CPI da Pandemia no WhatsApp

Relatório Monitor de WhatsApp

Maio de 2021 a Junho de 2021

O Projeto

O Monitor de WhatsApp e parte do projeto Eleicoes Sem Fake realizado por pesquisadores da UFMG com objetivo de dar mais transparencia a rede do WhatsApp e o universo de informacao que circula nesta rede, evidenciando os conteudos mais populares em grupos publicos monitorados pela pesquisa.

O monitoramento de grupos publicos de WhatsApp realizado neste trabalho contempla desde a criacao um total de mais de 800 grupos publicos relacionados a politica brasileira, monitorando dentre estes um diversificado espectro politico, abrangendo tanto grupos pro-governo como aqueles de oposicao e com outras visoes politicas. O sistemas pode ser acessado atraves do endereco <http://whatsapp-monitor.dcc.ufmg.br> e o processo de desenvolvimento e pesquisa publicado em (MELO et al., 2019a).

Mais informacoes e resultados desta pesquisa podem ser encontrados nos artigos publicados pelo grupo, tais como (MELO et al., 2019b) e (VIEIRA et al., 2019) que analisa o espalhamento e viralizacao das mensagens na rede. (RESENDE et al., 2019b), (RESENDE et al., 2019a) fazem uma analise mais aprofundada dos dados focado nas eleicoes de 2018. (REIS et al., 2020a) explora o beneficio do sistema para agencias de checagem no combate a fake news e (REIS et al., 2020b) disponibiliza uma parcela dos dados de imagem para analise.

Metodologia de seleção das mensagens

Para esse relatório, foram consideradas apenas as mensagens enviadas entre **1 de Maio de 2021** e **15 de Junho de 2021** Para determinar a temática e o contexto das mensagens, procuramos pela ocorrência de alguns termos epecíficos que se relacionam de alguma forma ao tema dado. Foram ignorados acentuação e a captalização das palavras, de modo a fazer com que a busca ficasse mais abrangente e ignorasse possíveis variações dos termos. Assim, dentre todas as mensagens do período, foram filtradas somente aquelas que continham as seguintes palavras-chaves:

Keywords

("CPI") ("CPI da COVID") ("CPI da Pandemia") ("CPI do Coronavirus") ("corona virus") ("coronavirus") ("pandemia") ("covid-19") ("covid19") ("vacina") ("coronavac") ("astrazeneca") ("astrazenica") ("butantan") ("butantã") ("imunidade de rebanho") ("Anvisa") ("Agência Nacional de Vigilância Sanitária") ("pfizer") ("Marcos Rogerio") ("Randolfe Rodrigues") ("Renan Calheiros") ("Omar Aziz") ("Ministro da Saude") ("Ministro da Saúde") ("Eduardo Pazuello") ("Pazuello") ("Luiz Henrique Mandetta") ("Mandetta") ("Nelson Teich") ("Teich") ("Marcelo Queiroga") ("Queiroga") ("Antonio Barra Torres") ("Fábio Wajngarten") ("Ernesto Araújo") ("Nise Yamaguchi") ("Yamaguchi") ("Carlos Murilo") ("Mayra Pinheiro") ("Dimas Covas") ("Luana Araújo")

Conjunto de Dados Analisados

Este relatório contempla os dados enviados nestes grupos públicos políticos de WhatsApp durante o período entre **1 de Maio de 2021** a **15 de Junho de 2021**. Neste período, um total de **384**

grupos ativos mandaram mensagens, somando um conjunto de **47338 mensagens** enviadas por **4945 usuários** distintos.

