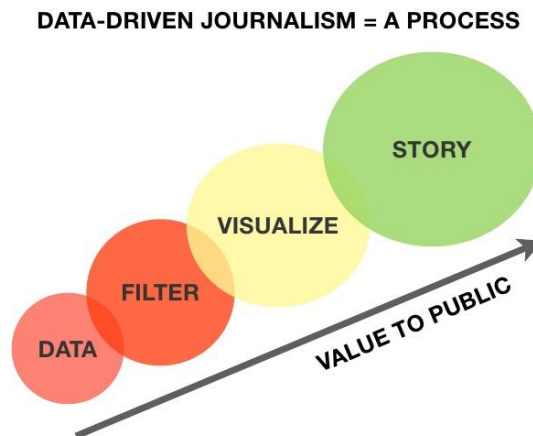
O cenário muda no ambiente digital, no qual, por força da convergência, as organizações de mídia passam a dividir esse controle sobre a cadeia com diversas outras empresas: de telecomunicações, responsáveis pela distribuição dos dados; de hardware, que fabricam os receptores desses dados (computadores, *tablets* e *smartphones*); de software e tecnologia, que fornecem as plataformas por meio das quais os dados serão processados, distribuídos e recebidos (sistemas operacionais, de análise de dados, processadores de texto, áudio e vídeo, etc.) e também processam esses conteúdos (sistemas de busca, redes sociais, plataformas multimídia); e com os próprios indivíduos que participam da criação de valor ao

⁴⁵ Do original “produzione, confezione, distribuzione, vendita”.

produzir, seleccionar, modificar e redistribuir esse conteúdo por meio de blogs, redes sociais e outros meios. Além disso, digitalização torna muito mais difícil o seu controle e, conseqüentemente, as organizações de mídia perdem valor.

O processo produtivo do JGD se caracteriza pelo fato de ser orientado pelas bases de dados, desde a coleta da informação, passando por sua manipulação e análise, até as diferentes formas de apresentação para as audiências. O diagrama elaborado por Lorenz (Figura 3) ilustra de forma bastante clara e simplificada as etapas de criação de valor nesse processo. O dado bruto é filtrado, apresentado de forma gráfica e visual e, por fim, transformado em uma narrativa.

Figura 3 - Processo de criação de valor no Jornalismo Guiado por Dados

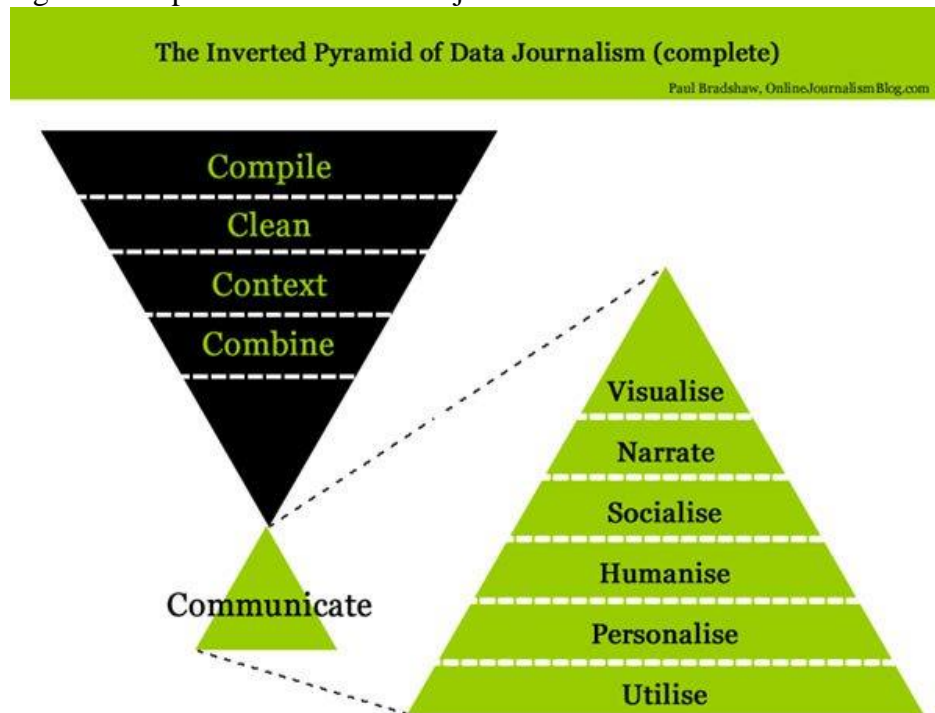


A criação de valor se inicia na coleta dos dados, uma vez que exige a aplicação de conhecimento e ferramentas especializadas e resulta em um material mais completo e refinado, que pode ser manipulado e combinado com outras bases de dados. A filtragem se assemelha ao processo de seleção, apenas com um volume muito maior de informações, o que significa dizer que não é facilmente realizado por aqueles que não têm conhecimento do tema e das ferramentas de análise. A visualização já envolve novas habilidades não só de programação, mas de design e infografia, além da habilidade narrativa para apresentar os dados de forma compreensível para as audiências. A história é a construção narrativa, que muitas vezes envolve o trabalho tradicional de reportagem, de fazer entrevistas, buscar personagens e novas informações provenientes de outras fontes que não as bases de dados.

Outro fluxograma elaborado por Bradshaw (2011b), batizado de “A pirâmide invertida do Jornalismo de Dados” (Figura 4), uma referência à técnica de redação jornalística que determina que as informações devem ser dispostas por ordem de importância na narrativa, detalha melhor os processos envolvidos na produção do JGD.

Há, portanto, um complexo trabalho de processamento e adequação dos dados para o consumo da audiência. Importante notar, porém, que cada etapa pode gerar um produto. Por exemplo, a compilação de bases de dados pode ser disponibilizada no site, como definido pela tipologia Fonte (BARBOSA; TORRES, 2013) e se tornar um produto, como na ProPublicaStore,⁴⁶ que vende bases de dados já trabalhadas e usadas nas reportagens da ProPublica.⁴⁷ Ou a comunicação pode parar na visualização dos dados, como em um Mashup, e assim por diante. Podem ser produzidos, portanto, diferentes conteúdos, com graus de valor distintos, de acordo com as etapas de processamento seguidas. ‘

Figura 4 - A pirâmide invertida do jornalismo de dados



Fonte: Bradshaw (2011b).

Outra questão relevante é que certas ferramentas e aplicativos se diferenciam das notícias tradicionais ao atrair audiência continuamente. Notícias podem até sobreviver como

⁴⁶ Ver em: <<https://projects.propublica.org/data-store/>>

⁴⁷ Ver em: <<http://www.propublica.org/>>

arquivo ou registro histórico, mas de maneira geral têm prazo de validade curto. No caso de uma ferramenta online, como o Basômetro, os leitores continuam voltando para fazer novas simulações. Outro exemplo é o aplicativo Rent or Buy⁴⁸ do jornal *The New York Times* não tem prazo, é sempre útil. É similar ao que ocorre com produtos de mídia voltados para o entretenimento, como filmes, músicas ou programas de televisão, que continuam sendo reproduzido mesmo décadas depois de prontos.

Com base na análise de projetos de JGD brasileiros e estrangeiros, foram identificadas novas formas por meio das quais a prática pode contribuir para a criação de valor nas organizações jornalísticas, seja por meio da renovação das funções que tradicionalmente criavam valor – acesso exclusivo a fontes, seleção e hierarquização de informações e distribuição eficiente –, seja por dar origem a novas funções e fontes de receita.

2.3.1 Acesso a fontes: dados públicos, informação exclusiva

O Jornalismo Guiado por Dados trabalha de maneira geral com bases de dados públicos, o que significa que o acesso não é restrito ao jornalista ou à organização de mídia e, portanto, a princípio não geraria valor. Existem, porém, diversas formas de se divulgar dados online, nem todas adequadas para a compreensão de pessoas não especializadas ou em formatos que permitam uma compreensão das informações e do tema a que se referem e nem todas estruturadas, ou seja, que possam seguir um modelo de organização pré-definido que possa ser lido e processado como uma base de dados por computadores. Não é incomum que governos e instituições forçados a publicar documentos sensíveis ou potencialmente prejudiciais àqueles que os controlam optem pelos métodos menos transparentes – como digitalizar textos em formato de imagem (impedindo que os softwares “entendam” as palavras), apresentar nas plataformas Java ou Flash (dificultando ou bloqueando a extração dos dados), usando *captcha*⁴⁹, entre outros. Como resultado, nem todos os dados públicos são efetivamente públicos, exigindo um conjunto de ferramentas e uma base de conhecimento específicas para serem manipulados e analisados.

⁴⁸ Ver em: <http://www.nytimes.com/interactive/business/buy-rent-calculator.html?_r=0/>

⁴⁹ Ferramenta que tem o objetivo de impedir a automatização do acesso aos dados, exibindo a imagem de conjunto de letras distorcidas – em tese compreensíveis apenas por humanos – que devem ser digitadas pelo usuário para que o acesso seja liberado.

O JGD exerce esse papel ao usar uma série de ferramentas e técnicas que permitem a extração de informações antes inacessíveis às pessoas comuns, como, por exemplo, raspagem⁵⁰, *crowdsourcing*⁵¹, *data mining*⁵², reconhecimento ótico de caracteres⁵³, mTurk⁵⁴, entre outros. Todas essas práticas exigem conhecimento, tempo e muitas vezes têm um custo, mesmo que a maior parte das ferramentas usadas seja gratuita. Este conjunto de práticas está na base da pirâmide invertida do jornalismo de dados (BRADSHAW, 2011b), associado às etapas de compilação e limpeza (correção de erros, padronização e adequação ao formato e modelo escolhido) de dados. É o primeiro passo para permitir a análise e compreensão que levará eventualmente à produção de sentido a partir desses dados.

Outra ferramenta, desta vez não tecnológica, é o uso das leis de acesso à informação para conseguir documentos e dados que não foram divulgados voluntariamente pelos governos e instituições – no caso do Brasil, a legislação existe desde 2011. Apesar de legalmente qualquer um poder fazer um pedido de informação com base na lei⁵⁵, nem sempre as pessoas têm conhecimento ou estão familiarizadas com o processo, que é bastante recente e desconhecido às vezes até por jornalistas. Ao aplicar esse conhecimento especializado, o jornalista produz informações exclusivas, uma vez que não são facilmente acessíveis. Esta não é uma técnica específica do JGD, mas é bastante usada por seus adeptos e, muitas vezes, exige a aplicação de técnicas e ferramentas como as descritas acima.

⁵⁰ Do inglês *web scraping*. Técnica de extração de dados de websites por meio da simulação do comportamento humano, que é quebrado em uma sequência de ações que são automatizadas por meio de algoritmos com o objetivo de transformar dados não estruturados em dados estruturados.

⁵¹ Nesse contexto pode ser entendido como a terceirização de tarefas, em geral mecânicas e tediosas, por companhias e instituições para um grupo externo e normalmente grande de pessoas por meio da internet, de forma colaborativa através de uma chamada pública

⁵² Análise automatizada ou semi-automatizada de grandes volumes de dados com o objetivo de encontrar padrões, usando métodos como inteligência artificial e estatística.

⁵³ Mais conhecido pela sigla em inglês OCR (Optical Character Recognition), é o uso de softwares para reconhecer de forma automatizada imagens e figuras gráficas transformando-os em textos e dados que podem ser processados por computador

⁵⁴ Amazon Mechanical Turk – ou mTurk – é uma plataforma de contratação de mão de obra online com 100 mil pessoas em 200 países, criada em 2005 pelo site de comércio online Amazon, na qual é possível encontrar pessoas para realizar tarefas simples por meio da internet. Entre os serviços oferecidos estão a coleta, compilação e redigitação de dados.

⁵⁵ Por meio do site <http://www.acessoinformacao.gov.br/acessoinformacaogov/>, mantido pela Controladoria Geral da União é possível fazer pedidos para qualquer órgão público brasileiro, bem como cobrar respostas.

2.3.2 Seleção e hierarquização de informações: pólo de dados confiáveis

Em um ambiente caracterizado pelo fluxo virtualmente ilimitado e não-hierarquizado de informações, como a web, selecionar aquelas que são relevantes e confiáveis é uma tarefa complexa, trabalhosa e, por isso mesmo, muito valorizada. Nesse contexto, a confiança é um bem escasso, afirmam Lorenz, Kayser-Bril e Mcghee (2011). As bases de dados de interesse públicos disponíveis online podem ser analisadas por diferentes instituições, que promovem a produção de sentido e divulgação das informações de acordo com agendas e interesses próprios. As organizações de mídia teriam então a credibilidade para analisá-las e interpretá-las segundo critérios reconhecidos e aceitos pela sociedade.

Um indício da demanda por informações confiáveis e pela confirmação de informações duvidosas é o sucesso de sites noticiosos voltados para a checagem de informações, como o Politifact (<http://www.politifact.com/>), criado pelo jornal norte-americano *Tampa Bay Times*, que é especializado na verificação de declarações feitas por políticos e venceu o prêmio Pulitzer, o mais importante de língua inglesa, em 2009. Há outros exemplos, como The Fact Checker (<http://www.washingtonpost.com/blogs/fact-checker/>), do *Washington Post*. O papel de gatekeeper é, portanto, substituído pelo de certificador de informações, ou pólo de dados confiáveis (Lorenz, Kayser-Bril e Mcghee, op.cit). E a objetividade dá lugar à transparência nos critérios adotados e nos interesses envolvidos nesse processo (HUNTER e WASSENHOVE, 2010). Há, dessa, uma mudança na função de selecionar e hierarquizar informações, que é mantido pelas organizações jornalísticas, mas não de forma a delimitar o acesso, mas chancelando as fontes e os dados e os colocando em contexto.

2.3.3 Distribuição: Personalização do conteúdo

Com a digitalização dos conteúdos e a popularização da internet, os meios de comunicação tradicionais perderam o monopólio da distribuição e passaram a enfrentar a concorrência de outras empresas do sistema de valor, que usam o conteúdo para potencializar e diferenciar seus produtos. Exemplos são as operadoras de telefonia e banda larga, que oferecem canais de televisão e serviços de informação pelo celular (Claro TV, Oi TV), os provedores de internet que fazem jornalismo (iG, Terra, Uol), empresas de tecnologia com

serviços de notícias (Google, Yahoo) e os próprios fabricantes de hardware, que investem em conteúdo para potencializar a venda de seus aparelhos eletrônicos e ter novas fontes de renda (Apple-iTunes, Sony Pictures, etc.), e os próprios consumidores que produzem conteúdo e o distribuem usando plataformas online.

Como se observa, a concorrência na distribuição se tornou muito mais vasta e diversificada, diminuindo o valor dos produtos das organizações de mídia tradicionais, em especial do jornalismo. Para retomar a relevância e a utilidade, as empresas precisam se diferenciar, oferecendo novos produtos e serviços que tenham mais valor para a audiência. O Jornalismo Guiado por Dados pode contribuir para essa estratégia ao se aproveitar das novas tecnologias e permitir a criação de novas formas de apresentação do conteúdo, como mashups, aplicativos, e criar novos produtos e serviços, como plataformas de distribuição de informações, aplicativos de serviços. O valor pode ser adicionado especialmente por meio da personalização, como no projeto Chicagocrime (<<http://crime.chicagotribune.com/>>), no qual é possível verificar o histórico de crimes cometidos em cada rua da cidade, o que se tornou inicialmente uma ferramenta de cobrança de políticas públicas de segurança, mas também passou a ser usado pelos moradores na hora de escolher um lugar para morar ou a escola dos filhos. Ou a iniciativa do jornal *Los Angeles Times*, também dos Estados Unidos, de mapear o tempo de atendimento a emergências pelo Corpo de Bombeiros e criar um aplicativo que permite a cada morador de Los Angeles verificar qual a demora para ser atendido em sua própria casa (<<http://graphics.latimes.com/how-fast-is-lafd/>>).

A personalização do conteúdo pode alcançar um nível ainda maior, se forem coletados os dados do usuário, como ocorre com a propaganda personalizada, característica de sites de busca e do comércio eletrônico, que sugere produtos de acordo com o histórico de pesquisas, compras e de navegação do usuário. Sabendo de antemão o perfil do usuário, os veículos jornalísticos podem não apenas sugerir reportagens teoricamente mais interessantes para cada perfil, mas apresentar conteúdos especificamente criados para cada pessoa. Uma ferramenta que analise bases de dados de saúde, por exemplo, poderia apresentar automaticamente as informações mais relevantes para o sexo, faixa etária e região geográfica do usuário.

2.3.4 Produção de sentido e contextualização de volumes de informações

Se a seleção e a hierarquização das informações de interesse da sociedade, que eram de domínio do jornalismo, passaram a ser feitas de maneira ampla e gratuita, retirando assim o seu valor, o mesmo não se pode dizer de bases de dados de interesse público, muitas vezes de difícil interpretação, seja pelo tamanho, a complexidade ou pelo tempo, ferramentas e habilidades que exige. Mesmo quando acessíveis a qualquer pessoa, a maioria não terá o conhecimento, a disposição e a capacidade para fazer esse trabalho.

A introdução de profissionais especializados em estatística eleva a interpretação das informações a outro nível, permitindo reconhecer padrões e identificar correlações em grandes volumes de dados. O domínio sobre a visualização gráfica de dados (infografia, mapas, design) pode, por sua vez, ajudar os jornalistas a perceber relações e ter uma compreensão holística das questões, ajudando a colocá-las em contexto. As atividades são possibilitadas e potencializadas pela programação, no desenvolvimento das ferramentas que auxiliarão todo esse trabalho.

Como mostrar, por exemplo, a real dimensão do que acontece na Amazônia, a maior floresta tropical do mundo, para uma audiência que não conhece a região? O jornalista Gustavo Faleiros, da organização de mídia sem fins lucrativos O Eco se sentia limitado pela tradicional combinação de texto e imagem usada no site, o que o levou a recorrer às ferramentas do JGD, criando o InfoAmazônia. A partir da combinação de imagens de satélites sobrepostas com várias camadas de informação – queimadas, unidades de conservação, urbanização, usinas, etc –, foi possível contextualizar os fenômenos que ocorrem naquele lugar. Com a criação de uma plataforma interativa, permitiu também o engajamento do público, que não só auxilia no entendimento sobre o assunto como amplia a criação de conteúdo.

A gente realmente cria um sentido ao que está um pouco disperso por aí. Essa idéia de mapas tem muito de conceitual por trás pela questão das camadas. Gerar camadas que tenham grande quantidade de dados por trás que podem ser lidas para gerar um contexto. Toda história por trás do InfoAmazônia é que seja uma ferramenta que, claro sirva para informar, mas que gere esse contexto mais aprofundado e de larga escala. Como é que eu consigo comunicar para um público que não está familiarizado com a Amazônia a dimensão de 7 milhões de quilômetros quadrados, nove países, a maior bacia hidrográfica do mundo? Como é que você visualiza a escala do que está acontecendo nessa região que é de suma importância global? A gente acredita que o InfoAmazônia, com todos os efeitos que ainda podem ser solucionados e todos os desafios, ainda possa trazer uma imagem, esse retrato da transformação que está ocorrendo na região. (FALEIROS, 2013)

Introduz-se dessa maneira um novo conjunto de habilidades no jornalismo, como se pode observar neste e no caso estudado. Habilidades que servem a uma determinada função e que não são dominadas pela maior parte das pessoas. Tornar a atividade multidisciplinar, por si só, já dificulta a execução, uma vez que, para fazer o mesmo seria necessário reunir um grupo de pessoas, o que não ocorre com o jornalismo tradicional, que sofre concorrência de blogs, redes sociais e outros meios.

2.3.5 Venda e licenciamento de aplicativos e ferramentas

Como foi mostrado, a prática do JGD pode resultar não apenas na criação de conteúdo, mas também de ferramentas e aplicativos que permitirão a organização, análise e visualização dos dados, a partir do uso de linguagens de programação pelos jornalistas ou pela inclusão de desenvolvedores de software no processo produtivo. O “efeito colateral” é o surgimento de novos produtos que podem gerar outras fontes de receita. Exemplo é o Guardian Open Platform⁵⁶, uma plataforma de dados abertos que oferece não só o conteúdo produzido pelo grupo de mídia The Guardian, como também os suportes de visualização e distribuição deste conteúdo, permitindo o desenvolvimento de aplicativos por terceiros. No Brasil, a equipe responsável pelo InfoAmazônia está mudando seu modelo de produção, adaptando a plataforma e negociando para terceiros: “A linha de produção que estou hoje é muito menos de notícias do que de desenvolver ferramentas e aplicativos que apóiem as notícias” (FALEIROS, 2013).

