

Relatório de maio de 2022

Por Ana Pinto-Martinho, Décio Telo e Gleice Luz, com supervisão de Gustavo Cardoso.

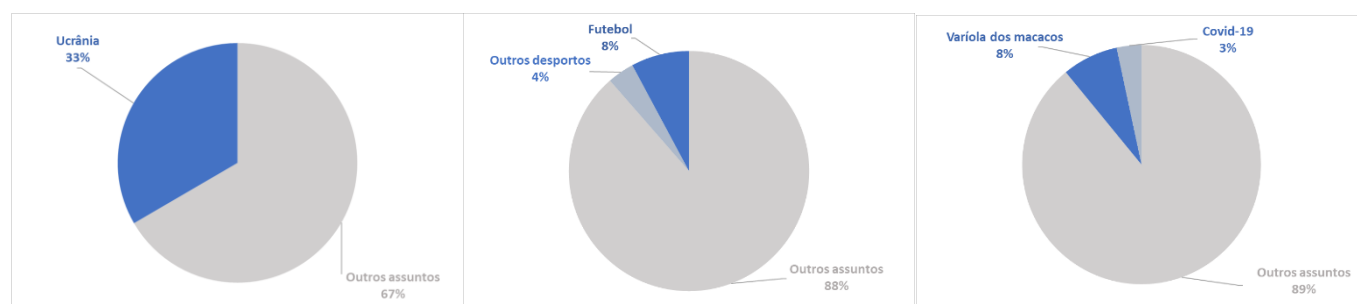
Síntese

A guerra na Ucrânia mantém-se como tema mais noticiado *online*, em órgãos de comunicação social generalistas. O massacre na escola Robb Elementary, no Texas, e a varíola dos macacos, em Portugal, geraram grande número de textos publicados mas, em maio, o campo das notícias caracteriza-se por uma certa fragmentação temática.



Da esquerda para a direita: Homenagem a vítimas em Uvalde, Texas, 28/05/2022 (EPA/TANNEN MAURY/LUSA); Volodymyr Zelensky em Kharkiv, Ucrânia, 29/05/2022 (EPA/Ukrainian Presidential Press Service); Marta Temido (D) e Lacerda Sales (E) na AR, Lisboa, 10/05/2022 (JOSÉ SENA GOULÃO/LUSA).

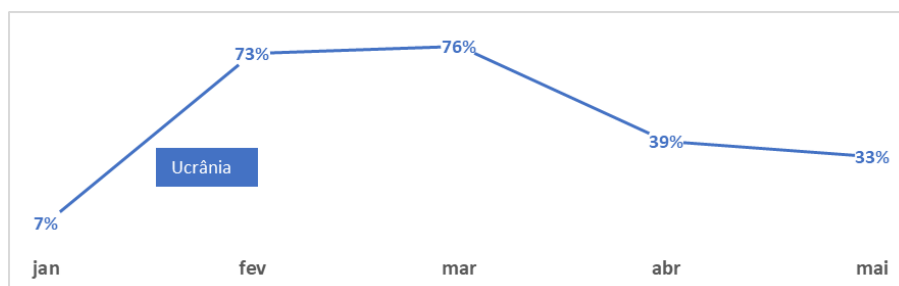
As ‘histórias’ que marcaram as notícias *online*



Esquerda: Artigos relacionados com Ucrânia. Centro: Artigos relacionados com desporto. Direita: Artigos relacionados com saúde. Percentagens referentes ao total de artigos considerados neste barómetro (N = 4719 documentos). Fonte: MediaLab/Priberam, dados recolhidos a 17/6/2022.

O destaque da guerra num mês pouco tematizado

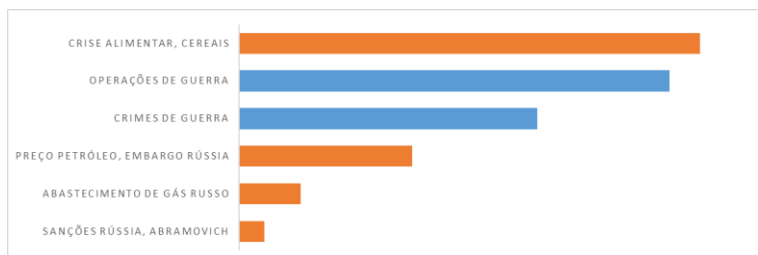
Estima-se que uma em cada três notícias publicadas *online* nos órgãos de comunicação social generalistas terá sido sobre a guerra na Ucrânia, com base nos dados disponíveis. Se considerarmos o peso relativo no conjunto dos documentos considerados na análise, podemos estar perante uma certa estabilização quanto ao número de textos publicados mensalmente sobre a Ucrânia.



