

## Relatório de abril de 2021

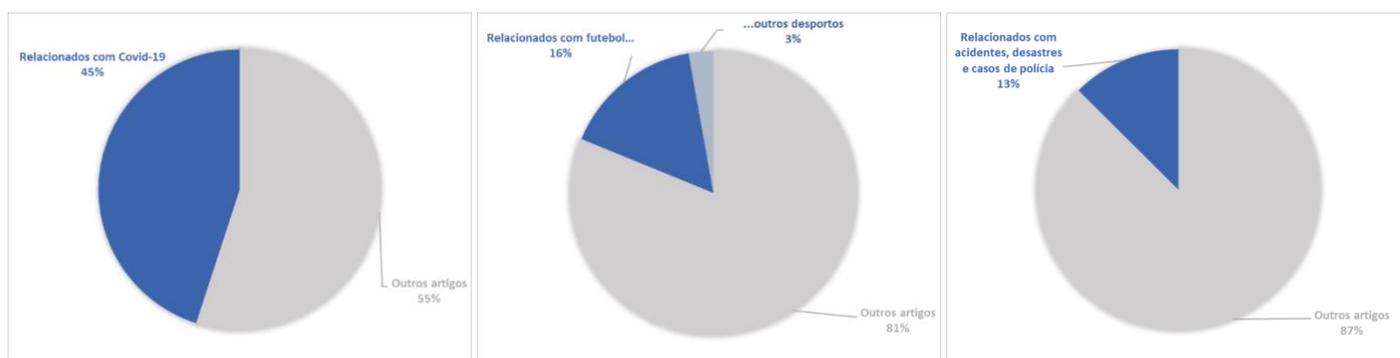
### Síntese

A **pandemia** por SARS-Cov-2, com especial incidência na divulgação de relatórios epidemiológicos, continuou no topo da agenda jornalística em abril, mantendo a tendência dos meses anteriores. Para além das notícias sobre Covid-19, o **futebol** foi alvo de uma atenção mediática relevante e o mês ficou, ainda, marcado por um grupo significativo de notícias relacionadas com **acidentes, desastres e casos de polícia** que coincidiram no tempo, marcando a cobertura noticiosa online.



Da esquerda: Cerimónia em memória de vítimas por Covid-19 em Copacabana, Brasil (ANTÓNIO LACERDA/LUSA), Descarrilamento de comboio em Taiwan (RITCHIE B. TONGO/LUSA), Manifestação contra a Super Liga Europeia antes de um jogo, no Reino Unido (FERNANDO VELUDO/LUSA)

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



### Pandemia por SARS-Cov-2: Estamos perante um decréscimo da atenção noticiosa?

A atenção dedicada à pandemia manteve-se dominante em abril<sup>1</sup>. Os dados disponíveis apontam para uma variação negativa na ordem dos -27% no que respeita o número de artigos publicados, o que pode indiciar um decréscimo da atenção noticiosa à medida que o número de casos diminui em Portugal, mas este número deve ser considerado com alguma reserva<sup>2</sup> e teremos que esperar pelos dados do próximo mês para perceber se se trata de uma tendência na cobertura mediática da pandemia.

A média diária de artigos publicados em abril sobre a pandemia (variação negativa de -25% para o mês de março) mantém a tendência do mês anterior (que registou uma variação de -5,9% face a fevereiro).

Em termos relativos, as notícias sobre Covid-19 representaram pouco menos de metade dos artigos considerados na amostra<sup>3</sup>.

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

<sup>2</sup> Os dados de abril podem ter sido influenciados pela ausência de publicações da SIC Notícias online, por motivos técnicos e pela eliminação de dois clusters na análise (320 artigos), por motivos metodológicos.

<sup>3</sup> No mês de abril foram selecionados, na plataforma Priberam, 13.441 artigos publicados em 15 empresas jornalísticas e noticiosas de âmbito nacional e generalista (imprensa não especializada). Poderá consultar mais detalhes sobre a metodologia no final do relatório.

### Futebol continua a preencher um espaço importante na cobertura generalista online

Os dados disponíveis permitem estimar que, em 20 artigos publicados no mês de abril em websites generalistas, 3 foram sobre futebol.