A partir do InfoAmazônia foi criado um laboratório de criação de produtos digitais, o EcoLab (<www.ecolab.org.br>). Seu primeiro produto foi um *template* para o site de hospedagem de blogs Wordpress chamado JEO, que permite a publicação de mapas de forma semelhante ao InfoAmazônia. O primeiro projeto a gerar receita foi um aplicativo desenvolvido para a organização de mídia Oxpeckers Center for Investigative Environmental Journalism,⁵⁷ que atua na África do Sul, para o monitoramento de caça a rinocerontes. Outros dois já foram elaborados a partir da plataforma JEO. O Ekuatorial,⁵⁸ feito para a Sociedade de

⁵⁶ Ver em: <<http://www.theguardian.com/open-platform>>

⁵⁷ <<http://oxpeckers.org/>>

⁵⁸ <<http://ekuatorial.com/en>>

Jornalistas Ambientais da Indonésia (SIEJ, em inglês)⁵⁹ e a Earth Journalism Network (EJN),⁶⁰ da Indonésia. Já o LandQuest,⁶¹ que mapeia os lucros do petróleo e de ajuda financeira entre a Europa e o Quênia, foi co-financiado pelo Sandbox Fund da African Media Initiative,⁶² e é liderado pela Internews no Quênia⁶³ em colaboração com o jornal espanhol *El Mundo*.⁶⁴ “O modelo que a gente criou agora é criar aplicações que a gente vende e gera receita para fazer os nossos próprios produtos” (FALEIROS, 2013).

2.4 Estratégias de criação de valor

Para lidar com o novo cenário de competição, as empresas de mídia devem adotar diferentes estratégias com o objetivo de maximizar o valor de seus produtos e, conseqüentemente, o retorno financeiro. Picard (2013) analisou as sete estratégias de criação de valor mais comuns adotadas pelas organizações jornalísticas: 1) Aumentar volume de notícias e informações; 2) Aumentar a velocidade de apuração e distribuição de notícias; 3) Aumentar a distância em relação aos fatos cobertos; 4) Fornecer notícias e informações exclusivas; 5) Fornecer notícias e informações personalizadas; 6) Fornecer notícias e informações multiplataforma; 7) Reembalar e reutilizar notícias e informações já existentes. E considerou seus impactos nas cinco principais partes interessadas: consumidores, anunciantes, investidores, jornalistas e a sociedade. Concluiu que a maioria falha em alcançar todas as partes ou perde seu efeito ao longo do tempo. A exceção é a especialização temática, que gera valor para todas as partes interessadas e tende a aumentar quanto maior é seu grau de aplicação. Com base nesse estudo sugeriu dez abordagens diferentes que acredita serem mais eficientes nesse contexto para aumentar o valor do conteúdo jornalístico:

- Enfatizar o jornalismo e o processamento de notícias
- Especializar-se ou centrar-se na localidade

⁵⁹ <<http://www.siej.or.id/>>

⁶⁰ <<http://earthjournalism.net>>

⁶¹ <<http://landquest.internewskenya.org/>>

⁶² <<https://www.facebook.com/pages/African-Media-Initiative-AMI/175369432567670>>

⁶³ <<https://www.facebook.com/internewsKE>>

⁶⁴ <<http://www.elmundo.es/>>

- Mudar ênfase da cobertura
- Escapar da enchente
- Tornar-se conselheiro confiável
- Usar as habilidades de certificação das organizações de mídia
- Enfatizar a facilidade de uso e conveniência
- Tornar-se o provedor principal em diferentes plataformas
- Fomentar a participação nos processos jornalísticos
- Criar experiências envolventes

Buscou-se verificar, a partir desse estudo, se a adoção das práticas do Jornalismo Guiado por Dados contribuiria ou não para o sucesso dessas estratégias dentro das organizações jornalísticas, com base nos exemplos observados e no estudo de caso. Pode-se afirmar que o JGD está ligado a quatro das estratégias mencionadas. “Escapar da enchente” se refere à produção de conteúdos específicos para cada público, chamados de *frontstream – feeds* personalizados de acordo com as levando em preferências e particularidades dos consumidores, como grupos socioeconômicos, estilos de vida – em oposição ao *backstream*, que seriam as notícias e informações brutas. O que se coaduna com a função de personalização oferecida pelo JGD.

“Mudar ênfase da cobertura” significa deixar de investir no aumento do volume e da velocidade de distribuição das notícias, que em última instância acabam as desvalorizando, o que seria uma medida quantitativa, e investir na qualificação das informações, por meio do processamento, contextualização e interpretação das notícias. Mais uma vez, o JGD tem entre suas potencialidades colocar em contexto grandes volumes de informação, e, por meio de conhecimento especializado, como o estatístico, pode levar a compreensão dos dados a outro nível. O JGD contribui com a estratégia de “Usar as habilidades de certificação das organizações de mídia” na medida em que promove a seleção e organização de grandes volumes de informação.

Outra estratégia que se coaduna com o JGD é “Especializar-se ou centrar-se na localidade”. A especialização temática e a concentração local são características presentes em diversos projetos de JGD, inclusive no Brasil, como o InfoAmazônia e o Retratos do Paraná (www.gazetadopovo.com.br/vidaecidadania/retratosparana/indicadores/), que reúne os principais indicadores do estado, oferece diversas formas de interação e reportagens baseadas nos dados. O recorte em nichos geográficos é favorecido pelo fato de os mapas e as ferramentas de visualização de dados geolocalizados estarem entre os recursos mais usados

pelos jornalistas guiados por dados. Já há inclusive um novo termo derivado da prática, o Geojornalismo⁶⁵.

Com relação aos temas, verifica-se uma certa especialização das equipes, que, à princípio, pode ter diversas razões não relacionadas às práticas do JGD. Uma delas é a própria especialização do veículo, como é o caso de O Eco, que originou o InfoAmazônia e é voltado para a cobertura de Meio Ambiente, e da *Gazeta do Povo*, que criou o Retratos do Paraná e é um jornal regional. Outra pode ser a origem dos jornalistas adeptos da prática, como é o caso do Estadão Dados, que foi iniciado na editoria de Política. Mas há indícios de que o JGD favorece a especialização, ao menos nessa etapa inicial de desenvolvimento da prática. Um fator de influência é a disponibilidade de dados e ferramentas. Há uma predominância de projetos voltados para a administração pública não só no Brasil como em outros países, por exemplo, como resultado dos movimentos de transparência governamental.

A exigência de conhecimento específico também pode ser relevante, pois, uma vez que o jornalista já está familiarizado com um determinado tema – quais as bases de dados disponíveis, como foram construídas, como interpretá-las, qual a confiabilidade, com quais fontes checar etc. –, é natural que fique inclinado a dar preferência ao assunto que domina. Mesmo porque, esse domínio é necessário para a criação de ferramentas mais sofisticadas e complexas. O Basômetro é um exemplo. Apenas um jornalista com um grau considerável de conhecimento do sistema político brasileiro e da dinâmica de poder entre o Congresso e o Executivo poderia concluir que uma das formas mais interessantes de se verificar o comportamento dos parlamentares era medir o seu grau de apoio ao governo federal. Nesse sentido, a limitação de recursos também deve ser considerada, pois, como se observa nos casos analisados, as equipes de dados ainda apresentam número reduzido. Dessa forma, faz mais sentido concentrar o trabalho em poucos temas, permitindo que sejam mais aprofundados e que haja maior produtividade. Por último, um fator determinante é que as bases de dados, por sua natureza, são relacionais e, quanto mais bases de dados forem coletadas e organizadas sobre um mesmo tema, mais ricas serão as possibilidades de análise e compreensão dos fenômenos a que se referem.

⁶⁵ Ver em: <<http://www.oeco.org.br/noticias/27613-o-eco-lab-lanca-guia-de-geojornalismo>>

3 ESTUDO DE CASO – ESTADÃO DADOS

3.1 Perfil do Grupo Estado

O Grupo Estado, controlado pela holding S.A. O Estado de S. Paulo, é um conglomerado de mídia brasileiro, sediado em São Paulo (SP), com atuação nas áreas de jornalismo impresso e online, rádio, televisão, logística, distribuição e serviços gráficos. Seu principal produto é o jornal diário *O Estado de S. Paulo*, um dos mais antigos e influentes do país, fundado em 1875. De 1966 a 2012, publicou também o diário *Jornal da Tarde*. A companhia, pertencente de forma majoritária à família Mesquita, controla um conjunto de empresas do qual se destacam: a Rádio Eldorado, responsável pelas rádios Estadão e Eldorado Brasil 3000, com abrangência no estado de São Paulo; a Agência Estado Ltda., maior agência de notícias do país, fundada em 1970; a OESP Mídia S.A., constituída em 1989, que tem como principais atividades a edição, produção e distribuição de listas telefônicas, guias e revistas empresariais, além de publicidade e serviços de Internet; e a São Paulo Distribuição e Logística Ltda., que presta serviços de gerenciamento de logística para distribuição de jornais e produtos editoriais. O grupo controla ainda o portal estadao.com.br, criado em 2000 a partir da união dos sites da Agência Estado, *O Estado de S. Paulo* e *Jornal da Tarde*. Em 2011, última data com informação divulgada, tinha 4.886 pessoas no quadro de funcionários.

O último balanço disponível, relativo a 2012, mostra que a companhia registrou receita líquida de R\$ 714,5 milhões, alta de 2% na comparação com os R\$ 700,7 milhões de 2011⁶⁶. O lucro líquido em 2012 foi de R\$ 38,7 milhões, similar ao de 2011. Apesar da diversificação das atividades, nota-se ainda uma concentração das receitas no meio impresso. O faturamento com a venda de espaço publicitário em 2012, considerando a soma dos jornais *O Estado de S. Paulo* e *Jornal da Tarde*, equivalia a uma participação de 12% do total nacional e de 32% do estado de São Paulo, de acordo com a companhia. Sabendo-se que o investimento publicitário no meio jornal em 2012 foi de R\$ 4,2 bilhões⁶⁷, pode-se inferir com base nos dados divulgados pela empresa que a receita com anúncios do impresso naquele ano foi de cerca de R\$ 504 milhões, o equivalente a 70,5% do faturamento total do grupo.

⁶⁶ Ver em: <<http://www.estadao.com.br/relatorio-da-administracao/DF%20Dez11%20SA%20consolidado%20com%20parecer.pdf>>

⁶⁷ Ver em: <<http://www.projetointermeios.com.br>>

O jornal *O Estado de S. Paulo* tinha em 2013 a quarta maior circulação do país, vendendo em média 232,3 mil exemplares diários, segundo o Instituto Verificador de Circulação (IVC), atrás do *SuperNotícia* (302,4 mil exemplares), de Minas Gerais; da *Folha de S. Paulo* (294,8 mil exemplares), de São Paulo, e de *O Globo* (267,5 mil exemplares), do Rio de Janeiro⁶⁸. A circulação de jornais cresceu significativamente nos últimos 10 anos no Brasil, um fenômeno contrário ao que se observa em outros países. De 6,9 milhões de exemplares diários, em 2003, para 8,8 milhões de exemplares por dia em 2012, representou um aumento de 25,5% na circulação das publicações, segundo estimativa da Associação Nacional de Jornais, baseada em dados do IVC. A receita publicitária dos jornais também aumentou, mas sua participação de mercado caiu, o que mostra o avanço da concorrência de outros meios, especialmente a internet. Em 2009, os jornais responderam por 14,1% do faturamento bruto com publicidade no Brasil, que somou R\$ 24,7 bilhões, segundo o projeto Inter-Meios⁶⁹. Já em 2012, o *market share* passou para 11,2%, de uma receita total de R\$ 37,7 bilhões⁷⁰. O faturamento do setor, portanto, subiu de R\$ 3,5 bilhões para R\$ 4,2 bilhões.

Apesar de os números mostrarem à primeira vista um cenário favorável ao mercado jornal na última década, o grupo Estado tem enfrentado dificuldade para ampliar suas receitas, que se mantiveram praticamente inalteradas desde 2005. Naquele ano, a receita líquida foi de R\$ 684 milhões, o que significa que, em sete anos, houve um crescimento de apenas 4,3% na receita, ante uma inflação de 49,9% medida pelo IPCA. Ou seja, houve uma perda real de receita de 26,5%. Ao fim de 2013, noticiou-se que o grupo teria sido colocado à venda, mas a informação não foi confirmada pela direção da empresa⁷¹.

A receita com publicidade online do grupo, com o portal estadao.com.br, não é divulgada. O site recebeu uma média de 85 milhões de *pageviews* por mês ao longo de 2012⁷², que o colocaram na 12ª posição entre os portais mais acessados e entre os três maiores de notícias. O faturamento publicitário do meio internet no país ainda não apresenta um crescimento forte em termos proporcionais, mas vem subindo de forma paulatina nos últimos

⁶⁸ Ver em: <<http://www.meioemensagem.com.br/home/midia/noticias/2014/01/27/Circulacao-de-jornais-cai-de-1-9-em-2013.html#ixzz2t1X6NOR9>>

⁶⁹ Ver em: <<http://www.projetointermeios.com.br>>

⁷⁰ Idem

⁷¹ Ver em: <<http://veja.abril.com.br/blog/radar-on-line/economia/braco-de-investimentos-do-itau-tenta-vender-o-estado-de-s-paulo/>>

⁷² Ver em: <<http://noticias.uol.com.br/politica/ultimas-noticias/2013/07/03/governo-beneficia-sites-progressistas-na-distribuicao-de-verbas.htm>>

anos, ao contrário do meio jornal. Segundo a pesquisa Inter-Meios, a participação das mídias digitais na receita total de publicidade cresceu de 4,3% para 5% entre 2009 e 2012. Em termos absolutos, no entanto, a alta foi significativa no período: de R\$ 1 bilhão para R\$ 1,8 bilhão, ou 80%.

Para tentar dar conta das mudanças ocorridas no mercado de mídia e enfrentar a concorrência com os meios digitais, o grupo começou em 2004 a fazer uma reformulação de seus processos produtivos⁷³ com o objetivo de ampliar a integração dos veículos e o conteúdo online. Dividido em três etapas, o plano foi elaborado pela consultoria catalã Cases i Associats, responsável pela mudança de outros jornais brasileiros, como *O Globo*. A fase 1 foi a reforma gráfica do Estadão, já em 2004. A fase 2, em janeiro de 2007, envolveu a mudança da arquitetura das redações, a integração das chefias de Estadão, JT e internet, além da implantação um estúdio de TV no meio da redação para transmissões online pelo estadao.com.br. Em 2008, foram iniciadas a fase 3 e o projeto Rumos, com o objetivo de revisar processos editoriais, potencializar a atuação multimídia e otimizar a alocação de recursos, contemplando um processo contínuo de produção e edição ao longo das 24 horas do dia. A terceira etapa alterou horários e rotinas dos jornais, aproximou as produções de todo o grupo e criou uma rádio-escuta integrada ao online.

Novas reformulações gráficas no jornal e no portal foram realizadas desde então. No caso do *site*, houve aumento e diversificação do conteúdo, especialmente por meio da inserção de produtos multimídia e blogs, páginas voltadas para nichos de audiência, além da criação de aplicativos para celular e *tablet*. Em dezembro de 2012 começou a vigorar o modelo *signwall*, no qual o internauta tem acesso livre a determinada quantidade de reportagens e artigos, tendo que fazer cadastro para leitura irrestrita do conteúdo disponibilizado pelo site, com direito a login e senha, mas sem ter que pagar por isso.⁷⁴ Em 2014 começará a cobrar pelo conteúdo online no modelo *pay wall*, já adotado por jornais como *The New York Times* e *The Washington Post* e, no Brasil, pelos sites de *O Globo* e *Folha de S. Paulo*.

A S.A. O Estado de S. Paulo é, portanto, um grupo de mídia tradicional e de grande porte que enfrenta desafios para ampliar sua audiência e receita motivados pelas mudanças tecnológicas e concorrenciais ocorridas a partir da popularização das mídias digitais e da internet. Sua estratégia aparenta dar maior peso ao meio online no processo produtivo, assim como ampliar, diversificar e qualificar o conteúdo, apostando nos pontos fortes de sua

⁷³ Ver em: http://www.observatoriodaimprensa.com.br/news/view/comunique-se__36524

⁷⁴ Ver em: <http://portal.comunique-se.com.br/index.php/entrevistas-e-especiais/73222-estadao-vai-cobrar-pelo-acesso-ao-conteudo-online-a-partir-de-2014>

cobertura, como as áreas de política e economia. É nesse contexto que se insere a criação de um núcleo de Jornalismo Guiado por Dados, dentro da editoria nacional, cujo conteúdo é integrado ao portal estadao.com.br.

3.2 O Estadão Dados

3.2.1 Perfil e contexto histórico

No Brasil, o primeiro grupo de mídia a criar um espaço em seu *site* dedicado ao conteúdo de Jornalismo Guiado por Dados foi a *Gazeta do Povo*, no início de 2012, com o projeto *Retratos do Paraná*, voltada para a divulgação de uma série de reportagens que desenharam o perfil do estado combinando o trabalho tradicional de reportagem com técnicas de coleta, análise e visualização de bases de dados públicas. No mesmo ano, a organização de mídia sem fins lucrativos O Eco, voltada para a produção e distribuição de conteúdo sobre o meio ambiente, lançou o projeto de JGD *InfoAmazônia*. Também em 2012, o Grupo Estado criou o primeiro núcleo multidisciplinar de profissionais voltado para a análise e a manipulação de bases de dados públicas com o objetivo de produzir jornalismo de precisão e *data-driven* (guiado por dados), o Estadão Dados. A produção do núcleo é destacada como uma nova seção do portal do Grupo na internet (acessado pelo endereço estadao.com.br), por meio da qual se acessa o blog da equipe. A primeira ferramenta baseada no conceito JGD desenvolvida pelo grupo foi o Basômetro⁷⁵. Todos os três exemplos se enquadram na tipologia Data/Dados do JGD, entendida como “denominação para editoriais/seções em alguns cibermeios nas quais são apresentados conteúdos produzidos a partir de dados disponibilizados pelos governos de países, estados, cidades, entre outros” (BARBOSA, 2012).

O Estadão Dados foi montado a partir da editoria Nacional, responsável pela cobertura de política no jornal *O Estado de S. Paulo*. A cobertura política sempre teve forte relação com o JGD desde sua emergência na última década, especialmente porque seu avanço ocorreu na esteira da tendência da divulgação de dados referentes à administração pública sob licenças abertas. Alguns dos exemplos mais destacados de coberturas guiadas por dados, que deram

⁷⁵ Ver em <http://estadaodados.com/basometro/>

notoriedade à prática, tiveram conotação política, como o escândalo das despesas dos parlamentares britânicos, pelo jornal *The Guardian*⁷⁶, em 2011. No Brasil, a produção de reportagens a partir de bases de dados da administração pública tem sido uma constante na última década, com o acompanhamento das despesas do Executivo, Legislativo e Judiciário, seja por meio de iniciativas oficiais – *sites* Portal da Transparência⁷⁷ e o Dados.gov.br, Lei de Acesso à Informação, entre outros – ou por esforço de organizações não governamentais, como Transparência Brasil⁷⁸ e Contas Abertas⁷⁹. Não é por acaso, portanto, que as primeiras iniciativas de inserção sistematizada do uso de bases de dados na produção jornalística tenham como objeto a cobertura política e de administração pública, já que oferecem o maior volume de matéria-prima para a elaboração de reportagens e são as áreas cuja cobertura concentra maior número de repórteres com mais experiência e conhecimento do manejo de bases de dados públicas.

