

## Relatório de agosto de 2021

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

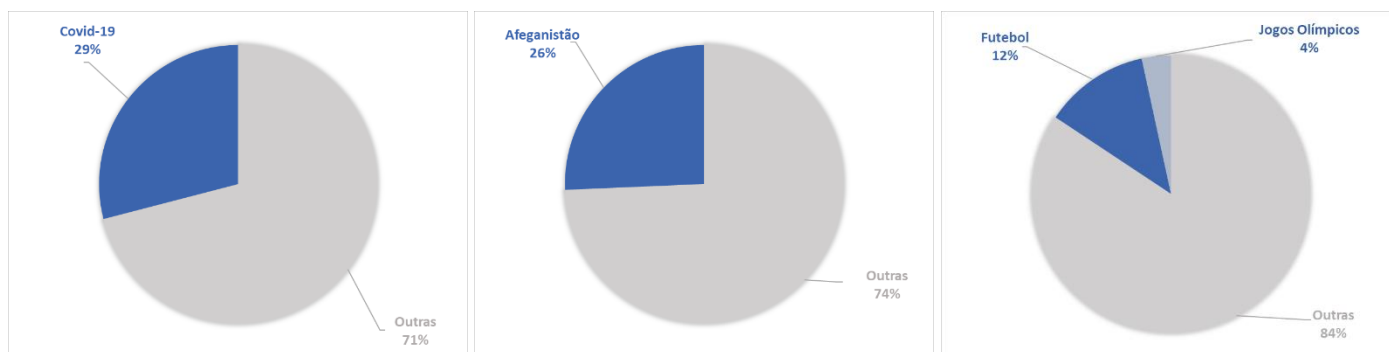
### Síntese

A vacinação foi o tema dominante nas notícias sobre Covid-19. A retirada norte-americana do Afeganistão quase superou a pandemia no que diz respeito à cobertura jornalística em agosto e, finalmente, uma boa notícia destacou-se pela sua ausência: os ‘incêndios florestais’ mantiveram-se fora dos principais destaques noticiosos.



Da esquerda: Jovens aguardam para serem vacinados em Cascais, 9/8/2021 (ANTÓNIO COTRIM/LUSA); Evacuação no aeroporto de Cabul, 23/8/2021 (Sgt. SAMUEL RUIZ/EPA/LUSA); Benfica em ação frente ao PSV Eindhoven no Estádio da Luz em Lisboa, 9/7/2021 (MÁRIO CRUZ/LUSA).

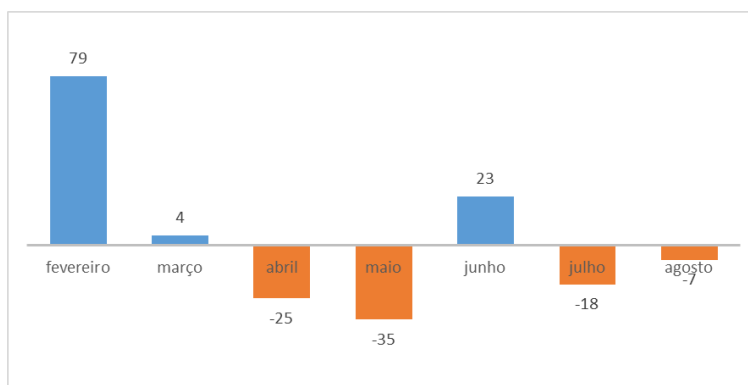
### As ‘histórias’ que marcaram as notícias online



Esquerda: Artigos relacionados com pandemia. Centro: Artigos relacionados com Afeganistão. Direita: Artigos relacionados com futebol e Tóquio 2020. Percentagens no conjunto de artigos considerados no âmbito deste barómetro (N=13.028 documentos). Fonte: MediaLab/Priberam (dados recolhidos a 6/9/2021).

### A pandemia volta a mostrar sinais de redução do seu valor-notícia

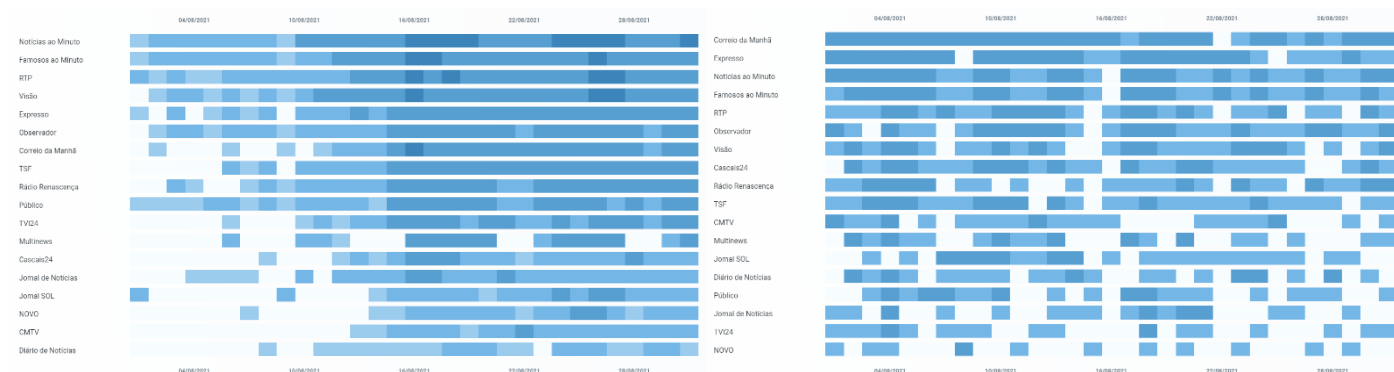
Em agosto, a pandemia provocada pelo vírus SARS-CoV-2 gerou, em média, 122 artigos por dia nos sites de notícias generalistas, menos 9 do que em julho. Mantém-se a tendência de descida na variação mensal face ao mês anterior.



