

Uma Análise de Tweets sobre o Coronavírus

Francisco Romes da Silva Filho¹, Emanuel Ferreira Coutinho¹

¹Universidade Federal do Ceará (UFC) – Quixadá – CE – Brasil

romes221b@gmail.com, emanuel.coutinho@ufc.br

Abstract. *With the growth of social media platforms, users started to publish thoughts and opinions on the most diverse subjects. In an environment of social networks, it is natural that there are discussions on current issues, such as elections, festivities and disasters. Currently, a subject that is widely discussed is the pandemic of COVID-19. Among the various services available on social media, one of them is microblogs, taking Twitter as an example. In this context of social networks, the objective of this work is to present an evolution of Coronavirus according to the posts of Twitter users in a period of time. As great results of this work, it was found that the words isolation, quarantine and cases were always present in queries to tweets when searching for Coronavirus and flexibilization, indicating a concern of society.*

Resumo. *Com o crescimento de plataformas de mídias sociais, os usuários passaram a publicar pensamentos e opiniões sobre os mais diversos assuntos. Em um ambiente de redes sociais, é natural que existam discussões sobre os temas da atualidade, como eleições, festividades e desastres. Atualmente, um assunto que está sendo bastante comentado é a pandemia do COVID-19. Dentre os diversos serviços disponíveis nas mídias sociais, um deles é o de microblogs, tendo como exemplo o Twitter. Nesse contexto de redes sociais, o objetivo deste trabalho é apresentar uma evolução do Coronavírus, no contexto brasileiro, a partir da análise das postagens dos usuários brasileiros do Twitter. Como exemplo de resultado relevante deste trabalho, verificou-se que as palavras isolamento, quarentena e casos sempre estiveram presentes nas consultas aos tweets quando se pesquisava por Coronavírus e flexibilização, indicando uma preocupação da sociedade.*

1. Introdução

Os últimos anos testemunharam o rápido crescimento de plataformas de mídia social (por exemplo, Facebook, Twitter e Google+) nos quais os usuários podem publicar pensamentos e opiniões sobre qualquer tópico [Giachanou e Crestani 2016]. A crescente popularidade das plataformas de mídia social mudou a *web* de um repositório estático de informações para um fórum dinâmico com informações em constante mudança. As plataformas de mídia social deram às pessoas a capacidade de expressar e compartilhar seus pensamentos e opiniões na *web* de uma maneira muito simples. Sendo assim, as redes sociais *online* são hoje um dos principais ambientes para debate, discussão e troca de informações entre as pessoas [Caetano et al. 2017].

Nesse ambiente de redes e mídias sociais é natural que existam discussões sobre os temas da atualidade, como eleições, festividades, desastres, dentre outros. E atualmente, um assunto que está sendo bastante comentado é a pandemia do COVID-19.

O COVID-19, sigla para “Doença de Coronavírus-2019”, é uma doença respiratória causada pela síndrome respiratória aguda grave Coronavírus-2 (SARS-CoV-2), um vírus contagioso pertencente a uma família de vírus de RNA positivo de cadeia simples conhecido como *coronaviridae* [Chamola et al. 2020].

Desde dezembro de 2019, a epidemia da doença de Coronavírus (COVID-19) varreu o mundo, causando um número e carga significativa de hospitalizações [Zhu et al. 2020][Wu e McGoogan 2020]. Segundo a Organização Mundial da Saúde - OMS (*World Health Organization* - WHO), a situação de surto da doença de Coronavírus (COVID-19) em 05 de maio de 2020 foi de 3.595.662 casos confirmados, 247.652 mortes confirmadas e 215 países, áreas ou territórios com casos [World Health Organization 2020].

As mídias sociais são um canal significativo para notícias e informações no ambiente de mídia moderno, com uma em cada três pessoas no mundo envolvidas em mídias sociais e dois terços dos usuários da Internet [Ortiz-Ospina 2019]. Desde o início da epidemia da doença de Coronavírus (COVID-19), a desinformação está se espalhando pelas mídias tradicionais e sociais em um ritmo acelerado [Kouzy et al. 2020]. Um exemplo de pesquisa nesse contexto foi desenvolvida em [Kouzy et al. 2020] com a análise da magnitude da desinformação que está sendo disseminada no Twitter em relação à epidemia de Coronavírus. Conforme [Pulido et al. 2020], pesquisas apontaram que, nas mídias sociais, a falsidade é compartilhada muito mais do que informações baseadas em evidências.

Um dos serviços disponível nas mídias sociais é o de *microblogging* ou microblogs. O microblog é um serviço de rede com o qual os usuários podem compartilhar mensagens, *links* para sites externos, imagens ou vídeos visíveis para os usuários inscritos no serviço [Giachanou e Crestani 2016]. As mensagens postadas em microblogs são curtas em contraste com os blogs tradicionais. Um dos microblogs mais populares é o Twitter, lançado em 2006, e desde então atrai um grande número de usuários. Devido ao fato de fornecer uma maneira fácil de acessar e baixar postagens publicadas, o Twitter é considerado um dos maiores conjuntos de dados de conteúdo gerado pelo usuário.

