

UNIVERSIDADE DE SANTO AMARO

Curso de Comunicação Social

Habilitação em Jornalismo

Ronaro Ramos Costa

**JORNALISMO DE DADOS COMO AGENTE DA CULTURA DA
TRANSPARÊNCIA E DA
CONSOLIDAÇÃO DEMOCRACIA NA ERA DIGITAL**

São Paulo

2016

Ronaro Ramos Costa

**JORNALISMO DE DADOS COMO AGENTE DA CULTURA DA
TRANSPARÊNCIA E DA
CONSOLIDAÇÃO DEMOCRACIA NA ERA DIGITAL**

Trabalho de conclusão de Curso de
Comunicação Social - Habilitação em
Jornalismo, da Universidade de Santo Amaro
- UNISA.

Orientador: Prof. Me. Guilherme Sardas

São Paulo

2016

RAMOS COSTA, RONARO

JORNALISMO DE DADOS COMO AGENTE DA CULTURA DA
TRANSPARÊNCIA E DA CONSOLIDAÇÃO DEMOCRACIA NA ERA
DIGITAL / RONARO RAMOS COSTA. -- São Paulo , 2016
81 f.

TCC Especialização (Comunicação Social) - Universidade de Santo
Amaro, 2016

Orientador(a): Guilherme Sardas

1.Jornalismo de dados. 2.cultura de transparência.
3.empreendedorismo. I.Sardas, Guilherme, orient. II.Universidade de
Santo Amaro III.Titulo

RONARO RAMOS COSTA

**JORNALISMO DE DADOS COMO AGENTE DE CONSOLIDAÇÃO DA
CULTURA DE TRANSPARÊNCIA**

Monografia apresentada para obtenção de título de Bacharel em
_____ ao Curso de _____ da
Universidade de Santo Amaro – UNISA sob orientação da Área de Concentração
_____.

Data de Aprovação: ___/___/___

BANCA EXAMINADORA

(nome do orientador e Titulação)

(nome do professor e titulação)

(nome do professor e titulação)

CONCEITO FINAL: _____

Aos filhos que não tenho,
e espero nunca ter.

AGRADECIMENTO

Ao meu pai e minha mãe,

Aos professores,

Ao orientador Guilherme Sardas,

A Mirna Tornus,

A Thiago Mali,

A Rodrigo Bugarelli,

A Sérgio Spagnuolo,

A todos que contribuíram com meu aprendizado,

Obrigado.

“No meio do caminho tinha uma pedra...”

Carlos Drummond de Andrade

RESUMO

Com o avanço tecnológico, o mercado jornalístico tem se alterado, a forma como se tem lido notícias está em constante mudança, o jornalismo impresso tem perdido cada vez mais espaço para o digital, a convergência é inevitável. O jornalismo de dados é parte desta mudança na forma de se ler e se fazer jornalismo. Com base nessa situação, foi feito um estudo acerca do jornalismo de dados e da cultura de transparência, para buscar entender as novas características do jornalismo e como o jornalismo de dados fortalece a cultura de transparência, fomenta a participação política da sociedade e por fim consolida o Estado democrático de direito.

Palavras-chave: Jornalismo de dados, Cultura de transparência, empreendedorismo.

ABSTRACT

With the technological advance, the journalistic market has changed, the way you read news is constantly changing, printed journalism has lost more and more space for digital, convergence is inevitable. Data journalism is part of this change in the way we read and do journalism. Based on this situation, a study was made of data journalism and a culture of transparency, in order to understand the new features of journalism and how data journalism strengthens the culture of transparency, promotes the political participation of society and ultimately consolidates The democratic rule of law.

Keywords: Journalism, Data, Culture of Transparency, Entrepreneurship.

LISTA DE FIGURAS

Figura 1 - Gráfico Crista de Galo	16
Figura 2 - Pirâmide Invertida de Jornalismo de Dados	24
Figura 3 - Pirâmide Invertida de Jornalismo de Dados (completo)	29
Figura 4 - Página inicial do Basômetro	50
Figura 5 - Projetos de Jorn. de dados da organização Transparência Brasil	51
Figura 6 - Mapa do desmatamento da floresta Amazônica	53
Figura 7 - Gráfico de reportagem Agência Pública	55
Figura 8 - Gráfico de reportagem do Estadão Dados	57

SUMÁRIO

1 INTRODUÇÃO	11
2 JORNALISMO DE DADOS NO CONTEXTO DA ERA DIGITAL	15
2.1 RAC e Jornalismo de Precisão	18
2.2 Jornalismo de dados: do Conceito à Prática	20
2.3 Empreendedorismo e Jornalismo de Dados	32
2.3.1 Descritivo do Ferramental Necessário para a Atuação com Jornalismo de Dados.....	34
2.4 Formação de Jornalistas	37
3 DEMOCRACIA E CULTURA DE TRANSPARÊNCIA	39
3.1 Big Data	42
3.2 Vazamento de Dados	43
3.3 Lei de Acesso à Informação	46
3.4 Projetos Jornalísticos Baseados na Cultura de Transparência na Era Digital	48
3.5 Exemplos de Reportagens	54
4 CONSIDERAÇÕES FINAIS	58
5 REFERÊNCIAS BIBLIOGRÁFICAS	63
6 ANEXOS	69

1 INTRODUÇÃO

O jornalismo tem passado por uma profunda mudança com o advento das tecnologias digitais nas últimas décadas. Os veículos impressos, rádios e emissoras de TV estão tendo que se adaptar ao mundo online. Mas essa não é a primeira vez que uma nova tecnologia revoluciona o jornalismo.

A grande revolução no jornalismo surge com a invenção do tipo móvel por Johannes Gutenberg em 1445 (SILVER, 2012, p.10). Antes do tipo móvel, o que existia era algo parecido com jornal, a gazeta manuscrita, feita à mão e disponível para poucas pessoas. Costella (1984, p.17) lembra que, antes do tipo móvel, o processo de comunicação era reduzido à quase exclusivamente à informação verbal. O que havia mais próximo de jornalistas eram os jograis, cantores de versos trovadores e outros artistas itinerantes que, de reino em reino, de vilarejo a vilarejo, literalmente, contavam as notícias, em versos, dos lugares por onde haviam passado, eram os jornalistas sem jornais. Muito antes, em Roma, havia algo que poderia ser considerado como jornal, também sem a figura do jornalista. O presidente do Colégio de Pontífices romanos redigia documentos políticos chamados “*Annales Maximí*”, a publicação era afixada anualmente diante da casa do “*Pontifex Maximus*” (COSTELLA, 1984, p17).

Costella (1984) afirma ainda que o jornal impresso surgiu século e meio após a criação da prensa de Gutenberg, pela convergência de vários fatores históricos, dos quais se destacam o papel, o tipo móvel e o correio. Cada um deles, sendo independentes quanto a suas origens, são componentes importantes do surgimento do jornal impresso. O correio, usado para o tráfego de informações, transportadas sob a forma de cartas, foram reunidas em coletâneas: compondo as chamadas gazetas manuscritas, isto é, jornais feitos à mão (COSTELLA, 1984, p.20).

Com a criação de Gutenberg, o mundo entrou em uma nova era comunicacional. Segundo Silver (2012, p.11) o custo de produção de livros diminuiu quase trezentas vezes, levando a produção de livros a aumentar cerca de trinta vezes no primeiro século pós-prensa. O tipo móvel se espalhou com rapidez pela Europa, da Alemanha de Gutenberg em 1445, até 1505, já haviam se instalado tipografias por quase toda Europa e, até 1700, o invento de Gutenberg já havia conquistado a México, em 1533; Lima, 1584; Guatemala, 1660; parte dos estados

que juntos formariam os Estados Unidos e Paraguai em 1700. No Brasil, a prensa chegou apenas em 1808 com a vinda da família real portuguesa (COSTELLA, 1984, p.62).

O invento de Gutenberg levou o mundo para a era do conhecimento. Se no mundo, antes da prensa não havia motivação para pesquisas científicas ou o registro massificado de pensamentos, com o advento da prensa, filósofos, cientistas e escritores puderam registrar seus feitos em livros, levando o conhecimento *a posteriori* (SILVER, 2012, p.11).

O tipo móvel trouxe benefícios para o mundo, mas teve seus efeitos colaterais. A qualidade da informação variava muito, “a lista de *best-seller* logo passou a ser dominada por textos hereges e pseudocientíficos”, como desta Silver:

Agora, os erros podiam ser produzidos em massa, como na chamada “*Wicked Bible*” [*A bíblia maldita*], que cometeu o mais infeliz erro tipográfico na história da página impressa: “Cobiçarás a mulher do próximo” (SILVER, 2012, p.11).

A revolução na comunicação causada pela prensa de Gutenberg foi um “divisor de águas” na sociedade, a civilização saiu do obscurantismo informacional medieval para se aventurar no mundo das ideias, criando e modificando tudo a sua volta. Os primeiros a tirarem proveito do tipo móvel a usaram para evangelizar. Como toda nova ideia encontra resistência, logo a Europa afundou em guerras, a maioria das cruzadas religiosas foram causadas pela Reforma Luterana, proposta por Lutero em suas 95 teses (SILVER, 2012, p.12)

Após a tipografia, o mundo teve outras invenções que modificaram as bases da comunicação e do jornalismo, como o rádio, a televisão e o cinema, mas a nosso ver, nenhuma foi e tem sido tão significativa quanto à internet, conforme as transformações propiciadas por esse meio aparecem em nosso tempo.

A *World Wide Web* (Rede de Abrangência Mundial) não foi criada de imediato, antes era necessário criar o suporte para o ambiente digital. Em 1945, surgem os primeiros computadores na Inglaterra e Estados Unidos. Grandes máquinas que, por muito tempo, foram usadas apenas por militares para cálculos científicos: os primeiros computadores eram operados por cientistas de uniforme branco em uma sala refrigerada. Apenas a partir dos anos 60, Estados e grandes empresas começaram a desfrutar dos benefícios do computador (LÉVY, 1999, p.31).

Nos anos 70, o desenvolvimento e a comercialização do microcomputador foi o início de uma nova era na comunicação, além das benfeitorias provocadas pela computação na produção industrial, como robótica, linhas de produção flexíveis, máquinas industriais com controles digitais. O computador pessoal passou a ser realidade nas universidades e nas residências, sendo usado tanto para trabalho, estudo e diversão (LÉVY, 1999, p.31).

O que revolucionaria a comunicação para a forma como conhecemos hoje foi concebida em 1969, nos Estados Unidos. A internet nasceu no período conhecido como Guerra Fria, como um projeto de comunicação do Departamento de Defesa estadunidense para garantir uma comunicação emergencial, caso o país fosse atacado por forças estrangeiras. O projeto foi concebido pelo *Advanced Research Projects Agency* (Arpa – Agência de Pesquisa e Projetos Avançados) para o serviço militar e recebeu o nome de Arpanet. Era a rede nacional de computadores. Após a Arpanet ter se popularizado nas universidades, sendo usada, principalmente, para pesquisas militares e transferências de arquivos extensos por *e-mail*, novas redes surgiram ao longo dos anos 1970 e 1980, como a Bitnet - *Because It's Network* e a CSNET - *Computer Science Network* (Rede de Ciência da Computação) (FERRARI, 2003, p.15).

Em 1980, Tim Berners-Lee inicia o projeto *World Wide Web* (WWW), baseado em hipertexto e sistemas de recursos para a internet. Em 1990, Berners-Lee, com a ajuda de Cailliau, cria a *WWW*, usada até hoje em *web sites* (FERRARI, 2003, p.16).

A internet parecia estar pronta, mas o designer e pesquisador Jean François Groff considerou que aquela interface pouco amigável poderia ser aperfeiçoada. Groff convidou Berners-Lee para ser seu aluno no projeto *InfoDesign*, que realizou inovações de *design*, arquitetura e protocolos, tornando a internet mais amigável (FERRARI, 2003, p.16).

A internet não demorou a crescer. Em 1996, já existiam 56 milhões de usuários no mundo. Em 2000, já havia 400 milhões de pessoas online e, em 2015 já são 3,2 bilhões de conectados, segundo a Organização das Nações Unidas (FERRARI, 2003, p.17).

Com o crescimento do número de usuários, cresceu também o tráfego de dados, em 1996 já haviam sido enviadas 95 bilhões de mensagens eletrônicas nos Estados Unidos (FERRARI, 2003, p.16).

O tipo móvel mudou o mundo em 1445, ao possibilitar a publicação de novas ideias com a produção industrial, sem a necessidade de copistas. Graças à tipografia, o jornal pode finalmente evoluir, século e meio depois. Com o advento do computador e internet no século XX, a comunicação voltou a mudar. O acesso à informação tornou-se cada vez mais acessível ao passo que a internet foi se popularizando pelo mundo. Novas ideias podem ser discutidas em escala global. As distâncias tornaram-se meros inconvenientes driblados pela possibilidade de se conversar por videoconferência ou mensagens de texto. Assim como a comunicação na sociedade mudou com a internet, o jornalismo também mudou.

A internet democratizou a comunicação, dando voz a qualquer um, em qualquer lugar, bastando estar conectado para poder se comunicar, ou buscar informação sem depender, unicamente, dos meios tradicionais de comunicação.

Uma série de evidências de mercado e números do consumo de comunicação no mundo, a mídia impressa tem sido fortemente abalada pela concorrência das mídias digitais perdendo boa parte dos seus leitores para o digital. A “Pesquisa Nacional de Mídia: 2015”¹, feita pela Secretaria de Comunicação Social, aponta que internet é o meio de comunicação que mais cresce entre os brasileiros, um quarto da população já acessa a rede diariamente, com intensidade de mais de quatro horas por dia. A pesquisa também mostra que setenta e cinco por cento dos entrevistados afirmaram não ler jornais. Cinquenta e um por cento dos entrevistados afirmaram não acessar a internet e quarenta e um por cento informaram que a razão pela qual não usa a internet é a falta de habilidade com o computador. Contudo, mídias impressa e audiovisual também têm se adaptado ao ciberespaço. O jornalismo tradicional tem sofrido mudanças radicais com o advento da internet e do jornalismo digital, mas o digital também tem feito surgir um novo jeito de fazer jornalismo, com a cultura de transparência e as novas tecnologias, está sendo possível produzir notícias mais assertivas, atenuando, assim, a dependência ao chamado jornalismo declaratório.

O jornalismo de dados é uma das novas modalidades jornalísticas que surgem das tecnologias computacionais, unidas ao ambiente digital, valendo-se de toda a infinidade de informação produzida na *web* e apresentando importantes contribuições para as sociedades democráticas.

1 In: <http://www.secom.gov.br/atuacao/pesquisa/lista-de-pesquisas-quantitativas-e-qualitativas-de-contratos-atuais/pesquisa-brasileira-de-midia-pbm-2015.pdf>

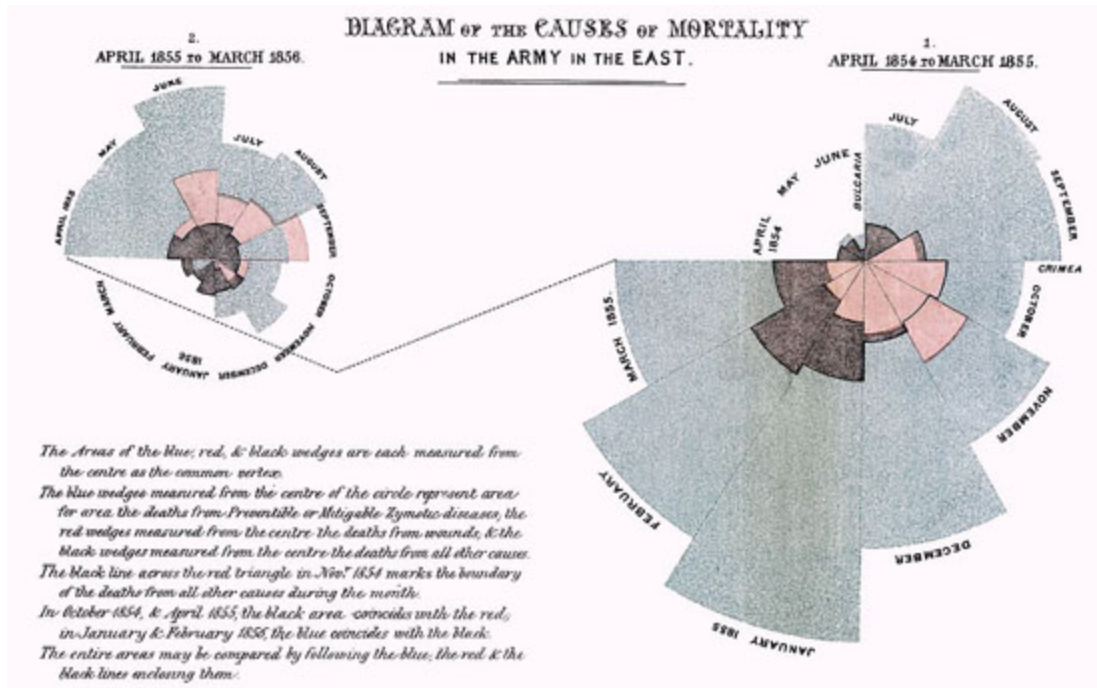
2 JORNALISMO DE DADOS NO CONTEXTO DA ERA DIGITAL

Jornalismo de dados é o termo que melhor define jornalistas que produzem reportagens utilizando fontes primárias, em especial informações em grande volume contidas em planilhas ou bancos de dados.

A partir dessa definição, o jornalismo de dados não é uma novidade surgida com o advento da internet. Muito antes de o mundo ficar conectado por redes de computadores, dados já eram usados para produzir reportagens, mesmo sem o auxílio de computadores. Em 1821, o jornal britânico *The Guardian* usou uma lista obtida de fonte não oficial, com informações sobre o número de alunos de escolas da cidade de Manchester e o custo de cada aluno. Com a lista, foi possível mostrar o número de alunos que recebiam educação gratuita, muito maior que os números oficiais revelavam (BOUNEGRU, 2012).

Outro exemplo de jornalismo de dados do século XIX, o artigo "*Mortalidade no Exército Britânico*", da enfermeira de guerra Florence Nightingale, publicado em 1858. Ela usou gráficos para defender, no parlamento inglês, o aperfeiçoamento do serviço de saúde do exército britânico. O gráfico mais famoso é conhecido como *crista de galo*, em uma espiral de seções em que cada uma representa as mortes a cada mês; no gráfico, Nightingale destacava que a maioria das mortes foi em decorrência de doenças que poderiam ser prevenidas (BOUNEGRU, 2012).

Figura 1 - Gráfico Crista de Galo



Fonte: Manual do Jornalismo de dados²

O jornalista e repórter Caco Barcellos (2006), usou o jornalismo guiado por dados para investigar, no início dos anos 1990, o assassinato de pessoas pela Polícia Militar de São Paulo. Em seu livro “*Rota 66 - A História da Polícia que Mata*”, Barcellos relata como criou um banco de dados sobre os indivíduos mortos durante o patrulhamento pela cidade a partir de 1970, quando a Polícia Civil e a Força Pública se juntaram para criar a Polícia Militar de São Paulo. As primeiras informações, não oficiais, foram colhidas de parentes ou amigos de pessoas mortas pela polícia que procuravam o corpo no Instituto Médico Legal (IML). Outra fonte não oficial, usada no banco de dados em construção, foram as notícias do jornal *Notícias Populares*, já que o diário tinha acesso aos Boletins de Ocorrência (B.O.) da PM para produzir reportagens. Como a PM não quis liberar os B.O.s para o jornalista, consultar as notícias do NP foi a alternativa encontrada para ter acesso aos B.O.s. A fonte oficial usada pelo jornalista foram os registros do IML. Após a pesquisa ter avançado, Barcellos usou um computador para criar e organizar um banco de dados. Com o cruzamento de informações, como cor da pele, faixa etária das

² In: <http://datajournalismhandbook.org/pt/>

vítimas e policiais que estavam nas ocorrências, ele pode identificar o perfil dos policiais que mais matam e o perfil das principais vítimas.