Variação mensal face ao mês anterior, em percentagem, de artigos sobre Covid-19. fonte: MediaLab/Priberam (dados atualizados em 6/9/2021).

## Afganistão foi o tema ‘quente’ no mês de agosto

As imagens dramáticas, durante a evacuação de cidadãos estrangeiros e pessoal de embaixadas no Afeganistão, representaram o culminar da decisão norte-americana de retirada das suas tropas do país, que teve início a 5 de julho com o abandono da base militar de Bagram<sup>1</sup>. No entanto, foi a partir de 16 de agosto que o tema atingiu um pico de atenção massiva nos sites noticiosos em Portugal.



Esquerda: Mapa de intensidade (heatmap) dos artigos com referência ao termo “Afeganistão” em agosto. Direita: Mapa de intensidade dos artigos com referência ao termo “covid” em agosto. Fonte: Priberam (dados recolhidos em 9/9/2021).

Em comparação, os artigos que referem a palavra “covid” apresentam maior dispersão e menos ‘picos’ de intensidade na cobertura jornalística, conforme ilustrado na visualização acima.

## Desporto quase que se resume a futebol nos sites de informação generalista

O desporto voltou a ser tema relevante nos sites de informação generalista, apesar da redução significativa da cobertura dos Jogos Olímpicos de Tóquio. Se, nos primeiros 8 dias de competição, os dados apontavam para uma média de 30 artigos por dia sobre a competição, no mês de agosto essa média baixou para 14.

Em contrapartida, o futebol esteve associado, em média, a 52 artigos/dia. Mais 24 artigos por dia do que em julho. Em grande parte, este recentrar da atenção deve-se ao regresso das competições, com destaque para a participação do Sport Lisboa e Benfica na fase de *play-off* da Liga dos Campeões da UEFA e para as transferências de Lionel Messi e Cristiano Ronaldo.

## Nota metodológica sobre a análise dos clusters de notícias

O quadro abaixo permite visualizar a proporção de cada agrupamento de notícias (*cluster*) no conjunto dos 44 *clusters* que mais se destacaram em agosto de 2021. Estes, resultam de um trabalho de análise e reagrupamento semântico realizado no conjunto mais alargado, de 50 *clusters*<sup>2</sup>, 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 algoritmo<sup>3</sup>.

<sup>1</sup> <https://apnews.com/article/bagram-afghanistan-airfield-us-troops-f3614828364f567593251aaaa167e623>.

<sup>2</sup> Para mais informação, consultar a secção de metodologia no final do relatório.

<sup>3</sup> 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.



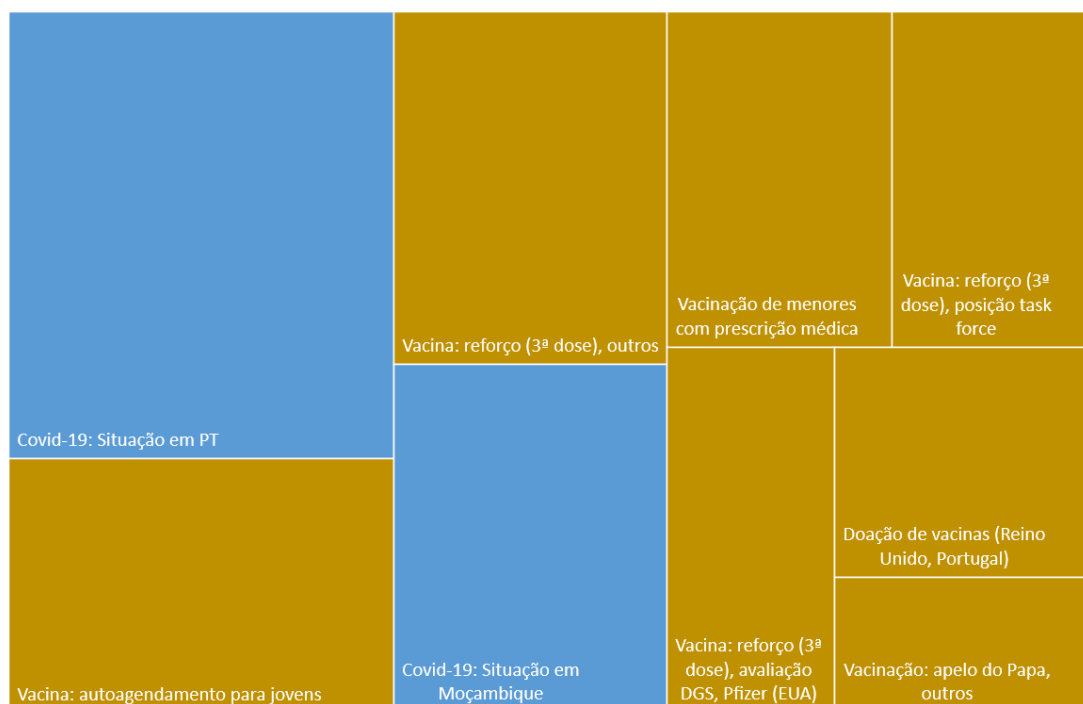
Os 'clusters' relacionados com a pandemia estão identificados pela cor azul, desporto com a cor verde e Afeganistão a laranja. Quadro construído com Índice de base = 100 em que a variável de interesse é o cluster mais publicado: "Covid-19: situação em Portugal". Fonte: MediaLab/Priberam (dados recolhidos em 6/9/2021)