O projeto Estadão Dados foi idealizado pelo editor e colunista político José Roberto de Toledo, um dos pioneiros no uso de bases de dados para a elaboração de reportagens no país, trabalhando desde a década de 1990 com técnicas de RAC, Jornalismo de Precisão e bases de dados públicas, especialmente na *Folha de S. Paulo*. Entre os anos de 1993 e 1994, começou “destrinchando” as tabelas das pesquisas do Instituto Datafolha⁸⁰. Nos anos seguintes, se debruçou sobre os dados estatísticos do Censo de 1991, que pela primeira vez eram disponibilizados online pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Também publicou um caderno especial que apresentava um indicador sintético de qualidade de vida para os 96 distritos paulistanos, criado em parceria com a universidade PUC-SP. Nesse época, o trabalho com dados era feito nos horários vagos, pois era editora da coluna Painel. Em 1996, pediu para deixar a função e se dedicar exclusivamente ao RAC e a “mastigar números” (TOLEDO, 2013).

Toledo se tornou referência no jornalismo com auxílio de computador no país e foi, em 2002, um dos fundadores da Associação Brasileira de Jornalismo Investigativo (Abraji),

⁷⁶ Ver em: www.guardian.co.uk/politics/mps-expenses

⁷⁷ Ver em: www.portaltransparencia.gov.br

⁷⁸ Ver em: www.transparencia.org.br

⁷⁹ Ver em: www.contasabertas.com.br/WebSite

⁸⁰ Criado em 1983, como departamento de pesquisas e informática do Grupo Folha da Manhã, controlador do jornal *Folha de S. Paulo*, o instituto de pesquisas Datafolha foi pioneiro no país, ao produzir conteúdo e servir como ferramenta de planejamento para veículos jornalísticos. Hoje é uma empresa independente e se tornou um dos mais importantes institutos de pesquisa de opinião do Brasil.

uma das principais promotoras do RAC no país, da qual é vice-presidente. Desde então, treinou mais de 1000 jornalistas em cursos de RAC promovidos pela Abraji, pelo Knight Center for Journalism in the Americas, pelo Master em Jornalismo da Fundação Getúlio Vargas, entre outras instituições. Se tornou colunista do Estado de S. Paulo em 2009, produzindo também o blog Vox Publica (<www.voxpublica.com.br>) no portal Estadão.

Para criar o Estadão Dados, se inspirou nas equipes de jornalismo guiado por dados criadas por organizações de mídia européias e norte-americanas, após participar do 11º Simpósio Internacional de Jornalismo Online, realizado pelo Centro Knight para o Jornalismo nas Américas, da Universidade do Texas, Austin, e uma visita na mesma época à redação do *The New York Times*, em Nova York. Como afirma o próprio jornalista, ao observar a experiência dos jornais estrangeiros, percebeu que o que faziam era diferente da prática tradicional de RAC e da simples inserção das bases de dados nas reportagens:

A linguagem que eles falavam era diferente, era de código. Todos eles codificavam ou tinham alguém em sua equipe que codificavam. (...) Percebi que fazer apenas o que a gente fazia, usar bancos de dados, planilhas, Excel, Open Office, não ia nos levar muito além de onde a gente tinha chegado. Para tentar imitar o que o *New York Times* estava fazendo, o que o *Guardian* estava fazendo, a agente ia precisar mudar a cultura mesmo. E essa mudança de cultura implicava necessariamente incorporar a programação ao jornalismo. (TOLEDO; TRASEL; ALVES, 2013)

As limitações orçamentárias e de mão-de-obra disponível, além das características locais não permitiriam a simples reprodução dos modelos adotados por redações estrangeiras – já naquela época a equipe voltada para a produção de ferramentas interativas e análise de dados do *The New York Times* somava mais de 20 pessoas –, o que exigia uma adaptação daquele modelo de trabalho para a realidade local. A proposta de Toledo de criar um núcleo a partir dos profissionais existentes foi bem aceita pela direção do jornal, que deslocou um repórter da editoria Nacional, um designer/infografista da editoria de Arte e um programador do estadao.com.br. Posteriormente foi integrada à equipe uma repórter exclusiva, Amanda Rossi, que acabou saindo por motivos pessoais. O primeiro projeto desenvolvido foi o Basômetro, ferramenta lançada em maio de 2012 que mede o apoio de parlamentares ao governo federal, e até hoje a mais complexa elaborada pelo núcleo.

Desde então foram criadas mais sete ferramentas interativas, chamadas de projetos especiais, e mais de duas centenas de gráficos, *posts* e reportagens publicadas no <<http://blog.estadaodados.com/>>, no portal estadao.com.br e no jornal *O Estado de S. Paulo*.

A criação do núcleo teve caráter de experimentação. Não visava nenhum resultado econômico ou aumento de receita. Também não havia nenhuma meta ou expectativa de

retorno de audiência. Ou seja, não foram exigidas contrapartidas imediatas e nem adotadas metas financeiras. O objetivo era o interesse público, ao introduzir um conjunto de habilidades e práticas inovadoras de produção jornalística para revelar novas informações e criar melhor conteúdo. Toledo acredita, porém, que adoção do JGD é inevitável:

É uma necessidade. Se o jornalismo quiser compreender o mundo de dados que são gerados diariamente ele precisa se adaptar a isso. Precisa emprestar as ferramentas necessárias para dar conta do trabalho. O jornalismo tradicional, de bloco e gravador, continua e continuará sendo necessário, mas ele não dá conta sozinho de atender essa nova demanda (TOLEDO, 2013).

3.2.2 Diferenças em relação às práticas tradicionais

Apesar de estar integrado à redação do jornal *O Estado de S. Paulo* e do portal estadao.com.br, o núcleo Estadão Dados tem diferenças em relação à formação da equipe e a dinâmica de produção. Além de produzir conteúdo, o núcleo Estadão Dados é responsável também por dar treinamento formal aos jornalistas iniciantes (focas) e informal aos repórteres que já atuam na redação, a pedido. O grupo também trabalha em conjunto com jornalistas de outras editorias em reportagens que exigem maior especialização para a análise e manipulação das bases de dados. A equipe atua, portanto, de forma transversal com outros núcleos de trabalho na redação, tanto na produção de conteúdo próprio, a exemplo do que ocorre normalmente em qualquer editoria, quanto prestando serviços ao público interno da redação – de maneira similar, mas não igual, ao que ocorre com áreas como arte, fotografia e diagramação, pois atua majoritariamente de forma independente e não acessória a outras editorias – além de fornecer treinamento para este mesmo público. Dessa forma, pode-se inferir que a dinâmica das funções do núcleo Estadão Dados é mais complexa que a de outros núcleos editoriais tradicionais.

A característica que talvez mais se destaque em relação a outras editorias é o caráter multidisciplinar. Além do coordenado, o núcleo é formado por um programador, Diego Rabatone, e três jornalistas, sendo que todos têm alguma noção de programação e visualização de dados. Um repórter é exclusivo e tem especialização em estatística, Lucas de Abreu Maia, um está ligado à editoria de Cidades/Metrópole, Rodrigo Burgarelli, e o terceiro, que esteve desde o início do núcleo é da editoria Nacional, Daniel Bramatti. A equipe fica localizada em uma sala com paredes de vidro anexa à redação. Esse tipo de sala é conhecida

popularmente entre os jornalistas como aquário. Nela ficam em horário integral o repórter Lucas e o programador Diego Rabatone. Bramatti ocupa uma mesa junto com os outros jornalistas da editoria Nacional, a qual continua vinculado. O mesmo acontece com Burgarelli.

A possibilidade de produção de sentido a partir de grandes volumes de dados é claramente percebida no núcleo Estadão Dados como um diferencial em relação à reportagem tradicional. O volume de informação com que trabalham é notadamente maior do que o manipulado no cotidiano pelos repórteres, assim como as habilidades e ferramentas necessárias para processar e analisar os dados de forma a se criar um entendimento, ainda que parcial, do que significam.

O jornalismo de dados é uma adaptação do RAC aos peta bytes de dados que o mundo produz diariamente. Não cabem mais numa planilha de Excel, travam programas de banco de dados mais simples como Access. É preciso recorrer a programadores/desenvolvedores, estatísticos, designers. É, por isso, um trabalho de equipe – acho (TOLEDO, 2013).

Coincidiu também com esse avanço tecnológico uma predisposição dos governos de abrir suas bases, o impulso da lei de acesso à informação, então temos muitos dados brutos que precisam ser lapidados. É um trabalho muito difícil, mas quando a gente consegue saem coisas muito legais. A gente fez isso com bases de convênios, por exemplo, que tinham 500 mil, 600 mil contratos para serem analisados e com a ferramenta X, Y você conseguia fazer uma leitura do que aquilo significava em termos macro. Para mim é isso, a diferença básica em relação à reportagem tradicional, em vez de você ficar naquela de contar com as pessoas, pegar o telefone e tal, você conseguir extrair a informação de números que são brutos e que não revelam nada se você não souber por onde abordá-los (BRAMATTI, 2013).

A partir de relatos dos jornalistas e da análise de conteúdo das reportagens produzidas pelo núcleo e as elaboradas tradicionalmente e publicadas pelo jornal O Estado de S. Paulo, pode-se afirmar que há diferenças na dinâmica de produção, nas habilidades e ferramentas utilizadas, nas fontes que originam as pautas e na abordagem científica adotada, além, é claro, do uso de bases de dados como fonte primária ou como principal apoio de argumentação em todas as matérias. A seguir, mostra-se o processo de produção do núcleo, buscando-se evidenciar as particularidades e as diferenças em relação as práticas tradicionais, separado por três etapas: pauta, produção, formatação e distribuição.

3.2.2.1 Pauta

A pauta das reportagens é independente e não está necessariamente relacionada com a do restante do jornal. As editorias, no entanto, buscam o núcleo para a realização de reportagens com dados que exijam maior conhecimento específico. O programador participa da pauta, apóia a apuração e os jornalistas também programam. As bases de dados públicas são a principal fonte para as pautas. A maior parte do conteúdo online produzido usa exclusivamente bases de dados e não inclui apuração tradicional, com entrevistas e “ida à campo”. Tradicionalmente nas redações dos grandes jornais e sites de notícias são feitas ao menos duas reuniões de pauta por dia. No caso de *O Estado de S. Paulo*, diariamente há uma reunião de abertura, com os pauteiros/chefes de reportagem ou subeditores de cada editoria, em geral feita às 10h, para definir a pauta do dia. Outra na parte da tarde, por volta das 14h, para relatar o andamento das reportagens e fazer correções e mudanças na cobertura, e mais uma por volta das 18h para definir a manchete e os temas de capa, com a participação da direção da redação e dos editores. Caso não haja nenhum fato que exija a postergação, entre 20h30 e 21h, a primeira edição é fechada e enviada para a gráfica. Às 23h, encerra-se a segunda edição. Na segunda-feira, é feita ainda uma reunião das pautas especiais e de fim de semana, em geral com a participação dos repórteres, e novamente na sexta-feira, após o primeiro fechamento, é feita uma reunião com os editores para definir as edições de domingo e segunda-feira, já que nos fins de semana é adotado esquema de plantão, com menos pessoas na redação.

No caso dos meios online, a dinâmica é diferente, pois não há um prazo determinado para o fechamento de cada edição de forma a permitir a impressão e a distribuição aos assinantes e bancas até a manhã do dia seguinte. Ainda assim, é comum fazer reuniões diárias para definir a cobertura e os temas que serão explorados naquele dia, além de reuniões semanais para pautas especiais. No caso do Estadão Dados, não há uma rotina formal de reuniões de pauta, que acontecem, mas não diariamente ou semanalmente, e sim por demanda de projetos. Boa parte da comunicação é feita por meios online, especialmente com o editor, que não fica diariamente na redação.

Os aplicativos e ferramentas, que exigem maior tempo de desenvolvimento e dedicação, são tratados como projetos especiais, a exemplo do que ocorre no meio impresso com edições temáticas ou históricas, sem serem incluídas na agenda de publicação, com prazos pré-determinados. Não há uma determinação de prazos e volume de produção para o

núcleo, mas há uma meta informal de publicar ao menos matéria por semana no jornal impresso, o que ocorre normalmente aos domingos, uma vez que a edição é preparada com antecedência, tem tamanho maior que as publicadas durante a semana e mais espaço para reportagens mais “frias”, ou seja, que não requer publicação imediata e é apresentada normalmente de forma menos factual e mais contextualizada.

As reportagens começam a ser elaboradas de maneira geral a partir da disponibilidade das bases de dados e não a partir da pré-apuração com fontes. Segundo Bramatti (2013), os dados públicos podem suprir a dependência ou a carência do relacionamento com fontes, que resulta em dicas e informações, e dão elementos para que se inicie a apuração de uma reportagem. À primeira vista, iniciar a reportagem com uma base de dados pode parecer inerente à prática do JGD, mas nem sempre é assim. Pode-se identificar um problema ou uma história e só depois buscar os dados que permitirão entendê-lo ou contá-la da melhor forma possível. Importante ressaltar que o JGD não prescinde da reportagem tradicional, de ir para a rua, ouvir fontes, conversar com pessoas, buscar as histórias e pessoas por trás dos números.

Os dados podem estar disponíveis em formatos que permitam seu processamento, como as extensões .csv ou .xls, mas muitas vezes exigem o uso de ferramentas raspagem, a conversão e reformatação dos arquivos, ou até mesmo a tabulação manual do conteúdo, o que é mais raro. A definição das ferramentas que serão usadas é feita a partir dos dados. Quando uma nova técnica ou ferramenta é aprendida, porém, os jornalistas podem buscar novas bases de dados que se adequem ou voltar a dados utilizados anteriormente de outra maneira, em um processo de experimentação.

Há um predomínio inicial de temas políticos, devido à origem do núcleo ser a editoria Nacional, mas observa-se ao longo do tempo uma maior variação para temas ligados à cidade de São Paulo – há um jornalista da editoria de Cidades no núcleo, Rodrigo Burgarelli – e a temas socioeconômicos, como saúde, educação, mortalidade, qualidade de vida, especialmente por conta da riqueza dos dados disponibilizados por fontes como o IBGE. Pode-se notar, então, uma clara influência da disponibilidade de dados na seleção das pautas, além da ligação editorial dos profissionais.

3.2.2.2 Produção

O núcleo tem como meta publicar ao menos uma visualização por dia no blog. Em geral, as visualizações publicadas diariamente online são simples, mapas e gráficos elaborados a partir normalmente de uma fonte de informação, sem o auxílio de reportagem tradicional. Além dos gráficos diários, produzem também ferramentas com diferentes graus de sofisticação, sem um prazo pré-definido, mas de acordo com a demanda e a disponibilidade de tempo e pessoal. Nesse caso, relacionam diferentes bases de dados de fontes distintas, usam recursos como mapas e gráficos interativos e *mashups*.

Além disso, todos os integrantes produzem reportagens para o impresso. No caso das matérias para o jornal seguem uma estrutura narrativa mais próxima da tradicional, com texto escrito em pirâmide invertida e apuração com fontes e autoridades ligadas ao tema. Os resultados são apresentados em textos corridos e infográficos estáticos, mas a coleta e análise dos dados, no entanto, é feita com o uso das técnicas e ferramentas de JGD, mas no jornal. Normalmente, há gráficos dinâmicos com os mesmos dados no website. Segue-se de maneira geral o processo descrito por Bradshaw (2011b) como “pirâmide invertida do jornalismo de dados”.

A produção das reportagens é feita majoritariamente de forma coletiva. Com exceção dos gráficos mais simples publicados na seção “Gráfico do dia” no blog, a equipe tende a trabalhar em conjunto, combinando os conhecimentos de programação, visualização e a apuração tradicional: “Geralmente pelo menos uma coisinha sempre tem alguma ajuda, sempre acaba virando um trabalho de equipe” (BRAMATTI, 2013).

O uso de bases de dados também permite adotar abordagens menos especulativas e mais científicas na apuração das informações, com a elaboração e o teste de hipóteses, de forma bem similar ao que preconiza o Jornalismo de Precisão (MEYER, 1973). Idéias consolidadas entre as fontes e o senso comum são contrastadas com os números de maneira a refutá-las ou comprová-las. Esse é um dos fundamentos do Estadão Dados, segundo seus realizadores, que pode ser observado nas reportagens produzidas pelo núcleo. A análise estatística é uma das bases de conhecimento usadas pelo JGD que qualifica a informação jornalística, tornando-a mais próxima ou equivalente à pesquisa científica, e é uma ferramenta indispensável para a produção de sentido a partir de grandes volumes de dados. Mesmo antes da chegada de um jornalista especializado à equipe já havia uma tendência a se adotar métodos mais científicos, na forma como é pretendido pelo Jornalismo de Precisão.

Um exemplo de aplicação da análise estatística é a reportagem *Mortalidade infantil está diretamente associada à falta de estudo dos pais*⁸¹, publicada em 26 de agosto de 2013 no jornal O Estado de S. Paulo e no site estadao.com.br. A partir da análise de 232 variáveis dos 5.565 municípios brasileiros coletadas no Censo 2010, a equipe do Estadão Dados concluiu que o baixo nível educacional dos pais é o principal fator que contribui para a mortalidade infantil, mais do que a pobreza ou falta de saneamento. A conclusão é surpreendente e pode inclusive balizar políticas públicas de uma maneira que não seria possível para uma reportagem tradicional, na qual essa informação certamente teria como fonte um estudo elaborado pelo próprio governo ou por organizações ligadas ao tema.

Uma das principais vantagens seria quebrar a dependência das fontes tradicionais de informação, tema discutido por Parasie e Dagiral (2013). Bramatti afirma que já usava esse recurso – recorrer aos dados por falta de acesso às fontes oficiais – ao produzir reportagens para a editoria de Política, pois ficava baseado em São Paulo, distante das fontes localizadas em Brasília. O uso de dados, associado a uma abordagem científica, como preconizada pelo Jornalismo de Precisão (MEYER, 1973), permite também fugir do jornalismo declaratório (RUBIM; COLLING, 2007) e desconstruir mitos, provocando novos entendimentos sobre diferentes assuntos, como afirma Bramatti:

É quase como um método científico: você levanta uma hipótese e aí vai e tenta comprovar ou derrubar. Eu sempre tive interesse em fazer jornalismo dessa maneira e eu já conhecia o Toledo, já tínhamos trabalhado juntos e ele me chamou e a gente começou a trabalhar na primeira ferramenta, que é a mais importante até hoje, que é o Basômetro (BRAMATTI, 2013)

3.2.2.3 Formatação

Os conteúdos produzidos pelo Estadão Dados são publicados em diferentes formatos. Nota-se uma diferença nas formas narrativas usadas pelo núcleo, além da tradicional composição texto/foto/gráfico estáticos do jornal impresso, que são características do JGD: os *mashups*, mapas e visualizações interativas que combinam diferentes bases de dados para criar um novo sentido àquelas informações e permitir que o próprio usuário desenvolva sua narrativa de forma independente. Exemplo é o projeto *São Paulo que balança*, que combina

⁸¹ Ver em: <http://www.estadao.com.br/noticias/impresso,mortalidade-infantil-esta-diretamente-associada-a-falta-de-estudo-dos-pais,1067807,0.htm>

os dados de voto em cada distrito da capital com as informações sócio-demográficas destes locais evidenciando uma clara distinção entre os moradores da periferia, que votam nos candidatos do PT, e os do centro expandido, que não votam no partido. O usuário pode visualizar os dados de cada distrito e até encontrar a sua própria rua no mapa.

Exemplo é a reportagem *Taxa de atraso de voos cai, mas tempo de viagem aumenta*⁸², publicada em 14 de dezembro de 2013 em O Estado de S. Paulo, cujos gráficos aparecem no blog, e que foi transformada em um videográfico publicado no site Estadão⁸³. Os videográficos combinam a visualização gráfica com a narrativa audiovisual. Gráficos dinâmicos em três dimensões são apresentados em uma ordem linear pré-determinada e seu conteúdo é descrito e analisado por um jornalista em gravação off. É interessante notar que, ao usar ferramentas digitais e interativas, o núcleo permite e facilita a análise das informações pelas próprias audiências:

Não parte-se de uma premissa, ou uma ideia pré-concebida. Parte-se dos dados brutos e, depois de muita depuração e consultas, chega-se a uma ou mais histórias que eles têm para contar. Isso não exige o jornalismo de dados de ter viés. Todos temos, mais de um. Mas permite um ângulo mais científico, no qual as conclusões podem ser demonstradas e reproduzidas - como na ciência. Também permite criar ferramentas online ou apps em que o usuário vai ter acesso a uma interface amigável para fazer suas próprias perguntas à base de dados, extrair suas próprias respostas e contar a história para si, sem intermediação do jornalista (TOLEDO, 2013).