Outra importante matéria guiada por dados foi a do jornalista Mário Rosa, publicada no *Jornal do Brasil*. Em 1992, Rosa usou o Sistema Integrado de Administração Financeira do Governo Federal (Siafi) para investigar o superfaturamento na compra de leite em pó pela Legião Brasileira de Assistência, presidida pela então esposa do presidente da República Fernando Collor de Melo, Rosane Collor. A matéria continha dados jamais vistos, como números de ordens bancárias e de empenhos (reservas feitas no Orçamento da União). O Siafi registra os gastos do governo federal com pagamentos ou repasses a empreiteiras, fornecedores, estados e municípios. O acesso do jornalista ao Siafi só foi possível porque o então senador Eduardo Suplicy (PT-SP) lhe entregou a sua senha pessoal de parlamentar. Após essa e outras reportagens, o Governo Federal liberou o acesso de jornalistas ao Siafi, e o sistema tornou-se uma das primeiras bases de dados públicas do Brasil (TRÄSEL, 2012).

O trabalho com dados mostrou beneficiar a construção da notícia pautada por dados. Com a popularização da internet e avanços na tecnologia computacional, a produção de dados aumenta e o mundo entra na era da informação. Segundo a IBM (2012), diariamente o mundo produz 2,5 quintilhões de *bytes*. Toledo (2014, p.7), afirma que se cada *byte* fosse um litro de água, seria possível encher o oceano Pacífico de zeros e uns todo ano. Para explicar a grande quantidade de dados sendo produzida diariamente, cientistas e engenheiros da computação cunharam um novo termo para o fenômeno: *Big Data* (JUNIOR, 2012, p.57). Esses dados ainda não são acessíveis ao público, em sua grande maioria; são pulsos magnéticos, sinais eletrônicos, *bits* e *bytes* ou zeros e uns, dispostos em linguagem binária que só podem ser compreensíveis aos olhos do grande público através de interfaces que transformam zeros e uns em planilhas, imagens, vídeos, sons. Mesmo assim precisam ser organizados em planilhas, tabelas e banco de dados para serem considerados informação (TOLEDO, 2014, p.8).

Dessa forma, o jornalismo de dados surge com o propósito de analisar e organizar os dados separando a informação útil da informação irrelevante, usando o faro jornalístico para encontrar notícias de interesse público ao, digamos, “entrevistar” planilhas.

2.1 RAC e Jornalismo de Precisão

A Reportagem Assistida por Computador (RAC) surgiu em 1952, nos Estados Unidos, quando a rede de TV estadunidense CBS usou a técnica para prever o resultado da eleição presidencial daquele ano (BOUNEGRU, 2012). A RAC consiste em incorporar a prática da ciência social com a do jornalismo tradicional visando obter precisão na informação que será publicada nos jornais e para criar um processo de produção jornalístico mais eficaz. Para tanto, é necessário a utilização de programas de computador para aproximar o jornalista da fonte primária, possibilitando-o de ter melhor compreensão da realidade (PENA, 2013, p.174). No início dos anos 1970, o termo “Jornalismo de precisão” foi designado para descrever esse tipo de apuração jornalística.

“[...] banco de dados e planilhas de calculo são as ferramentas mais utilizadas pela RAC. Por meio delas, os jornalistas contextualizam informações, conferem números, encontram novos indícios, ilustram reportagens com gráficos e tabelas, checam dados e até produzem pesquisas qualitativas”. (PENA, 2013, p.175).

O jornalista Philip Meyer estudou Ciências Sociais em Harvard por um ano. Em 1967, Meyer volta ao seu trabalho no conglomerado Knight Newspapers e pede licença aos seus superiores para participar da cobertura distúrbios sociais que acontecia na cidade de Detroit e se junta com a equipe do jornal *Detroit Free Press*, pertencente ao grupo *Knight*. Meyer sugeriu à equipe a aplicação de um questionário representativo dos bairros afetados pelos distúrbios, com o objetivo de pesquisar suas raízes. O psicólogo Nathan Caplan, do Instituto de Pesquisa Social da Universidade de Michigan (ISR), que também estava interessado em investigar as causas que levaram à revolta popular, colaborou com Meyer no estudo (TRÄSEL, 2014, p.97).

Meyer e Caplan rascunharam o questionário, contrataram trinta professores negros para fazer as entrevistas. O pesquisador do ISR, John Robinson, recrutado por Caplan, traçou uma amostra aleatória de endereços na área das manifestações.

Com amostra representativa de 437 residentes negros, Caplan e Meyer começaram a analisar os dados dos questionários (TRÄSEL, 2014, p.97).

Com os dados analisados, Meyer pode constatar que não havia relação entre condição econômica e participação nos distúrbios, os nativos da cidade eram mais propensos a participar de saques, agressões, homicídios e incêndios do que moradores que migraram do Sul. Os dados da pesquisa revelaram também que as razões principais para a revolta era a brutalidade da polícia, más condições de moradia e a escassez de empregos. No dia 20 de agosto de 1967, jornais do grupo *Knight* publicaram o relatório de Meyer (TRÄSEL, 2014, p.97).

Em 1973, a partir da experiência em Detroit e outras experiências, Meyer publicou o livro *Precision Journalism*. Para Meyer, o uso de técnicas de ciências sociais, como recursos metodológicos da sociologia, aliado à rapidez de cálculo dos computadores podia ajudar jornalistas a errar menos (TRÄSEL, 2014, p.98).

O jornalismo de precisão ganhou destaque novamente nas eleições presidenciais americanas de 2012, quando Nate Silver, o então jornalista do jornal *The New York Times*, conseguiu prever com exatidão que Barack Obama e senadores de 49 estados americanos seriam eleitos, surpreendendo especialistas em sondagens de opinião pública (CASTILHO, 2013).

Pena (2013, p.174) descreve a RAC como o uso de instrumentos tecnológicos para aproximar o jornalista da informação mais confiável, dando-lhe condições para interpretar a realidade, possibilitando com isso um jornalismo comprometido com os fatos. A RAC, aliada a métodos científicos de investigação social e psicossocial, foi chamada de “Jornalismo de Precisão” pelo professor Philip Meyer. Para Meyer, o jornalista não é apenas intérprete e transmissor, mas é administrador e analista de dados.

“O que se espera, então, de um jornalista inserido no contexto da RAC é um rigor científico capaz de fornecer uma organização da informação realmente alinhada com a necessidade do público, além da substituição da objetividade baseada em depoimentos contraditórios pela objetividade baseada em investigação contextualizada”. (PENA, 2013, p.175)

O jornalismo de precisão, para Träsel (2014), tem origem na necessidade de tornar o jornalismo independente de *press-releases* e de declarações de

personalidades públicas. O jornalismo de precisão surge para aperfeiçoar o jornalismo já praticado, possibilitando maior precisão ao analisar dados colhidos a partir de técnicas já usadas por sociólogos em pesquisas científicas. Träsel (2014) explica que Meyer vê no jornalismo de precisão uma forma de recuperar a objetividade nas redações, para Meyer muitos jornalistas passaram a usar abordagens interpretativas com o advento do Novo Jornalismo (*New Journalism*).

O jornalismo de precisão, aliado às técnicas de RAC, possibilitou ao jornalista investigar com maior rapidez. Com computadores, o trabalho de apuração ficou mais ágil, mais fácil de evitar erros humanos de tabulação de resultados de questionários, e, portanto, mais objetivo. Ressalta-se, contudo, que essa modalidade não é dependente de computadores, haja vista que as técnicas de pesquisa sociológica já eram usadas e analisadas em papel, portanto, é possível para um jornalista realizar experiências empíricas, baseadas em métodos de sociologia, sem o uso do computador, porém perde-se em agilidade.

2.2 Jornalismo de Dados: do Conceito à Prática

O jornalismo guiado por dados (JGD), ou simplesmente jornalismo de dados, é o uso de práticas computacionais aplicadas em conjunto de métodos científicos para a produção de notícia. Sua origem, segundo Träsel (2014, p.15), é derivada da Reportagem Assistida por Computador (RAC) e do jornalismo de precisão.

“As práticas de JGD envolvem técnicas de reportagem assistida por computador (RAC), visualização de dados, infografia, criação e manutenção de bases de dados e a política de acesso à informação e transparência pública de governos” (TRÄSEL, 2014, p.106).

O jornalismo de dados vai além de trabalhar com dados para fazer reportagens, é uma forma de contar histórias, descobrir e organizar fatos de maneira que seja possível retratar a realidade em todos os seus cenários: pelos números, com os dados; e complementando com comparações que podem ser feitas *in loco* pelo repórter.

Há várias expressões usadas para se referir ao jornalismo de dados, como “jornalismo guiado por dados” (TRÄSEL, 2014), “jornalismo computacional” (ANDERSON, 2012; LIMA JR., 2011), “jornalismo de dados” (GRAY, CHAMBERS e BOUNEGRU, 2012; PARASIE e DAGIRAL, 2013) e “jornalismo digital em base de dados” (BARBOSA, 2007, 2008, 2009, 2011; FIDALGO, 2004, 2007; MACHADO, 2006; RAMOS, 2011a, 2011b).

Bounegru (2012), lembra que o tema não era muito conhecido e pouco se falava sobre o ele. O tema, jornalismo de dados, começou a ganhar visibilidade a partir do vazamento de grande quantidade de dados pelo *WikiLeaks*.

“Um dos mais importantes passos para dar visibilidade ao termo foi a forma como grupos de mídia como The Guardian e The New York Times lidaram com a imensa quantidade de dados divulgados pelo WikiLeaks. Nesse período, o termo passou a ser usado de maneira mais ampla (ao lado de Reportagem com Auxílio do Computador, ou RAC) para descrever como jornalistas estavam usando dados para melhorar suas reportagens e para aprofundar investigações sobre um tema” (BOUNEGRU, 2012).

De acordo com a autora, a primeira definição de *Data Journalism* (Jornalismo de Dados) foi feita em 2006, pelo programador Adrian Holovaty, fundador do site *EveryBlock*, um serviço de informação que traz notícias ao usuário sobre o que está acontecendo no seu bairro ou até mesmo quarteirão. Em um ensaio publicado em seu *site* pessoal intitulado “Uma maneira fundamental na qual sites de jornais precisam mudar”³, Holovaty, propõe que jornalistas incorporem técnicas de gerenciamento de bases de dados no cotidiano das redações, para facilitar o reaproveitamento das informações coletadas nas reportagens e publiquem dados estruturados e compreensíveis por máquinas, ao lado do tradicional “borrão de texto dos jornais”:

“Muito do que os jornalistas locais coletam no cotidiano é informação estruturada: o tipo de informação que pode ser picada, de forma automatizada, por computadores. No entanto, a informação é destilada numa grande bolha de texto - um artigo de jornal - que não tem chance de ser reaproveitado” (HOLOVATY, 2006).⁴

3 “A fundamental way newspapers sites need to change”. In:
<http://www.holovaty.com/writing/fundamental-change/>

Träsel (2014, p.110), sugere como opção ao problema exposto por Holovaty, criar rotinas produtivas para o desenvolvimento de bases de dados estruturados sobre assuntos de interesse de uma redação:

“Desta forma, os jornalistas produziram dados estruturados independentes de formatos narrativos e suportes materiais. Esses dados livres, armazenados em bases, poderiam ser posteriormente agregados, cruzados, analisados e transformados em visualizações gráficas” (TRÄSEL, 2014, p.110).

Um exemplo do que Träsel (2014, p.110) propõe, como alternativa para melhor aproveitar informações coletadas numa redação seria o caso hipotético de um repórter que atua na editoria de crimes, o jornalista começaria a manter um banco de dados sobre vítimas de assédio sexual na cidade de São Paulo. A cada novo caso de assédio sexual registrado em uma delegacia, ou relatado em redes sociais, o repórter buscaria, junto à vítima, informações para registrar em uma planilha de dados como dia, horário, local, número de agressores, aparência dos agressores, tipo de abordagem etc. A partir de uma determinada quantidade de dados seria possível encontrar tendências entre os crimes, com isso o repórter poderia chegar a um perfil de local e horário onde mais ocorre o crime, tática para atacar do agressor, entre outros.

Para Holovaty (2006), qualquer fato de interesse público é passível de expressão em formatos de dados estruturados. Por conseguinte, Holovaty (2006) afirma que as “redações são lentas para aceitar novos pensamentos. Jornalistas não são o grupo mais *tech-savvy*⁵, eles não são o grupo mais inovador, e eles são (apenas um pouco) resistentes à mudança”⁶.

Em contrapartida, é uma prática antiga das redações o arquivamento de notícias para posterior consulta e reuso. Träsel (2014, p.110) lembra que pesquisa em documentos oficiais e registros, físicos ou digitais, é uma prática comum de

4 “So much of what local journalists collect day-to-day is structured information : the type of information that can be sliced-and-diced, in an automated fashion, by computers. Yet the information gets distilled into a big blob of text — a newspaper story — that has no chance of being repurposed”. In: <http://www.holovaty.com/writing/fundamental-change/>

5 Especialista tecnologia moderna, especialmente computadores. In: <http://dictionary.cambridge.org/pt/dicionario/ingles/tech-savvy>

6 “Newspapers are slow to accept this sort of thinking. Journalists aren't the most tech-savvy bunch, they're not the most innovative bunch, and they're (just a tad) resistant to change”. In: <http://www.holovaty.com/writing/fundamental-change/>

apuração. O uso de técnicas computacionais para armazenar e organizar arquivos em banco de dados é uma solução para que os jornais consigam arquivar uma grande quantidade de informação sem ocupar muito espaço (TRÄSEL, 2014, p.114).

Concomitantemente, o arquivamento de notícias gera um banco de dados que pode ser útil em futuras reportagens. A digitalização dessas notícias possibilita o arquivamento de uma grande quantidade de jornais e, por conseguinte, esse banco de dados é útil e de mais fácil manuseio utilizando técnicas de jornalismo de dados.

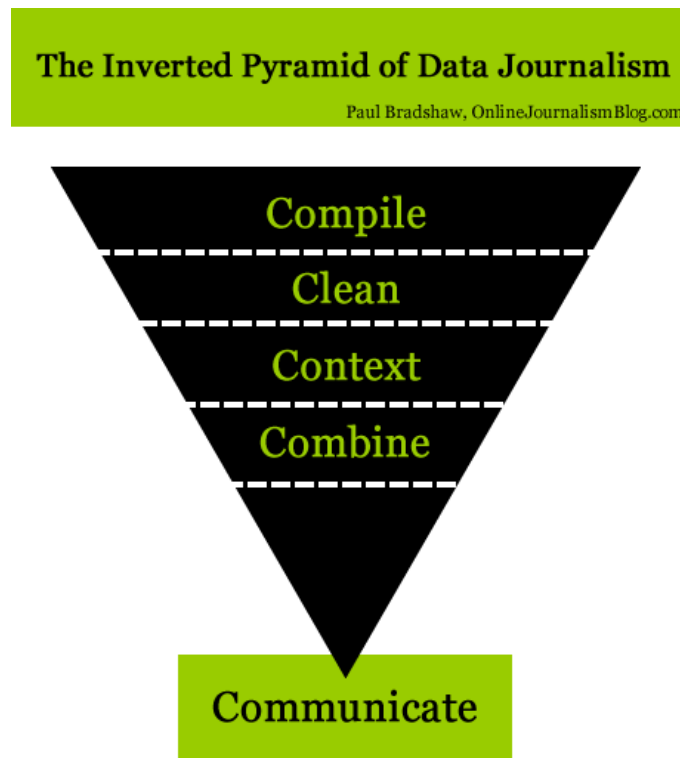
Para Barbosa e Torres (2013), o jornalismo de dados é aquele feito a partir de dados, gerados e disponibilizados por fontes públicas, privadas e por organizações jornalísticas do *mainstream*⁷, e podem estar em seu estado bruto, em planilhas de cálculo *Excel* (ou similares), disponíveis em banco de dados na internet.

A construção da notícia a partir de técnicas de jornalismo de dados se dá em cinco etapas, o jornalista Paul Bradshaw descreveu essas etapas no artigo “A pirâmide invertida do jornalismo de dados”⁸ publicado em seu *weblog* em 2011. Para Bradshaw (2011), o trabalho com jornalismo de dados pode começar de duas formas: ou o repórter tem uma pergunta que requer o uso de dados, ou tem uma base de dados que precisa ser interrogada. A pirâmide invertida de Bradshaw tem uma lógica simples de afunilamento da informação, começando no topo com a compilação de grande quantidade de informação, por conseguinte limpando, contextualizando e por fim comunicando os dados.

7 Organizações tradicionais dominantes. In: <http://dictionary.cambridge.org/pt/dicionario/ingles/mainstream>

8 “The inverted pyramid of data journalism”. In: <https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>

Figura 2 - Pirâmide Invertida de Jornalismo de Dados



Fonte: Paul Bradshaw (2011)

Compilar dados é reunir o conteúdo necessário para começar o trabalho com dados, ou seja, compilar é pesquisar, juntar informação para ser processada. Bradshaw (2011) explica que a compilação de dados é a busca da informação que será interrogada. Essa busca pode ocorrer de várias formas:

- “1. Fornecida diretamente a você por uma organização (quanto tempo até que vejamos 'divulgações de dados' ao lado de press releases?);
2. Encontrada com o uso de técnicas de busca avançada para mergulhar nas profundezas dos sites governamentais;
3. A partir da raspagem de bases de dados escondidas online usando ferramentas como *OutWit Hub*⁹ e *Scrapewriter*¹⁰;
4. Da conversão de documentos em algo que possa ser analisado usando ferramentas como *DocumentCloud*¹¹;
5. Obtendo informação de APIs¹²;

9 Aplicativo para explorar *deepweb* (profundidade da internet) e fazer extração de dados.

10 Plataforma de extração e análise de dados.

11 Aplicativo para trabalhar com diversos formatos de documentos.

12 Geolocalização.

6. Da coleta de dados feita por você mesmo a partir de observação, pesquisas, formulários online ou *crowdsourcing*^{13,14} (Bradshaw, 2011).

Das seis maneiras de encontrar conteúdo para trabalho com dados descrita por Bradshaw, obter dados fornecidos por organizações privadas ainda é o caminho desafiador para a maior parte dos jornalistas, mas há outras formas de se conseguir dados, como por exemplo por busca avançada no buscador *Google*. O *Google*, e outros buscadores, tem disponível, em seu mecanismo de busca, uma série de códigos que podem ser usados para aprimorar a busca por bases de dados na *web*, garantido mais assertividade. Abaixo, são apresentados, alguns termos que podem facilitar na busca por bases de dados:

- Pesquisar por formato de arquivo e incluir termos relativos ao conteúdo que deseja encontrar. É possível bucar planilhas inserindo o termo de busca "filetype:XLS filetype:CSV", ou dados geocodificados¹⁵ com "filetype:shp", ou bancos de dados com "filetype:MDB, filetype:SQL, filetype:DB" e, por fim, pode buscar apenas arquivos PDF usando o termo "filetype:pdf";
- Outra forma de pesquisar é pela URL¹⁶, inserindo "inurl:downloads filetype:xls", com esse termo o *Google* buscará todos os arquivos em Excel que tem a palavra "*downloads*¹⁷" em seu endereço. É possível limitar a busca a resultados a um único endereço, por exemplo "site:agency.gov".

13 Contribuição coletiva. Tradução Livre.

14 "1. supplied directly to you by an organisation (how long until we see 'data releases' alongside press releases?),

2. found through using advanced search techniques to plough into the depths of government websites;

3. compiled by scraping databases hidden behind online forms or pages of results using tools like OutWit Hub and Scaperwiki;

4. by converting documents into something that can be analysed, using tools like DocumentCloud;

5. by pulling information from APIs;

6. or by collecting the data yourself through observation, surveys, online forms or crowdsourcing". In: <https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>

15 Busca por coordenadas ou códigos.

16 Endereço disponível em uma rede, como por exemplo na internet, de um recurso, por exemplo um site.

17 Baixar um arquivo.

- Ao buscar o conteúdo diretamente pode ser que ele não esteja disponível, por isso é interessante tentar a busca por lugares em que podem estar disponíveis dados em massa. O termo "site:agency.gov Directory Listing" pode encontrar várias listas geradas automaticamente pelo servidor com fácil acesso aos dados brutos, já "site:agency.gov Database Download" buscará apenas listas criadas intencionalmente para serem encontradas.

