

## Relatório de maio de 2021

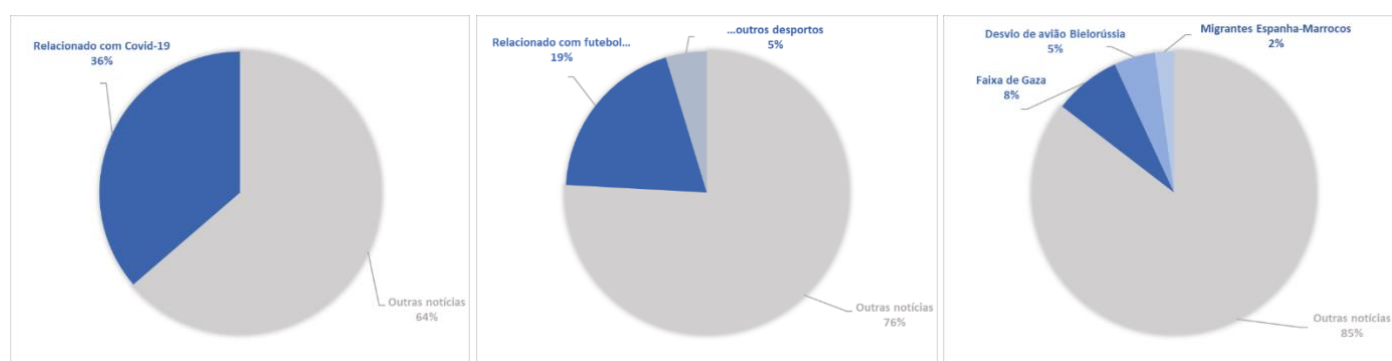
### Síntese

A **pandemia** está a perder valor noticioso. Desde março que o número de notícias publicadas online tem vindo a decrescer. Em sentido inverso, o **futebol** tem vindo a receber atenção mediática crescente. O mês de maio fica, também, marcado por três acontecimentos mediáticos internacionais, de grande impacto: O regresso da violência na **Faixa de Gaza** aos *media* portuguesas, o fluxo inesperado de **migrantes e refugiados** em Espanha e o desvio ilegal de um avião comercial pelas autoridades da **Bielorrússia**.



Da esquerda: Vice-Almirante Henrique Gouveia e Melo na sessão de apresentação da situação epidemiológica (MIGUEL A. LOPES/LUSA), adeptos ingleses no Porto (EPA/MANUEL FERNANDO ARAUJO), palestinianos em torno de uma fogueira junto das suas casas destruídas em Beit Hanoun (EPA/MOHAMMED SABER).

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



Confirma-se a tendência, há menos notícias online sobre a pandemia

A pandemia continua a ser o tema noticioso que recebe mais atenção mediática<sup>1</sup> mas os dados demonstram que, desde o início do desconfinamento, assiste-se a uma redução do número de artigos publicados. Os dados disponíveis apontam para uma variação negativa na ordem dos -35% no que respeita o número de artigos publicados. Em abril tinha sido na ordem dos -27%.

Em média, foram publicados 131 artigos por dia sobre a pandemia, o que, apesar de ser um valor absoluto significativo, corresponde a uma variação negativa de -37% face ao mês anterior.

Em termos relativos, as notícias sobre Covid-19 representaram pouco mais de um terço dos artigos considerados na amostra<sup>2</sup>. Em abril tinha representado, aproximadamente, metade dos artigos.

<sup>1</sup> A atenção jornalística é entendida, neste contexto, pelo número de artigos publicados em órgãos de comunicação social online.

<sup>2</sup> No mês de abril foram selecionados, na plataforma Priberam, 11.207 artigos publicados em 16 empresas jornalísticas e noticiosas de âmbito nacional e generalista (imprensa não especializada). Poderá consultar mais detalhes sobre a metodologia no final do relatório.

### O peso do futebol cresceu na cobertura generalista online

Os dados disponíveis permitem estimar que, em 10 artigos publicados no mês de abril em websites de notícias generalistas, 2 foram sobre futebol. Este valor corresponde a um crescimento da atenção noticiosa em comparação com o mês anterior, embora estejamos a falar de uma variação ténue (variação de 4%).

O mês de maio coincidiu com finais de torneios e de época desportiva e, em especial, com a polémica final da Champions League, o mais importante troféu europeu de futebol, que se realizou no Estádio do Dragão, no Porto.