3.2.2.4 Distribuição

Esse conteúdo é distribuído em diferentes plataformas, tanto online, através do site estadao.com.br ou do blog do núcleo, quanto pelo jornal impresso, e em diferentes formatos, como narrativas textuais, mapas, ferramentas interativas e videográficos. No estadao.com.br há 297 menções ao Estadão Dados, enquanto na TV Estadão, há outras sete.

As publicações no blog do núcleo são divididas em três seções: “Gráfico do dia”, onde são publicadas visualizações de dados de diversas fontes diariamente, com mais de 120 gráficos; “Permanentes”, com mashups, gráficos e visualizações de séries estatísticas que são atualizadas periodicamente; e “Séries especiais”, com visualizações de temas específicos. Há

⁸² Ver em: <www.estadao.com.br/noticias/cidades,taxa-de-atraso-de-voos-cai-mas-tempo-de-viagem-aumenta,1108655,0.htm>

⁸³ Ver em: <<http://tv.estadao.com.br/videos,VOOS-NO-BRASIL-ATRASAM-CADA-VEZ-MENOS-MAS-TEMPO-DE-VIAGEM-AUMENTA-ANO-A-ANO,220643,250,0.htm>>

ainda a área “96XSP”, com mapas dos 96 distritos da capital paulista, elaborados a partir de tabulações especiais da amostra do Censo 2010 feitas pelo Ibope em parceria com o Estadão Dados. Por último, a seção “Projetos especiais”, voltada para as ferramentas mais sofisticadas, incluindo o *Basômetro* e mais sete projetos abrangendo temas variados: *Ringue 2014* (potencial de votos de candidatos nas eleições presidenciais), *Coligações* (visualização interativa das alianças entre os partidos brasileiros nas eleições de 2012), *Eleições 2012* (visualização interativa dos resultados das eleições de 2012), *São Paulo que balança* (ferramenta interativa relacionando os eleitores que votam e não votam no PT com os distritos da cidade de São Paulo e dados socioeconômicos), *Cotas* (gráfico animado com dados de pesquisa de opinião sobre cotas), *Religiões* (visualizações dos dados sobre religiosidade nos estados brasileiros coletados no Censo 2010) e *Biomassas* (visualizações de dados de desmatamento nos biomas brasileiros em 2012).

3.2.3 Impacto no impresso

Parte integrante da redação do jornal *O Estado de S. Paulo*, o núcleo publica regularmente reportagens no impresso. Desde que foi criado até 22 de fevereiro, o Estadão Dados foi mencionado sete vezes na manchete do jornal e outras 103 vezes nos cadernos internos do periódico. Na análise foi considerada apenas a edição São Paulo, que circula no estado, pois é fechada por último e normalmente é mais completa que a Edição Brasil, que é distribuída para os outros estados. As matérias creditadas ao núcleo apareceram no caderno de Política 59 vezes, em Metrópole/Cidades 18 vezes, Economia, 9 vezes, Geral, 6 vezes, Informática outras 3 e uma em Internacional.

Esse número, porém, não reflete o total de matérias feitas pelos jornalistas do núcleo, e mesmo a partir de bases de dados, uma vez que nem todas são atribuídas nominalmente ao Estadão Dados, e sim aos repórteres. No caso das capas, muitas vezes não há a identificação dos repórteres, que só aparece internamente. Se forem consideradas as menções a José Roberto Toledo, Daniel Brammati e Rodrigo Burgarelli, os números ficariam distorcidos, pois eles produzem conteúdo não relacionado ao núcleo. A verificação completa da produção do Estadão Dados veiculada no impresso exigiria, então, a análise individual de cada reportagem, o que não é factível e nem é o objetivo desta pesquisa. Para apresentar uma quantificação mais aproximada, foram incluídas as menções aos integrantes exclusivos do núcleo (Amanda

Rossi e Lucas de Abreu Maia), excluindo-se, porém, os casos que também mencionam o Estadão Dados, para evitar redundâncias.

Tabela 1 - Menções ao Estadão Dados e integrantes exclusivos no jornal impresso O Estado de S. Paulo (Mai/2012 a Fev/2014)

Jornalistas\Editoria	Política	Metrópole	Economia	Geral	Link	Editorial	Capa	Total
Estadão Dados	59	18	9	6	3	1	7	103
Amanda Rossi	22	3	0	4	1	0	1	31
Lucas de Abreu Maia	38	0	0	0	0	0	0	38
Total	119	21	9	10	4	1	8	172

Fonte: elaboração do autor

3.2.4 Basômetro

Lançado em 12 de maio de 2012, o Basômetro é o primeiro projeto do Estadão Dados. Trata-se de uma ferramenta digital interativa que tem o objetivo de verificar o nível de governismo do Congresso, ou seja, quantas vezes os parlamentares votam de acordo com a orientação do governo federal. Para isso, permite a visualização dos dados de 216 votações da Câmara dos Deputados, entre 28/6/2011 e 11/12/2013, e 157 votações do Senado, entre 23/2/2011 e 17/12/2013.

Visualmente, apresenta um gráfico de eixos cartesianos simples. No eixo vertical, os pontos são ordenados de acordo com o percentual de vezes em que o parlamentar votou com o governo (percentual de governismo). No eixo horizontal, seguem a ordem dos partidos aos quais os parlamentares pertencem e, dentro de um mesmo partido, a ordem alfabética. Os parlamentares são representados por bolas coloridas localizadas nos eixos de acordo com essas duas informações. Cada partido tem uma cor. O gráfico é dinâmico e permite selecionar individualmente cada ponto nos eixos. Ao clicar no ponto, aparece uma janela com a foto do político, seu percentual de governismo e os números de votos contra e a favor do governo, abstenções e faltas, mostrando o percentual de governismo dos parlamentares de cada estado. Um botão deslizável (slider) pode ser arrastado para cima ou para baixo, permitindo a visualização em linha da taxa de governismo.

Figura 5 - Reprodução da ferramenta Basômetro



Fonte: Site do Basômetro (<<http://estadaodados.com/basometro/>>)

Na barra superior, há quatro menus e uma caixa para pesquisa. No primeiro, é possível selecionar uma das 11 ferramentas JGD disponíveis; no segundo, qual gestão é abrangida pelos dados; no terceiro, se os votos são da Câmara ou Senado; e, no último, se os votos são por bancada ou individuais. A caixa de pesquisa permite buscar políticos pelo nome. Do lado esquerdo, uma barra vertical mostra os 18 partidos do Congresso e os dados agregados com seus respectivos percentuais de governismo. É possível gerar gráficos para cada partido, além de três bancadas: ruralistas, evangélicos e ambientalistas. Do lado direito, outra barra reúne dados agregados das 27 unidades federativas representadas na Câmara e no Senado. Ao clicar em cada uma, um gráfico é gerado.

Abaixo do gráfico, uma barra permite selecionar o período das votações a partir da qual o gráfico comparativo é gerado. Pode-se selecionar o recorte de tempo e, apertando a tecla play, dá-se início a uma animação mostrando o comportamento dos parlamentares nas votações ocorridas ao longo do período selecionado.

Os dados usados são abertos e disponibilizados pela equipe por meio dos sites Github e Google Docs. O Basômetro foi inspirado no projeto Camaraws, de Leonardo Leite e Saulo Trento (PoliGNU - Grupo de Estudos de Software Livre da Poli-USP), que usou os resultados de votações na Câmara dos Deputados para avaliar o grau de semelhança entre os partidos. O site do Basômetro tem ainda dois vídeos, um tutorial e um mostrando o encolhimento do núcleo duro durante o governo Dilma, além de links para as reportagens e análises feitas a partir da ferramenta, além de um texto explicativo.

Do ponto de vista narrativo, há uma grande modificação, que se coaduna com a mudança nas convenções narrativas a partir da digitalização e da emergência das bases de dados, como identificado por Manovich (2001) e aplicado ao jornalismo por Barbosa (2007). Nesse caso, o leitor constrói a própria narrativa ao assumir o controle dos recortes de tempo e de agentes que deseja observar com a ferramenta Basômetro, como destaca Toledo:

É também uma revolução na narrativa jornalística. Em vez de o jornalista contar para o leitor/espectador/ouvinte o que aconteceu, o Basômetro transfere ao usuário o poder de narrar a história para si próprio. O jornalista perdeu a exclusividade de descrever o que se passou. Qualquer um pode fazer isso - sem intermediários, preferências ou preconceitos que não os seus (in: GRAY; CHAMBERS; BOUNEGRU, 2014).

Para verificar o impacto da ferramenta na cobertura do tema governismo na cobertura política do Estadão, foi realizada uma análise de conteúdo (BAUER; GASKELL, 2011), comparando as matérias que se apoiam e as que não se apoiam no Basômetro no período entre 12 de maio de 2012, quando a ferramenta foi lançada, e 1º de março de 2013. Usando-se a ferramenta de busca da versão digital do jornal impresso *O Estado de S. Paulo*, verificou-se que, de 1.247 textos publicados sobre a base aliada do governo Dilma Rousseff, 33 eram relativos ao Basômetro. A busca foi feita a partir do uso da palavra-chave “base aliada”. Posteriormente foram eliminadas as redundâncias e feita uma nova filtragem pela palavra-chave Basômetro. Daí foi realizada a análise de conteúdo das 33 reportagens que citavam o Basômetro e de outras 33 selecionadas dentre as restantes de forma aleatória.

Fazendo-se uma comparação entre reportagens que utilizaram e as que não utilizaram o Basômetro nota-se que a diferença mais óbvia é a fonte das informações. Verificou-se nas matérias não-Basômetro analisadas uma predominância das declarações dos próprios parlamentares, em on e em off, e de fontes ligadas ao Congresso e ao Executivo como fonte das análises do comportamento dos deputados e senadores e a conjuntura política. Já as feitas com auxílio da ferramenta partem das ações concretas dos parlamentares, as votações, para elaborar análises e avaliar a dinâmica de apoio ao governo federal e a conjuntura política, sem consulta prévia explícita aos políticos.

Como resultado, enquanto as matérias tradicionais levam em conta a base aliada “nominal” – formada pelo conjunto de partidos formalmente coligados ao PT – ou aquela construída a partir de declarações dos próprios parlamentares aos jornalistas e observadores, as que foram produzidas com a ferramenta apresentam a base aliada “real”, ou seja, o retrato

do apoio ao Executivo em um determinado período, baseado no número de parlamentares que efetivamente seguiu a orientação do governo federal nas votações do Congresso.

Pode-se ilustrar essa diferença com duas reportagens, uma anterior e outra posterior ao lançamento da ferramenta. A primeira, publicada em 18 de março de 2012 e intitulada “Em crise com base aliada, governo desconhece tamanho real da coalizão”⁸⁴, afirmava que a própria presidente, Dilma Rousseff, não sabia o tamanho de sua coalizão, em meio a uma crise entre o Executivo e o Legislativo. Já a segunda, publicada no dia 12 de maio de 2012 com o título “Núcleo duro’ de Dilma na Câmara encolhe em 2012, revela Basômetro”⁸⁵, mostra que os deputados que ficaram ao lado do governo federal em pelo menos 90% das votações, o que chamam de “núcleo duro”, vinha caindo desde o ano anterior e havia passado de 306 parlamentares em 2011, para apenas 118 nos primeiros meses do ano seguinte. Observa-se, portanto, que há uma precisão maior ao retratar o apoio parlamentar da presidência.

Os indícios apontam também para uma menor influência dos congressistas na definição do agendamento (McCOMBS; SHAW, 1972) quando se toma como ponto de partida as bases de dados na elaboração das pautas. As matérias feitas com o Basômetro, de maneira geral, no entanto, não parecem ter como ponto de origem as fontes, mas sim a análise de dados. Isso pode ser reforçado pelo fato de os integrantes do Estadão Dados não estarem baseados em Brasília ou participarem da cobertura regular das atividades do Congresso, mas sim, na sede do grupo em São Paulo. Ou seja, não há um relacionamento diário ou regular entre fonte, diminuindo uma possível influência – ou a suscetibilidade do repórter à influência – dos políticos. Não que seja necessária a presença física na capital federal para entrar em contato com essas fontes, mas a proximidade é um fator de influência.

Se confirmada como uma característica recorrente do JGD, essa relativa independência em relação aos políticos é salutar para a credibilidade das informações, pois, como afirma Lage (2000): “Fontes oficiais, como comprovam autores de todas as épocas, falseiam a realidade. Mentem para preservar interesses estratégicos e políticas duvidosas, para beneficiar grupos dominantes, por corporativismo, militância, em função de lutas internas pelo poder”.

Por outro lado, se for centrada apenas nos dados, reportagens não dão conta das movimentações e acordos paralelos que tenham levado a um determinado comportamento da

⁸⁴ Ver reprodução do texto em: www.estadao.com.br/noticias/impresso,em-crise-com-base-aliada-governo-desconhece-tamanho-real-da-coalizao-,850002,0.htm

⁸⁵ Ver reprodução do texto em: www.estadao.com.br/noticias/nacional,nucleo-duro-de-dilma-na-camara-encolhe-em-2012-revela-basometro,872159,0.htm

base aliada. Essa falta parece ser suprida com o apoio de correspondentes. As fontes funcionariam, portanto, como checagem e esclarecimento das informações extraídas das bases de dados, que podem revelar dinâmicas não verbalizadas pelos políticos, como uma retaliação ou distanciamento dos interesses do Executivo.

Outra característica interessante identificada na leitura dos textos é que, enquanto as reportagens tradicionais tendem a destacar em votações específicas de grande visibilidade, como a Lei Geral da Copa, o Código Florestal, as matérias produzidas com auxílio do Basômetro têm como foco principal identificar padrões de comportamento dos políticos ao longo de um determinado período. Exemplo é a reportagem intitulada “Senado supera governismo da Câmara em 2012”⁸⁶ e publicada no dia 20 de maio de 2012. A matéria não parte de votações relevantes ou declarações de políticos que evidenciassem a diferença de posicionamento entre as duas Casas, mas de uma análise estatística do conjunto de votações daquele ano.

Essa preferência vai de encontro a uma das convenções narrativas consolidadas na cobertura jornalística política ao longo do século XX identificadas e analisadas por Schudson (1982): Que uma notícia deve se concentrar em um único evento em vez de acontecimentos contínuos e repetidos, ou que, se a ação for repetida, a atenção deve se concentrar na novidade e não no padrão. Caso se confirme essa tendência, isso pode significar uma mudança significativa no enquadramento das reportagens. Uma explicação para esse enquadramento pode ser o fato de o JGD usar quase sempre a análise estatística, o que favorece a identificação de padrões, inerente a essa prática.

Por basear-se em dados estatísticos, a informação oferecida pela ferramenta é mais precisa do que a maneira tradicional de se avaliar a adesão dos parlamentares ao executivo, adotada tradicionalmente pela mídia: declarações dos próprios políticos, interlocutores e outras fontes oficiais. Não que essas declarações tenham perdido valor – continuam sendo importantes para se entender as motivações por trás do aumento ou redução do “governismo” –, mas são menos confiáveis por terem outros interesses que podem levá-las a desinformar e mentir para o jornalista (LAGE, 2000).

A partir da análise comparativa das reportagens feitas com o auxílio da ferramenta JGD Basômetro, verifica-se indícios de que o uso de bases de dados como suporte e fonte primária para a produção de reportagens pode incentivar modificações na forma narrativa, no enquadramento e no próprio agendamento das pautas relacionadas à cobertura jornalística das

⁸⁶ Ver reprodução online em: www.estadao.com.br/noticias/impreso,senado-supera-governismo-da-camara-em-2012-,875257,0.htm

votações da Câmara dos Deputados e do Senado, ao privilegiar análises estatísticas independentes de dados concretos e a identificação de padrões de comportamento ao invés de declarações de políticos e fatos novos ou inusitados. O estudo, porém, permite apenas apontar possíveis tendências surgidas com a emergência do JGD, sem pretender comprovar modificações reais na cobertura jornalística, o que exigiria um estudo mais abrangente e aprofundado.

CONCLUSÃO

A atividade jornalística passa por um momento de transição como resultado de um contexto mais abrangente de mudanças provocadas pelo surgimento das novas tecnologias de informação e comunicação, que tem fortes reflexos sociais, políticos, culturais e econômicos em todo o mundo, tornando a sociedade contemporânea mais global, horizontal e transparente (CASTELLS, 1999, 2003; JENKINS, 2008, MANOVICH, 2001). Entre os efeitos desse fenômeno está o crescimento exponencial no volume de informações que são produzidas, armazenadas e distribuídas, bem como uma transformação na maneira como são recebidas, manipuladas e compreendidas, especialmente com a popularização da Internet e das mídias digitais (GOLDHABER, 2006, 2012; SIMON, 1972; LANTHAM, 1994).

Dessa forma, as organizações jornalísticas – que por mais de 100 anos preservaram o mesmo modelo de negócio, baseado na venda de audiências para anunciantes, e por quase 50 anos mantinham seus processos produtivos praticamente inalterados, assim como sua relevância e influência política e social – precisam lidar agora com concorrência onipresente pela coleta, seleção, produção e distribuição de notícias e informações, sobre a qual detinham praticamente monopólio (ALBARRAN, 2010; DOYLE, 2013; PICARD, 2009, 2013; RICHERI, 2012). Isso se reflete na perda de audiência e de anunciantes, sua principal fonte de receita e pilar da sustentabilidade da indústria de mídia.

As estratégias adotadas por estas instituições para lidar com as mudanças não estão sendo bem-sucedidas de maneira geral, provocando uma queda da qualidade dos profissionais nas redações, das condições de trabalho e do conteúdo jornalístico em si. O foco no aumento do volume de notícias e na rapidez de sua distribuição tornaram os jornalistas e os conteúdos que produzem superficiais, padronizados e facilmente intercambiáveis, de pouco ou nenhum valor relativo para a sociedade (PICARD, 2013). Como resultado, foi iniciado um ciclo vicioso que leva à perda de mais receita, novos cortes de pessoal e menor qualidade no conteúdo.

Nesse contexto, os jornalistas começam a incorporar novas técnicas e habilidades da computação (Reportagem com Auxílio de Computador, jornalismo computacional) e das ciências sociais (Jornalismo de Precisão) para auxiliar na busca por informações relevantes e de interesse público, assim como para facilitar a sua compreensão pelas audiências. O uso de computadores e de metodologia científica já era uma realidade nas redações norte-americanas desde os anos 1960 e nas brasileiras desde os anos 1990, mas ganhou uma nova dimensão a

partir do surgimento uma plataforma global de circulação de bases de dados, ferramentas e conhecimento necessário para usá-las, que multiplicou esses três elementos em escala inédita.

O Jornalismo Guiado por Dados (JGD) emerge nesse contexto de crise financeira e estrutural das organizações jornalísticas, abaladas pelo aumento da concorrência e pela perda da hegemonia e, conseqüentemente do valor, das funções tradicionais que a sustentaram, notadamente o acesso exclusivo às informações, a seleção e hierarquização dessas informações e sua distribuição eficiente (PICARD, 2009). O aumento exponencial da produção e divulgação de bases de dados de interesse público por governos, empresas, instituições e indivíduos criou uma nova oportunidade para que o jornalismo assuma a frente no exercício dessas funções, dessa vez com o papel mais valorizado de produtor de sentido a partir desses grandes volumes de dados pouco acessíveis para as pessoas comuns.

A definição do que é JGD ainda é fluida. Träsel (2013b) é quem chega mais próximo de um conceito que abarque os objetos analisados ao longo desta pesquisa. Sugere-se aqui uma ampliação de sua definição com o objetivo de ressaltar a influência dos fatores simbólicos e subjetivos no processo produtivo jornalísticos, que tendem a ser esquecidos especialmente quando a análise envolve o uso da tecnologia: *o jornalismo guiado por dados é a aplicação da computação e dos saberes das ciências sociais na produção de sentido a partir de bases de dados com o objetivo de ampliar a função da imprensa com defensora do interesse público.*

Isso porque analisar o JGD sob o aspecto da produção de sentido abre espaço para se reconhecer um leque de fatores que influenciam a prática e, portanto, afetam o seu resultado. Conforme já foi observado, uma gama de fatores influencia a construção das reportagens desde a apuração até sua edição, quando se considera o caráter simbólico e referencial da atividade jornalística (LAGE, 2009). No curso desta pesquisa foram identificados cinco fatores que afetam especificamente a produção de sentido a partir de bases de dados, além daqueles comuns à produção jornalística tradicional. São eles: disponibilidade de dados; confiabilidade dos dados; disponibilidade de ferramentas e habilidades necessárias para coletar, analisar e visualizar dados; limitações e potencialidades das ferramentas; e capacidade de construir narrativas a partir de dados. Importante ressaltar que não se trata de uma lista exaustiva e que pode haver outros elementos que afetam o resultado final específicos do JGD.