Segundo uma estatística sobre as redes sociais mais populares do mundo em abril de 2020, classificadas por número de usuários ativos¹, o número de usuários do Twitter estava em torno de 386 milhões de usuários. Com esta quantidade de usuários ativos, espera-se que o alcance desta rede social seja grande, com impacto significativo.

O objetivo deste trabalho é apresentar uma evolução do Coronavírus, no contexto brasileiro, a partir da análise das postagens dos usuários brasileiros do Twitter. E como as redes sociais são muito presentes no cotidiano popular, este trabalho evidenciou os efeitos da pandemia mediante as discussões no Twitter. Para isto, o software de análise qualitativa, Atlas.ti² foi utilizado para extrair dados do Twitter e produzir nuvens de palavras para a realização da análise do conteúdo das discussões presentes nos *tweets*.

Este trabalho está dividido, além desta introdução, nas seguintes seções: a Seção 2 apresenta alguns trabalhos relacionados ao tema; a metodologia aplicada ao trabalho é apresentada na Seção 3; na Seção 4, os resultados e análises são apresentados; e por fim,

¹Global social media ranking 2019 — Statista - <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

²ATLAS.ti - <https://atlasti.com/>

a Seção 5 descreve as conclusões do trabalho.

2. Trabalhos Relacionados

Como as pesquisas sobre a pandemia atual do Coronavírus estão na maioria dos casos ainda em andamento, há certa dificuldade na busca por trabalhos científicos já publicados. Muitos artigos estão na situação de disponíveis para a comunidade acadêmica e científica, mas ainda sem terem passado por um processo de revisão, como por exemplo [Singh et al. 2020], [Kleinberg et al. 2020] e [Alshaabi et al. 2020]. Entretanto, diversos *datasets* estão sendo disponibilizados com dados sobre o Coronavírus. Alguns deles foram extraídos especificamente do Twitter, como em [Chen et al. 2020], [Alqurashi et al. 2020] e [Lopez et al. 2020].

Kouzy et al. (2020) conduziram uma pesquisa no Twitter sobre o COVID-19, incluindo 673 *tweets*. Dessa quantidade, 548 *tweets* (81,4%) incluíram informações genuínas referentes à epidemia de COVID-19. A pesquisa concluiu que a desinformação médica e conteúdo não verificável pertencente à epidemia global de COVID-19 estão sendo propagados a um ritmo alarmante nas mídias sociais. Foi indicada uma quantificação antecipada da magnitude da disseminação de informações erradas, com destaque à importância de intervenções precoces, para conter esse fenômeno que põe em risco a segurança pública em um momento em que a conscientização e as ações preventivas apropriadas são fundamentais [Kouzy et al. 2020].

Pulido et al. (2020) apresentaram um estudo sobre o tipo de *tweets* que circulavam no Twitter em torno do surto de COVID-19 por dois dias, com o objetivo de analisar como as informações falsas e verdadeiras foram compartilhadas. Para isto, 1000 *tweets* foram analisados, e os resultados mostraram que informações falsas são mais tuitadas, porém menos retuitadas que evidências científicas ou *tweets* de verificação de fatos, enquanto evidências baseadas em ciência e *tweets* de verificação de fatos capturam mais envolvimento do que meros fatos. Esses achados trazem informações relevantes para as políticas de saúde pública [Pulido et al. 2020].

A pesquisa proposta neste trabalho é semelhante aos trabalhos relacionados no sentido em que se busca identificar informações relevantes em *tweets* de usuários que comentaram sobre o Coronavírus durante um período de tempo, assim como nos outros trabalhos. Contudo, com o objetivo de apresentar uma evolução do Coronavírus no Brasil.

3. Metodologia

O objetivo deste trabalho é apresentar uma evolução do Coronavírus, no contexto brasileiro, a partir da análise das postagens dos usuários brasileiros do Twitter. O Twitter é um meio de comunicação bastante utilizado para discutir assuntos com caráter importante e urgente. Portanto, é possível afirmar que o Coronavírus, dado seu impacto mundial, é uma questão bastante discutida nesta rede social pelos seus usuários.

Para compartilhar informações de forma mais ampla, a rede social fornece a empresas, desenvolvedores e usuários acesso programático a dados do Twitter através das APIs (Interfaces de Programação de Aplicações) da empresa. Uma aplicação que faz uso das APIs do Twitter é o Atlas.ti, uma ferramenta de análise qualitativa, que foi utilizada neste estudo para o atendimento do objetivo. O software permite a pesquisadores produzir

vários tipos de análise de dados, desde texto a imagens. E a integração da ferramenta com a API do Twitter possibilita importar e analisar os dados da rede social.

A operação de importar dados é feita com a indicação dos parâmetros (palavras-chaves), a quantidade máxima de *tweets* desejada, tipos de *tweets* (recentes e/ou populares), inclusão de imagens e/ou *retweets* (republicação de um mesmo *tweet*), ou ainda, somente *tweets* (texto). A consulta foi elaborada solicitando duas palavras-chaves: “Coronavírus” e “Flexibilização”. Portanto, utilizando o operador aspas duplas, as publicações encontradas continham as duas expressões entre aspas ou variações, como o exato termo sem a presença de acentuação. Ainda apareciam resultados que fossem classificados como *tweets* do tópico, classificação feita pelo Twitter. Os resultados foram selecionados de forma misturada, ou seja, apresentando *tweets* populares e recentes, e com o limite máximo de 1000 *tweets* por busca, excluindo *retweets*, impedindo repetições e imagens que não foram analisadas neste estudo.

