

**METODOLOGIAS APLICADAS ÀS REDES SOCIAIS ON-LINE:  
a importância de considerar a unidade de análise para a pesquisa**

***METHODOLOGIES APPLIED TO ONLINE SOCIAL NETWORKS:  
the importance of considering the unit of analysis for research***

---

Emerson Cervi\*

Luiza Jardim\*\*

**Resumo**

Este *paper* filia-se à área de trabalhos sobre técnicas e ferramentas metodológicas utilizadas para a análise de objetos empíricos em áreas de fronteira científica, no caso, as Redes Sociais On-line (RSO) e explora a importância da escolha da unidade de análise na pesquisa RSO. A partir da análise de produções recentes em alguns congressos brasileiros, foram identificados três grupos de análise: micro, meso e macro. O grupo micro considera como unidade de análise as contas/produtores de conteúdos, visando a identificar “quem diz” e “como se diz”. O grupo meso analisa as interações entre as contas, buscando compreender as relações entre elas e a formação de comunidades. Já o grupo macro se preocupa com a dimensão temporal da presença e intensidade do tema, identificando a sobrevivência do tema no Twitter e sua intensidade. O artigo destaca a importância de escolher a unidade de análise adequada para os objetivos da pesquisa, evitando buscar respostas certas para perguntas de pesquisa erradas. Além disso, a análise de conteúdo automatizada, uma das ferramentas mais utilizadas no objeto Redes Sociais On-line, pode ser útil para identificar a homogeneidade ou heterogeneidade dos textos publicados nas discussões sobre o tema. O estudo também demonstra como a análise no nível macro pode ser importante para compreender a presença e intensidade do tema ao longo do tempo. A fim de apresentar o emprego dessas diferentes técnicas, o artigo analisa as discussões no Twitter sobre uma reunião do ex-presidente Jair Bolsonaro com embaixadores sobre as urnas eletrônicas, em julho de 2022.

**Palavras-chave:** Redes sociais on-line. Twitter. Metodologia. Bolsonaro.

**Abstract**

The paper situates itself with the field of studies on techniques and methodological tools used for the analysis of empirical objects in scientific frontier areas, specifically focusing on Online Social Networks (OSN). It explores the importance of selecting the unit of analysis in OSN research. Through the analysis of recent works presented at some Brazilian conferences, the paper identifies three distinct analysis categories: micro, meso, and macro. The micro group considers user accounts/content producers as the unit of analysis, aiming to identify "who says" and "how it is said." The meso group investigates interactions between accounts, seeking to understand the relationships between them and the formation of communities. The macro group is concerned with the temporal

---

\*Emerson Urizzi Cervi é doutor em Ciência Política pelo Instituto Universitário de Pesquisas do Rio de Janeiro (IUPERJ), com estágio pós-doutoral em eleições na América Latina pela Universidad de Salamanca. Professor do departamento de Ciência Política da Universidade Federal do Paraná (UFPR). Professor permanente do programa de pós-graduação em Ciência Política (PPGCP) e do programa de pós-graduação em Comunicação (PPGCOM) na UFPR. Coordenador do grupo de pesquisa em Comunicação Política e Opinião Pública ([www.cpop.ufpr.br](http://www.cpop.ufpr.br)). E-mail: [ecervi7@gmail.com](mailto:ecervi7@gmail.com).

\*\*Luiza Jardim é doutoranda e graduada em Administração Pública pela Fundação Getulio Vargas (FGV-EAESP) e mestre em Ciência Política pela Universidade Federal de Minas Gerais (UFMG). E-mail: [luizabsjardim@gmail.com](mailto:luizabsjardim@gmail.com)

dimension of the presence and intensity of the theme, identifying the theme's survival on Twitter and its level of intensity. The article emphasizes the importance of choosing the appropriate unit of analysis for research objectives, avoiding seeking correct answers to the wrong research questions. Additionally, automated content analysis, one of the most widely used tools in the study of Online Social Networks, can be useful for identifying the homogeneity or heterogeneity of texts published in discussions on the topic. The study also demonstrates how macro-level analysis can be crucial for understanding the presence and intensity of the theme over time. To exemplify the application of these diverse analytical techniques, the article analyzes discussions on Twitter about former president Jair Bolsonaro's meeting with ambassadors about electronic voting machines, which took place in July 2022.

**Keywords:** Online social networks. Twitter. Methodology. Bolsonaro.

## Introdução

O avanço das Tecnologias da Informação e da Comunicação (TICs) proporcionou novas formas de relações, e as Redes Sociais On-line (RSO) configuraram-se como um terreno fértil para discussões e manifestações da sociedade. Na medida em que essas redes sociais foram ganhando protagonismo, cresceram também os estudos sobre a participação política dentro das ferramentas.

Neste artigo, investiga-se a relação entre a unidade de análise e os objetivos da pesquisa sobre RSO, com o objetivo de apresentar as principais técnicas utilizadas por pesquisadores do campo. Dentre essas técnicas, estão as utilizadas para descrever o comportamento on-line de manifestações públicas e seus efeitos na relação entre a intimidade e o público (Neves, 2017); a forma como temas públicos, como a pandemia, são tratados por autoridades políticas (Santos *et al.*, 2021); a disputa política, a partir de diálogos no Twitter<sup>1</sup> (Santos *et al.*, 2020); estudos de rede de desinformação formadas em campanhas eleitorais (Soares, 2019); a circulação de informação e os filtros-bolha (Recuero; Zago; Soares, 2017); as narrativas e a permanência de redes on-line (Segurado; Chicarino; Malina, 2017).

Não é nosso objetivo tratar do tema como novidade. A maior parte dos trabalhos empíricos sobre redes sociais on-line utiliza alguma técnica para chegar aos resultados. Também não pretendemos defender o uso de uma técnica única, excluindo as demais. Defendemos que técnicas de análise são complementares e cumprem uma função aditiva nas análises, desde que exista uma predominante. Nossa questão central é que a qualidade dos resultados de uma pesquisa cresce quando fica clara a relação entre a técnica de análise, objetivos e objeto empírico. Sem isso, e dada a diversidade e volume de informações facilmente acessadas pelos pesquisadores, até a simples descrição de um corpus empírico pode se perder na confusão entre técnicas, objetivos e objeto empírico que não dialogam entre si.

---

<sup>1</sup> Neste artigo, utilizamos a nomenclatura previamente utilizada em estudos sobre a rede social Twitter, que posteriormente ao período da coleta de dados passou a chamar-se X. Esta nomenclatura inclui tanto se referir à rede como Twitter, como entender as publicações como *tweets* e as repostagens como *retweets*.

A importância de considerar a unidade de pesquisa na seleção da técnica de análise está diretamente relacionada com a qualidade dos resultados obtidos. Devido ao grande número de técnicas disponíveis, selecionamos os três principais grupos de análise a partir de diferentes unidades de análise. O que diferencia é o nível da unidade de análise. Estes três grupos são: micro, ao considerar as contas/produtores de conteúdo como unidade de análise; macro, que pode ser interno ou externo às RSO; e, também macro, relacionado à passagem do tempo.

Na próxima seção, discorreremos sobre a utilização de técnicas destes três grupos de análise a partir de artigos apresentados nas principais conferências que tratam da temática no Brasil, a saber: ANPOCS, Compolítica e Compós. Foram coletados todos os artigos que fazem parte de Grupos de Trabalho (GTs) sobre ambientes digitais, apresentados entre 2017 e 2021. O objetivo do *paper* não é fazer uma revisão exaustiva da literatura que analisa empiricamente as características das redes sociais de forma exploratória ou descritiva. Utilizamos trabalhos já publicados para identificar e exemplificar as técnicas utilizadas para a análise de RSO, em especial quando a unidade de análise, objeto e objetivos não estão claramente relacionados.

Na seção seguinte, usamos uma base de dados coletada pelo grupo de pesquisa Observatório de Conflitos na Internet, liderado pela Universidade Federal do ABC, sobre a reunião organizada por Jair Bolsonaro em 18 de julho de 2022, que convocou embaixadores de pelo menos 40 países para debater o sistema eleitoral brasileiro. Na ocasião, Bolsonaro reuniu os participantes no Palácio da Alvorada e fez ataques, sem provas, às urnas eletrônicas, e o assunto repercutiu nas RSO. A base analisada é de dados do Twitter, entre os dias 7 e 29 de julho, e reúne 5,6 mil *tweets*.

### **Três grupos de análise**

A democracia digital é um dos temas centrais dos estudos de “Internet e Política” (Sampaio *et al.*, 2022), que pode ser entendido como um subcampo dentro de Comunicação e Política (Sampaio; Bragatto; Nicolás, 2016). A democracia digital pode ser definida, de maneira ampla, como uma articulação entre o mundo digital e o físico, podendo englobar qualquer iniciativa que faça uso de tecnologias digitais de comunicação para corrigir ou melhorar processos democráticos (Almada *et al.*, 2019; Mendonça; Pereira; Filgueiras, 2016).

