

Relatório de julho de 2021

Por Décio Telo, Ana Pinto-Martinho e Livino Neto, com supervisão de Gustavo Cardoso.

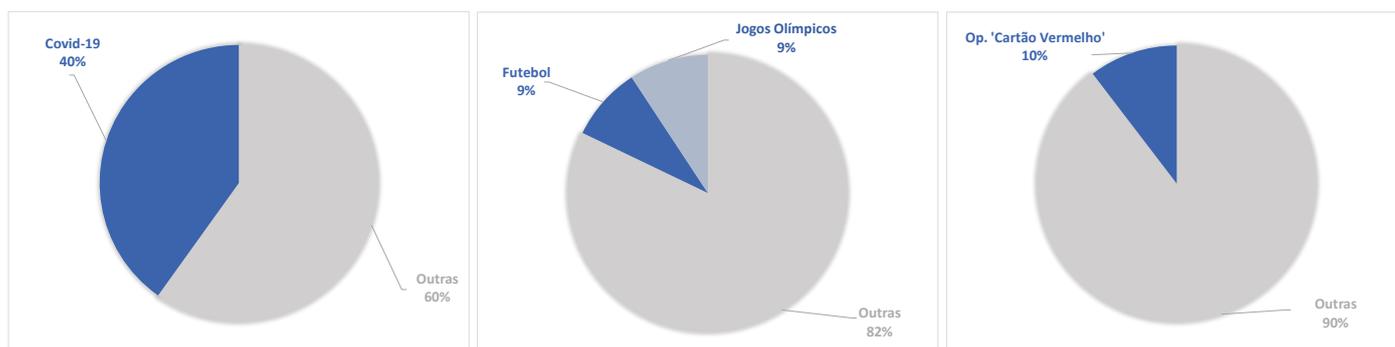
Síntese

A 4.ª vaga da pandemia por SARS-CoV-2 atingiu Portugal em força e a chama olímpica chegou, finalmente, a Tóquio, para o início dos Jogos Olímpicos de 2020. Em julho, estes temas tiveram ampla cobertura jornalística com grande incidência na narrativa sobre vacinação da população mais jovem em Portugal.



Da esquerda: Marta Temido e Isaltino Morais em Carnaxide, Oeiras, 16/7/2021 (JOSÉ SENA GOULÃO/LUSA); Chama olímpica no estádio Olímpico de Tóquio, Japão, 23/7/2021 (JOSÉ COELHO/LUSA); Transporte de pessoas detidas no caso "Cartão Vermelho" em Lisboa, 9/7/2021 (MANUEL DE ALMEIDA/LUSA).

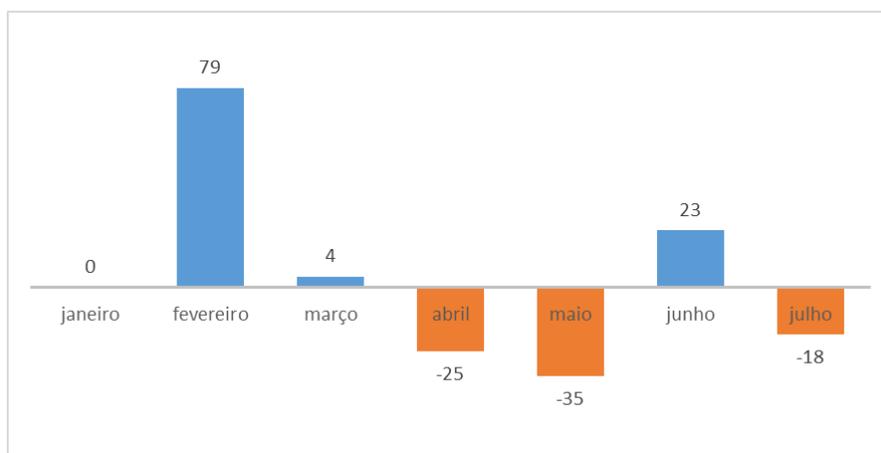
As 'histórias' que marcaram as notícias online



Percentagem de artigos relacionados com pandemia (gráfico da esquerda), Jogos Olímpicos (ao centro) e com operação Cartão Vermelho (à direita) no conjunto dos artigos analisados no âmbito deste barómetro (N=10.138 documentos). Fonte: MediaLab/Priberam (dados recolhidos a 3/8/2021).

Ligeira variação na cobertura da pandemia

A pandemia provocada pelo vírus SARS-CoV-2 gerou, em média, 131 artigos por dia nos sites de notícias generalistas, no mês de julho de 2021. Não obstante este indicador, houve uma ligeira descida na variação mensal face ao mês de junho.