Outros desportos, incluindo a cobertura do Campeonato da Europa de Judo, que se realizou em Lisboa e na qual quatro atletas portugueses saíram medalhados, não ultrapassaram uma percentagem estimada de 3%, ou seja, em 30 artigos publicados pode estimar-se que um terá sido sobre outros desportos.

### Abril: Mês de acidentes, desastres e ‘casos de polícia’

Os dados revelam que o mês de abril foi também marcado por vários *clusters* de notícias relacionados com acidentes, desastres naturais, incêndios e vários tipos de ‘casos de polícia’, na sua maioria detenções por agressão ou assalto.

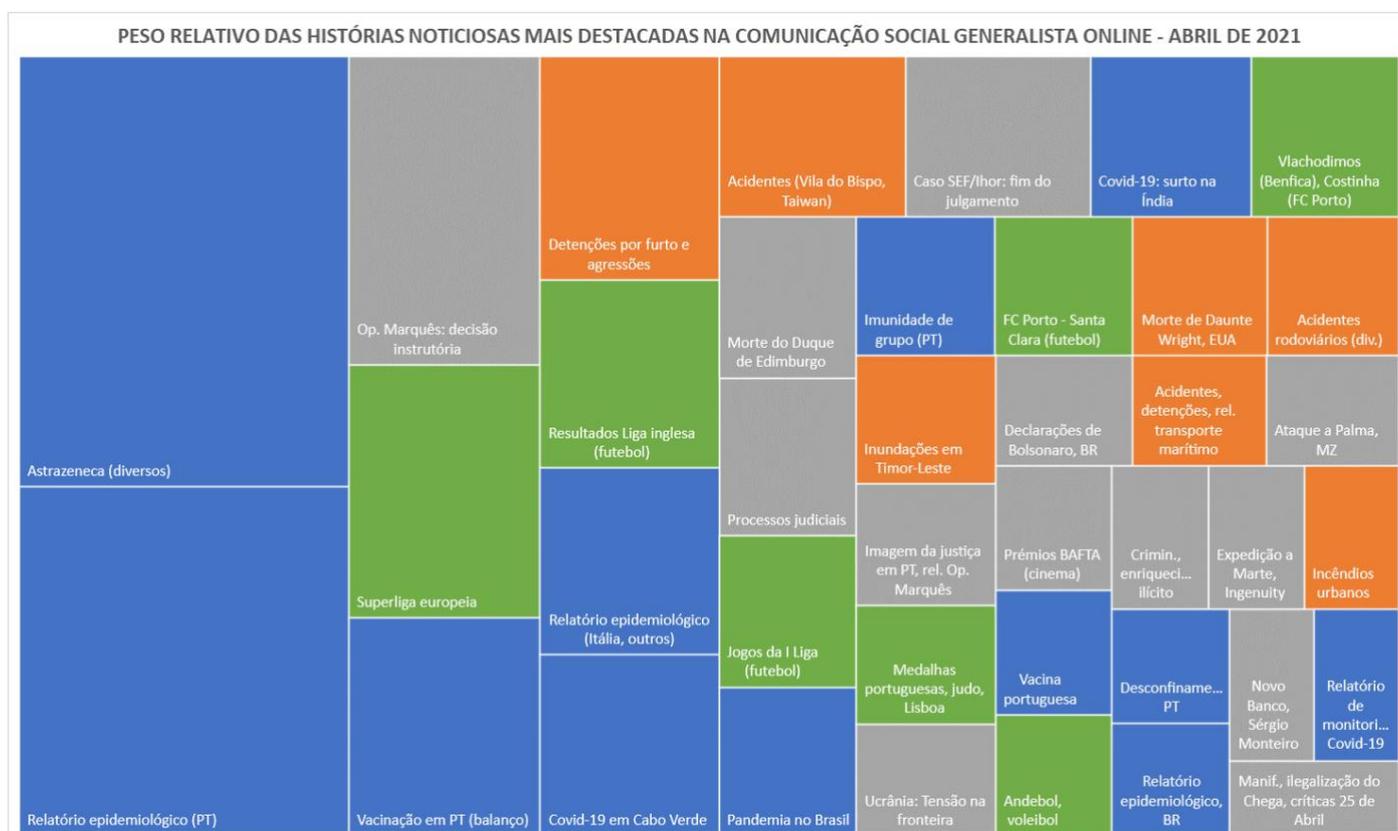
Esta situação resulta, essencialmente, da coincidência temporal de um conjunto de acontecimentos disruptivos com elevado ‘valor-notícia’ e estima-se que, em 30 artigos publicados online no mês de abril, 4 enquadraram-se nesta ‘categoria’.