A raspagem de bases de dados usando softwares¹⁸ é outra forma de extrair dados da internet. O *OutWit Hub* é um programa feito para explorar a *Deep Web*¹⁹ e extrair conteúdo de bases de dados. Outra ferramenta usada para extração de dados é o *Scrapewiki*, com ele é possível fazer raspagem de dados online de maneira que possa ser reutilizados em outros aplicativos. Há também o *DocumentCloud*, ferramenta que permite transformar uma planilha que esteja no formato PDF em formato de planilha (.txt ou .csv) aceito em diversos programas de planilha.

Há também outra forma de obter dados para uma reportagem é o próprio repórter construir a base de dados para sua reportagem. O jornalista pode fazer uma pesquisa por conta própria, usar formulários, algo como o que Meyer fez em Detroit em 1967, ou como Barcellos fez no início dos anos 1990, ao investigar os assassinatos da Polícia de São Paulo.

Obter dados via Lei de Acesso à Informação (LAI) é uma maneira de conseguir informação direto com a fonte - os aspectos da LAI serão tratados mais à frente.

A compilação de dados é a parte mais importante do trabalho com jornalismo de dados, é a partir dela que se obtém material para trabalho, é a parte mais revisitada do processo, sendo usada nas outras fases, caso haja necessidade de compilar mais dados.

A fase subsequente à compilação é a limpeza de dados. Após ter obtido o conteúdo para trabalho, é necessário extrair dele o que não será necessário para compor uma reportagem. Para Bradshaw há dois modos de limpar: primeiro, é preciso remover o erro humano e depois converter o dado em um formato consistente com outra base de dados utilizada.

¹⁸ Programa de computador, o mesmo que aplicativo.

¹⁹ "Internet profunda". Tradução Livre. É formada por inúmeros sites que não são acessíveis por links padrões em WWW e não estão disponíveis em buscadores, como o Google.

“Por exemplo, as bases de dados, muitas vezes, incluem alguns ou todos os erros a seguir: entradas duplicadas; entradas vazias; utilização de valores padrão para economizar tempo; formatação incorreta (por exemplo, palavras em vez de números); entradas corrompidas ou entradas com código HTML; vários nomes para a mesma coisa (por exemplo, B.B.C., BBC e British Broadcasting Corporation); e falta de dados”. (Bradshaw, 2011)²⁰

A ferramenta ideal para fazer a limpeza dos dados é um programa de planilha, como o Excel, Calc e Google Docs. É importante excluir todos os espaços vazios ou caracteres estranhos que possam atrapalhar na hora de fazer cálculos.

A contextualização dos dados é o momento em que o jornalista “fará perguntas”, ou seja, o jornalista vai “inquirir os dados” para encontrar uma história de interesse público. Contudo, é necessário ter cautela, dados são fontes e como qualquer fonte não são totalmente confiáveis. Bradshaw (2011) explica que os dados vêm com suas próprias histórias, preconceitos e objetivos. Por isso, é necessário questionar os dados:

“Quem os coletou, quando e por qual razão. Como foram coletados (a metodologia)? O que exatamente eles querem dizer com isso?”²¹ (Bradshaw, 2011).

Para contextualizar, é necessário entender também a linguagem usada para representar categorias, além de ter que compreender códigos, lugares, classificações e terminologias especializadas. É nesta fase que provavelmente o jornalista tende a compilar mais informação. Pois, com os dados extraídos podem surgir novas perguntas que necessitem de mais informações. Por exemplo, descobrir que uma cidade tem o número de crimes elevado em comparação com outras cidades pode tornar necessário um comparativo de populações, para se chegar a uma porcentagem por habitantes. Ao final, pode ser que a cidade com

20 “For example, datasets will often include some or all of the following: duplicate entries; empty entries; the use of default values to save time or where no information was held; incorrect formatting (eg words instead of numbers); corrupted entries or entries with HTML code; multiple names for the same thing (eg BBC and BBC and British Broadcasting Corporation); and missing data”. Idem.

21 “Who gathered it, when, and for what purpose? How was it gathered? (The methodology). What exactly do they mean by that?”. In: <https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>.

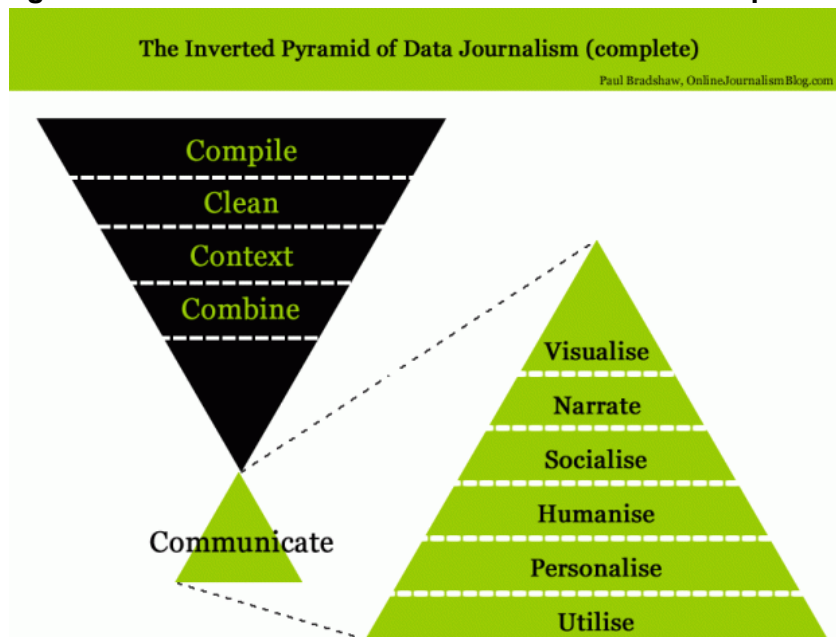
muitos números de crime seja proporcional às outras cidades, comparando-se os números de crimes por cem mil habitantes.

Após contextualizar os dados, ou seja, encontrar “uma história para contar”, pode ser que o jornalista de dados já tenha uma boa reportagem nas mãos, mas que pode ser aperfeiçoada a partir da combinação de dados. Ao combinar dados, o jornalista, estará cruzando informações, enriquecendo a reportagem e, por conseguinte, se aprofundando nos fatos. Com o aprofundamento feito a partir da combinação de dados de fontes distintas, o jornalista pode usar o tradicional “faro jornalístico” para buscar personagens *in loco* e, com isso, com personagens que possam humanizar a história, a reportagem fica mais atraente.

O último passo da pirâmide invertida de jornalismo de dados de Bradshaw é comunicar os dados. A comunicação dos dados é feita por visualização em gráficos, mapas, infográficos ou animação. Contudo, segundo Bradshaw (2011), é necessário considerar para visualização pontos como a narrativa, aplicativos de notícias, entre outros. Para explicar os pontos a se considerar na visualização da notícia de jornalismo de dados, Bradshaw (2011b) escreveu o artigo “6 maneiras de comunicar jornalismo de dados (a pirâmide invertida de jornalismo de dados parte 2)”²².

22 “6 ways of communicating data journalism (The inverted pyramid of data journalism part 2)”. In: <https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>

Figura 3 - Pirâmide Invertida de Jornalismo de Dados parte 2



Fonte: Paul Bradshaw (2011b)

Os esquemas dos seis tipos de visualização de notícias de jornalismo de dados elaborado por Bradshaw (2011b) tem por objetivo demonstrar as diferentes formas de se comunicar dados: visualizar, narrar, Comunicação social, humanizar, personalizado, utilizar.

A visualização (*Visualisation*), segundo Bradshaw (2011b), é a maneira mais rápida para comunicar os resultados do jornalismo de dados. Ferramentas como o Google Docs, Excel, Calc, entre outros, possibilitam a criação de gráficos para visualização com apenas alguns cliques. Já sites especializados, como o infogr.am, oferecem *templates*²³ de infográficos animados para visualização em outros sites, ou o *download*²⁴ de infográficos sem animação. Tais sites são de fácil utilização, sem grande necessidade de conhecimento em programação.

A publicação ou comunicação dos dados pela perspectiva da narrativa (*Narration*) ainda é muito eficaz. Falar sobre superfaturamento em uma obra com prejuízo de meio milhão de reais não é muito eficaz, afinal de contas, a maioria das pessoas não fazem ideia de quanto seja meio milhão de reais, mas se o repórter

²³ Modelos. Tradução Livre.

²⁴ Baixar arquivo.

explicar que com meio milhão de reais é possível comprar, em um exemplo hipotético, quatorze carros populares no valor de trinta e cinco mil reais cada, ou mil cento e onze cestas básicas no valor de quatrocentos e cinquenta reais cada, o leitor entenderá mais facilmente a dimensão do desvio de verba pública. Unir a narração dos fatos ao infográfico pode ser uma forma de tornar a notícia ainda mais compreensível.

O terceiro ponto da segunda pirâmide de Bradshaw (2011b) é a comunicação social (*Social communication*), que segundo o autor, seria o ato de o leitor participar da produção de notícias a partir da disponibilidade de planilhas usadas por jornalistas para fazer matérias. Bradshaw (2011b) cita como exemplo o portal de notícias britânico *The Guardian* que disponibiliza, em seu blog de dados, planilhas para que os leitores trabalhem os dados e façam matérias de acordo com suas visões. Os leitores enviam as planilhas de volta ao jornal que publica o que o leitor fez.

Os números são frios, transmitir ao receptor acréscimos que os concretizem pode deixar a reportagem desumanizada, portanto, o autor sugere como quarto ponto de visualização a humanização (*Humanise*) da notícia. Se em um jornal, principalmente televisivo, o âncora tratar da demora para atendimento hospitalar para enfermos que esperam para fazer uma cirurgia valendo-se apenas de dados, a notícia pode parecer sem muita importância. Mas se nessa mesma matéria for apresentada, além dos dados numéricos e estatísticos da espera, o depoimento de pessoas que estão há meses esperando para fazer uma cirurgia, a matéria pode fazer com que o telespectador compreenda melhor a importância da denúncia.

“Então, depois de ter sido enterrado na abstração de dados, precisamos lembrar que sair e gravar uma entrevista com uma pessoa cuja vida foi afetada, podem fazer uma grande diferença para o poder da nossa história”. (Bradshaw, 2011b)²⁵

A interatividade é, sem dúvida, uma das tendências fundamentais do jornalismo a partir do advento da internet. Personalizar (*Personalise*), quinta

25 “So after being buried in abstract data we need to remember that going out and recording an interview with a person whose life has been affected by that data can make a big difference to the power of our story”. In: <https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>

proposta do autor refere-se à possibilidade do leitor interagir com a notícia construindo seu próprio entendimento. No Brasil, um exemplo claro de interatividade, é o Basômetro, do Estadão Dados, que permite ao usuário analisar o comportamento dos parlamentares, escolhendo pelo partido, ou um parlamentar em especial, ou uma bancada, e assim verificar como o político tem votado, por exemplo, ou se um parlamentar da esquerda tem apresentado comportamento mais condizente com a direita, entre muitas outras possibilidades.

O último ponto é a utilização (*utilize*), que envolve o uso de ferramentas ou serviços de informação a partir de dados coletados.

O processo de produção de notícia utilizando técnicas de jornalismo de dados é simples, apesar da quantidade de etapas, é possível fazer uma reportagem sobre gastos públicos em menos de uma hora, para um jornalista experiente em lidar com dados. Porém, trabalhos mais aprofundados requerem maior investimento de tempo e capital intelectual. Com a crise que jornais tem enfrentado, que os leva a dispensar cada vez mais jornalistas, fazer grandes reportagem que consumam tempo pessoal, ainda é visto como algo pouco atraente para muitos veículos. Thiago Mali, coordenador de cursos da Abraji, em entrevista coletada para esta pesquisa, destaca as dificuldades enfrentadas por jornais brasileiros para implementar jornalismo de dados ou até mesmo bases de dados, como sugere Holovaty (2006, s. p.), nas redações:

“Nas redações, tem o problema de como teve muitos cortes de pessoal, e jornalismo de dados exige um certo tempo e dedicação para fazer, muitas vezes os chefes os editores não querem deixar uma pessoa dedicada muito tempo a analisar dados para conseguir de repente uma ou duas matérias. A partir dali, preferem usar essa mesma pessoa para escrever várias notas que podem render mais audiência no mesmo tempo. Existe uma demanda de público e jornalistas para tentar fazer mais isso e saber mais isso, mas existe um problema circunstancial. Hoje, o mercado não permite que isso avance mais rapidamente no Brasil”.²⁶

A partir das propostas de jornalismo de dados de Holovaty (2006, s. p.) e Bradshaw (2011a), observa-se uma fusão do jornalismo com as ciências da computação. Portanto, seria necessário que jornalistas assumissem um papel multidisciplinar trabalhando com dados, não apenas como jornalistas, mas como

26 Anexo A

programadores para preparar formatos de visualizações de dados, criação de aplicativos e mecanismos de recomendação de notícias, caso contrário, programadores assumiriam o papel central das redações. Elaborando ferramentas de visualização, enquanto jornalistas ficariam incumbidos de contextualizar a notícia.

2.3 Empreendedorismo e Jornalismo de Dados

Verifica-se, paulatinamente, que o jornalismo embasado na lógica dos dados tem ganhado força nas redações de grandes conglomerados e de iniciativas independentes de jornalismo. Com abertura de dados por órgãos governamentais, trabalhar com dados tornou-se uma regra para muitas redações e, com isso, o jornalismo de dados passou a ser também uma oportunidade para empreender.

O mercado está receptivo para modelos de negócios com foco na transformação de dados visíveis e compreensíveis por humanos em conteúdo exclusivo e de teor analítico. Empresas de tecnologia estão seguindo o movimento do “*Big Data*” e estão trabalhando com imensa quantidade de dados, usando softwares de código aberto, para extrair informação em segundos. No mercado jornalístico, o *The Guardian* e o *The New York Times* foram pioneiros em implantar o jornalismo de dados em suas redações.

A *Bloomberg* é um exemplo de sucesso ao investir em dados. A empresa opera cerca de 300 mil terminais e fornece dados financeiros aos seus usuários. Os ganhos da Bloomberg com esse negócio são de cerca de 6,3 bilhões de dólares por ano (Lorenz, 2012). Outras empresas resolveram investir em agências de notícias como a canadense Thomson Reuters que há duas décadas saiu do negócio de jornais para se tornar uma agência de notícias. A companhia cresceu com base em serviços de informação, dando uma visão mais aprofundada aos clientes de diversas áreas.

No Brasil, o jornal *Estado de São Paulo* foi o primeiro a aderir ao jornalismo de dados. Em 2012, o grupo Estadão criou o núcleo Estadão Dados que logo apresentou o seu primeiro trabalho, o Basômetro.

O primeiro projeto de jornalismo de dados no Brasil, o *J++*, foi criado em 2012 pelos jornalistas Thiago Mali, Marco Túlio Pires, Natália Mazotte, Juan Torres e o

designer e programador Thomaz Rezende. Em entrevista, Mali afirma ser difícil, no momento, obter retorno financeiro com jornalismo de dados:

“O mercado está muito ruim, o mercado jornalístico em geral, tem as redações demitindo muita gente, sobram poucas pessoas dentro das principais redações jornalísticas e tem algumas pessoas tentando caminhos alternativos, fazer de maneira mais independente jornalismo de dados, o que é muito interessante, muito louvável. Mas, até agora, não existe nenhum modelo que seja considerado super bem-sucedido, o mais conhecido que está fazendo isso é o *Volt Data Lab*, mas até onde eu sei é algo embrionário, eles estão testando para ver se funciona”.²⁷

O projeto *J++* de jornalismo de dados foi descontinuado no fim de 2015, a página na internet ainda está disponível.

O projeto *Volt Data Lab*, do jornalista Sérgio Spagnuolo é proveniente de um blog criado no fim de 2014, que logo se tornou uma agência de jornalismo independente focada em investigação a partir de dados. O Volt Data Lab oferece cursos de formação em jornalismo de dados e reportagens à mídia e a organizações não governamentais.²⁸

Spagnuolo começou desenvolvendo projetos em parceria com o projeto de jornalismo independente *Aos Fatos*, que faz checagem de notícias e declarações de políticos e autoridades de expressão nacional. O jornalista atuou, sem cobrar pelo serviço prestado, por seis meses para o *Aos Fatos*, tornou-se parceiro da plataforma.²⁹

Sem concorrentes no mercado nacional, a *Volt Data Lab* presta serviço para grandes empresas da mídia tradicional, como *Estadão Dados* e *Folha de São Paulo*.

A agência possui apenas quatro colaboradores: Heloisa Pizzo, administrativo e marketing; Juliana Elias, jornalista; Fábio Aleixo, consultor de desenvolvimento *web*; e Raphael Gadelha, designer e artista gráfico. Spagnuolo não possui funcionários com registro de trabalho pelo ministério do Trabalho. Os quatro colaboradores atuam em outras empresas e prestam serviço à agência *Volt Data Lab* de acordo com a demanda. Contudo, para o jornalista é difícil ter muitos funcionários, apesar da demanda de trabalho estar crescendo cada vez mais. Em

27 Anexo A

28 In: <http://www.voltdata.info/>

29 Spagnuolo concedeu entrevista para a pesquisa, mas por falha técnica a mesma não foi gravada

conversa informal o jornalista disse não sabe exatamente quando investiu até o momento, mas arrisca dizer que foi algo em torno de dez mil reais, além do tempo de dedicação. Diferente de outras *startups*³⁰ criadas por jornalistas, a *Volt Data Lab* não recebeu investimento de empresários, nem mesmo arrecadou dinheiro por *crowdfunding*³¹, como, por exemplo, a *ProPublica*, organização estadunidense de jornalismo investigativo, sem fins lucrativos, que sobrevive de doações.³²

Outra iniciativa empreendedora voltada para reportagens usando as técnicas de jornalismo de dados é o site *Fiquem Sabendo*. O site de notícias de jornalismo de dados é independente e trabalha com informações obtidas via Lei de Acesso à Informação e se propõe a revelar informações de interesse social que o poder público não revela³³. Seu fundador, Léo Arcoverde, é repórter há treze anos.

A agência *Volt Data Lab* e o portal *Fiquem Sabendo* são as primeiras empresas no Brasil dedicadas somente a jornalismo de dados, no entanto, o uso da modalidade já é comum na mídia digital, impressa e audiovisual. Há também muitas organizações sem fins lucrativos voltadas ao trabalho exclusivo com dados, como o *InfoAmazonia*, o portal *Transparência Brasil* e *Meu Município*, os quais analisaremos no próximo capítulo.

2.3.1 Descritivo do Ferramental Necessário para a Atuação com Jornalismo de Dados

Trabalhar com dados exige ter um leque de *softwares* capazes de possibilitar ao jornalista executar cada uma das cinco fases da pirâmide invertida de Bradshaw. Buscar ferramentas de código aberto para trabalhar é a melhor opção para iniciar um negócio sem ter que gastar com *softwares* pagos.

Código Aberto, ou *Open Source*, é um *software* que disponibiliza a redistribuição e alteração do seu código de fonte gratuitamente, sendo proibido a cobrança de *royalties* ou qualquer outro tipo de venda³⁴.

30 Empresa em início de atividade.

31 Contribuição coletiva.

32 Spagnuolo concedeu entrevista para a pesquisa, mas por falha técnica a mesma não foi gravada.

33 In: <<http://www.fiquemsabendo.com.br/quem-somos/>>. Acesso em 04 de nov. 2016

34 In: <<http://brasilescola.uol.com.br/informatica/codigo-aberto.htm>>. Acesso em 24 de out de 2016

A opção por *software* de código livre diminui as despesas de quem está iniciando um negócio. O *Windows 10 Pro*, um dos sistemas operacionais mais populares do mundo, é vendido no *site* da *Microsoft* por mais de oitocentos reais³⁵, em sua versão mais recente. Já o *Linux Ubuntu*, a distribuição de código aberto mais popular, é ofertado gratuitamente³⁶, sendo facultativa a doação de qualquer quantia para os desenvolvedores do programa.

