



Associação Brasileira de Jornalismo Investigativo

I Seminário de Pesquisa em Jornalismo Investigativo
Universidade Anhembi-Morumbi, 24 a 26 de julho de 2014

Jornalismo Guiado por Dados e sua contribuição para a agenda pública no Brasil¹

Um estudo de caso sobre as
publicações online do Globo e do
Estadão

Fábio Vasconcellos²

Journalism Guided by Data and its Contribution to the public agenda in Brazil

A case study of online publications by
Globo and Estadão

¹ Trabalho apresentado no I Seminário de Pesquisa em Jornalismo Investigativo, realizado na Universidade Anhembi-Morumbi, cidade de São Paulo, entre 24 e 26 de julho de 2014

² Jornalista e professor da Escola Superior de Propaganda e Marketing/Rio (ESPM). Doutor em Ciência Política. E-mail: fabio.vasconcellos10@gmail.com. Twitter:@fvas

Resumo: Este estudo de caso analisa o conteúdo das publicações online dos jornais O Globo e Estado de São Paulo, com o objetivo de mapear a temática utilizada nas reportagens com o uso de dados. Sofisticadas técnicas de captura, análise e apresentação de dados qualificam a informação, aspecto importante no processo de *accountability* político das democracias. Este estudo contribui para o debate sobre a agenda pública no Brasil e os artificios que os jornais online utilizam para incentivar o engajamento dos leitores. Os resultados apresentados não devem ser generalizados, mas vistos como ponto de partida da discussão sobre o Jornalismo Guiado por Dados no país. Os achados demonstram que os assuntos mais abordados são política, desigualdade social, saúde, transporte e eleições. As reportagens focaram assuntos nacionais, e muitas publicações trazem recursos interativos que ampliam as possibilidades de compartilhamento das informações. Outro achado é a grande diferença do volume de publicações do Estadão quando comparado como o Globo, variação que pode ser explicada pela criação de um grupo específico de jornalistas no diário paulista para localizar bancos de dados públicos e transformá-los em notícia.

Palavras-Chave: Jornalismo Guiado por Dados. Informação. Agenda Pública.

Abstract: This case study analyzes the content of online material published by Brazilian newspapers O Globo and Estado de São Paulo with the aim to map the themes used in news with the use of data. Sophisticated techniques in coverage, analysis and presentation of data that qualifies information, an important aspect in the process of political accountability in democracies. This study contributes to debate concerning the public agenda in Brazil and the tools used by online journals to increase the engagement of its readership. The results presented must not be generalized but rather understood as a starting point in the discussion concerning Journalism Guided by Data in Brazil. The findings show that the most frequently approached topics are politics, social inequality, health, transportation and elections. The news focused on national subjects and many publications offered interactive resources that increased the possibility of sharing information. Another finding was the large difference in terms of the volume of publications by the Estado de São Paulo newspaper compared to Globo, a variation that can be explained by the creation of a specific group of journalists in the São Paulo daily with the purposed of locating data public bases and transforming them into news

Keywords: Journalism Guided by Data, Information, Public Agenda.

Democracia, informação e jornalismo

Desde os anos 90, o jornalismo vem passando por uma série de mudanças e a expansão da internet tem sido apontada como uma das principais razões. A forte migração do público jovem para o ambiente online, associada à queda do número de assinaturas das suas versões impressas, à redução de receita e ao fechamento de jornais e revistas, indicaria um futuro devastador para essa atividade. Em 2012, C.W. Anderson, Emily Bell e Clay Shirky, da Escola de Jornalismo de Columbia, publicaram o estudo “Post-Industrial Journalism: adapting to the present”³ no qual defendem que o modelo de negócio do jornalismo, que classificam como industrial, de fato não sobreviverá.

Apesar da afirmação, o argumento da Escola de Columbia está longe de ser pessimista. A partir da análise do que denominam de o “ecossistema do jornalismo”, os pesquisadores sustentam que um novo jornalismo está surgindo, com uma estrutura menor, mais dinâmico, mais acessível e tão ou mais importante para a sociedade como foi no período industrial. Em outras palavras, embora estejam em curso mudanças significativas de ordem tecnológica, com o seu devido impacto financeiro nas empresas jornalísticas tradicionais, Anderson, Bell e Shirky sustentam, quase em tom de aposta, que a internet deve ser vista como uma aliada, e não uma inimiga do jornalismo.

Para os pesquisadores que assinam “Post-Industrial”, o jornalismo ocupará de forma crescente o ambiente online, em razão das suas novas práticas, acesso ao leitor e novas formas de financiamento. Para realizá-lo, contudo, será necessário que os repórteres saibam utilizar de maneira eficiente os inúmeros recursos da rede para capturar, analisar e contextualizar as informações, bem como as possibilidades da internet para chegar aos leitores com notícias que de fato importam para as suas vidas e que promovam o seu engajamento.

A premissa geral dos autores de Columbia é a de que a prática do jornalismo sobreviverá não apenas em razão da nova tecnologia, mas porque essa atividade é importante para funcionamento das democracias. Está presente aí uma crença valiosa da contribuição do jornalismo para o processo de *accountability* político nas democracias representativas

³ Disponível em http://towcenter.org/wp-content/uploads/2012/11/TOWCenter-Post_Industrial_Journalism.pdf, acessado em 12 de março de 2014

(LATTMAN-WELTMAN 2003; MAIA 2006; MIGUEL 2014). Sem saber o que as instituições e o poder político fazem ou deixam de fazer, restringimos e muito a nossa capacidade de tomar decisões, como a de punir políticos que não atendem às nossas expectativas, ou a de premiar aqueles que avaliamos positivamente. Em outras palavras, o jornalismo é importante, segundo Columbia, porque “expõe a corrupção, chama a atenção para a injustiça, cobra políticos e empresas por promessas e obrigações assumidas. Informa cidadãos e consumidores, ajuda a organizar a opinião pública, explica temas complexos e esclarece divergências” (ANDERSON, BELL, SHIRKY, pág.3)⁴

Sem dúvida, a informação é um elemento central no processo de desenvolvimento e afirmação das democracias representativas, e existe um dado aqui nada desprezível. A maneira pela qual ela está à disposição do cidadão e como ele pode se apropriar das informações afetam a sua capacidade de decisão em assuntos públicos. Em linhas gerais, e até com certo exagero, poderíamos dizer que a qualidade e a diversidade da informação produzida pelo jornalismo interferem direta ou indiretamente na relação entre representante e representado, espinha mestra da democracia tal qual a conhecemos.

