



UNIVERSIDADE DE BRASÍLIA (UnB)
FACULDADE DE CIÊNCIA DA INFORMAÇÃO (FCI)
PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO (PPGCINF)

LARISSA DE ARAÚJO ALVES

**VANTAGENS DE CITAÇÃO E ALTMÉTRICAS EM ARTIGOS E PREPRINTS NA
TEMÁTICA COVID-19**

Brasília

2022

LARISSA DE ARAÚJO ALVES

**VANTAGENS DE CITAÇÃO E ALTMÉTRICAS EM ARTIGOS E PREPRINTS NA
TEMÁTICA COVID-19**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação (PPGCINF), da Faculdade de Ciência da Informação (FCI), da Universidade de Brasília (UnB) como requisito para obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Gestão da Informação

Linha de pesquisa: Comunicação e Mediação da Informação

Orientador: Prof. Dr. João de Melo Maricato.

Brasília
2022



FOLHA DE APROVAÇÃO

Título: "Vantagens de citação e alométricas em priprints e artigos na temática COVID-19"

Autor (a): Larissa de Araújo Alves

Área de concentração: Gestão da Informação

Linha de pesquisa: Produção, Socialização e Usos da Informação e do Conhecimento

Dissertação submetida à Comissão Examinadora designada pelo Colegiado do Programa de Pós-graduação em Ciência da Informação da Faculdade em Ciência da Informação da Universidade de Brasília como requisito parcial para obtenção do título de MESTRE em Ciência da Informação.

Dissertação aprovada em: 09 de agosto de 2022.

Presidente (UnB/PPGCINF): Pro. Dr. João de Melo Maricato

Membro Externo (UFSCar): Profa. Dra. Ariadne Chloe Mary Furnival

Membro Interno (UnB/PPGCINF): Prof. Dr. Fernando César Lima Leite

Suplente (UnB/PPGCINF): Prof. Dr. Dalton Lopes Martins

Em 01/07/2022.



Documento assinado eletronicamente por **Joao de Melo Maricato, Professor(a) de Magistério Superior da Faculdade de Ciência da Informação**, em 11/08/2022, às 15:05, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Ariadne Chloe Mary Furnival, Usuário Externo**, em 12/08/2022, às 17:32, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Fernando César Lima Leite, Professor(a) de Magistério Superior da Faculdade de Ciência da Informação**, em 30/08/2022, às 10:09, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



A autenticidade deste documento pode ser conferida no site http://sei.unb.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **8352270** e o código CRC **5FC8B155**.

A474v Alves, Larissa de Araújo.

Vantagens de citação e altmétricas em artigos e preprints na temática COVID-19 / Larissa de Araújo Alves. -- 2022
158 p.: il. color. ; 30 cm.

Orientador: João de Melo Maricato.

Dissertação (Mestrado em Ciência da Informação) -
Universidade de Brasília, Faculdade de Ciência da Informação,
Brasília, 2022.

Inclui bibliografia.

1. Acesso Aberto. 2. Altmétrie. 3. Bibliometrie. 4. Preprint. I.
Maricato, João de Melo, orient. II. Título.

Dedico este trabalho àqueles que me deram a vida e as oportunidades para desfrutá-la em busca dos meus sonhos, meus amados e incentivadores pais.

AGRADECIMENTOS

No filme Harry Potter e o prisioneiro de Azkaban, Alvo Dumbledore expressa que “a felicidade pode ser encontrada mesmo nas horas mais sombrias, se a pessoa se lembrar de acender a luz”¹. Durante o desenvolvimento da pesquisa de mestrado, nem sempre foi possível acender a luz sozinha. Para trilhar o caminho até o interruptor e, enfim, encontrar a luz acesa situada na sala de defesa, contei com o apoio de diversos personagens neste capítulo de minha história, a quem agradeço:

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pela concessão da bolsa de estudos que viabilizou o desenvolvimento desta pesquisa;

À Universidade de Brasília (UnB), em especial, a todo corpo docente e de funcionários da Faculdade de Ciência da Informação (FCI) com que tive contato, pelo acolhimento e nobres ensinamentos, pelas oportunidades e pelos recursos fornecidos durante todo o mestrado e graduação;

Ao Prof. João de Melo Maricato, pela orientação atenciosa e dedicada, pelos desafios propostos, pela confiança em meu trabalho e pela pronta disposição em esclarecer minhas dúvidas e sanar incertezas;

Aos membros das bancas de qualificação e defesa, Prof^a Ariadne Chloe Mary Furnival e Prof. Fernando César Lima Leite pela disponibilidade e pelas valiosas colocações, contribuições e sugestões de aprimoramento da pesquisa. A participação de vocês foi essencial para o desenvolvimento e refinamento desta dissertação;

A minha orientadora de graduação, Prof^a Michelli Costa, por ter me apresentado o acesso aberto, o mundo da pesquisa em CI e pela torcida e incentivo para que eu ingressasse no Programa de Pós-Graduação em Ciência da Informação (PPGCinf);

