

Consumo de notícias e informações políticas no Brasil: Mapeamento do primeiro turno das eleições presidenciais brasileiras de 2018 no Twitter

MEMORANDO DE DADOS [COMPROP](#) 2018.4 / 4 DE OUTUBRO DE 2018
comprop@oii.ox.ac.uk

Caio Machado
Oxford University
[@caiocvm](#)

Beatriz Kira
Oxford University
[@beakira](#)

Gustavo Hirsch
Oxford University
[@gustavoh1985](#)

Nahema Marchal
Oxford University
[@nahema_marchal](#)

Bence Kollanyi
Oxford University
[@bencekollanyi](#)

Philip N. Howard
Oxford University
[@pnhoward](#)

Thomas Lederer
Graphika

Vlad Barash
Graphika
[@vlad43210](#)

RESUMO

No Brasil, existe uma preocupação crescente sobre o uso de propaganda computacional e a polarização política que ela pode causar. Neste memorando, analisamos dados sobre notícias e informações políticas compartilhadas no Twitter no período que antecedeu a eleição presidencial brasileira de 2018. Descobrimos que: (1) no Brasil, o debate político nas redes sociais é altamente partidário, e nas conversas no Twitter há visível dominância do candidato à frente nas pesquisas Jair Bolsonaro; (2) contas associadas às hashtags de Luiz Inácio Lula da Silva e Fernando Haddad apresentam as mais altas frequências de tweets; (3) os usuários do Twitter brasileiros estão compartilhando mais conteúdo político profissional nessa rede social do que junk news - a maior proporção em todas as eleições que estudamos; (4) ao passo que os partidários de Bolsonaro foram responsáveis por compartilhar a variedade mais ampla de fontes de junk news, os partidários de Lula da Silva e Haddad foram responsáveis pelo maior volume de compartilhamento.

INTRODUÇÃO

Os brasileiros estão entre os usuários de redes sociais e aplicativos de mensagens mais ávidos do mundo, e as redes sociais têm se tornado uma plataforma fundamental para o compartilhamento de notícias e informações políticas pelos cidadãos. Como em outras partes do mundo, o tipo de notícias e informações aos quais os eleitores estão expostos nas redes sociais varia consideravelmente, compreendendo desde conteúdo produzido por fontes profissionais de notícias até conteúdo altamente polarizador e com forte apelo emocional. Fontes responsáveis pela produção e pelo compartilhamento de informações enganosas ou inexatas muitas vezes imitam fontes de notícias reconhecidas, e atores de todos os lados do espectro político potencializam o uso de informações falsas para capturar a atenção. Em momentos de maior interesse público, os algoritmos das redes sociais podem promover e disseminar material conspiratório mais do que informações verdadeiras.¹ Isso traz preocupações relativas à manipulação da opinião pública, especialmente em momentos politicamente sensíveis, como eleições, referendos e plebiscitos.²

No Brasil, há um crescente debate sobre o impacto da desinformação na polarização e na disseminação da violência, tanto online quanto offline. Pela primeira vez na história eleitoral do país, o Brasil testemunhou a invasão de grupos políticos populares do Facebook e a perseguição, ameaça e violência voltados contra líderes políticos, ativistas e celebridades nas redes sociais.³

De acordo com o Relatório Mídia Digital da Reuters de 2018, as plataformas online continuam sendo a principal fonte de notícias nos centros urbanos brasileiros, com 66% dos entrevistados relatando que usam as redes sociais para buscar notícias, incluindo 14% que dizem usar o Twitter. O Centro de Estudos sobre as Tecnologias da Informação e da Comunicação (Cetic.br) relata que 50% da população usa a Internet para ler notícias e que 68% dos usuários de Internet compartilharam algum tipo de conteúdo online em 2016.⁵ O WhatsApp também se tornou uma fonte importante e relevante de notícias. De acordo com dados fornecidos pela empresa, com mais de 120 milhões de usuários ativos, o Brasil é responsável por 10% da base de usuários do mundo.⁶

No contexto político altamente polarizado do Brasil, analisamos fontes de notícias e informações compartilhadas nas redes sociais por um período de 10 dias que antecedeu o primeiro turno da eleição presidencial brasileira de 2018. Nossas perguntas de pesquisa são: (1) Quais candidatos lideraram as conversas no Twitter e quais estavam associados a tweets de alta frequência? (2) Quais tipos de conteúdo estão sendo compartilhados no Twitter? (3) Como grupos de públicos diferentes compartilharam junk news no Twitter?