Há uma preferência por jornalistas e programadores em usar *softwares* de código aberto para trabalhar. Rodrigo Burgarelli, repórter do *Estadão Dados*, por exemplo, diz trabalhar apenas com *Linux* e ferramentas de código aberto, como banco de dados e *LibreOffice*:

“Além de ser gratuito, elas são as melhores ferramentas, você pode instala no seu computador de casa ou do trabalho, a documentação na internet é muito ampla, então se você tem alguma dúvida, tem uma base de usuários gigante mexendo, então você consegue tirar sua dúvida muito rápido. Como sou jornalista e não fiz curso sobre isso (programação), é muito ter muita dúvida já respondida na internet”.³⁷

Burgarelli explica que os fóruns disponíveis na internet, onde usuários trocam informações entre si, é bom para jornalistas que não sabem programar resolverem, de forma rápida, dificuldades técnicas que possam aparecer.

O *LibreOffice* é uma ferramenta de escritório similar ao *Microsoft Office*, oferecem documentos de texto (*writer*), planilha (*Calc*), apresentação em slide (*Impress*), editoração eletrônica e construção de vetoriais (*Draw*), e editor de fórmula (*Math*). O *LibreOffice* está disponível para distribuições *Linux*, *Microsoft Windows* e *IOS*. Para o mesmo fim, também é possível usar o *Google Docs*.

Banco de dados é algo essencial para o jornalismo de dados. Com uma ferramenta de banco de dados é possível organizar e estruturar dados. O mais popular de código aberto é o *SQLite*, mas há muitos outros, como o *MongoDB*, *CouchBase*, *OrientDB* e *Cassandra*³⁸.

35 In <https://www.microsoftstore.com/store/msbr/pt_BR/pdp/Windows-10-Pro/productID.320406200>. Acesso em 24 de out de 2016

36 Disponível em: <<https://www.ubuntu.com/download/desktop>>. Acesso em 24 de out de 2016

37 Anexo B

38 Disponível em: <<https://www.propus.com.br/os-4-melhores-bancos-de-dados-de-codigo-aberto-que-voce-deve-considerar-para-seu-proximo-projeto/>>. Acesso em 24 de out de 2016.

Há ferramentas disponíveis gratuitamente, mesmo não sendo de código aberto, como o *Google Sheets*, disponível no *Google Docs*. Com ele, é possível extrair tabelas de *sites* e criar *hiperlinks* para que as informações contidas no *site* atualizem na planilha do *Docs*. Outra ferramenta para extrair dados é o IFTTT, com ele é possível compilar dados de redes sociais. A ferramenta conecta conta de e-mail com, por exemplo, a conta do *Twitter* do presidente da república, cada vez que o presidente da República. Sempre que o presidente publicar algo nesta rede social, o e-mail cadastrado receberá uma notificação. IFTTT é uma sigla em inglês que significa “*If this, than that*”, ou seja, “Se isso, então aquilo”.

Outra ferramenta de extração gratuita é a *Webscraper*, uma extensão disponível no *Google Chrome*. O *Webscraper* cria robôs de extração de dados de qualquer site. *Convextra*³⁹ também possibilita a raspagem de dados de qualquer site, seu método de raspagem é mais amigável, ideal para quem não tem muito conhecimento em programação. O *Convextra* oferece serviço gratuito e pago: o pago oferece serviços extras, como maior armazenagem, agendamento e maior quantidade de extração por mês. O *OutWit Hub* e *Scraperwiki* já citados no subcapítulo anterior, também são ferramentas gratuitas de compilação de dados.

Spagnuolo (2015) fez uma lista das melhores ferramentas com as melhores ferramentas para se trabalhar com dados. A primeira dica é para quem não sabe qual melhor gráfico usar para publicar dados, para isso ele indica o site *The Data Visualisation Catalogue*⁴⁰ para aprender a montar uma visualização.

Para visualização de dados, Spagnuolo (OBSERVATÓRIO DA IMPRENSA, 2015) recomenda como melhor ferramenta o *Plot.ly*⁴¹. A plataforma online é gratuita e permite complexos tipos de análises e cruzamento de dados.

*Carto*⁴² é uma plataforma de localização que possibilita criar visualizações de mapas. O site oferece planos gratuito e pago. Os planos pagos oferecem privacidade, mais espaço de armazenamento, sincronização de dados e suporte técnico.

39 In: <<http://convextra.com/prices.php>>. Acesso em 05 de nov. de 2016

40 <http://www.datavizcatalogue.com/>

41 In: <<https://plot.ly/>>. Acesso em 05 de nov. de 2016

42 In: <<https://carto.com/pricing/>>. Acesso em 05 de nov. de 2016

*Color Brewer*⁴³ é uma ferramenta gratuita para selecionar esquemas de cores para mapas.

*CometDocs*⁴⁴ é uma ferramenta gratuita que disponibiliza serviços como armazenamento, compartilhamento e conversão de dados. A funcionalidade de conversão de dados é a mais interessante para o jornalismo de dados, porque, possibilita converter dados em formato PDF para *Excel*, *Word*, imagem, *PowerPoint*, entre outros formatos. A plataforma pode ser usada sem cadastro, mas para ter acesso a todo serviço disponível é necessário fazer cadastro.

*Creatavist*⁴⁵ é uma plataforma multimídia para construir histórias compostas de texto, vídeo, imagens, mapas e muito mais.

Como se pode verificar, há grande diversidade de ferramentas de código aberto, ou gratuitas, disponíveis para quem deseja iniciar um negócio com jornalismo de dados.

2.4 Formação de Jornalistas

Todos os dias são produzidos cerca de 2,5 bilhões de *gigabytes* de dados, o suficiente para preencher 156 milhões de *iPads* (PORTAL EXAME, 2016). Ainda segundo o portal *Exame*, o *déficit* de profissionais aptos a lidar com dados chegará a trezentos e oitenta mil vagas no mundo em 2018.

Diante desse quadro, para suprir a demanda por jornalistas de dados, as universidades terão que capacitar alunos para lidarem com dados. Ou seja, técnicas de apuração, análise de dados, contextualização e visualização de dados trouxeram à tona a necessidade de formar jornalistas multidisciplinares, capazes de serem jornalistas e programadores como, segundo Träsel (2014, p. 98), Meyer havia proposto em 1973, em *Precision Journalism*.

Mali acredita que as universidades ainda estão muito atrasadas em relação ao ensino de jornalismo de dados. Para ele, “o primeiro passo é o jornalista ter familiaridade com softwares de planilha, para conseguir filtrar, classificar e separar

43 In: <<http://colorbrewer2.org/#type=sequential&scheme=BuGn&n=3>>. Acesso em 05 de nov. de 2016

44 In: <<http://www.cometdocs.com/>>. Acesso em 05 de nov. de 2016

45 Disponível em: <<https://atavist.com/cms/project>>. Acesso em 05 de nov. de 2016

informações identificando os fora da curva e dentro da curva”⁴⁶. Já Alberto Cairo, em entrevista a Ângela Pimenta (OBSERVATÓRIO DA IMPRENSA, PIMENTA, 2015), acredita que o jornalista não pode ficar preso ao perfil “escritor-humanista” e, sim, ser um “pesquisador-cientista”:

“O jornalista brasileiro, como muitos outros de tradição mediterrânea (não se esqueçam que sou espanhol) é, em geral, um escritor-humanista, não um pesquisador-cientista. Como ter os dois perfis é fundamental em qualquer redação, a mídia brasileira precisa hoje menos do primeiro e mais do segundo. Em algumas palestras no país, enquanto comentava exemplos de gráficos ou histórias que poderiam ser melhoradas, falei casualmente: “Aqui podem ver um caso claro de quando é melhor usar a mediana e não a média”, só para ficar chocado pelos olhares de confusão de uma parte da audiência. Se nós não sabemos algo tão básico como o que é uma mediana, o que dizer de desvio padrão, análises de regressão, valor-p, ou métodos bayesianos, tão em moda hoje graças ao sucesso de Nate Silver no The New York Times” (OBSERVATÓRIO DA IMPRENSA, PIMENTA, 2015).

O ensino de estatística básica ou de planilhas eletrônicas, conhecimento necessário em redações, na maioria das faculdades de jornalismo, não está presente. Ao consultar a matriz curricular de jornalismo de treze universidades brasileiras foi constatado que onze não têm disciplina de jornalismo de dados, doze não oferecem disciplina de estatística. Apenas a Escola Superior de Propaganda e Marketing (ESPM) e o Centro Universitário de Maringá (UNICESUMAR) possui disciplina de jornalismo de dados, sendo que, na última o nome da disciplina é jornalismo em base de dados. Apenas a ESPM tem estatística em sua matriz curricular.

Tonus, presidente do Fórum Nacional de Professores de Jornalismo (FNPJ), gestão 2012-2016, explica, em entrevista⁴⁷, que a construção da matriz do curso de jornalismo é feita a partir das Diretrizes Curriculares para o Curso de Jornalismo⁴⁸. As diretrizes estabelecem seis eixos para serem seguidos pelas universidades como conteúdos curriculares: eixo de fundamentação humanista, eixo de fundamentação específica, eixo de fundamentação textual, eixo de formação profissional, eixo de

46 Anexo A

47 Anexo C

48 In: <http://portal.mec.gov.br/dmdocuments/documento_final_cursos_jornalismo.pdf>. Acesso em 26 de out de 2016.

aplicação processual e eixo de prática laboratorial. Apenas no último eixo se fala em webjournal, mas não se descreve o que seria importante estudar para trabalhar no ambiente digital. Embora existam diretrizes a serem seguidas, não há uma grade fixa de disciplinas que devem ser ensinadas pelas universidades, portanto, é parcialmente facultativa a decisão de cada universidade sobre quais disciplinas ofertar.

Mali, avalia o ensino de jornalismo de dados nas universidades como deficientes, para ele o que se ensina está muito distante do que se vê no dia a dia das redações:

“É ridículo que as universidades ainda hoje na maior parte não tenham incorporado [jornalismo de dados] no seu currículo. Muitas incorporam e dão o nome de uma matéria, mas na prática o que se ensina lá é algo muito distante do que se vai usar no dia a dia, correndo atrás de apuração e notícias”.⁴⁹

Segundo Mali, a busca por cursos de jornalismo de dados já é significativa, para ele há “um interesse crescente, a gente abre e em dois dias esgotam as vagas. Existe uma demanda de muitos jornalistas querendo entender melhor como fazer isso”.⁵⁰ Para Mali, o jornalismo de dados já não é um diferencial para o jornalista concorrer a uma vaga de emprego, já se tornou algo básico, que todo jornalista deve ter.

3 DEMOCRACIA E CULTURA DE TRANSPARÊNCIA

Uma das principais fontes de informações do jornalismo é o poder público, por conseguinte, as bases de dados também das instituições governamentais têm grande relevância na tarefa de atendimento do interesse público. Ter acesso a tais informações, porém, nunca foi fácil. Apesar do artigo 5º, inciso XXXIII, da Constituição brasileira⁵¹ garantir o direito à informação, sempre houve pouco engajamento dos agentes públicos em praticar esse direito.

49 Anexo A

50 Idem

51 In: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm>. Acesso 27 de out. de 2016

“Todos têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado” (CONSTITUIÇÃO FEDERAL, 1988)

Nesse contexto, a cultura da transparência surge como uma forma de consolidar a participação da sociedade civil na democracia. A partir dos anos 2000, com a popularização da internet, a busca por informação aumenta e é criada a organização sem fins lucrativos *Transparência Brasil* com o objetivo de manter um banco de dados com informações sobre financiamento eleitoral, histórico de vida pública e processos sofridos por parlamentares. Três anos depois foi criado o Fórum de Direito de Acesso a Informações Públicas com o objetivo de pressionar o governo pela regulamentação do direito de acesso à informação pública (TRÄSEL, 2012). Em 2005, é criada a Associação Contas Abertas com objetivo de acompanhar o processo de execução orçamentária e financeira da União, por meio do monitoramento pelo Siafi (CAIRO, 2012). Em 2009, a Lei Complementar 131 (LC131) altera a redação da Lei Complementar 101 de Responsabilidade Fiscal (LRF), no que se refere à transparência da gestão pública. O novo texto determina disponibilização, em tempo real, de informações sobre a execução orçamentária e financeira da União, dos Estados, do Distrito Federal e dos Municípios⁵². A nova lei passou a assegurar transparência mediante:

“I – incentivo à participação popular e realização de audiências públicas, durante os processos de elaboração e discussão dos planos, lei de diretrizes orçamentárias e orçamentos; II – liberação ao pleno conhecimento e acompanhamento da sociedade, em tempo real, de informações pormenorizadas sobre a execução orçamentária e financeira, em meios eletrônicos de acesso público; III – adoção de sistema integrado de administração financeira e controle, que atenda a padrão mínimo de qualidade estabelecido pelo Poder Executivo da União e ao disposto no art. 48-A.”⁵³

52 In: <<http://www.leidatransparencia.cnm.org.br/>>. Acesso em 29 de out 2016.

53 Idem.

A partir da LC131, os poderes institucionais começam a discutir maior transparência na gestão pública e em 7 de dezembro de 2011, a Câmara Federal lança o projeto Dados Abertos, com o objetivo de disponibilizar dados legislativos em formato digital para a sociedade (BRAMATTI, p.220). Dado aberto é o dado que pode ser livremente utilizado, reutilizado e redistribuído por qualquer um⁵⁴.

“O projeto Dados Abertos tem como objetivo disponibilizar dados brutos em formato aberto para utilização em aplicações digitais. Isto significa que cidadãos e entidades da sociedade civil poderão acessar os dados públicos da Câmara e, com eles, desenvolver ferramentas inteligentes que permitam desde a percepção mais efetiva da atuação parlamentar e os gastos detalhados da Casa, até o resultado das votações do dia, de forma simples e automática pela Internet”.⁵⁵

A transparência pública é a obrigação imposta ao administrador público de prestar contas à população, divulgando o que faz, como faz, por que faz, o quanto gasta e o planejamento para o futuro⁵⁶. A transparência pública pode ocorrer de duas formas: ativa e passiva. A primeira ocorre quando o poder público divulga dados por iniciativa própria, sem que tenha sido solicitado⁵⁷. A transparência passiva ocorre quando os dados são disponibilizados mediante solicitação⁵⁸. A disponibilização de informações pública pode ser usada para fiscalização da gestão pública, avaliar as políticas públicas, o controle social e influencia no controle da corrupção.

Logarezzi (2016) afirma existir mais de noventa países que possuem uma legislação que garante o acesso à informação, sendo 23 na América Latina. Segundo a autora, há um crescente reconhecimento do direito à informação como um direito fundamental, com respaldo de importantes órgãos internacionais como a Organização das Nações Unidas (ONU), a Organização dos Estados Americanos (OEA), o Conselho Europeu (CoE) e a União Africana (UA) (LOGAREZZI, 2016, p.5).

54 In: <https://commons.wikimedia.org/wiki/File:Manual_Dados_Abertos.pdf>. Acesso em 29 de out. 2016.

55 In: <<http://www2.camara.leg.br/transparencia/dados-abertos/perguntas-e-respostas>>. Acesso em 29 de out. 2016.

56 In: <<http://gestaolai.sp.gov.br/glossario-transparencia>>. Acesso em 28 de out. de 2016.

57 Idem

58 In: <<http://gestaolai.sp.gov.br/glossario-transparencia>>. Acesso em 28 de out. de 2016.

3.1 Big Data

A era da informação ou era digital, segundo Pena (2016b), é o termo usado para descrever avanços tecnológicos advindos da Terceira Revolução Industrial e que resultaram no meio de comunicação instrumentalizado pela informática e internet.

O termo *Big Data* surgiu para descrever uma imensa quantidade de dados sendo produzida diariamente na internet (JUNIOR, 2012, p.57). Chede (IBM, 2012), explica que *Big Data* é volume, variedade, velocidade, veracidade e valor. Volume é a grande quantidade de dados gerados por dia, segundo o autor, são gerados *pentabytes* de dados por dia e estima-se que o volume dobre a cada dezoito meses. Variedade, porque os dados são gerados de fontes distintas em vários formatos, de dados estruturados, dados numéricos em bancos de dados, documentos de textos não estruturados, e-mails, vídeos, áudios, mídias sociais etc. Velocidade, porque os dados fluem em uma velocidade independente e devem ser tratados em tempo hábil, como um controle automático de tráfego na rua. Veracidade, porque é necessário ter certeza que os dados fazem sentido e são autênticos. Valor, porque é importante que projetos de *Big Data* tenha retorno positivo.

A era do *Big Data* se concretizou como ferramenta eficaz para a consolidação da democracia na era digital com o primeiro grande vazamento de dados.

“O WikiLeaks ficou famoso por colocar à disposição de sete organizações noticiosas informações sobre a guerra do Iraque, liberando o acesso a cerca de 400 mil documentos militares”. (JUNIOR, 2012, p.59).

Além dos vazamentos de dados com quantidades crescentes de informação, há também outro fator a se considerar na era do Big Data: a cultura de transparência. Para Junior (2012, p.62), a divulgação de informações se dá por dois modos, o primeiro é conhecido como “*Open Data*”⁵⁹, pela liberação de dados por instituições públicas, de forma passiva ou ativa. O segundo é feita de forma

59 “Dados Abertos”. Tradução Livre.

colaborativa por instituições, ou pessoas, que criam bancos de dados na internet para que todos possam compartilhar informação. Outro conceito é o *Linked Data*⁶⁰, que tem o objetivo de cruzar dados online.

“Open Data possui o conceito de propriedade pública dos dados, rejeitando a ideia econômica de valor imposta pela escassez da informação oriunda dos processos iniciados desde a Revolução Industrial. O Linked Data é um dos conceitos mais importantes do mundo dos dados digitais. É o Linked Data que permite o uso da web para relacionar dados que não estavam previamente relacionados”. (JUNIOR, 2012, p.62)

Chede (IBM, 2012), avalia que as tecnologias que sustentam o Big Data podem ser analisadas de duas formas: as envolvidas com *Analytics*, *Hadoop* e *MapReduce*⁶¹, modelos de programação que permite o processamento massivo de dados; e as tecnologias de infraestrutura que armazenam os dados. Chede (2012, IBM) explica que “*Big Data* é a simples constatação prática que o imenso volume de dados gerados a cada dia excede a capacidade das tecnologias atuais de os tratarem adequadamente”⁶².

3.2 Vazamento de Dados

Vazamento de dados ocorre quando informações que não deveriam ser públicas são disponibilizadas na *web* por um invasor, ou agente interno que teve acesso não autorizado aos dados (G1, ROHR, 2016). É possível constatar dois tipos de vazamentos de dados na *web*: o vazamento feito por cibercriminosos, conhecidos também como *crackers*⁶³, que visam lucrar com informações roubadas; e o segundo tipo de vazamento é feito com o intuito tornar público informações de relevância

60 “Dados interligados”. Tradução Livre.

61 In: <<http://www.devmedia.com.br/hadoop-mapreduce-introducao-a-big-data/30034>>. Acesso em 28 de out. 2016

62 Você Realmente sabe o que é Big Data? In: <https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/voce_realmente_sabe_o_que_e_big_data?lang=en>.

63 “Pirata virtual”. Indivíduo aficionado por informática que utilizam seu conhecimento e habilidade para cometer crimes. In: <<http://www.tecmundo.com.br/o-que-e/744-o-que-e-cracker-.htm>>. Acesso em 30 de out. 2016

social, neste caso o *hacker*⁶⁴ tem interesse ideológico e não o financeiro, o vazamento pode acontecer por um agente interno ou por invasão de sistema.

Uma das primeiras organizações dedicadas à publicação de dados secretos foi a *Cryptome.org*. Fundada em 1996⁶⁵, por John Young, um ativista de liberdade de expressão, *Cryptome* virou manchete por ter publicado em 1999 e 2005 com nomes dos agentes da MI6, o serviço secreto britânico para assuntos estrangeiros. “*Cryptome.org* publica documentos de pessoas que queriam revelar segredos sem correr o risco de serem expostas como delatores” explica Domscheit-Berg (2011, p.5).