Aos Doutores que, desde o início, acreditaram, incentivaram e acompanharam as mais variadas fases do processo de investigação, Tiago Braga e André Appel. O empolgante e assertivo estímulo de vocês me proporcionou energia para começar e continuar me dedicando a empreitada de buscar respostas, por mais de dois anos, para um único problema: o de pesquisa;

Ao Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), por ter sido, para mim, muito mais que a casa da informação científica e tecnológica no Brasil, mas também, por ter sido a minha segunda casa e escola logo no início de minha vida profissional e acadêmica. Dos frutos que o Ibict me trouxe, agradeço à todos os colegas da CGTI e COTEC, que se mostraram solícitos em ajudar e me ofereceram apoio e refúgio durante a caminhada de um mestrado pandêmico;

¹ *"Happiness can be found, even in the darkest of times, if one only remembers to turn on the light".*

Às amigas Denise Oliveira e Nathaly Leite, que me ofereceram escuta atenta em tantos momentos e dividiram comigo, simultaneamente, as dores e as alegrias da experiência de ser mestranda;

Com tamanha importância e amor, agradeço aos meus pais, Ednaldo e Helena, pelo apoio incondicional, pelo exemplo de força e perseverança e pela confiança primária e inigualável em meu potencial. Vocês são o motivo para que tudo fizesse sentido!

Por fim, com muito carinho, registro meu agradecimento aos meus queridos irmãos, Davi e Nathália, pela amizade, o companheirismo e a mais verdadeira torcida. Com vocês, não divido somente o sobrenome, mas a admiração e o amor. Muito obrigada por tudo e por tanto!

The speed of progress in science has always been strongly dependent on how efficiently scientists can communicate their results to their peers and to lay-persons willing to implement these results in new technologies and practices.

- Bo-Christer Björk, 2005

RESUMO

ALVES, Larissa de Araújo. *Vantagens de citação e altmétricas em artigos e preprints na temática COVID-19*. Orientador: Prof. Dr. João de Melo Maricato. Brasília, 2022. 158 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2022.

Questões associadas à comunicação científica também tendem a ser influenciadas pelas especificidades de períodos de Emergência em Saúde Pública de Importância Internacional (ESPII), como por exemplo, a pandemia de *corona virus disease* (COVID-19), que aumentou a necessidade por informações e acelerou a dinamicidade de produção científica. O volume de publicações científicas, diferenças entre áreas do conhecimento, uso de *preprints* para disseminação de descobertas e tempo necessário para evolução da ciência são pontos que chamam atenção no que se refere à produção científica de COVID-19. Dessa forma, a presente pesquisa busca identificar as dinâmicas de impacto bibliométrico de citação e altmétrico em artigos e *preprints* na temática COVID-19. O desenvolvimento de estudos por meio da aplicação de abordagens quantitativas bibliométricas e altmétricas fornece colaborações significativas para resolução de crises de saúde, uma vez que propiciam a visualização dos esforços empenhados para a produção de conhecimento voltado ao entendimento dos surtos causados por doenças. Quanto ao aspecto metodológico, a pesquisa caracteriza-se como quantitativa e lança mão de análises descritivas, comparativas e de correlação de Spearman de dados relacionados a amostra de 90.000 publicações, 83.477 artigos e 6.523 *preprints*, sobre COVID-19. Apresenta resultados sobre a observação da influência da variação entre tipos de acesso (aberto e fechado), subtipos de acesso (bronze, dourado, híbrido, verde e fechado) e tipos de publicações nos valores de indicadores bibliométricos e altmétricos calculados para os dados de citação, leitores/capturas no Mendeley e menções no Twitter. Os resultados expressam que artigos de acesso aberto demonstraram maior potencial de citação (82,3%) que artigos de acesso fechado (62,5%) e *preprints* (59,2%). No aspecto cobertura, o Mendeley demonstrou maior abrangência entre os artigos (valores de cobertura superiores a 88%), enquanto que o Twitter foi a fonte com a melhor cobertura para *preprints* (66%). Observa, no geral, vantagens de acesso aberto em comparação com o acesso fechado entre as publicações da amostra. Identifica, dentre os artigos da amostra, um destaque dos maiores valores de indicadores bibliométricos e altmétricos obtidos para os subtipos de acesso bronze e híbrido. Quanto às correlações calculadas, o estudo aponta que todas podem ser interpretadas como significativas, pois nenhum valor obtido para os coeficientes de correlação Spearman demonstram correlações analisadas como fracas ou perfeitas. Para os artigos, a melhor correlação identificada foi entre citações e leitores/capturas. Já entre os *preprints*, o melhor coeficiente de correlação se situa entre citações e menções no Twitter.

Palavras-chave: Acesso aberto. Altmetria. Bibliometria. Citação. Comunicação científica. Preprint. Vantagens de acesso aberto.