A disponibilidade dos dados está diretamente relacionada ao conteúdo gerado pelo JGD e envolve não apenas o fato de estarem acessíveis, o que é relevante, mas de que maneira e em que formatos são apresentados e podem ser utilizados. Podem ser conseguidos com o uso de leis de acesso à informação, raspagem de dados ou vazamento, por exemplo. Os

arquivos podem estar em diversos formatos, inclusive não estruturados, ou seja, que não podem ser entendidas, processadas e relacionadas com outras bases pelo computador. Nesse caso só podem ser extraídos manualmente, o que leva tempo, ou com o desenvolvimento de algoritmos mais sofisticados. Tudo influencia na escolha dos jornalistas que irão utilizá-los.

Por exemplo, a lista de funcionários públicos e seus respectivos cargos e salários pode ser divulgada em formato PDF⁸⁷ como se fosse uma imagem, caso da prefeitura de Jaboticabal, em São Paulo⁸⁸. Esse formato impede a tabulação das palavras e números, pois o software entende os dados como uma foto. Caso não tenha conhecimento das ferramentas, o jornalista tem poucas opções. Uma delas é digitar os dados em uma tabela, o que pode tomar muito tempo ou ser inviável dependendo do tamanho. Outra é analisar os dados visualmente e procurar informações relevantes, como seria feito com um documento de papel. Essa opção limita a capacidade de relacionar os salários, verificar quais são os mais altos, relacionar com outras bases, como conselhos de estatais, por exemplo. Se o jornalista conhecer os softwares de reconhecimento óptico, no entanto, pode converter o conteúdo em dados estruturados em poucos segundos.

A confiabilidade dos dados, por sua vez, influencia diretamente no conteúdo final e é determinante para sua adoção ou não pelos profissionais. Essa capacidade de selecionar e verificar inclusive é um dos diferenciais do jornalismo. Um tema ou área podem deixar de ser abordados caso não haja dados confiáveis a seu respeito. Ou suscitar uma coleta pelos próprios jornalistas, seja individualmente ou coletivamente. De qualquer maneira, é sempre levada em consideração.

Como já foi demonstrado, a prática do JGD exige um novo conjunto de habilidades e ferramentas. Na maioria das vezes se trabalha com profissionais que não são jornalistas, como desenvolvedores e designers. A disponibilidade das ferramentas e das habilidades e bases de conhecimento necessárias para coletar, analisar e visualizar os dados, portanto, determina quais são as possibilidades de trabalho. Um jornalista com conhecimento de estatística, por exemplo, poderá chegar a informações muito mais relevantes ao analisar conjuntos de dados do que aquele que não domina essa área.

As limitações e potencialidades dessas ferramentas também influenciam a produção. É natural que o jornalista após aprender a usar um novo aplicativo, habilidade ou conhecimento,

⁸⁷ O PDF (Portable Document Format) é um formato de arquivo aberto, que comporta textos, gráficos e imagens, criado pela Adobe Systems para visualizar documentos mantendo a mesma padronização independentemente do sistema operacional e do *software* em que for usado.

⁸⁸ Ver em: <http://www.jaboticabal.sp.gov.br/2010/index.php/conteudo/visualizar/salarios-dos-servidores-publicos>

ou mesmo durante o processo de aprendizado, queira aplicar na prática. Da mesma forma, se tiver maior domínio sobre um determinado software tenderá a usá-los mais do que outros, buscando dados que se adéquem ou adequando os dados à ferramenta.

Por fim, a capacidade de construir narrativas a partir dos dados afeta a forma como serão apresentados à audiência e sua atratividade para o público. Nesse caso, envolve não apenas a formatação dos dados, mas a contextualização e humanização desses dados para que se tornem mais palatáveis e compreensíveis. Assim como um escritor pode elaborar um texto sedutor, mas baseado apenas em ficção, um programador pode criar um aplicativo extremamente sofisticado e factualmente preciso de visualização de dados, mas que não atraia ninguém.

Ao se voltar para um tema que, apesar de importante, não é facilmente compreendido e absorvido pelas pessoas não especializadas, o JGD retoma em parte o papel de mediador perdido pelo jornalismo tradicional por conta da concorrência com outras instâncias da sociedade contemporânea. Quando se direciona para a coleta, análise e distribuição de bases de dados e incorpora um novo conjunto de habilidades e ferramentas, o profissional do JGD se diferencia das práticas tradicionais e padronizadas, e conseqüentemente se torna mais valorizado. O mesmo vale, em maior escala, para as organizações jornalísticas.

O JGD contribui para a criação de valor ao renovar e recuperar a relevância das funções tradicionais do jornalismo. O acesso exclusivo à informação se dá, agora, por meio dos recursos necessários para que seja coletada, como o uso de ferramentas e técnicas de raspagem de dados, *crowdsourcing*, *data mining*, reconhecimento ótico de caracteres, mTurk e da aplicação das leis de acesso a informação. O papel de gatekeeper é substituído pelo de “pólo de dados confiáveis” em um ambiente no qual a informação existe em excesso, mas a informação confiável é escassa (LORENZ; KAYSER-BRIL; MCGHEE, 2011). Além da escassez, o valor é criado na medida em que o processo de coleta, manipulação e visualização de dados, usando habilidades e ferramentas especializadas, amplia sua utilidade para o público. O mesmo acontece na distribuição. Em vez da criação de valor ser oriunda do controle, passa a ser construída pela nova formatação de produtos e serviços de forma que tenham mais utilidade para os consumidores de informação. A possibilidade de personalização dos conteúdos jornalísticos oferecida pelas ferramentas e aplicativos de JGD confere uma nova relevância e utilidade para as audiências a que se destinam. O mesmo acontece com o uso de novas narrativas, mas atraentes e que simplificam o entendimento de temas complexos, economizando tempo e esforço.

A incorporação do JGD pelas organizações jornalísticas se apresenta, então, como uma forma de recuperar parte da relevância e do valor perdido com a ascensão dos meios digitais e da internet. Por meio do uso de novas bases de conhecimento, permite a criação de produtos e serviços diferenciados e inovadores com maior utilidade para o consumidor.

A partir da análise de casos exemplares de JGD e de pesquisas sobre o tema, levanta-se a possibilidade de a prática, quando inserida nas organizações jornalísticas, poder atuar de maneira a renovar essas funções, recuperando assim, parte da relevância perdida pelo jornalismo e contribuindo para a criação de valor. Foram identificadas seis possíveis contribuições do JGD para a criação de valor dentro das organizações jornalísticas: busca de informações exclusivas; produção de sentido e contextualização de grandes volumes de informações; personalização do conteúdo; redução do custo de investigação (PARASIE; DAGIRAL, 2013); pólo de dados confiáveis (LORENZ; KAYSER-BRIL; MCGHEE, 2011); e venda e licenciamento de aplicativos e ferramentas.

Cinco dessas contribuições foram encontradas no estudo de caso do núcleo Estádio Dados, com exceção da venda e licenciamento de aplicativos e ferramentas – esta foi verificada no estudo de caso de apoio InfoAmazônia. A busca de informações exclusivas é primordial e orienta a prática da equipe cotidianamente, sendo uma dos fatores que justifica sua existência. Sem conteúdo original, é difícil acreditar que seria mantida. Há vários exemplos de furos, como o levantamento que mostrou que os governadores empregavam 105 mil funcionários nos estados sem concurso público⁸⁹, ou que só 8% dos municípios arrecadam mais do que gastam, sustentando, portanto, todos os outros⁹⁰.

Essa exclusividade deriva majoritariamente da produção de sentido e contextualização de grandes volumes de informações. Isso porque o núcleo quase não trabalha com bases de dados exclusivas, mas produz novos sentidos a partir de dados públicos, seja colocando-os em um novo contexto, relacionando com outras bases de dados, analisando de forma mais rigorosa o que passou em branco por outros veículos e instituições, ou simplesmente perguntando aos números o que ninguém havia pensado antes. Exemplo é a quebra de crenças arraigadas no terreno fértil para especulação que é a política, como o mito que os brasileiros não se querem saber do assunto e que o interesse está concentrado nas grandes cidades. Após analisar 18 milhões de registros que os partidos entregaram em outubro ao Tribunal Superior

⁸⁹ Ver em: <<http://www.estadao.com.br/noticias/nacional,governadores-controlam-maquina-de-105-mil-cargos-sem-concurso-publico,1015021,0.htm>>

⁹⁰ Ver em: <<http://economia.estadao.com.br/noticias/economia-geral,so-8-dos-municipios-brasileiros-arrecadam-mais-do-que-gastam,174415,0.htm>>

Eleitoral (TSE), o Estadão Dados descobriu que são os pequenos municípios – aqueles com menos de 20 mil votos válidos nas eleições para prefeito – que determinam a lógica das filiações partidárias no Brasil e que o índice de filiação partidária brasileiro é maior do que todos os países europeus, com exceção da Áustria e do Chipre⁹¹.

Esse tipo de tratamento da informação, aliado à marca e a tradição do veículo, contribuem para a sua transformação em um “pólo de dados confiáveis” (LORENZ; KAYSER-BRIL; MCGHEE, 2011), ou seja, um centro difusor e certificador de informações e dados. Isso é especialmente verdade quando se trata da cobertura política, pois, com o Basômetro, o Estadão Dados atraiu um público especializado: cientistas políticos que buscam dados confiáveis para seus estudos. A resposta foi tão positiva que está sendo elaborado um livro com as análises feitas por acadêmicos com o uso da ferramenta. No próprio site do Basômetro foram publicados 10 artigos de cientistas políticos, o que demonstra engajamento deste público com a ferramenta.

A personalização do conteúdo se apresenta em grau menor do que o observado em projeto de JGD criados por equipes estrangeiras, mas está presente nas ferramentas interativas, que permitem ao usuário fazer recortes e análises de forma customizada. O uso de mapas também permite um certo grau de personalização por local, como acontece no projeto “96 vezes São Paulo”⁹², que traça o perfil dos 96 distritos de São Paulo com base nos dados do Censo 2010 do IBGE.

Outra contribuição observada é a redução do custo de investigação, conforme observado por Parasie e Dagiral (2013, p.58). Efetivamente, o uso de ferramentas gratuitas de coleta, formatação e visualização de dados reduz o investimento necessário em deslocamento e mão-de-obra, já que o tempo gasto entre conseguir os dados e divulgá-los é bem menor do que se fosse seguida a rotina tradicional. Além da automação de procedimentos antes feitos manualmente, como a coleta e a tabulação de dados, pode-se colocar nessa conta os gráficos mapas e visualizações – mesmo os mais simples era feitos normalmente por designers com *software* caros, mas agora podem ser elaborados pelos próprios jornalistas com o uso de aplicativos gratuitos ou de baixo preço. Entre os aplicativos usados estão: infogr.am, plot.ly, Datawrapper, Google Docs, Google Maps e CartoDB.

Observou-se no caso estudado que o trabalho do JGD é multidisciplinar e majoritariamente coletivo, envolvendo a aplicação de conhecimentos das áreas de computação

⁹¹ Toledo, J.R.;Bramatti, D.;Maia, L.A.;Burgarelli, R. Eleição municipal e pequenas cidades induzem filiação partidária no Brasil. *O Estado de S. Paulo*, 5 de Janeiro de 2014, p. A4.

⁹² Ver em: <http://blog.estadaodados.com/96xsp/>

e estatística, além da aplicação de lógica científica na abordagem das reportagens, com formulação e teste de hipóteses. O investimento na criação do núcleo envolveu a contratação de um jornalista e um programador e a alocação em período não integral de mais dois jornalistas e o coordenador, além das instalações físicas, que têm custo marginal para a empresa. Em retorno, além das contribuições para a criação de valor no processo produtivo como já mencionado pode-se incluir o treinamento de outros profissionais da redação e da diretoria de arte nas ferramentas utilizadas, além do prestígio de se posicionar como uma empresa inovadora, ainda que esse reconhecimento se origine menos da audiência do que no meio jornalístico.

De maneira geral, observou-se que a inserção do JGD nas redações brasileiras ainda acontece de maneira informal e experimental. O Estadão Dados, objeto do estudo de caso, é a equipe mais avançada existente hoje no país, tem reconhecimento interno e externo, mas ainda não possui os recursos e a visibilidade de outras equipes de dados em publicações estrangeiras

As redações dos grandes veículos têm profissionais treinados para lidar com dados, especialmente por conta do esforço de instituições como Abraji, Transparência Brasil e Contas Abertas, assim como designers, nas editorias de arte, e programadores, geralmente lotados nos websites, que, juntos, poderiam formar equipes de JGD, a exemplo do que foi feito com o Estadão Dados. Ao menos em termos de prestígio o retorno tem sido alto, como mostram os prêmios recebidos por adeptos do uso de bases de dados, como Gazeta do Povo, Contas Abertas, Transparência Brasil, InfoAmazônia, etc.

As organizações jornalísticas vivem de maneira geral no Brasil um ciclo vicioso de perda de receita publicitária, precarização do trabalho para contenção de despesas e queda da qualidade do conteúdo, que se tornou superficial, padronizado e declaratório. A mudança do conteúdo é resultado do processo de profissionalização do jornalismo, que ocorre desde a década de 1960-70 e que promove a commoditização do trabalho jornalístico e, conseqüentemente, dos conteúdos. Esse fenômeno se agrava com o jornalismo online, premido ainda mais pela pressão do tempo, agora com disputa de segundos entre os veículos concorrentes, que favorece ainda mais a padronização, a superficialidade e o jornalismo declaratório.

O JGD no Brasil surge em meio a um esforço de jornalistas de resgatar e valorizar o bom jornalismo, voltado para a revelação e verificação de informações com o objetivo de defender o interesse público, e quem é denominado Jornalismo Investigativo, marcando sua

oposição ao declaratório. O JGD oferece ferramentas que ampliam essa capacidade de coletar, verificar e compreender

A adoção do JGD, portanto, seria uma das formas de as organizações jornalísticas adaptarem suas estratégias à nova realidade da mídia, criando conteúdos, produtos e serviços diferenciados e atraentes para o consumidor de informações. Há, claro, outras práticas e estratégias que podem contribuir, talvez até de maneira mais incisiva, para a criação de valor nesse cenário, mas essa se apresenta como uma das opções do leque de alternativas que estão a disposição daqueles que buscam reverter o anacronismo dos modelos produtivos e de negócio tradicionais do jornalismo.

A busca pelo jornalismo de qualidade se interlaça com a necessidade de garantir a sustentabilidade da atividade. Percebe-se uma coincidência na busca por conteúdo exclusivo e de qualidade, voltado para os interesses dos cidadãos, com as estratégias de valorização econômica do produto jornalístico. O jornalismo deve voltar para suas raízes, para suas funções originais, se quiser manter sua relevância e seu valor, pois só assim se diferenciará da massa de produtores e distribuidores de dados e informação que existe hoje no mundo. O JGD surge como a prática que incorpora novos conhecimentos e novas tecnologias ao processo jornalístico para alcançar esses objetivos.

Ao mesmo tempo, pode contribuir para aumentar a relevância e o valor dos produtos jornalísticos, ao ampliar o potencial de coleta, análise, compreensão e comunicação de informações das organizações de mídia, ao criar novos produtos e serviços, ao permitir a personalização do conteúdo, e abrir espaço para o desenvolvimento de novos modelos de negócio que contribuem para a sustentabilidade das organizações jornalísticas.

A incorporação de novas tecnologias e áreas de conhecimento ao processo jornalístico podem ser grandes aliadas nesse processo se forem enxergadas não como ameaça, mas como ferramentas para fazer melhor jornalismo. O mesmo vale para a gestão das organizações e o conhecimento das forças econômicas que são determinantes para a manutenção da atividade. Ao reconhecer aquilo que realmente tem valor, as organizações e os profissionais podem centrar esforços e evitar erros, como o esvaziamento em talentos e experiência das redações.

Esse é um processo que ainda está em andamento. As tecnologias da informação e comunicação continuam evoluindo, assim como sua influência sobre a dinâmica social e econômica. Ao mesmo tempo em que os jornalistas buscavam meios para compreender a massa crescente de dados que é produzida no mundo digital, ocorria outro fenômeno ligado aos dados: o uso da inteligência artificial para o desenvolvimento de ferramentas capazes de produzir textos em linguagem natural sem a intervenção humana. Ao que tudo indica a

automação promete ser responsável pelo próximo choque tecnológico nas organizações jornalísticas, com efeitos perturbadores sobre os processos produtivos e os modelos de negócio. Projeto e experimentos começam a tomar corpo, especialmente nos Estados Unidos e na Europa com o objetivo de substituir os jornalistas por algoritmos na produção de notícias. Serão substituídos exatamente aqueles cujo trabalho é padronizado e pode ser reproduzido pelo uso de inteligência artificial.

Adotada primeiro na gestão de empresas, para auxiliar e acelerar a tomada de decisões a partir de grandes bases de dados, a tecnologia acabou chegando à mídia. A agência de notícias Thomson Reuters foi a primeira a anunciar, em 2006, que iria substituir jornalistas por algoritmos na produção de informes financeiros, mas sem alarde. O tema ganhou repercussão em 2010, quando a companhia de mídia digital Statsheet lançou uma rede de 345 websites cobrindo todos os times da liga de basquete universitária norte-americana. Posteriormente, foi a vez de a Forbes anunciar que estava trocando repórteres por algoritmos na elaboração de notícias financeiras. Os relatórios financeiros das companhias e previsões de analistas passaram a ser lidos e transformados em matérias por um software desenvolvido pela empresa Narrative Science. Já em 2009, Peter Kirwan (2009) alertava para o risco de obsolescência da atividade:

Journalists remain artisans in an era of industrialization. Inside newsrooms, the old craft methods remain dominant. Outside, across the vast expanse of the web, algorithms are automating the information industry. None of this would matter if newsrooms were hermetically sealed enclaves. But they're not. Increasingly, journalists find themselves on the business end of a high-velocity feedback loop created by digital distribution and reader interaction (KIRWAN, 2009).

O uso da inteligência artificial para automatizar o trabalho jornalístico não foi bem recebido pela categoria, que teme pelo impacto na oferta de trabalho e em questões éticas, como a responsabilidade pelo conteúdo publicado, direitos autorais e transparência nos critérios adotados na construção dos algoritmos. Realmente, sua adoção integra uma tendência de precarização dos profissionais e de redução do valor atribuído a seu trabalho:

This development can best be observed on the Internet where automated aggregation websites take over the gatekeeping function and content farms produce low-cost content to attract page views. Robot journalism fits in with this trend, producing content for the long tail, in virtually no time and with low additional costs for articles which can be produced in large quantities (VAN DALEN, 2012).

Por outro lado, pode dar mais tempo para que os jornalistas se concentrem naquilo que tem realmente valor, a criatividade, a apuração, a elaboração de reportagens de fôlego, tudo

aquilo que é mais humano (VAN DALEN, 2012). Além disso, pode expandir o alcance da cobertura jornalística para nichos hoje ignorados e que são muito pequenos para justificar o investimento da mídia, o que é de interesse público e pode atrair novas fontes de receita.

Além disso, mais do que substituir jornalistas, a inteligência artificial tem o potencial de expandir seu potencial ao ampliar sua capacidade de compreender grandes volumes de dados e torná-las mais acessíveis ao público ao transformar números em textos de forma rápida, barata e em uma escala inimaginável. Um bom exemplo é a parceria feita pela organização de mídia sem fins lucrativos ProPublica com a Narrative Science para melhorar seu projeto de JGD Opportunity Gap⁹³, que reúne e compara os dados de desempenho de 52 mil escolas dos Estados Unidos com informações socioeconômicas de cada local para verificar a disparidade de oportunidades oferecidas a cada criança. Inicialmente, os dados eram apresentados em gráficos e tabelas que exigiam certo grau de instrução para que fossem compreendidos, o que impedia o acesso à informação exatamente das pessoas mais prejudicadas pela desigualdade no ensino. Foi criado um algoritmo que traduziu os dados em textos com linguagem natural, um para cada escola, ampliando a compreensão do problema e alcançando todos por ele afetados.