AS ELEIÇÕES GERAIS BRASILEIRAS E A MÍDIA

O primeiro turno das eleições gerais brasileiras ocorrerá em 7 de outubro de 2018. A população brasileira vai às urnas eleger presidente, governadores

e membros do Congresso Nacional, Assembleias Estaduais e Câmaras Municipais. Os cargos de presidente, vice-presidente e governador são preenchidos por meio de um sistema de dois turnos. Se nenhum candidato conseguir mais de 50% dos votos válidos no primeiro turno, um segundo turno com os dois candidatos mais votados ocorrerá no dia 28 de outubro de 2018.

O Brasil tem um sistema representativo presidencial proporcional, no qual o cenário partidário é composto por múltiplos partidos organizados em coalizões amplas. O presidente é o chefe de Estado e é eleito para um mandato de quatro anos, com a possibilidade de reeleição para um segundo mandato consecutivo. Durante 13 anos seguidos, de 2003 a 2016, o Brasil foi governado pelo Partido dos Trabalhadores (PT), primeiro com Luiz Inácio Lula da Silva por dois mandatos, depois por Dilma Rousseff, por um mandato e meio. O atual presidente é Michel Temer, do Movimento Democrático Brasileiro (MDB), que substituiu Dilma Rousseff depois de seu controverso processo de impeachment em 2016.

Em março de 2018, o ex-presidente Lula da Silva, que até então liderava as pesquisas de intenção de votos para presidente, foi condenado à prisão. No início de setembro de 2018, ele foi impedido de concorrer às eleições pelo Tribunal Superior Eleitoral do Brasil (TSE), com base na Lei da Ficha Limpa, a qual impede pessoas com condenações criminais de concorrerem a cargos políticos eletivos. Lula foi forçado a abandonar a corrida eleitoral e foi substituído pelo candidato a vice-presidente do PT Fernando Haddad. Há agora 13 candidatos inscritos na disputa presidencial, o segundo maior número de candidatos concorrendo ao cargo na história democrática do Brasil, desde o fim do regime militar em 1985.

Em 6 de setembro de 2018, o candidato Jair Bolsonaro do Partido Social Liberal (PSL) foi esfaqueado no abdômen durante um comício eleitoral. O golpe quase fatal deu grande visibilidade a Bolsonaro e provocou ampla circulação nas redes sociais de teorias conspiratórias que alegavam que o ataque havia sido arquitetado por pela esquerda.^{7,8} Apesar da agitação na mídia, as intenções de voto a Bolsonaro não aumentaram significativamente logo após o evento, e sua taxa de aprovação de fato diminuiu naquele contexto.⁹ Ao passo que Bolsonaro segue liderando as pesquisas de intenção de voto, a popularidade de seu rival Fernando Haddad cresceu significativamente desde sua indicação oficial, apesar de pesquisas mais recentes apontarem certa estagnação do segundo colocado a poucos dias do primeiro turno.¹⁰

PROPAGANDA COMPUTACIONAL E JUNK NEWS NO BRASIL

Historicamente, no Brasil, o sucesso nas urnas anda de mãos dadas com a transmissão na mídia. Durante

o período oficial da campanha eleitoral, os canais abertos de televisão e rádio são obrigados a ceder aos partidos e candidatos um horário diário para a propaganda eleitoral gratuita. O tempo de televisão é alocado com base no número de representantes eleitos de cada partido no Congresso Nacional e é usado como moeda de troca na formação coalizões e alianças eleitorais. Nessas eleições, como resultado desses acordos, há enormes discrepâncias no tempo concedido para cada candidato: enquanto Geraldo Alckmin, do Partido da Social Democracia Brasileira (PSDB), desfruta de cinco minutos e 32 segundos para sua campanha, o candidato líder nas pesquisas Jair Bolsonaro recebeu apenas oito segundos de tempo total.

A eleição deste ano marcou um aumento nas táticas de campanha online, o que oferece novas oportunidades para analisar como as ferramentas digitais podem influenciar o eleitorado. De acordo com uma pesquisa recente, 56% dos eleitores brasileiros dizem que as redes sociais influenciariam a sua escolha de candidato à presidência.^{11, 12} Especialistas brasileiros acreditam que a eleição geral de 2018 será a mais digital da história do país e que o uso da Internet e das redes sociais será crucial para o êxito dos candidatos e partidos políticos. No ano passado, foi aprovado um projeto de lei que autoriza o impulsionamento de conteúdo político na Internet, e, neste ano, o TSE publicou resolução que detalha as regras para propaganda política paga nas redes sociais. Pela primeira vez, candidatos poderão pagar para impulsionar conteúdos, mas o uso de *bots* e perfis falsos para aumentar visibilidade é estritamente proibido. Nessa mesma resolução, o TSE também estabeleceu que a liberdade de expressão não protege a circulação deliberada de informações ou afirmações falsas ou manifestações feitas apenas com a intenção de prejudicar a imagem de alguém.