### A narrativa da pandemia nas notícias

A análise identificou, em agosto, 9 tópicos<sup>4</sup> relacionados com a pandemia, o valor mais baixo até à data. Esta concentração temática traduziu-se, ainda, numa atenção muito focada no processo de vacinação.

Os dados disponíveis parecem apontar para a redução sistemática da atenção jornalística. A hipótese de uma terceira dose manteve-se muito presente nas notícias sobre a pandemia, mas a situação epidemiológica, maioritariamente em Portugal, voltou a ser o *cluster* que reuniu o maior número de artigos.

<sup>4</sup> Para mais informação consulte a metodologia no final do relatório.



Menos temas, maior concentração na vacinação (cor dourada) e evolução da Covid-19 em Portugal e no Mundo (cor azul).  
 Fonte: MediaLab/Priberam (dados recolhidos em 6/9/2021)

No que respeita a vacinação, manteve-se uma dupla abordagem, do ponto de vista narrativo. Por um lado, a situação da vacinação entre as camadas mais jovens da população portuguesa gerou muitos artigos orientados para informação factual, sobre regras e processo (“Autoagendamento da vacinação exclusivo para jovens com 16 ou 17 anos”, Diário de Notícias, 03/08/2021; “Vacinação de menores de 12 a 15 anos saudáveis só com prescrição médica”, Multinews, 02/08/2021). Por outro lado, o reforço da vacinação através da administração de uma terceira dose colocou em evidência o papel da validação científica como argumento para a tomada de decisão em Portugal (“Task force afasta terceira dose sem “evidência científica” da sua necessidade”, Diário de Notícias, 06/08/2021) ou para a escolha do público-alvo num cenário de reforço da vacinação (“Anticorpos em imunodeprimidos crescem mais 43% após terceira dose da vacina”, Jornal de Notícias, 14/08/2021; “Alemanha prepara dose de reforço para as pessoas mais vulneráveis”, Expresso, 01/08/2021).

### Descrição da amostra

No mês de agosto foram considerados 13.028 artigos, de 16 fontes de informação jornalística<sup>5</sup> para a realização deste relatório, selecionadas a partir de um universo de mais de 160 entidades (aproximadamente 237.336 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*<sup>6</sup> apresentado neste relatório.

<sup>5</sup> CMTV, Correio da Manhã, Diário de Notícias, Expresso, Jornal de Notícias, Jornal SOL, Multinews, Notícias ao Minuto, NOVO, Observador, Público, Rádio Renascença, RTP, TSF, TVI24, Visão.

<sup>6</sup> Para mais informações sobre metodologia consultar secção no final do relatório.

Do ponto de vista temático, a ‘política’ mantém-se a área mais representada nas notícias online. Dos cinco temas mais relevantes nas notícias online, apenas as ‘questões sociais’ subiram uma posição por troca com os artigos sobre ‘saúde’.

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

Fonte: MediaLab/Priberam (dados recolhidos em 7/9/2021)

António Costa perdeu a ‘liderança’ no que diz respeito à personalidade mais referida em artigos. A responsabilidade pelas alterações verificadas no ranking das ‘personalidades’ recai, sobretudo, sobre dois acontecimentos que marcaram o mês: A situação no Afeganistão e a saída de Lionel Messi do Barcelona (futebol). Quanto a entidades coletivas, a situação manteve-se inalterada, face ao mês anterior.

Personalidades	Ranking
Joe Biden	4
António Costa	-1
Marcelo Rebelo de Sousa	0
Lionel Messi	17
Osama bin Laden	16

Fonte: MediaLab/Priberam (dados recolhidos em 7/9/2021)

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

Fonte: MediaLab/Priberam (dados recolhidos em 7/9/2021)

## 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.

Este relatório está disponível em versão para Web em <https://medialab.iscte-iul.pt/barometro/noticias/>.

## 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.