A definição dos parâmetros foi feita após alguns testes. O termo “Coronavírus” foi utilizado em todos os testes, no teste que esteve sozinho na consulta apresentou vários memes - publicações com o objetivo de divertir o interlocutor - como resultados, o que atrapalha na análise. Então, foi feita uma busca de um outro parâmetro para a consulta. No teste, em que foi utilizado a palavra “Covid” a maioria dos resultados eram de *tweets* estrangeiros, ou seja, se distanciam do contexto brasileiro. “Isolamento” e “Quarentena” também foram testadas, porém, os resultados apresentavam ainda uma grande quantidade de memes. Outras palavras foram consideradas, contudo, apresentavam pequena quantidade de resultados ou direcionam bastante a discussão, limitando as possibilidades da análise.

Tendo em vista, a escolha de “Coronavírus” e “Flexibilização” se deu devido ao fato que o primeiro termo é a temática principal do trabalho e também um filtro necessário para encontrar publicações que abordam a doença. Já o segundo termo é uma palavra que se apresenta bastante em discussões sobre ações e impacto do vírus, ou seja, a presença da palavra é apenas para selecionar *tweets* que disponham de dados, ações, medidas, notícias e opiniões sobre o tema. E ainda, assim, eliminando a possibilidade de ocorrência de resultados que apresentem publicações com teor que fuja da importância e do contexto da temática, como memes.

A análise dos resultados após a operação de importação de dados do Twitter foi realizada por meio de uma nuvem de palavras, palavras que se destacam com maiores ocorrências nos textos presentes. Para isso, era feito uma exclusão prévia de símbolos especiais e palavras, como artigos, conjunções, preposições, etc, para facilitar a análise. E ainda foi preciso definir um número mínimo de ocorrências dessas palavras, a frequência da palavra tinha que ser maior ou igual a 5.

Por conseguinte, era feito a visualização da nuvem observando as palavras, seus contextos e tendências. A parte inicial da análise foi feita identificando as palavras presentes, as suas frequências e quais as palavras mais frequentes, com gráficos de barras. Em seguida, a observação da nuvem foi realizada identificando os contextos das palavras presentes, os padrões e os tipos de situações, analisando o conteúdo do *tweet*. A partir disso, as palavras foram separadas em categorias que representassem as áreas discutidas.

O software consegue retornar dados de até 7 dias antes do dia da busca realizada.

Dado esse intervalo, foi definido neste trabalho analisar as consultas de 5 dias, a cada 3 dias, pelo fato da grande quantidade de informações novas e o caráter urgente presente em redes sociais, verificando os padrões, novos contextos e abordagens.

Para complementar a análise dos dados, consultou-se os dados de outras ferramenta, a *Reports on the COVID-19 Outbreak*³. Ela apresenta um estudo sobre as curvas de crescimento do COVID-19, com a possibilidade de geração de gráficos relacionados a países e aos estados do Brasil. Os dados desta ferramenta são atualizados diariamente a partir das bases da Johns Hopkins University e do Ministério da Saúde.

4. Resultados e Análises

As subseções a seguir apresentam os resultados das consultas sobre o Twitter. Os dias em que a busca foi aplicada foram: 1, 4, 7, 10 e 13 de maio de 2020. Uma consulta extra foi realizada, para servir como referência no tempo dos *tweets*, no dia 26 de fevereiro de 2020, correspondendo ao dia com o primeiro caso de Coronavírus (COVID-19) registrado no Brasil.

4.1. Resultados

Segundo a base de dados sobre COVID-19 disponibilizada pela Johns Hopkins University⁴, o primeiro caso de Coronavírus (COVID-19) registrado no Brasil foi em 26 de fevereiro de 2020. Apenas para fim de referência, geramos a nuvem de palavra (Figura 1(a)) e o gráfico de barras com as 10 mais frequentes palavras (Figura 1(b)) para o dia 26 de fevereiro de 2020, obtidos a partir de 74 *tweets*. Nessa geração em particular, utilizar a palavra “flexibilização” não era relevante, pois não era discutido na época. Percebe-se que a quantidade de palavras da nuvem é baixa em relação à quantidade de menções à “Coronavírus”. Isso se deve ao início dos registros de casos no Brasil, e ainda por não ter ocorrido mortes registradas por COVID-19 até então. Também até aquele momento não havia menções relevantes sobre ações de combate à pandemia, como isolamento e distanciamento social.