Não obstante, há notícias de envolvimento de políticos e partidos em atividades que visam manipular a opinião pública através das redes sociais, seja com a ajuda de sua própria equipe de campanha ou por meio de empresas contratadas para realizar campanhas na Internet. A operação e a produção de conteúdo automatizado são práticas difundidas no Brasil em todo o espectro político, e estudos sugerem que grupos com interesses diferentes têm usado conteúdo gerado automaticamente para influenciar as discussões no Twitter e beneficiar determinados candidatos.^{13,14}

Por essa razão, diferentes órgãos governamentais brasileiros estão buscando maneiras de identificar, acompanhar e punir a disseminação deliberada de notícias falsas e o uso de outros mecanismos que podem influenciar indevidamente a forma pela qual os cidadãos recebem informações no período que antecede as eleições gerais de 2018. O TSE organizou um seminário internacional sobre junk news em junho deste ano, que resultou em uma série de resoluções sobre campanhas eleitorais na Internet.

Ao mesmo tempo, mais de 20 projetos sobre notícias falsas estão sendo avaliados pelo Congresso Nacional.^{15,16}

MÉTODOS E AMOSTRAGEM

O banco de dados do Twitter contém 1.432.000 tweets postados por 204.097 usuários do Twitter únicos e que foram coletados entre 19 de agosto e 28 de agosto de 2018, usando uma combinação de hashtags de partidos políticos relevantes, hashtags específicas relacionadas às eleições e contas dos partidos e candidatos. Dos 13 candidatos registrados para concorrer à presidência, apenas nove pertencem a partidos com mais de cinco representantes no Congresso Nacional, número que é o limiar adotado pela Lei Eleitoral Brasileira para decidir quais candidatos devem participar dos debates na TV. Também adotamos essa regra para selecionar candidatos para a coleta de dados do Twitter. No caso do PT, considerando a alta probabilidade de que a candidatura de Luiz Inácio Lula da Silva fosse cassada (o que de fato aconteceu), coletamos dados relacionados tanto a Lula como a Haddad. A lista de hashtags associadas à eleição brasileira foi compilada por uma equipe de três codificadores treinados, que são falantes nativos de português brasileiro e bastante familiarizados com a política brasileira. Antes de iniciar a coleta de dados, o conjunto de hashtags foi refinado em um período de teste, que revelou as hashtags usadas com mais frequência, para que a lista fosse revisada adequadamente.

A API (*application program interface*) do Twitter foi usada na modalidade *Streaming* para coletar tweets publicamente disponíveis. O método preciso de amostragem da plataforma não é divulgado, no entanto, o Twitter informa que os dados disponíveis por meio da API Streaming são, no máximo, 1% do tráfego público global total no Twitter a qualquer momento. Os tweets foram coletados se eles: (1) contivessem pelo menos uma das hashtags relevantes ou pelo menos uma das contas do Twitter dos partidos políticos ou políticos; (2) contivessem a hashtag na URL compartilhada ou o título da página da Web; (3) fossem um retweet de uma mensagem que contivesse uma hashtag relevante ou uma menção na mensagem original; ou (4) fossem a citação de um tweet referindo-se a um tweet com uma hashtag ou menção relevante.

A lista de hashtags associadas à eleição brasileira foi aperfeiçoada por meio de um processo iterativo do tipo ‘bola de neve’, o que incluiu a coleta inicial de dados ao longo de cinco dias de testes para identificação das hashtags mais utilizadas relacionadas às eleições. Nós rastreamos hashtags que eram a favor e contra candidatos e seus partidos. Cada tweet foi contabilizado se contivesse pelo menos uma das hashtags relevantes. Se a mesma hashtag fosse usada várias vezes em um tweet, ela era contada apenas uma vez. Se um tweet contivesse mais de uma das hashtags rastreadas, ele era creditado em cada

grupo de hashtag relevantes (ver *Tabela 1*). Durante a análise do tráfego do Twitter relacionado ao partido, cada tweet era contado uma vez se continha pelo menos uma das hashtags ou das menções associadas a um partido político. Se o mesmo tweet continha hashtags ou menções a partidos diferentes, era incluído na lista de cada um dos partidos relevantes. Se um tweet incluía mais de uma hashtag ou menção relevante para o mesmo partido, ainda era contabilizado apenas uma vez por grupo.