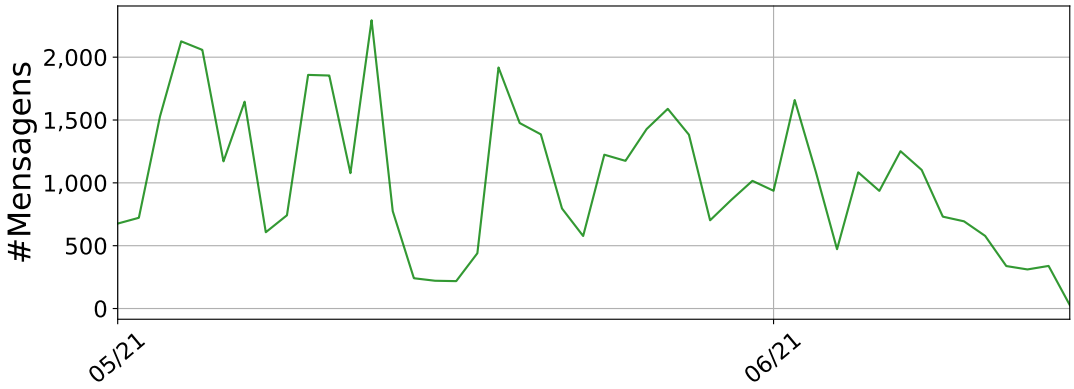
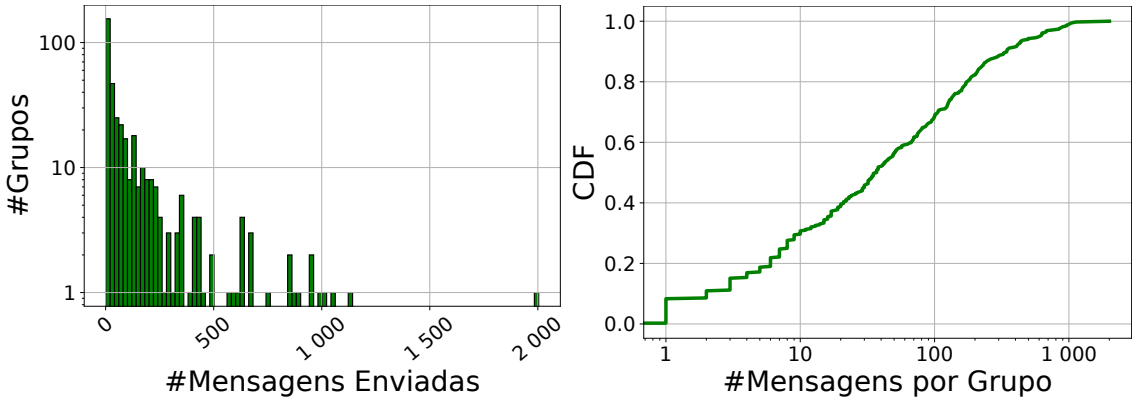


Figura 1: Mensagens por dia

Quem envia as mensagens?

Tanto os grupos como os usuários possuem comportamento distinto dentro WhatsApp. Portanto, para detalhar melhor este relatório, primeiramente voltamos nossa atenção sobre os grupos e usuários mais ativos.

A Figura explora a quantidade de mensagens enviadas por grupo. No histograma, os grupos são divididos pelo seu volume de mensagens, sendo possível observar quantos grupos existem para cada parcela de volume. A barra mais a esquerda representa quantos grupos enviaram poucas mensagens no período, enquanto a barra mais a direita representa os grupos mais ativos.

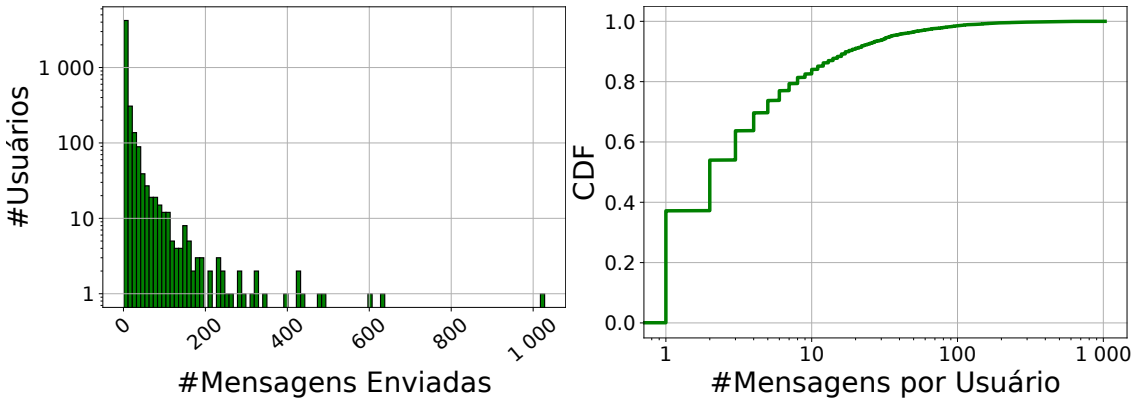


(a) Histograma

(b) Distribuição (CDF)

Figura 2: Mensagens por Grupo

De forma semelhante, a Figura explora a quantidade de mensagens enviadas por usuário. Note que a quantidade de usuários é expressivamente maior que a de grupos, porém a quantidade de mensagens por usuários é menor que de grupos. O histograma também mostra o volume enviado por cada usuário separadamente, com a quantidade de usuários com poucas mensanges representado pela barra mais a esquerda e os mais ativos à direita.