ABSTRACT

Issues associated with scientific communication also tend to be influenced by the specificities of periods of Public Health Emergencies of International Importance, such as the Corona Virus Disease Pandemic (COVID-19), which increased the need for information and accelerated the dynamics of scientific production. The volume of scientific publications, differences between areas of knowledge, the use of preprints for dissemination of discoveries, and the time needed for the evolution of science are points that draw attention concerning the scientific production of COVID-19 and that can be observed from various perspectives. Thus, the present research seeks to identify the dynamics of bibliometric citation impact and altmetrics in articles and preprints on the subject of COVID-19. The development of studies through the application of quantitative bibliometric and altmetric approaches provides significant collaborations for the resolution of health crises since they provide the visualization of the efforts made to produce knowledge aimed at understanding outbreaks caused by diseases. As for the methodological aspect, the research is characterized as quantitative and makes use of descriptive, comparative, and Spearman correlation analyses of data related to a sample of 90,000 publications, 83,477 articles, and 6,523 preprints on COVID-19. Presents results on the observation of the influence of variation between access types (open and closed), access subtypes (bronze, gold, hybrid, green and closed) and publication types on the values of bibliometric and altmetric indicators calculated for citation data, readers/captures in Mendeley and Twitter mentions. The results express that open access articles showed higher citation potential (82.3%) than closed access articles (62.5%) and preprints (59.2%). In the coverage aspect, Mendeley demonstrated greater coverage among articles (coverage values greater than 88%), while Twitter was the source with the best coverage for preprints (66%) In general, it observes advantages of open access compared to closed access among sample publications. Among the articles on the theme of COVID-19, we identify a highlight of the highest values obtained for the bronze and hybrid access subtypes. As for the calculated correlations, the study points out that all can be interpreted as significant, however, no value obtained for Spearman correlation coefficients demonstrates correlations interpreted as weak or perfect. For articles, the best correlation identified was between citations and readers/captions. For preprints, the best correlation coefficient is between citations and Twitter mentions.

Keywords: Open access. Altmetrics. Bibliometrics. Citation. Scientific communication. Preprint. Open access advantages.

LISTA DE FIGURAS

Figura 1 - Comparação entre as diferenças do modelo tradicional e o modelo atual de comunicação científica científica na área da saúde.

Figura 2 - Fluxo de surgimento do *preprint* no contexto de desenvolvimento da pesquisa até a publicação dos resultados em periódicos posteriormente disseminados.

Figura 3 - Mapeamento das influências dos *preprints* no contexto acadêmico.

Figura 4 - Quantitativo de periódicos científicos e artigos em acesso aberto registrados no DOAJ (março de 2022).

Figura 5 - Subtipos utilizados pela Scopus para classificação de acesso aberto em publicações indexadas.

Figura 6 - Subtipos utilizados pela Dimensions para classificação de acesso aberto em publicações indexadas.

Figura 7 - Caracterização da pesquisa.

Figura 8 - Interface do módulo Dimensions Analytics.

Figura 9 - Expressão de busca e filtros utilizados para pesquisar artigos e *preprints* na Dimensions.

Figura 10 - Etapas e procedimentos de coleta e tratamento de dados.

Figura 11 - Categorização das variáveis da pesquisa quanto à natureza e dependência.

Figura 12 - Interpretação do intervalo de valores possíveis para o coeficiente de correlação de Spearman.

Figura 13 - Fluxo de identificação do tipo de acesso de um periódico (completamente aberto/dourado ou fechado) pelo Unpaywall.

Figura 14 - Fluxo de identificação dos subtipos de acesso de um artigo (completamente aberto/dourado ou fechado) pelo Unpaywall.

LISTA DE GRÁFICOS

Gráfico 1 - Crescimento do número de submissões mensais registradas no arXiv entre 1992 e 2019.

Gráfico 2 - Crescimento do número de repositórios de acesso aberto registrados no OpenDOAR (dezembro de 2005 a março de 2022).

Gráfico 3 - Histograma de citações por tipo de publicação (2020-2021)

Gráfico 4 - Histograma de leitores/capturas por tipo de publicação (2020-2021).

Gráfico 5 - Histograma menções no twitter por tipo de publicação (2020-2021).

Gráfico 6- Dispersão dos dados de citação por tipo de publicação.

Gráfico 7- Dispersão dos dados de leitores/capturas por tipo de publicação.

Gráfico 8- Dispersão dos dados de menções no Twitter por tipo de publicação.

Gráfico 9 - Quantidade de publicações sobre COVID-19 recuperados na base Dimensions (2020-2021) por tipo de acesso (aberto e fechado)

Gráfico 10 - Quantitativo de artigos sobre COVID-19 recuperados na base Dimensions (2020-2021) distribuído por subtipo de acesso (bronze, dourado, híbrido, verde, fechado).

LISTA DE QUADROS

Quadro 1 - Diferenças entre *preprints* e artigos tradicionais.

Quadro 2 - Marcos relevantes para o MAA.

Quadro 3 - Principais aspectos das declarações de Budapest, Bethesda e Berlim.

Quadro 4 - Categorização de subtipos relacionados ao acesso aberto.

Quadro 5 - Benefícios do acesso aberto.

Quadro 6 - Fontes de dados.

Quadro 7 - Seleção de dados de identificação e citação das publicações coletados a partir da Dimensions.

Quadro 8 - Detalhamento do quantitativo de registros, por tipo de publicação, componentes da amostra da pesquisa.