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

Outros temas e acontecimentos obtiveram atenção jornalística considerável em abril de 2021. Desde já deve assinalar-se o ‘regresso’ do “Caso Sócrates” ou “Operação Marquês”. De acordo com os dados disponíveis, podemos estimar que cada um dos 15 órgãos de comunicação social generalistas considerados no barómetro publicou 50 artigos sobre a decisão instrutória deste processo judicial.

Outros processos judiciais estiveram presentes na agenda mediática, em abril (por exemplo, a conclusão do julgamento do caso “Ihor/SEF”) e a própria imagem da justiça em Portugal foi tema noticioso que surgiu na sequência da decisão instrutória no “Caso Sócrates/Op. Marquês”.

O quadro abaixo (índice de base = 100 em que a variável de interesse é o cluster mais publicado, “AstraZeneca-diversos”) permite visualizar a variação relativa de cada um dos clusters noticiosos que mais se destacaram.



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.

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

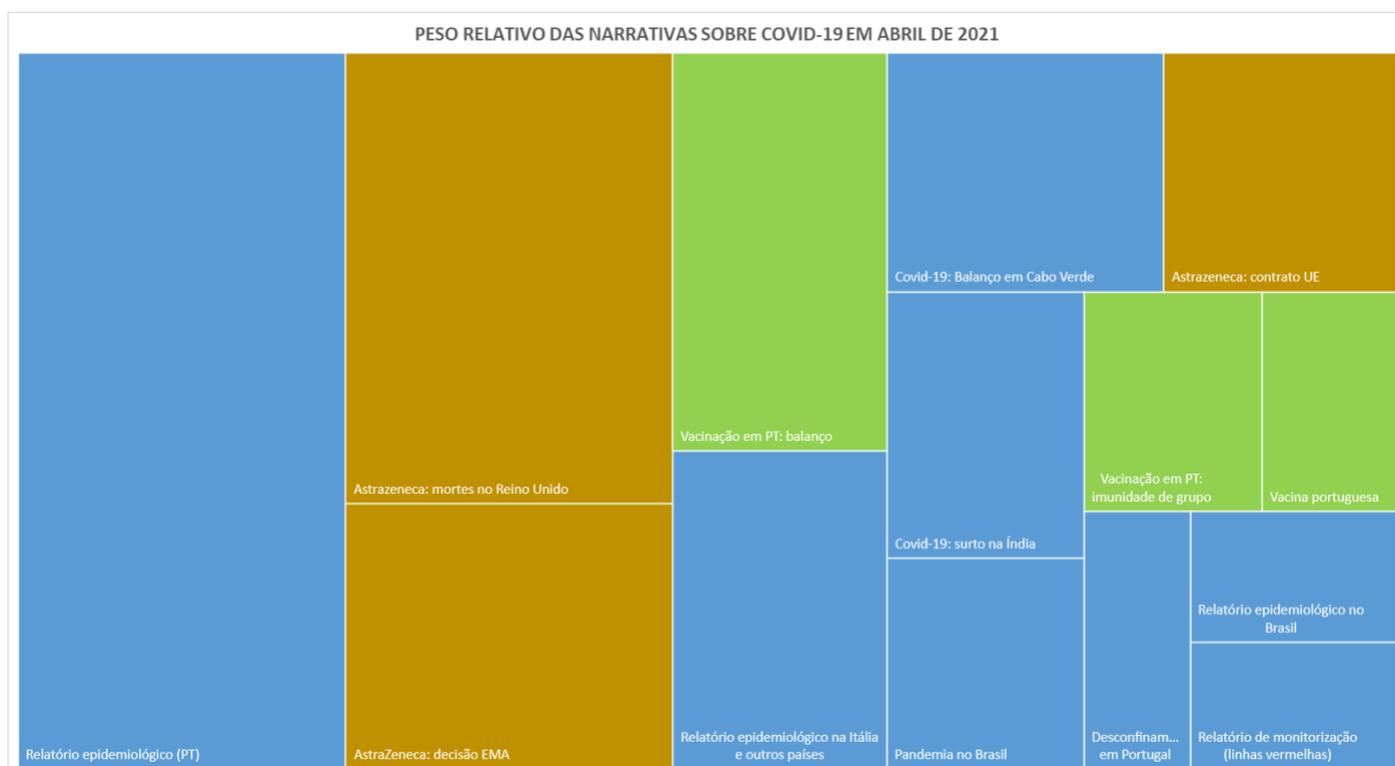
O quadro acima permite visualizar a proporção de cada agrupamento de notícias (*cluster*) no conjunto dos 39 *clusters* que mais se destacaram em abril de 2021. Estes, resultam de um trabalho de análise e reagrupamento semântico realizado num conjunto mais alargado, de 50 *clusters*<sup>4</sup>, identificados 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>5</sup>.