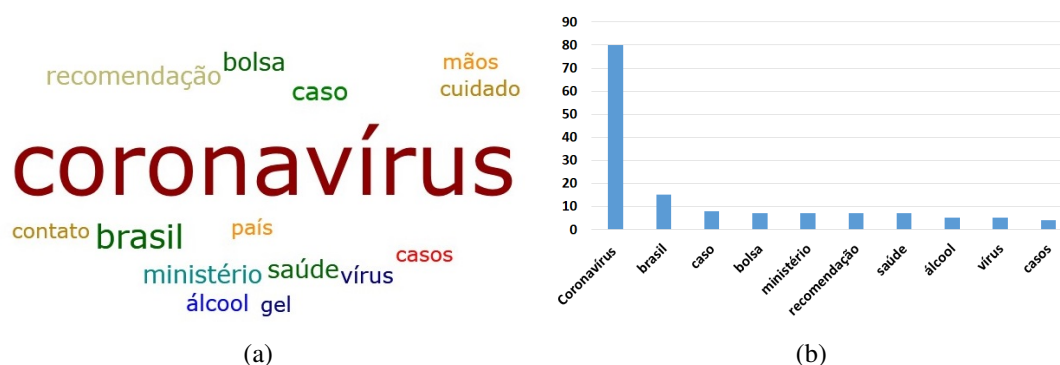


Figura 1. Nuvem de palavras geradas para referência. (a) Nuvem de palavras para o dia 26/02/2020 e (b) Gráfico de barras com as 10 mais frequentes palavras para o dia 26/02/2020

³*Reports on the COVID-19 Outbreak* - <https://sistemas.quixada.ufc.br/covid/>

⁴COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University - <https://github.com/CSSEGISandData/COVID-19>

Com mais de um mês após a data do primeiro registro de contágio de COVID-19 no Brasil, coletamos algumas amostras de *tweets*, de dias pré-determinados. A quantidade de *tweets* coletados foi 272. A primeira coleta foi no dia 01 de maio de 2020. A Figura 2(a) exibe uma nuvem de palavras, e a Figura 2(b) exibe um gráfico de barras com as 10 mais frequentes palavras. Destaca-se a palavra “flexibilização” com mais menções que o próprio “Coronavírus”. A palavra “pandemia” também está entre as mais citadas. Palavras de ações de combate à pandemia estão em destaque: isolamento e quarentena.

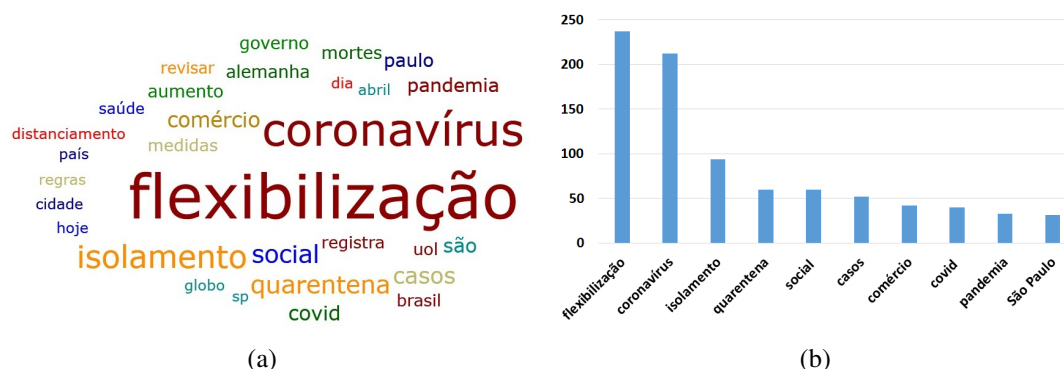


Figura 2. Coleta para o dia 1. (a) Nuvem de palavras para o dia 01/05/2020 e (b) Gráfico de barras com as 10 mais frequentes palavras para o dia 01/05/2020

A segunda coleta ocorreu o dia 04 de maio de 2020, com 84 *tweets* coletados. A Figura 3(a) exibe uma nuvem de palavras para o dia 04 de maio de 2020, e a Figura 3(b) exibe um gráfico de barras com as 10 mais frequentes palavras para o dia 04 de maio de 2020. Novamente, a palavra “flexibilização” obteve mais menções que o próprio “Coronavírus”. A palavra “decreto” surgiu entre as mais citadas. Palavras de ações de combate à pandemia continuam em destaque: isolamento e quarentena.

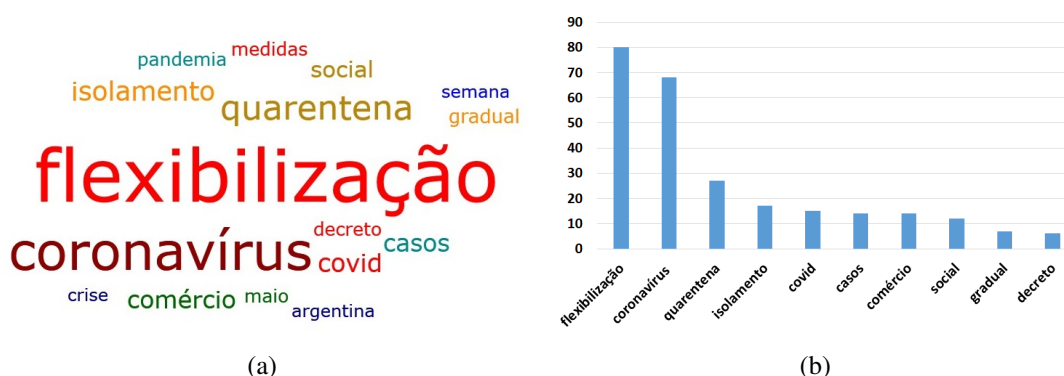


Figura 3. Coleta para o dia 2. (a) Nuvem de palavras para o dia 04/05/2020 e (b) Gráfico de barras com as 10 mais frequentes palavras para os dia 04/05/2020