Há pouco mais de uma década, Lupia e McCubbins (1998) propuseram um debate no mundo da Ciência Política que indicava, contudo, um aparente problema nessa relação. Sem mencionar especificamente o jornalismo, até porque o foco da análise era o cidadão, os autores fizeram uma pergunta bem objetiva: “como pode haver boas democracias se os cidadãos não têm informação suficiente para tomar as melhores decisões?”. O chamado “Dilema da Democracia”, que acabou se transformando no título da obra de Lupia e McCubbins, refletia sobre algo bastante discutido entre cientistas políticos atuais, porque atacava o coração do argumento dos pensadores clássicos que ajudaram a fundar as democracias representativas entre os séculos XVII e XVIII. Para os clássicos, a democracia, para sobreviver, dependia da participação política de cidadãos bem informados. A questão

⁴ Para o propósito deste projeto de pesquisa, *accountability* político é compreendida aqui não apenas como uma forma de controle social, de sujeição do poder público a estruturas formais e institucionalizadas de constrangimento às suas ações à frente da gestão pública, mas como um atributo da sociedade civil como forma de fiscalizar, independentemente do poder público, não só em anos eleitorais, como também ao longo do mandato dos seus representantes. Nesse sentido, o jornalismo/imprensa é um dos agentes responsáveis pela produção e circulação de informações sobre os agentes públicos, fazendo funcionar, desse modo, o processo de *accountability* político.

formulada por Lupia e McCubbins indicava, contudo, outra direção. Aquilo que os clássicos pensavam seria apenas uma idealização. Na prática, o cidadão comum tem um reduzido número de informações, mas, curiosamente, isso não faz das democracias um sistema pior. A explicação dos autores era bastante simples: as democracias vão bem apesar da baixa quantidade de informações a que os indivíduos têm acesso porque estes elaboram um cálculo de atenção, isto é, recorrem a alguns estímulos informacionais que ajudam a reduzir os erros nas suas tomadas de decisão. Em outras palavras, quando percebem que a atenção dedicada a uma fonte de informação ajudou a evitar um erro, os cidadãos canalizam a sua atenção para essas fontes, reduzindo o esforço de recorrer a inúmeras outras fontes. Resumo: não é a quantidade de fontes que importa, mas o tipo de informação que cada cidadão avalia como positiva para a sua decisão.

Parte do estudo de Lupia e McCubbins, vale ressaltar, foi inspirado na pesquisa de Downs (1999), que problematizou a questão da seguinte forma: manter-se informado sobre a política numa democracia representa um alto custo cognitivo e de tempo para o cidadão comum, quase sempre envolvido com uma série de outras atribuições da vida privada. Nesse sentido, seria irracional a tentativa do cidadão comum acompanhar em detalhes o que se passa na política. Segundo Downs, as escolhas políticas dos indivíduos são baseadas em princípios ideológicos, e os partidos teriam um importante papel nesse processo porque são atalhos que ajudam a reduzir os custos de informação sobre a política. A razão para isso seria aparentemente simples: como estão em disputa eleitoral e estão interessados no poder político, os partidos procuram ajustar as suas propostas a princípios ideológicos que os ajudariam a se diferenciar dos demais concorrentes, transformando-se assim em atalhos informacionais importantes para os eleitores.

Com sua pesquisa, Downs influenciou outros autores a desenvolverem melhor o argumento, especialmente aqueles que viam nos partidos um atalho de importância decrescente em muitas democracias. Para esses novos autores, não só os partidos, mas, e principalmente, as campanhas eleitorais e a mídia, leia-se a imprensa, deveriam ser vistos como elementos centrais no processo de decisão dos cidadãos, porque oferecem atalhos com baixo custo compreensivo aos leitores (IYENGAR e KINDER, 1987; POPKIN, 1994; HOLBROOK, 1996). Simplificando: a premissa geral que está presente nesses novos estudos

defende que o cidadão vive uma situação de incerteza ao ter que tomar a sua decisão, já que costuma ser pouco informado sobre a política. Como tem aversão ao risco e normalmente não aceita apoiar quem ele não conhece, a informação passa a ser uma *commodity* preciosa oferecida pelas campanhas eleitorais e pela imprensa (BARTELS, 2006). É a partir do acesso a informações de baixo custo cognitivo que os cidadãos podem participar do processo de *accountability* político, isto é, com a compreensão sobre as responsabilidades dos atores públicos, e daí beneficia-los ou puni-los.

Temos aqui, portanto, um cenário muito interessante. 1) Um sistema político que pressupõe a existência da informação como requisito da participação dos representados, aliado ao reconhecimento de que o cidadão comum não tem condições de acompanhar o governo em toda a sua complexidade; 2) um modelo no qual a mídia é vista como um canal de acesso a informações com baixo custo cognitivo e, por fim, 3) as facilidades e riquezas do uso de computadores na apuração, apresentação e disseminação de informações sobre dados públicos, bem como das possibilidades de engajamento do cidadão através da internet. Em outras palavras, as novas formas de investigação e produção de conhecimento promovido pelo jornalismo no ambiente da web indicam o quão significativa pode ser a contribuição dessas novas práticas no processo de *accountability* político nas modernas democracias.

Democracia e Jornalismo Guiado por Dados

A hipótese elaborada pela Ciência Política abordado na seção anterior é um excelente ponto de partida para pensarmos o papel do jornalismo. Embora seja relativamente recente, a prática do chamado Jornalismo Guiado por Dados é parte de uma tradição iniciada na década de 60 pelo jornalista Philippe Meyer (1991). Ele foi um dos primeiros a adotar o conceito de “Jornalismo de Precisão”, que consistia no uso do computador associado a métodos da ciência social para produzir reportagens com menores chances de erro. O estudo de Meyer acabou por incentivar o surgimento nos anos 90 do termo Reportagem com Auxílio de Computador (RAC), ainda hoje utilizado nos fóruns especializados como o Investigative Reporters e Editors (IRE) e a Associação Brasileira de Jornalismo Investigativo (Abraji).

No fim dos anos 90, a campanha pela abertura de dados governamentais em vários países incentivou o jornalismo a desenvolver novas técnicas para capturar, analisar e publicar essas informações através de notícias baseadas em dados públicos sobre áreas de governo e políticas públicas. Basicamente, as técnicas consistem na produção, tratamento e cruzamento de grande quantidade de dados, permitindo uma maior eficiência na recuperação da informação, na apuração da reportagem a partir do conjunto de dados, na circulação em diferentes plataformas e na geração de visualizações e infografias (TRÄSEL 2013). Esses novos procedimentos ajudaram os jornalistas a aprimorarem o seu próprio conhecimento acerca da realidade social e política, reduzindo a dependência de fontes externas ao processo de produção e análise da informação. Tudo isso em um contexto no qual o cenário de escassez foi substituído pela fase da abundância da informação (FLEW et al 2012).

“Quando a informação era escassa, a maior parte de nossos esforços estava voltada a caçar e reunir dados. Agora que a informação é abundante, processá-la tornou-se mais importante. O processamento acontece em dois níveis: 1) análise para entender e estruturar um fluxo infinito de dados e 2) apresentação para fazer com que os dados mais importantes e relevantes cheguem ao consumidor. Como acontece na ciência, o jornalismo de dados revela seus métodos e apresenta seus resultados de uma forma que possam ser replicados”. (MEYER apud GRAY et al 2013)⁵

Em suma, o chamado Jornalismo Guiado por Dados passaria a ser um vertente do RAC, porque envolveria não apenas o uso de computadores, mas o conhecimento mínimo em estatística, sistemas computacionais e métodos das ciências sociais (FLEW et al 2012; HAMILTON and TUENER 2009; GRAY et al 2013; FLEW et al 2012; HACKETT 2013). Essa nova prática jornalística pode ser vista, portanto, como aquela capaz de produzir atalhos informativos ao cidadão comum a partir de bases de dados que ajudam a entender e a explicar a realidade política e social das democracias. Mas, como dito, não é qualquer prática. Trata-se do desenvolvimento de habilidades específicas de buscar e analisar dados, transformá-los em notícia e apresentá-los na web com o uso de recursos que podem ser fatores decisivos na

⁵ Declaração publicada In GRAY, J.; BOUNEGRU, L.; CHAMBERS, L. (org) Manual de Jornalismo de Dados. 2013. Disponível: <http://datajournalismhandbook.org/pt/>

disseminação do conhecimento produzido, bem como de redução do esforço cognitivo necessário que o cidadão deve empreender para compreender a mensagem.

