

Relatório de outubro de 2021

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

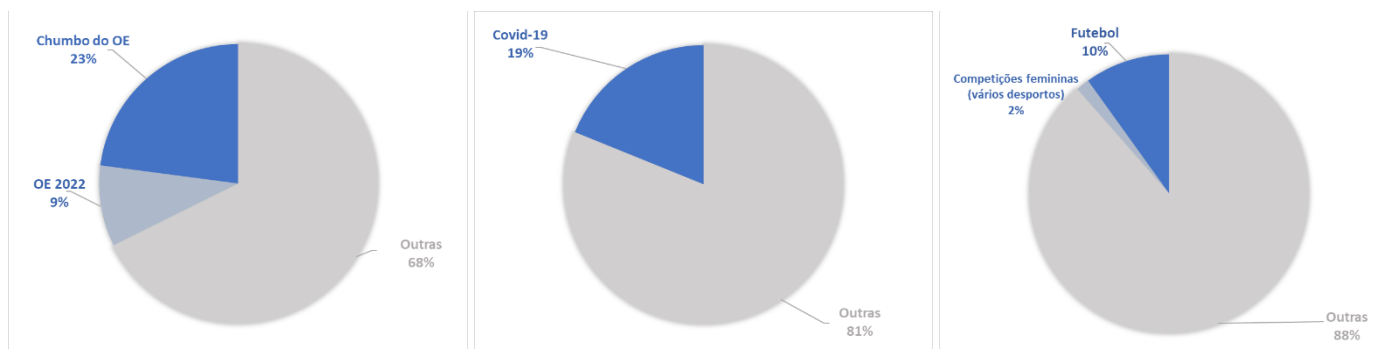
Síntese

O chumbo do Orçamento do Estado (OE2022) na Assembleia da República (AR) dominou a atenção na comunicação social online e faz com que, pela primeira vez este ano, a situação pandémica não seja o tema dominante do mês.



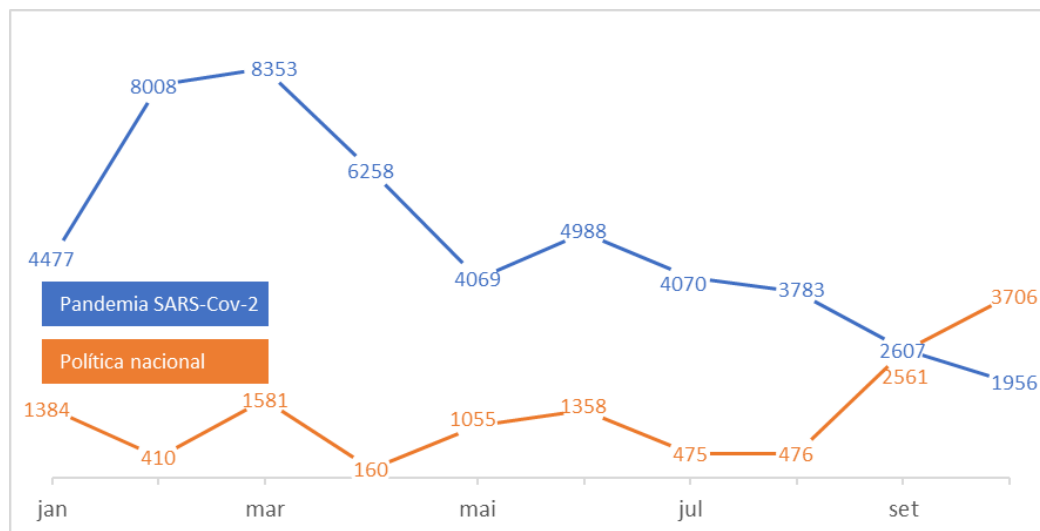
Da esquerda para a direita: Jerónimo de Sousa durante conferência de imprensa, Lisboa, 25/10/2021 (TIAGO PETINGA/LUSA); Graça Freitas em conferência de imprensa, 8/10/2021 (TIAGO PETINGA/LUSA); Comissão do BE no Centro Hospitalar de Setúbal, 12/10/2021 (RUI MINDERICO/LUSA).

