



**Como fizemos isso: o caso da transparência do Núcleo
Jornalismo na cobertura de plataforma**
**How we did it: the case of Núcleo Jornalismo's transparency
on the platform beat**

Alexandro Mota

Palavras-chave: Jornalismo, Transparência, Platform Beat, Plataformização

As cobranças para que grandes empresas de tecnologia, as *Big Techs*, tornem seus processos e negócios mais transparentes têm vindo de toda parte, parlamentos, governos, sociedade civil e imprensa. No Brasil, a transparência é um pilar do poder público, que tem sido provocado para regular e reduzir a opacidade que em parte caracteriza o funcionamento das plataformas (D'ANDRÉA, 2020; POELL *et al.*, 2020; VAN DIJCK *et al.* 2018).

Esta comunicação discute a intersecção dos temas transparência, plataformas e jornalismo. Parte-se da necessidade de uma transparência de mão dupla: a busca por maior regulamentação e transparência das plataformas, que por sua vez vai permitir um jornalismo de maior precisão e acompanhamento público, pode também ser seguido de uma maior transparência do próprio jornalismo.

O objetivo da presente pesquisa de caráter exploratório é compreender como tem se dado a transparência na cobertura cotidiana das plataformas a partir do estudo de caso da seção “Como fizemos isso” das reportagens do Núcleo Jornalismo.



1. *Platform Beat* e transparência

O jornalismo que se concentra na investigação do funcionamento das organizações de plataformas e motores de buscas e seus impactos na sociedade tem sido chamado de *Platform Beat*. Os jornalistas cuja especialização é a plataformização denunciam aspectos irregulares do funcionamento das plataformas, dos usos ilegais ou contra as regras das comunidades das redes sociais, dos impactos e manobras dessas empresas na vida socioeconômica, além de suas relações com a própria mídia, com a curadoria de informações e com o cenário de desinformação.

Reportagens desta área podem se debruçar sobre a calibragem de algoritmos para fins específicos; como os dados dos usuários são protegidos ou negociados; como funcionam os sistemas de publicidade, inclusive em períodos eleitorais; as conversações de impacto social nas mídias sociais, como disseminação e grupos de discurso de ódio, além da relação com a própria mídia em termos de distribuição e remuneração de notícias. Um dos principais veículos especializados em plataforma é o The Markup, com sede em Nova York e com foco em jornalismo investigativo, a organização sem fins lucrativos tem como slogan “Big Tech Is Watching You. We’re Watching Big Tech”.

O caráter ainda tentativo da cobertura das plataformas e limitações impostas pelas características dessas empresas exigem uma maior transparência do jornalismo. Para Napoli (2021), os jornalistas são desafiados nessa cobertura pela necessidade de domínios técnicos muito específicos, a inacessibilidade dos dados e pela “caixa preta” das plataformas. “Essas complexidades técnicas associadas às plataformas digitais podem potencialmente minar o rigor com que as organizações de notícias cobrem a *platform beat*”. (Napoli, 2021, p. 9, tradução nossa). Assim, essa limitação operacional pode induzir a erros ou gerar vieses excessivos, o que, argumenta-se aqui, exige deixar o público ciente. Outra complexidade da *Platform Beat* que precisa ser contornado



Anais de Resumos Expandidos

V Seminário Internacional de Pesquisas em Midiatização e Processos Sociais

ISSN 2675-4169

Vol. 1, N. 5 (2022)

com transparência é a frequente dependência financeira do setor de notícias dessas plataformas.

A transparência nos processos é uma forma de reivindicação do jornalismo por autoridade e enfrentamento da crise de credibilidade, é um elemento caracterizador do jornalismo de dados e de verificação. Guerra (2014) argumenta que a transparência eleva a confiabilidade das notícias e se constitui como um dos elementos possíveis de análise e enquadramento de qualidade do jornalismo. A transparência também é a marca de uma produção de jornalismo independente, como apontado por Patrício e Batista (2020). Assim, a transparência é um valor do jornalismo contemporâneo (GEHRKE, 2018). Do ponto de vista organizacional, a tese de Marília Gehrke (2021), ao propor um Protocolo de Transparência no Jornalismo Guiado por Dados, sugere como elementos de observação os documentos, fontes de dados, Graus de incerteza, marcas de apuração, metodologia, reprodutibilidade e código aberto.