O banco de dados final contém links para fontes de notícias compartilhadas cinco vezes ou mais no Twitter e inclui também links para conteúdo no YouTube e no Facebook. Links apontando para o próprio Twitter foram excluídos da nossa amostra. Essa abordagem resultou em 97,6% de cobertura, o que significa que a equipe codificou 97,6% de todas as URLs compartilhadas. O processo de classificar as URLs, contas, canais e páginas de origem, com base na avaliação das fontes, foi feito de acordo com um processo de codificação rigoroso e iterativo, usando uma tipologia desenvolvida e refinada por meio de estudos prévios de seis eleições, em cinco democracias ocidentais e vários países da América Latina.^{17,18}

Para medir a confiabilidade entre os codificadores, calculamos o alpha de Krippendorff, que era 0,84. A literatura existente indica que esse resultado proporciona um alto nível de confiabilidade.¹⁹ Em seguida, rastreamos como essas URLs eram compartilhadas pelo Twitter. Usamos o pacote de visualização da Graphika para mapear contas que seguiam aquelas associadas a fontes de junk news conhecidas. A visualização de dados de redes sociais é uma ferramenta poderosa que ajuda a entender como as pessoas compartilham informações e se associam umas com as outras. Usando palavras-chave selecionadas, contas iniciais e links conhecidos para conteúdos específicos, é possível construir grandes visualizações de rede que podem ser examinadas para encontrar comunidades de contas ou “grupos” que compartilham tipos de conteúdo muito semelhantes entre si. Em nossa análise, calculamos as pontuações de cobertura e consistência para cada grupo identificado. A *Cobertura* reflete o número de domínios de propaganda postados por determinado grupo sobre o número total de domínios da nossa lista de junk news identificadas. A *Consistência* de um grupo se refere à percentagem total de links sobre o total de domínios de propaganda identificados em nossa lista de fontes de junk news compartilhada pelo grupo. Um valor alto para a cobertura mostra que o grupo está compartilhando uma ampla gama de propaganda, enquanto um alto valor para a consistência mostra que o grupo está desempenhando um papel fundamental na disseminação de tal propaganda. As pontuações de cobertura e consistência foram calculadas a partir do número de links compartilhados entre os grupos e as fontes de

junk news. A tipologia que explica nossa classificação de conteúdo é a seguinte:

Conteúdo de fontes profissionais de notícias

- Grandes veículos de notícias. Notícias políticas e informações publicadas pelos principais jornais, emissoras de TV ou rádio, bem como agências de notícias.
- Noticiários locais. Conteúdo proveniente de jornais locais ou regionais, ou afiliadas de grandes veículos de notícias.
- Mídia nova e startups. Conteúdo proveniente de mídias novas e publicações digitais nativas, novas marcas, e startups.
- Tabloides. Reportagens com foco em sexo, crime, astrologia e celebridades, incluindo publicações da imprensa marrom.

Conteúdo político profissional

- Governo. Links para sites de órgãos e agências do governo.
- Especialistas. Conteúdo retirado de notas técnicas, relatórios ou trabalhos acadêmicos publicados por pesquisadores baseados em universidades, *think tanks* ou outras instituições de pesquisa.
- Partido político ou candidato. Links para conteúdo oficial produzido por partidos políticos ou campanhas de candidatos, bem como comitês dos partidos políticos.

Conteúdo polarizante ou conspiratório

- Notícias e informações falsas (junk news). Fontes que deliberadamente publicam informações enganosas, equivocadas ou incorretas que visa se passar por notícias reais sobre política, economia ou cultura. Esse conteúdo inclui várias formas de propaganda, notícias e informações ideologicamente extremas, hiperpartidárias ou conspiratórias. Para ser classificado como junk news, a fonte deve se enquadrar em pelo menos três dos cinco critérios listados abaixo:

- *Profissionalismo*: Fontes que não fazem uso de padrões e boas práticas de jornalismo profissional. Elas se abstêm de fornecer informações claras sobre autores, editores, e proprietários reais. Faltam a elas transparência e *accountability*, e não publicam correções ou retificações de informações.
- *Estilo*: Fontes que usam linguagem emocionalmente carregada, com expressões emotivas, hipérbolos, ataques *ad hominem*, manchetes enganosas, uso excessivo de maiúsculas, generalizações arriscadas e falácias, uso de imagens comoventes e muitas fotos e memes mobilizadores.
- *Credibilidade*: Fontes que se baseiam em informações falsas e teorias conspiratórias, que muitas vezes empregadas estrategicamente. Elas noticiam informações sem consultar fontes diversas e não realizam a verificação dos fatos. As fontes utilizadas e os padrões de produção geralmente não são confiáveis.
- *Viés*: Textos dessas fontes são altamente enviesados, ideologicamente distorcidos ou hiperpartidários e as notícias frequentemente incluem comentários fortemente opinativos.
- *Falsificação*: Essas fontes imitam notícias de fontes reconhecidas. Elas falsificam fontes, marcas e padrões de estilo. Comentários e notícias falsas são estilisticamente disfarçados de notícias, com referências a agências de notícias e fontes confiáveis, além de manchetes escritas em

um tom de notícia com carimbos de data, hora e local.

- Rússia. Conteúdo produzido por fontes de notícias e informações reconhecidamente russas.