A terceira coleta ocorreu o dia 07 de maio de 2020. A quantidade de *tweets* coletados foi 97. A Figura 4(a) exibe uma nuvem de palavras para o dia 07 de maio de 2020. A Figura 4(b) exibe um gráfico de barras com as 10 mais frequentes palavras para o dia 07 de maio de 2020. Mais uma vez, a palavra “flexibilização” obteve mais menções que o próprio “Coronavírus”. A palavra “aumento” surgiu entre as mais citadas. Palavras de ações de combate à pandemia continuam em destaque: isolamento e quarentena.

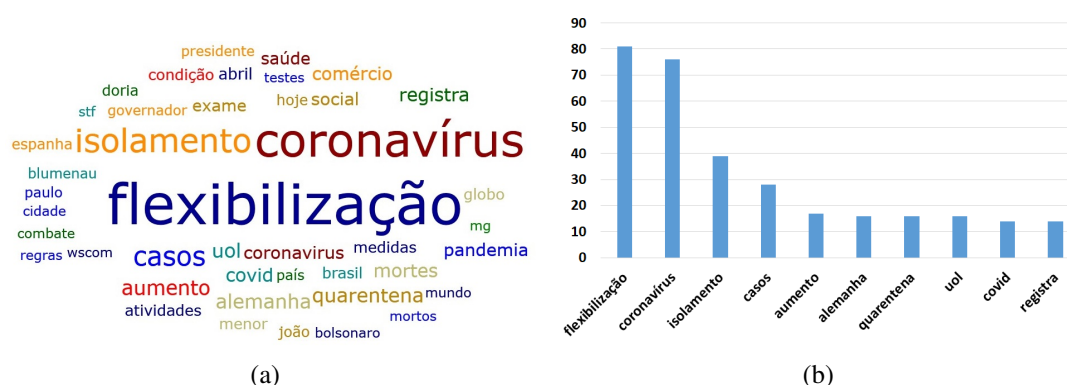


Figura 4. Coleta para o dia 3. (a) Nuvem de palavras para o dia 07/05/2020 e (b) Gráfico de barras com as 10 mais frequentes palavras para os dia 07/05/2020

A quarta coleta ocorreu o dia 10 de maio de 2020, resultando em 185 *tweets* coletados. A Figura 5(a) exibe uma nuvem de palavras para o dia 10 de maio de 2020. A Figura 5(b) exibe um gráfico de barras com as 10 mais frequentes palavras para o dia 10 de maio de 2020. A palavra “flexibilização” obteve mais menções que o próprio “Coronavírus”. As palavras “comércio” e “pandemia” apareceram entre as mais citadas. Palavras de ações de combate à pandemia permanecem em destaque: isolamento e quarentena.

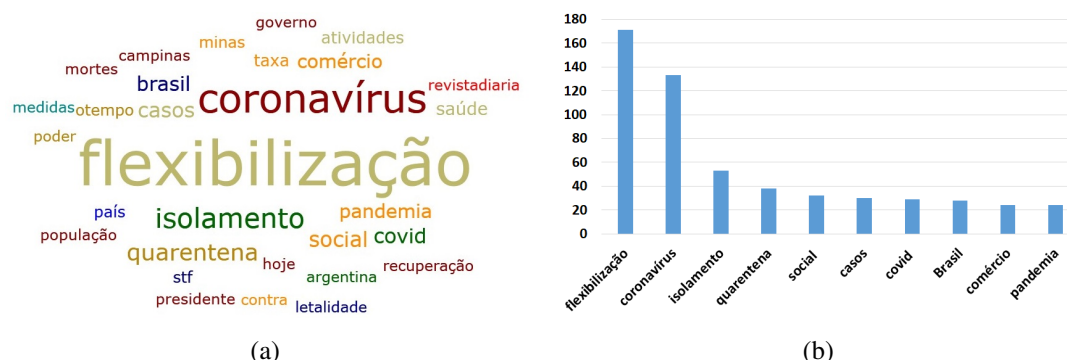


Figura 5. Coleta para o dia 4. (a) Nuvem de palavras para o dia 10/05/2020 e (b) Gráfico de barras com as 10 mais frequentes palavras para os dia 10/05/2020

A quinta e última coleta ocorreu o dia 13 de maio de 2020, resultando em 220 *tweets* coletados. A Figura 6(a) exibe uma nuvem de palavras para o dia 13 de maio de 2020. A Figura 6(b) exibe um gráfico de barras com as 10 mais frequentes palavras para o dia 13 de maio de 2020. A palavra “flexibilização” continuou com mais menções que o próprio “Coronavírus”. As palavras “comércio” e “pandemia” permaneceram entre as mais citadas. Por fim, palavras de ações de combate à pandemia continuam em destaque: isolamento e quarentena.