A importância do Jornalismo Guiado por Dados como instrumento com potencial de qualificar o processo de *accountability* político, como este estudo defende, contudo, remete necessariamente à pergunta: que tipo conteúdo tem sido apresentado a partir do uso dessas técnicas? Essa questão nos leva à discussão sobre a relevância e a amplitude da informação produzida pelo Jornalismo Guiado por Dados. O debate sobre relevância, a nosso ver, exige necessariamente focar a análise na agenda predominante dessas reportagens. Pouco adiantaria uma defesa do papel do jornalismo no processo de *accountability* se a temática preferida pelos veículos forem temas alheios à construção de atitudes cívicas por parte dos leitores.

Com relação à amplitude, é importante saber se as reportagens adotam recursos online para despertar a atenção e o envolvimento dos leitores. Este ponto é central em razão do seguinte suposto. Os temas que favorecem a atitude cívica por parte dos cidadãos ampliam suas potencialidades quando as reportagens utilizam apresentações que ajudam a reduzir os custos cognitivos sobre os temas e, mais importantes ainda, quando apresentam canais por meio dos quais os leitores podem se engajar na disseminação dessas mensagens.

A nosso ver, portanto, a qualificação dos procedimentos do jornalismo nesse novo ecossistema pode contribuir para o processo de *accountability* político num duplo sentido. Por exigir um maior preparo dos jornalistas e das empresas nas quais trabalham, os novos procedimentos acabam por induzir uma maior independência no procedimento de apuração, fator de enorme valor para as democracias. Por outro lado, a utilização do ambiente online para a distribuição de reportagens com dados, com todas as técnicas de apresentação interativa, incentiva à participação e o envolvimento dos leitores em temas complexos que exigiriam, em plataformas tradicionais, um maior esforço para absorver o conteúdo. A disposição dos leitores em compartilhar as informações seria uma evidência direta desse envolvimento e das potencialidades do Jornalismo Guiado por Dados no ambiente web.

O cenário do Jornalismo Guiado por Dados no Brasil: prática e pesquisa

A cultura dos novos procedimentos de investigação jornalística e seu potencial de produzir narrativas acessíveis ao cidadão tem despertado a atenção de agências internacionais preocupadas com a qualidade da informação apresentada pela imprensa. Em 2013, a Unesco divulgou o documento curricular para os cursos de jornalismo no qual apresenta diretrizes gerais para o que considera uma boa formação nessa área. O tópico “Jornalismo de Dados” foi incluído no programa e classificado como um procedimento de alta “especialização no campo do jornalismo investigativo, embora as suas técnicas possam ser também muito bem aplicadas no jornalismo diário”. Também em 2013, foi traduzido para o português o livro “The Data Journalism Handbook” organizado Jonathan Gray, Lilian Bounegru e Lucy Chambers após um seminário realizado em Londres, em 2011. A obra apresenta o conceito de Jornalismo Guiado por Dados, além das técnicas e habilidades que os profissionais precisam adquirir para desenvolver suas investigações. Já existem versões traduzidas também para o Russo, Espanhol, Francês e Georgiano.

A constituição de fóruns especializados nessas técnicas é outro elemento que indica a relevância do tema. Desde 1989, a ONG Investigative Reporters e Editors (IRE) tem promovido cursos para jornalistas interessados em utilizar técnicas de investigação com o uso de computador. Em 2010, o European Journalism Centre (EJC) promoveu o primeiro congresso específico de Jornalismo Guiado por Dados, em Amsterdã. Outra importante instituição que tem trabalhado na disseminação dessas técnicas tem sido o Knight Center Journalism, baseado na University of Texas.

Os preceitos do jornalismo de precisão e, posteriormente, do Jornalismo Guiado por Dados, também chegaram ao Brasil. Desde a sua fundação em 2002, a Abraji tem promovido cursos de RAC, nos moldes indicados por Meyer nos anos 60 e, mais recentemente, tem patrocinado encontros e treinamentos sobre o Jornalismo Guiado por Dados. O mais recente deles, o 8ª *Global Investigative Journalism Conference*, realizado em 2013, na PUC Rio, reuniu mais de 4 mil profissionais, que participaram de cursos e discutiram temas vinculados não só à investigação jornalística, mas às técnicas e ferramentas que podem ser utilizadas em reportagem com o uso de dados.

Embora os pesquisadores de Columbia sustentem que as empresas menores terão maior facilidade de sobreviver no novo ecossistema do jornalismo, temos visto no Brasil iniciativas de alguns jornais tradicionais para investir nas novas ferramentas da web e no Jornalismo Guiado por Dados. Em maio de 2012, o jornal Estado de São Paulo lançou o EstadãoDados, um setor da redação dedicado à captura, análise e produção de visualizações interativas de reportagens feitas a partir de dados. Quando foi criado, o EstadãoDados contava com três jornalistas e um programador de web. No mesmo ano que começou a funcionar, o EstadãoDados produziu o “Basômetro”, uma ferramenta online que permite a qualquer pessoa consultar 480 mil votos dos deputados federais em 1.025 votações, observando como as bancadas se organizam e como os campos da situação e oposição se comportam em cada matéria discutida.

Embora o Estadão tenha sido pioneiro ao criar uma equipe dedicada exclusivamente ao Jornalismo Guiado por Dados e, mais do que isso, um canal para suas publicações diárias, outras experiências começaram a surgir em outros estados. Em junho de 2012, o jornal carioca O Globo publicou uma série de reportagens intitulada “Veneno em doses diárias”, na qual relatava a associação entre taxas de mortes por câncer e uso de agrotóxicos nos 92 municípios do estado. A reportagem foi realizada com a compilação e análise de bancos de dados do Sistema Único de Saúde (SUS), órgãos do governo estadual e do IBGE. Para a web, o jornal produziu um infográfico interativo, no qual os leitores podiam visualizar a associação entre produção agrícola com uso de agrotóxicos e taxas de mortes por câncer. Outras reportagens com dados foram produzidas em 2013 com análises sobre temas relativos à educação e a acidentes de trânsito, sempre buscando compreender relações entre variáveis. Desde então, o uso de dados na produção e apresentação de reportagens têm se repetido não só no Globo e Estadão, mas em outros jornais como a Folha de SP, Gazeta do Povo, no Paraná, e a revista Época.

O interesse crescente dos jornais brasileiros no Jornalismo Guiado por Dados segue a dinâmica do contexto internacional. Resumidamente, três mudanças foram fundamentais para o início desse processo globalmente, segundo Flew et al (2012): a primeira delas diz respeito ao maior acesso aos bancos de dados, que passaram a ficar disponíveis na internet com a cultura do *Open Data* e *Open Government*. Esses dois processos se interligam e têm sido

adotados por diversos países, facilitando o acesso à base de informações, principalmente de órgãos de governo, que antes ficavam restritas ou eram de difícil acesso à imprensa ou ao cidadão comum.