## A narrativa da pandemia nas notícias

A análise identificou, em março, 14 tópicos<sup>6</sup> relacionados com a pandemia, menos 3 do que no mês anterior.

O foco nos relatórios epidemiológicos foi dominante na cobertura mediática da pandemia nas notícias online. Não apenas sobre a situação em Portugal, mas também sobre países de língua portuguesa, em particular o Brasil.

A atenção à evolução epidemiológica em Portugal foi prevalente, com aproximadamente 1500 artigos publicados. Por outro lado, a Índia - que se tornou, em abril, o país do mundo com maior número de casos e de mortes diárias – obteve, em média, menos de um artigo em cada 25 publicações sobre a situação epidemiológica da Covid-19.



As notícias sobre Covid-19 subdividem-se em três grupos temáticos: a evolução da pandemia em Portugal e no Mundo (cor azul), Vacina AstraZeneca (cor dourada) e Processo de vacinação em Portugal (cor verde).

A vacina AstraZeneca recebeu atenção redobrada face ao mês anterior. Foram identificados três grandes *clusters* relacionados com a vacina: ‘mortes no Reino Unido’, ‘decisão da Agência Europeia do Medicamento (EMA)’, ‘contrato entre a farmacêutica e a União Europeia’.

<sup>4</sup> Para mais informação, consultar a secção de metodologia no final do relatório.

<sup>5</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>6</sup> Para mais informação consulte a metodologia no final do relatório.

Destaque, ainda, para um terceiro grupo de *clusters* associados à vacinação em Portugal, que podemos distinguir, semanticamente, em três domínios: ‘balanço do processo de vacinação’, ‘previsão de imunidade de grupo’, ‘desenvolvimento da vacina portuguesa’.

## Descrição da amostra

No mês de abril foram considerados 13.441 artigos, de 15 fontes de informação jornalística<sup>7</sup> para a realização deste relatório, selecionadas a partir de um universo de 166 entidades (aproximadamente 300.000 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 39 *clusters*<sup>8</sup>.

Do ponto de vista temático, a ‘**política**’ foi a área temática mais representada nas notícias online, em abril.

Áreas temáticas (IPTC)	Índice 100
Política	100
Economia, negócios e finanças	80
Problemas sociais	56
Saúde	56
Desporto	49

O Notícias ao Minuto foi a entidade com mais artigos publicados com, aproximadamente, 9.000 publicações no período em análise. Seguiram-se os websites da RTP, com menos 30% de publicações, e Correio da Manhã com menos 33%.

Fonte (órgão de comunicação social)	Índice 100
Notícias ao Minuto	100
RTP	70
Correio da Manhã	67
Visão	52
Observador	46

António Costa e Marcelo Rebelo de Sousa foram as personalidades mais referidas em abril e as entidades mais referidas foram o Governo e seus ministérios, com uma larga margem relativamente a outras entidades.

Personalidade	Índice 100
António Costa	100
José Sócrates	75
Marcelo Rebelo de Sousa	71

Entidade	Índice 100
Governo	100
Partido Socialista	30
Comissão Europeia	27

<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, TSF, TVI24, Visão. Por motivos técnicos não foi possível incluir artigos da SIC Notícias.

<sup>8</sup> Para mais informações sobre metodologia consultar secção no final do relatório.

## Metodologia do Barómetro de Notícias 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.