Peso relativo do tema 'Ucrânia' no total de artigos considerados na análise. Fonte: MediaLab/Priberam.

Identificámos duas matérias predominantes na narrativa mediática sobre a guerra na Ucrânia, em maio. Metade dos documentos incidem sobre operações e ações militares (*clusters* destacados a azul no gráfico abaixo). A outra metade é constituída por quatro *clusters* povoados com ‘histórias’ sobre as consequências e efeitos da guerra a uma escala

global, com especial foco nas crises alimentar e de abastecimento de gás e petróleo, tendo como pano de fundo sanções e embargos à Rússia.



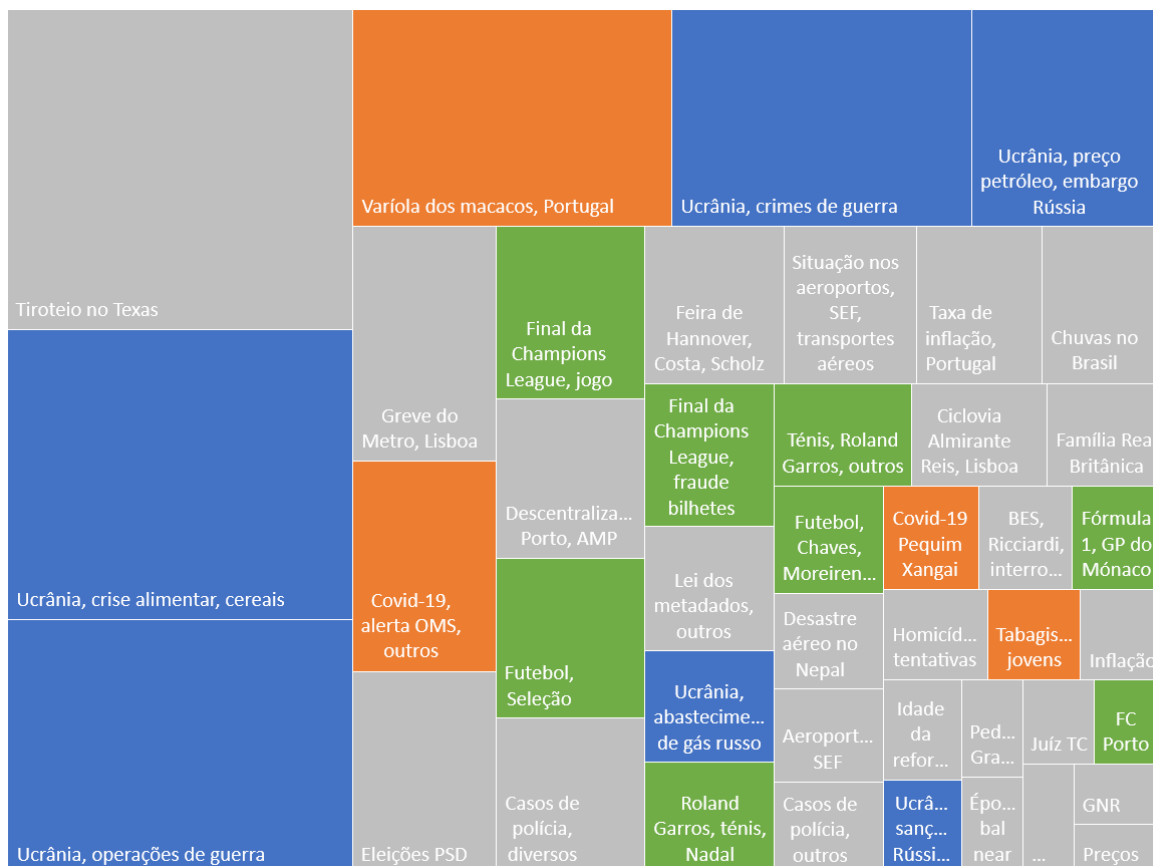
Clusters de 'histórias' noticiosas com proximidade semântica. Fonte: MediaLab/Priberam.