Com parte desse processo, o governo brasileiro sancionou em 2012 a Lei de Acesso à Informação (Lei 12.527). A nova legislação determina que os órgãos públicos, nos três níveis de governo, passem a divulgar seus dados e a atender aos pedidos de informação de qualquer cidadão. Novas leis foram criadas em nível estadual e municipal regulamentando o processo de pedido e entrega das informações solicitadas. O segundo aspecto da cultura do Jornalismo Guiado por Dados no mundo diz respeito à combinação dos recursos da web com a queda dos custos dos serviços online de produção gráfica e análise de dados. Como exemplo, podemos citar o Tableau Public, CartoDB, Google Drive, Infogr.am, DocumentCloud, OutHit Hub, Scraper Wiki, Open StreetMap entre outros. O terceiro aspecto, segundo Flew et al (2012), para a disseminação do Jornalismo Guiado por Dados é a existência dos recursos da internet e das redes sociais que permitem uma maior interação e engajamento dos leitores com a notícia. Esse modelo ampliou as possibilidades das reportagens, que passaram a circular em diversos outros canais e muito em função da participação dos leitores na disseminação do conhecimento produzido pelos jornais.

O crescente uso de dados nas reportagens publicadas no Brasil e a existência dos fóruns que passaram a discutir esses novos métodos levaram também pesquisadores brasileiros a se debruçarem sobre esse tema. Como ainda são muito recentes, os estudos sobre o Jornalismo Guiado por Dados ainda não foram capazes de constituir um conhecimento claro sobre o uso dessa técnica nas redações no país. Apesar disso, vale destacar os trabalhos de Barbosa (2007), Lima Júnior (2011) e Träsel (2013a; 2013b) que discutiram o impacto do modelo computacional no processo de investigação jornalística.

Em seu trabalho, Barbosa analisou cronologicamente o uso do modelo de dados a partir da análise de 22 casos brasileiros. Uma das melhores contribuições de Barbosa foi compreender historicamente as mudanças dos websites dos jornais com o crescente uso das técnicas digitais. O uso de banco de dados, segundo autora, representaria o início da quarta-geração do jornalismo na web, classificado como Jornalismo Digital em Bases de Dados (JDBD), que consistiria numa “fase de base tecnológica avançada, conexão banda larga,

proliferação de plataformas móveis, equipes especializadas, uso expandido de base de dados, algoritmos, linguagem de programação” (BARBOSA 2007, pág. 150). A autora trata do conceito de JDBJ no seu sentido mais amplo, incluindo a produção, conteúdo e distribuição do material jornalístico.

Dessa forma, as principais vantagens do modelo para Barbosa (2007) podem ser assim resumidas: 1) asseguram o padrão dinâmico para a oferta efetiva de conteúdos originais; 2) habilitam a organização das informações jornalísticas, permitindo a automação dos processos; 3) conferem maior agilidade para o tratamento das informações nas etapas de pré-produção, produção e pós-produção, assegurando rapidez para a recuperação e para o compartilhamento dos conteúdos; 4) favorecem a expansão de atuação das empresas, pois permitem a distribuição do conteúdo jornalístico para plataformas variadas; 5) propiciam a ampliação do eixo temático e a consequente diversidade temática.

Träsel (2013a), por sua vez, utiliza o método etnográfico para compreender como são realizadas as reportagens realizadas pela equipe do EstadãoDados. A principal preocupação do autor é saber como são organizados esses procedimentos e as diversas funções e papéis dos profissionais que fazem parte da equipe. Em outro trabalho, Träsel (2013b) defende que os procedimentos do Jornalismo Guiado por Dados poderiam ser caracterizados como parte de um intercâmbio com cultura hacker, compreendida como uma “tendência à apropriação de tecnologia, à valorização da liberdade de informação e a disposição para o trabalho colaborativo” (TRÄSEL 2013b, pág. 14).

Em seu trabalho, Lima Júnior (2011) faz uma análise do contexto das transformações tecnológicas, em outras palavras, ao crescente uso da Tecnologia de Informação e Comunicação (TIC) como um elemento fundamental do “Jornalismo Computacional”, como o autor prefere chamar. O potencial no uso das TIC num contexto de Big Data seria talvez a principal razão para uma mudança no tipo de habilidade dos futuros jornalistas. Lima Júnior defende também a ideia de que está em gestação um “hacking journalism”, cuja principal é explorar as tecnologias filtrando informações e colocando-as de forma visual. Nesse cenário, o argumento segue o de outros pesquisadores, segundo os quais, o contexto atual favorece também o uso das tecnologias para encontrar e contar histórias.

Como podemos notar, os trabalhos brasileiros apontados procuram, nesse momento, definir o que seria o Jornalismo Guiado por Dados e suas características, podendo variar de uma compreensão geral sobre o uso de dados pelas empresas jornalísticas ou o nível de aprofundamento no uso de ferramentas e procedimentos na apuração e apresentação das reportagens. O esforço dos pesquisadores passa, inclusive, pela própria discussão sobre o melhor termo a ser utilizado. Num caso, ele pode ser chamado de Jornalismo Guiado por Dados (TRÄSEL 2013a, 2013b), Jornalismo Computacional (LIMA JÚNIOR 2011) ou ainda Jornalismo Digital em Bases de Dados (JDBD) (BARBOSA 2007).

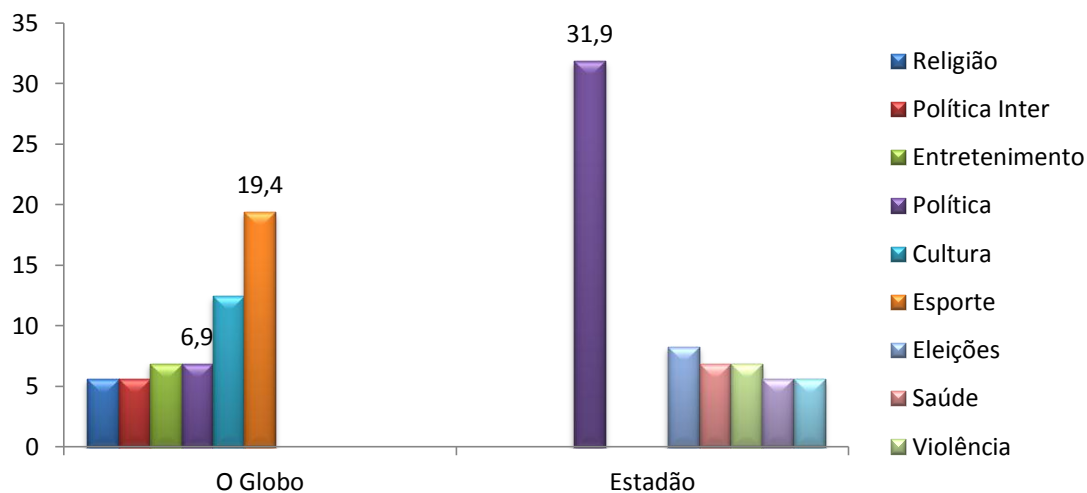
Embora louvável o trabalho dos pesquisados brasileiros, pouco se analisou o conteúdo das reportagens guiadas por dados no Brasil e, especialmente, a sua conexão com a questão da *accountability* político. Em outros termos, se partimos do suposto de que o processo de *Open Data* e *Open Government* são centrais no processo de incentivo à proliferação do Jornalismo Guiado por Dados, sem entrar aqui necessariamente no mérito de qual tipo de tecnologia deva ser utilizada, parece ser fundamental saber se esses procedimentos têm sido adotados pela imprensa e quais temáticas ela tem se ocupado.

Resultados

Neste estudo construímos uma amostra com 144 publicações do Estadão (72) e do Globo (72). As publicações foram colhidas no weblog do EstadãoDados e da seção Infografia do site do jornal O Globo, entre os dias 11 e 25 de março de 2014. O material foi colhido de forma regressa, isto é, da última publicação online até a primeira. A análise foi limitada a 144 publicações porque não havia mais conteúdo disponível no site do Globo, na seção Infografia. Desse modo, optou-se por manter uma mostra de 72 casos para os dois jornais.