Variação mensal, em percentagem, de artigos sobre Covid-19. fonte: MediaLab/Priberam (dados atualizados em 3/8/2021).

Substituição de Luís Filipe Vieira na presidência do Benfica

Os dados disponíveis indicam que a substituição de Luís Filipe Vieira na presidência do Sport Lisboa e Benfica, na sequência da sua detenção no âmbito da Operação Cartão Vermelho, foi o tópico que reuniu o maior número de artigos publicados online e representou cerca de 10% das notícias analisadas pelo barómetro. Este indicador não inclui artigos publicados em *sites* desportivos. É o primeiro mês, este ano, em que a pandemia não se encontra associada a *cluster* com maior número de artigos publicados nos *sites* de comunicação social generalistas.

O desporto, entre jogos olímpicos e futebol

A cerimónia de abertura dos Jogos Olímpicos de Tóquio (2020) aconteceu, finalmente, no dia 23 de julho de 2021. Em 8 dias de competição, registou-se uma média de 30 artigos por dia, publicados nos sites de notícias generalistas. Dois *clusters* temáticos destacaram-se, nesta fase: a cerimónia de abertura e o desempenho dos atletas portugueses que competiram na primeira semana dos jogos. No futebol, ainda em fase de preparação da nova época, a atenção foi, sobretudo, para contratações em diversos clubes nacionais e internacionais, com algum destaque para a de João Mário pelo Sport Lisboa e Benfica.



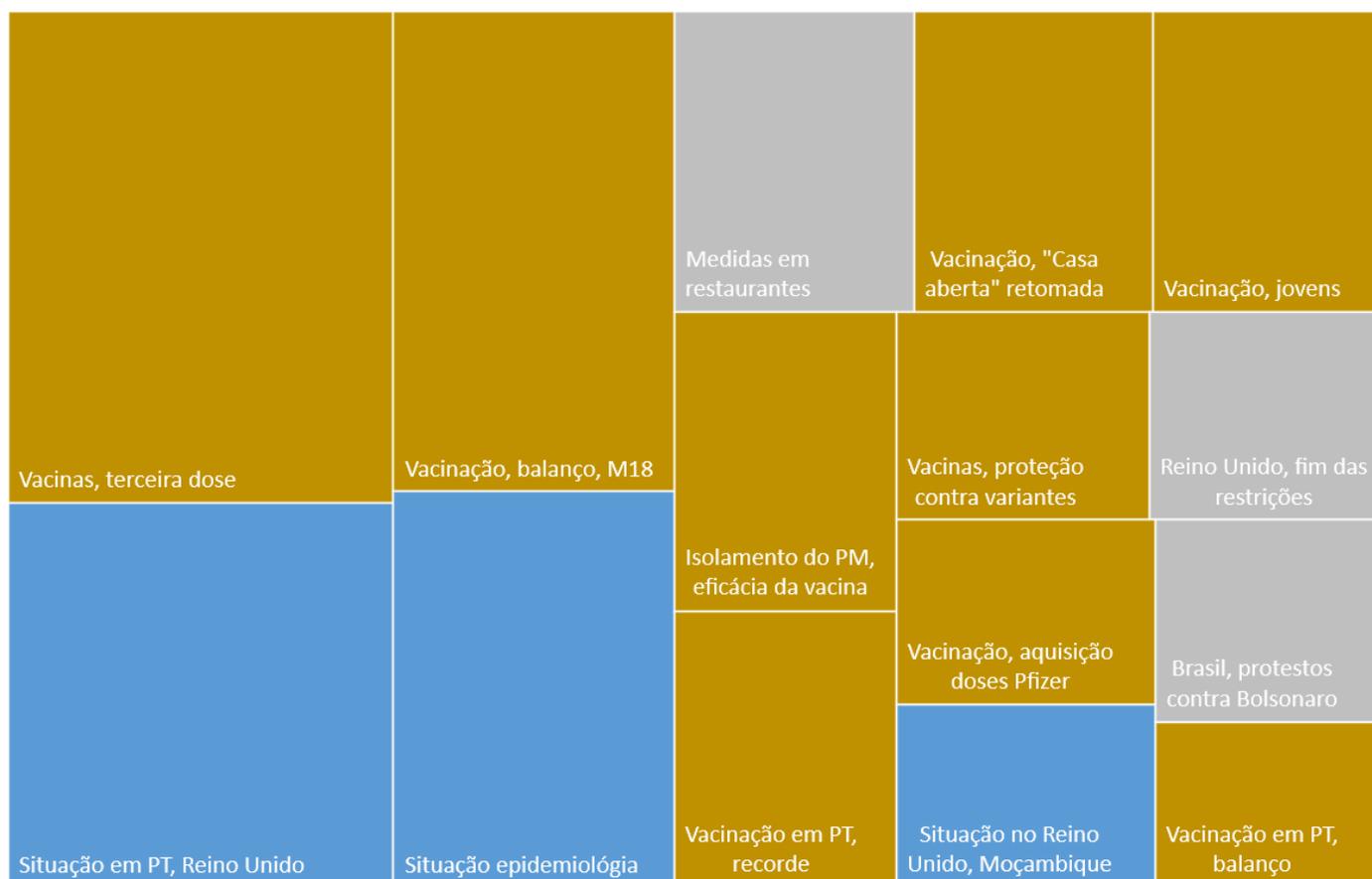
Os 'clusters' relacionados com a pandemia estão identificados pela cor azul, os associados ao desporto com a cor verde e os relacionados com a Operação "Cartão Vermelho" a laranja. Os restantes estão preenchidos a cinzento. Quadro construído com índice de base = 100 em que a variável de interesse é o cluster mais publicado: "Operação Cartão vermelho...". fonte: MediaLab/Priberam (dados recolhidos em 3/8/2021)