Segundo Domscheit-Berg (2011), Julian Assange baseou-se no *Cryptome.org* para fundar o *WikiLeaks*, em 2006⁶⁶, na Suécia. O *WikiLeaks* é uma organização sem fins lucrativos, descrita em seu website⁶⁷ como uma organização multinacional de biblioteca associada, especializada em análise e publicação de grandes conjuntos de dados de materiais censurados ou que tenha qualquer outra restrição que envolva guerra, espionagem e corrupção, já tendo publicado mais de dez milhões de documentos.

Em novembro de 2007 o *WikiLeaks* faz o seu primeiro grande vazamento de dados da história, mais de 1,2 milhão de documentos secretos foram disponibilizados na internet para consulta gratuitamente (ROGERS, 2012). Os documentos vazados foram os manuais de procedimento da Baía de Guantánamo e revelavam que, nos centros de detenção cubanos, os Estados Unidos violavam os direitos humanos e as convenções de Genebra (DOMSCHEIT-BERG, 2011, p.5). Em 2008, o *WikiLeaks* publicou novos documentos vazados por agente interno, nos quais Domscheit-Berg (2011) descreve o trabalho feito antes de disponibilizar os arquivos no *site*:

“Qualquer um teria feito o upload de uma infinidade de números e cálculos, organogramas, fluxos de trabalho e contratos em nossa caixa postal digital. Para que tudo isso? Julian e eu precisávamos de alguns dias até verificarmos o material. Em centenas de páginas, havia a representação da troca de correspondência interna, memorandos e cálculos da instituição bancária Julius Bär – um dos

64 “Modificar”. Indivíduo aficionado por informática que utiliza seu conhecimento para aprimorar softwares. In: <<https://www.significados.com.br/hacker/>>. Acesso em 30 de out. de 2016

65 In: <<https://cryptome.org/>>. Acesso em 01 de nov. de 2016

66 In: <<https://wikileaks.org/What-is-Wikileaks.html>>. Acesso em 01 de nov. de 2016

67 In: <<https://wikileaks.org/What-is-Wikileaks.html>>. Acesso em 01 de nov. de 2016

maiores bancos privados da Suíça” (DOMSCHEIT-BERG, 2011, p.15).

Domscheit-Berg (2011) explica que a partir dos documentos analisados foi possível entender como patrimônios milionários eram escondidos antes da investigação de fraude tributária.

Os vazamentos de dados do *WikiLeaks* eram em sua maioria de “funcionários que queriam divulgar as injustiças em suas próprias empresas, mas eram impedidos por contratos de adesão internos e cláusulas de confidencialidade” descreve Domscheit-Berg (2011, p.18).

Outros vazamentos surgiram no *WikiLeaks* a partir de funcionários insatisfeitos, mas o site recebeu vazamentos de ativistas digitais, o principal deles é o grupo *Anonymous*. Os ciberativistas buscaram o *WikiLeaks* para denunciar a seita Cientologia, uma seita religiosa poderosa nos Estados Unidos, que tem entre os membros estrelas de Hollywood (Domscheit-Berg, 2011).

Desde 2012, Julian Assange, fundador e redator responsável pelo *WikiLeaks*, está refugiado na embaixada do Equador, em Londres, após o Equador lhe conceder asilo político. Assange é acusado, na Suécia, de abuso sexual, mas nunca apresentaram provas, ou fizeram uma denúncia formal contra ele. Em entrevista a Amy Goodman (2015, d. n.), Assange sustenta que há uma conspiração em que a Suécia estaria atuando como substituta dos Estados Unidos e que os governo americano deseja vê-lo preso para impedir o trabalho do *WikiLeaks*. O vice-presidente dos Estados Unidos, Joe Biden, chegou a classificar Assange como “terrorista de alta tecnologia” e representantes eleitos e comentaristas políticos de ambos os partidos afirmam que Assange deveria ser assassinado⁶⁸. Mesmo sob pressão, o *WikiLeaks* continua a publicar documentos secretos, entre os seus noventa parceiros, estão os brasileiros *Carta Capital*, *Agência Pública* e *O Globo*.

Outros vazamentos de dados têm surgido ao longo dos anos, com cada vez mais informações para ser tratadas por jornalistas. O mais recente é o caso *Panamá Papers*, revelado em abril de 2016, que tiveram vazados 2,6 *terabytes* ou 11,5 milhões de documentos do escritório de advocacia Mossak Fonseca. Jornalistas de

68 In: <<http://www.cartamaior.com.br/?/Editoria/Internacional/O-castigo-de-Julian-Assange-Condenado-sem-ter-sido-acusado/6/33638>>. Acesso em 01 de nov. de 2016

vários países foram convocados pelo Consórcio Internacional de Jornalismo Investigativo (ICIJ, na sigla em inglês) para participar da organização dos dados. Os documentos têm nomes de pessoas que criaram *offshores*⁶⁹ em paraísos fiscais para fugir de impostos ou esconder dinheiro obtido de forma ilícita (PORTAL O ESTADO DE SÃO PAULO, 2016).

3.3 Lei de Acesso à Informação (LAI)

A Lei de Acesso à Informação (LAI) nº12.527/11, foi sancionada pela presidente Dilma Rousseff em novembro de 2011 e entrou em vigor seis meses depois, maio de 2012. Apesar de já existirem outras leis que aproximam o Estado da sociedade, como a Lei Complementar 131, a LAI estabeleceu a obrigatória prestação de contas pelos três poderes: Judiciário, Legislativo e Executivo. A lei é válida para todas as esferas de governo: municipal, estadual, distrital e federal, incluindo empresas públicas, sociedade de economia mista e outros entes controlados direta ou indiretamente pela União. A LAI abrange também entidade privada sem fins lucrativos que receba recursos públicos⁷⁰. Assim, diz o artigo 6º:

“I - gestão transparente da informação, propiciando amplo acesso a ela e sua divulgação; II - proteção da informação, garantindo-se sua disponibilidade, autenticidade e integridade; e III - proteção da informação sigilosa e da informação pessoal, observada a sua disponibilidade, autenticidade, integridade e eventual restrição de acesso.”⁷¹

Com isso, a LAI estabelece que a publicidade e a transparência de informações é um dever e o sigilo é a exceção. Ressalta-se, porém, que, como destaca Abramo (2012), antes da LAI, a noção jurídica do sigilo só existia para informações sob poder da administração pública federal. Após a LAI, qualquer

69 Contas bancárias em paraísos fiscais.

70 Disponível em: <<http://www.justica.gov.br/Acesso>>. Acesso em 29 de out. de 2016.

71 Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm>. Acesso em 29 de out. de 2016

estado, município, Tribunal de Contas, ente legislativo passou a gozar da prerrogativa de definir, arbitrariamente, o que colocar em sigilo e por quanto tempo (ABRAMO, 2012). O que não compromete, a nosso ver, o importante avanço da transparência pública possibilitada pela LAI.

As informações públicas classificadas com restrição de acesso são classificadas em três instâncias de prazo de sigilo: Ultrassecreta, com prazo de sigilo de vinte e cinco anos, pode ser renovada uma única vez; Secreta, quinze anos de sigilo; e Reservada, com cinco anos de sigilo⁷². Abramo (2012) vê a possibilidade de sigilo como um convite à opacidade:

“Esse gênero de oportunismo da opacidade está sendo praticado em todos os cantos do país. Como a nova legislação define que cada poder, em cada esfera, define seu próprio mecanismo de recurso contra negativas de prestação de informação, o que acaba por ocorrer é que o mesmo indivíduo que definiu que determinado tipo de dado deve permanecer secreto é aquele que dá a palavra final a qualquer recurso”. (ABRAMO, 2012).

Abramo (2012) avalia que uma regulamentação não cria demanda por dados, mas se houver uma busca cada vez maior por informação, o objetivo de transparência nas três esferas de poder será atingido. Porém, segundo o autor, a busca por informação é feita por algum interesse, “em qualquer país, os demandantes por informação do Estado são, pela ordem: o setor privado; a imprensa; organizações não governamentais; acadêmicos; cidadãos”⁷³.

Os órgãos sujeitos à LAI têm obrigação de criar e manter um serviço de informações ao cidadão em local com condições apropriadas para atender e orientar o público. É dever também dos órgãos protocolar documentos e requerimentos de acesso à informação, além de informar sobre a tramitação (LOGAREZZI, 2016, p.7).

Em casos de transparência passiva, a LAI estabelece que o órgão responsável tem o prazo de vinte dias para responder a solicitação, sendo prorrogável por mais dez dias, mediante justificativa. Não é obrigatório ao solicitante se identificar ou explicar como os dados serão utilizados (LOGAREZZI, 2016, p.7).

72 Disponível em: <<http://www.justica.gov.br/Acesso>>. Acesso em 29 de out. de 2016.

73 In: http://datajournalismhandbook.org/pt/getting_data_2.html

Logarezzi (2016), explica que é dever da Controladoria Geral da União (CGU) organizar uma campanha nacional de fomento à cultura de transparência.

Segundo Logarezzi (2016, p.), a punição mínima ao desrespeito à LAI para agentes públicos e militares é a suspensão. Sendo que em casos mais graves, o infrator pode responder por improbidade administrativa.

3.4 Projetos Jornalísticos Baseados na Cultura de Transparência na Era Digital

Com a LAI em vigor, a mídia jornalística brasileira iniciou projetos de jornalismo de dados. O Jornal *O Estado de São Paulo* criou o *Basômetro*, primeiro núcleo de dados das redações brasileiras. A plataforma foi lançada com a LAI em vigor, em 2012, mas o projeto é baseado em dados que já eram disponibilizados antes da lei. A proposta do *Basômetro* é medir o grau de apoio dos deputados e senadores ou o apoio partidário ao governo federal no Congresso Nacional. Ele permite ao usuário tirar suas próprias conclusões a partir dos votos de deputados e senadores aprovando ou rejeitando as proposições do Executivo, permite também analisar se um político é mais conservador ou liberal ao votar (TOLEDO, 2012).

O *Basômetro* é uma ferramenta que permite ao usuário analisar como quiser o comportamento dos parlamentares. Sua principal fonte de informação são as bases de dados dos sites da Câmara e do Senado Federal, que estão disponíveis para consulta de qualquer cidadão (TOLEDO, 2012).

Bramatti (2014, p.219), explica que antes da existência de um banco de dados digital, nos primeiros anos do governo Fernando Henrique Cardoso, da Câmara dos Deputados e do Senado Federal, para se saber o voto de cada deputado sobre um determinado projeto, era necessário ao interessado comparecer a uma sala abaixo do Plenário da Câmara, entrar em uma fila e solicitar ao atendente um relatório emitido pelo painel eletrônico. Bramatti (2012, p.219) descreve que o processo de análise do comportamento dos parlamentares e das bancadas nas votações era feita de forma manual com lápis ou caneta.

Bramatti (2012, p.220) comenta que, em 1997, Câmara anunciou uma reforma para agilizar o sistema de votação, com os objetivos de evitar a ação de

políticos que votavam no lugar dos colegas e agilizar a publicação dos resultados das votações na internet.

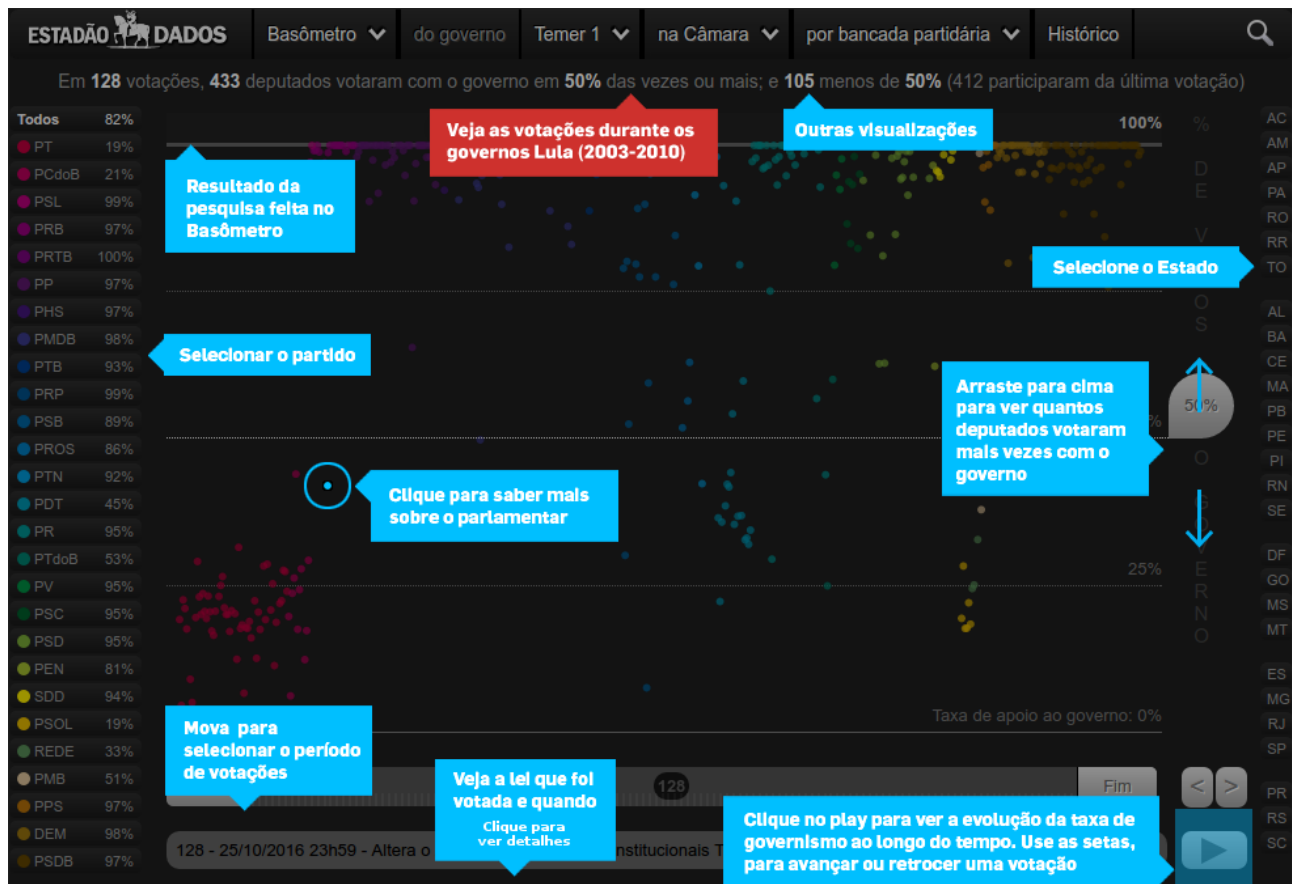
“A publicação de dados na rede facilita a vida de jornalistas, pesquisadores e analistas políticos – mas não muito. Os resultados apareciam em arquivos de texto (TXT) ou de banco de dados (DBF), algo com que muitos leigos não a menor familiaridade. Como a composição dos presentes variava a cada votação, nenhuma lista era igual à outra, o que, mais uma vez, praticamente inviabilizava a análise de séries históricas”. (BRAMATTI, 2012, p.220)

Segundo Bramatti (2012), o cenário só melhorou uma década depois, quando a Câmara dos Deputados lançou o projeto Dados Abertos, em 2011, disponibilizando dados em formato digital e em formatos mais acessíveis, possibilitando a criação de ferramentas como o *Basômetro*.

O *Basômetro*, como conta Bramatti (2012), foi inspirado no projeto *Caramaws*, criado logo após a Câmara iniciar o seu projeto de transparência com o Dados Abertos. Atualmente, intitulado *Radar Parlamentar*, o site ilustra as semelhanças entre partidos políticos com base na análise matemática dos dados das votações no Legislativo.

Em outubro de 2015, a ferramenta atingiu setecentos mil registros de votos de parlamentares, em mil e quinhentos e quarenta e três votações da Câmara e do Senado Federal, nos últimos treze anos e dez meses (PORTAL ESTADÃO DADOS, 2012).

Figura 4 - Página inicial do Basômetro



Fonte: Site Estadão Dados

Como se verifica, cada parlamentar é representado com uma bolinha na cor do seu partido. Quanto mais próxima a bolinha estiver do governo, no alto, maior é a taxa de governismo. O usuário pode selecionar o tipo de análise que deseja fazer: do comportamento dos parlamentares durante o governo de Dilma Rousseff, ou Luiz Inácio Lula da Silva (apenas Câmara); por bancadas partidárias ou votações específicas; por partido, ou conjunto de legendas. O *Basômetro* é atualizado sempre que ocorre votação na Câmara ou Senado.

Outro projeto que fiscaliza o poder público, a partir de bases de dados, é a instituição sem fins lucrativos *Transparência Brasil*. Fundada em abril de 2000, a instituição é comprometida com o combate à corrupção. Com dados recolhidos dos três poderes, nas três esferas, *Transparência Brasil* disponibiliza três ferramentas

ativas de monitoramento de instituições públicas à sociedade: *Àsclaras*, *Excelências* e *Meritíssimos*.

Figura 5 - Projetos de Jornalismo de dados da organização Transparência



Fonte: Site Transparência Brasil

Àsclaras trata-se de um banco de dados com informações e análises sobre o financiamento eleitoral em todas as campanhas municipais, estaduais e nacionais desde 2002⁷⁴. Com a plataforma, é possível fazer análises sobre o financiamento de campanha por meio de doações de empresas.

A ferramenta *Excelências* contém informações sobre todos os parlamentares em exercício na Câmara e no Senado Federal. Com dados recolhidos dos sites da Câmara e do Senado, dos Tribunais de Justiça, dos Tribunais de Contas, por cadastros mantidos por ministérios e outras fontes, é possível buscar informações sobre qualquer parlamentar. O projeto disponibiliza espaço para que os políticos apresentem argumentos referentes às informações divulgadas no projeto⁷⁵.

⁷⁴ Disponível em: <<http://www.transparencia.org.br/projetos>>. Acesso em 03 de nov. de 2016

⁷⁵ Disponível em: <<http://www.excelencias.org.br/>>. Acesso em 03 de nov. de 2016

O projeto *Meritíssimos* analisa o desempenho dos ministros do Supremo Tribunal Federal. A ferramenta ainda é um piloto, ou seja, está em fase de testes, mas já é possível analisar as expectativas de tempo de resolução de processos⁷⁶.

Outro projeto independente que surgiu a partir da cultura de transparência é *Meu Município*. O projeto disponibiliza dados dos municípios brasileiros. É possível entender como as prefeituras arrecadam e gastam o dinheiro do cidadão. No portal, está disponível contas de receitas e despesas e indicadores gerenciais de cada cidade. O site foi desenvolvido com o objetivo de comparar despesas, receitas e indicadores gerenciais entre os municípios que tenham o perfil parecidos. Os municípios foram agrupados de acordo com a proximidade das variáveis que influenciam diretamente no resultado estudado⁷⁷. Os dados apresentados pelo site foram extraídos do Ministério da Fazenda e do IBGE.

*Gastos Abertos*⁷⁸ é outro projeto que surgiu com a abertura de dados públicos. O projeto acompanha todas as transações financeiras de empresas e governos. O objetivo é apresentar os dados de maneira interessante e de fácil compreensão. Atualmente o projeto trabalha com gastos da prefeitura de São Paulo, do governo do estado de São Paulo e do governo federal. *Gastos Abertos* é um projeto da *Open Knowledge Brasil* (OKBr).⁷⁹

A *Open Knowledge Internacional* foi fundada em maio de 2004, em Cambridge, no Reino Unido, sob a crença que o conhecimento deve ser livre, ou seja, a informação deve ser acessível a todos⁸⁰. A *Open Knowledge Brasil*, também chamada de Rede pelo Conhecimento Livre, faz parte da *Open Knowledge Internacional*, sendo uma Organização da Sociedade Civil (OSC) sem fins lucrativos, que iniciou os trabalhos no Brasil em 2013. Apartidária, desenvolve ferramentas de visualização de dados com o objetivo de promover o conhecimento livre para tornar a relação entre governo e sociedade mais transparente⁸¹.