Uma pesquisa realizada por Sampaio *et al.* (2022) concluiu que, no sentido de Bourdieu (1988) de campo acadêmico, democracia digital não se configura como um campo no Brasil, porque não há um conjunto consolidado e considerável de associações, congressos ou

periódicos especializados na temática, e tampouco é uma subárea reconhecida pelos principais órgãos de pesquisa do país – a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

No entanto, os autores afirmam que, se for realizada uma flexibilização do conceito de campo acadêmico, é possível associar democracia digital ao subcampo da Comunicação e Política, dado que as principais áreas que produzem sobre o tema são a Comunicação Social e a Ciência Política. Ao analisar os principais congressos dessas áreas, identificam-se GTs relevantes que reúnem produções acadêmicas sobre o tema. Para este estudo, selecionamos artigos a partir dos anais dos seguintes GTs: Fenômenos e Práticas da Política Online (Compolítica, 2021); Internet e Política (Compolítica, 2019 e 2017); Comunicação e Cibercultura (Compós, 2020, 2019 e 2017); Internet, Política e Cultura (Anpocs, 2021 e 2020); e Ciberpolítica, ciberativismo e cibercultura (Anpocs, 2017). Foram analisados 10 artigos, que compõem o corpus deste trabalho.

A partir disso, selecionamos os três principais grupos de análise a partir das diferentes unidades de análise apresentadas nos artigos analisados. O primeiro é micro, ao considerar como unidade de análise as contas/produtores de conteúdos. A análise de contas de produtores, replicadores, aqueles que interagem, parte da identificação da intensidade de presença dessas unidades nas RSO. O objeto, portanto, é “quem”. Aqui, o objetivo é analisar “quem diz”, e um possível desdobramento desta análise é a Análise da Ação, que visa a identificar o comportamento dos usuários, sendo o “como se diz”. Os comportamentos são identificados, na prática, pelo volume e heterogeneidade de interações. No caso do Twitter, a medição se dá no número de curtidas, replicações e respostas que uma conta recebe ou fornece. Aqui, o foco está nas ações que geram interações.

O segundo nível de unidade de análise é macro, e pode ser interno ou externo às RSO. A técnica mais utilizada para analisar internamente as conexões feitas para além das contas é a Análise de Redes. Esta é uma abordagem focada sobretudo na estrutura dos grupos sociais, buscando compreender o formato dessas estruturas e a posição dos atores dentro da rede (Recuero; Zago; Soares, 2017). Não se busca entender quem são os integrantes individuais do conjunto analisado, mas sim as características gerais do grupo, além das características formadas pelas interações entre eles. Essas características são representadas empiricamente na forma de medidas das redes ou representadas na forma de grafos, em que os atores são nós e as arestas são as conexões que se estabelecem entre eles. Aqui, o objetivo é identificar “como se organiza o que é dito”.

Já quando a análise é externa às redes, usa-se a Análise de conteúdo. Esse tipo de ferramenta reúne o conjunto de textos publicados na RSO com o objetivo de analisar sobretudo o que se diz, ao invés de quem diz ou da forma como é dito. A análise de conteúdo é considerada a mais antiga técnica de pesquisa social empírica. Desde o século XVII, há relatos de pesquisa sistemática a partir de termos em textos oficiais, religiosos ou não. Um desdobramento da análise de conteúdo nas RSO, ou seja, sobre “o que se diz” é a Análise Temática. Aqui, o objeto de análise é um tema, representado por um termo específico ou um conjunto de termos conectados.

Por fim, o terceiro grupo de unidade de análise – também macro – não está relacionado aos produtores ou aos conteúdos, mas sim à passagem do tempo. Aqui, o objetivo é analisar “qual a duração do que se diz”, e o objeto de análise é a passagem do tempo. As diferentes durações temporais de determinadas interações, padrões de comportamento, proeminências de contas ou usuários ou temáticas exige uma abordagem metodológica específica. Os resultados esperados são relativos à duração e permanência nos ambientes digitais. Para tanto, a análise temporal oferece um conjunto de técnicas próprias para medir as características de determinado fenômeno ao longo do tempo. No caso de RSO, a passagem do tempo pode ser medida em segundos, minutos, horas ou dias. Neste artigo trabalhamos com a unidade temporal “hora”, a partir da qual agregamos os totais e médias/hora em períodos do dia (manhã, tarde e noite. O quadro a seguir sumariza os tipos de análise apresentados aqui por objeto e objetivo a ser cumprido.

Quadro 1 – Sumarização das unidades de análise, objetivos e objetos em RSO

Nível	Unidade	Objetivo	Objeto
Micro	Conta/produtor	Quem diz	Contas
	Tipo de Ação	Como se diz	Interações
Macro	Rede	Como se organiza o que é dito	Forma das interações
	Conteúdo/Tema	O que se diz	Texto
	Tempo	Qual a duração/intensidade do que se diz	Unidade temporal

Fonte: Elaboração própria.

Como afirmado anteriormente, não se espera em uma pesquisa a apresentação de uma única unidade de análise ou objeto para cumprir determinado objetivo. Não raras vezes, a análise de interações é mesclada com análise temporal. Ou, a análise que propõe uma tipologia de usuários que interagem em determinado tema também utiliza algum tipo de análise textual. O que distingue o uso adequado das técnicas de pesquisa dos demais usos é a capacidade de deixar claro qual a técnica predominante no trabalho e, para isso, defendemos que é preciso relacionar a unidade de análise, objeto e objetivos para a definição das ferramentas analíticas adequadas.

Entre os exemplos de trabalhos com análise de conteúdo no Twitter, temos análises empíricas sobre os produtores de conteúdos nas RSO. Santos *et al.* (2021) estudaram as

publicações de dois usuários do Twitter: O presidente da república e o governador de São Paulo no período de um ano. Eles buscavam analisar como os produtores tratavam do tema covid-19. Antes disso, Santos *et al.* (2020) tinham analisado a disputa política a partir de contas no Twitter de grupos de direita, em especial o que os autores chamam de “clã bolsonaro” e MBL. Este é um clássico estudo de produtores que buscou entender as convergências e divergências entre os usuários pesquisados.

Neves (2017) analisa as contas que deram visibilidade à hashtag #primeiroassédio, e aponta como a adesão à campanha de combate ao machismo gerou um espaço comum entre a intimidade e o público. Como ferramenta de coleta de dados, a autora usa a observação de relatos publicados no Twitter. Apesar de não detalhar a maneira como a coleta e a análise de dados são feitas, pode-se inferir que a unidade de análise é o conteúdo, e o objeto é o texto, com o objetivo de analisar o que se diz associado à *hashtag* analisada.

Outra unidade macro muito utilizada em análises de RSO é a rede. Soares (2019) analisa as redes de contas formadas em campanhas de desinformação durante as eleições brasileiras de 2018. Ele usa as redes para conectar ações de atores com grande visibilidade que identificam a forma das conexões que têm em comum a desinformação. O objeto é a forma das interações, consistindo em uma análise da rede enquanto unidade. O mesmo autor, em 2020, utiliza uma abordagem de métodos mistos para discutir os impactos da recirculação de notícias no Twitter para verificar se a maior participação de usuários impulsiona a visibilidade de determinados conteúdos. O autor faz uso dos dois métodos mencionados no grupo 2, a Análise de Redes e a Análise de conteúdo (Soares, 2020). A primeira é utilizada com o objetivo de compreender os processos de circulação de informação no Twitter e a segunda é utilizada para compreender o papel dos líderes de opinião na ressignificação de acontecimentos políticos, por meio do conteúdo de suas mensagens. Sendo assim, na primeira análise o objeto são as interações, enquanto na segunda, os textos.

Recuero, Zago e Soares (2017) exploram a circulação de informações em redes próprias, o que permite a formação de filtros-bolha. Assim como Soares (2020), os autores agregam à análise de redes a análise de conteúdo para questionar o caráter democratizante das RSO. A análise de redes visa a identificar como são constituídas as redes de conversação no Twitter, enquanto a análise de conteúdo verifica quais informações circulam e quais não circulam.

Em outra abordagem, Segurado, Chicarino e Malina (2017) investigam as redes formadas em eventos distintos: o impeachment de Dilma Rousseff e a eleição para prefeitura de São Paulo, em 2016. A partir de fluxos e conexões, os autores chegam às narrativas que se mantêm

e as que são alteradas nos dois acontecimentos. Trata-se de um estudo comparativo da formação de redes de conversações.

Dienstbach *et al.* (2021) utilizam a técnica de Análise de conteúdo com o objetivo de descrever e analisar a natureza e o funcionamento de discursos em defesa do “tratamento precoce” para o novo coronavírus. Os autores adotam uma abordagem mais próxima da análise semântica, que é diferente do que apresentaremos a seguir, que é centrada na descrição léxica. A unidade de análise é o conteúdo e seu objeto, portanto, é um conjunto de publicações no Twitter, filtradas a partir de seu conteúdo. Os autores aprofundam a discussão teórica sobre as técnicas utilizadas e declaram adotar uma investigação interna à Análise de conteúdo, que seria o foco de interesse na semântica argumentativa, que “recai antes sobre os possíveis usos da língua durante interação entre os interlocutores do que no seu ‘conteúdo’” (Dienstbach *et al.*, 2021, p. 4).

O estudo de Dienstbach *et al.* (2021) evidencia o que foi apontado neste artigo, sobre a utilização de mais de uma técnica dentro do mesmo estudo. Apesar de declararem fazer uso de Análise de conteúdo, no próprio objetivo da pesquisa percebe-se uma inclinação à Análise de Redes, com foco na discussão sobre “como se organiza o que é dito”. Os autores analisam mapas de interação da rede como uma etapa prévia à análise do conteúdo. Os resultados apontam para padrões argumentativos linguísticos, reforçando a centralidade da Análise de conteúdo.