4.2. Análises e Discussões

Apesar das palavras “flexibilização” e “Coronavírus” estarem presentes em todas as consultas feitas ao Twitter, “flexibilização” sempre esteve com mais ocorrências que “Coronavírus”. Isto se deve ao fato que no período de coleta (01 de maio de 2020 a 13 de

- Instituições: governo (estaduais e federais), STF. Em geral, ações dos órgãos brasileiros estiveram em pauta.

Algumas observações podem ser destacadas conforme o resultado das consultas dos *tweets* em relação à presença de países nos *tweets* que não eram o Brasil: Alemanha, Argentina e Espanha. Considerando os dados disponibilizados pelo site Worldometer⁵ em relação à pandemia, algumas relações foram discutidas. Os dados foram consultados em 18 de maio de 2020. A Alemanha, apesar do número de mortes por 1 milhão de habitantes ser superior, 96 mortes, ao número do Brasil, 76 mortes, o país europeu se encontra em outra fase de enfrentamento do vírus com a sua curva de evolução sendo descendente e já realizando o retorno da suas atividades. E a Argentina, como sendo um país vizinho, sul-americano, que tem bons índices no combate ao vírus. Em comparação com o Brasil, pode ser observado a quantidade de mortes por 1 milhão de habitantes, Argentina com 8 mortes e o Brasil com 76 mortes, até a data 18 de maio de 2020, ou seja, um número bem inferior. Por último temos a Espanha, quarto país europeu com o maior número de mortes no continente, 27.650 mortes, atrás apenas de Reino Unido (Inglaterra, Escócia, País de Gales e Irlanda do Norte), Itália e França. Portanto, o impacto mundial do vírus se apresenta constante nas discussões, gerando repercussão nos *tweets* com o acompanhamento da evolução em outros países.

E na última coleta, 13 de maio de 2020, surgiu o termo “*lockdown*”, que corresponde a um confinamento, a versão mais rígida do distanciamento social. É uma imposição do Estado que significa bloqueio total. Considerando a evolução do vírus no país, provavelmente esta seja a medida mais discutida para o enfrentamento neste momento, refletido na rede social.

Muitas ferramentas atualmente estão divulgando dados sobre a pandemia. Geralmente são informações de casos confirmados, mortes e casos em que o paciente se recuperou. Esse trabalho utilizou a ferramenta *Reports on the COVID-19 Outbreak* [de Vasconcelos et al. 2020] para obtenção desses dados. A Figura 7 apresenta uma visão dos 10 países com mais casos de COVID-19 confirmados no mundo, sendo gerada diretamente na ferramenta. A imagem aponta o Brasil como o sexto no *ranking*. Observando a curva, percebe-se que o Brasil está próximo a vários países que reportaram casos antes, o que implica que a pandemia foi se ampliando pelo mundo independente de região geográfica. Considerando que muitos países estão desenvolvendo ações de isolamento social como forma de contingência nesta pandemia, os *tweets* indicados pelas busca na pesquisa deste artigo refletiram este aspecto de impacto social também.

Conforme dados obtidos a partir da ferramenta, a Figura 8 apresenta a partir do dia em que foi contabilizado a primeira morte no Brasil (17 de março de 2020) até o dia 13 de maio de 2020 a quantidade de casos confirmados, mortes e pacientes que se recuperaram. A ideia desse gráfico é visualizar a evolução dos casos no Brasil. Percebe-se claramente pela curva que os números estão aumentando. Em todas as nuvens de palavras geradas a partir de maio, a palavra “casos” sempre esteve presente. Essa é uma preocupação pois a contabilização dos casos confirmados e óbitos está aumentando. Ameniza um pouco o fato que muitos casos de recuperação também estão sendo obtidos. Esse aumento nos casos pode também ser um indicativo da preocupação nos *tweets*, principalmente

⁵Reported Cases and Deaths by Country, Territory, or Conveyance - Worldometer - <https://www.worldometers.info/coronavirus/#countries/>