As ‘histórias’ que marcaram as notícias online



Esquerda: Artigos relacionados com chumbo do OE na AR. Centro: Artigos relacionados com Covid-19. Direita: Artigos relacionados com Desporto. Percentagens referentes ao total de artigos considerados no âmbito deste barómetro (N=10.357 documentos). Fonte: MediaLab/Priberam (dados recolhidos a 5/11/2021).

Do Orçamento do Estado à dissolução do Parlamento, a política nacional é o tema mais noticiado. Outubro de 2021 tem grande potencial para ficar registado na história pela reprovação, na generalidade, do OE 2022 – pela primeira vez na história da democracia – e pela dissolução do Parlamento por iniciativa presidencial.

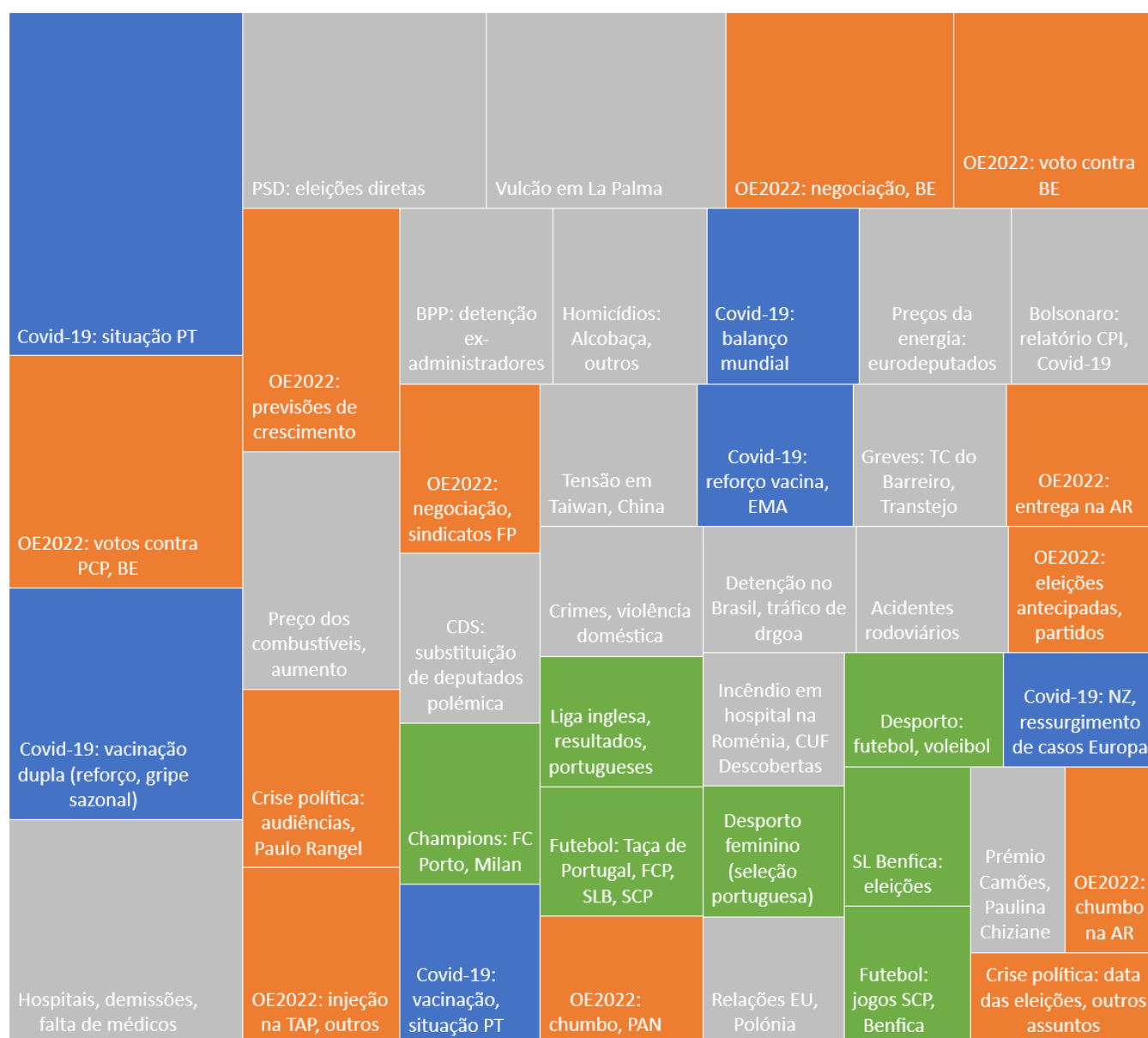


Estimativa do número de artigos publicados online sobre Covid-19 e política nacional desde janeiro de 2021. Fonte: MediaLab/Priberam.