Santos Jr. (2021) também combina múltiplas técnicas em seu estudo sobre desinformação com abordagem multiplataformas. O objetivo do estudo é responder a perguntas sobre a interação nas plataformas (grupo 2), a dinâmica temporal da circulação de informações (grupo 3) e os principais atores que compartilharam as mensagens (grupo 1). O autor declara que são utilizados dois objetos no estudo, *links* e *hashtags*, portanto, compreendidos na nossa classificação enquanto texto. Na prática, o estudo parte de um objeto textual – os *links* e *hashtags* – e analisa a articulação das mensagens e a sua dimensão temporal. O artigo usa também as interações enquanto objeto para compreender a ressonância e o tempo para medir se há coordenação entre diferentes atores na disseminação das informações.

### **Aplicação prática dos três grupos de análise**

À luz dos três grupos de análise de RSO, apresentamos neste artigo três abordagens distintas para análise de uma mesma base de dados, orientadas pelos objetivos e unidades de análise. Para tanto, utilizamos uma base de dados sobre a reunião organizada pelo presidente Jair Bolsonaro em 18 de julho de 2022, na qual ele convocou embaixadores de ao menos 40

países para debater o sistema eleitoral brasileiro. Analisamos a discussão no Twitter, com base em 5,6 mil *tweets* coletados entre os dias 7 e 29 de julho de 2022, compreendendo a data da reunião do presidente com os embaixadores no Brasil, que foi dia 18 de julho.

A coleta foi realizada pela equipe do Observatório de conflitos na Internet<sup>2</sup> a partir da entrada API do Twitter no período descrito acima. Os *tweets* selecionados continham os termos relacionados à reunião. Para as análises micro, foram utilizados pacotes estatísticos para sumarização de ocorrências, seja em número de contas, seja em número de postagens. Para a unidade macro de análise, na formação de redes a partir das conexões de contas, foi utilizado um plugin do Excel, o NodeXL, na sua versão gratuita. Para a análise de conteúdo automatizada aplicou-se a Interface do R para Análise Multidimensional de Textos (Iramuteq). Por fim, para a análise das séries temporais usou-se um pacote de código aberto para análise de regressões econométricas de séries temporais (Gretl). A seguir são apresentadas cada uma das técnicas, ferramentas metodológicas, unidades de análise e objetivos a serem alcançados.

### Unidade 1 – Conta/produtor para identificar disposição em participar

O primeiro nível de análise é o micro. No nosso caso, trata-se da análise das contas que se encontram na amostra aleatória daqueles que publicaram sobre a reunião com os embaixadores. Aqui, o objetivo é descrever quem tratou do assunto. Para tanto, primeiro descrevemos o comportamento geral dos participantes e depois classificamos a parte mais ativa das contas, a partir do número de postagens no período. Optamos por considerar as contas que fazem parte do grupo de 1% mais ativo, ou seja, dos que mais postaram sobre o tema. A Tabela 1 a seguir sumariza as principais estatísticas para todo o conjunto e para os mais ativos. O total de 5.605 postagens contidas na amostra compreende 3.608 contas no Twitter. Assim, o percentil mais ativo inclui as 36 contas que mais postaram no período de análise sobre o tema.

Tabela 1 – Estatísticas descritivas das contas

<b>Estatística</b>	<b>Total</b>	<b>1% mais ativo</b>
Número de contas	3.608	36
Total de postagens (Perc.)	5.605(100%)	484(8,6%)
Média (des. Pad.) de postagens	1,5 (1,71)	13,8 (5,58)
Mediana	1	12
Máximo	35	35
Mínimo	1	10

Fonte: Elaboração própria.

<sup>2</sup> <https://observa.pesquisa.ufabc.edu.br/>

No conjunto total, a média de postagem por conta ficou em 1,5 tuítes, com desvio padrão de 1,71. Isso indica grande variabilidade interna. No entanto, a mediana ficou em 1 tuíte por conta, mostrando que pelo menos metade das contas consideradas postou uma única vez sobre o tema no período de análise. O número máximo de tuítes foi de 35, da conta @cnnbrasil, e o mínimo foi de apenas uma postagem. A concentração de atividade em algumas contas fica mais clara quando consideramos apenas o percentil mais alto, ou seja, o conjunto de 1% de contas com mais atividade. Embora sejam apenas 36 contas (1%), esse grupo representa 8,6% das postagens (484) que compõem a amostra. A média de *posts* por conta no percentil mais alto fica em 13,8, com desvio padrão de 5,58, indicando maior homogeneidade em relação ao grupo total. Isso também pode ser verificado na proximidade entre média e mediana, que está em 12 postagens por conta. O número máximo é 35 e o mínimo no percentil superior foi de 10 postagens por conta.

A segunda etapa da análise de quem participa das discussões parte da categorização de contas. Aqui, categorizamos as contas do percentil superior de postagens. Das 36 contas, duas estavam suspensas no momento da análise e foram desconsideradas. As demais foram categorizadas a partir da indicação do nome e da descrição na bio da conta. Ao final, as contas foram agregadas em seis categorias. São elas: i) Imprensa – contas de veículos de comunicação tradicionais (impressos ou eletrônicos) e nativos digitais; ii) Militante de esquerda – quando se apresenta como pessoa física que defende pautas e de políticos de esquerda; iii) Militante de direita – quando se apresenta como pessoa física que defende pautas e de políticos de direita; iv) Movimento – quando se identifica como entidade que defende pautas específicas (na amostra analisada foram encontrados apenas movimentos de direita); v) Jornalista – quando se apresenta como jornalista/comunicador, pessoa física, e não meio de comunicação e vi) Usuário comum – quando se trata de pessoa física, mas não é possível identificar filiação a algum movimento ou posição política explícita no nome, foto e bio de apresentação. O quadro 1 a seguir sumariza os nomes das contas por categoria.

Quadro 1 – Categorização das contas no percentil superior de postagens

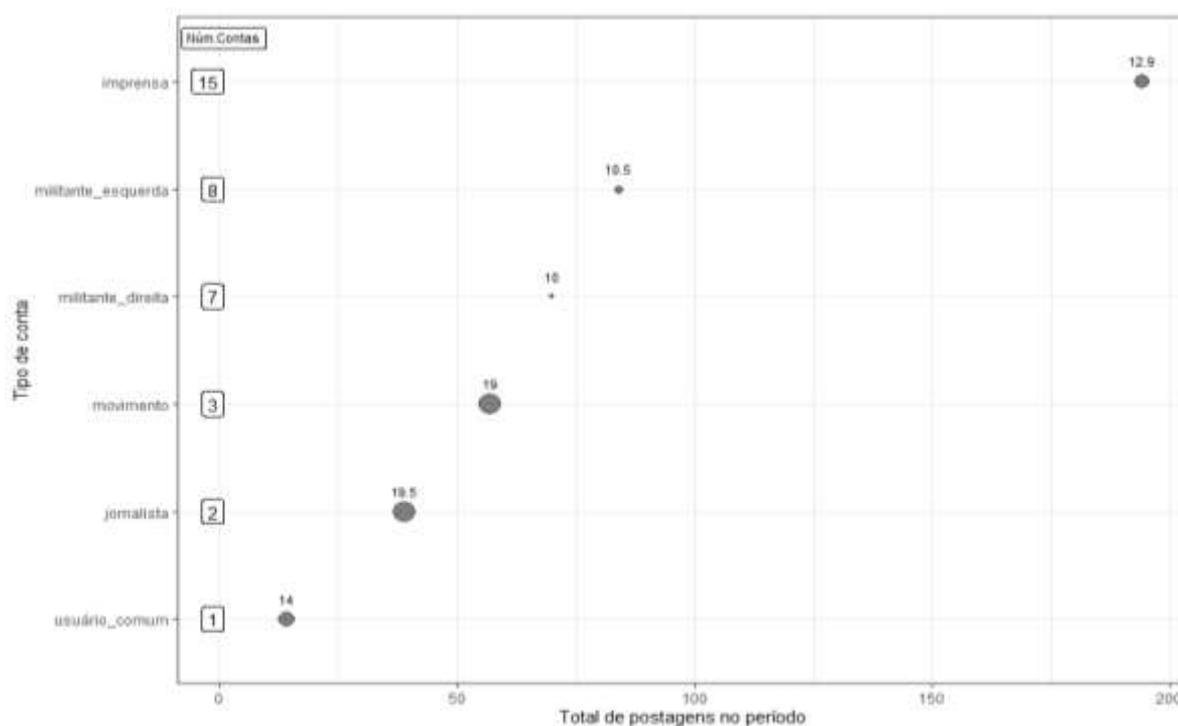
<b>Categoria</b>	<b>Contas (número de postagens)</b>
Imprensa	@cnnbrasil(35), @correio(17), @jovempannews(15), @globonews(14), @revistaforum(14), @feedynewsbr(12), @jornaloglobo(12), @o_antagonista(12), @brasil247(11), @folha(11), @jornaldarecord(11), @camarotעדacpi(10), @em_com(10), @estudioi(10)
Militante de esquerda	@e70004251de(15), @joaofernandesf1(3), @mortarischmidt(13), @tarlouze(12), @tiosamnews(11), @claudio_livre(10), @fenam924(10)
Militante de direita	@everthon_garcia(13), @conservadora191(12), @deborabpaim_sp(12), @nebels(12), @walsilva_wal(11), @jjairobrasil(10)
Movimento	@brazilfight(31), @bolsonarobrazil(13), @consdobrasil(13)
Jornalista	@vonivar(26), @blogdonoblat(13)
Usuário comum	@duscpcp2016(14)

Nota: as contas @sourobodolso e @todoscontraocom aparecem no percentil superior de postagens, porém, no momento da análise elas estavam suspensas pelo Twitter, o que impede a sua categorização.