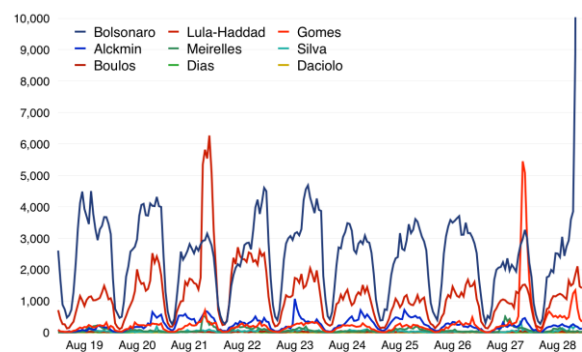
Outras notícias políticas e informações

- Blogs de comentário político. Blogs políticos que empregam padrões de produção de conteúdo profissional, como edição de texto, além de empregarem escritores e equipe editorial. Esses blogs geralmente se concentram em comentários de notícias, em vez da reportagem neutra de ciclos de notícias, e são frequentemente opinativos ou partidários.
- Cidadãos, sociedade civil, e conteúdo cívico. Links para conteúdo produzido por cidadãos independentes, grupos cívicos, organizações da sociedade civil, observatórios, agências verificadoras de fatos, grupos de interesse e grupos de pressão representando interesses ou agendas políticas específicas. Isso inclui blogs e sites dedicados ao jornalismo cidadão, ativismo pessoal e outras formas de expressão cívica que exibem originalidade e criação que vão além da curadoria ou agregação. Esta categoria inclui o Medium, o Blogger e o WordPress, a menos que uma fonte específica hospedada em qualquer uma dessas páginas possa ser identificada.

ACHADOS DE PESQUISA E ANÁLISE

Para nossa análise dos dados do Twitter, examinamos o volume de tweets, o grau de tweets de alta frequência e os tipos de conteúdo de notícias compartilhados no Twitter durante as eleições presidenciais brasileiras.

Figura 1: Conversação por Hora no Twitter Sobre os Candidatos a Presidência do Brasil com Base no Uso de Hashtags



Fonte: Cálculos dos autores a partir de amostra coletada entre 19/08/18 — 28/08/18. Nota: Ver suplemento online para a lista completa de hashtags.

A Figura 1 mostra a conversação por hora no Twitter sobre a eleição com base no uso de hashtag. A Tabela 1 mostra a conversação sobre as eleições por candidato no Twitter com base no uso da hashtag de cada candidato. Também identificamos os níveis de tuitagem de alta frequência de hashtags pertencentes a candidatos específicos. Para medir isso, estabelecemos o critério de 50 ou mais tweets com essas hashtags em um período de 24 horas. Como os tweets geralmente contêm várias hashtags, há alguma sobreposição entre os grupos de candidatos, portanto, o número total de tweets na Tabela 1

(1.428.771) não representa o número total de tweets exclusivos em todo o conjunto de dados (1.432.000).

Tabela 1: Conversação no Twitter e Tuitagem de Alta Frequência sobre as Eleições Brasileiras

<i>Partido político ou tráfego geral</i>	<i>N</i>	<i>%</i>	<i>N de tweets de alta frequência</i>	<i>% de tweets de alta frequência</i>
Geral	61,187	4	2,020	3
Jair Bolsonaro	645,950	45	23,893	32
Lula & Haddad	482,472	34	35,354	47
Ciro Gomes	67,549	5	8,490	11
Geraldo Alckmin	61,532	4	1,885	2
Marina Silva	52,150	4	1,560	2
Guilherme Boulos	33,742	2	496	1
Henrique Meirelles	17,000	1	1,691	2
Alvaro Dias	2,686	0	6	0
Cabo Daciolo	4,503	0	44	0
Total	1,428,771	100	75,439	100

Fonte: Cálculo dos autores a partir de amostra coletada entre 19/08/18 — 28/08/18. Nota: Ver complemento online para a lista completa de hashtags. As percentagens foram arredondadas para o valor inteiro mais próximo, a não ser que sejam menores que um percento, caso em que foram arredondadas para uma casa decimal. Tweets de alta frequência se refere ao número de tweets de contas que tuitam com alta frequência.

Jair Bolsonaro, o primeiro colocado nas pesquisas, dominou a conversa no Twitter, respondendo por 56% do tráfego total baseado em hashtag, enquanto Lula e Fernando Haddad representaram a maior percentagem de tweets de alta frequência (45%) (Ver *Figura 1* e *Tabela 1*). Bolsonaro teve 34 pontos percentuais a mais de conteúdo diário no Twitter do que Lula ou Haddad, e cerca de 10 vezes mais do que *Ciro Gomes* e *Geraldo Alckmin*. Os tweets de alta frequência também foram mais concentrados, com 90% da atividade girando em torno de três candidatas: *Lula & Haddad*, *Bolsonaro* e *Ciro Gomes*. Em 21 e 27 de agosto de 2018, observamos um aumento na atividade no Twitter associada a hashtags relacionadas a *Ciro Gomes* (*Figura 1*), o que pode ser atribuído principalmente às suas longas entrevistas em canais de TV. O pico é ainda maior em 28 de agosto, quando *Jair Bolsonaro* foi entrevistado no telejornal mais assistido do Brasil e o Supremo Tribunal Federal (STF) começou a análise de um caso de racismo que o envolvia. Naquela noite, o tráfego por hora para *Bolsonaro* atingiu um pico de 31.700 tweets por um curto período de uma hora.

