

Relatório de setembro de 2021

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

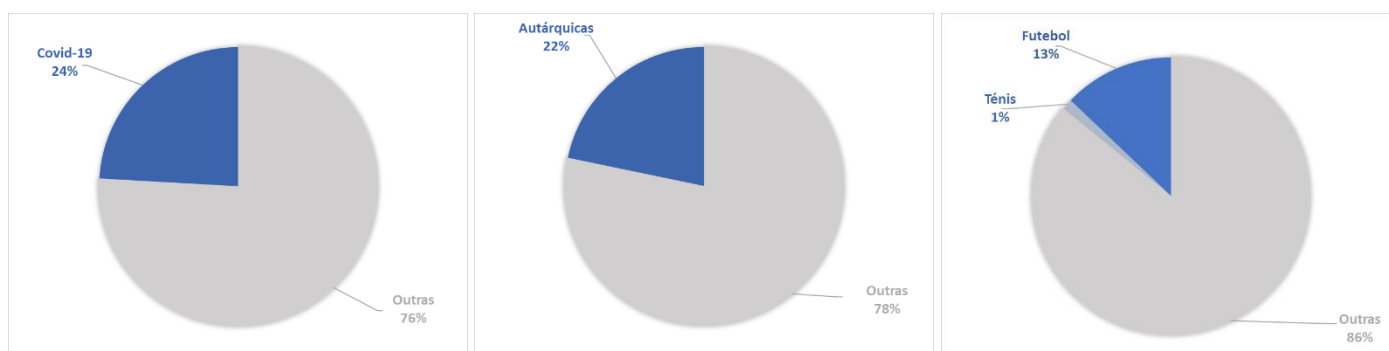
Síntese

A morte de Jorge Sampaio foi o acontecimento de setembro que gerou maior número de artigos na imprensa *online*. A atenção jornalística à evolução da pandemia mantém-se em curva descendente mas superou, ainda assim, as eleições autárquicas de 26 de setembro.



Da esquerda para a direita: Cerimónia fúnebre de Jorge Sampaio no Alto de São João, Lisboa, 12/9/2021 (EPA/RODRIGO ANTUNES/LUSA); Intervenção da ministra da Saúde no Infarmed, Lisboa, 16/9/2021 (MIGUEL A. LOPES/LUSA); Votação na freguesia da Sé no Funchal, 26/9/2021 (HOMEM DE GOUVEIA/LUSA).

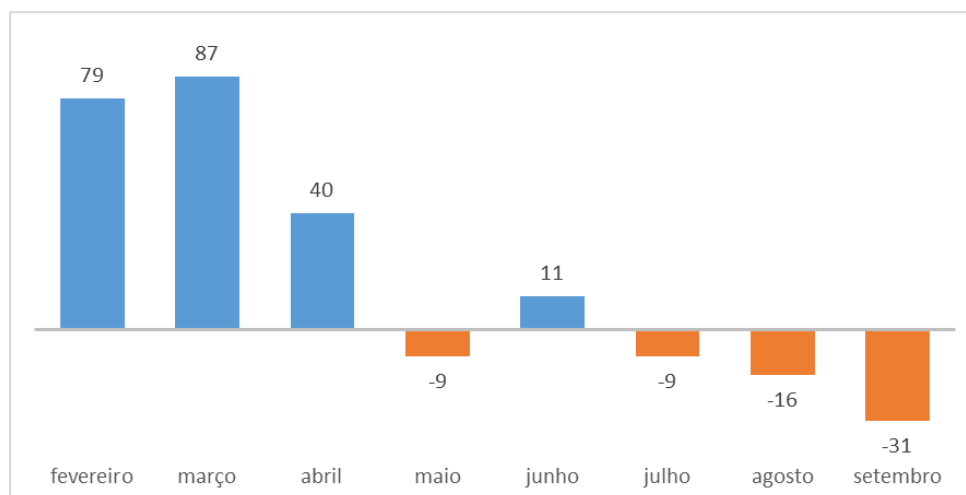
As ‘histórias’ que marcaram as notícias online



Esquerda: Artigos relacionados com a pandemia. Centro: Artigos relacionados com eleições autárquicas. Direita: Artigos relacionados com futebol. Percentagens no conjunto de artigos considerados no âmbito deste barómetro (N=10.816 documentos). Fonte: MediaLab/Priberam (dados recolhidos a 14/10/2021).

O valor-notícia da pandemia continua num ciclo descendente

Em setembro, a situação da Covid-19 gerou, em média, 87 artigos por dia nos *sites* de notícias generalistas, menos 35 do que em agosto. A tendência de descida na variação mensal face ao mês janeiro, acentuou-se significativamente.



Variação mensal, em percentagem, face a janeiro de 2021, de artigos sobre Covid-19. fonte: MediaLab/Priberam (dados atualizados em 14/10/2021).

As eleições autárquicas estiveram presentes em múltiplos *clusters* de ‘histórias’ noticiosas

A cobertura noticiosa das eleições autárquicas foi multifacetada, com grande dispersão temática e os dados organizados pelo algoritmo confirmam-no. As ações de campanha surgem associadas a muitos clusters e, de uma forma geral, a observação permite confirmar a forte presença de líderes dos principais partidos políticos envolvidos em disputas partidárias ‘nacionais’.

Desporto foi, quase exclusivamente, futebol nos sites de informação generalista

Foram publicados 47 documentos por dia sobre futebol, em setembro, de acordo com os dados da plataforma Priberam. Para além do desporto mais popular em Portugal, apenas o ténis, por via do Open dos Estados Unidos, conseguiu marcar presença nos 50 clusters de histórias noticiosas mais relevantes.