Com relação aos temas de todas as 144 publicações, a distribuição nos dois jornais apresentou o seguinte resultado. No caso do Globo, os temas vinculados ao Esporte apresentaram maior frequência: 19,4%. Muitas das publicações eram especiais interativos, isto é, ambientes criados pela equipe de Infografia Digital. No caso do EstadãoDados, o tema mais frequente foi Política (31,9%).

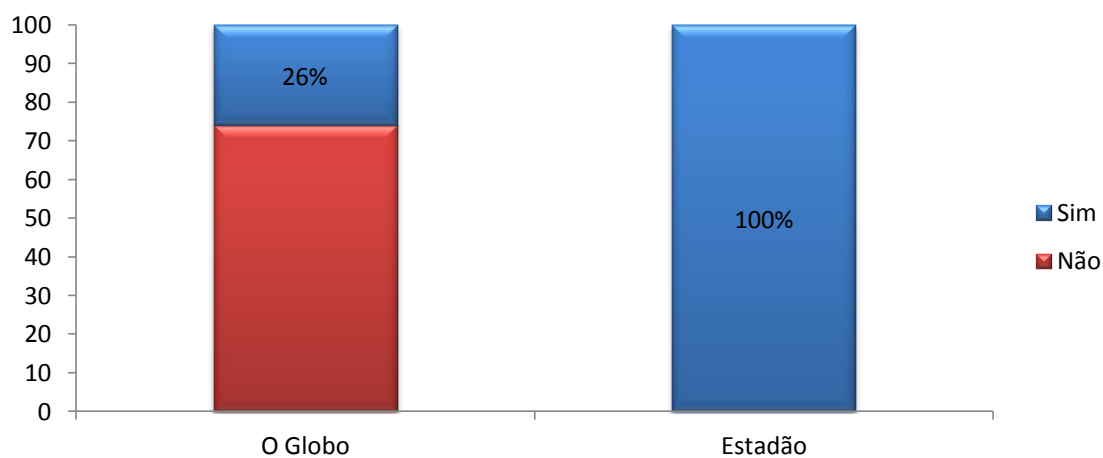
Gráfico 01: Percentual do conteúdo de todas as publicações com ou sem uso de dados



Nota: N=72 Globo; N=72 EstadãoDados

Do total de 144 publicações da amostra, 91 (63%) tinham alguma utilização de dados, distribuídos da seguinte forma: 100% (72 casos) das publicações do Estadão utilizavam dados, enquanto no caso do Globo, apenas 26,% (19 casos).

Gráfico 02: Percentual de uso de dados nas publicações analisadas



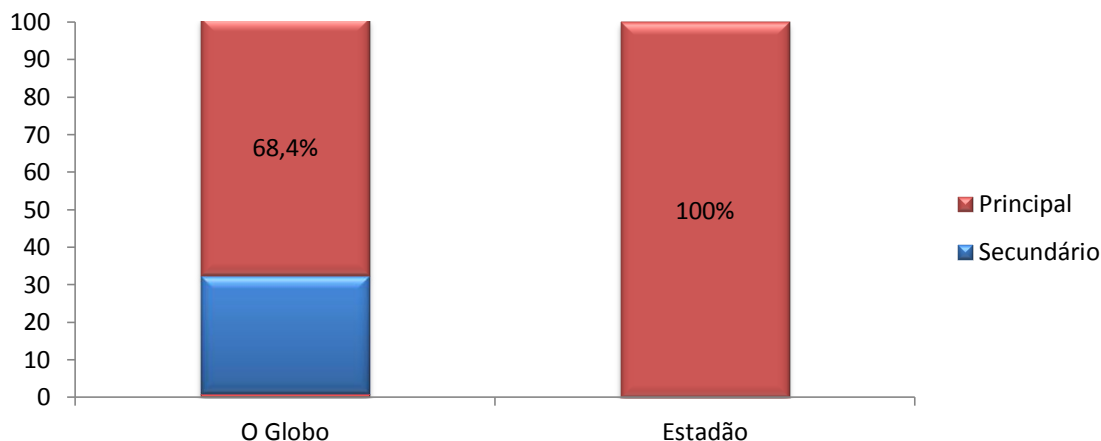
Nota: N=72 O Globo; N=72 EstadãoDados

Essa grande diferença entre Estadão e o Globo ocorre pelo tipo de estrutura das duas redações. O Estado de São Paulo mantém uma equipe permanente responsável pela captura e análise de dados sobre diversos temas. Todo o material é publicado no weblog EstadãoDados. No caso do Globo, não há uma equipe específica para a produção de bancos e análise. As publicações que utilizam dados dividem espaço com especiais interativos produzidos pela equipe de Infografia Digital do jornal. Nessa equipe há profissionais com conhecimento em estatística e programação, mas como precisam atender às demandas de toda a redação, isto é, na produção de ambientes digitais sobre inúmeros temas, o foco em dados se restringe.

Outra característica que afetou o volume de casos analisados do Globo. Há mais publicações com o uso de dados do que o registrado na seção Infografia do jornal carioca. Como relatado por profissionais da área, nem todo material produzido é indexado a esta seção, ficando sua veiculação restrita ao dia da publicação das reportagens e às páginas das respectivas editoriais. Como este estudo estava interessado especificamente nas reportagens com dados e não havia condições de identificar todo o material já publicado pelo jornal carioca, optou-se por analisar tudo que estava disponível apenas na seção Infografia do Globo, única forma de localizar e coletar o conteúdo armazenado. No caso do jornal paulista, como já dito, optou-se pela coleta de material do weblog EstadãoDados, cujo foco são as publicações com o uso de dados, tema central desta pesquisa.

Com relação apenas as publicações com dados, a análise revelou o seguinte. Em 100% dos casos, o EstadãoDados utilizou dados como elemento central, isto é, os dados eram o próprio foco da informação. Já o Globo, 68,4% das publicações tinham os dados como elemento central, enquanto outras 31,5%, como elemento secundário, isto é, os dados apareciam mais para compor a matéria do que como a própria informação central. A definição de foco central e secundário é bastante simples. Se a publicação informa um dado como resultado de uma análise ou compilação de um banco de dados feita pelos jornalistas, sua categoria é “foco central” porque o lide, no jargão jornalístico, é o próprio conhecimento produzido pela equipe. Por outro lado, se a publicação traz uma informação não centrada em dados, seu foco é secundário. Uma característica desse tipo é a utilização de um gráfico mais como ilustração, ou seja, o lide não está no próprio gráfico. Ele também não ocupa espaço hierárquico na página que possa ser entendido como a informação mais importante.