Fonte: Elaboração própria.

O Gráfico 1 a seguir mostra a distribuição dos tipos de contas por total de postagens e média de tuítes por conta. Por exemplo, na categoria “imprensa” estão 15 contas individuais, que, juntas, somaram 197 postagens no período, o que resulta em um média de 12,9 tuítes por conta.

Gráfico 1 – Tipos de Contas por número de postagens e média de tuítes por conta Fonte:



Elaboração própria.

Apesar de ser a categoria com maior número de postagens, a “imprensa” não apresentou as maiores médias por conta. As duas contas de “jornalistas” representaram 19,5 postagens de média, ficando em primeiro lugar, com um total de 39 tuítes. Em seguida vem a categoria “movimentos”, com 19 postagens de média por conta, totalizando 57 tuítes. Militantes de esquerda, com 8 contas e média de 10,5 postagens, representaram 84 tuítes, e militantes de direita, com 7 contas e média de 10 postagens, representaram 70 postagens. A única conta de “usuário comum” fez 14 postagens no período.

Em conclusão, ao analisar as contas produtoras de conteúdos sobre determinado assunto no Twitter, buscamos identificar quem tem mais disposição em participar. No exemplo, percebemos que a imprensa tem maior número de contas com disposição em participar, o que resulta em maior presença dela no número de postagens. No entanto, isso é diferente da

disposição individual em participar. Percebemos que contas de movimentos e contas de jornalistas são as mais “dispostas”, pois apresentaram as maiores médias. A partir daí é possível comparar as categorias mais dispostas em participar desse tema com outros temas que são tratados pelos usuários do Twitter.

Importante é não confundir a disposição em participar com notoriedade digital. Como apresentado na seção anterior, pesquisadores da área costumam tratar principalmente da notoriedade, ou seja, capacidade de se fazer presente, ainda que com pouca atividade individual. Esta é a diferença para a disposição em participar, que se dá pela maior participação relativa na produção de postagens, ainda que elas tenham menor alcance. As duas características – disposição e notoriedade – são medidas no nível individual, mas a disposição se dá pelo número de postagens, enquanto notoriedade é identificada com a capacidade de mobilizar digitalmente, que é medido pelo número de interações. É possível ter muita disposição e baixa notoriedade ou o contrário, dependendo do tema em discussão. No próximo tópico apresentamos uma forma de análise da notoriedade de contas digitais.

### **Unidade 1 – Conta/interações para identificar notoriedade digital**

Nem sempre a disposição de participar coincide com notoriedade e, na maior parte dos casos, como apresentado a seguir, elas apresentam sinais trocados: mais notoriedade é acompanhada de menos disposição em participar. Se disposição em participar é medida pelos números de intervenções/postagens, notoriedade é identificada pelo efeito que causa no ambiente digital a partir do volume de interações. No Twitter é possível interagir com uma postagem com resposta a ela, replicando-a ou curtindo-a. Embora cada reação tenha um significado distinto, para fins de identificar notoriedade, aqui vamos somar os três tipos para chegar a um total de reações. Partimos do princípio que quanto mais reações por postagem, maior a notoriedade da conta.

A Tabela 2, mostra as estatísticas descritivas das interações e permite identificar que as diferenças entre o percentil superior aqui são maiores que as indicadas na disposição de participação (Tabela 1). Considerando o total de 5.605 postagens, foram coletadas 810.198 interações entre respostas, *retweets* e curtidas. A média geral de postagens por conta ficou em 1,55, enquanto a média de interações foi de 104,3 por postagem, com mediana de uma reação por postagem. Isso significa que pelo menos metade das postagens teve no máximo uma reação. O máximo foi de 48.454 interações e o mínimo foi de zero.

Se considerarmos apenas o percentil superior do total de interações, encontraremos 94 postagens, o que representa 1,6% do total, distribuídas nas 36 contas. A média de postagem por conta ficou em 2,61. Essas postagens tiveram 500.829 interações, representando 61,8% do total de interações. Aqui já é possível perceber como a reputação se distribui de maneira muito desigual. Apenas 1,6% das postagens, publicadas por 1% das contas, representam 61,8% das interações. A média de interações por postagem no percentil superior é de 7,3 mil reações, com mediana acima de 4 mil por postagem. O post com maior número de reações teve 48.454 e o com menor ficou em 2.078.

Tabela 2 – Estatísticas Descritivas das interações

Estatística	Total	1% com mais interações
Número de contas	3.608	36
Total de postagens (Perc.)	5.605 (100%)	94(1,6%)
Média postagem/conta	1,55	2,61
Total de interações (Perc.)	810.198 (100%)	500.829(61,8%)
Média (des. Pad.) de reações/postagem	104,3 (1.134,18)	7.311,53 (8.712,79)
Mediana de reações/postagem	1	4.086
Máximo de reações/postagem	48.454	48.454
Mínimo de reações/postagem	0	2.078

Fonte: Elaboração própria.

As estatísticas descritivas mostram como a reputação digital se distribui de maneira muito desigual e não tem relação com disposição de participação. O percentil que consegue as maiores interações tem menos que o dobro de postagens que o total da amostra. No entanto, concentra quase 2/3 do total de interações, com média 70 vezes superior de reações quando comparado à média geral. Para dar continuidade à análise por reputação digital, categorizamos as contas a partir do critério anterior. Aqui, apareceu uma nova categoria, a de “político” e não apareceram as categorias “militante de esquerda” e “usuário comum”. O quadro 2 a seguir sumariza todas as contas que compõem o percentual superior de reações, por categoria, e média de interações na amostra.

Quadro 2 – Categorização das contas no percentil superior de interações

	Contas (média de reações por postagem)
Político (11,7mil)	@jairbolsonaro(48,4mil), camargodireita(21,7mil), @edraposo_(12,9mil), @dfdanielfreitas(10mil), @marcofeliciano(9,5mil), glauber_braga(6,6mil), @_luzdocarmo(5,8mil), @bolsonarosp(5mil), @joubberth19(4mil), @ivanvalente(3mil), @mariadorosario(2mil).
Jornalista (6,2mil)	@ricardo_abreu_(25,1mil), @gabrielaprioli(6,1mil), @andrezamatais(3mil), @veramagalhaes(2,7mil), @helenachagas(2,3mil), @miriamleitaio(2,3mil), @andreshalders(2,1mil).
Militante de direita (5,2mil)	@clauwild1(10,4mil), elainnefaria1(9,3mil), @rafalougou(8,8mil), @mitags(7mil), @rosane15734055(5,9mil), @brom_elisa(3,9mil), @freu_rodrigues(3,4mil), @dimacgarcia(3,2mil), @profpaulamarisa(3mil), @luizcor24841840(2,4mil), @sandramorenah22(2,4mil), @saulchristos(2,3mil).
Movimento (3,5mil)	@brazilfight(3,8mil), @slpgn_giants_pt(3,2mil)
Imprensa (3,4mil)	@globones(4,3mil), @conexaopolitica(2,5mil).

Nota: a conta @guedinhoexplica não foi encontrada e a conta @sfirminofilho estava suspensa pelo Twitter no momento da análise. As duas foram excluídas.
---