O ‘chumbo’ na Assembleia da República ofuscou o debate público sobre as medidas do Orçamento do Estado. A cobertura noticiosa do Orçamento do Estado (OE) começou o mês conforme previsto na agenda noticiosa, com notícias relacionadas com o protocolo e negociações habituais (“Governo apresenta proposta de Orçamento do Estado para 2022”, rtp.pt, 12/10/2021) mas rapidamente foi dominada pela ameaça e confirmação da reprovação na Assembleia da República (“Orçamento do Estado para 2022 chumbado no Parlamento”, cmjornal.pt, 25/10/2021).

Desporto no feminino

O algoritmo da plataforma Priberam identificou um *cluster* de notícias relacionadas com desporto feminino. O facto é merecedor de destaque, ainda mais tratando-se de um resultado a partir de uma amostra de sites de notícias generalistas (ou seja, não inclui órgãos de comunicação social ‘desportivos’). A notícia mais representativa deste *cluster* foi sobre a conquista da medalha de bronze no campeonato europeu de ténis de mesa (“Portugal conquista bronze no Europeu feminino”, rr.sapo.pt, 02/10/2021) mas houve também destaque para o andebol (“Seleção portuguesa de andebol feminino perde na Hungria 34-24 rumo ao Euro2022”, observador.pt, 6/10/2021) e, obviamente, para o futebol (“Seleção lusa de futebol feminino sub-17 vence Eslováquia em jogo de preparação”, cmjornal.pt, 5/10/2021).



Os ‘clusters’ relacionados com a pandemia estão identificados pela cor azul, desporto com a cor verde e OE2022-AR a laranja. Quadro construído com Índice de base = 100 em que a variável de interesse é o cluster mais publicado: “Covid-19: situação PT”. Fonte: MediaLab/Priberam (dados recolhidos em 5/11/2021)

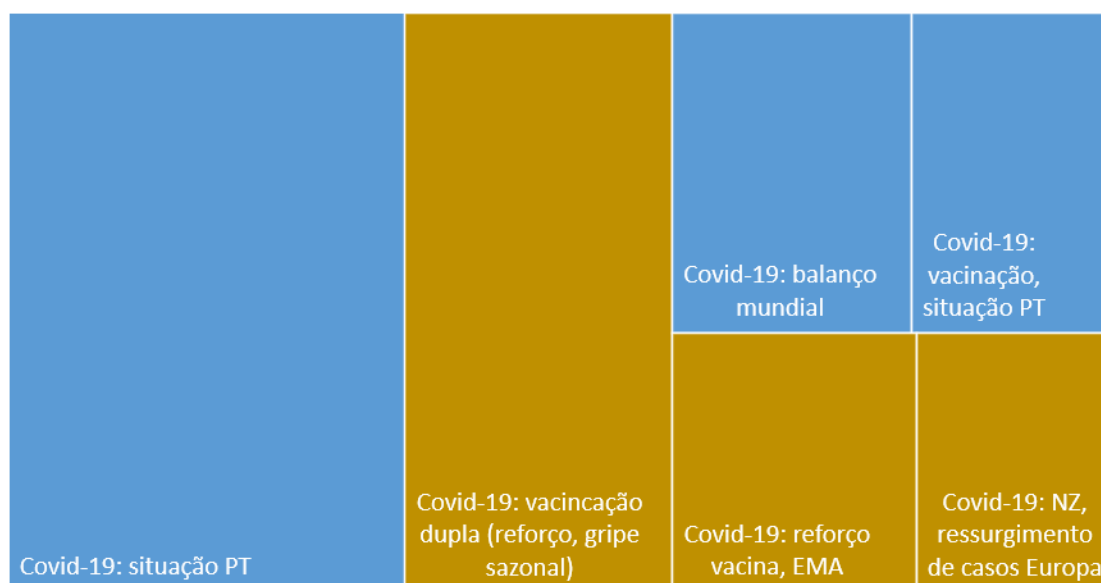
Nota metodológica sobre a análise dos *clusters* de notícias