(a) Histograma (b) Distribuição (CDF)

Figura 3: Mensagens por Usuário

Para investigar mais a fundo os usuários que enviam as mensagens, também levantamos o volume de mensagens de acordo com localização geográfica de onde elas foram enviadas. a Table **??** mostra a quantidade de mensagens enviadas por estado, além de quantos usuários daquela região estão presentes nos nossos dados. Como o WhatsApp funciona com base no número de celular de cada usuário, estimamos esta localização a partir do DDD de cada número.

Tabela 1: Quantidade de usuários por estado (DDD).

DDD	# Mensagens	# Usuários
BA	5680	457
SP	5459	800
RJ	4784	565
MG	4296	471
RN	3454	92
DF	3102	266
CE	2463	240
PB	2114	146
SC	1796	123
PE	1720	203
PI	1467	98
ES	1296	116
GO	1185	138
MA	1077	177
PR	1004	191
PA	912	119
MS	862	72
AM	788	97
RS	738	128
MT	533	79
TO	472	66
AL	459	68
SE	316	70
AP	306	20
RO	297	40
RR	197	19
AC	167	23

Conteúdo das Mensagens

Partindo para o conteúdo das mensagens, a primeira observação é sobre o tipo de mensagem enviada. Como a plataforma do WhatsApp permite envio de mensagens multimídia, é comum nos grupos que este recurso seja explorado para envio das diversas formas de mensagens que o aplicativo permite. Portanto, a Figura apura os tipos de mensagens enviadas entre texto, imagem, vídeo e áudio.

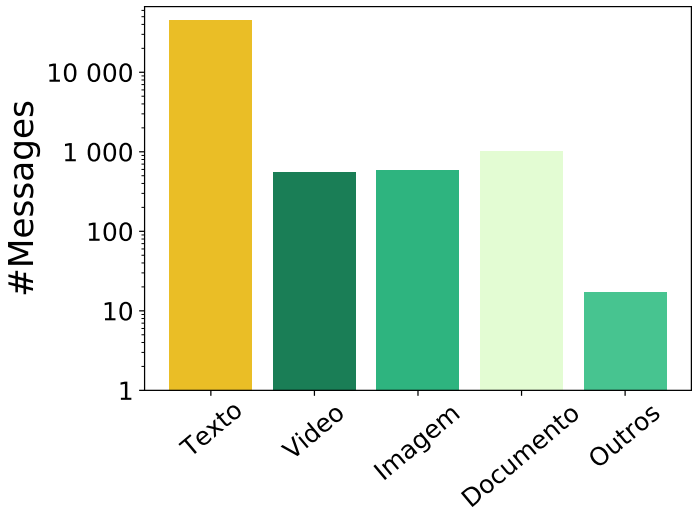


Figura 4: Tipo de Mensagem enviada.

O Monitor de WhatsApp processa e agrupa os conteúdos enviados nos grupos contabilizando aqueles mais compartilhados para cada tipo de mídia. Com isso, é possível adentrar naquilo que se tornou mais popular no período de tempo analisado. A seguir, este relatório busca auxiliar no entendimento dos tópicos e assuntos mais relevantes do período.

Texto

As mensagens selecionadas para gerar este relatório foram filtradas com base num conjunto de palavra-chaves. Com intuito de enxergar melhor como essas *keywords* estão presentes e distribuídos nas mensagens, a Tabela resume a quantidade de ocorrências de cada *keyword* e quantos usuários e grupos publicaram alguma mensagem contendo-as.

Tabela 2: Número de ocorrências das palavras-chaves utilizadas para filtragem de dadoos.