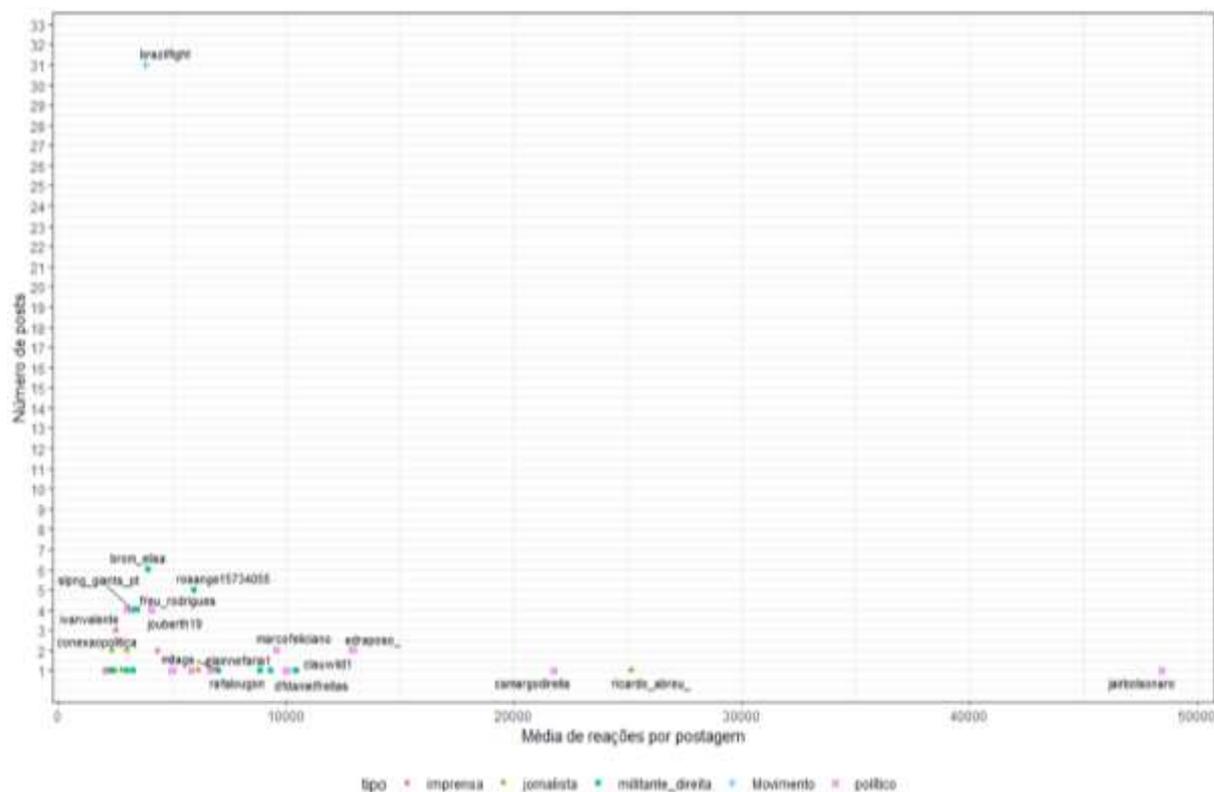
Fonte: Elaboração própria

O quadro 2 acima mostra que os políticos apresentam maior notoriedade na discussão do tema. Além de apresentarem o maior número de contas no percentil superior de interações, também têm a maior média de reações, com 11,7/postagem. A conta @jairbolsonaro passa de 48 mil interações, sendo a que recebeu maior número de reações. Em seguida estão as contas de jornalistas, com média de 6,2 mil interações. Porém, aqui há uma conta que se destaca das demais: @ricardo\_abreu\_ apresenta média de 25 mil interações, enquanto as demais contas pessoais de jornalistas giram em torno de quatro mil de média. Os militantes de direita apresentam maior número de contas no percentil superior, embora com menor média de interações (5,2 mil/postagem). Por fim, dentre as duas categorias institucionais, estão duas contas de movimentos, com média de 3,5 mil interações, e duas contas de meios de imprensa, com 3,4 mil interações de média.

Se compararmos os Quadros 1 e 2 percebemos uma evidente diferença entre disposição em participar e notoriedade do participante. Entre as categorias com disposição destaca-se a de meios de imprensa, além disso, militantes de direita e de esquerda apresentam participações próximas entre eles no que diz respeito à produção de postagens. Já quanto à notoriedade, medida em volume de interações, os políticos se destacam, seguidos de jornalistas e militantes de direita. Contas de figuras públicas externas às redes sociais, como políticos ou jornalistas, e figuras públicas nativas das redes sociais, militantes de direita, são as que apresentam maior notoriedade.

O Gráfico 2 a seguir mostra a distribuição das contas no percentil superior de interações entre média de reações pelo número de postagens presentes na amostra. As categorias estão indicadas por diferentes cores. A principal informação da imagem é que as contas com grande notoriedade costumam fazer poucas postagens. Há um *outlier*, o @brazilfight, com 31 postagens. Tirando isso, todos os demais casos ficam abaixo de sete *posts*. No caso das médias de reações por *posts*, outro *outlier* é a conta @jairbolsonaro, que fica muito distante de todas as demais. Também é importante perceber que, entre as contas com maior notoriedade e apenas uma postagem, estão principalmente políticos e militantes de direita. Das 36 contas do percentil superior, 24 fizeram apenas uma postagem e, dessas, apenas cinco são de jornalistas. Contas com maior notoriedade em apenas uma postagem são de políticos ou de militantes de direita.

Gráfico 2 – Relação entre número de postagens e média de reações por conta do percentil superior



Fonte: Elaboração própria.

A análise das contas que participam da discussão sobre a reunião com embaixadores exemplifica que há pelo menos duas grandes dimensões distintas de análise dos produtores de conteúdos em ambientes digitais. Uma delas é a do interesse em participar da discussão, com mais produção de conteúdos. A outra é a notoriedade de quem participa, medida pela capacidade de gerar interações nos conteúdos produzidos. Não se trata de diferenças individuais, entre contas, mas sim de categorias. Jornalistas e militantes dos dois extremos ideológicos são mais interessados em participar, enquanto políticos apresentam menos interesse. No entanto, militantes de esquerda praticamente não contam com notoriedade. Jornalistas apresentam menos notoriedade do que interesse em participar. Políticos e militantes de direita são os que gozam de mais notoriedade, quando medida pela capacidade de gerar interações digitais. O ganho que os estudos de redes digitais on-line têm é a capacidade de diferenciar o interesse em participar da notoriedade do participante. Até aqui, foram apresentadas as análises no nível micro. No próximo tópico passaremos ao nível macro.

**Unidade macro - Análise de rede como forma de identificar padrões de relações**

Também é possível analisar redes sociais on-line de forma macro, sem dar atenção para as ações individuais e sim aos comportamentos agregados. Uma dessas formas é usar técnicas de análise de rede, que são muito comuns em várias abordagens que visam a identificar conexões entre atores. No caso do Twitter, o formato e a intensidade das conexões acontecem entre contas de usuários. Existem várias formas de medir conexões e formatar redes no Twitter. Pode ser pelas respostas, pelos *retweets* ou por citações. Aqui, aplicamos a técnica para os casos em que contas são citadas no corpo do texto de uma postagem. Com isso, limitamos a rede ao conjunto de conexões que acontecem pela presença direta do “@” no texto publicado e não apenas por marcações da rede. Além disso, é possível que em um *tweet* seja citada mais de uma conta. Aqui, usaremos apenas a primeira citação, quando houver mais de uma. Com isso, será possível criar a rede de conexões mais fortes entre citações na amostra selecionada.

Dos 5.605 *tweets* que compõem a amostra, em 1.093 deles há citação direta de outra conta no corpo do texto. Então, aqui usaremos esse subconjunto como base para construir a rede de conexões. Seguindo o princípio do destaque do percentil superior, a Tabela 3 a seguir indica as dez contas mais citadas nos textos sobre a reunião com embaixadores. Elas foram agregadas em categorias, por definição da bio ou função institucional. Aqui, aparece pela primeira vez uma categoria de “instituições de Estado”, que não tinham tido destaque nem na análise das contas por disposição de participar, nem por notoriedade. Também é a primeira vez que ganha destaque o “Youtube”, que é outra rede e aqui está categorizado como repositório de conteúdo digital, junto com “Googlenews”.

Quadro 3 – Principais contas citadas em textos por categorias

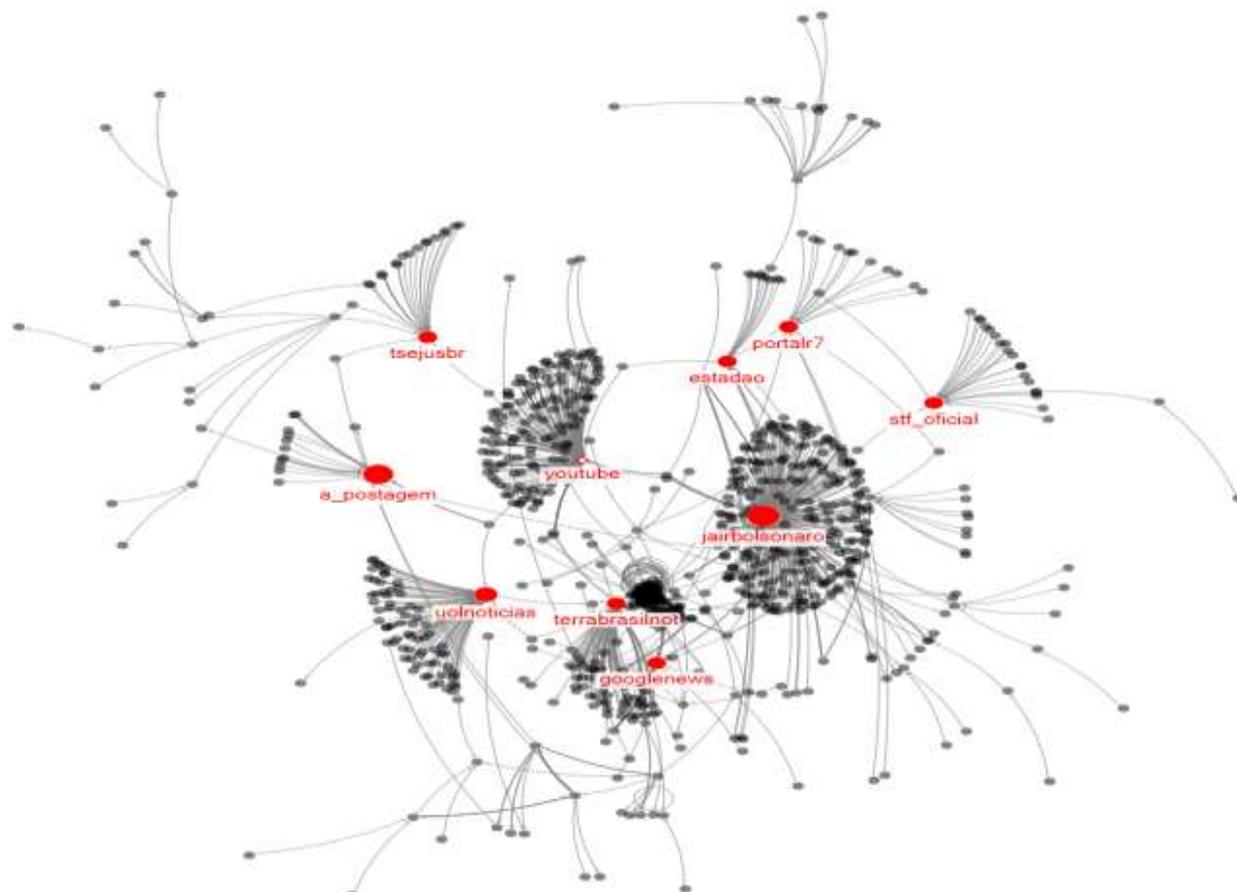
	Contas (número de primeiras citações)
Político	@jairbolsonaro(327).
Instituição de Estado	@stfoficial(22), @tsejusbr(21).
Repositório digital	@youtube(189), @googlenews(29).
Movimento	@Terrabrasilnot(68).
Imprensa (3,4mil)	@uolnoticias (81), @a_postagem(18), @estado(15), portalf7(15).