### Menos acidentes, mais grandes acontecimentos internacionais

Em maio o peso dos acidentes e ‘casos de polícia’ nas notícias online desceu ligeiramente, para um valor médio de 44 notícias por dia, mas, em contrapartida, três grandes acontecimentos internacionais negativos – violência na Faixa de Gaza, desvio de avião na Bielorrússia, vaga de migrantes e refugiados em Espanha – representaram 15% dos artigos considerados na análise. De entre eles, o regresso da violência na Faixa de Gaza foi a ‘história’ noticiosa que se destacou com três clusters temáticos a marcar a narrativa jornalística: os bombardeamentos israelitas, a reunião de emergência da União Europeia e as detenções e número de feridos palestinianos em Jerusalém, na sequência de confrontos com a polícia israelita.

### Outras ‘histórias’ noticiosas destacadas pela comunicação social online

A política nacional esteve presente em diversos clusters identificados pelo algoritmo da plataforma Priberam. Apesar de não ter sido possível identificar uma ‘história’ noticiosa específica, com grande destaque, o conjunto de artigos relacionados com política nacional e partidária foi, foi em termos estimados, cerca de 1 em cada 10 artigos publicados em maio.

Dois ‘casos’ saíram da agenda mediática com a mesma velocidade com que entraram: os casos “Ihor-SEF” e “Sócrates-Op. Marquês” não foram alvo de qualquer desenvolvimento relevante em maio.



Os ‘clusters’ relacionados com a pandemia estão identificados pela cor azul, os quadros associados ao desporto com a cor verde e os que estão relacionados com acidentes, desastres e casos de polícia, foram identificados com cor 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: “Boletim epidemiológico PT”.

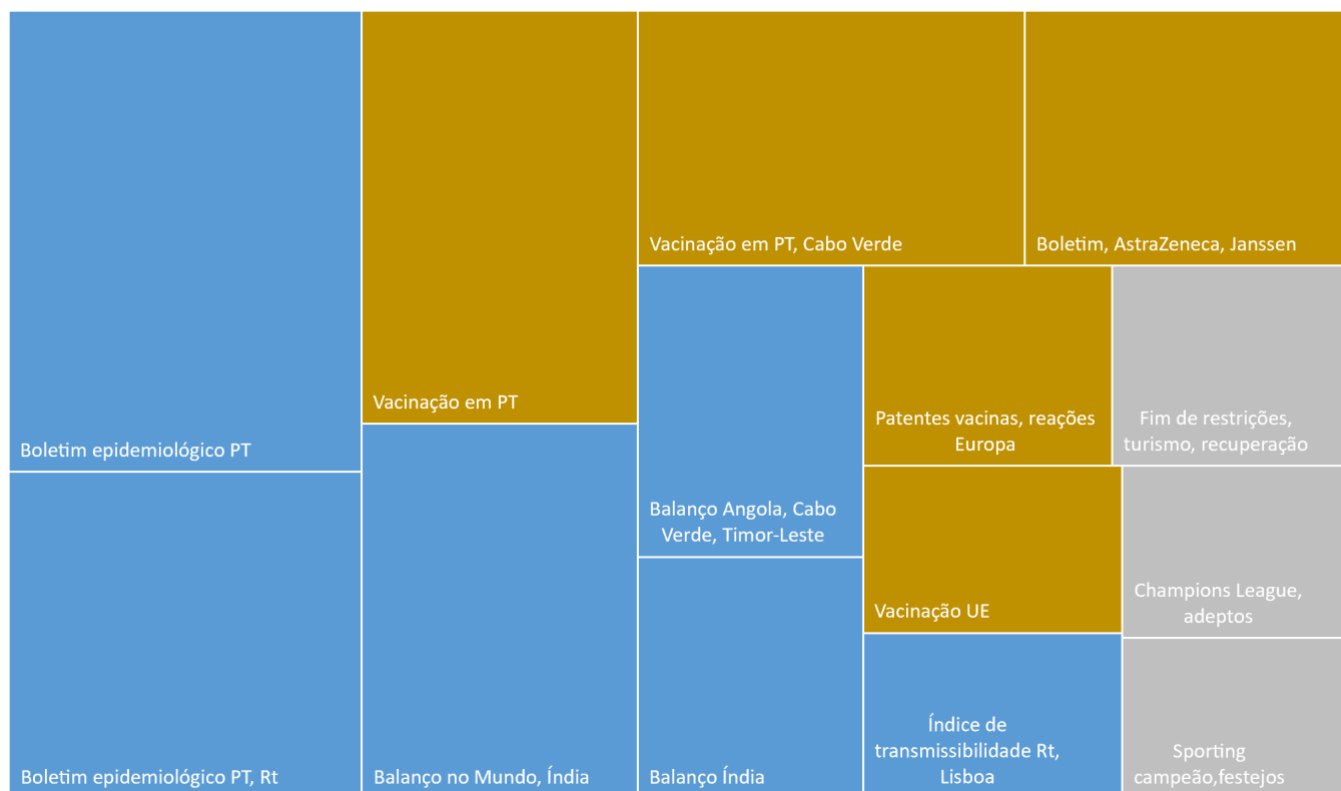
## Nota metodológica sobre a análise dos *clusters* de notícias