Keyword	# Ocorrências	# Grupos	# Usuários
vacina	15145	343	2710
CPI	14141	338	2705
covid-19	11296	318	2086
pandemia	9180	337	2375
coronavirus	4085	288	1323
Pazuello	3474	243	940
CPI da COVID	3072	245	913
pfizer	2695	263	958
Renan Calheiros	2265	236	894
coronavac	1750	205	639
Anvisa	1618	224	570
Mandetta	1432	218	535
Queiroga	1070	188	357
Ministro da Saude	1048	183	378
Omar Aziz	987	192	490
CPI da Pandemia	910	190	401
astrazeneca	897	197	383
Yamaguchi	890	159	453
butantan	780	152	314
covid19	722	176	354
Nise Yamaguchi	694	153	384
Eduardo Pazuello	612	164	314
Marcelo Queiroga	396	133	157
Mayra Pinheiro	362	99	159
Marcos Rogerio	308	102	133
Ernesto Araujo	264	101	106
Randolfe Rodrigues	232	96	126
imunidade de rebanho	214	58	91
Luana Araujo	209	76	110
Dimas Covas	196	65	68
Fabio Wajngarten	179	72	61
Luiz Henrique Mandetta	162	85	68
Nelson Teich	159	81	65
Agencia Nacional de Vigilancia Sanitaria	149	72	63
corona virus	117	68	86
Antonio Barra Torres	69	36	24
astrazenica	22	18	19
Carlos Murilo	3	3	3

A seguir é exibida a lista das 5 mensagens que mais circularam nos grupos monitorados durante o período observado. Nesta análise apenas mensagens com acima de 140 caracteres foram consideradas. Para definir se duas mensagens são iguais, os textos foram comparados utilizando a similaridade de Jaccard.

Top 1ª Mensagem mais compartilhada A 1ª mensagem mais compartilhada foi enviada um total de **193** vezes por **154** usuários distintos e apareceu em **103** grupos. Ela apareceu pela primeira vez nos dados em **9 de Maio de 2021**.

Top 1ª Mensagem

BOMBA: DOCUMENTOS MOSTRAM CORONAVÍRUS SENDO TESTADO COMO ARMA BIOLÓGICA 5 ANOS ANTES DA PANDEMIA POR CHINESES https://terrabrasilnoticias.com/2021/05/bomba-documentos-mostram-coronavirus-sendo-testado-como-arma-biologica-5-anos-antes-da-pandemia-por-chineses/

Top 2ª Mensagem mais compartilhada A 2ª mensagem mais compartilhada foi enviada um total de **137** vezes por **110** usuários distintos e apareceu em **80** grupos. Ela apareceu pela primeira vez nos dados em **20 de Maio de 2021**.