Se adotada apenas com o propósito de reduzir custos e precarizar o trabalho, a automação promete criar pouco ou nenhum valor para as organizações jornalísticas. Se aplicada com o objetivo de melhorar e potencializar o trabalho, no entanto, pode se tornar uma evolução do JGD e contribuir para a sustentabilidade da atividade. Mais do que nunca é preciso compreender as forças que impactam o jornalismo na contemporaneidade, especialmente avanço das tecnologias da informação e comunicação e o novo papel das organizações de mídia na economia. Para que possam ser influenciadas e direcionadas a servir os propósitos do bom jornalismo, a defesa do interesse público, e não subordiná-lo aos interesses outros, como vem ocorrendo.

⁹³ Ver em: <<http://projects.propublica.org/schools/>>

REFERÊNCIAS

ALBARRAN, Alan B. *The Media Economy*. New York: Routledge, 2010.

_____. *The SAGE Handbook of Media Studies*. Londres: SAGE, 2004. p. 291-307. Disponível em: <<http://www.sagepub.com/mcquail6/PDF/Chapter%2014%20-%20The%20SAGE%20Handbook%20of%20Media%20Studies.pdf>>. Acesso em: 17 de Novembro de 2013.

ALVES-MAZZOTTI, A. J. Usos e abusos dos estudos de caso. *Cadernos de Pesquisa*, v. 36, n. 129, p. 637-651, set./dez. 2006.

ARRUDA, Angela. Teoria das representações sociais e teorias de gênero. *Cadernos de Pesquisa*, n. 117, p. 127-147, nov. 2002. Disponível em: <<http://www.scielo.br/pdf/cp/n117/15555.pdf>>. Acesso em julho de 2011.

BARBOSA, S.; TORRES, V. Extensões do paradigma JDBD no jornalismo contemporâneo: modos de narrar, formatos e visualização para conteúdos. In: ENCONTRO DA COMPÓS, 11., 2012. Juiz de Fora. *Anais...* Juiz de Fora, 2012.

_____. *Jornalismo Digital em Base de Dados (JDBD) - Um paradigma para produtos jornalísticos digitais dinâmicos*. 2007. (Tese de Doutorado) – PósCom/UFBA, Salvador, 2007. Disponível em: <http://www.facom.ufba.br/jol/pdf/tese_suzana_barbosa.pdf>.

BAUER, Martin W.; GASKELL, George (orgs.). Trad: Pedrinho A. Guareschi. *Pesquisa qualitativa com texto, imagem e som: um manual prático*. Petrópolis: Vozes, 2011.

BEAUVAIS, Zach. *Journalism Needs Data in the 21st Century*. Readwrite, 4 de Agosto de 2009. Disponível em: <http://readwrite.com/2009/08/04/journalism_needs_data_in_21st_century#awesm=~oxxpEJevRX6s6F>. Acesso em: 12 de Janeiro de 2014.

BORELLI, Viviane. *Jornalismo como atividade produtora de sentidos*, 2005. In: Biblioteca Online de Ciências da Comunicação. Disponível em: <<http://www.bocc.ubi.pt/pag/borelli-viviane-jornalismo-atividade-sentidos.pdf>>. Acesso em: 28 de Janeiro de 2014.

BRADSHAW, P. *Gathering data: a flowchart for data journalists*. Online Journalism Blog. Publicado em 6 de Setembro de 2011a. Disponível em: <<http://onlinejournalismblog.com/2011/09/06/gathering-data-a-flow-chart-for-data-journalists>>. Acesso em: 13 de Abril de 2013.

The inverted pyramid of data journalism. Online Journalism Blog. Publicado em 7 de Julho de 2011b. Disponível em: <<http://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>>. Acesso em: 13/ de Abril de 2013.

_____. *The US election was a wake up call for data illiterate journalists*. Online Journalism Blog, 2012. Disponível em: <<http://onlinejournalismblog.com/2012/11/07/the-us-election-was-a-wake-up-call-for-data-illiterate-journalists/>>.

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

DAVENPORT, Thomas H; BECK, John C. *A economia da atenção*. Tradução de Alfonso Celso da Cunha Serra. Rio de Janeiro: Campus, 2001.

CARNEIRO, Jorge Manoel Teixeira. CAVALCANTI, Maria Alice Ferreira Deschamps; SILVA, Jorge Ferreira da. Porter revisitado: análise crítica da tipologia estratégica do mestre. *Rev. adm. contemp.* Curitiba, v. 1, n.3, Sept./Dec. 1997. Disponível em: <<http://dx.doi.org/10.1590/S1415-65551997000300002>>

CASTELLS, Manuel. *A sociedade em rede: a era da informação: economia, sociedade e cultura*. São Paulo: Paz e Terra, 1999.

_____. *A Galáxia da Internet: reflexões sobre a internet, os negócios e a sociedade*. Trad.: Maria Luiza Borges. Revisão: Paulo Vaz. Rio de Janeiro: Jorge Zahar Ed., 2003.

CODREA-RADO, A. The politics of data. *Columbia Journalism Review*, 10 de Outubro de 2012. Disponível em: <http://www.cjr.org/data_points/the_politics_of_data.php>. Acesso em: 8 de Janeiro de 2014.

COHEN, S.; HAMILTON, J. T.; TURNER, F. Computational journalism. *Communications of the Association for Computing Machinery*, v. 54, n. 10, p. 66, 1 de Outubro de 2011.

COHEN, Sarah; LI, Chengkai; YANG, Jun; YU, Cong. *Computational Journalism: A Call to Arms to Database Researchers 5th Biennial Conference on Innovative Data Systems Research (CIDR '11)*. 9-11 de Janeiro de 2011, Asilomar, California, Estados Unidos. Disponível em: <http://www.cidrdb.org/cidr2011/Papers/CIDR11_Paper17.pdf>. Acesso em 19 de Fevereiro de 2014.

COX, Meli. *The development of computer-assisted reporting*. Newspaper Division, Association for Education in Journalism and Mass Communication, Southeast Colloquium, 17 de Março de 2000, University of North Carolina, Chapel Hill. Disponível em: <<http://com.miami.edu/car/cox00.htm>>. Acesso em: 2 de dezembro de 2013

DIAKOPOULOS, N. *Cultivating the Landscape of Innovation in Computational Journalism*. CUNY Graduate School of Journalism Tow-Knight Center for Entrepreneurial Journalism, 2012. Disponível em: <<http://pt.scribd.com/doc/87863948/Cultivating-the-Landscape-of-Innovation-in-Computational-Journalism>>. Acesso em 15 de Fevereiro de 2014

DOYLE, G. *Understanding Media Economics*. [E-book]. Londres: SAGE (Kindle edition), 2013.

DUARTE, Jorge; BARROS, Antonio (org.). *Métodos e Técnicas de Pesquisa em Comunicação*. 2. ed. São Paulo: Atlas, 2012.

EISENHARDT, K. M. Building Theories From Case Study Research. *The Academy of Management Review*, v. 14, n. 4, Outubro de 1989, pp. 532-550

FAUSTINO, Paulo. Gestão, Tendências e Estratégias no Mercado dos Media em Portugal. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 33., 2010. Caxias do Sul. *Anais...* Caxias do Sul, 2010. Disponível em: <<http://www.intercom.org.br/papers/nacionais/2010/resumos/R5-3088-1.pdf>>.

FLEW, T.; SPURGEON, C.; DANIEL, A. *The promise of computational journalism. Media, Democracy and Change: Refereed Proceedings of the Australian and New Zealand Communications Association Annual Conference*. Canberra: 2011. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/17512786.2011.616655>>.

FRANCO, Maria Laura P. B. Representações sociais, ideologia e desenvolvimento da consciência. *Cadernos de Pesquisa*, n. 121, p. 169-186, jan./abr. 2004. Disponível em: <<http://www.scielo.br/pdf/cp/v34n121/a08n121.pdf>>. Acesso em: 20 de julho de 2011.

GOLDHABER, Michael H. The Attention Economy and the Net. *First Monday*, Chicago, v.2 n.4, Abril 1997. Disponível em: <<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/issue/view/79>>. Acesso em: 10 de julho de 2012.

_____. The value of openness in an attention economy. *First Monday*, Chicago, v.11 n.6, Junho 2006. Disponível em: <<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/1334/1254>>. Acesso 10 de julho de 2012.

GORDON, Rich. *Want to build a data journalism team? You'll need these three people*. Knight lab Northwestern University blog, 28 de Junho de 2013. Disponível em: <<http://knightlab.northwestern.edu/2013/06/28/want-to-build-a-data-journalism-team-youll-need-these-three-people/>>. Acesso em 6 de Janeiro de 2014.

GRAY, J.; CHAMBERS, L.; BOUNEGRU, L. *The Data Journalism Handbook: How Journalists Can Use Data to Improve the News* [E-book]. Londres: O'Reilly Media, 2012.

_____. *Manual de Jornalismo de Dados: como os jornalistas podem usar dados para melhorar suas reportagens* [E-book]. Londres: O'Reilly Media, 2014. Disponível em: <<http://datajournalismhandbook.org/pt/>> Acesso em: 03 de Fevereiro de 2014.

GRAY, J. *Open Data and Data-driven journalism*. Data-driven journalist roundtable, Amsterdam, European Journalism Centre (EJC), 2010.

_____. *The future of data journalism*. Publicado em datajournalism.net em 28 de julho de 2012a. Disponível em: <http://datadrivenjournalism.net/news_and_analysis/The_future_of_data_journalism>. Acesso em 27 de dezembro de 2013

_____. *What data can and cannot do*. Publicado em Guardian Datablog em 28 de Maio de 2012b. Disponível em: <<http://www.theguardian.com/news/datablog/2012/may/31/data-journalism-focused-critical>>.

GUARESCHI, Pedrinho; JOVCHELOVITCH, Sandra. *Textos em representações sociais*. Petrópolis: Vozes, 1995.

HAAMER, Kris. *What is Media Economics?* Krishhaamer.co, 2012. Disponível em: <<http://krishaamer.com/2012/01/21/media-economics>>. Acesso em: 10 de dezembro de 2013

HARPER, Christopher. *Journalism in a Digital Age*. MIT Communications Forum, 1998.

HILBERT, Martin e LÓPEZ, Priscila. The world's technological capacity to store, communicate, and compute information. *Science*; v. 332, n. 6025, p. 60-5, Apr 1- Feb 10, 2011.

HOHLFELDT, Antonio; MARTINO, Luiz C.; FRANÇA, Vera Veiga (org.). *Teorias da Comunicação - conceitos, escolas e tendências*. Petrópolis: Vozes, 2001.

HOLOVATY, A. *A fundamental way newspaper sites need to change*. Adrian Holovaty, 6 set. 2006. Disponível em: <http://www.holovaty.com/writing/fundamental-change>.

HUME, Ellen. *Resource Journalism: a New Model for News*. MIT Communications Forum, 1998.

HUNTER, Mark Lee; WASSENHOVE, Luk N. Van. Disruptive News Technologies: Stakeholder Media and the Future of Watchdog Journalism Business Models. *INSEAD Working Paper Series*, Março de 2010. Disponível em: <http://markleehunter.free.fr/documents/IJ_business.pdf>

JENKINS, Henry. *Cultura da convergência*. Tradução Susana Alexandria. São Paulo: Aleph, 2008.

JENKINS, H. FORD, S. And GREEN, J. *Spreadable Media – Creating value and meaning in a networked culture*. New York: New York University Press, 2013.

KIRWAN, Peter. *The rise of machine-written journalism*. Wired Magazine, December 2009. <<http://www.wired.co.uk/news/archive/2009-12/16/the-rise-of-machine-written-journalism>>

KOVACH, Bill; ROSENSTIEL, Tom. *The Elements of Journalism* [E-book]. Nova York; Crown Publishers, 2001.

LAGE, N. Relacionamento do repórter com as fontes: procedimentos e teoria. In: COMPÓS, 9., 1994, Porto alegre. *Anais...* Porto Alegre/RS, 2000. Disponível em: <http://www.compos.org.br/data/biblioteca_1432.pdf>

LANHAM, Richard A. The Economics of Attention. *Proceedings of 124th Annual Meeting*, Association of Research Librarians, Austin, Texas, 1994. Disponível em: <<http://www.arl.org/resources/pubs/mmproceedings/124mmlanham>>.

LIMA JUNIOR, W. T. Jornalismo computacional em função da “Era do Big Data”. *Revista Líbero*, v. 14, n. 28, p. 45-52, 2011.

LINS, Bernardo E. O estudo econômico da mídia: origens e tendências recentes. In: SIMPÓSIO REGIONAL DE PESQUISA EM COMUNICAÇÃO, 7., 2005, Goiânia. *Anais...* Goiânia: UFG, 2005.

LOPES, Maria Immacolata Vassallo de. Pesquisa de Comunicação. *Revista Brasileira de Ciência da Comunicação*, n.1, p. 12-40, 2004.

LORENZ, Mirko. *Training data driven journalism: Mind the gaps*. Data Driven Journalism. Publicado em 21 de Setembro de 2011. Disponível em: <http://datadrivenjournalism.net/news_and_analysis/training_data_driven_journalism_mind_the_gaps>

_____. *Data-driven journalism: Status and Outlook*. Data-driven journalism roundtable. Amsterdam, European Journalism Centre (EJC), 2010.

LORENZ, Mirko; KAYSER-BRIL, Nicolas; MCGHEE, Geoff. *News organizations must become hubs of trusted data in a market seeking (and valuing) trust*. Nieman Journalism Lab. Publicado em 1º de Março de 2011. Disponível em: <<http://www.niemanlab.org/2011/03/voices-news-organizations-must-become-hubs-of-trusted-data-in-an-market-seeking-and-valuing-trust/>>

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

MACHADO, Elias; PALACIOS, Marcos. Um modelo híbrido de pesquisa: a metodologia aplicada pelo GJOL. In: LAGO, Claudia; BENETTI, Marcia (Org.). *Metodologia de pesquisa em jornalismo*. São Paulo: Paulus, 2006. v. 1

MAGRETTA, J. *What Management Is: How It Works and Why It's Everyone's Business* [E-book]. London: Profile books, 2013.

MANOVICH, Lev. *The Language of New Media*. Boston: MIT Press, 2001. Disponível em: <<http://www.manovich.net/LNM/Manovich.pdf>>

_____. *From reading to pattern recognition*. Manovich.net, 2011. Disponível em: <http://manovich.net/DOCS/reading_patterns.2011.pdf>

MCADAMS, Mindy. *Don't just teach skills, train young journalists to be lifelong learners*. Nieman Journalism Lab. Publicado em 17 de Setembro de 2012. Disponível em: <<http://www.niemanlab.org/2012/09/mindy-mcadams-dont-just-teach-skills-train-young-journalists-to-be-lifelong-learners/>>

MEYER, P. *Precision Journalism: a reporter's introduction to social science methods*. Bloomington: Indiana University Press, 1973.

_____. *The New Precision Journalism* [versão online]. Indiana University Press, 1991. Disponível em: <<http://www.unc.edu/~pmeyer/book>>. Acesso em: 7 de Setembro de 2013

_____. The Future of CAR: Declare Victory and Get Out! In: _____. *When nerds and words collide: reflections on the development of Computer Assisted Reporting*. São Petersburgo: The Poynter Institute for Media Studies, 1999. Disponível em: <www.unc.edu/~pmeyer/carfuture.doc>. Acesso em 27 de Dezembro de 2013.

MCCOMBS, M. E.; SHAW, D. L. The Agenda-Setting Function of Mass Media. *The Public Opinion Quarterly*, v. 36, n. 2, 1972, p. 176-187.

MCLUHAN, Marshall. *Os meios de comunicação como extensões do homem* (Understanding media). São Paulo: Cultrix, 2012.

MICHENER, Robert Gregory. *The Surrender of Secrecy: Explaining the Emergence of Strong Access to Information Laws in Latin America*. University of Texas at Austin, abril 2010. Dissertação de Mestrado. Disponível em: <<http://gregmichener.com/Dissertation.html>>

MIELNICZUK, L. *Jornalismo na Web: uma contribuição para o estudo do formato da notícia na escrita hipertextual*. Tese (Doutorado) – FACOM/UFBA, Salvador, 2003.

MILLER, Claire. C.. Abril, 2013. *Data Science: The Numbers of Our Lives*. New York Times. Disponível em: <http://www.nytimes.com/2013/04/14/education/edlife/universities-offer-courses-in-a-hot-new-field-data-science.html?_r=0>

MOREIRA, Sonia Virgínia. *International Media Concentration Research Project, Brazil chapter*. Université Paris-Dauphine, Janeiro, 21-22, 2011. Disponível em: <http://www4.gsb.columbia.edu/filemgr?file_id=738017>.

_____. Brazil: Media and Telecommunication Conglomerates. In: NOAM, E (org.). *Media Concentration and Ownership Around the World: Who Owns the World's Media?*. New York: Oxford University Press, 2014 (no prelo).

NIELSEN, Rasmus K. (2012). *Ten Years that Shook the Media World: Big Questions and Big Trends in International Media Developments*. Oxford: Reuters Institute for The Study of Journalism.

PARASIE, Sylvain; e DAGIRAL. Éric. Des journalistes enfin libérés de leurs sources? Promesse et réalité du “journalisme de données”. *Sur le journalisme, About journalism, Sobre jornalismo [En ligne]*, v. 2, n. 1, 2013. Disponível em: <<http://surlejournisme.com/rev/index.php/slj/article/view/70>> Acesso em: 5 de Janeiro de 2014.

PARKIN, M.; POWELL, M.; MATTHEWS, K. Economics, 2008. In: DOYLE, G. *Understanding Media Economics*. [E-book]. Londres: SAGE (Kindle edition), 2013.

PEARSON, K. *How Computational Thinking is Changing Journalism & What's Next*. Disponível em: <<http://www.poynter.org/how-tos/digital-strategies/e-media-tidbits/95941/how-computational-thinking-is-changing-journalism-whats-next>>

PICARD, Robert. *Criação de Valor e o futuro das Organizações Jornalísticas*. Porto: Media XXI, 2013.

_____. *Why Journalists Deserve Low Pay*. Presentation to the Reuters Institute for the Study of Journalism, RISJ Seminar Series, University of Oxford, 6 de Maio, 2009

PORTER, M. E. *Competitive advantage: creating and sustaining competitive performance*. New York: Free Press, 1985.

_____. The competitive advantage of nations. *Harvard Business Review*, p.73-93, Mar./Apr. 1990.

_____. What is strategy? *Harvard Business Review*, p.61-78, Nov./Dec. 1996.

POZZEBON, M.; FREITAS, H. M. R. Pela aplicabilidade com um maior rigor científico dos estudos de casos em sistemas de informação. *Revista de Administração contemporânea*, v. 2, n. 2, maio/ ago, p. 143-170. 1998.

REID, Alastair. *Are journalist being “intellectually outgunned”?*. Disponível em: <<http://www.journalism.co.uk/news/are-journalists-being-intellectually-outgunned-when-it-comes-to-data/s2/a555684/>>. Acessado em: 7 de Fevereiro de 2014.

RICHERI, G. *Economia dei media*. [E-book]. Roma-Bari: Editori Laterza, 2012.

ROGERS, S. *Anyone can do it. Data journalism is the new punk*. Datablog, 2012. Disponível em: <<http://www.theguardian.com/news/datablog/2012/may/24/data-journalism-punk>>. Acesso em: 20 de julho de 2013.

_____. *Facts are Sacred: The power of data*. [E-Book]. Londres: Guardian Books (Kindle Edition), 2011.

_____. *Wikileaks' Afghanistan war logs: how our datajournalism operation worked*. The Guardian Datablog, Julho 2010. Disponível em: <<http://www.theguardian.com/news/datablog/2010/jul/27/wikileaks-afghanistan-data-datajournalism>>. Acesso em: 12 de maio de 2013.

ROWLEY, J. Using Case Studies in Research. *Management Research News*, v. 25, n. 1, p. 16-27, 2002.

RUBIM, Antonio Albino Canellas; COLLING, Leandro. Cobertura jornalística e eleições presidenciais de 2006 no Brasil. *Política & Sociedade*, n. 10, p. 173-193, Abril de 2007.