Em seguida, analisamos os tipos de conteúdo de notícia compartilhados durante os 10 dias de nossa coleta de dados. Nossos resultados são apresentados na *Tabela 2*.

Tabela 2: Tipos de Notícias e Informações Compartilhadas no Twitter

<i>Tipo de Fonte</i>	<i>N</i>	<i>%</i>
Fontes profissionais de notícias		
Veículos profissionais	13,956	20.4
Tablóides	65	0.1
Subtotal	14,021	20.5
Conteúdo Político Profissional		
Partido Político ou Candidato	18,772	27.4
Estatais Pro-Governo	1,885	2.8
Governo	357	0.5
Especialistas	30	0.0
Subtotal	21,044	30.7
Conteúdo polarizante ou conspiratório		
Junk News	781	1.1
Conteúdo Russo	25	0.0
Subtotal	806	1.2
Outras Notícias e Informações Políticas		
Vídeo/Imagem	9,583	14.0
Cidadão ou Sociedade Civil	3,098	4.5
Blogs Políticos	1,751	2.6
Petições e Arrecadação	1,814	2.7
Categorias Remanescentes*	2,695	3.9
Subtotal	18,941	27.7
Outros		
Plataformas de Redes Sociais	10,641	15.6
Categorias Remanescentes**	2,956	4.3
Subtotal	13,597	19.9
Total	68,409	100

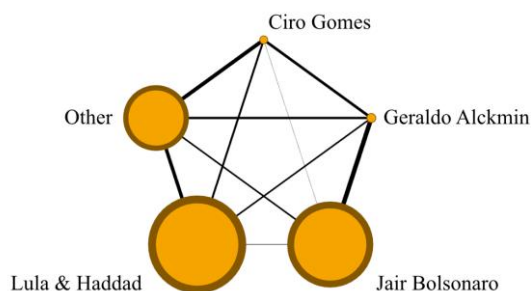
Fonte: Cálculos dos autores a partir de amostra coletada entre 19/08/18 — 28/08/18. Ver complemento online para a lista completa de hashtags. Nota: Grandes veículos de notícias, noticiários locais e mídia nova e startups foram reunidos na categoria Veículos Profissionais para esta tabela. Na categoria Outras Notícias e Informações Políticas, o grupo Categorias Remanescentes reúne Humor Político, Estilo de Vida, Religião, Portais Online, Serviços de Nuvem, e Outros, por conta do baixo percentual de compartilhamento. Na categoria Outros, as subcategorias Não Disponível, Compras, Serviços e Aplicações, Encurtadores de Links, e Outros Não Político, foram agrupadas sob Categorias Remanescentes pela mesma razão.

O conteúdo produzido por fontes profissionais (Veículos Profissionais e Conteúdo Político Profissional combinado) foi o mais compartilhado no Twitter, constituindo aproximadamente 50% do total de compartilhamentos, enquanto o Conteúdo Polarizante e Conspiratório representou menos de 2% dos compartilhamentos. Isso segue a tendência que observamos em outros países latino-americanos.²⁰ Conteúdo político profissional foi o mais compartilhado, totalizando 31% de todos os compartilhamentos, seguido de fontes profissionais de notícias, com 21% dos compartilhamentos. Na categoria Notícias Políticas e Informações, conteúdos de Vídeo/Imagem representaram 14% de todas os compartilhamentos, reforçando a dominância de

conteúdo audiovisual durante a campanha. Dos links categorizados como Outros, 16% estão vinculados a outras redes sociais, incluindo o Facebook, indicando um alto grau de uso cruzado das plataformas.

Além disso, após identificarmos cinco grupos principais de contas que seguem outros usuários conhecidos por compartilharem fontes de junk news, calculamos uma pontuação de heterofilia para cada combinação de pares de grupos (veja o suplemento online para o índice de heterofilia). Uma pontuação de heterofilia de 1,0 indica uma conexão neutra. Qualquer coisa acima disso indica uma empatia forte, enquanto qualquer coisa abaixo indica falta de conexão. Uma pontuação mais alta de um grupo em relação a ele mesmo indica mais conexões dentro do próprio grupo. Observamos uma alta pontuação de heterofilia (1,2) entre os grupos Alckmin e Bolsonaro, sugerindo um alto grau de interação entre os dois grupos. Isso pode ser atribuído ao fato de que ambos os candidatos se apresentam como opositores do Partido dos Trabalhadores (PT) e de seus seguidores, em geral, rejeitarem candidatos de esquerda.