Quadro 9 - Medidas de tendência central (variáveis quantitativas por tipo de publicação).

Quadro 10 - Medidas de forma (variáveis quantitativas por tipo de publicação).

Quadro 11 - Escala de interpretação para os valores do coeficiente de correlação de Spearman (r_s).

Quadro 12 - Síntese dos procedimentos de análise conforme objetivos específicos da pesquisa.

Quadro 13 - Quantitativo de artigos e *preprints* sobre COVID-19 recuperados na base Dimensions (2020-2021).

Quadro 14 - Análise dos metadados do *dataset* gerado pela Dimensions na Etapa 1 da coleta de dados.

LISTA DE TABELAS

Tabela 1 - Indicadores de citação de artigos e *preprints* por tipo de acesso (aberto e fechado).

Tabela 2 - Indicadores de leitores/capturas de artigos e *preprints* por tipo de acesso (aberto e fechado).

Tabela 3 - Indicadores de menções no Twitter de artigos e *preprints* por tipo de acesso (aberto e fechado).

Tabela 4 - Indicadores bibliométricos de citação e altmétricos (leitores/capturas e menções no Twitter) de artigos por subtipo de acesso (bronze, dourado, híbrido, verde, fechado).

Tabela 5 -Matriz de correlação das variáveis citações, leitores/captura e menções no Twitter para artigos.

Tabela 6 -Matriz de correlação das variáveis citações, leitores/captura e menções no Twitter para *preprints*.

LISTA DE ABREVIATURAS E SIGLAS

ABC - Academia Brasileira de Ciências
ABI - Associação Brasileira de Imprensa
APC - Article Processing Charge
API - Application Programming Interface
ARL - Association of Research Libraries
BOAI - Budapest Open Access Initiative
CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CI - Ciência da Informação
CONFOA - Conferência Luso-Brasileiro sobre Acesso Aberto
COVID-19 - Corona Virus Disease 2019
DeCS - Descritores em Ciências da Saúde
DOAJ - Directory of Open Access Journals
DOI - Digital Object Identifier
DORA - Declaração de São Francisco sobre Avaliação da Pesquisa
ESPII - Emergência em Saúde Pública de Importância Internacional
FCR - Field Citation Ratio
H1N1 - Influenza A
IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia
JASIST - Journal of the American Society for Information Science and Technology;
JMIR - Journal of Medical Internet Research
MAA - Movimento de Acesso Aberto
MIT - Massachusetts Institute of Technology
OA - Open Access
OAAA - Open Access Altmetric Advantage
OACA - Open Access Citation Advantage
OAI - Open Archives Initiative
OAI-PMH - Open Archives Initiative Protocol for Harvesting
OECD - Organization for Economic Cooperation and Development
OJS - Open Journal System
OMS - Organização Mundial da Saúde
OpenDOAR - Directory of Open Access Repositories
PHEIC - Public Health Emergency of International Concern
PLoS - Public Library of Science
PMC - Lançamento da PubMed Central
PMCID - Número de referência do PubMed Central
PMID - PubMed Indexing Number
RCR - Relative Citation Ratio
RoMEO - Rights METadata for Open archiving
rs - Coeficiente de correlação de Spearman
SARS - Severe Acute Respiratory Syndrome
SBPC - Sociedade Brasileira para o Progresso da Ciência
SEER - Sistema Eletrônico de Editoração de Revistas
SPARC - Scholarly Publishing and Academic Resources Coalition
TICs - Tecnologias de Informação e Comunicação
TRLN - Triangle Research Libraries Network
UNESCO - United Nations Educational, Scientific and Cultural Organization
WoS - Web of Science

SUMÁRIO

1 CONSIDERAÇÕES INICIAIS	18
1.1 HIPÓTESES DE PESQUISA	23
1.2 OBJETIVOS	23
1.2.1 Objetivo geral	24
1.2.2 Objetivos específicos	24
1.3 JUSTIFICATIVA	24
2 REFERENCIAL TEÓRICO	28
2.1 A COMUNICAÇÃO E PRODUÇÃO CIENTÍFICA	29
2.1.1 Comunicação e produção científica em tempos de COVID-19	32
2.1.2 Preprint	35
2.2 ACESSO ABERTO	43
2.2.1 Os subtipos de acesso aberto	51
2.2.2 Argumentos a favor do acesso aberto	62
2.3 INDICADORES PARA MENSURAR AS VANTAGENS DO ACESSO ABERTO	65
2.3.1 Indicadores bibliométricos de citação	66
<i>2.3.1.1 Estudos sobre vantagens de citação de acesso aberto</i>	72
2.3.2 Indicadores altmétricos e vantagens altmétricas de acesso aberto	75
2.3.3 Correlação entre indicadores de citação e altmétricos	78
3 METODOLOGIA	80
3.1 CARACTERIZAÇÃO DA PESQUISA	80
3.2 FONTES DE DADOS	81
3.3 COLETA E TRATAMENTO DE DADOS	83
3.4 ANÁLISE DE DADOS	91