Top 2ª Mensagem

URGENTE!!! *PAZUELLO ENCERRA SEU DISCURSO COM FRASE PERFEITA:* "A pandemia escancarou uma crítica situação da saúde pública no Brasil. Depois de anos de negligência."Ex Ministro Gen. Eduardo Pazuelo começou desmontando a primeira mentira da oposição: de que ia ficar em silêncio. Está falando e respondendo tudo e com dados os questionamentos. *PAZUELLO DEIXA RENAN SEM AÇÃO.* "Senador, seria um absurdo o presidente da República não ouvir opiniões divergentes para poder formar sua decisão. Ouvir Osmar Terra e a Dra. Nise Yamaguchi não é crime e eles não tinham nenhum papel no ministério."*PAZUELLO DETONA RENAN* "Perguntar a um general se ele tem competência de gestão e liderança é como o senhor perguntar se chuva molha. Se o senhor achar que os generais do Exército não tem competência para gestão e liderança, temos que recriar o nosso Exército."*PAZUELLO REBATE PRIMEIRA TENTATIVA DE PRESSÃO DE RENAN E AZIZ* "TEMOS UMA POPULAÇÃO INTEIRA NOS ASSISTINDO SENADORES, ME DESCULPEM, MAS PERGUNTAS SIMPLÓRIAS TIPO SIM OU NÃO NÃO SERÃO RESPONDIDAS SEM CONTEXTUALIZAÇÃO DE MINHA PARTE E EU VOU RESPONDER A TODAS AS PERGUNTAS."*PAZUELLO DEFENDE A SOBERANIA NACIONAL NA CPI* "A OMS e a OPAS estavam presentes constantemente conosco, mas elas não impõem decisões ao nosso país. O Brasil é soberano na sua tomada de decisão em todas as ações no combate à pandemia."*PAZUELLO ENCERRA SEU DISCURSO DETONANDO:* "A Pandemia escancarou a crítica situação da saúde pública no Brasil, depois de anos de negligência. O governo Bolsonaro foi o que mais investiu em saúde e espero que isso continue após a pandemia sair dos holofotes."*PAZUELLO* "Nós não somos obrigados a seguir orientações de ONU, OMS, de lugar nenhum. Nós somos soberanos"Eduardo Pazuello *Pazuello acaba de enterrar a narrativa sobre a cloroquina* "Caso os senadores não queiram passar vergonha, sugiro começarem a investigar os desvios praticados por governadores. Na Zika e na Chikungunya, havia protocolos do Ministério da Saúde para o uso da Cloroquina em altas doses, inclusive para grávidas. Hoje temos no mundo 29 países com protocolo para uso da Cloroquina contra a Covid, incluindo China, Cuba, México e Venezuela" *Pazuello derruba a narrativa de que o governo não fez campanhas publicitárias de combate à Covid* Pazuello apresenta todos os dados das campanhas e chama a atenção para os mais de *2 bilhões* de acessos às campanhas divulgadas pelas redes sociais do governo. *PAZUELLO FAZ RENAN PASSAR VERGONHA* "Osmar Terra não tinha nenhuma influência oficial no Ministério da Saúde senador. O presidente nunca discutiu essa tese de imunização de rebanho comigo senador."*Pazuello dando show na CPI da Pandemia. Esclareceu sobre o processo de compra de vacinas da Pfizer:* "A resposta do Brasil sempre foi: SIM, queremos comprar. Mas até a assinatura do contrato houve uma série de reuniões e discussões para discutir os termos do contrato" Renan Calheiros pergunta: qual o papel do Osmar Terra? *Pazuello:* Deputado Federal! *PAZUELLO COLOCA RENAN NO SEU LUGAR* RENAN: O presidente tomou a decisão de não responder a Pfizer? *PAZUELLO:* "Não houve decisão de responder a Pfizer, estou lhe respondendo pela quinta vez, senador. O senhor não vai conseguir conduzir minhas respostas" "O senhor deveria saber disso". "O senhor não parece entender as próprias perguntas". *10 X 0 PAZUELLO!!!*

VERGONHA ALHEIA!!! Acharam que o General ficaria quietinho? Engoliu Renan que agora está com o rabo entre as pernas. *PARABÉNS PAZUELLO!!!* É DISSO QUE ESSA CORJA PRECISA! *VAMOS COMPARTILHAR!*

Top 3ª Mensagem mais compartilhada A 3ª mensagem mais compartilhada foi enviada um total de **109** vezes por **2** usuários distintos e apareceu em **2** grupos. Ela apareceu pela primeira vez nos dados em **4 de Maio de 2021**.

Top 3ª Mensagem

Veja esta matéria: https://www.portalsbn.com/artigo/bahia-recebe-remessa-com-351-750-doses-de-vacinas-contra-a-covid-19 - É notícia? Tá no Portal SBN | Sistema Brasileiro de Notícias!

Top 4ª Mensagem mais compartilhada A 4ª mensagem mais compartilhada foi enviada um total de **106** vezes por **86** usuários distintos e apareceu em **75** grupos. Ela apareceu pela primeira vez nos dados em **2 de Maio de 2021**.

Top 4ª Mensagem

Vamos espalhar esse vídeo da esposa do Omar Aziz. Ele é o presidente da CPI que vai investigar a atuação do Bolsonaro e governadores na pandemia...

Top 5ª Mensagem mais compartilhada A 5ª mensagem mais compartilhada foi enviada um total de **76** vezes por **19** usuários distintos e apareceu em **33** grupos. Ela apareceu pela primeira vez nos dados em **10 de Maio de 2021**.