Figura 2: Público Brasileiro de Junk News no Twitter



Fonte: Cálculos dos autores a partir de amostra coletada entre 22/08/2018 - 21/09/2018. Nota: Grupos são determinados a partir de associações de rede. Esse é uma visualização básica (ver complemento online para uma visualização completa)

Tabela 3: Tamanho, Cobertura, e Consistência de Grupos de Públicos de Junk News no Twitter				
Pontuação	Número Usuários	Usuários %	Cobertura	Consistência
Lula & Haddad	4,549	39	54	65
Jair Bolsonaro	3,946	34	81	27
Other	2,930	25	78	7
Geraldo Alckmin	126	1	24	0
Ciro Gomes	70	1	8	0
Total	1,1621			

Fonte: Cálculo dos autores a partir de amostra coletada entre 22/08/2018 - 21/09/2018. Percentagens foram arredondadas para o valor inteiro mais próximo, a não ser em casos em que fossem menores do que um.

Finalmente, a partir da Tabela 3 identificamos que os apoiadores de Jair Bolsonaro

têm uma cobertura de 81%, indicando que essas contas compartilhavam o maior número de fontes de junk news no Twitter, enquanto os apoiadores de Lula e Haddad compartilhavam o maior volume de fontes de notícias gerais, com 65%.

CONCLUSÕES

Nossas principais conclusões são: (1) o debate político no Brasil no Twitter é altamente partidário, com o principal candidato presidencial, Jair Bolsonaro, dominando a conversa nessa rede social; (2) contas associadas às hashtags de Lula e Haddad mostram o mais alto nível de tuitagem de alta frequência; (3) Os usuários brasileiros do Twitter estão compartilhando mais conteúdo político profissional do que junk news - a maior proporção em todas as eleições que estudamos; (4) enquanto os partidários de Bolsonaro espalharam a mais ampla gama de conteúdos produzidos por fontes reconhecidas de junk news, os partidários de Lula e Haddad foram responsáveis pelo maior volume de compartilhamento de fontes de junk news.

A alta proporção de notícias políticas profissionais compartilhadas por usuários brasileiros pode ser atribuída ao fato de que apenas uma fração pequena, mais politicamente educada e alfabetizada da população brasileira usa o Twitter para acessar notícias. No entanto, notamos que à medida que o compartilhamento de informações e os debates políticos se afastam progressivamente de plataformas públicas como o Twitter e se movem para espaços de discussão mais privados como o WhatsApp ou o Facebook Messenger, uma análise de outros serviços populares poderia produzir resultados diferentes sobre a disseminação de desinformação nas redes sociais no Brasil.

COMPLEMENTO ONLINE E DADOS COLETADOS

Acesse o link http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/10/brazil_supplement_por.pdf para material adicional relacionado a esta análise, incluindo (1) visualizações em alta resolução das redes para o Twitter, (2) a lista completa de segmentos e grupos, (3) cálculo das pontuações de heterofilia, (4) explicação detalhada do algoritmo de agrupamento hierárquico aglomerativo usado para criar grupos, (5) a redução do k-core usada para reduzir o conjunto de usuários do Twitter, (6) uma lista dos 10 principais sites de junk news que encontramos no banco de dados.

SOBRE O PROJETO

O Projeto sobre Propaganda Computacional, baseado no Oxford Internet Institute, é formado por uma equipe interdisciplinar de cientistas sociais e cientistas de dados que pesquisam como os atores políticos manipulam a opinião pública através das redes sociais. Este trabalho inclui a análise da interação de algoritmos, automação, política e redes

sociais para amplificar ou reprimir conteúdo político, desinformação, discurso de ódio e junk news. Os memorandos de dados são projetados para apresentar retratos rápidos da análise de eventos atuais em um formato curto e, embora reflitam a experiência metodológica e a análise considerada, eles não passam por revisão por pares. Os documentos de trabalho apresentam análises mais aprofundadas e argumentos ampliados que foram revisados de maneira colegiada e se engajam com questões públicas. Os artigos, capítulos de livros e livros do COMPROP são textos mais detalhados que passaram por revisão por pares e foram formalmente publicados.