Fonte: Elaboração própria.

A conta que mais aparece na rede é “jairbolsonaro”, com quase um terço do total de primeiras citações no texto. Ele é o único político que aparece no percentil superior. As instituições de Estado citadas são o “STF” e o “TSE”. Entre os repositórios digitais, os que aparecem no percentil superior são Youtube e Googlenews. Nesse caso, o Youtube é citado,

provavelmente, por conta do vídeo da reunião. Existe a conta de um movimento, a “Terrabrasilnot”, que, embora passe a impressão de ser um portal noticioso, é um movimento conservador em seu conteúdo e na própria definição. Os portais noticiosos que aparecem no percentil superior são “UOL”, “Estadão” e “R7”, entre os tradicionais. A conta “A\_postagem” é de uma iniciativa de jornalismo colaborativo.

Gráfico 3 - Rede de primeiras citações no texto para o percentil superior de contas citadas



Distância geodésica máxima = 13	Densidade = 0,000958
Distância geodésica média = 4,35	Razão de reciprocidade = 0,016

Fonte: Elaboração própria

A análise de redes sociais possui uma série de estatísticas próprias que servem para identificar as características de cada grupamento. Em geral, elas são usadas para comparar densidade, formato e intensidade de conexões em diferentes redes. Aqui, apenas descreveremos as principais características da rede de primeiras citações sobre a reunião com embaixadores. A distância geodésica máxima de 13 indica que é necessário passar por 13 nós/contas para chegar ao ponto mais distante da rede em relação a qualquer ponto inicial. A distância geodésica média indica a média de distância entre dois nós/contas quaisquer na rede. Nesse caso, a distância média é de 4,32 nós/contas. Densidade é uma medida geral da rede. O valor máximo

é 1 e ocorre quando todos os nós/contas estão conectados entre si. No caso da rede acima a densidade de 0,000958 indica baixa conexão entre os nós, em outras palavras, contas que citam determinados nomes tendem a não citar outros. Por fim, a razão de reciprocidade mostra a capacidade que um nó/conta que cita algum nome tem de intermediar a relação de outro nó/conta com aquele nome. Como a densidade foi baixa, a tendência é que a razão de reciprocidade também fique próxima de zero, como foi o caso.

Ao contrário das abordagens micro iniciais, a análise macro a partir das técnicas de rede permite identificar o formato e a intensidade de interações, assim como o direcionamento e, em algumas vezes, a busca por validação de posições determinadas. Usando o mesmo critério do percentil mais alto de primeiras citações em textos, encontramos uma característica da rede que até então não tinha aparecido: a presença de instituições de Estado, especificamente do judiciário e da justiça eleitoral.

A análise de redes é uma das formas de análise macro, que trata das conexões, proximidades e distâncias entre conjuntos de contas. Também é possível analisar os textos de forma macro, ou seja, independente da conta de origem, para identificar padrões léxicos presentes na conversação sobre o tema.

### **Unidade macro – análise léxica para descrever o conteúdo das conversações**

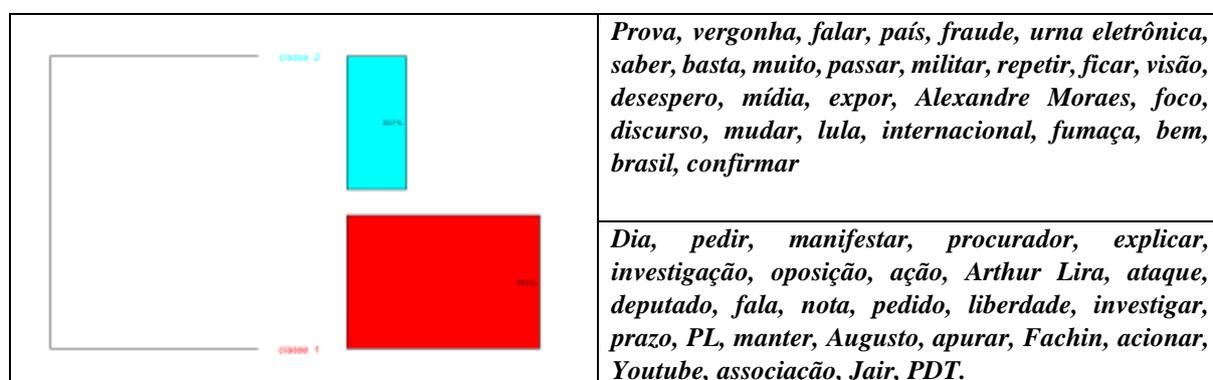
Outra abordagem macro nas análises de conversações de redes sociais é a partir do conjunto de textos publicados. A análise de conteúdo é considerada uma das mais antigas técnicas de pesquisa social e serve para identificar o que os atores dizem e como eles se manifestam. Aqui, o objetivo não é identificar os participantes dos debates, a partir de suas características originais, mas sim, agregar os autores por características presentes nos textos. O exemplo utilizado aqui é o de análise léxica de conteúdo, que foca na presença, volume e proximidade entre termos que aparecem nos textos. Outra abordagem seria a semântica, cujo objetivo seria analisar o sentido das construções textuais. Como as redes sociais produzem um volume muito grande de textos, a análise semântica tende a ser indireta e a técnica mais comum é a partir da análise de sentimentos.

Como exemplo dos léxicos usados nas discussões sobre a reunião com embaixadores, usaremos o algoritmo de Reinert para estabelecer *Clusters* Hierárquicos Descendentes (CHD) de termos. Segundo essa forma automatizada de análise, o algoritmo estabelece

limites de proximidade entre termos para identificar agrupamentos (*clusters*) entre os que aparecem em conjunto. O número de *clusters* é a primeira informação relevante. Ele indica quão heterogêneo (maior número) ou homogêneo (menor número) foi o conjunto de léxicos no texto. Além disso, os termos específicos de cada *cluster* também mostram quais são as possíveis direções das discussões. Seguindo a mesma abordagem dos tópicos anteriores, organizamos dois conjuntos textuais, sendo o primeiro com os textos de todos os 5.605 *tweets*. Em seguida, a técnica de análise automatizada é replicada, mas apenas para os textos publicados pelas contas que integram o percentil superior de interações. O objetivo é comparar as características do conjunto geral com a dos textos que tiveram maior volume de reações na rede.

O Gráfico 4 a seguir mostra o resultado do CHD para o conjunto total de *tweets*. A primeira informação importante é o pequeno número de *clusters*: foram apenas dois. Isso indica certa homogeneidade no uso de termos do início ao fim do período de análise. Além disso, o *cluster* 1 (vermelho) representa aproximadamente 3/4 do total. O *cluster* 2 tem apenas 22,7% do conjunto de léxicos. Por fim, os termos agregados em cada um dos *clusters* não apresentam grandes diferenças entre si. O *cluster* 2 (azul) tende a ser mais crítico à reunião, com termos como “desespero”, “vergonha” ou de personalidades públicas como “Alexandre Moraes” e “Lula”, que são identificados como contrários à reunião. Já o *cluster* 1, o maior em volume, tende a apresentar termos mais próximos de Bolsonaro, como “PL”<sup>3</sup>, “Jair”, “deputado”, “liberdade”. O volume de texto mais próximo de uma versão a favor da reunião representou três vezes mais que os léxicos críticos a ela.