Nota metodológica sobre a análise dos *clusters* de notícias

O quadro acima permite visualizar a proporção de cada agrupamento de notícias (*cluster*) no conjunto dos 41 *clusters* que mais se destacaram em julho de 2021. Estes, resultam de um trabalho de análise e reagrupamento semântico realizado no conjunto mais alargado, de 50 *clusters*¹, identificados pelo algoritmo da plataforma Priberam. Esta operação de reagrupamento é realizada com recurso a metodologias de análise de conteúdo. Deste trabalho resulta que a quantidade final de *clusters* no relatório final é sempre igual ou inferior ao ‘top 50’ gerado pelo algoritmo².

A narrativa da pandemia nas notícias

A análise identificou, em junho, 15 tópicos³ relacionados com a pandemia, menos 4 do que no mês anterior. Este mês, o processo de vacinação dominou por completo a cobertura noticiosa nos sites generalistas. A hipótese de uma terceira dose constituiu o *cluster* com mais artigos publicados. A narrativa jornalística colocou em destaque duas ‘histórias’: Na primeira, a informação de que a DGS está a admitir uma terceira dose “para pessoas mais vulneráveis” (Expresso online, 23/7/2021) e, na segunda, a perspetiva da Agência Europeia EMA, que “acredita que não será preciso tomar uma terceira dose da vacina contra a covid-19” (Jornal i online, 9/7/2021) ou do Infarmed que, de acordo com notícia da Rádio Renascença, comunica “que, para já, não é necessário” uma terceira dose da vacina (RR online, 23/7/2021).



As notícias sobre Covid-19 com destaque para 2 grupos temáticos: evolução da pandemia em Portugal e no Mundo (cor azul), Vacinação (cor dourada). Outros temas estão identificados com a cor cinza. fonte: MediaLab/Priberam (dados recolhidos em 3/8/2021)

¹ Para mais informação, consultar a secção de metodologia no final do relatório.

² Apesar da elevada precisão dos clusters classificados automaticamente pelo algoritmo (cf. [Miranda et al, 2018](#)), alguns deles podem apresentar características polissémicas, por diversos motivos, tornando difícil a tomada de decisão final quanto à inclusão num determinado *cluster*, pelo que a opção metodológica que nos parece mais adequada é considerar, apenas, os *clusters* que demonstram homogeneidade do ponto de vista do assunto identificado, acontecimento ou ‘história’ relatada.

³ Para mais informação consulte a metodologia no final do relatório.

Descrição da amostra

No mês de junho foram considerados 10.138 artigos, de 18 fontes de informação jornalística⁴ para a realização deste relatório, selecionadas a partir de um universo de mais de 160 entidades (aproximadamente 215.098 artigos).

A recolha de dados foi realizada na plataforma Priberam e a organização das histórias noticiosas em *clusters* é realizada em dois passos: Um primeiro agrupamento é feito pelo algoritmo da plataforma, que devolve um 'TOP 50' com os *clusters* mais publicados. Num segundo momento, foi realizado um reagrupamento desse 'TOP 50' com recurso a técnicas de análise de conteúdo que aplicam uma análise 'fina' aos dados, resultando no quadro final de *clusters*⁵apresentado neste relatório.