AGRADECIMENTOS E TRANSPARÊNCIA

Os autores agradecem o apoio do European Research Council, “Computational Propaganda: Investigating the Impact of Algorithms and Bots on Political Discourse in Europe,” Projeto 648311, 2015-2020, Philip N. Howard, Pesquisador Principal. As atividades do projeto foram aprovadas pelo Comitê de Ética na Pesquisa da Universidade de Oxford, CUREC OII C1A 15-044. Pelo apoio ao nosso Observatório de Eleições e nossa pesquisa conduzida na Europa, agradecemos a Adessium Foundation, Microsoft Research e Omidyar Network. Todas as opiniões, achados, conclusões ou recomendações expressas neste material são as dos próprios autores e não refletem os posicionamentos da Universidade de Oxford, nem do European Research Council, Adessium Foundation, Microsoft Research, ou Omidyar Network. Também agradecemos a Dr. Vidya Narayanan e Lisa-Maria Neudert pelas suas contribuições para este documento.

REFERÊNCIAS

- [1] P. N. Howard, S. Woolley, and R. Calo, “Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration,” *Journal of Information Technology & Politics*, vol. 15, no. 2, pp. 81–93, Apr. 2018.
- [2] Jason Gilmore and P. N. Howard, “Does social media make a difference in political campaigns? Digital dividends in Brazil’s 2010 national elections,” *Center for Communication and Civic Engagement*, vol. Working Paper No. 2013-2.
- [3] R. Jansen, “Organizadora de grupo contra Bolsonaro relata agressão no Rio,” *Estado de S. Paulo*, 25-Sep-2018.
- [4] N. Newman, R. Fletcher, A. Kalogeropoulos, D. A. L. Levy, and R. K. Nielsen, “Reuters Institute Digital News Report 2018,” Reuters Institute, University of Oxford, 2018.
- [5] Cetic.br, “ICT households - Survey on the Use of Information and Communication Technologies in Brazilian Households,” Brazilian Internet Steering Committee, São Paulo, 2017.
- [6] P. Higa, “Facebook tem mais usuários que WhatsApp no Brasil e chega a dois terços da população,” *Tecnoblog*, 19-Jul-2018.
- [7] F. Molica, “Teorias da conspiração,” *Revista VEJA*, Juiz de Fora, 14-Sep-2018.
- [8] D. Söğür Hous, “Boatos e teorias da conspiração sobre atentado a Bolsonaro se espalham,” *Folha de S. Paulo*, São Paulo, 09-Jul-2018.
- [9] “Eleitor discerne entre comoção e voto,” *Valor*, São Paulo, 09-Nov-2018.
- [10] B. Brooks, “Brazil’s Bolsonaro leads presidential race but Haddad closing gap: poll,” *Reuters*, São Paulo, 24-Sep-2018.
- [11] G. Sá Pessoa, “Anúncios pagos em redes sociais ampliam recursos para candidatos,” *Folha de S. Paulo*, São Paulo, 29-Jan-2018.
- [12] “Redes sociais e mídias tradicionais são as fontes de informação com mais influência na escolha do presidente em 2018,” *Ibope Inteligência*, 13-Jun-2017. [Online]. Available: <http://www.ibopeinteligencia.com/noticias-e-pesquisas/redes-sociais-e-midias-tradicionais-sao-as-fontes-de-informacao-com-mais-influencia-na-escolha-do-presidente-em-2018/>.
- [13] FGV DAPP, “Robôs, redes sociais e política no Brasil: Estudo sobre interferências ilegítimas no debate público na web,” 24-Aug-2017.
- [14] InternetLab, “Bot or not: who are the followers of our candidates for president?,” *InternetLab*, 03-Jul-2018. .
- [15] Tribunal Superior Eleitoral, “Seminário Internacional Brasil - União Europeia Fake News: Experiências e Desafios,” *Tribunal Superior Eleitoral*. [Online]. Available: <http://www.tse.jus.br/hotsites/fakenews/>.
- [16] F. Brito Cruz, H. F. de C. da Silveira, J. de S. Abreu, M. S. de P. Andrade, R. S. Vieira, and T. D. Oliva, *Direito Eleitoral na Era Digital*. São Paulo: InternetLab, 2018.
- [17] V. Narayanan, V. Barash, J. Kelly, B. Kollanyi, L. M. Neudert, and P. N. Howard, “Polarization, Partisanship and Junk News Consumption over Social Media in the US,” *COMPROP Data Memo 2018.1*.
- [18] P. N. Howard, B. Kollanyi, S. Bradshaw, and L. M. Neudert, *Social Media, News and Political Information during the US Election: Was Polarizing Content Concentrated in Swing States?* Oxford: Internet Institute, Oxford University, 2017.
- [19] K. Krippendorff, “Reliability in content analysis,” *Human communication research*, vol. 30, no. 3, pp. 411–433, 2004.
- [20] M. Glowacki *et al.*, “News and Political Information Consumption in Mexico: Mapping the 2018 Mexican Presidential

Election on Twitter and Facebook,” Oxford
Internet Institute, University of Oxford, Data
memo 2018.2, Jun. 2018.