O *Artigo 19*⁸² é uma organização não-governamental de direitos humanos e sem fins lucrativos que defende e promove à liberdade de expressão e de acesso à

76 In: <<http://www.meritissimos.org.br/stf/index.php>>. Acesso em 03 de novembro de 2016

77 In: <<https://meumunicipio.org.br/sobre-o-portal>>. Acesso em 03 de nov. de 2016

78 In: <http://gastosabertos.org/> Acesso em 03 de nov. de 2016

79 In: <<http://br.okfn.org/projetos/>>. Acesso em 03 de nov. de 2016

80 In: <<http://blog.okfn.org/2004/05/24/open-knowledge-foundation-launched/>>. Acesso em 03 de nov. 2016

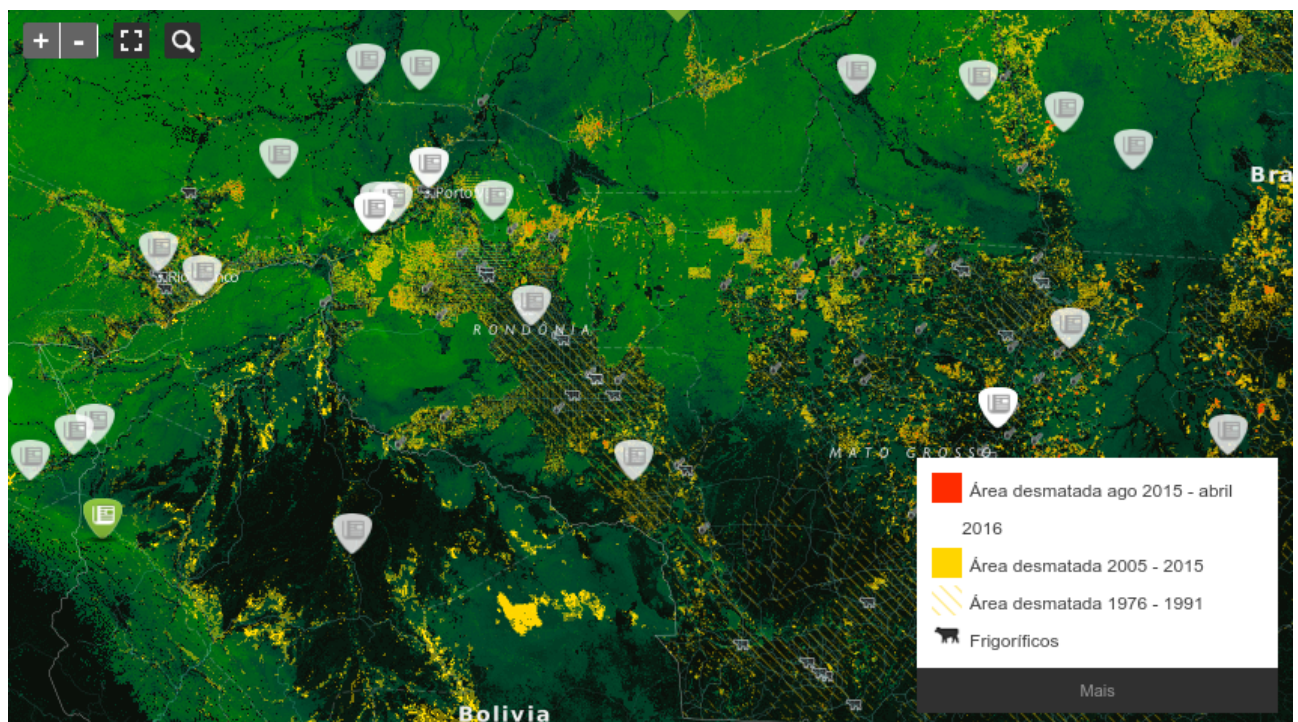
81 In: <<http://br.okfn.org/sobre/>>. Acesso em 03 de nov. de 2016

82 In: artigo19.org. Acesso em 03 de nov. de 2016

informação em todo o mundo. Fundada em 1987, em Londres, seu nome tem origem no 19º artigo da Declaração Universal dos Direitos Humanos da ONU. A organização tem escritórios em nove países e está no Brasil desde 2007⁸³.

InfoAmazônia é um diálogo de jornalismo com dados Geográficos que permite ao usuário conhecer os principais focos de incêndios florestais em parques e reservas do Brasil. Utilizando dados próprios e ferramentas do *Google*, como *Google Earth*, marcar os focos de incêndio deixando a experiência de navegação do público mais fácil pelo site (FALEIROS, 2012).

Figura 6 - Mapa do desmatamento da floresta Amazônica



Fonte: Site InfoAmazônia

Faleiros (2012) explica que uma grande quantidade de dados gratuitos disponíveis em formato aberto no site da NASA ou no site do Instituto de Pesquisas Espaciais (INPE), como séries históricas sobre fogo e desmatamento, reforçaram a inspiração para a criação do *InfoAmazônia*. Ao jornalismo feito no *InfoAmazônia*, Faleiros (2012) explica que deu o nome de geojornalismo por usar geolocalização para registrar incêndios e desmatamento.

83 In: <<http://artigo19.org/a-organizacao/>>. Acesso em 03 de nov. de 2016

“Nossa ideia, logo de início, era usar o dado de satélite como contexto e guia para reportagens que deveriam ser feitas em campo pelos jornalistas. Assim surgiu um nome para a prática, o geojornalismo - uma espécie de galho dentro da frondosa árvore do jornalismo de dados.” (FALEIROS, 2012).

O projeto foi lançado em 2012 em parceria com O Eco e *Internews*, uma organização dos Estados Unidos que fomenta o jornalismo em países subdesenvolvidos. O *InfoAmazônia* é financiado pelo Centro Internacional de Jornalistas (ICFJ) por meio das bolsas *Knight*.

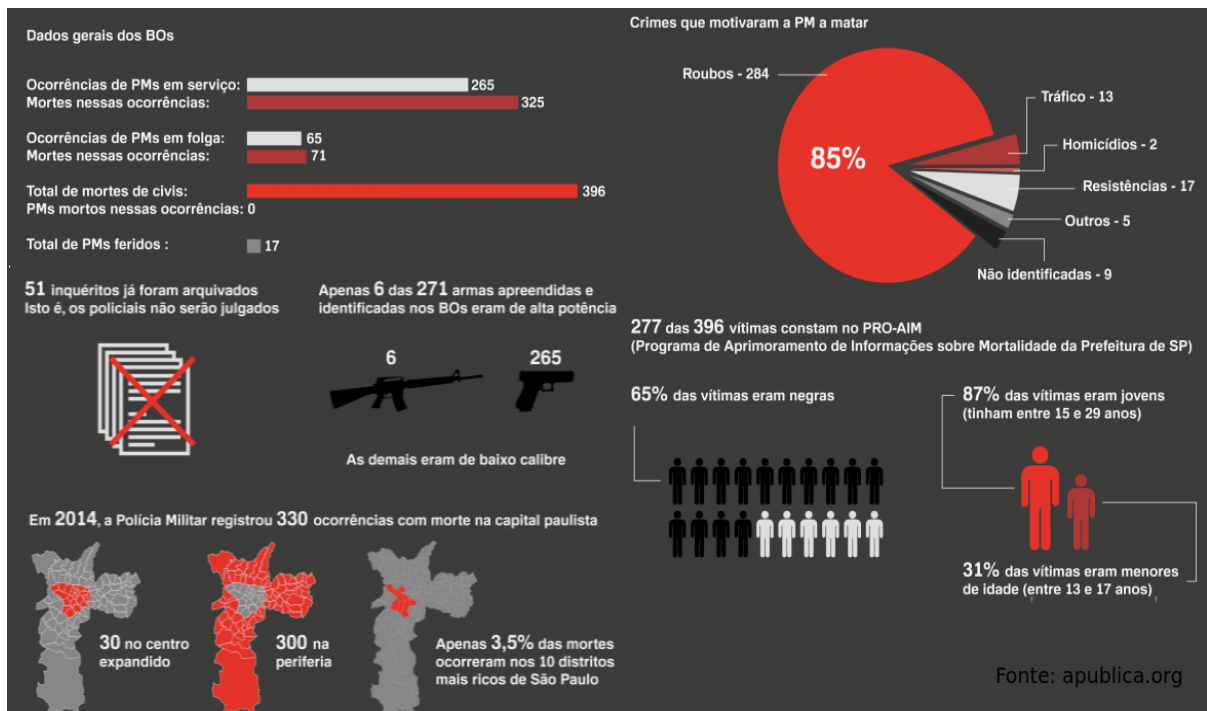
3.5 Exemplos de Reportagens

Com o fomento da cultura de transparência pelo poder público, os jornais e portais de notícia independente têm produzido reportagens de jornalismo de dados. A *Agência Pública* ou *aPublica* publicou em dezembro de 2015 a reportagem “396 mortes pela PM paulista: as histórias por trás dos B.O.s”⁸⁴. Os jornalistas Ciro Barros, Iuri Barcelos e José Cícero da Silva analisaram todos os boletins de ocorrência das mortes cometidas por polícias em 2014. Segundo a reportagem, durante seis meses de solicitações ao Departamento de Inteligência da Polícia Civil de São Paulo (Dipol), via Lei de Acesso à Informação, foram obtidos trezentos e trinta boletins de ocorrência. Por dois meses, foram analisadas e tabuladas informações que revelaram padrões de atuação em ocorrência que a polícia mata.⁸⁵

84 Disponível em: <<http://apublica.org/2015/12/396-mortes-e-o-padroa-da-pm-paulista/>>. acesso em 05 de nov. de 2016

85 Disponível em: <<http://apublica.org/2015/12/396-mortes-e-o-padroa-da-pm-paulista/>>. acesso em 05 de nov. de 2016

Figura 7 - Gráfico de reportagem Agência Pública



Fonte: Agência Pública

A partir dos dados analisados os repórteres encontraram um enredo padrão para as intervenções da PM:

“O enredo de uma intervenção letal da Polícia Militar (PM) em São Paulo começa com um homem jovem e negro suspeito do crime de roubo nas ruas da capital paulista. A PM sai em perseguição e, quando o encontra, os policiais são supostamente recebidos a tiros. Os PMs então “revidam a injusta agressão”, no jargão dos boletins de ocorrência – ou seja, atiram de volta. E são certos: poucos personagens dessa história sobrevivem. As armas das vítimas da PM costumam ser de baixo calibre: apenas seis entre as 271 supostamente apreendidas eram de alta potência, como fuzis ou escopetas. Percebemos também que as intervenções ocorrem principalmente em locais afastados do centro expandido, região que concentra as áreas mais nobres de São Paulo.” (AGÊNCIA PÚBLICA, BARROS..., 2015).

Após analisar os B.Os, os repórteres observaram que os únicos depoimentos dos crimes são dos PMs envolvidos, e que, apesar estes de serem sempre recebidos a tiros, conforme os boletins, poucos policiais são feridos nas operações.

Outra reportagem da *Agência Pública*, foi realizada a partir de informações obtidas por vazamento de dados feito pela organização *WikiLeaks*. Intitulada “Ligações perigosas: a DEA e as operações ilegais da PF brasileira”⁸⁶, a reportagem revela telegramas, de 1973, do embaixador dos Estados Unidos no Brasil, John Crimmins, ao Departamento de Estado chefiado por Henry Kissinger. Os documentos vazados mostram que o ex-diretor da Polícia Federal brasileira, general Canepa, efetuava prisões e extradições ilegais a pedido do departamento anti-drogas estadunidense.

O site de notícias *Fiquem Sabendo*, usando dados obtidos por meio da LAI, publicou a reportagem intitulada “Cai número de livros retirados de bibliotecas públicas de SP”⁸⁷. Na reportagem, o repórter Léo Arcoverde apresenta dados que mostram diminuição no empréstimo de livros nas bibliotecas municipais de São Paulo. Segundo dados obtidos pelo jornalista via LAI, caiu 3% a retirada de livros entre 2014 e 2015. A reportagem apresenta também uma lista dos livros mais retirados em bibliotecas paulistanas. O jornalista justifica a importância social da reportagem, assim como de todas do site, com informações da Constituição Federal.

O *Estado de São Paulo*, em reportagem sobre o câncer, intitulada “Câncer já é a principal causa de morte em 476 das 5.570 cidades brasileiras”⁸⁸, mostra levantamento feito pelo departamento de jornalismo de dados do jornal, o *Estadão Dados*, sobre as estatísticas de mortalidade em 2014. Os dados usados na matéria são da base de dados do Ministério da Saúde, o *Datasus*. Na reportagem, assinada pela jornalista Fabiana Cambricoli e Rodrigo Burgarelli, foram usados gráficos para ilustrar as informações passadas em texto.

86 Disponível em: <<http://apublica.org/2013/04/dea-caneppa-policia-federal-operacao-condo/>>. acesso em 05 de nov. de 2016

87 Disponível em: <<http://www.fiquemsabendo.com.br/2016/02/cai-numero-de-livros-retirados-de-bibliotecas-publicas-de-sp/>>. acesso em 05 de nov de 2016

88 Disponível em: <<http://saude.estadao.com.br/noticias/geral,cancer-ja-e-a-principal-cao-de-morte-em-476-das-5570-cidades-brasileiras,10000073577>>. acesso em: 05 de nov. de 2016

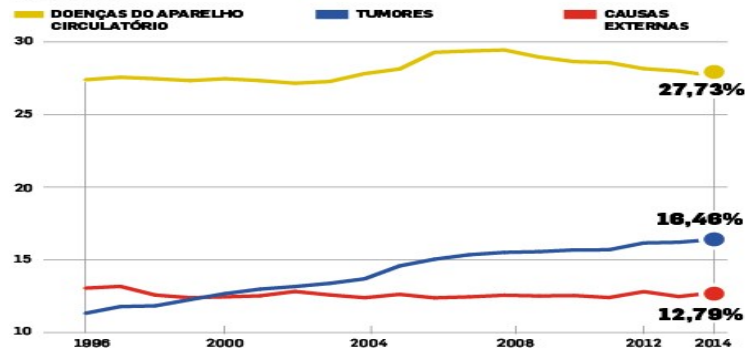
Figura 8 - Gráfico de reportagem do Estadão Dados

AVANÇO

● Câncer já é a maior causa de mortalidade em 476 cidades brasileiras

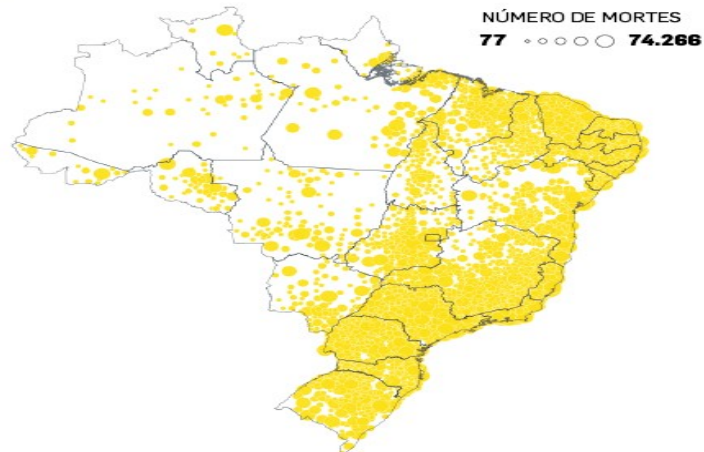
Evolução de óbitos pelas três principais causas

EM PORCENTAGEM



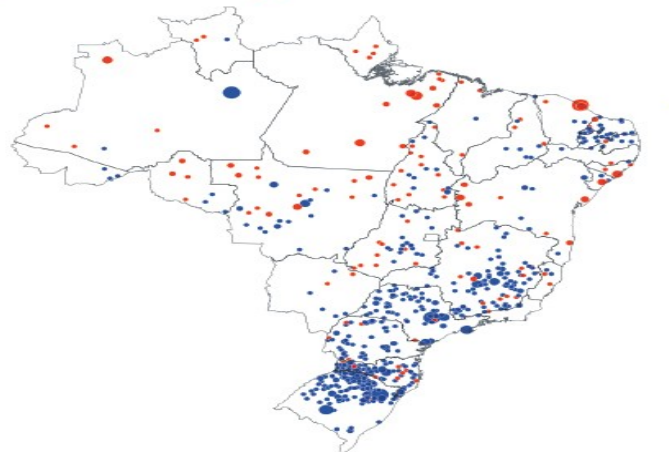
Maior causa de morte em cada município em 2014

● DOENÇAS DO APARELHO CIRCULATÓRIO



● TUMORES

● CAUSAS EXTERNAS



FONTE: DATASUS

INFOGRÁFICO/ESTADÃO

Fonte: Portal Estado de São Paulo

A organização *InfoAmazônia* de jornalismo ambiental, publicou uma grande reportagem intitulada “A política de desmatamento”⁸⁹, que explica quanto o governo federal gasta para combater o desmatamento e preservar a Amazônia. A reportagem contém análise de dados orçamentários do último mandato do ex-presidente Luís Inácio Lula da Silva e do primeiro mandato da ex-presidente Dilma Rousseff. A reportagem apresenta mapas animados demonstrando o avanço do desmatamento e gráficos animados com informações obtidas do *Siga Brasil*, sistema de informações sobre orçamento público desenvolvido pelo Senado Federal. Segundo Vedum⁹⁰, em nota explicativa da metodologia usada para fazer a reportagem, o *Siga Brasil* integra o Sistema Integrado de Administração Financeira do Governo Federal (SIAFI), Sistema Integrado de Planejamento e Orçamento (SIOP), Sistema de Elaboração da Lei Orçamentária (SELOR), Sistema de Informação das Estatais (SIEST), Sistema de Gestão de Convênios e Contratos de Repasse do Governo Federal (SICONV), dentre outras.

4 Considerações finais

Esta pesquisa objetivou discutir, por meio de uma abordagem teórica, com referencial empírico sempre que essa articulação se mostrou possível, as características de base do jornalismo de dados, bem como suas principais atuações no jornalismo atual. A abordagem visou compreender essa prática como uma modalidade jornalística dotada de rotinas particulares e investigar a eficácia do jornalismo de dados como meio de fiscalização dos poderes institucionais, públicos e privados, de consolidação da democracia e transparência pública no fluxo global de informação.

Em breve síntese, este percurso teórico-empírico foi realizado, primeiramente, por meio de fontes bibliográficas diversas voltadas ao tema, por meio de trabalhos acadêmicos e publicações avulsas em jornais, revistas e sites. Somou-se a isso a coleta de depoimento de profissionais com ampla experiência no tema, ligados algumas das atividades mais recentes e atualizadas do jornalismo de dados do

89 Disponível em: <<http://desmatamento.infoamazonia.org/analise/results/>>. acesso em 05 de nov. de 2016

90 Disponível em: <<http://desmatamento.infoamazonia.org/metodologia/>>. acesso em 05 de nov. de 2016

Brasil, além de estarem entre os precursores da modalidade no Brasil. Para situar a discussão no contexto do debate acadêmico, o que é de extrema relevância para esse tema, a pesquisa contou ainda com a colaboração de uma acadêmica de grande experiência, coordenadora do curso de Jornalismo na Universidade Federal de Uberlândia (UFU) e presidente do Fórum Nacional de Professores de Jornalismo (FNPJ), gestão 2012-2016. Por fim, procurou-se realizar uma atualização aprofundada das ferramentas mais usadas na construção da notícia e publicações jornalísticas com dados, com a finalidade de elucidar a rotina técnica do jornalismo de dados e apresentar um acervo básico de consulta para o jornalista interessado, ainda que, em se tratando de recursos tecnológicos, haja necessidade inevitável de revisão dessa listagem em breve.