O quadro anterior permite visualizar a proporção de cada agrupamento de notícias (*cluster*) no conjunto dos 42 *clusters* que mais se destacaram em outubro de 2021. Estes resultam de um trabalho de análise e reagrupamento semântico realizado no conjunto, mais alargado, de 50 *clusters*¹, identificado 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 cobertura noticiosa da pandemia tem vindo a perder peso nas agendas das redações, um pouco por todo o mundo. A análise do barómetro ilustra a redução sistemática de temas e assuntos destacados na comunicação social.

Em outubro, o algoritmo identificou 6 tópicos³ relacionados com a pandemia, aproximadamente metade dos que tinha identificado no mês anterior. A cobertura está, agora, praticamente reduzida a dois grandes focos de atenção: o processo de vacinação e a situação epidemiológica em Portugal e no Mundo. No primeiro caso, o artigo “Covid-19: R(t) chega a 1 e incidência também sobe, em dia com sete mortes e 766 infetados em Portugal” (expresso.pt, 15/10/2021) foi o mais representativo do conjunto de notícias sobre a situação epidemiológica. No caso da vacinação, o artigo com o título “Vacinação simultânea contra a gripe e covid-19. Marta Temido garante que «eficácia e segurança permanecem idênticas»” (expresso.pt, 18/10/2021) surge como referência do *cluster* sobre vacinação simultânea contra a gripe e Covid-19.



A situação epidemiológica volta a concentrar mais notícias (cor azul) mas a atenção continua repartida com as próximas fases do processo de vacinação em Portugal (cor dourada). Quadro construído com Índice de base = 100 em que a variável de interesse é o cluster mais publicado: “Covid-19: situação PT”. Fonte: MediaLab/Priberam (dados recolhidos em 5/11/2021).

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 outubro foram considerados 10.357 artigos, de 18 fontes de informação jornalística⁴ para a realização deste relatório, selecionadas a partir de um universo de mais de 160 entidades (aproximadamente 243.044 artigos).

A recolha de dados foi realizada na plataforma Priberam e a organização das histórias noticiosas em *clusters* é realizada no MediaLab 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 em que se aplica uma observação 'fina' aos dados, resultando no quadro final de *clusters*⁵ apresentado neste relatório.

Do ponto de vista temático, a 'política' cedeu a posição às 'questões sociais' que é, agora, a área mais representada nas notícias *online* (ver quadro abaixo, direita). Os restantes temas do 'top 5' mantêm-se nas mesmas posições.

Os cinco órgãos de comunicação social online com maior número de artigos publicados, na amostra, mantêm-se nas mesmas posições (quadro abaixo, esquerda) com exceção do site da revista Visão que subiu uma posição por troca com o Observador.

Órgão de comunicação social	Ranking
Notícias ao Minuto	0
RTP	0
Correio da Manhã	0
Visão	1
Observador	-1

Fonte: MediaLab/Priberam (dados recolhidos em 8/11/2021)

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

Fonte: MediaLab/Priberam (dados recolhidos em 8/11/2021)

António Costa e Marcelo Rebelo de Sousa foram (novamente) as personalidades mais presentes nos artigos estudados (quadro abaixo, esquerda). O contexto noticioso explica, também, a presença de 5 personalidades políticas no 'top 5' de nomes citados na imprensa online, em outubro.

Personalidades	Ranking
António Costa	0
Marcelo Rebelo de Sousa	0
Rui Rio	1
João Leão	17
Paulo Rangel	14

Fonte: MediaLab/Priberam (dados recolhidos em 8/11/2021)

Entidades	Ranking
Governo	0
Partido Socialista (PS)	0
Assembleia da República (AR)	7
Bloco de Esquerda (BE)	2
Partido Comunista Português (PCP)	4

Fonte: MediaLab/Priberam (dados recolhidos em 8/11/2021)

O mesmo acontece com as entidades coletivas (quadro acima, direita). A Assembleia da República subiu 8 posições e os partidos políticos que mais diretamente estiveram envolvidos nas negociações com o governo, e consequente reprovação do OE 2022 na Assembleia da República, receberam mais atenção do que no mês anterior.

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.

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.