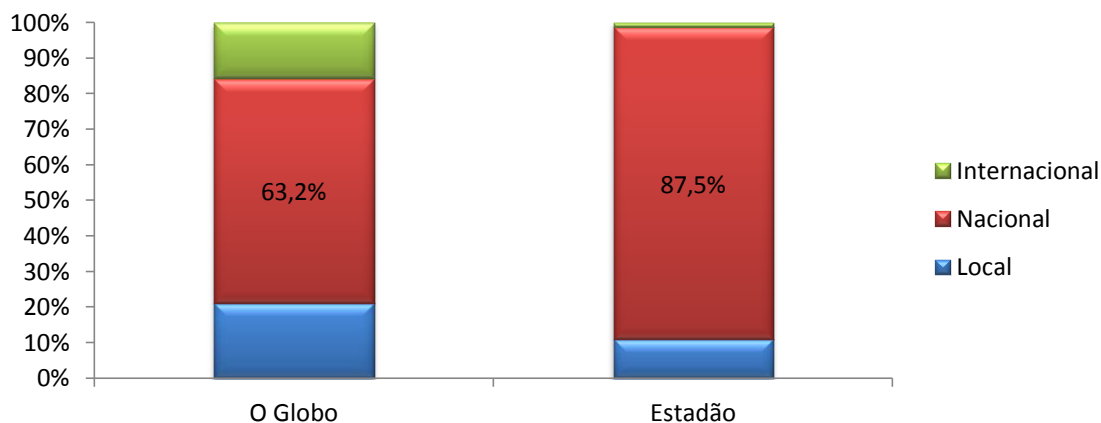
Gráfico 03: Foco do uso dos dados nas publicações



Nota: N=19 O Globo; N=72 EstadãoDados

O EstadãoDados como o Globo focam em assuntos nacionais, quando utilizam dados nas suas publicações. No caso do Globo, há uma maior distribuição entre temas internacionais, nacional e local.

Gráfico 04: Foco das publicações com dados

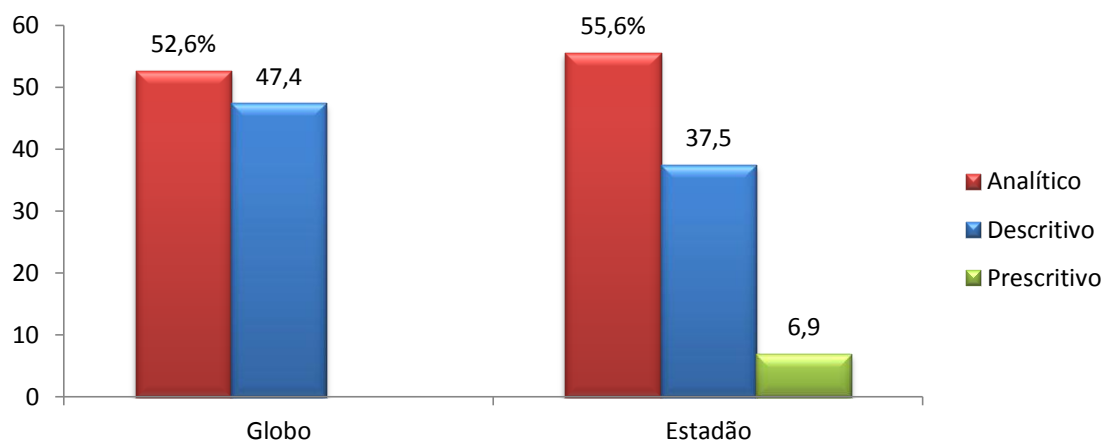


Nota: N=19 O Globo; N=72 EstadãoDados

Nos dois veículos, há uma tendência em apresentar os dados de forma analítica, isto é, com o cruzamento de mais de uma variável (numérica ou categórica), procurando encontrar

padrões ou relações entre essas variáveis; em outros casos, os dados são apresentados apenas de maneira descritiva, sem maior aprofundamento. No caso do EstadãoDados, há, contudo, o uso de dados prescritivos, ou seja, que procuram de alguma maneira informar ao leitor a probabilidade de um evento acontecer. Um exemplo claro desse tipo de abordagem foi a publicação do dia 4 de setembro de 2013 sobre a probabilidade de alguém ser preso após cometer um crime na cidade de São Paulo.

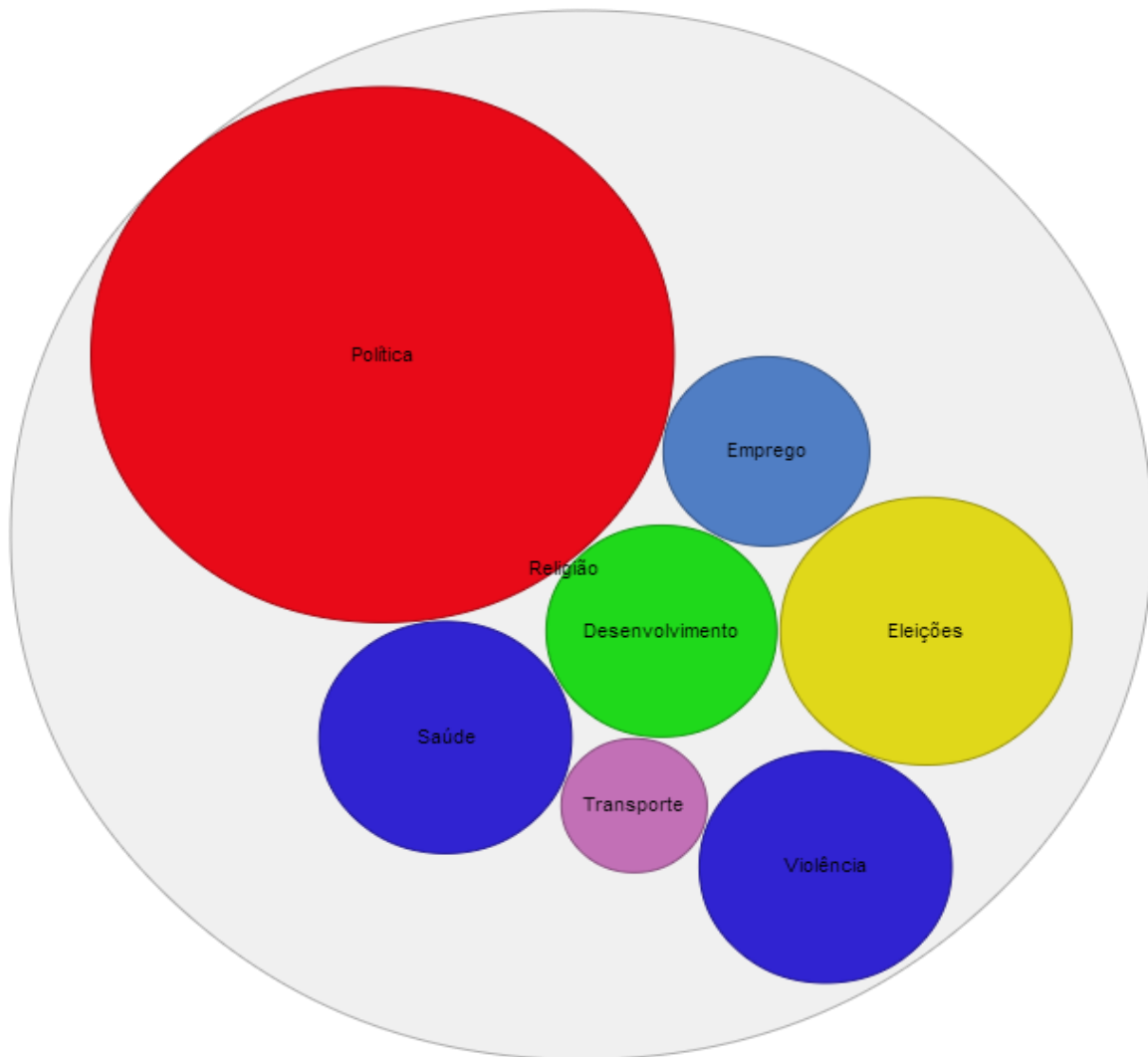
Gráfico 05: Tipo de apresentação dos dados



Nota: N=19 O Globo; N=72 EstadãoDados

Entre os temas mais publicados pelo EstadãoDados com o uso de dados estão Política (31,9%), seguido de Saúde (6,9%), Eleições (6,9%), Violência (5,8%) e Desigualdade Social (5,8%). No caso do Globo, os temas mais abordados foram Esporte (21,4%) e Religião (14,2%) seguido de Política (7,1%) e Saúde (7,1%).