Do ponto de vista temático, a 'política' mantém-se a área mais representada nas notícias online, com aproximadamente 15.648 artigos publicados. Dos cinco temas mais relevantes nas notícias online, apenas o desporto subiu uma posição por troca com as 'questões sociais'.

Categoria temática (IPTC)	Ranking
Política	0
Economia, negócios e finanças	0
Desporto	1
Questões sociais	-1
Crime, lei e justiça	0

Fonte: MediaLab/Priberam (dados recolhidos em 3/8/2021)

António Costa continuou a ser a personalidade mais referida em maio, seguido de Marcelo Rebelo de Sousa obteve metade das referências nas notícias online, em comparação com o Primeiro-Ministro. Luís Filipe Vieira surge de forma abrupta, em segundo lugar, como consequência da sua detenção no âmbito da 'Operação Cartão Vermelho'. No que respeita a entidades coletivas, Governo e Estado mantiveram-se no topo das referências nas publicações online, notando-se a subida do S. L. Benfica, clube de futebol, pelos mesmos motivos que explicam a presença de Luís Filipe Vieira nos rankings.

Personalidades	Ranking
António Costa	0
Luís Filipe Vieira	19
Marcelo Rebelo de Sousa	-1
Eduardo Cabrita	7
Joe Biden	-2

Fonte: MediaLab/Priberam (dados recolhidos em 3/8/2021)

Entidades	Ranking
Governo	0
Estado	0
Direção-Geral de Saúde	2
S. L. Benfica	6
Partido Socialista	-1

Fonte: MediaLab/Priberam (dados recolhidos em 3/8/2021)

⁴ CMTV, Correio da Manhã, Diário de Notícias, Expresso, Jornal de Notícias, Jornal i, Jornal SOL, Multinews, Notícias ao Minuto, NOVO, Observador, Público, Rádio Renascença, RTP, SIC Notícias, TSF, TVI24, Visão.

⁵ Para mais informações sobre metodologia consultar secção no final do relatório.

Metodologia do Barómetro MediaLab-Priberam

O Barómetro de Notícias analisa, periodicamente, as grandes ‘histórias’ mediáticas que recebem maior atenção na comunicação social online, em Portugal. Desta forma, pretende-se identificar, sistematicamente e com uma metodologia consistente ao longo do tempo, a tematização da agenda de debate público através da comunicação social.

A noção de grande ‘história’ mediática provém do termo anglo-saxónico "top story". Pode ser mais facilmente identificável quando se trata de um acontecimento imprevisto que se impõe pela sua natureza disruptiva e valor-notícia associado, mas também pode ser um assunto ou problema público que os media mantêm na agenda durante algum tempo.

A opção pela comunicação social online prende-se com a importância crescente e forma de circulação das notícias online, que corresponde a mudanças nas práticas jornalísticas - maior diversidade de plataformas, conteúdos, estilo e formas de destaque - e nos hábitos de consumo, também eles em mudança e que passam, em boa medida, pela circulação de notícias publicadas em plataformas online, quer sejam redes sociais ou aplicações ‘mobile’.

Como construímos o barómetro

Todo o trabalho de análise é feito sobre os dados fornecidos pela plataforma [Priberam](#). O sistema desenvolvido pela Priberam permite agregar artigos noticiosos em ‘histórias mediáticas’ (no sentido anglo-saxónico, conforme referido anteriormente) agrupadas em clusters. Este processo tem a dificuldade adicional de exigir a análise automática, completamente online e escalável de um fluxo ininterrupto de dados em contínua atualização. Este método implica a capacidade de autocorreção de decisões passadas de uma forma eficiente por forma a manter a coerência dos clusters.

A plataforma capta, de forma dinâmica, todas as publicações noticiosas disponíveis em mais de mil entidades (fontes de informação). A partir da organização automática, pelo algoritmo, das 50 ‘histórias’ mais significativas, agrupadas em clusters, procede-se a uma validação e análise semântica no MediaLab do CIES-Iscte, com recurso a técnicas tradicionais de análise de conteúdo.

Para o relatório final, publicado mensalmente, considera-se o conjunto de fontes que correspondem a órgãos de comunicação social generalista (publicações periódicas, empresas jornalísticas e noticiosas não especializadas) de âmbito nacional. Esta seleção das fontes é feita através de um filtro aplicado após a recolha dos dados.