Top 5ª Mensagem

URGENTE: Documento chinês vazado aponta coronavírus como arma biológica cinco anos antes da pandemia *https://www.folhadapolitica.com/2021/05/urgente-documento-chines-vazado-aponta.html?m=1*

Também foi feita uma análise geral de todas mensagens enviadas. Embora seja difícil verificar manualmente o grande volume de mensagens enviadas no WhatsApp, algumas ferramentas facilitam o processo de identificar topicos de maior interesse, como a nuvem de termos. Esta visualização permite enxergar os principais termos presentes nas mensagens enviadas, assim como sua popularidade, representada pelo tamanho de cada termo dentro da nuvem. A Figura mostra a nuvem de termos gerada com todas as mensagens enviadas nos grupos políticos de WhatsApp.

Figura 5: Nuvem de termos das mensagens.

Mesmo sendo um aplicativo de mensagens de conversa, muitos usuários uutilizam o esquema de hashtags para definirem os tópicos de algumas mensagens e divulgar seu conteúdo. A Figura foi gerada utilizando apenas as hashtags(#) extraída das mensagens.

Figura 6: Nuvem de hashtags mais populares.

URLs e Domínios

Os domínios web de maior circulação podem ser vistos de forma resumida na Tabela ,

Tabela 3: Número de ocorrências do top 20 domínios presentes nas mensagens.

Domínio	# Total	# Distintas	# Usuários
whatsapp.com	12154	78	247
terrabrasilnoticias.com	3339	556	634
youtube.com	3332	1049	601
t.me	2876	197	274
instagram.com	1717	533	348
gazetabrasil.com.br	1623	865	81
novaiguacu24h.com.br	1446	102	108
facebook.com	1277	592	348
globo.com	920	590	357
uol.com.br	912	597	291
twitter.com	853	404	200
boatos.org	834	140	24
dunapress.org	708	48	24
jornaldacidadeonline.com.br	656	169	272
contrafatos.com.br	551	169	87
ultimasnoticiasdobrasil.com	550	65	18
portalcidade.news	524	179	69
blogspot.com	517	285	69
conteudoms.com	491	485	4
empregabilidadebrasil.com	441	34	24

De forma mais detalhada, observamos como alguns dos domínios das plataformas online e redes sociais mais tradicionais foram compartilhados no WhatsApp ao longo do tempo na Figura .

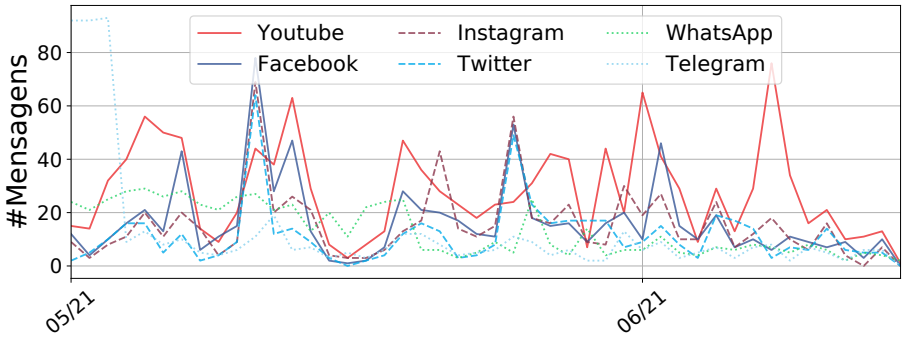


Figura 7: Compartilhamento de plataformas tradicionais no WhatsApp ao longo do tempo.

Já os endereços web específicos mais compartilhados estão apresentados a seguir com as top 5 URLs mais compartilhadas.

Top 1ª URL mais compartilhada A 1ª URL mais compartilhada foi enviada um total de **295** vezes por **214** usuários distintos e apareceu em **128** grupos.
<https://terrabrasilnoticias.com/2021/05/bomba-documentos-mostram-coronavirus-sendo-testado-como-ar
>

Top 2ª URL mais compartilhada A 2ª URL mais compartilhada foi enviada um total de **168** vezes por **51** usuários distintos e apareceu em **31** grupos.
<https://t.me/apostagem>

Top 3ª URL mais compartilhada A 3ª URL mais compartilhada foi enviada um total de **148** vezes por **134** usuários distintos e apareceu em **88** grupos.
<https://www.novartis.com/news/media-releases/novartis-commits-donate-130-million-doses-hydroxychlor

Top 4ª URL mais compartilhada A 4ª URL mais compartilhada foi enviada um total de **148** vezes por **134** usuários distintos e apareceu em **88** grupos.
<https://sahelstandard.com/2020/04/06/anti-malaria-drug-can-kill-coronavirus-novartis-chief/
>