Gráfico 06: Distribuição dos temas das publicações com uso de dados EstadãoDados



Nota: N=72 EstadãoDados

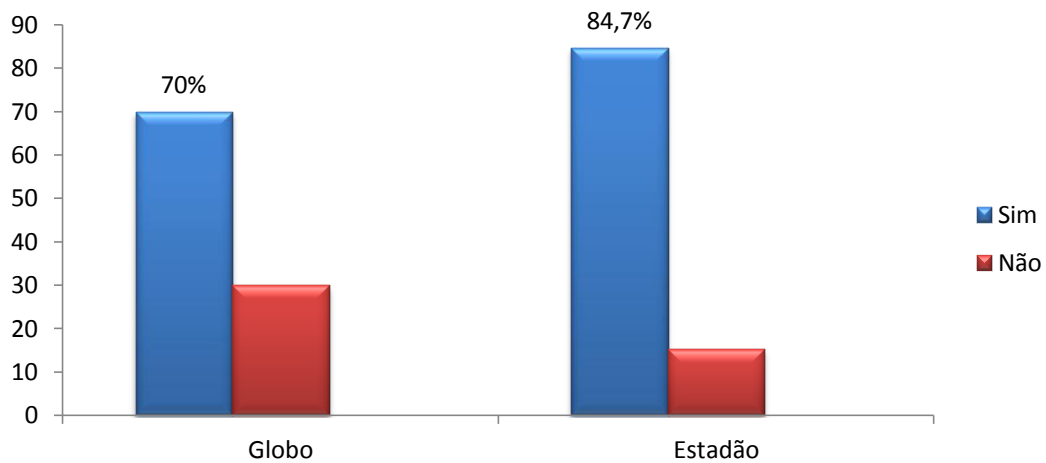
Gráfico 07: Distribuição dos temas das publicações com uso de dados Globo



Nota: N=19 O Globo

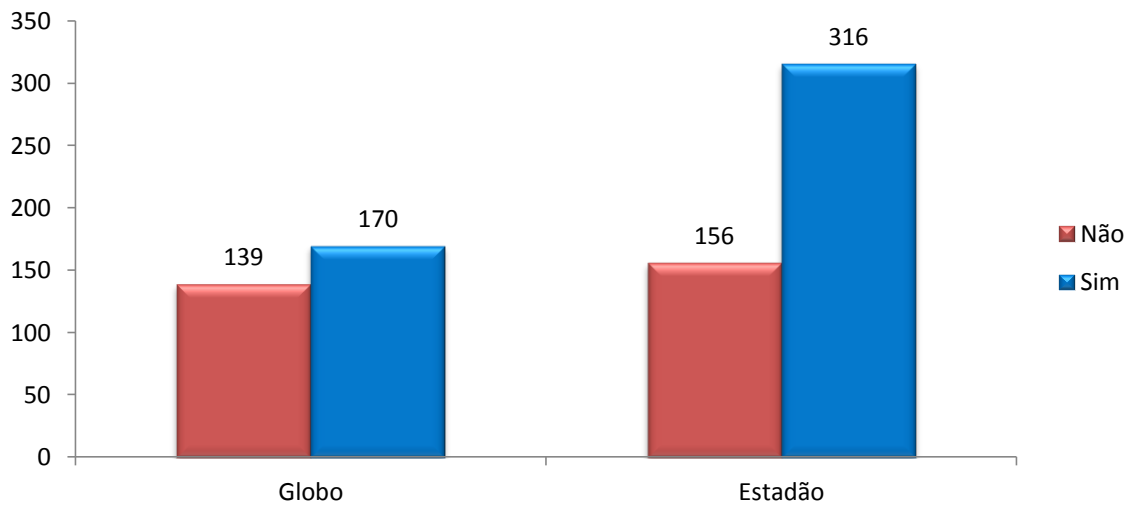
A maioria das publicações dos dois jornais com o uso de dados apresentou ao menos um recurso interativo, que permitia ao leitor manipular a organização dos dados ou mesmo identificar, com o uso do mouse, os percentuais e categorias dos gráficos. Como podemos observar no Gráfico 09, uso de recursos interativos aumentou em 100% os compartilhamentos das publicações do Estadão, enquanto no Globo, 22%.

Gráfico 08: Percentual de publicações com o uso de recursos interativos



N=19 Globo/ N=72 Estadão

Gráfico 09: Uso de recursos interativos x média do número de compartilhamentos das publicações

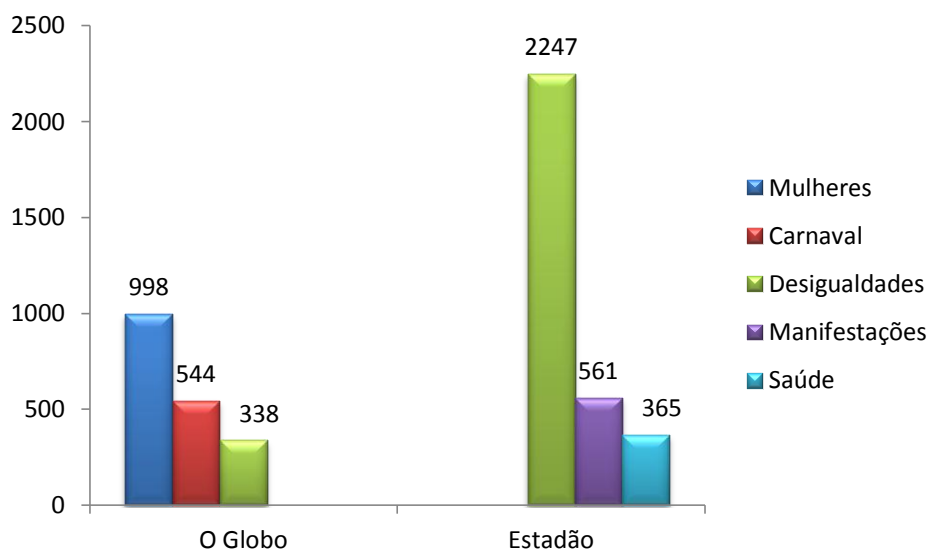


Nota: N=19 O Globo; N=72 EstadãoDados

Com relação aos conteúdos mais compartilhados nas publicações dos dois veículos, podemos dizer que aqueles com o uso de dados ainda estão longe de alcançar os índices obtidos por publicações online sem dados. Contudo, a média das publicações com o uso de dados apresenta números animadores. Embora o tempo de publicação interfira no volume de compartilhamentos, é possível ter uma compreensão aproximada de como os leitores se envolvem com os temas que utilizaram dados nos dois jornais. Não foi possível chegar a um índice de compartilhamentos médio por tempo de publicação do conteúdo porque na seção do Globo não havia a informação de quando o material foi veiculado.