O barómetro tem como objetivo secundário, registar a variação em percentagem e longevidade das histórias noticiosas mais relevantes ao longo dos meses. Os dados quantitativos fornecidos pelo algoritmo, apesar de se tratar de um indicador muito importante, não devem ser considerados como valores definitivos, dada a natureza dinâmica da rede, com a impossibilidade de captar a totalidade dos conteúdos publicados, com um grau de certeza absoluta. Desta forma, o barómetro privilegia frequências relativas e variações em percentagem no processo de definição da respetiva importância relativa.

Quem somos

O barómetro de notícias é fruto de uma parceria entre o [MediaLab CIES-Iscte](#) e a [Priberam](#).

O [MediaLab](#) é uma unidade integrada no [Centro de Investigação e Estudos em Sociologia](#) (CIES-Iscte). Na sua atividade, desenvolve e apoia projetos em fase de recolha ou tratamento de dados, que estejam relacionados com Estudos dos novos media (Internet e redes sociais), produção de conteúdos multimédia, estudos dos media ditos tradicionais (rádio, televisão e imprensa).

A [Priberam](#) é líder mundial na oferta de processamento de linguagem natural e tecnologias de pesquisa para a língua portuguesa e líder de mercado em sistemas de gestão do conhecimento jurídico em Portugal. Equipas multidisciplinares de engenheiros de software, linguistas e advogados, entre outros, desenvolvem o portfólio de produtos e serviços da empresa que são usados por milhões de pessoas em todo o mundo.

Bibliografia essencial sobre o algoritmo Priberam

Rúben Cardoso, Zita Marinho, Afonso Mendes and Sebastião Miranda, [Priberam at MESINESP Multi-label Classification of Medical Texts Task](#), CLEF 2021.

Rúben Cardoso, Afonso Mendes and André Lamurias, [Priberam Labs at the NTCIR-15 SHINRA2020-ML: Classification Task](#), Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies, December 8-11, 2020 Tokyo Japan..

Afonso Mendes, Shashi Narayan, Sebastião Miranda, Zita Marinho, André F. T. Martins and Shay B. Cohen, [Jointly Extracting and Compressing Documents with Summary State Representations](#), Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2019), Minneapolis, USA

Zita Marinho, Afonso Mendes, Sebastião Miranda and David Nogueira, [Hierarchical Nested Named Entity Recognition](#), The 2nd Clinical Natural Language Processing Workshop (within NAACL 2019), Minneapolis, USA

Sebastião Miranda, Arturs Znotins, Shay B. Cohen and Guntis Barzdins. [Multilingual Clustering of Streaming News](#). Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), Brussels, Belgium, 2018.

Toda a bibliografia relevante na página de publicações [Priberam Labs](#).

Bibliografia relevante sobre o Barómetro de Notícias do MediaLab

Livino Neto, Gustavo Cardoso, Décio Telo, [A montra jornalística na estação pandémica: análise das capas do Correio da Manhã, Jornal de Notícias e Público](#). Lisboa: Observatorio (OBS*) Journal, 2021. Special issue, pp. 24-44.

Paulo Couraceiro, Miguel Paisana, José Moreno, Décio Telo, Ana Pinto-Martinho, Gustavo Cardoso, [Estado das Notícias 2018](#). Lisboa, 2018. MediaLab e OberCom.

Ana Pinto-Martinho, Gustavo Cardoso, Décio Telo, Barómetro de Notícias ISCTE-IUL: 4 anos de análise ao destaque noticioso em Portugal. Lisboa, 2017. Comunicação no 4.º Congresso dos Jornalistas.

Gustavo Cardoso, Susana Santos, Décio Telo (Orgs.), [Jornalismo em tempo de crise](#). Lisboa, 2016. Mundos Sociais.

António Manuel Pinheiro Santos, [Imprensa como indicador: a representação dos temas económicos nos media portugueses](#). Lisboa: ISCTE-IUL, 2016. Dissertação de mestrado.

Décio Telo, [A problemática do acontecimento na produção jornalística: uma análise a partir das notícias sobre a 'crise no PS'](#). Lisboa: ISCTE-IUL, 2014. Dissertação de mestrado.

Susana Costa Santos, Carlota Pina Bicho, [Eleições autárquicas 2.0: análise das estratégias de comunicação online de candidatos, partidos e movimentos independentes](#), Lisboa: SOCIOLOGIA, PROBLEMAS E PRÁTICAS, n.º 81, 2016, pp. 189-210.

Licença de utilização



Este trabalho está licenciado sob a Licença Creative Commons Atribuição 4.0 Internacional. Para ver uma cópia desta licença, visite <http://creativecommons.org/licenses/by/4.0/> ou envie uma carta para Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.