3.4.1 Definição e caracterização das variáveis da pesquisa	91
3.4.2 Procedimentos de análise	93
3.5 LIMITAÇÕES DA PESQUISA	103
4 RESULTADOS E DISCUSSÕES	104
4.1 DESCRIÇÃO GERAL DOS DADOS	104
4.2 RESULTADOS ORIENTADOS AOS OBJETIVOS ESPECÍFICOS	105
4.2.1 OE1 Comparar as variações entre o acesso aberto e os impactos bibliométrico de citação e altmétricos (menções e leitores/capturas) em artigos e preprints sobre COVID-19	105
4.2.2 OE2 Identificar a relação entre os indicadores de citações, menções e leitores/capturas e os subtipos de acesso (verde, bronze, dourado, híbrido e fechado)	114
4.2.3 OE3 Mensurar correlações entre indicadores bibliométricos de citação e altmétricos em artigos e preprints	125
5 CONCLUSÕES E CONSIDERAÇÕES FINAIS	128
APÊNDICE A	148

1 CONSIDERAÇÕES INICIAIS

De forma cíclica e contínua a ciência evolui e produz conhecimentos objetivos e racionais por meio da aplicação de métodos validados. O processo de desenvolvimento da ciência envolve etapas de observação, análise, discussão, consulta, produção, comunicação e avaliação. O fazer científico demanda tempo e caracteriza-se como um esforço corporativo e coletivo em busca de novos conhecimentos (MUELLER; PASSOS, 2000).

Como um fenômeno global capaz de impulsionar e direcionar o desenvolvimento do conhecimento e da evolução humana, a ciência pode ser compreendida a partir de uma multiplicidade de perspectivas. Não há resposta universal e objetiva para a pergunta “o que é ciência?”. A definição de ciência varia conforme o ponto de vista adotado, a disciplina considerada e o aspecto selecionado (social, instrumental, político, econômico, histórico, filosófico e documental). Conforme esclarece Ziman (1984), por exemplo, quando a ciência é tratada como um meio para resolução de problemas, o aspecto evidenciado é o instrumental. Por outro lado, quando vista como conhecimento organizado, têm-se ênfase nas características arquivísticas, documentais e históricas da definição.

Mesmo que sejam vastas as possibilidades de interpretação sobre tal fenômeno, é legítimo considerar que “a ciência precisa de tempo para pensar”. A ciência precisa de tempo para ler e para falhar. A ciência nem sempre sabe o que pode ser agora. A ciência se desenvolve de forma instável, com movimentos bruscos e saltos imprevisíveis à frente [...]” ²([SLOW SCIENCE ACADEMY](#), 2010, *online*, tradução nossa). Por meio da ciência, problemas são investigados, hipóteses são testadas e adversidades são superadas. Dentre essas estão as Emergências em Saúde Pública, eventos que colocam em risco a vida e o bem-estar público e demandam por respostas efetivas e coordenadas ([WORLD HEALTH ORGANIZATION](#), 2019).

Declarada pela Organização Mundial da Saúde (OMS) como uma *Public Health Emergency of International Concern* (PHEIC)³ (em português, Emergência em Saúde Pública de Importância Internacional - ESPII), a *Corona Virus Disease 2019* (COVID-19) representa um dos maiores desafios já enfrentados pela humanidade nas mais diversas esferas (científica,

² “Science needs time to think. Science needs time to read, and time to fail. Science does not always know what it might be at right now. Science develops unsteadily, with jerky moves and unpredictable leaps forward” (SLOW SCIENCE ACADEMY, 2010, *online*)

³ “An extraordinary event which is determined to constitute a public health risk to other States through the international spread of disease and to potentially require a coordinated international response”. Disponível em: <https://www.who.int/ihr/procedures/pheic/en/>. Acesso em: 15 jul. 2021.

social, política, cultural, ética, tecnológica, jurídica e econômica). No contexto científico, o cenário mundial decorrente da divulgação dos primeiros casos de infecção de humanos pela cepa (2019-nCoV) do novo coronavírus (Sars-CoV-2) no final de 2019, impôs aos pesquisadores dificuldades para acessar dados e informações e produzir conhecimentos em pouco tempo, demandando adaptações no processo de produção, disseminação e validação das descobertas científicas (JORGE; ALBAGLI; SOUZA, 2020).

Em decorrência do alastramento geográfico de transmissão da doença, em 11 de março de 2020, a COVID-19 foi caracterizada pela OMS como uma pandemia. Na data da declaração do estado pandêmico, havia notificação de casos da doença em 114 países. Durante o pronunciamento sobre a declaração da pandemia de COVID-19, Tedros Adhanom Ghebreyesus, diretor-geral da OMS a época, evidenciou aos países alguns pontos chave para a prevenção das infecções pelo vírus e para a minimização de danos: “primeiro, preparem-se e estejam prontos. Segundo, detectem, protejam e tratem. Terceiro, reduzam a transmissão. Quarto, inovem e aprendam” ([ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE, 2020, online](#)).