O quadro acima permite visualizar a proporção de cada agrupamento de notícias (*cluster*) no conjunto dos 46 *clusters* que mais se destacaram em maio de 2021. Estes, resultam de um trabalho de análise e reagrupamento semântico realizado num conjunto mais alargado, de 50 *clusters*<sup>3</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 da plataforma Priberam<sup>4</sup>.

## A narrativa da pandemia nas notícias

A análise identificou, em maio, 14 tópicos<sup>5</sup> relacionados com a pandemia, o mesmo número do mês anterior. Dois grandes grupos temáticos destacaram-se, em maio: os boletins com balanços diários de casos a nível nacional e internacional e o processo de vacinação (igualmente em duas vertentes: cobertura nacional, situação no mundo).

O foco nos relatórios epidemiológicos, na cobertura mediática da pandemia nas notícias online, foi ainda mais expressivo do que o verificado no mês anterior. A preocupação com países de língua portuguesa manteve-se, apesar de não ter sido identificado nenhum *cluster* sobre a pandemia no Brasil. A situação da pandemia na Índia recebeu alguma atenção, com foco nos balanços e evolução de casos.



As notícias sobre Covid-19 subdividem-se em três grupos temáticos: a evolução da pandemia em Portugal e no Mundo (cor azul), Vacinação (cor dourada) e outros temas (cor cinza).

<sup>3</sup> Para mais informação, consultar a secção de metodologia no final do relatório.

<sup>4</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.

<sup>5</sup> Para mais informação consulte a metodologia no final do relatório.

Quanto às polémicas com a vacina AstraZeneca, ao contrário do que se verificou no mês anterior<sup>6</sup>, não se identificou nenhuma ‘história’ noticiosa relevante e o tema surgiu num cluster que reuniu notícias sobre a vacina Janssen, entre outras.

Os *clusters* associados à vacinação em Portugal relacionam-se, sobretudo, com balanços e pontos de situação do processo de vacinação.

### Descrição da amostra

No mês de abril foram considerados 11.207 artigos, de 16 fontes de informação jornalística<sup>7</sup> para a realização deste relatório, selecionadas a partir de um universo de mais de 160 entidades (aproximadamente 215.859 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 46 *clusters*<sup>8</sup>.

O Notícias ao Minuto foi a entidade com mais artigos publicados com, aproximadamente, 8.300 publicações no período em análise. Seguiram-se os websites do Correio da Manhã, com menos 23% de publicações, e RTP com menos 45%.

Do ponto de vista temático, a ‘política’ foi a área temática mais representada nas notícias online, tal como em abril, com aproximadamente 16.700 artigos publicados. Em comparação com o mês anterior, verificou-se a saída do tema ‘saúde’ do ‘top 5’.

Categoria temática (IPTC)	Índice 100
Política	100
Economia, negócios e finanças	85
Questões sociais	56
Desporto	56
Crime, lei e justiça	45

António Costa continuou a ser a personalidade mais referida em maio, seguido de Marcelo Rebelo de Sousa com menos 40% de referências nas notícias online. No que respeita entidades coletivas, Governo e Estado lideraram as referências em artigos publicados online.

Fonte (órgão de comunicação social)	Índice 100
António Costa	100
Marcelo Rebelo de Sousa	60
Joe Biden	48

Entidades	Índice 100
Governo	100
Estado	56
Sporting CP	34

<sup>6</sup> Em abril foram identificadas 3 narrativas associadas à vacina AstraZeneca: ‘mortes no Reino Unido’, ‘decisão da Agência Europeia do Medicamento (EMA)’, ‘contrato entre a farmacêutica e a União Europeia’

<sup>7</sup> CMTV, Correio da Manhã, Diário de Notícias, Expresso, Jornal de Notícias, Jornal i, Jornal SOL, Notícias ao Minuto, Observador, Público, Rádio Renascença, RTP, SIC Notícias, TSF, TVI24, Visão.

<sup>8</sup> 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, sejam elas 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.