Gráfico 4 – *Clusters* por CHD para o conjunto total de textos

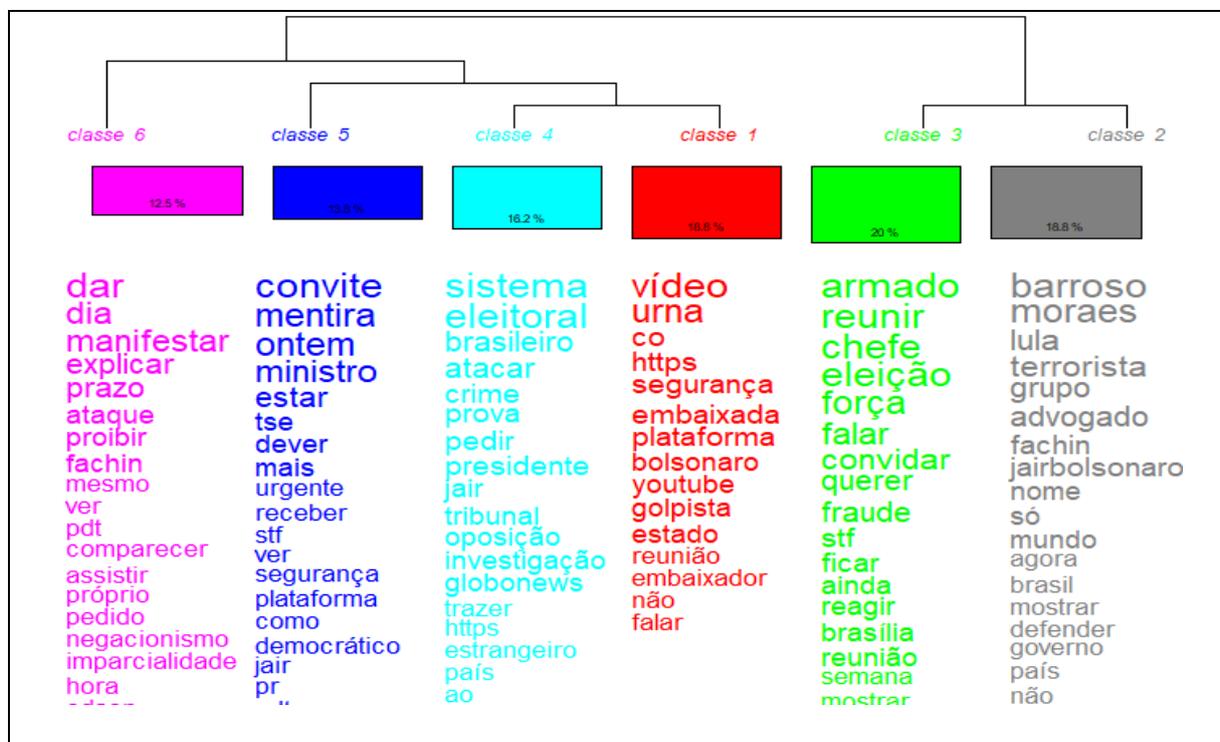


Fonte: Elaboração própria.

<sup>3</sup> Aqui em referência ao Partido Liberal, partido ao qual Bolsonaro está filiado.

A comparação com o conjunto de textos das contas com maior interação nos permite identificar diferenças entre aqueles que têm maior capacidade de interação. O Gráfico 5 abaixo mostra, em primeiro lugar, que o número de *clusters* é maior, chegando a seis *clusters* ao todo, que podem ser reagrupados em três ou até dois grupos. Os grupamentos 1, 4, 5 e 6 subdividem partes do corpus textual mais próximo de uma visão positiva da reunião. Já os *clusters* 2 e 3 indicam especificidades de léxicos mais críticos à reunião. Aqui, os *clusters* têm tamanhos mais semelhantes, variando de 12% a 20% do total do corpus. Ademais, assim como no conjunto total de textos, não há grandes marcadores léxicos que destaquem determinado *cluster* em relação aos demais.

Gráfico 5 - *Clusters* por CHD para textos do percentil superior de interações



Fonte: Elaboração própria.

A abordagem macro da análise de conteúdo léxica não tem o objetivo de particularizar os resultados até as contas ou autores. Pelo menos não em uma fase inicial. É possível, a partir dos termos encontrados nos *clusters*, classificar as contas e identificar os autores que integram majoritariamente cada grupamento, mas, neste caso, a classificação seria ponto de chegada e não de partida. Essa é uma das principais diferenças entre a abordagem micro e macro. A análise léxica permite identificar quão heterogênea, desigual, extremista ou homogênea, distribuída em proporções similares e tendendo ao centro é o conjunto de textos publicados em torno de um tema nas redes sociais. É o formato geral da discussão que interessa aqui e não as características

dos que estão discutindo. Por fim, no próximo tópico apresentamos outra abordagem macro da análise de discussões no Twitter, que é a dinâmica ao longo do tempo.

### Unidade macro – série temporal para descrever a sobrevivência da discussão

A última unidade de análise macro apresentada aqui é a mais externa. Ela não trata dos tuítes, das redes de conexões formadas ou dos usuários. Ela se ocupa da dinâmica de atenção destinada ao tema ao longo do tempo. Assim como a análise de redes, as séries temporais formam um campo específico dos estudos de fenômenos empíricos que se sucedem ao longo do tempo. Aqui, o objetivo é identificar fatores externos que influenciaram na direção da curva, aumentando ou diminuindo o número de manifestações sobre o tema em determinado momento do tempo. O objetivo é descrever a dinâmica da série e não as características individuais das manifestações. Para o caso da reunião com os embaixadores, procuramos saber quando o tema entrou na rede do Twitter e como foi a dinâmica de atenção destinada a ele. Para isso, agregamos os tuítes por hora, entre o primeiro momento em que o tema entrou no Twitter, que foi em 7 de julho de 2022, até o dia 29 de julho, quando ele praticamente desapareceu.

A reunião aconteceu no dia 18 de julho, às 14h. Mas, na noite do dia 7 de julho, às 20h, aparece um primeiro tuíte sobre o tema, informando que o presidente havia anunciado em sua “live” semanal que faria uma reunião com embaixadores no Brasil para falar sobre as falhas do sistema de votação do país. Para identificar a dinâmica temporal, dividimos o período em dois momentos, antes e depois da reunião. Além disso, também dividimos o dia em quatro períodos distintos: madrugada, entre 0h e 6h; manhã, das 7h às 12h; tarde, das 13h às 18h; e noite, das 19h às 24h. Isso porque se espera que exista variação de atividade não apenas entre os dias, mas também entre períodos dentro do dia. A Tabela 3 a seguir mostra as médias de tuítes por hora antes e depois da reunião e por período do dia.

Tabela 3 – Média de tuítes por hora por período de tempo

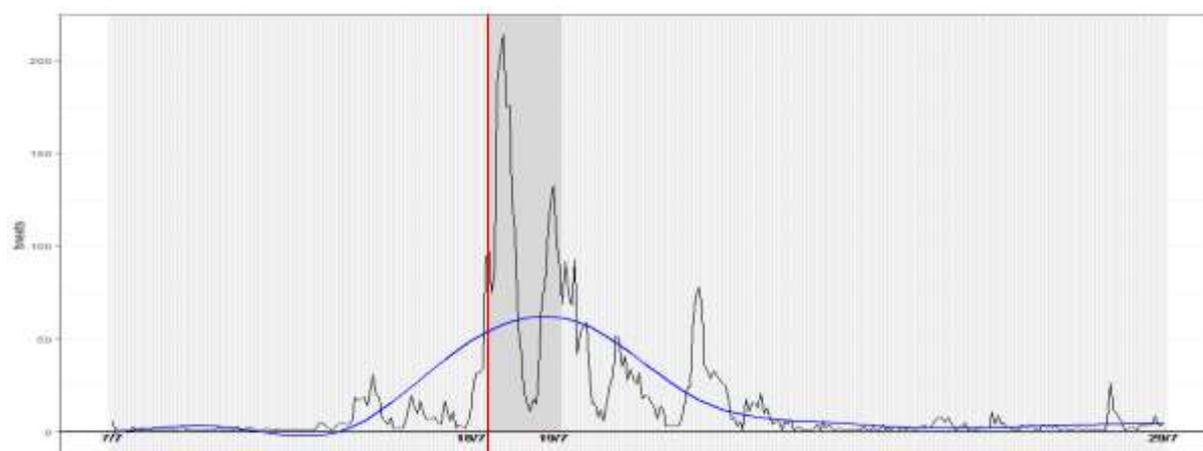
Período de análise	Período do dia	Média	Média Geral
Antes da reunião 7/7 20h a 18/7 14h	Manhã	8,8	6,8
	Tarde	7,5	
	Noite	5,8	
	Madrugada	2,4	
24h após a reunião			94,0
Depois da reunião 18/7 15h a 29/7 14h	Manhã	19,2	20,5
	Tarde	28,3	
	Noite	22,6	
	Madrugada	6,9	

Fonte: Elaboração própria.

A média entre 7 e 18 de julho foi de 6,88 tuítes por hora sobre a reunião, com variação entre 8,78 de média no período da manhã e 2,43 de madrugada. Após a reunião, entre 18 e 29 de julho, a média por hora sobe para 20,5 tuítes, com variação entre 28,37 no período da tarde e 6,95 de madrugada. O evento multiplicou por três a média de tuítes por hora. Mas, a maior concentração foi nas primeiras 24 horas após a reunião, que apresentou uma média de 94 tuítes por hora.

O Gráfico 6 a seguir mostra a série temporal entre 7 e 29 de julho. A linha vermelha marca a hora em que a reunião aconteceu. A área sombreada indica as 24 horas posteriores ao evento e a linha azul é a tendência suavizada das ocorrências por hora. Há duas informações importantes na imagem. A primeira é a sazonalidade, já apontada na tabela com as médias. Os períodos do dia importam para análise da presença do tema na rede. Dentro do dia, durante a noite e madrugada o número de tuítes é reduzido, para voltar a subir na manhã e tarde seguintes. O comportamento tende a se repetir com menor intensidade com a passagem do tempo. Fica evidenciado que o período de sobrevivência do evento no Twitter foi de 24 horas após a reunião, com menos intensidade no período da madrugada, embora essa madrugada tenha apresentado um número maior de tuítes do que quase todos os períodos dos dias anteriores ao evento.