O sofrimento humano e as armas nos EUA

No dia 24 de maio, 19 crianças morreram num ataque com arma de assalto em Uvalde, Texas, EUA. O atentado aconteceu numa escola primária e morreram, também, duas professoras e o atacante armado, de 18 anos. Este acontecimento e o problema da venda de armas de assalto nos EUA provocaram, como seria de esperar, uma cobertura jornalística muito intensa, em todo o Mundo. Portugal não foi exceção e este acontecimento gerou, aproximadamente, 17 artigos por dia.

Os temas que restam, fragmentados

O mapa visual das 'histórias' noticiosas surge com muitos clusters tematicamente desassociados (quadros destacados em tom cinza, abaixo), resultado, em parte, da ausência de acontecimentos agregadores nos domínios da política nacional e da pandemia, *per se*, ao contrário do que tem sucedido em períodos anteriores.



Os 'clusters' relacionados com Ucrânia estão identificados com cor azul, desporto a verde e saúde, a laranja. Quadro construído com Índice de base = 100 em que a variável de interesse é o cluster com mais publicações: "Tiroteio no Texas". Fonte: MediaLab/Priberam (dados recolhidos em 17/6/2022)

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 43 *clusters* que mais se destacaram neste mês. 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 de *clusters* no relatório final é sempre igual ou inferior ao ‘top 50’ gerado pelo algoritmo².

Descrição da amostra

Para a redação deste relatório foram considerados 4719 artigos, publicados em 17 fontes de informação jornalística³, selecionados a partir de um total de, aproximadamente, 53256 artigos.

A recolha de dados é executada na plataforma Priberam e a organização das histórias noticiosas, em *clusters*, processa-se em dois passos: Um primeiro agrupamento é efetuado pelo algoritmo, que devolve um ‘TOP 50’ com os *clusters* mais publicados. Num segundo momento procede-se ao reagrupamento desse ‘TOP 50’, com recurso a técnicas de análise de conteúdo, que permitem uma observação ‘fina’ dos dados, resultando no quadro final de *clusters*⁴ apresentado neste relatório.

A RTP (notícias) superou o ‘Correio da Manhã’ e foi, em maio, fonte de maior número de artigos publicados *online*. Observador, Visão e CNN Portugal mantiveram as posições relativas (quadro abaixo, esquerda).

Do ponto de vista temático, a categoria “crime, lei e justiça” voltou a subir no ranking e foi, em maio, a terceira categoria com mais artigos publicados. “Guerra e conflitos armados” é, agora, o quarto tema no ranking (quadro abaixo, direita).

Órgão de comunicação social	Ranking
RTP	▲ 2
Correio da Manhã	▼ -1
Observador	■ 0
Visão	■ 0
CNN Portugal	■ 0

Fonte: MediaLab/Priberam (dados recolhidos em 17/6/2022)

Categoria temática (IPTC)	Ranking
Política	■ 0
Economia, negócios e finanças	■ 0
Crime, lei e justiça	▲ 1
Guerra e conflitos armados	▼ -1
Questões sociais	■ 0

Fonte: MediaLab/Priberam (dados recolhidos em 17/6/2022)

O ranking de personalidades e entidades mais referidas nos textos publicados mantém-se praticamente inalterado. Únicas exceções para a subida de ‘Joe Biden’ (quadro abaixo, esquerda), muito por força das notícias relacionadas com o tiroteio em Uvalde, e para a ‘Assembleia da República’ (quadro abaixo, direita), fonte de conteúdos noticiosos na esfera política que foi, aliás, a categoria temática com mais publicações, embora muito diferenciadas entre si.

Personalidades	Ranking
Vladimir Putin	■ 0
Volodymyr Zelensky	■ 0
António Costa	■ 0
Marcelo Rebelo de Sousa	■ 0
Joe Biden	▲ 3

Fonte: MediaLab/Priberam (dados recolhidos em 17/6/2022)

Entidades	Ranking
Governo	■ 0
Estado	■ 0
ONU	■ 0
Assembleia da República	▲ 5
Partido Socialista	■ 0

Fonte: MediaLab/Priberam (dados recolhidos em 17/6/2022)

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 CNN Portugal, Correio da Manhã, Diário de Notícias, Expresso, Jornal de Notícias, Jornal i, Jornal SOL, Multinews, Notícias ao Minuto, Observador, Público, Rádio Renascença, RTP, SIC Notícias, TSF, TVI-Notícias, Visão.

⁴ 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.