A necessidade de inovar e aprender para salvar vidas gerou uma corrida científica em busca de métodos de prevenção, tratamento e desenvolvimento de vacinas ([TORRES-SALINAS; ROBINSON-GARCIA; CASTILLO-VALDIVIESO, 2020](#)). Os riscos à saúde da população mundial impostos pelo alastramento do novo coronavírus alteraram a ideia de tempo para a ciência. Cada dia passado representava o aumento de casos confirmados de contaminação e a elevação dos óbitos registrados em decorrência da doença. Dessa forma, a pandemia catalisou os processos formais de produção do conhecimento científico e demandou um contínuo ajuste de métodos para a geração de respostas ao surto de COVID-19 de forma mais eficiente e efetiva. Tais adaptações provocaram mudanças nos sistemas científicos e tecnológicos ([ZHANG *et al.*, 2020](#)).

Entre os sistemas afetados como consequência do estágio emergencial, menciona-se o de comunicação científica, que se caracteriza pela integração de processos vinculados à elaboração, avaliação, disseminação, circulação, uso e preservação da informação científica pelos participantes da comunidade acadêmica ([ASSOCIATION OF COLLEGE & RESEARCH LIBRARIES, 2003, online](#)). Na perspectiva de Kupferschmidt (2020), o surto de COVID-19 rompeu com os moldes convencionais e vigentes de comunicação entre os cientistas, impulsionando o aumento da colaboração no compartilhamento de informações e evidenciando

problemas e fragilidades do sistema tradicional, como o tempo necessário para efetuar os processos de avaliação e publicação dos resultados de pesquisa.

Bermúdez-Rodríguez *et al.* (2020) adicionam à reflexão sobre as mudanças na comunicação científica decorrentes da pandemia a questão do acesso aberto. De acordo com esses autores, no contexto pandêmico, tornou-se mais latente na comunidade científica o apelo direcionado às editoras e às agências de financiamento a pesquisa “para implementação de políticas de acesso aberto (*open access*), visando incentivar a disponibilização em modo *online* e sem restrições dos resultados de investigação científica, especialmente aqueles relacionados ao novo coronavírus” (BERMÚDEZ-RODRÍGUEZ *et al.*, 2020, p. 2, grifo dos autores).

O acesso aberto às publicações científicas é estudado mundialmente por meio do Movimento de Acesso Aberto (MAA), o qual é visto por Mueller (2006) como um marco importante a ser considerado no que se refere às mudanças ocorridas no sistema tradicional de comunicação científica. Os desdobramentos e as investigações voltadas ao entendimento das possíveis aplicações e benefícios do acesso aberto aos processos de comunicação da ciência fizeram eclodir, a partir da formalização do MAA pela [Budapest Open Access Initiative \(BOAI\)](#) em 2002, discussões na literatura voltadas às mudanças nos modelos de negócio adotados por editoras.

Outro desdobramento do MAA foi o surgimento da hipótese caracterizada na literatura como *Open Access Citation Advantage* (OACA) ou vantagens de citação do acesso aberto, a qual considera que as publicações disponibilizadas sem barreiras financeiras ou de acesso possuem maiores possibilidades de serem citadas e gerarem impacto quando comparadas com as publicações em acesso fechado ou restrito, que se caracterizam pelas barreiras de acesso e de pagamento exigidas para que haja a liberação do conteúdo integral dos documentos publicados ([PIWOWAR *et al.*, 2018](#)). Para a OACA, o tipo de acesso é uma variável relevante e que pode promover influências no que tange às citações recebidas pelos documentos publicados.

A observação e o estudo da hipótese de OACA têm sido realizado, nos mais variados contextos temáticos e disciplinares, por meio da aplicação de metodologias inseridas no âmbito dos estudos métricos da informação, que visam analisar, utilizando uma abordagem quantitativa, as questões de produtividade, colaboração, coocorrência, citação e interações, que estão vinculadas à comunicação científica. Uma das frentes tratadas pelos estudos métricos é a Bibliometria, que diz respeito aos métodos matemáticos e estatísticos que buscam avaliar

aspectos quantitativos da produção científica objetivando a compreensão sobre o uso e a geração de documentos (NORONHA; MARICATO, 2008).

A utilização da abordagem bibliométrica para observação da produção científica tem corroborado também para o desenvolvimento de pesquisas com o foco nas publicações sobre a temática da COVID-19. Estudos como os de [Bermúdez-Rodríguez et al. \(2020\)](#), [Nasab e Rahim \(2020\)](#), [Ortiz Núñez e Stable Rodríguez \(2021\)](#), [Torres Pascual e Torrell-Vallespin \(2020\)](#) e [Zhai, et al. \(2020\)](#) recorrem às análises bibliométricas para investigarem questões de produção, disseminação, colaboração e impacto do conhecimento científico gerado sobre a doença causada pelo novo coronavírus. Cada uma das pesquisas, com seus recortes de cunho temporal, geográfico, disciplinar, temático e de fontes de informação contribuem para a visualização dos desdobramentos provocados pela pandemia na comunicação científica. Observa-se que as pesquisas bibliométricas realizadas até o momento com o foco na temática apresentada buscam explorar e compreender aspectos básicos a respeito do desenvolvimento do conhecimento sobre COVID-19: quem produziu; de que modo produziu; onde foi produzido e disseminado; como se deu a questão da colaboração e coautoria; quais as principais publicações e áreas de pesquisa; entre outros.