Gráfico 6 – Série Temporal – Tuítes por hora



Fonte: Elaboração própria.

A análise de séries temporais permite verificar a dinâmica da presença de determinado tema nas redes sociais, além de identificar o tempo de “sobrevivência” dele, por meio da atenção dada pelos usuários da rede. Assim, é possível comparar a sobrevivência de diferentes temas, quando eles surgem, qual o período de duração e quando tendem a desaparecer. A sazonalidade é um traço permanente das conversações nas redes sociais, embora se saiba da circulação de conteúdos não orgânicos. O fato é que em muitos casos, as primeiras 24 horas são as mais

importantes para qualquer tema nas RSO, e não foi diferente no caso da reunião com os embaixadores.

### **Considerações finais**

Defendendo que a unidade de análise é tão importante para a compreensão do fenômeno empírico quanto a clara identificação dos objetivos a serem cumpridos, este estudo organiza cada conjunto de técnicas disponíveis para a análise de RSO em função de seus objetivos e da unidade de análise. Cada conjunto de técnicas foi descrito a partir de suas características e objetivos, exemplificado a partir de seu uso na academia brasileira com base em uma seleta coleção de artigos, e aplicado à análise de uma mesma base de dados sobre um evento ocorrido em julho de 2022. Essas três etapas permitiram refletir sobre a utilização das técnicas e o que cada uma revela em termos de contribuição ao campo de conhecimento.

Tanto nos artigos analisados quanto na aplicação demonstrada, fica claro que a análise no nível micro tem dois objetivos principais, que são complementares. O primeiro deles é saber quem fala, ou seja, quais contas têm mais interesse em participar das discussões sobre determinado tema. Isso é medido através do número de intervenções no período. No caso aplicado, descobrimos que há uma categoria de usuários que busca participar mais, jornalistas e imprensa. Já o segundo objetivo, que é identificar a notoriedade dos participantes, é medido pela atividade gerada com outros usuários a partir de interações digitais, revelando que o mesmo grupo apontado pelo alto número de intervenções não necessariamente tem notoriedade entre os participantes. As categorias que tiveram mais notoriedade no caso analisado foram políticos e militantes. Os artigos de Santos *et al.*, (2020, 2021) são também ilustrativos do nível micro de análise.

O nível macro de análise, apresentado aqui a partir de três objetivos distintos, permite descrever as relações entre usuários não mais pelas vontades individuais ou notoriedade dos participantes, mas sim de forma agregada. O primeiro objetivo macro descrito, que é identificar como se organiza o que é dito, é feito a partir da análise de redes. Essa técnica, empregada por Soares (2019), Segurado, Chicarino e Malina (2017) e Recuero, Zago e Soares (2017), foi aplicada na base de dados utilizada neste artigo para identificar as principais características da rede de interações formadas nas conversações sobre a reunião com os embaixadores. Aqui, foi possível encontrar outro tipo de conta que não havia aparecido antes. Esse tipo não tem vontade de participar e também não se destacou entre as contas com notoriedade, porém, aparece muito

citada nos textos de tuítes: as instituições de Estado. Ou seja, embora as redes sejam formadas em torno de políticos e meios de imprensa, elas também contam com a participação, via citações, de instituições de Estado.

O segundo objetivo macro é identificar quais são as principais características dos textos publicados nas discussões sobre o tema, a partir da análise de conteúdo ou temática, utilizada por Soares (2020), Ituassu *et al.* (2019), Dienstbach *et al.* (2021) e Neves (2017). No caso ilustrado, a análise de conteúdo automatizada permitiu identificar grande homogeneidade nos tuítes sobre o tema e, em geral, os textos formaram dois *clusters*, um mais favorável e outro crítico à reunião. Porém, quando olhamos apenas para os textos que apresentaram maiores interações – aqueles com notoriedade – percebemos que há maior complexidade e heterogeneidade no uso de léxicos.

Por fim, demonstramos como pode ser importante a análise na dimensão mais externa do nível macro, que não se preocupa com quem publica, quais são as reações geradas, como se comportam as redes de interações ou que léxicos caracterizam a discussão. É a dimensão temporal da presença e intensidade do tema que é enfatizada, tipo de análise utilizada por Santos Jr. (2021). No caso da reunião dos embaixadores, o uso desta técnica permitiu identificar que o tempo de sobrevivência do tema no Twitter foi de 24 horas, com alta intensidade, embora tenha tido alguma presença desde uma semana antes e duas semanas depois da reunião.

Este *paper* exploratório teve o objetivo de demonstrar como é importante atentar para o nível de análise, se micro ou macro, relacionando-o com os objetivos da pesquisa, para evitar a busca de respostas certas para perguntas de pesquisa erradas ou perguntas certas que são seguidas de respostas erradas.

## Referências Bibliográficas

- ALMADA, Maria P. *et al.* Democracia digital no Brasil. **Matrizes**, São Paulo, v. 13, p. 161-181, 2019.
- BOURDIEU, Pierre. **Homo academicus**. Stanford: Stanford University Press, 1988.
- DIENSTBACH, Dalby *et al.* Kit Covid entre aspas: a arquitetura discursiva da desinformação nas mídias sociais. *In: ENCONTRO ANUAL DA ANPOCS*, 45., 2021. **Anais do 45º Encontro Anual da Anpocs**, 2021. p. 1-17.
- ITUASSU, Arthur *et al.* De Donald Trump a Jair Bolsonaro: democracia e comunicação política digital nas eleições de 2016, nos Estados Unidos, e 2018, no Brasil. *In: ENCONTRO COMPOLÍTICA*, 8., 2019, Brasília. **Anais do 8º Encontro da Compolítica**. Brasília, UnB, 2019. p. 1-25.
- MENDONÇA, Ricardo F.; PEREIRA, Marcus A.; FILGUEIRAS, Fernando. (org.). **Democracia Digital: Publicidade, Instituições e Confronto Político**. Belo Horizonte: Editora UFMG, 2016.
- NEVES, Gláucia S. O Caso Hashtag #primeiroassédio: narrativas testemunhais como semântica coletiva nas lutas por reconhecimento no contexto digital. *In: ENCONTRO COMPOLÍTICA*, 7., 2017, Porto Alegre. **Anais do 7º Encontro da Compolítica**. Porto Alegre, UFRGS, 2017. p. 1-23.

RECUERO, Raquel; ZAGO, Gabriela; SOARES, Felipe Bonow. Mídia social e filtros-bolha nas conversações políticas no twitter. *In: ENCONTRO ANUAL DA COMPÓS*, 26., 2017, São Paulo. **Anais do 26º Encontro Anual da Compós**. São Paulo: Faculdade Cásper Líbero, 2017. p. 1-27.

SAMPAIO, Rafael Cardoso; BRAGATTO, Rachel C.; NICOLÁS, Maria Alejandra construção do campo de internet e política: análise dos artigos brasileiros apresentados entre 2000 e 2014. **Revista Brasileira de Ciência Política**, Brasília, v. 4, n. 21, p. 285- 320, 2016.

SAMPAIO, Rafael Cardoso *et al.* O campo da democracia digital brasileira: uma análise cientométrica de artigos publicados entre 1999 e 2020. **Revista Brasileira de Ciência Política**, Brasília, v. 37, p. 1-38, 2022.

SANTOS, Marcelo B. P. *et al.* Os usos do twitter pela família Bolsonaro e o campo da direita brasileira. *In: ENCONTRO ANUAL DA ANPOCS*, 44., 2020. **Anais do 44º Encontro Anual da Anpocs**, 2020. p. 1-24.

SANTOS, Marcelo B. P. *et al.* Twitter e a politização da vacina covid-19 no campo da direita. *In: ENCONTRO ANUAL DA ANPOCS*, 45., 2021. **Anais do 45º Encontro Anual da Anpocs**, 2021. p. 1-36.

SANTOS JR., Marcelo A. Fabricando e coordenando a desinformação: análise da circulação multiplataformas da #LaranjaldoBoulos. *In: ENCONTRO COMPOLÍTICA*, 9., 2021. **Anais do 9º Encontro da Compolítica**, 2021. p. 1-30.

SEGURADO, Rosemary; CHICARINO, Tatiana; MALINA, Pedro. Metodologia de monitoramento de redes digitais aplicada ao Brasil: um estudo sobre política e eleições. *In: ENCONTRO ANUAL DA ANPOCS*, 41., 2017, Caxambu, MG. **Anais do 41º Encontro Anual da Anpocs**. Caxambu, MG: Hotel Glória, 2017. p. 1-33.

SOARES, Felipe B. Esfera pública e desinformação: estratégias de circulação e legitimação da desinformação. *In: ENCONTRO ANUAL DA COMPÓS*, 28., 2019, Porto Alegre. **Anais do 28º Encontro Anual da Compós**. Porto Alegre: PUC, 2019. p. 1-22.

SOARES, Felipe B. Circulação de informação no twitter: como líderes de opinião ressignificam as notícias. *In: ENCONTRO ANUAL DA COMPÓS*, 29., 2020, Campo Grande. **Anais do 29º Encontro Anual da Compós**. Campo Grande: UFMS, 2020. p. 1-24.

Recebido em: 04/03/2024

Aceito em: 10/09/2024