SCHUDSON, M. The Politics of Narrative Form: The Emergence of News Conventions in Print and Television. *Daedalus*, v. 111, n. 4, p. 97-112, 1982. Disponível em: <<http://www.jstor.org/stable/20024819>>. Acesso em 15 de novembro de 2012

SILVA, Daniela B. *Transparência na esfera pública interconectada*. São Paulo: Faculdade Cásper Líbero, 2011. Dissertação de Mestrado.

SILVER, Nate. *O Sinal e o ruído: Por que tantas previsões falham e outras não*. Rio de Janeiro: Intrínseca, 2103.

SIMON, Herbert A. *Designing Organizations for an Information-Rich World.*, 1972 Disponível em: <<http://diva.library.cmu.edu/webapp/simon/item.jsp?q=/box00065/fld05029/bdl0001/doc0001/>>. Acesso em 23 de julho de 2012

SONNAC ; GABSZEWICZ. *L'industrie des médias à l'ère numérique*. Paris: La Découverte, 2013.

STAKE, R. E. The Case study method in social inquiry. *Educational Researcher*, v.7, n.2, p.5-8, 1978

_____. *The Art of Case Study Research*. London: Sage Publications, 1995.

STEINER, C. *Automate This: how algorithms came to rule the world*. London: Portfolio / Penguin, 2012.

ANEXO A - Entrevista José Roberto De Toledo

Quando começou a usar bases de dados para produzir reportagens?

Por volta de 1993/4, destrinchando tabelas das pesquisas do Datafolha, e a partir de 1994/5, compilando dados estatísticos sobre o Censo de 1991 que, pela primeira vez, o IBGE começava a disponibilizar online. Me lembro de um caderno especial sobre a cidade de São Paulo, em que criamos um indicador sintético de qualidade de vida para os 96 distritos paulistanos em parceria com a PUC-SP). Nesse época eu ainda era editor do Painel, da Folha, e fazia isso meio que no tempo vago. A partir de 1996, pedi para deixar o Painel e me dedicar essencialmente ao jornalismo com auxílio do computador e a mastigar números.

O que diferencia o jornalismo de dados do RAC?

A meu ver são duas coisas: quantidade e método. O jornalismo de dados é uma adaptação do RAC aos petabytes de dados que o mundo produz diariamente. Não cabem mais numa planilha de Excel, travam programas de banco de dados mais simples como Access. É preciso recorrer a programadores/desenvolvedores, estatísticos, designers. É, por isso, um trabalho de equipe - acho.

Quais os benefícios do Jornalismo de Dados em relação a outras formas e métodos de produção jornalística?

Não parte-se de uma premissa, ou uma ideia pré-concebida. Parte-se dos dados brutos e, depois de muita depuração e consultas, chega-se a uma ou mais histórias que eles têm para contar. Isso não exime o jornalismo de dados de ter viés. Todos temos, mais de um. Mas permite um ângulo mais científico, no qual as conclusões podem ser demonstradas e reproduzidas - como na ciência. Também permite criar ferramentas online ou apps em que o usuário vai ter acesso a uma interface amigável para fazer suas próprias perguntas à base de dados, extrair suas próprias respostas e contar a história para si, sem intermediação do jornalista.

Como avalia a inserção desse método nas organizações jornalísticas brasileiras?

Muito incipiente, mas com grande interesse.

Quais as principais barreiras para que se popularize nos jornais brasileiros da mesma forma como ocorre em outros países, como Estados Unidos e Inglaterra?

Recursos, treinamento, prioridade pelas chefias.

O que o motivou a criar o Estadão Dados?

Eu participei de um seminário da Universidade do Texas, em Austin, a convite do Centro Knight para Jornalismo nas Américas, de um seminário internacional que reuniu os 60 melhores jornalistas de dados do mundo (exclua-me dessa conta). Depois visitei o NY Times em vi como funcionava o time de dados deles. Na volta ao Brasil, propus ao Estadão criar algo semelhante e tanto o Gandour quanto a Claudia toparam.

Quais argumentos usou para convencer o Grupo Estado a investir no projeto? A diretoria do jornal aceitou rapidamente ou criou barreiras?

A aceitação foi rápida, imagino, porque eu já fazia algo nessa linha em parceria com repórteres do jornal de modo informal. Creio que eles já estavam convencidos de que o investimento valia a pena. E não houve gastos adicionais, apenas remanejamento de vagas e recursos - priorização.

Qual foi o investimento no projeto?

Deslocou-se uma vaga de repórter e um infografista que também programava para montar o núcleo comigo e o Daniel Bramatti, que continua sendo repórter de política.

Como é a dinâmica de produção do Estadão Dados? Há uma reunião de pauta diária ou semanal? Todos participam?

É muito informal. As reuniões existem, mas não há um calendário fixo. Procuramos selecionar as melhores ideias e nos dedicar a elas, enquanto atendemos demandas mais quentes e prosaicas da redação.

Como é feito o desenvolvimento? O repórter apura as informações e passa para o programador? Todo o material é editado ou é publicado diretamente?

A tendência é, cada vez mais, todos fazerem quase tudo. Os repórteres Rodrigo Burgarelli e Lucas Maia aprendem a programar. O Bramatti se tornou um especialista em GIS e mapas. O Diego, o programador, também ajuda com o conteúdo. É um samba do programador doido.

Qual a origem das pautas? Os repórteres escolhem determinadas bases de dados e buscam correlações a partir da análise ou partem de uma apuração para buscar dados que permitam testar a informação apurada?

Não há regra. Pode ser uma base de dados nova a que temos acesso, pode ser uma regressão estatística que deu certo e descobriu uma novidade, pode ser uma demanda da redação. Varia muito.

Como é o relacionamento da editoria com o restante da redação? Há interação, treinamento, produção de conteúdo em conjunto com repórteres de outras editorias?

Sim, cada vez mais. Damos treinamento formal para os focos da redação a cada semestre. E ensinamos quem tem interesse alguns truques e técnicas, quando nos pedem. O blog do Estadão Dados já virou muitas pautas em diversas editorias. E seu modelo foi copiado por pelo menos um blog do próprio jornal, com sucesso. Essa é a ideia: evangelizar e conquistar corações e mentes para o jornalismo de dados.

O Basômetro foi o primeiro e mais complexo aplicativo desenvolvido pelo Estadão Dados. Quanto tempo levou para ser feito? Há a expectativa de produzir outros aplicativos como esse?

Demorou cerca de 4 meses. E ainda toma tempo de vez em quando para ser atualizado. Isso é bacana, porque o projeto se mantém vivo e útil, mas é também um problema, porque cada nova ferramenta que colocamos no ar é uma espécie de cria, que precisa ser alimentada. Sim, temos projetos.

Por que vale a pena investir em jornalismo de dados? O que o jornalismo de dados tem que outros métodos e formatos jornalísticos não oferecem?

É uma necessidade. Se o jornalismo quiser compreender o mundo de dados que são gerados diariamente ele precisa se adaptar a isso. Precisa emprestar as ferramentas necessárias para dar conta do trabalho. O jornalismo tradicional, de bloco e gravador, continua e continuará sendo necessário, mas ele não dá conta sozinho de atender essa nova demanda.

Os projetos do Estadão Dados geraram audiência significativa?

Sim, mas não estamos atrás disso. A audiência vem quando o tema é de interesse público e é apresentado de forma a captar o interesse do público. Nem sempre é fácil juntas as duas coisas. Mas a prioridade é o interesse público e a experimentação de novas formas e técnicas.

Quais benefícios identificou até agora para os leitores, a empresa e os jornalistas envolvidos no projeto?

Todos aprendemos muito - e os leitores/usuários tiveram acesso a informações que não teriam de outra forma.

ANEXO B - Entrevista Daniel Bramatti

Gostaria de saber como foi o processo de criação do Estadão Dados e qual a sua participação.

Vou fazer um pequeno histórico. O Estadão Dados foi criado pelo meu colega José Roberto de Toledo. Ele é um pioneiro na Reportagem com Auxílio de Computador. Há alguns anos ele foi convidado para um evento nos EUA, realizado pelo Knight Center, que reuniu todos os cabeças do jornalismo de dados, os pioneiros ali, naquele momento. E para ele ficou muito claro que o trabalhar com números, ferramentas, computador estava passando por uma revolução muito grande pelo volume de dados que a internet oferta, pelas ferramentas que se multiplicam e pela necessidade de que, ou o jornalista dá um pulo do gato de também entrar no mundo da programação ou de agregar programadores às redações. Ele voltou para o Brasil com essa ideia, trouxe para o Estadão, e ela foi bem acolhida pela cúpula. Eu já tinha feito algumas coisas com ele, já tinha trabalhado como subeditor durante um tempo, voltei para a reportagem, estava em um momento em que não tinha muitas fontes e tinha decidido trabalhar com números, um momento que acabou catapultando essa parceria. Eu trabalho na editoria de Política e também no Estadão Dados.

Conte um pouco do seu histórico até chegar ao Estadão Dados.

Eu trabalhei na Folha durante 12 anos na sucursal de Brasília, cobri congresso, fiz algumas coisas de economia e tenho essa facilidade de fazer contas, tenho interesse por isso. Tenho algumas matérias antes do Estadão Dados que eu já fazia no espírito do Estadão Dados. Por exemplo, todo mundo fala que tem que fazer aliança com o PMDB porque o PMDB tem maior capilaridade nos municípios. Então vamos fazer a conta para ver se nas cidades em que o PMDB governa o candidato que a presidente do partido apoia tem significativamente mais votos do que as cidades que o PMDB não governa. Não, não se encontra nenhuma evidência de que o palanque do PMDB valha alguma coisa. É quase como um método científico: você levanta uma hipótese e aí vai e tenta comprovar ou derrubar. Eu sempre tive interesse em fazer jornalismo dessa maneira e eu já conhecia o Toledo, já tínhamos trabalhado juntos e ele me chamou e a gente começou a trabalhar na primeira ferramenta, que é a mais importante até hoje, que é o Basômetro, não sei se você conhece...

Conheço. É a que eu acho mais interessante também.

Realmente, a gente começou muito bem e depois foi caindo...(risos). A gente tinha duas pessoas de programação e mais uma que trabalhava parcialmente, mas depois essas pessoas saíram e a gente teve dificuldade de recompor a equipe. Hoje temos um programador que tem muitas tarefas e fica sobrecarregado e, enfim, a gente não consegue se dedicar tanto a uma única ferramenta. O desenvolvimento de uma ferramenta dessas demora meses, então tem muita coisa que a gente tem planos mas ainda não conseguiu recursos para fazer...

Aproveitando que você falou da equipe, quantos eram no início e como é a formação hoje?

Nesse embrião, éramos eu, o Toledo, quando o Basômetro já existia foi contratada mais uma jornalista para trabalhar na equipe do Estadão Dados, foi a primeira contratação específica para o Estadão Dados, a Amanda Rossi. E os que trabalhavam com programação, um veio da editoria de arte do Estadão e os outros dois trabalhavam no portal do Estadão. Eles não

foram diretamente para o Estadão Dados, mas deram o suporte para a gente criar o Basômetro. Eles contribuíram muito, inclusive com a interface, é um desenho que ele foi adaptado de um projeto prévio, que já existia e o Toledo teve essa ideia de usar o apoio ao governo como balisa, porque no Brasil você não tem uma divisão direita/esquerda, como em outros países, uma coisa tão claramente segmentada, ou oposição/situação, o que dá a movimentação do congresso é se o cara vota com o governo ou não mesmo participando de um partido de oposição, um governista no DEM, por exemplo, e isso fica muito claro somente com esse tipo de visualização. Depois a equipe sofreu algumas mudanças. Essa jornalista que tinha sido contratada saiu para se dedicar a um livro. Chamamos outro repórter que tem muita facilidade para mexer com números, tem alguma noção de programação. Ele trabalhava na editoria de Metrópole, então agora a gente está com uma ponta na Política, que eu faço, e com Metrópole. Temos mais uma aquisição também, que é o Lucas, que tinha trabalhado no Estadão há alguns anos, depois ele foi para os Estados Unidos, fez um curso na Universidade de Chicago voltado para estatística e agora está usando esse conhecimento todo...essa ciência que está nos ajudando muito a perceber alguns fenômenos, é muito legal. Então hoje nós temos o Toledo, o Lucas, o Burgarelli, o Diego, que é o nosso programador, e eu. Mas a ideia é que o Estadão Dados seja um núcleo no qual as pessoas transitem. Estou dando um curso para o pessoal da editoria de arte, um curso de mapas online, então a ideia é que seja um lugar de difusão de conhecimento, não uma panelinha que vai fazer as matérias com dados, mas o legal é agregar essa cultura de trabalhar com dados à redação inteira.

Esse modelo foi inspirado em algum exemplo internacional?

Quando o Toledo voltou, ele já trouxe sabendo de algumas experiências internacionais, mas a gente não sabe exatamente como funciona o do Guardian, a gente sabe que a área do New York Times tem uma estrutura que nós nunca vamos ter, tem quase 40 pessoas. A gente sabe que tem que trabalhar com pouca gente e que pelo fato de trabalhar com uma equipe pequena tenta compensar com essa disponibilidade para levar o jornalismo de dados para a redação inteira, não queremos ter o monopólio dessa atividade. O Toledo e o Lucas acabaram de dar um curso para os focas. Eu acho que no momento as experiências internacionais não nos valeriam muito porque a gente se vê numa situação peculiar, tem muitos programadores, não dá para se valer muito. A gente está tentando encontrar um modelo possível, com poucos recursos humanos, mas com sorte, com apoio da cúpula do jornal. No momento em que o jornal não contrata para nenhuma editoria, a gente conseguiu contratar o Lucas.

Isso é algo que gostaria de saber um pouco mais. Pelo visto, a receptividade da direção foi positiva, mas como foi o processo de convencimento, foi fácil convencer do valor dessa prática para a empresa?

Foi muito bem acolhido desde o começo. Contribuiu para isso o fato de o Basômetro ter tido uma estreia muito boa, rendeu muitas matérias legais, teve impacto nas redes sociais, na academia os cientistas políticos já usaram o Basômetro, vai ser lançado um livro com 14 artigos de cientistas políticos com base no Basômetro, então a gente tem conseguido gerar material exclusivo para o jornal, que muitas vezes vai para a primeira página. Eu diria que tem um espaço consolidado e tem uma valorização interna muito grande. E acho que externa também, tem encontrado boas referências de colegas em outros jornais. Estão acompanhando o que a gente faz e gostando.

Quais as diferenças principais entre o jornalismo de dados em relação às formas tradicionais de jornalismo na sua opinião?

Eu acho que – no meu caso, por exemplo, eu não sou programador, mas tenho aprendido uma ferramenta nova por semana, praticamente – surgem muitas coisas legais a todo momento, em termos de raspagem de dados, visualização de dados, construção de mapas. No fazer uma matéria você acaba aprendendo a mexer numa ferramenta nova que é valiosíssima e, outra coisa, as pauta já pululam na sua cabeça de acordo com a ferramenta que você está mexendo. Coincidiu também com esse avanço tecnológico uma predisposição dos governos de abrir suas bases, o impulso da lei de acesso à informação, então temos muitos dados brutos que precisam ser lapidados. É um trabalho muito difícil, mas quando a gente consegue saem coisas muito legais. A gente fez isso com bases de convênios, por exemplo, que tinham 500 mil, 600 mil contratos para serem analisados e com a ferramenta X, Y você conseguia fazer uma leitura do que aquilo significava em termos macro. Para mim é isso, a diferença básica em relação à reportagem tradicional, em vez de você ficar naquela de contar com as pessoas, pegar o telefone e tal, é você conseguir extrair informação de números que são brutos e que não revelam nada se você não souber por onde abordá-los.

Você acredita que foi uma coincidência o fato de haver maior liberação de informações ao mesmo tempo em que se desenvolvia essa nova prática ou uma coisa levou a outra, pois foi necessário desenvolver uma prática para processar esses dados?

Acho que não é uma coincidência no sentido de que foi um acaso. São dois movimentos que passaram a andar paralelos, não sei até que ponto tem intenção nisso, mas que um depende do outro. Não adiante ter muitas ferramentas se não tem dados para trabalhar. E não adianta ter um monte de dados se você não tem como trabalhar com eles. Então eu digo coincidência no sentido de que os dois movimentos aconteceram quase ao mesmo tempo.

Você já tinha uma noção de programação? O que você aprendeu que destacaria de mais importante no seu trabalho no Estadão Dados?

Eu não tinha a menor noção de programação. Não digo que hoje eu tenha, mas hoje sei o que é um código, o que ele faz. Se abrir um código na minha frente, sei que se eu mexer aqui ou ali vai ter um efeito ali e acolá. Meu domínio de planilha não vou dizer que é 100%, mas coisas muito além do básico eu consigo fazer hoje com Excel, Google Docs. Eu fiz um curso de cartografia online, o CartoDB, que é muito bacana e que hoje eu me considero um usuário intermediário, porque ele tem muitos recursos que eu ainda não usei e que posso usar no futuro, pois dá para fazer umas coisas muito legais, mas tem uma curva de aprendizado aí. É legal, se tiver interesse, dá para criar uma conta gratuita, é limitada, mas tem muitos tutoriais, é uma ferramenta muito pratica para criar um mapa, analisar os dados ou já para publicar, porque já está na nuvem, já tem um link para embedar no seu site ou blog ou para disseminar nas redes. Então se você olhar nos trabalhos que o Estadão faz, praticamente todos os nossos dados são feitos no CartoDB. O que mais? Ferramentas para limpar dados, como o antigo Google Refine, que agora é o Open Refine, é uma ferramenta importante também. Raspagem de dados, o GitHub...ontem mesmo eu estava testando um negócio novo, que ainda nem sei para que serve direito, que é chamado import.io, é como se fosse uma ferramenta de raspagem de dados para transformar qualquer site que tenha dados estruturados em uma planilha. Você tem uma tabela num site que mostra quantos terremotos aconteceram na Colômbia e onde. Aí usando o import.io e usando o CartoDB em 20 minutos eu faço um mapa mostrando os dados de uma forma visual. Tem outros exemplos mil. A gente

usa também algumas coisas que servem para publicação de gráficos na internet, o Datawrapper, o infogram também. Tem muita coisa que a gente precisa fazer no manual ainda ou investir muito tempo para criar um coisa que consiga raspar esses dados.

Como surgem as pautas? Vocês analisam os dados que aparecem e daí surgem as pautas ou vocês já têm uma pauta e buscam os dados que tenham a ver com o assunto?

A gente um blog, onde a gente coloca as coisas que...às vezes até são desenvolvidas como pauta posteriormente...mas tem uma divulgação do IBGE, por exemplo, que a gente fica sabendo com antecedência e no dia até dá alguma coisa, ou pega alguma base que é conhecida mas que é pouco explorada, DATASUS, base de comércio exterior, a gente fez algumas coisas com os dados do Censo cruzando com os distritos de São Paulo, mostrando vários cortes. Tem algumas coisas que são mais de curto prazo e outras coisas mais de longo prazo que a gente pendura na parede e vai vendo em que vai investir de semana a semana. Geralmente eu tenho como meta entregar pelo menos uma matéria de política que envolva dados uma vez por semana para a editoria. Às vezes eu faço sozinho, às vezes envolvo a equipe. Geralmente pelo menos uma coisinha sempre tem alguma ajuda, sempre acaba virando um trabalho de equipe, mas não que isso mobilize a equipe inteira só nisso. O Burgarelli tem feito coisas para Metrópole também, coisas que ele vai desenvolvendo sem ajuda. Tem o blog e tem coisas mais de fôlego que ainda a gente pensa em fazer, coisas do tipo do Basômetro, mas que a gente ainda depende de capacidade, de tempo. E além disso tem todo o preparo para eleições. Ano que vem certamente a gente vai dar total atenção para isso e trabalhar com pesquisas, com visualizações de como no dia foram os resultados e tal. Tem bastante coisa aí no nosso mapa, digamos assim.

Em termos de retorno, não financeiro, mas você falou do Basômetro, que teve uma boa receptividade do público, por exemplo, você teria os dados de audiência e de resposta do público?