A partir de tais procedimentos pode-se, assim, chegar a algumas considerações conclusivas e outras necessariamente em construção, em se tratando de um tema atual. O jornalismo de precisão não é dependente de computadores, por ser possível para um jornalista realizar experiências empíricas, baseados em métodos de sociologia, sem o uso de um computador. Já o trabalho com auxílio de computadores seguiu o fluxo do desenvolvimento tecnológico com a consolidação das redes de computadores. O jornalismo de precisão, aliado as técnicas de RAC, possibilitou ao jornalista investigar com maior rapidez e qualidade. Com computadores, o trabalho de apuração ficou mais acelerado, mais fácil evitar erros humanos de tabulação de resultados de questionários, otimizando a objetividade do jornalismo – um preceito tão fundamental da profissão.

Costella (2002, p.22) nos conta que o papel não gerou a tipografia, nem o correio, mas foram condições de eficiência que acarretaram a possibilidade do seu uso. Ou seja, sem o papel, o que seria da tipografia que necessitava de papel para poder imprimir conteúdo?: ou sem o correio, que serviu de agente de aproximação entre os homens, ao incrementar o tráfego de informações:

“O jornalismo impresso tipograficamente subiu ao palco da História no momento em que a experiência tipográfica, calejada nos livros, se somou à experiência jornalística desenvolvida com a gazeta manuscrita”. (COSTELLA, 2012, p.22).

São vários os fatores que ocorreram em momentos distintos da história do mundo, e que, juntos, foram o ponto de partida para o nascimento do jornalismo. Portanto, o desenvolvimento do trabalho com RAC, aliado ao jornalismo de precisão, podem ser considerados os propulsores do desenvolvimento do jornalismo de dados. Por conseguinte, o jornalismo de dados se diferencia da RAC no uso de vários elementos para construir a notícia. A RAC é o trabalho com planilhas digitais, para cálculos estatísticos. Já o jornalismo de dados vai além, trabalhando com dados obtidos de fontes públicas ou privadas; usando técnicas de extração de dados, em grande escala, antes comuns apenas para programadores; analisando esses dados para dar significados e interpretações diretas pelo instrumental jornalístico, para então trabalhar formas de visualização, como a configuração interativa do *Basômetro*, que dá ao usuário o poder de escolher o que quer saber, ou uma visualização sem interação do leitor com a notícia, mas que seja de fácil compreensão.

Os avanços em liberação de dados governamentais, antes privados ao acesso da sociedade, foram de fundamental importância para o fomento do jornalismo de dados no Brasil. Bramatti (2014, p.220), explica que a publicação dos dados governamentais na internet facilitou a vida de jornalistas, porque não é mais necessário esperar em uma fila em Brasília para pegar o registro das votações do dia na Câmara Federal. A disponibilidade dos dados na rede não só facilita e agiliza o trabalho jornalístico, como também torna acessível a informação pública a jornais de menor porte e ao cidadão comum, que não precisa mais esperar pelo resultado das votações legislativas e outros dados nos jornais, já que está disponível, para acesso na internet, informações dos três poderes. Com isso, pode-se dizer que a cultura de transparência é instrumento importante para a consolidação da democracia e o jornalismo de dados é agente que corrobora para o Estado democrático de Direito. Com a afirmação da cultura de transparência, observou-se que houve aumento na produção de matérias pautadas por dados, assim como a criação de diversos sites de jornalismo independente, que usam dados obtidos institucionalmente, agora públicos, para informar com maior objetividade e clareza.

O empreendedorismo com jornalismo de dados, no Brasil, se mostrou ainda modesto, os pioneiros como a *J++*⁹¹, empresa que ofertava serviços de jornalismo de

91 Anexo A

dados, não obtiveram sucesso no empreendimento, decorrência de um mercado instável, o que tem demissões em massa e cortes substanciais de despesas nos veículos jornalísticos. Já o *Volt Data Lab*⁹², do jornalista Sérgio Spagnuolo, é um projeto em desenvolvimento. Com a necessidade de se informar da população e uma grande quantidade de dados sendo disponibilizados pelos poderes institucionais, o jornalismo de dados pode ser um mercado potencial para trabalhar com notícias para o grande público, assim como para públicos específicos, como faz o *The Economist*, o *Bloomberg*, a *Thomson Reuters* e tantos outros veículos que deixaram de ser agências de comunicação para prestar serviços de informação mais aprofundados, como conta Lorenz:

“E observem a *Economist*. A revista tem construído uma marca excelente, influente em seu aspecto de mídia. Ao mesmo tempo, a "Economist Intelligence Unit" agora é mais uma empresa de consultoria, elaboração de relatórios sobre tendências relevantes e previsões para quase todos os países do mundo. Eles estão empregando centenas de jornalistas e alegam servir cerca de 1,5 milhão de clientes em todo o mundo.” (LORENZ, 2012).

No entanto, a *Volt Data Lab* foca seu negócio apenas em jornalismo, deixando de lado o mercado em ascensão como o de análise de dados para estratégia empresarial e estratégia de marketing – algo menos comum no cenário midiático internacional, como informa Lorenz:

“Em todo o mundo, existe uma onda de empresas iniciantes no setor, cobrindo uma vasta gama de áreas, por exemplo, a Timetric, que tem por objetivo "reinventar a pesquisa em negócios", OpenCorporates, Kasabi, Infochimps, e Data Market. Muitas delas são, indiscutivelmente, experimentos, mas juntas, podem ser consideradas um importante sinal de mudança” (LORENZ, 2012).

É inegável que o jornalismo de dados seja um mercado em expansão, e que deva ser explorado, em todas as áreas e meios, mas, para tal, o jornalista deve parar de pensar apenas como jornalista somar às habilidades técnicas e humanísticas uma atitude empreendedora.

92 Spagnuolo concedeu entrevista para a pesquisa, mas por falha técnica a mesma não foi gravada

Dessa maneira, faz-se necessário formar um jornalista mais bem preparado para o mercado de trabalho, não só para empreender, mas também para exercer o jornalismo de dados em toda sua complexidade tecnológica. Tonus⁹³ deixa claro que é de responsabilidade das universidades organizarem o currículo acadêmico do curso de Jornalismo, ficando elas responsáveis por entender as demandas e preparar o profissional adequadamente para o mercado de trabalho. No entanto, nota-se clara dificuldade das universidades em acompanhar os avanços tecnológicos da profissão. O FNPJ poderia sinalizar nas diretrizes curriculares a importância do jornalismo de dados e da matemática para o jornalismo.

Dados disponíveis para serem analisados são o cerne do jornalismo de dados. Sem dados, enfim, não há jornalismo de dados. Isto posto, conclui-se que a transparência dos poderes institucionais, abrindo suas bases de dados para o público, é fundamental para fomentar o jornalismo de dados. Desta maneira, o jornalismo de dados é parte fundamental para consolidar o Estado democrático e a cultura de transparência, levando ao cidadão informação analisada e contextualizada para melhor compreensão da realidade. Como exemplo da colaboração do jornalismo de dados para o fomento da participação da sociedade, está na ferramenta *Basômetro* que possibilita a visualização, de forma clara, do comportamento dos parlamentares nas votações nas casas legislativas. Outros projetos, sem fins lucrativos, trabalham o jornalismo de dados com o objetivo de levar informações importantes para participação política do cidadão comum, com recortes segmentados, como o portal *Transparência Brasil* e seus projetos *Meritísimos*, *Àsclaras* e *Excelências*. É importante salientar também o papel fundamental dos vazamentos de dados na consolidação do Estado democrático, ao permitir a publicação de informações de interesse social, e comprometedoras da atuação pública em diversas esferas, e que são mantidos em sigilo.

Ao jornalista cabe, portanto, o papel de importantíssimo de traduzir os dados disponibilizados por governos ou vazados por ciberativistas, para que assim o jornalismo possa alcançar a sua função de atender o interesse público e formar cidadãos de forma cada vez mais objetiva. Uma gama de agentes compartilha a responsabilidade de contribuir para que essa modalidade jornalística se torne uma prática efetiva e ainda mais relevante no jornalismo atual, e para a democracia: o

93 Anexo C

profissional da área, as universidades, o poder público e os veículos de comunicação.

5 REFERÊNCIAS BIBLIOGRÁFICAS

ABRAMO, Cláudio Weber. **Lei de Acesso à Informação no Brasil: Um longo caminho a percorrer.** In: GRAY et al. (orgs.). Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em: <http://datajournalismhandbook.org/pt/getting_data_2.html>. Acesso em: 22 set. 2016.

ALI, Tanveer. **Produtores de notícias locais aderem ao jornalismo de dados.** 2014. Disponível em: <http://observatoriodaimprensa.com.br/grande-pequena-imprensa/_ed812_produtores_de_noticias_locais_aderem_ao_jornalismo_de_dados/>. Acesso em: 23 out. 2016.

BARBOSA, S. A.; TORRES, V. **Paradigma 'Jornalismo Digital em Base de Dados' em expansão contínua:** modos de narrar, formatos e visualização para conteúdo. São Paulo: Galaxia, 2013.

BARROS, Ciro; BARCELOS, Iuri; SILVA, José Cícero da. **396 mortes pela PM Paulista:** as histórias por trás dos BOs. 2015. Disponível em: <<http://apublica.org/2015/12/396-mortes-e-o-padrao-da-pm-paulista/>>. Acesso em: 6 de out. 2016.

BOUNEGRU, Liliana. **Jornalismo de dados em perspectiva.** In: GRAY et al. (orgs.). Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em: <http://datajournalismhandbook.org/pt/introducao_4.html>. Acesso em: 13 set. 2016.

BRADSHAW, Paul. **O que é jornalismo de dados?**. In: GRAY et al. (orgs.). Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em:

<http://datajournalismhandbook.org/pt/introducao_0.html>. Acesso em: 22 set. 2016.

BRASIL. Presidência da República. Secretária da Comunicação Social. **Pesquisa Brasileira de Mídia 2015**: hábitos de consumo de mídia da população brasileira.

Brasília: Secom, 2014. Disponível em:

<<http://www.secom.gov.br/atuacao/pesquisa/lista-de-pesquisas-quantitativas-e-qualitativas-de-contratos-atuais/pesquisa-brasileira-de-midia-pbm-2015.pdf>>. Acesso em: 23 de out. 2016

CAIRO, Alberto. **Existe jornalismo de dados e visualização no Brasil?**. In: GRAY et al. (orgs.). Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em:<http://datajournalismhandbook.org/pt/introducao_6.html>. Acesso em 13 out. 2016.

CASTILHO, Carlos. **Jornalismo de dados cria primeiro repórter milionário**. 2013. Disponível em: < <http://observatoriodaimprensa.com.br/codigo-aberto/jornalismo-de-dados-cria-primeiro-reporter-milionario-na-imprensa/>>. Acesso em: 13 out. 2016.

CHEDE, Cezar. **Você realmente sabe o que é Big Data?** 2012. Disponível em: <https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/voce_realmente_sabe_o_que_e_big_data?lang=en>. Acesso em 25 de out. 2016.

COSTELLA, Antônio. **Comunicação – do grito ao satélite** (história dos meios de comunicação). São Paulo: ed. Mantiqueira, 1984.

DANTAS, Humberto; TOLEDO, Jose Roberto; TEIXEIRA, Marco Antonio Carvalho. **Análise política & jornalismo de dados**: ensaios a partir do basômetro. Rio De Janeiro: FGV, 2014.

DOMSCHEIT-BERG, Daniel. **Os bastidores do WikiLeaks**: A história do site mais controverso dos últimos tempos escrita seu ex-porta-voz. Rio de Janeiro: Elsevier, 2011.

ÉPOCA. **IBGE diz que cometeu "erros graves" na pesquisa sobre desigualdade**. 2014. Disponível em: <<http://epoca.globo.com/tempo/noticia/2014/09/ibge-diz-que-cometeu-berros-gravesb-na-pesquisa-sobre-desigualdade.html>>. Acesso em: 11 out. 2016.

ESCOLA DE DADOS. **Raspagem de dados**: ferramentas gratuitas e fáceis de usar. Disponível em: <<http://escoladedados.org/manual/cursos/raspagem/>>. Acesso 24 de out. 2016.

ESTADÃO DADOS. **Ferramenta analisou 700 mil registros de votos no congresso**. 2015. Disponível em: <<http://politica.estadao.com.br/noticias/geral,ferramenta-analisou-700-mil-registros-de-votos-no-congresso,1789713>>. Acesso em: 22 out. 2016.

ESTADO DE SÃO PAULO. **11,5 milhões de documentos expõem corrupção global**. 2016. Disponível em: <<http://politica.estadao.com.br/noticias/panama-papers,11-5-milhoes-de-registros-financeiros-expoem-corrupcao-global,10000024510>>. Acesso em: 11 set. 2016.

ESTADO DE SÃO PAULO. **Estadão dados une transparência e interatividade**. 2012. Disponível em: <<http://politica.estadao.com.br/noticias/geral,estadao-dados-une-transparencia-e-interatividade,872163>>. Acesso em: 23 de out. 2016.

FALEIROS, Gustavo. **InfoAmazônia**: o diálogo entre jornalismo e dados geográficos. In: GRAY et al. (orgs.). Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em: <http://datajournalismhandbook.org/pt/estudos_de_caso_1.html>. Acesso em: 13 set. 2016.

FEIJÓ, Bruno Vieira. **Para os cientistas de dados não há desemprego**. 2016. Disponível em: <http://exame.abril.com.br/ciencia/para-os-cientistas-de-dados-nao-ha-desemprego/>. Acesso em: 26 de out. 2016.

GOODMAN, Amy. **"Pretrial Punishment"**: Julian Assange Remains in Ecuadorean Embassy Fearing Arrest If He Leaves. 2015. Disponível em: http://www.democracynow.org/2015/5/28/pretrial_punishment_julian_assange_rema ins_in>. Acesso em: 28 de out. 2016.

GRAY, Jonathan; CHAMBERS, Lucy; BOUNEGRU, Liliana (orgs.). **Manual de jornalismo de dados**: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em: <http://datajournalismhandbook.org/pt/>. Acesso: 10 de maio 2016.

HOLOVATY, Adrian. **A fundamental way newspaper sites need to change**. Adrian Holovaty, 6 set. 2006. Disponível em: <http://www.holovaty.com/writing/fundamental-change/>>. Acesso: 25 out. 2016.

KAYSER-BRIL, Nicolas. **Guia rápido de trabalho de campo**. In: GRAY et al. (orgs.). Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em: http://datajournalismhandbook.org/pt/getting_data_0.html>. Acesso em: 13 set. 2016.

LAGE, Nilson. **Jornalismo de precisão e RAC**. 2012. Disponível em: <http://www.observatoriodaimprensa.com.br/artigos/da090420033.htm>>. Acesso em: 1 maio 2016.

LÉVY, Pierre. **Cibercultura**. São Paulo: ed. 34, 1999.

LOGAREZZI, Lia. **Guia prático da lei de acesso à informação** [livro eletrônico]. São Paulo: Artigo 19, 2016. Disponível em: <http://artigo19.org/blog/2016/10/24/guia-pratico-da-lei-de-acesso-a-informacao/>>. Acesso em 29 de out. 2016

LONGHI, Raquel; D'ANDREA, Carlos. **Jornalismo convergente: reflexões, apropriações, experiências**. Florianópolis: Insular, 2012.

LORENZ, Mirko. **Modelos de Negócios para Jornalismo de Dados**. In: GRAY et al. (orgs.). Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em: <http://datajournalismhandbook.org/pt/na_redacao_9.html>. Acesso em: 29 set. 2016.

LORENZ, Mirko. **Por que jornalistas devem usar dados?** In: GRAY et al. (orgs.). Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em: <http://datajournalismhandbook.org/pt/introducao_4.html>. Acesso em: 13 set. 2016.

MARCONDES, Carlos Henrique. **“Linked data” – Dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web**. 2012. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17n34p171/22782>>. Acesso em: 19 de set. 2016.

MENÁRGUEZ, Ana Torres. **Não é preciso ser matemático para ser especialista em ‘big data’**. 2015. Disponível em: <http://observatoriodaimprensa.com.br/e-noticias/_ed839_nao_e_preciso_ser_matematico_para_ser_especialista_em_big_data/>. Acesso em: 21 jun. 2016.

MORALES, Carlos. **Jornalismo de dados: Recursos para ajudar repórteres a coletar, organizar e analisar dados**. 2011. Disponível em: <<https://knightcenter.utexas.edu/pt-br/blog/jornalismo-de-dados-recursos-para-ajudar-reporteres-coletar-organizar-e-analisar-dados>>. Acesso em: 03 set. 2016.

O GLOBO. **Entenda o que é o SwissLeaks**. 2015. Disponível em: <<http://oglobo.globo.com/brasil/entenda-que-o-swissleaks-15568429>>. Acesso em: 15 maio 2016.

OPEN KNOWLEDGE FOUNDATION. **Manual de dados abertos** [livro eletrônico].

São Paulo: OKF, 2011. Disponível em:

<https://commons.wikimedia.org/wiki/File:Manual_Dados_Abertos.pdf>. Acesso em 30 de out. 2016

PENA, Felipe. **Teoria do jornalismo**. 3. ed. São Paulo: Contexto, 2013.

PENA, Rodolfo F. Alves. **Era da Informação**. Disponível em:

<<http://mundoeducacao.bol.uol.com.br/geografia/era-informacao.htm>>. Acesso em: 28 de out. 2016.

PERCÍLIA, Eliene. **"Código Aberto"**; Brasil Escola. Disponível em

<<http://brasilecola.uol.com.br/informatica/codigo-aberto.htm>>. Acesso em 24 de outubro de 2016.

PIMENTA, Ângela. **Redações multidisciplinares**. 2015. Disponível em:

<<http://observatoriodaimprensa.com.br/tendencias/redacoes-multidisciplinares/>>. Acesso em: 27 de out. 2016.

RAMOS, Angélica Peralta. **Alguns exemplos selecionados**. In: GRAY et al. (orgs.).

Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em:

<<http://datajournalismhandbook.org/pt/>>. Acesso em: 5 out. 2016.

ROGER, Simon. **Alguns exemplos selecionados**. In: GRAY et al. (orgs.). Manual

de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens. São Paulo: Abraji/EJC, 2012. Disponível em:

<<http://datajournalismhandbook.org/pt/>>. Acesso em: 5 abr. 2016.

SILVER, Nate. **O Sinal e o Ruído**. 1. Ed. Rio de Janeiro, 2013.

SIQUEIRA, Ethevaldo. **Chegou a era do zettabyte**. 2012. Disponível em: <https://www.ibm.com/developerworks/community/blogs/ethevaldosiqueira/entry/chegou_a_era_do_zettabyte?lang=en>. Acesso em: 6 abr. 2016.

TOLEDO, José Roberto de. **Basômetro**: Passando o poder da narrativa para o usuário. In: GRAY et al. (orgs.). **Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens**. São Paulo: Abraji/EJC, 2012. Disponível em: <http://datajournalismhandbook.org/pt/estudos_de_caso_0.html>. Acesso em: 29 set. 2016.

TOLEDO, José Roberto de. **Basômetro**: passando o poder da narrativa para o usuário. In: GRAY et al. (orgs.). **Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens**. São Paulo: Abraji/EJC, 2012a. Disponível em: <<http://datajournalismhandbook.org/pt/>>. Acesso: 30 set. 2016.

TRÄSEL, Marcelo Ruschel. **Entrevistando Planilhas**: Estudo da crença e do ethos de um grupo de profissionais de jornalismo guiado por dados. Disponível em: <<http://tede2.pucrs.br/tede2/bitstream/tede/4590/1/461784>>.pdf. Acesso em: 01 de out. 2016.

TRÄSEL, Marcelo Ruschel. **O jornalismo guiado por dados numa perspectiva brasileira**. In: GRAY et al. (orgs.). **Manual de jornalismo de dados: como os jornalistas podem usar dados para melhorar suas reportagens**. São Paulo: Abraji/EJC, 2012. Disponível em: <http://datajournalismhandbook.org/pt/introducao_5.html>. Acesso em: 29 set. 2016.

6 ANEXOS

ANEXO A – Entrevista Thiago Mali

Como o jornalismo de dados colabora com o jornalismo?