Conforme destaca [Boetto et al. \(2021\)](#), em decorrência das possibilidades fornecidas pelas tecnologias de informação e comunicação atuais, é a primeira vez na história humana que o enfrentamento de uma pandemia pode ser descrito, debatido e investigado de forma simultânea ao desenvolvimento das pesquisas por meio do uso de ambientes e ferramentas *online*. Além dos canais convencionais de comunicação do conhecimento científico, os pesquisadores têm recorrido aos canais informais *online*, a exemplo das mídias sociais (Twitter, Youtube) e das redes sociais científicas, para disseminar resultados de investigações na temática da COVID-19.

Nessa conjuntura, abre-se espaço para o desenvolvimento de estudos alométricos sobre o assunto em questão. Tal tipologia de pesquisa tem por escopo a mensuração da disseminação e uso da literatura científica no contexto *online* a partir de indicadores calculados por meio de dados de atividades registradas em redes sociais, Wikipédia, blogs, gerenciadores de referência, como por exemplo: compartilhamentos, salvamentos, menções, curtidas, comentários e outros ([PRIEM et al., 2010](#); [SOUZA, 2014](#)). [Bermúdez-Rodríguez et al. \(2020\)](#), [Torres-Salinas, Robinson-Garcia e Castillo-Valdivieso \(2020\)](#) e [Boetto et al. \(2021\)](#) aproximam seus trabalhos das análises alométricas para explorar questões vinculadas à produção científica de COVID-19.

O uso de indicadores alométricos também pode favorecer o desenvolvimento de pesquisas voltadas para verificação da hipótese de *Open Access Altmetric Advantage* (OAAA), ou vantagens alométricas de acesso aberto, que surgiu na literatura em decorrência de trabalhos já existentes sobre OACA e da ampliação da Almetria enquanto área de estudo. Uma das pesquisas que tratam da questão da OAAA é o artigo de [Holmberg et al. \(2019\)](#).

Além das possibilidades isoladas de estudos realizados a partir do uso de indicadores bibliométricos ou alométricos, existem abordagens que consideram os últimos como complementares às análises viabilizadas por meio da Bibliometria ([COSTAS; ZAHEDI; WOUTERS, 2014](#); [WEI, 2020](#)). É nessa conjuntura que emerge a investigação associada às possíveis correlações entre indicadores alométricos, gerados a partir da presença de informações científicas em mídias e redes sociais e indicadores bibliométricos de citação, que mensuram as interações tradicionais quantificáveis com as publicações científicas em canais formais de comunicação científica.

As análises quanto à existência de relações e correlações entre os indicadores alométricos e as citações, bem como o grau de intensidade dessas, são realizadas a partir da aplicação de métodos quantitativos e estatísticos em diferentes áreas do conhecimento. [Wei \(2020\)](#) é um exemplo de estudo que buscou compreender em que medida indicadores de citação e alométricos permitem a identificação de impacto em periódicos científicos de acesso aberto utilizando o Coeficiente de Correlação de Pearson para analisar as correlações entre duas variáveis. Esse também é adotado por [Araújo, Caran e Souza \(2016\)](#) para investigar relações entre dados de citação e alométricos associados às publicações do periódico DataGramZero. Dessa forma, mostra-se possível aproximação entre essas abordagens para investigação das publicações científicas sobre COVID-19, tanto na perspectiva bibliométrica quanto alométrica.

Na concepção de [Razumova \(2020\)](#), a pandemia de COVID-19 construiu um cenário favorável ao avanço da ciência, da comunicação científica e do acesso aberto. Como a mais grave PHEIC do século XXI, a doença em questão é um fenômeno recente que provoca discussões, reflexões e impulsiona mudanças na busca pelas melhores soluções. No contexto descrito, um ponto latente é o apelo da comunidade acadêmica pela aproximação das práticas científicas aos preceitos defendidos pelo Movimento pela Ciência Aberta, como o acesso, os dados e a avaliação aberta (JORGE; ALBAGLI; SOUZA, 2020).

“Preprints são um passo em direção ao futuro da Ciência Aberta”, coloca [Pulverer \(2016, p. 2617, tradução nossa\)](#). Caracterizados como uma possibilidade para atenuar o período decorrido entre a descoberta científica e sua disseminação formal, os *preprints* receberam

destaque no contexto de combate a pandemia de COVID-19, uma vez que foram um canal de comunicação bastante utilizado pelos pesquisadores com intuito de registrar e divulgar os resultados dos estudos desenvolvidos sobre a temática em voga.

Por meio dos *preprints*, as pesquisas em andamento puderam ser documentadas e acessadas, gerando um maior alcance das informações aos interessados, fato que dialoga com o ponto apresentado por [Pulverer \(2016\)](#) quando expõe que, ao serem arquivados, datados e citados ou mencionados, os *preprints* geram registros formais sobre a evolução de investigações científicas que ficam acessíveis de forma aberta. Em algumas áreas do conhecimento, a exemplo da Física, a disponibilização dos *preprints* para que possam receber *feedbacks* e avaliações prévias, em repositórios de manuscritos, constitui uma prática de comunicação científica que busca o aprimoramento das publicações que serão submetidas aos periódicos científicos ([PULVERER, 2016](#)).