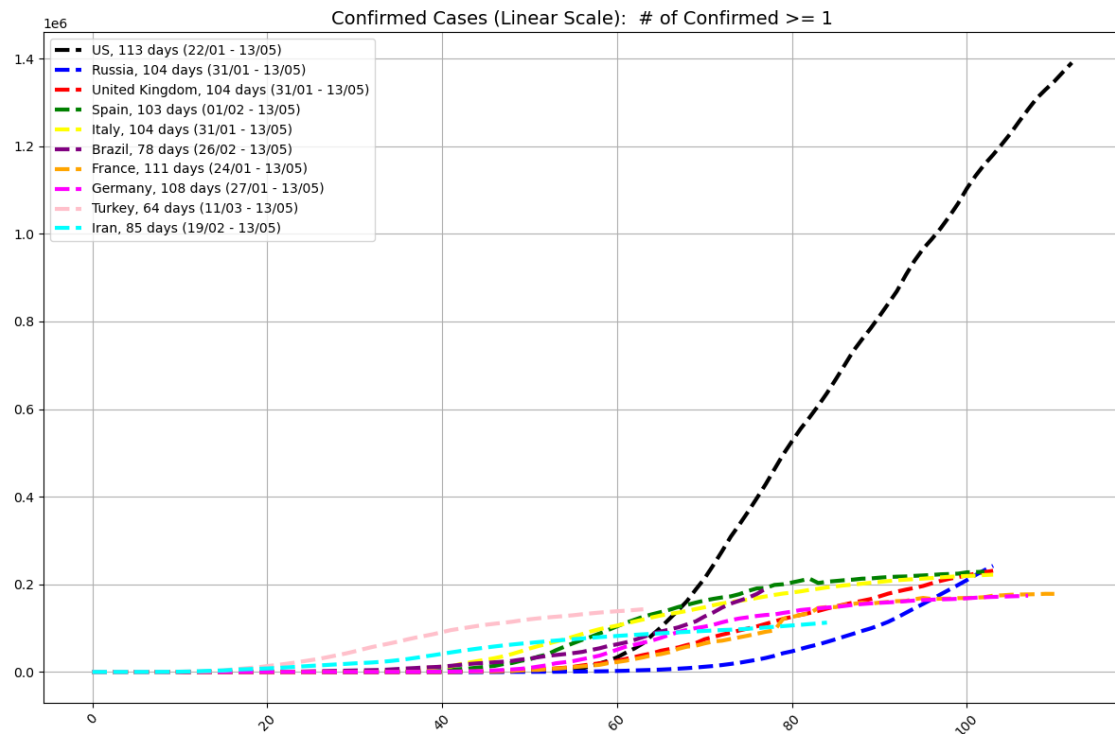


Figura 7. Dez países com maiores quantidades de casos confirmados [de Vasconcelos et al. 2020]

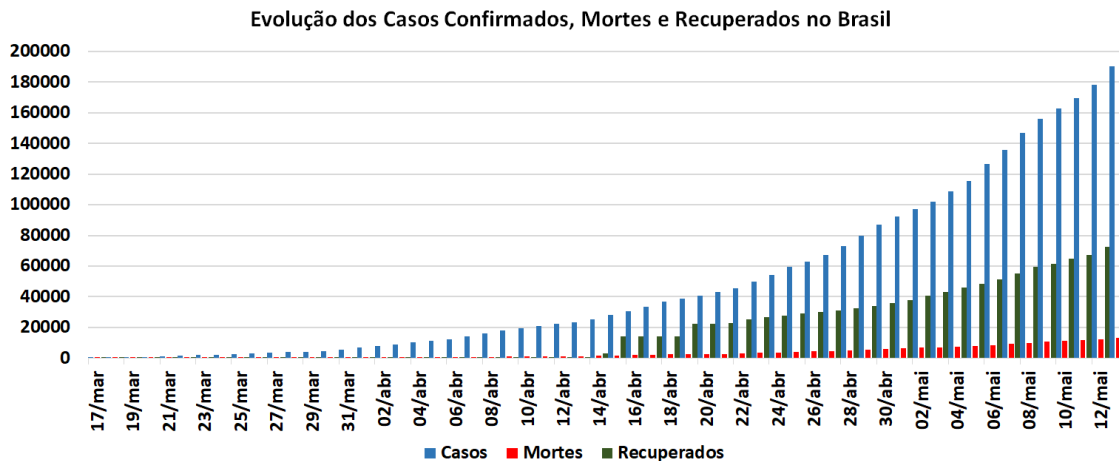


Figura 8. Quantidade de casos confirmados, mortes e pacientes que se recuperaram no Brasil (adaptado dos dados obtidos a partir de [de Vasconcelos et al. 2020])

com comentários sobre *lockdown*, que surgiram apenas na última consulta, e as palavras “isolamento”, “quarentena” e “casos” que surgiram em todas as consultas.

5. Conclusão

Este trabalho procurou apresentar uma evolução do Coronavírus, no Brasil, a partir das postagens dos usuários do Twitter. Para isto, dados dos *tweets* de usuários foram coletados

em alguns dias específicos com o uso das palavras “Coronavírus” e “Flexibilização”, que serviram como um filtro. Os *tweets* foram consolidados em nuvens de palavras e gráficos de barras. Além disso, alguns gráficos e dados de outras fontes foram utilizados para expandir as discussões.

A partir dos resultados, observamos a capacidade do Twitter de ser um termômetro para um momento de crise, como a da pandemia do Coronavírus. Percebemos pelos resultados que palavras como “isolamento”, “quarentena” e “casos” estiveram presentes em todas as coletas, indicando que há uma preocupação sobre como está sendo tratada a estratégia de redução de casos, mostrando o claro impacto na saúde e na sociedade.

Houve também certa preocupação em relação à economia, considerando-se a palavra “comércio” sendo também bastante citada. Outra demonstração dos grandes efeitos da pandemia foi a apresentação de discussões que citam diferentes cidades e países, sinalizando as consequências globais, além das internas, com as palavras “São Paulo”, “Blumenau”, “Alemanha”, “Argentina”, “Espanha”, etc.