O jornalismo de dados é mais uma ferramenta do jornalismo do que algo separado, o fato de você adquirir habilidades de tabular dados, entrevistar bases de informações, de extrair conclusões a partir de tabelas grandes, auxilia na verdade todo um ferramental que auxilia um jornalista a conseguir obter informações mais relevantes para sua matéria e muitas vezes mais precisas. Jornalismo de dados de maneira nenhuma substitui jornalismo tradicional, um repórter ruim que usar ferramentas de jornalismo de dados continuará sendo um repórter ruim. Agora um repórter bom, que tiver conhecimento de como usar esse tipo de ferramenta pode fazer reportagens muito mais interessantes, muito mais precisas, muito mais aprofundadas. Junto com as técnicas tradicionais do jornalismo, as técnicas do jornalismo de dados possibilita que você não fique tão a reboque de declarações oficiais ou de assessoria de imprensa, quando você consegue ir direto às bases de dados e verificar por você mesmo quais são as informações, quais são as respostas para as perguntas que você tem.

O que diferencia JGD da RAC?

Essa é uma discussão que na verdade não existe um consenso sobre isso. O que existe é que a maior parte dos teóricos coloca dentro de um histórico que a RAC surge ali no meio da década de 50 para 60 nos Estados Unidos, quando alguns jornalistas e repórteres começam de uma maneira mais sistematizada a usar essas técnicas de dados dentro do jornalismo. Então, tem primeiro pesquisas, por exemplo, de opinião para prever qual vai ser o candidato eleito que era algo que era novidade naquela época. Isso faz parte de uma maneira das técnicas de RAC de jornalismo de dados, com o tempo essas técnicas vão se aprofundando e vão melhorando. O jornalismo de dados em si, é um termo que se populariza por volta de 2009/2010. A principal corrente dos teóricos que estudam isso colocam como uma evolução um certo tipo de evolução, não exatamente natural da RAC para o jornalismo de dados, no sentido que o jornalismo de dados permite uma customização muito maior da maneira como o leitor acessa as informações. Então por termos nessa virada ali de 2009 até 2012, uma virada da década muito forte em termos de transparência de abertura de bases de dados governamentais e órgãos

públicos em todo o mundo, brasil inclusive, você possibilitou muito mais informação a disposição dos jornalistas e possibilitou que ao invés deles pegarem uma tabela, como se fazia na RAC, e analisar a exaustão para retirar dessa tabela informações mais exclusivas, como agora essas bases estão abertas em grande volume, em grande quantidade para todo mundo, possibilitou fazer outros tipos de análises, um pouco mais globais, então, ao invés de você pegar e extrair grandes conclusões de uma tabela, o que você faz é possibilitar que o leitor também consulte essa tabela. Eu posso pegar e mostrar a taxa de desemprego hoje no brasil e eu posso permitir que o leitor, a partir de todos os dados da tabulação da taxa de desemprego, ele tenha acesso e consulte qual é a taxa de desemprego para pessoas, por exemplo, na mesma categoria que ele, então pessoas que tenham o mesmo nível de escolaridade, pessoas do sexo masculino ou feminino, negras, brancas, ou pardas, enfim, você tem essa possibilidade de segmentação de customização de organizar as informações de uma maneira a permitir que o leitor consiga por si mesmo ler e dizer até para as aplicações jornalísticas que usam jornalismo de dados o que ele quer entender. O RAC, ele está dentro desse conceito maior de jornalismo de dados, é uma parte disso, enfim, de você usar as técnicas para obter as conclusões mais gerais, mas o jornalismo de dados, agora como um todo, passa a englobar esse conceito de customização das notícias, conceito de você deixar o leitor entender um pouco melhor o que está em jogo ali, selecionar o que ele se interessa mais naquele momento.

O jornalismo de dados seria uma evolução da RAC?

A rac está dentro do jornalismo de dados, o jornalismo de dados seria uma maneira de usar técnicas de dados que termite maior customização, vem depois da RAC.

Com o acesso à informação pública disponível ao cidadão, dá para qualquer um, independente da formação, fazer jornalismo de dados?

Na verdade, da mesma maneira que tem qualquer pessoa fazendo jornalismo hoje, em blogs, canais independentes, se pode ter pessoas fazendo jornalismo com técnicas de dados também, você tendo esse letramento, nas técnicas de dados,

sabendo usar corretamente planilhas, formulas, tabulações, programas de bases de dados, da mesma maneira que se pode fazer jornalismo sem ser formado em jornalismo, não há nenhum obstaculo.

A produção de matérias a partir de dados pode fazer com que o jornalista deixe de observar a sociedade para trabalhar apenas com dados, planilhas?

Pelo contrário, se você for usar da maneira mais indicada as bases de dados do jornalismo, você na verdade consegue olhar com um olhar muitas vezes mais humanizado e mais aprofundado interessado. Para dar um exemplo, tem uma matéria que foi feita pelo pessoal do estadão, Daniel bramatti, onde ele pegou dados de homicídios no brasil inteiro, jogou no mapa e foi ver quais são as regiões e microrregiões do brasil onde se tinha mais homicídios, aí ele encontrou uma região ali no sul do Pará onde tinha uma taxa completamente absurda de homicídios comparável a de guerras. E daí, ao invés dele simplesmente publicar isso, ele foi até a região entender que havia um conflito entre madeireiros e que estava deixando um monte de gente morta. Descobriu uma história, aí contou do ponto de vista do delegado que não conseguia combater o crime ali, que estava sendo ameaçado pelos madeireiros, os madeireiros brigando uns com os outros. A partir de uma base de dados ele encontrou uma história muito mais humanizada do que ele não encontraria de outra maneira. Aí cabe ao jornalista de não ficar no papel raso de encontrar os dados e mostrar “é isso ou aquilo” e também ir pra campo. O jornalismo de dados é uma ferramenta que permite que o jornalista aprofunde suas matérias, agora como usar essa ferramenta cabe ao jornalista, ele pode ficar simplesmente em mostrar gráficos e dados ou ele pode a partir desses dados que ele recolheu ir a campo entender o que está acontecendo de uma maneira bem humanizada inclusive.

Dá para ganhar dinheiro com jornalismo de dados?

Se você conhecer alguém que está ganhando dinheiro no brasil você me avisa porque até agora não conheço ninguém que está realmente ganhando dinheiro com jornalismo de dados, com dados obviamente tem uma gente ganhando dinheiro em

outros tios de mercado. O mercado está muito ruim, o mercado jornalístico em geral, tem as redações demitindo muita gente, sobram poucas pessoas dentro das principais redações jornalísticas e tem algumas pessoas tentando caminhos alternativos, fazer de maneira mais independente jornalismo de dados, o que é muito interessante, muito lovável, mas até agora não existe nenhum modelo que seja considerado super bem sucedido, o mais conhecido que está fazendo isso é o Volt Data Lab, mas até onde eu sei é algo embrionário, ele está testando pra ver se funciona.

A demanda por trabalhos de jgd aumentou?

Você pode considerar a demanda de duas maneiras: dos leitores sim aumentou, eles querem saber mais, tem muita gente interessada, dos cursos que a gente dá tem muito jornalista interessado, cada vez mais, em aprender esse tipo de técnica. Porém nas redações têm o problema de como teve muitos cortes de pessoal e jornalismo de dados exige um certo tempo e dedicação para fazer, muitas vezes os chefes os editores eles não querem deixar uma pessoa dedicada muito tempo a analisar dados para conseguir de repente uma ou duas matérias, a partir dali prefere usar essa mesma pessoa para escrever várias notas que podem render mais audiência no mesmo tempo, existe uma demanda de público e jornalistas para tentar fazer mais isso e fazer saber mais isso, mas existe um problema circunstancial, hoje o mercado não permite que isso avance mais rapidamente no Brasil.

É necessário maior tempo que uma matéria comum?

Se você for fazer uma matéria mais aprofundada com dados, muito possivelmente, principalmente quando se trata de visualizações ou dessas matérias que a gente chama de jornalismo de dados em que tem contexto mais customizado onde o leitor possa ir e vir, você vai precisar de um jornalista com tempo para analisar, tabular e limpar os dados, você vai precisar de um programador que crie uma ferramenta para conseguir mostrar esses dados de uma maneira mais interessante, ou até pode usar ferramentas gratuitas que tem suas limitações e exige tempo. Precisa de bastante tempo para ordenar, para fazer essas análises, muito mais tempo do que ligar para

alguém e fazer uma entrevista. Exige uma certa dedicação sim para fazer coisas mais aprofundadas, coisas mais básicas e triviais em jornalismo de dados dá fazer e tem sido feito, como por exemplo, comparar bens de candidatos, comparar arrecadação de campanha, isso já começa a ser adotado por varias redações, isso começa a fazer parte do dia a dia.

A busca por cursos de jgd cresceu?

Em seis edições de cursos de jornalismo de dados na abraji, em todas preencheu todas as vagas e em todas edições sempre fica uma grande fila de espera, a fila de espera agora tem trezentas pessoas querendo fazer o curso que deve abrir inscrição em dezembro. A gente percebe um interesse crescente por conta disso, que a gente abre em dois dias esgotam as vagas. Existe uma demanda de muitos jornalistas querendo entender melhor como fazer isso, o que me parece um pouco obvio no sentido que é uma clara vantagem para quem hoje sabe rapidamente entrar em um site de base de dados público, abrir aquilo, olhar , baixar as tabelas e entender as informações que estão por trás daquilo, a pessoa consegue rapidamente ter uma vantagem sobre outro jornalista que não sabe fazer isso e fica ligando para a assessoria de imprensa para pedir as mesmas informações. Isso é encarado e visto pelo jornalista como uma vantagem competitiva no sentido de que possa produzir melhor e mais rápido e mais preciso alguns tipos de notícias, de informações, a tendencia é que o interesse só cresça e a tendencia é que em breve, na verdade e isso é agora saber usar tecnicas de jornalismo de dados deixe de ser um diferencial e comece a ser um básico, algo que todo jornalista tem que saber, afinal de contas temos muitas bases de dados hoje, é um tipo de conhecimento e habilidade que já não pode mais faltar entre os jornalistas. Infelizmente uma boa parte das faculdades de jornalismo brasileiro hoje não tem isso bem colocado dentro do seu curriculo, deve está correndo atras de se atualizar nesse conhecimento.

Para trabalhar com jgd o que o estudante de jornalismo tem saber/aprender?

É ridículo que as universidades ainda hoje na maior parte não tenham incorporado [jornalismo de dados] no seu currículo. Muitas incorporam e dão o nome de uma

matéria, mas na prática o que se ensina lá é algo muito distante do que se vai usar no dia a dia, correndo atrás de apuração e notícias. O que a gente recomenda é primeiro a pessoa ter familiaridade de mexer com tabelas e excel, de softwares, planilhas eletrônicas, para conseguir a partir delas filtrar informações, classificar informações, separar quais são os pontos fora da curva e dentro da curva. Voltar ao que muitos jornalistas esqueceram porque não tiveram mais contato desde o colegio, que é o cálculo percentual, cálculo de evolução de dados com o passar de tempo. Então são coisas mais básicas que todo jornalista que sai da faculdade já deveria saber qual que é o lide, qual que é a notícia mais importante que está ali dentro daquelas informações, é necessário que se aprenda a entrevistar as bases de dados e não é ensinado na maior parte das faculdades de jornalismo que eu conheço no Brasil.

Como funciona a LAI na prática nas redações?

Há bastante dificuldade, a LAI é muito bem vinda, sancionada em 2012 e permitiu que a gente tenha um arcabouço legal, pra defesa da transparência e para permitir mais acesso à informação públicas pelos jornalistas. A LAI funciona dois jeitos: uma é a que a gente chama de transparência ativa, porque os municípios, os poderes legislativos, estaduais, municipais, federais, todos os órgãos públicos são obrigados pela lei ativamente divulgar uma série de informações na internet, ou seja, eles precisam ter um site, dentro deste site precisa ter determinadas informações, como por exemplo, quadro de servidores, salários que é pago a eles. O fato de estar previsto em lei já ajuda bastante, embora não seja todos os órgãos públicos que cumpram cem por cento infelizmente isso. Existe a LAI passiva também que é um canal que todos os órgãos públicos são obrigados a abrir para receber pedidos de informação dos cidadãos em geral incluindo aí jornalistas. O fato da lei obrigar esses entes públicos a fazer isso melhora muito a vida do jornalista no sentido que se os caras se negam a fazer isso, o jornalista e o cidadão comum está aparado pela lei para pedir essas informações, para exigir que esse tipo de informação seja entregue, você pode recorrer, você pode usar os mecanismos legais que estão a disposição. Ainda falta bastante para que o Brasil, em geral, atingir os níveis bons de cumprimento da própria lei, se pega a controladoria geral da união, por exemplo, tem

um ranking onde ela mostra os municípios que...eles criam uma escala de zero a dez e dão uma nota, dez significa que o município está cumprindo, está fazendo mais que o básico e zero que o município não está cumprindo a lei. Cinquenta por cento dos municípios do Brasil ficam com a nota entre zero e um, ou seja, ainda é muito ruim, muito mal feito o cumprimento da lei na maior parte das cidades brasileiras, isso é algo que vem melhorando.

ANEXO B - Entrevista Rodrigo Bugarelli

Jornalismo de dados é muito em momentos de crise, dados te ajuda a otimizar alguns processos que demora muito para fazer, então se antes para saber o resultado de uma eleição tinha que ir um a um, então precisa de muita gente. Se você domina excel, banco de dados, você consegue fazer muito mais trabalho em menos tempo. Um trabalho de melhor qualidade, de mais controle sobre o resultado. Tendo a crise, contrate um jornalista de dados porque consegue fazer um mais aprofundando em menos tempo.

Como tem sido fazer reportagem mais aprofundadas com pouco pessoal?

Essa é a pior parte, mas depende do que quer fazer com o veículo de comunicação. Se você quer cobrir o dia. Vai precisar de muita gente, com jornalismo de dados para fazer matérias rápidas para o dia. Mas a ideia é fazer trabalhos mais aprofundados e muito bem explicados. Isso requer um jornalismo diferente, jornalismo de fôlego.

O jornalismo de dados ele aproxima muito o jornalista dos dados e não tem como você analisar um conjunto de dados sem visualizar ele de alguma maneira, então jornalistas de dados cada vez mais estão aprendendo jornalismo de dados. As pessoas estão sendo acostumadas ter mais acesso ao jornalismo de dados e estão gostando. A nossa principal ferramenta de jornalismo de dados é o basômetro que você mesmo consegue usar, como o deputado votou, se ele é governista ou não em relação ao presidente. E as pessoas acessam, tem muito acesso, porque com a

visualização de dados hoje, você consegue aproximar as pessoas da matéria-prima sabe, então você diminui a distancia delas com os dados.

Eu uso apenas software livre, linux, banco de dados livre, livre office. Tenho essa preferencia porque, além de ser gratuita, elas são as melhores ferramentas, elas tem muita documentação na internet, se você tem alguma duvida tem uma base de usuarios muito grande para dar assistencia.

ANEXO C – Entrevista Mirna Tonus

primeiro você precisa ter uma... antes das diretrizes, que era muito generico, você tinha la o que você considerava importante para os estudantes e dividia nos semestres, carga horaria, enfim. Então esse planejamento é feito por um grupo de professores que pensam em disciplinas possíveis e ai fazem todos os calculos e ai pensam nas imentas das diciplinas, nos objetivos, nas metodologias, bibliografia, enfim. Então esse é um processo normal de você pensar um projeto pedagogico. Antes das diretrizes você não tinha orientõaçs para trabalhar com a formação especifica em jornalismo. Com as diretrizes você tem seis eixos nos quais você pode pensar mais especificamente as disciplinas. Não dispensando disciplinas de formação humanistica, naõa dispensando disciplina de formação em comunicação de modo geral por causa da ciencia da comunicação, mas com eixo que permitem em pensar essencialmente a formação profissional, as atividades laboratorias, então teve uma mudança significativa a partir das diretrizes.

Não esta consistente, pelo que eu avalio, pelo que eu vejo de projetos pedagogicos você jornalismo de dados hora como disciplina, hora como partes de disciplinas, no nosso caso aqui nos temos jornalismo digital e dentro do jornalismo digital temos o jornalismo de dados. O jornalimso de dados acaba sendo ensinado em tecnologias da comunicação. Nos temos ele tanto quanto em disciplina, como em conteúdo. Como não existe uma grade minima, ou um curriculo minimo, não existe isso. As universidades ficam livres para pensar como irão trabalhar. Agora eu acho essencial trabalhar com jornalismo de dados hoje. Alias hoje não, faz tempo. Não é agora

que os dados se tornaram importantes, demorou para o jornalismo pensar isso. Pensar nos dados como fonte de informações, isso já estava esboçado na pesquisa jornalística que é uma técnica importante, é uma área importante dentro do jornalismo. Está entre as técnicas de apuração você fazer uma pesquisa e a pesquisa pode ser de dados.

Jornalismo de dados sempre esteve bastante presente, isso as tecnologias de maneira geral e jornalismo de dados no projeto pedagógico. Porque foi se vendo a necessidade de os professores trabalharem com tecnologias digitais e dados. As discursos de projetos pedagógicos nos grupos de professores do fórum sempre envolveram essa discussão, desde que eu estou dentro do fórum.

Vejo estatística como algo importante, mas tem correntes que defendem o uso das palavras, principalmente os que são avessos a tecnologia. Quando você entra em contato com a tecnologia você como ela pode te auxiliar nos cálculos. Você não precisa ser um exímio em matemática para ser jornalista, mas se você tem conhecimento e domínio tecnologia, ou conhecimento básico de tecnologia, que vai te ajudar a lidar com números de uma maneira diferente, que vai te ajudar a fazer a matéria, a exercer o jornalismo

ANEXO D – Grade Curricular das Universidades

Universidade de São Paulo – USP. Disponível em:

<https://uspdigital.usp.br/jupiterweb/listarGradeCurricularcodcg=27&codcur=27011&codhab=402&tipo=N> Acesso em: 25 de out. 2016

Mackenzie. Disponível em: <http://up.mackenzie.br/graduacao/sao-paulo/jornalismo/matriz-curricular/> Acesso em: 25 de out. 2016

Universidade Federal Rio Grande do Sul – UFRGS.

http://www.ufrgs.br/ufrgs/ensino/graduacao/cursos/exibeCurso?cod_curso=313
Acesso em; 25 de out. 2016

Casper Libero. Disponível em:

<http://casperlibero.edu.br/graduacao/jornalismo/matriz-curricular-de-jo/> acesso em: 25 de out. de 2016

Universidade Paulista – UNIP.

http://www.unip.br/ensino/graduacao/tradicionais/hum_com_social_jornalismo_grade.aspx acesso em 25 de out. 2016

Universidade Federal de Uberlândia – UFU. Disponível em

<http://www2.faced.ufu.br/comsoc/grade.php?set=2> acesso em 25 de out. 2016

Metodista. Disponível em: <http://portal.metodista.br/jornalismo/modulos/modulos> acesso em 25 de out. 2016

Universidade Federal de Minas Gerais. Disponível em

<http://www.fafich.ufmg.br/colcom/wp-content/uploads/2016/04/HOR%C3%81RIO-COM.-SOCIAL-2016.1.pdf> acesso em 25 de out de 2016

CESJF. <http://www.cesjf.br/jornalismo-matriz-curricular.html> Acesso em 25 de out. 2016

UNASP. Disponível em:

<http://www.unasp.edu.br/NewGraduacao/ProjetoPedagogico/17> acesso em 25 de out. 2016

CESMAC. Disponível em: [http://www.cesmac.edu.br/admin/wp-](http://www.cesmac.edu.br/admin/wp-content/uploads/2014/12/matrizCurricularjornalismo-20131.pdf)

[content/uploads/2014/12/matrizCurricularjornalismo-20131.pdf](http://www.cesmac.edu.br/admin/wp-content/uploads/2014/12/matrizCurricularjornalismo-20131.pdf) acesso em 25 de out de 2016

UNICESUMAR. Disponível em: <https://www.unicesumar.edu.br/presencial/cursos-graduacao/jornalismo/> acesso em 25 de out. de 2016