*No caso do Basômetro a gente acabou meio que perdendo esse histórico porque ele mudou de hospedagem e a gente tinha milhares de compartilhamentos do Facebook e do Twitter que sumiram. Foi um erro. Zerou acho que duas vezes já. Então a gente perdeu um pouco essa contabilidade de visualizações e usuários. O que a gente tem ‘
E esse retorno de as pessoas falarem, os cientistas políticos falando como é uma ferramenta importante inovadora, é a ferramenta mais completa de acesso aos dados do Congresso.*

O maior retorno então seria de credibilidade?

Sim. Você consegue chegar em públicos que são muito sofisticados, digamos, na verdade você pega todos os públicos, mas exige um grau de sofisticação e interesse, um cara que está muito ligado, provavelmente um difusor de idéias, um cara que vai comentar sobre a ferramenta para varias pessoas. É um público que a gente considera muito importante, mesmo que não seja numericamente grande. Não vamos competir com a matéria sobre o Corinthians porque não vamos por aí. Chega a um publico sofisticado que talvez não chegasse antes.

Que outros benefícios você observa que são criados pelo jornalismo de dados?

O tipo de jornalismo mais baseado em estatísticas em evidências mais sólidas é muito positivo no sentido de derrubar mitos. Citei o caso do PMDB que todo mundo fala, mas não

tem a menor base. Dizem que o PMDB é importante porque tem mil prefeitos, mas foram perguntar aos moradores daquelas cidades se eles sabem em quem o prefeito vai votar quem apóia? Eles sabem quem o partido do prefeito, que às vezes nem sabem qual é, está apoiando? Não sabem. E se souberem, não influi no voto deles. O eleitor claramente faz uma eleição nacional e dois anos depois faz uma eleição local e desvinculada. Não vota sempre do mesmo jeito, o que esta em jogo é diferente e ele tem noção disso. O jornalismo de dados, por trabalhar não com impressões – “ah, especialistas dizem que”, quem são os especialistas? Geralmente é um que cita o outro, que cita o outro – e fica muito claro na reeleição do Obama o caso do Nate Silver. Todos os comentaristas políticos dizendo que a eleição ia ser apertada e ele pegando todas as pesquisas, olhando objetivamente, dizia: “o cara tem 96% de chances de ser eleito”. Antes da eleição o cara era piada e depois saiu absolutamente consagrado, porque acertou o resultado da eleição em 50 estados. Por quê? Porque deixou de lado a impressão a opinião e foi olhar objetivamente o que tinha ali. Eu acho isso legal no jornalismo de dados, você poder trabalhar com dados mais concretos e poder derrubar mitos que nos assombram aí no jornalismo.

A política seria um terreno mais fértil para esse jornalismo declaratório?

Com certeza. Às vezes é inevitável. Você não tem como analisar pura e objetivamente um fenômeno político. O caso das manifestações em massa por exemplo nenhuma pesquisa antecipou, não tinha como prever aquilo. Foi um caldo de cobertura política que entrou em ebulição e ninguém muito bem entender o que estava acontecendo. Acho bacana você colocar na política essa noção de que tem que botar um pouquinho de ciência. Quem disse isso? Com base em quê? Qual é a fonte? Tentar botar um pouco mais o pé no chão e sair do etéreo, do declaratório.

Vale a pena um jornal investir em JGD? Por quê?

Acho que vale muito a pena investir em jornalismo de dados. Primeiro por ser um tipo de jornalismo que não teria muito espaço em outros meios, como televisão, rádio. Jornal e também jornal online acho que é o meio ideal para fazer esse tipo de coisa, porque agrada e impressiona um público mais exigente, que não está interessado no jornalismo por click, aquelas matérias que você clica tem três parágrafos e não diz nada. É um público que está interessado em observar fenômenos, dar dois passos atrás e observar ali o big picture, olhar a floresta e não o galho da árvore e isso você consegue com o jornalismo de dados. Acho que reforça a imagem dos meios. Acho que o Estadão gostou disso. O Estadão Dados é uma marca que eles estão promovendo, é uma coisa que reforça a marca do Estadão como algo interessado em ir fundo nas questões em mostrar as coisas como elas são e não como se pensa que elas são.

Em alguns casos se encontra outras formas de receita com o JGD. Hoje o Estadão Dados tem alguma receita envolvida, de anúncio ou comercialização de conteúdo?

Não. Não tem no momento. Em algumas reuniões a gente já tocou no assunto, mas a gente não é cobrado. Vem mais da gente, que pensa “isso aqui poderia render algo além do interesse jornalístico para o público geral e gerar um produto para o público X”. Mas ainda não fizemos nenhum avanço nesse sentido. Talvez nos próximos meses isso ocorra com uma iniciativa da Agência Estado, o Broadcast Político. Tem o Broadcast Financeiro que é um produto para o mercado financeiro e agora tem o Broadcast Político, voltado para o mundo político. Talvez tenha um produto do Estadão dados que acabe vinculado a isso e não seria

disponível para todo o público, mas só para os pagantes daquele produto específico, e aí a gente estaria levando um pouco de valor também. Mas não é algo que tenha sido cobrado. A ideia toda foi de melhorar o conteúdo do jornal. O online também, mas a gente está produzindo para o papel mesmo. Nas edições de domingo são raras às vezes em que o Estadão Dados não está na edição dominical. É a edição que traz um conteúdo diferenciado, não está calcado necessariamente no noticiário da semana.

ANEXO C - Entrevista Gustavo Faleiros

Como começou a trabalhar com Jornalismo Guiado por Dados?

Quando comecei a trabalhar com jornalismo de dados ainda não se chamava jornalismo de dados. Eu estava trabalhando no O Eco, é o meu veículo-mãe, minha casa. Comecei a trabalhar em 2006 como repórter, hoje sou editor. Trabalha com o modelo de mídia sem fins lucrativos, como a Agência Pública e o Repórter Brasil. Hoje sou parte da associação e por isso sou tão ligado, os projetos estão todos ligados à associação. Quando eu estava trabalhando como repórter e editor uma das coisas que mais cobria eram queimadas, envolvia muito essa questão de dados, planilhas e principalmente imagens de satélite, de fogo. A gente fazia visualização com imagens de satélite e isso foi o que me levou a trabalhar com mapas interativos, com Google Maps e Google Earth e a partir daí não parou mais. Comecei descobrindo outras bases de dados ligadas àquilo, como limites de unidades de conservação, dados de desmatamento e assim fomos. Ainda em 2008 a gente fez um projeto de monitor de fogo dentro das áreas protegidas. A gente olhava os boletins de queimadas do IBAMA e do INPE e identificava quais áreas protegidas e unidades de conservação estavam sofrendo impacto e a partir disso gerava reportagens, ligava para os chefes dos parques, pedia fotos, para qualificar a informação que a gente tinha achado nos boletins e com os dados.

As bases de dados surgiram naturalmente então dentro do processo de apuração?

Exato. A gente foi usando e descobrindo que, por exemplo, juntando bases de dados que é disponibilizada diariamente sobre queimadas e outra com os limites de unidades de conservação a gente conseguia criar uma visualização interessante mostrando fogo dentro da Chapada Diamantina, fogo dentro de Itatiaia. Essas foram as primeiras experiências. A gente fazia coisas muito simples, bem básicas, mapas e no máximo gráficos para mostrar tendências. Foi uma coisa que partiu principalmente de mim. Logo de cara já fiquei viciado, fascinado por isso, descobri novas imagens de satélite, o que me puxou muito foi descobrir as imagens de satélite como uma forma de comunicar uma nova dimensão do problema. Então tinha muito slideshow com imagens de satélite, colocava imagens de satélite sobrepostas com os limites das áreas de conservação. Mais ou menos por aí que as coisas começaram a rolar. E nesse projeto do monitor a gente criou uma pequena força-tarefa dentro de O Eco que olhava os dados e que cada uma ficava responsável por ligar para um chefe de cada área protegida. Disso surgiu a ideia de que a gente poderia fazer um geoblog, mapear correspondências em distintas áreas da Amazônia. Já naquele momento a gente estava fazendo experimentos e propostas para criar uma coisa que a gente imaginava que seria um dia o infoamazonia, não com esse nome, a gente imaginava que seria monitor. Já em 2008 enviamos uma proposta, mas o recurso não foi aprovado. E na primeira vez que eu vi o termo jornalismo de dados, acho que em 2010 imediatamente identifiquei que era exatamente o que a gente estava fazendo e que estava ligado também a um movimento de transparência de dados. E que a gente estava num momento interessante. Aquilo que a gente estava pensando há dois anos estava enquadrado naquele momento.

Sobre o financiamento, quais as maiores dificuldades que enfrentou para tirar o InfoAmazônia do papel?

Sempre foi mostrar como isso podia ser um projeto de relevância jornalística, não só de visualização ou gráfica. E a outra questão, que continua sendo uma questão, além da relevância jornalística, é mostrar o que você está provendo para a sociedade que é diferente ou mais informativo que um texto ou alguma outra forma jornalística tradicional.

E qual foi o seu argumento?

O argumento é que a gente tem uma capacidade com essa ferramenta que...do lado jornalístico já tem toda uma questão de desenvolver tecnologia e experimentar novas ferramentas que, em si, dependendo de quem são os doadores já é atrativo, mas para outras ONGs ligadas ao Meio Ambiente ou à transformação social, que atuam como doadores nesse projeto, o argumento é que a gente está gerando uma capacidade de organização da informação, que é tão vasta hoje e que a gente realmente cria um sentido que esta um pouco disperso por aí. Essa idéia de mapas tem muito de conceitual por trás pela questão das camadas. Gerar camadas que tenham grande quantidade de dados por trás que podem ser lidas para gerar um contexto. Toda história por trás do InfoAmazônia é que seja uma ferramenta que, claro sirva para informar, mas que gere esse contexto mais aprofundado e de larga escala. Como é que eu consigo comunicar para um público que não está familiarizado com a Amazônia a dimensão de 7 milhões de quilômetros quadrados, nove países, a maior bacia hidrográfica do mundo? Como é que você visualiza a escala do que está acontecendo nessa região que é de suma importância global? A gente acredita que o InfoAmazônia, com todos os efeitos que ainda podem ser solucionados e todos os desafios, ainda possa trazer uma imagem, esse retrato da transformação que está ocorrendo na região.

A forma tradicional jornalística então não estava dando conta de mostrar essa dimensão?

Acho que não. Tem coisa que o jornalismo tradicional nunca deixará de servir, contar histórias, criar empatias, quantas reportagens foram feitas estes anos de extrema qualidade. Mas vejo nosso serviço em um espectro diferente. A gente agrega, organiza e distribui dados. A gente distribui muitos dados. Hoje o InfoAmazônia tem no mínimo 10 GB de dados organizados. Depois a gente organiza por visualização de mapas ou buscas e geralmente a gente distribui. A gente permite que as pessoas baixem os dados, retirem e customizem os dados e mapas e distribua na própria página ou nas redes sociais. Porque aí a gente também está pensando num aspecto de serviço também. Quem gostaria de usar o InfoAmazônia? Pode ser um jornalista que cobre essa área e não tem essa capacidade de gerar visualizações que a gente já em praticamente prontas.

Como foi feita essa transição? Exigiu o aprendizado de novas habilidades?

Se a gente tiver pessoas que tenham experiência em criar e vender projetos isso seria sempre muito benéfico, mas não vejo isso como uma habilidade que deva ser considerada no espectro do jornalista. Se for considerar dentro dos processos normais do jornalista, vender a sua pauta é básico. A única coisa é que num projeto como o InfoAmazônia a gente está vendendo uma plataforma em torno de um assunto. Tem muita dificuldade. Eu fui repórter até dois anos atrás quando comecei a me tornar um gestor de projetos. A partir daí comecei a conhecer como é que se encontra fonte de financiamento, como faz propostas, orçamento, cronograma. Não é nenhuma ciência maluca. É mais prática e disposição. É chato ficar escrevendo propostas, relatórios de cumprimento de metas, mas acho que isso não vai ser diferente para maioria de nós daqui para frente. Todo mundo que quiser fazer alguma coisa que goste tem

que se comprometer um pouco a vender seu peixe. O jornalismo pós-industrial é a melhor definição mesmo. Não tem mais uma linha muito clara em que você só tem uma função, fazendo uma coisa por dia. Tive essa experiência no Valor Econômico e foi fantástico, eu fico às vezes saudosista, mas a verdade é que mudou bastante e bem rápido.

O envolvimento sem dúvida é maior. Me sinto muito envolvido no que eu faço em tentar fazer a coisa vingar de uma forma que eu não tinha trabalhando para uma empresa.

Como foi a formação da equipe? Você sentiu que seriam necessários outros profissionais?

Com certeza, a gente teve que encontrar outros profissionais e hoje o time que eu trabalho é basicamente de programadores e designers. Tem o time do Eco ainda que é baseado no Rio de Janeiro, mas aqui em São Paulo eu tenho um time que é só de desenvolvedor e designers e programadores. São cinco profissionais e um estagiário, mas a gente contrata outros por projeto. Antes que o InfoAmazônia decolasse e eu começasse esse esquema de bolsista do Centro Internacional de Jornalistas, que ajudam o InfoAmazônia, eu lancei com uma amiga minha com o Instituto Claro um projeto chamado MapasColetivos.com. A gente criou esse projeto de colaboração cidadão e esse foi o primeiro que eu fiz com essa visão de que a gente precisava trabalhar com programadores e não só com jornalistas. Então a gente teve uma experiência dolorida, teve muito menos dinheiro que o InfoAmazônia, então a gente sofreu muito para colocar esse projeto de pé e foi o meu primeiro aprendizado. Depois dessa experiência foi um pouco mais fácil fazer o InfoAmazônia, a gente já tinha uma noção de custo. Então inclusive quando eu fui fazer a captação do InfoAmazônia eu joguei bem para o alto a questão de desenvolvimento, a dimensão que a gente queria que tivesse. A gente trabalhou diretamente com o time do MapBox, que faz a tecnologia de mapas. Hoje é a empresa que mais cresce na área de mapas, são competidores diretos do Google. E no ano passado eles ainda eram uma empresa pequena. Então a gente contratou eles para fazer o primeiro desenvolvimento. E foi uma experiência muito interessante, porque eles são muito profissionais, talvez por isso estejam crescendo tanto, e a gente aprendeu muito com ele, principalmente essa questão de projeto, quantas pessoas têm que estar envolvidas, não trabalhar somente com um programador, mas vários que possam um cobrir as dificuldades dos outros. E ter uma pessoa específica para o design e não ter um programador que você vai jogar tudo em cima dele. Então foi bem legal.

Nossa experiência com o InfoAmazônia, ao contrário do Mapas Coletivos que foi muito traumática essa foi muito legal mesmo, a gente teve essa experiência legal com os caras e, depois disso, a gente tinha recursos para trabalhar com uma equipe de manutenção, que é basicamente a equipe que está comigo até hoje, são uns caras muito talentosos, da Casa de Cultura Digital de São Paulo, do estúdio Kazume e do Memelab. A gente se juntou para criar o Laboratório de Inovação e Jornalismo Ambiental, que é o projeto que estou tocando no Eco este ano. A gente conseguiu recursos para essa equipe meio que para acelerar o processo de criação de certos produtos. A linha de produção que estou hoje é menos de produzir notícias do que de fazer produtos que apoiem as notícias.

Como subproduto do InfoAmazônia a gente criou um tema de wordpress que permite aos jornalistas criar coisas produtos semelhantes ao InfoAmazônia, no EcoLab você pode procurar o JEO. Outro produto é o Oxpeckers, primeiro projeto que a gente fez fora do escopo do eco, pessoas contrataram gente para fazer a tecnologia faz monitoramente de caça de rinocerontes na África do Sul, é um grupo de jornalistas da África do Sul. E uma outra fonte de receita Um modelo que a gente criou agora, criar aplicações que a gente vende e gera receita para fazermos os nossos próprios produtos. Nos próximos meses a gente vai fazer outro para um grupo de jornalistas da Indonésia e depois para um grupo de jornalistas do Quênia. Tudo baseado na tecnologia que a gente criou por trás do

InfoAmazônia. Foi muito interessante trabalhar com outros profissionais que não jornalistas por causa disso. Tenho as idéias e vou buscar os recursos para executá-las baseado na opinião técnica das pessoas que trabalham comigo, se são viáveis ou não.

Dos seis, a estagiária é de programação também?

Falei seis, mas são sete. Tem eu, tem outro jornalista que é o editor do InfoAmazônia, que faz moderação de histórico e tudo, a estagiária é analista de computação, mas a gente força ela a fazer um pouco do trabalho jornalístico para aprender como são organizadas as informações. Tem um engenheiro que trabalha como programador e cuida principalmente de bases de dados. Todos esse projetos têm alguma complexidade de lidar com as bases de dados. Tem o programador sênior, que é quem desenvolveu os projetos e tem um gerente de tecnologia que é quem me ajuda a decidir quantos programadores vão ser necessários, se a gente vai precisar contratar mais pessoas ou não. Tem um designer fixo e vira e mexe a gente contrata outra designer que faz layout tátil, interativo.

E se tiver uma demanda maior, contrata outras pessoas?

Geralmente a gente contrata um designer para fazer o inicial e depois a gente usa nossa capacidade interna para fazer a implementação. Mas a partir do ano que vem a gente pretende fazer projeto com móbile também aí vamos ter que contratar gente, porque não é a nossa especialidade, não está dentro da capacidade que a gente tem internamente.

Qual o investimento no InfoAmazônia e quem financiou?

É um projeto que tem três financiadores. Um que paga o meu trabalho que é o Centro Internacional de Jornalistas. Tem um financiador inicial que foi uma rede chamada CDKN, Climate Development Knowledge Network, que financiou nosso parceiro norte-americano, na verdade, que se chama Internews, e finalmente a gente tem um doador anônimo, que faz projetos, mas pede que não seja divulgado o nome da empresa. Mas não é nada secreto, é só o modelo que eles têm nos Estado Unidos, de ONGs que por questões as vezes até religiosos que preferem não fazer propaganda sobre filantropia. De forma geral, teve um custo de mais ou menos US\$ 150 mil.

É acumulado?

É um custo acumulado. Depois a gente manteve uma manutenção pontual que, se muito adicionou US\$ 20 mil-30 mil, mas é o custo de pagar os jornalistas que fazem atualização. Depois, até a fase final de desenvolvimento não teve nenhum custo de grande monta para manutenção. O que a gente está buscando agora é buscar mais recursos para criar uma área de colaboração.

Como é a divisão entre jornalistas e os que não são jornalistas?

Diria que 70% é tecnologia e desenvolvimento e 30% para coordenação e mão de obra jornalística. Mas tem outra coisa a ser levada em conta que é o fato de que o que está agregado em termos de reportagens ao InfoAmazônia teve custos externos. Para a gente começar, teve projeto anterior de US\$ 200 mil que basicamente financiava reportagens em nove países sobre a Amazônia. E eu não computo isso. Pagava viagens para ir para mata, jornalistas do Suriname ou Guiana Francesa para fazer matéria para a gente e assim por

diante. Ou a Pública. A gente trabalha em parceria com a Agência Pública, aqui em São Paulo e eles financiaram várias reportagens, com viagens para a Amazônia, que a gente está agregando na nossa base. O custo jornalístico a gente delegou. O que o InfoAmazônia faz é delegar reportagens. A gente não tem uma produção primária de reportagens, nossa produção primária é de mapas.

Como surgem os temas para as matérias? Parte-se do que está disponível ou tem uma ideia e buscam as bases de dados?

Tem os dois processos. Várias bases de dados que estão aqui são de grandes temas que a gente já sabia que fazem parte dos principais vetores de mudança na Amazônia, geração de energia, petróleo e gás, etc. no ano passado, em dezembro a gente trabalhou muito próximo do Instituto Socioambiental, que tem a rede RAISG. E essa rede está analisando os grandes vetores de transformação da Amazônia. No ano passado, além do desmatamento e fogo, que a gente tinha mapeado, colocamos também estradas, elétrica, petróleo. Então a gente conseguiu, e eu tenho muito orgulho, porque não conheço nenhuma outra plataforma que permita ver com tanta clareza e atualizados os dados dos nove países da Amazônia em uma só plataforma. E aí tem o outro processo que foi dentro do que você me perguntou. Agente começar, por exemplo, que tem muita reportagem feita com dados do Imazon, a ONG que faz os dados do desmatamento em paralelo ao governo brasileiro. Então a gente fez os mapas do Imazon. Fez os mapas específicos para histórias, por exemplo.