De qualquer modo, no caso do Globo, a reportagem sobre mulheres presas grávidas, que ganhou um ambiente especial na internet e utilizou dados de maneira secundária, obteve 998 compartilhamentos, maior número registrado pelo jornal. O conteúdo com o segundo maior número de compartilhamentos foi o especial sobre o carnaval (544), seguido das publicações que trataram sobre a desigualdade no Brasil (338). No caso do Estadão, a publicação com o maior volume de compartilhamentos foi o mapa interativo sobre os índices de desigualdade do país Brasil (2.247), seguido de publicações sobre as manifestações no país (561) e temas relativos à saúde (365).

Gráfico 07: Média do número de compartilhamentos dos temas das publicações



Nota: N=19 O Globo; N=72 EstadãoDados

Conclusões

Como se trata de um estudo preliminar com o uso de uma amostra reduzida, é impossível fazer generalizações sobre as publicações feitas a partir do Jornalismo Guiado de Dados praticado no Brasil. Contudo, há indicações interessantes nesta análise. Procuramos observar sistematicamente as publicações do Globo e do Estadão, com o objetivo de mapear os temas e a maneira pela qual os dois veículos que, como sabemos, ocupam um espaço privilegiado na agenda pública do país, utilizam o Jornalismo Guiado por Dados. Como podemos ver, os temas escolhidos apresentam relevância do ponto de vista da qualificação do processo de *accountability* político, com destaque para o Estadão. Sem dúvida, a disposição dos dois veículos em preparar suas equipes para lidar com banco de dados e, posteriormente, aplicar essas técnicas para gerar informação a partir de temas que podem ser relevantes para agenda pública, é um ponto positivo que merece ser mencionado.

Entretanto, a diferença de modelo de organização da produção adotada pelos dois jornais impactou no volume de publicações com o uso de dados. Enquanto a equipe de Infografia do Globo atende a demandas de todo o jornal, reduzindo o espaço para temas exclusivamente vinculado a dados, a equipe do Estadão pode se dedicar apenas a essas publicações. Essa diferença aparece claramente na proporção de publicações com o uso de dados nos dois jornais, bem como na maior variedade de temas adotados pelo Globo quando comparado com o Estadão. Em resumo, a maneira como os veículos se organizam para a sua produção diária impacta diretamente o volume e o tipo das suas publicações com dados que ficam disponíveis aos leitores nas respectivas seções online dos dois veículos.

Outra conclusão igualmente importante foi com relação ao foco das publicações. Como vimos, elas são majoritariamente sobre temas nacionais, o que indica a grande necessidade de os dois veículos atentarem para assuntos locais, que afetam mais diretamente o dia a dia dos seus moradores, e que podem ser centrais para o exercício da cidadania política nessas duas cidades. Outro achado foi com relação ao tipo de abordagem dos dados. Na maioria das vezes, os dois veículos apostaram numa abordagem mais analítica, ou seja, numa clara tentativa em encontrar padrões e relações entre as variáveis. Embora demonstre o esforço de produzir conhecimento com os dados, fica evidente a necessidade de se investir

mais em publicações de natureza prescritiva, aquelas que também podem auxiliar os leitores nas suas tomadas de decisão.

Por fim, outra conclusão deste estudo é que indica o quão importante é a utilização das ferramentas da web pelas equipes que realizam Jornalismo Guiado por Dados. Como foi possível notar, a utilização de recursos interativos predomina nas publicações dos dois jornais. Ficou claro que o seu uso amplia a chance de a publicação ser compartilhada pelos leitores. Ou seja, o engajamento dos leitores aumenta à medida que as publicações utilizam esses recursos, efeito que possibilita que outros cidadãos tenham acesso à informação pelas redes sociais e e-mail. Efeito que, como defendemos, é de suma importância para disseminação do conhecimento gerado pelo jornalismo, com todas as suas vantagens de ser um canal de baixo custo cognitivo aos cidadãos sobre temas relevantes da agenda pública.

Referências

- ANDERSON, C.W.; BELL, E.; SHIRKY, C. **Post-Industrial Journalism: Adapting to the present**. Columbia Journalism School, 2013. Disponível: http://towcenter.org/wp-content/uploads/2012/11/TOWCenter-Post_Industrial_Journalism.pdf
- BARBOSA, S. **Jornalismo Digital em Base de Dados (JDBD): um paradigma para produtos jornalísticos digitais dinâmicos**. Tese de doutorado. Universidade Federal da Bahia, 2007.
- BARTELS, L. Priming and Persuasion in Presidential Campaigns. In: BRADY, H.; JOHNSTON, R. **Capturing Campaigns Effects**. [S.l.]: The University of Michigan Press, 2006. p. 78-112.
- COHEN, S.; HAMILTON, J. e TURNER, F. Computational journalism: How computer scientists can empower journalists, democracy's watchdogs, in the production of the news in the public interest. **Communication of the ACM**, vol. 54, n.10, p.66-71, 2011.
- DOWNS, A. **Uma teoria econômica da democracia**. São Paulo: EdUSP, 1999
- FLEW, T.; SPURGEON, C.; DANIEL, A. e SWIFT, A. The promise of computational journalism. **Journalism Practice**, vol. 6, n. 2, p. 157-171, 2012.

- GRAY, J.; BOUNEGRU, L.; CHAMBERS, L. (org) **Manual de Jornalismo de Dados. 2013.** Disponível: <http://datajournalismhandbook.org/pt/>
- HACKETT, K. **Journalism data age.** Quill Magazine. Setembro p. 35-38, 2013.
- HAMILTON, JAMES T. and TURNER, FRED. **Accountability Through Algorithm: developing the field of computational journalism.** Report from Center for Advanced Study in the Behavioural Sciences, julho 27-31, 2009.
- HOLBROOK, T. **Do campaigns matter?** Thousand Oaks: Sage, 1996.
- IYENGAR, S.; R., K. D. **News that matters: Television and American opinion.** Chicago: University Of Chicago Press, 1987
- JÚNIOR LIMA, W. **Jornalismo Computacional em função da Era do Big Data.** Paper 9º Encontro Nacional de Pesquisadores em Jornalismo. Rio de Janeiro, novembro, 2011.
- LORENZ. M. Por que jornalistas devem usar jornalismo de dados? In: **Manual de Jornalismo de Dados.** GRAY, J.; BOUNEGRU, L.; CHAMBERS, L. (org), 2013
- LATTMAN-WELTMAN, F. Mídia e transição democrática: a (des) institucionalização do pan-óptico no Brasil. In **Mídia e política no Brasil.** ABREU, A., LATTMAN-WELTMAN, F. e KORNIS, M. (org) FGV, Rio de Janeiro, pp 129-183, 2013.
- LUPIA, A.; MCCUBBINS, M. D. **The Democratic Dilemma: Can Citizens Learn What They Need To Know?** USA: Cambridge University Press, 1998.
- MEYER, P. **The new precision journalism.** Indiana University Press, 1991.
- MIGUEL. F. L. **Democracia e representação: territórios em disputa.** Unesp, São Paulo, 2014.
- POPKIN, S. **The Reasoning Voter: Communication and Persuasion in Presidential Campaign.** Chicago: The University of Chicago Press, 1994
- TRÄSEL, M. (a) **Jornalismo Guiado por Dados: relações da cultura hacker com a cultura jornalística.** Paper XXII Encontro Anual da Compós. Universidade Federal da Bahia de 4 a 5 de junho de 2013, p. 1-16
- _____. (b) **Learning to be Guided by Data: The Training of Journalists Team Estadão Data.** Papel Global Investigative Journalism Conference, Pontifícia Universidade Católica do Rio de Janeiro, outubro, 2013.