Por fim, a presença do termo “*lockdown*”, presente na última coleta, expôs uma nova estratégia, o que indica uma reflexão acerca das estratégias e sua efetividade, a partir do acompanhamento da evolução da pandemia no país, demonstrada pelos gráficos utilizados.

Como apresentado ao longo do trabalho, as redes sociais refletem bastante momentos de crise. Então, este trabalho contribui com a disseminação de informações sobre a pandemia, através de um estudo que acompanhou a evolução da discussão do Coronavírus, no contexto brasileiro, na rede social Twitter.

Ressaltamos que os resultados aqui apresentados são resultados de consultas a *tweets* nas datas selecionadas. Apenas foi apresentado o que foi identificado. Entretanto, dado às limitações da ferramenta de coleta, é possível que mais dados seriam retornados em cada consulta, provocando uma variação nas palavras e em suas quantidades. Além disso, não se pode generalizar tendências, pois apenas foram analisados cinco dias. E a dinâmica das redes sociais varia muito conforme os acontecimentos.

Como trabalhos futuros, seria interessante investigar por mais tempo os dados do Twitter, pois apenas cinco dias de coleta não é suficiente para a obtenção de conclusões mais precisas e mais científicas. A análise dos dados em outras rede sociais também pode ser um trabalho que pode apontar diversas situações de impacto social neste momento de pandemia. E a correlação dos dados obtidos a partir do Twitter com outras redes sociais também pode colaborar com uma melhor e mais precisa descrição da situação real da pandemia.

Agradecimentos

Este trabalho foi realizado com recursos do projeto **NÚCLEO DE ANÁLISE QUALITATIVA PARA ESTUDOS DE COMPUTAÇÃO**, EDITAL Nº 4/2019 - PREX, da Universidade Federal do Ceará (UFC).

Referências

Alqurashi, S., Alhindi, A., e Alanazi, E. (2020). Large arabic twitter dataset on covid-19. <https://arxiv.org/abs/2004.04315>.

- Alshaabi, T., Minot, J. R., Arnold, M. V., Adams, J. L., Dewhurst, D. R., Reagan, A. J., Muhamad, R., Danforth, C. M., e Dodds, P. S. (2020). How the world's collective attention is being paid to a pandemic: Covid-19 related 1-gram time series for 24 languages on twitter. <https://arxiv.org/abs/2003.12614>.
- Caetano, J., Lima, H., dos Santos, M., e Marques-Neto, H. (2017). Utilizando análise de sentimentos para definição da homofilia política dos usuários do twitter durante a eleição presidencial americana de 2016. In *Anais do VI Brazilian Workshop on Social Network Analysis and Mining*, Porto Alegre, RS, Brasil. SBC.
- Chamola, V., Hassija, V., Gupta, V., e Guizani, M. (2020). A comprehensive review of the covid-19 pandemic and the role of iot, drones, ai, blockchain and 5g in managing its impact. *IEEE Access*, pages 1–1.
- Chen, E., Lerman, K., e Ferrara, E. (2020). Covid-19: The first public coronavirus twitter dataset. <https://arxiv.org/abs/2003.07372>.
- de Vasconcelos, D. R., de Souza, C. P., e Filho, R. N. C. (2020). Reports on the covid-19 outbreak. <https://sistemas.quixada.ufc.br/covid/>. Acessado: 09-05-2020.
- Giachanou, A. e Crestani, F. (2016). Like it or not: A survey of twitter sentiment analysis methods. *ACM Comput. Surv.*, 49(2).
- Kleinberg, B., van der Vegt, I., e Mozes, M. (2020). Measuring emotions in the covid-19 real world worry dataset. <https://arxiv.org/abs/2004.04225>.
- Kouzy, R., Jaoude, J. A., Kraitam, A., Alam, M. B. E., Karam, B., Adib, E., Zarka, J., Traboulsi, C., Akl, E. W., e Baddour, K. (2020). Coronavirus goes viral: Quantifying the covid-19 misinformation epidemic on twitter. *Cureus*, 12(3).
- Lopez, C. E., Vasu, M., e Gallemore, C. (2020). Understanding the perception of covid-19 policies by mining a multilanguage twitter dataset. <https://arxiv.org/abs/2003.10359>.
- Ortiz-Ospina, E. (2019). The rise of social media. <https://ourworldindata.org/rise-of-social-media>. Acessado: 05-05-2020.
- Pulido, C. M., Villarejo-Carballido, B., Redondo-Sama, G., e Gómez, A. (2020). Covid-19 infodemic: More retweets for science-based information on coronavirus than for false information. *International Sociology*, 0(0):0268580920914755.
- Singh, L., Bansal, S., Bode, L., Budak, C., Chi, G., Kawintiranon, K., Padden, C., Varnarsdall, R., Vraga, E., e Wang, Y. (2020). A first look at covid-19 information and misinformation sharing on twitter. <https://arxiv.org/abs/2003.13907>.
- World Health Organization (2020). Coronavirus disease (covid-19) pandemic. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Acessado: 06-05-2020.
- Wu, Z. e McGoogan, J. M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*, 323(13):1239–1242.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., e Tan, W. (2020). A novel coronavirus from patients with pneumonia in china, 2019. *New England Journal of Medicine*, 382(8):727–733. PMID: 31978945.