2. Metodologia

A presente pesquisa tem como matriz metodológica os Estudos de Caso, ideal para entender de modo amplo um fenômeno em desenvolvimento. A seleção das reportagens foi feita através do Google – o buscador interno do site não retornou resultados para o nome da seção de interesse. Foram usados os seguintes operadores e termos de pesquisa: "como fizemos isso" site:nucleo.jor.br. A busca se deu no dia 29 de junho de 2022, com a conta do Google do pesquisador logada. Foram selecionados os 19 resultados das duas primeiras páginas de busca, sendo ignorada a sugestão de imagens relacionadas (por isso 19, não 20). A observação se deu com técnicas inspiradas pelo o que Yin (2001) prevê como uma análise guiada pela descrição dos casos. Definiu-se uma matriz de categorias que foi inspirada nos elementos do Protocolo de Transparência no Jornalismo Guiado por Dados (GEHRKE, 2021), sem se ater necessariamente a todas as suas variáveis. Quanto ao *corpus*, uma matéria foi de 2020, 10 de 2021 e oito deste ano. Foram identificados 13 autores no período e quatro



Anais de Resumos Expandidos

V Seminário Internacional de Pesquisas em Midiatização e Processos Sociais

ISSN 2675-4169

Vol. 1, N. 5 (2022)

editores – mais da metade (55,6%) das edições foram assinadas pelo próprio Spagnuolo. 84,2% das unidades de análise foram arquivadas em “Reportagens”, 10,5% em “Interativos” e 5,3% em “Jogo Rápido” – avalia-se aqui como baixa a representatividade no corpus da seção Interativos, geralmente guiada por dados. As *tags* das reportagens apontam para a plataforma a que a reportagem se refere. O Facebook e seus produtos compreendem mais da metade da amostra e o Twitter tem o segundo maior número de menções, com 13% da amostra. 63,2% da amostra foi guiada por dados, 31,6% por uma apuração convencional e uma reportagem (5,3%) baseou-se na análise de releases. A tabulação dos dados e suas categorias estão disponíveis em bit.ly/comofizemosisso.

3. Como fizemos isso: primeiros olhares

O Núcleo foi fundado pelos jornalistas Sérgio Spagnuolo e Alexandre Orrico. Parte da equipe inicialmente atuava para a *startup* Vortex Media, antes do seu fechamento. Incubada pela agência Volt Data Lab, a iniciativa foi lançada em 6 de janeiro de 2020. A empresa já recebeu investimento do Google News Initiative e do Instituto Serrapilheira, aceita doação e lançou em julho de 2022 seu programa de membros. O Núcleo tem uma seção dedicada a transparência da organização em nucleo.jor.br/transparencia/ e hoje se posiciona como “uma iniciativa que cobre o impacto das redes sociais nas vidas das pessoas”.

A seção “Como fizemos isso” fica ao final das reportagens e consolida e resume o processo de apuração das informações, por vezes completando dados anteriormente apresentados ao longo do texto. Os textos do Núcleo são diretos, a média de palavras é de 1.300 palavras e a seção em análise representa, também em média, 8,8% do texto total (mínimo de 1,6% e máximo de 26,3%) ou uma média de 108 palavras. Os textos não se apegam a terminologias técnicas, mas parece remeter a um leitor com algum conhecimento em análise de dados ao não detalhar do que se trata, por exemplo, API, busca por palavra-chave, Coeficiente de Gini etc., mas tem particular preocupação em



Anais de Resumos Expandidos

V Seminário Internacional de Pesquisas em Midiatização e Processos Sociais

ISSN 2675-4169

Vol. 1, N. 5 (2022)

apresentar possíveis conflitos de interesse, como a menção de propriedade da ferramenta CrowdTangle (Meta).

Os textos da seção têm particular preocupação em detalhar os métodos de coleta dos dados, apontar as fontes e por vezes mencionar os termos de busca, mas é menor a preocupação no detalhamento sobre como se deu a análise em si, quais possíveis ferramentas estatísticas ou tecnológicas foram empregadas, quantos e quem foram os codificadores ou quais os critérios de exclusão de itens da amostra – 22,4% menciona algum critério de inclusão ou de exclusão ou de ambos de dados. Mesmo sendo produções recentes (menos de dois anos), muitos dos *links* oferecidos para os bancos de dados estão indisponíveis. Visto que a transparência se liga também com aspectos da memória e arquivamento, essa falha não permite futuras recuperações e mesmo replicação das análises. Alguns dos arquivos apontam para os serviços do Google, o que sugere dependência dessas plataformas na armazenagem de dados.