Os ‘clusters’ relacionados com a pandemia estão identificados pela cor azul, desporto com a cor verde e eleições autárquicas a laranja. Quadro construído com Índice de base = 100 em que a variável de interesse é o cluster mais publicado: “Jorge Sampaio”. Fonte: MediaLab/Priberam (dados recolhidos em 14/10/2021)

Nota metodológica sobre a análise dos *clusters* de notícias

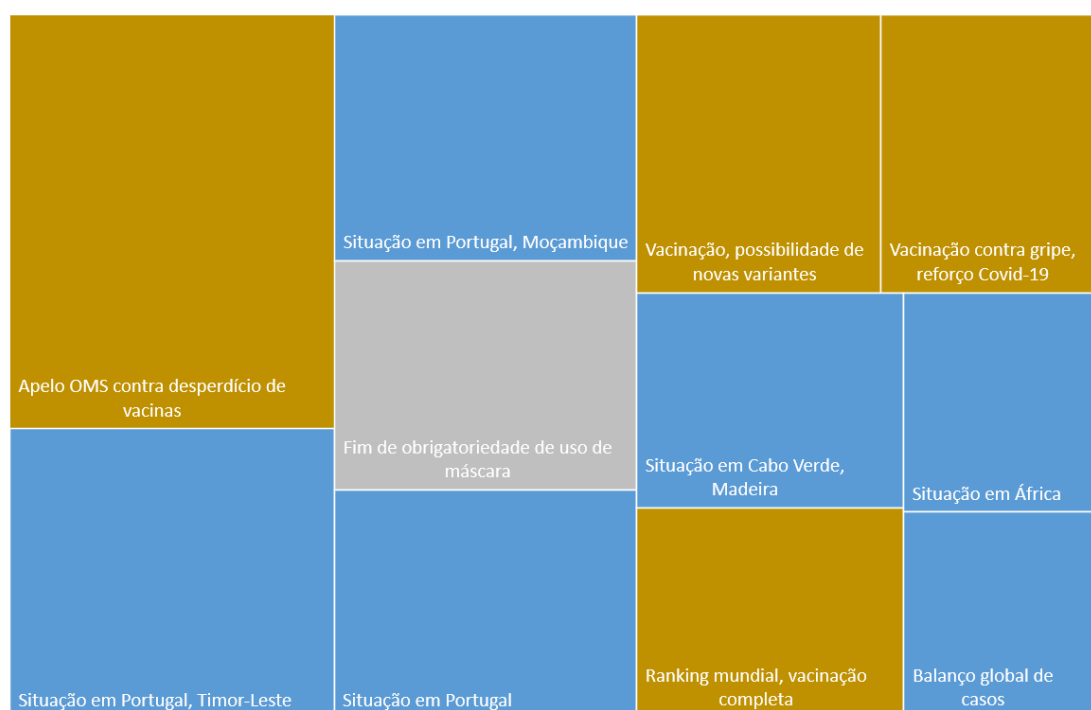
O quadro abaixo permite visualizar a proporção de cada agrupamento de notícias (*cluster*) no conjunto dos 49 *clusters* que mais se destacaram em setembro de 2021. Estes, resultam de um trabalho de análise e reagrupamento semântico

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 agosto, 13 tópicos³ relacionados com a pandemia. Contrariando a tendência do mês anterior, a atenção mediática foi menos focada no processo de vacinação, sendo que a situação epidemiológica em Portugal, Timor-Leste e África (especialmente em países de língua oficial portuguesa) foram alvo de grande quantidade de artigos.

Não obstante, o apelo da Organização Mundial de Saúde contra o desperdício de vacinas de reforço para pessoas saudáveis, com o objetivo de reduzir a desigualdade mundial na distribuição de vacinas, foi o *cluster* que reuniu, isoladamente, o maior número de artigos.



Distribuição mais equitativa, em comparação com o mês anterior, entre clusters sobre vacinação (cor dourada) e evolução da pandemia em Portugal e no Mundo (cor azul). Fonte: MediaLab/Priberam (dados recolhidos em 14/10/2021).

Descrição da amostra

No mês de setembro foram considerados 10.816 artigos, de 18 fontes de informação jornalística⁴ para a realização deste relatório, selecionadas a partir de um universo de mais de 160 entidades (aproximadamente 248.520 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

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.

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

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 clusters⁵ apresentado neste relatório.

Do ponto de vista temático, apenas a 'política' mantém-se como área mais representada nas notícias online. Registou-se a subida de notícias associadas a questões sociais, em parte por via da morte de Jorge Sampaio e relacionados com crime, lei e justiça.

Categoria temática (IPTC)	Ranking
Política	0
Questões sociais	2
Economia, negócios e finanças	-1
Crime, lei e justiça	2
Desporto	-2

Fonte: MediaLab/Priberam (dados recolhidos em 14/10/2021)

António Costa voltou a 'liderar' o ranking no que diz respeito à personalidade mais referida em artigos. O contexto das eleições autárquicas explica, também, a subida de outros líderes políticos nacionais no ranking e, ainda, a presença de Jorge Sampaio, relacionadas com a sua morte.

Personalidades	Ranking
António Costa	1
Marcelo Rebelo de Sousa	1
Jorge Sampaio	18
Rui Rio	18
Joe Biden	-4

Fonte: MediaLab/Priberam (dados recolhidos em 14/10/2021)

Entidades	Ranking
Governo	0
Partido Socialista	3
Partido Social Democrata	4
CDS - Partido Popular	7
Direção-Geral de Saúde	-2

Fonte: MediaLab/Priberam (dados recolhidos em 14/10/2021)

O mesmo acontece com as entidades coletivas. Os partidos políticos subiram no ranking, fruto do aumento de referências nos textos jornalísticos em mês de eleições e o Governo mantém-se, invariavelmente, como a entidade mais referida nas notícias online.

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.

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.