Top 5ª URL mais compartilhada A 5ª URL mais compartilhada foi enviada um total de **133** vezes por **13** usuários distintos e apareceu em **13** grupos.
<https://gazetabrasil.com.br/contribua/>

Vídeos do Youtube

Além das URLs mais populares, também observamos quais links de vídeos do Youtube foram mais compartilhados no período analisado. Como Youtube é uma plataforma muito popular, é de grande importância levantar os conteúdos mais populares que as mensagens redirecionam. A seguir listamos os 5 top videos de Youtube

Top 1ª Youtube video mais compartilhado O 1º video de Youtube mais compartilhado foi enviada um total de **83** vezes por **8** usuários distintos e apareceu em **7** grupos.
<https://www.youtube.com/channel/UCsOkGiMUsY_GNNvj_zJd8tw>

Top 2ª Youtube video mais compartilhado O 2º video de Youtube mais compartilhado foi enviada um total de **68** vezes por **2** usuários distintos e apareceu em **19** grupos.
<https://www.youtube.com/watch?v=meBZpWX5Zmo&feature=youtu.be>

Top 3ª Youtube video mais compartilhado O 3º video de Youtube mais compartilhado foi enviada um total de **58** vezes por **2** usuários distintos e apareceu em **19** grupos.
<https://www.youtube.com/watch?v=tsf_4jIuseY&feature=youtu.be>

Top 4ª Youtube video mais compartilhado O 4º video de Youtube mais compartilhado foi enviada um total de **48** vezes por **9** usuários distintos e apareceu em **19** grupos.
<https://www.youtube.com/watch?v=MkJh5nsdUOs&feature=youtu.be>

Top 5ª Youtube video mais compartilhado O 5º video de Youtube mais compartilhado foi enviada um total de **37** vezes por **3** usuários distintos e apareceu em **17** grupos.
<https://www.youtube.com/watch?v=ud-puw6gdHY&feature=youtu.be>

Referências

MELO, P.; MESSIAS, J.; RESENDE, G.; GARIMELLA, K.; ALMEIDA, J.; BENEVENUTO, F. Whatsapp monitor: A fact-checking system for whatsapp. In: *ICWSM*. [S.l.: s.n.], 2019.

MELO, P.; VIEIRA, C.; GARIMELLA, K.; MELO, P. O. V. de; BENEVENUTO, F. Can whatsapp counter misinformation by limiting message forwarding? In: *Proceedings of the the International Conference on Complex Networks and their Applications*. [S.l.: s.n.], 2019. (Complex Networks'19).

REIS, J. C.; MELO, P.; GARIMELLA, K.; BENEVENUTO, F. Can whatsapp benefit from debunked fact-checked stories to reduce misinformation? *Harvard Kennedy School Misinformation Review*, 2020.

REIS, J. C. S.; MELO, P.; GARIMELLA, K.; ALMEIDA, J. M.; ECKLES, D.; BENEVENUTO, F. A dataset of fact-checked images shared on whatsapp during the brazilian and indian elections. In: *Proceedings of the Int'l AAAI Conference on Weblogs and Social Media (ICWSM)*. [S.l.: s.n.], 2020.

RESENDE, G.; MELO, P.; REIS, J. C. S.; VASCONCELOS, M.; ALMEIDA, J.; BENEVENUTO, F. Analyzing textual (mis)information shared in whatsapp groups. In: *WebScience*. [S.l.: s.n.], 2019.

RESENDE, G.; MELO, P.; SOUSA, H.; MESSIAS, J.; VASCONCELOS, M.; ALMEIDA, J.; BENEVENUTO, F. (mis)information dissemination in whatsapp: Gathering, analyzing and countermeasures. In: *WWW*. [S.l.: s.n.], 2019.

VIEIRA, C.; MELO, P.; MELO, P. O. V. de; BENEVENUTO, F. O paradoxo da viralização de informação criptografada no whatsApp. In: *Anais do Simpósio Brasileiro de Redes de Computadores (SBRC)*. Gramado, Brasil: [s.n.], 2019.

Equipe:

Philipe Melo
Vitor Mafra
Daniel Kansaon
Kaio Sá
Fabrício Benevenuto