Em relação a replicabilidade, outro aspecto a se observar é o oferecimento de banco de dados. 44% das matérias não apresentam o banco de dados utilizado nas matérias e em 16,7% dos casos o banco de dado está indisponível, há uma linkagem parcial do banco de dados ou casos em que o que é oferecido é um *link* com a base geral dos dados e não os que foram propriamente utilizados. Quando os dados foram apresentados, em 44,4% das vezes tratou-se de uma planilha, 33,3% um arquivo PDF e 22,2% um link para a base geral. Em apenas três reportagens há menção direta à coleta por API e em quatro matérias da amostra é disponibilizado algum código utilizado na coleta/análise – em um dos casos o arquivo está indisponível. 73,3% das matérias não cita qualquer ferramenta utilizada no processo. A mais citada é o CrowdTangle.

Em geral, os dados são apresentados sem que se aponte possíveis vieses da coleta, das plataformas em si ou mesmo que se apresente limitação da análise. Há quatro casos em que as limitações são mencionadas, em um deles admite-se a possibilidade de falso negativo; um texto menciona que tuítes podem também ser produzidos por robôs; em outro há uma indicação de que os termos de uso do CrowdTangle impedem divulgar



Anais de Resumos Expandidos

V Seminário Internacional de Pesquisas em Midiatização e Processos Sociais

ISSN 2675-4169

Vol. 1, N. 5 (2022)

os dados integralmente; e em um quarto e último caso a reportagem é corrigida a seguinte mensagem: “Texto atualizado às 16h34 de 23.nov.2021 para refletir que trata-se de coleta dados (reclamações) feitas espontaneamente por usuários e não necessariamente de uma pesquisa”. Apenas uma matéria, sobre a relação da mídia com as assessorias de imprensa das plataformas, indica que há alguma dificuldade explícita com as empresas (elas oferecem apenas respostas genéricas para os questionamentos).

4. Considerações finais

O esforço de transparência do Núcleo Jornalismo é um passo importante e um exemplo para as demais redações. Mesmo que limitado pelo espaço e pelas condições de produções habituais do jornalismo cotidiano – grande parte da amostra é de reportagens não aprofundadas ou especiais, a dedicação de uma seção para detalhar as produções têm um caráter público de compromisso com o leitor. Como apresentado, não há um padrão ou elementos mínimos que precisem ser apresentados nessa seção pelos repórteres, bem como a oferta de códigos e banco de dados está em menos da metade das matérias analisadas, embora fortemente encorajada pela política editorial da empresa. Como frequentemente observado no jornalismo de dados, o trabalho do Núcleo sofre de certo fetichismo dos grandes volumes de dados, com frequentes valorizações na linha de apoio de menções a milhões ou milhares de publicações analisadas. Isso pode apontar o porquê do foco em uma maior transparência na coleta de dados, mas não explica certa negligência na apresentação de ferramentas e técnicas de análises.

Neste estudo, entendeu-se os contornos mais amplos de como se dá a transparência no processo da cobertura de plataforma feito pelo Núcleo Jornalismo através de um estudo inicial, exploratório e cuja limitação evidente ainda é ter uma amostra pequena como única fonte de evidência. Como mencionado, é um primeiro esforço de compreensão do tema e da dinâmica deste veículo, um dos primeiros do país a se posicionar estrategicamente no que compreendemos como *Platform Beat*.



Referências

- D'ANDRÉA, C. **Pesquisando plataformas online: conceitos e métodos**. 1. ed. Salvador: Edufba, 2020.
- GEHRKE, M. **Transparência no método como valor para o jornalismo**. Anais do 16º SBPJOR. **Anais...**São Paulo: 2018.
- GEHRKE, M. **Os elementos de transparência no jornalismo guiado por dados**. Universidade Federal do Rio Grande do Sul, 2021.
- GUERRA, J. Transparência editorial: a credibilidade jornalística à luz dos sistemas de gestão da qualidade. **Revista Latinoamericana de Ciencias de la Comunicación**, v. 11, n. 20, p. 196–209, 2014.
- NAPOLI, P. M. The platform beat: Algorithmic watchdogs in the disinformation age. **European Journal of Communication**, v. 36, n. 4, p. 376–390, 2021.
- PATRÍCIO, E.; BATISTA, R. Elementos de identidade em iniciativas de jornalismo independente. **Revista Extraprensa**, v. 13, n. 2, p. 217–231, 2020.
- POELL, T.; NIEBORG, D.; DIJCK, J. VAN. Plataformização. **Fronteiras - Estudos Midiáticos**, v. 22, p. 2–10, 2020.
- VAN DIJCK, J.; POELL, T.; WAAL, M. DE. **The Platform Society: Public Values in a Connective World**. 1ª ed. New York: Oxford, 2018.