Em face do exposto, o presente trabalho busca corroborar para a ampliação das investigações sobre as aproximações entre o tema COVID-19, o acesso aberto e os estudos métricos da informação. Nesse sentido, com o intuito de nortear o desenvolvimento da pesquisa, levanta-se o seguinte questionamento: *Qual a correlação entre as dinâmicas de impacto bibliométrico de citação e alométrico em artigos e preprints na temática COVID-19?*

1.1 HIPÓTESES DE PESQUISA

Diante da contextualização delineada nas considerações iniciais do presente trabalho e tendo em vista as reflexões existentes na literatura a respeito da existência de possíveis vantagens de acesso aberto (OACA e OAAA) em relação ao acesso fechado, esta pesquisa levanta duas hipóteses:

- H1)** Artigos de acesso aberto tendem a apresentar maiores valores para indicadores bibliométricos de citação e alométricos quando comparados aos artigos de acesso fechado;
- H2)** Os *preprints* na temática da COVID-19 registram maiores impactos de citações, menções e leitores/captura em relação aos artigos de acesso aberto e fechado.

1.2 OBJETIVOS

A seguir, encontram-se elencados os objetivos definidos para orientar o desenvolvimento desta pesquisa.

1.2.1 Objetivo geral

Identificar as dinâmicas de impacto bibliométrico de citação e altmétrico em artigos e *preprints* de acesso aberto na temática COVID-19.

1.2.2 Objetivos específicos

OE1 Comparar as variações entre o acesso aberto e os impactos bibliométrico de citação e altmétricos (menções e leitores/capturas) em artigos e *preprints* sobre COVID-19;

OE2 Identificar a relação entre os indicadores de citações, menções e leitores/capturas e os subtipos de acesso (verde, bronze, dourado, híbrido e fechado);

OE3 Mensurar correlações entre indicadores bibliométricos de citação e altmétricos em artigos e *preprints*;

1.3 JUSTIFICATIVA

O estudo proposto busca investigar aspectos inseridos ou relacionados ao arcabouço da comunicação científica, área relevante para o desenvolvimento da Ciência da Informação (CI) que, consoante Borko (1968, p. 3, tradução nossa), preocupa-se “com o corpo de conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação, e utilização da informação”. O autor elenca também, alguns fatores que contribuem para justificar a necessidade da CI enquanto disciplina capaz de melhorar os processos vinculados à informação, dentre esses, se relacionam ao interesse desta pesquisa: a) a existência de um volume grande de cientistas, de periódicos científicos e publicações técnicas; b) o aumento do grau de especialização que interfere na comunicação e intercâmbio de informações entre as disciplinas; e, c) a pressão por informações imediatas em virtude do tempo decorrido entre a realização da pesquisa básica e sua aplicação. Mesmo que tais fatores tenham sido enumerados em 1968, eles ainda são atuais e demonstram aproximações com questões de interesse da comunicação científica.

Volume de publicações científicas, diferenças entre áreas do conhecimento e tempo necessário para evolução da ciência são pontos que chamam a atenção no que se refere à

produção científica de COVID-19 e que podem ser observados sob os mais diversos prismas. Isto posto, busca-se aproximar a discussão sobre o fenômeno mundial que tem afetado a sociedade nas mais variadas áreas da perspectiva informacional vigente na CI.

A realização desta pesquisa também visa contribuir com a reflexão e a discussão sobre os aspectos associados às hipóteses de OACA e OAAA no contexto nacional. De acordo com [Cintra \(2017, p. 20\)](#), existe uma carência de aprofundamento de estudos sobre OACA no campo das Ciências Sociais Aplicadas, em especial, na CI. Assim, abordagem sobre as vantagens de acesso aberto almejada pelo presente trabalho busca incrementar as pesquisas desenvolvidas no âmbito do MAA, acrescentando novas interpretações e análises vinculadas a esse fenômeno significativo para a comunicação científica.

Ao buscarem compreender e testar a hipótese de OACA, as investigações realizadas tendem a estabelecer comparações entre os artigos em acesso aberto de maneira geral (desconsiderando a gama de variações de subtipos e possibilidades de abertura) com os documentos publicados de forma fechada. Quando consideradas as variações de subtipos de acesso aberto, observa-se nos estudos o foco nos subtipos tradicionais, como o verde e o dourado ([LANGHAM-PUTROW; BAKKER; RIEGELMAN, 2021](#)).

Tendo em vista a evolução do MAA e os múltiplos modelos de negócios e ramificações oriundos da adaptação das primeiras estratégias propostas para o alcance de acesso aberto (abordadas na seção 2.2.1 deste trabalho), aliada às possibilidades atuais de coleta de dados relativos à classificação dos subtipos de acesso aberto dourado, verde, bronze e híbrido em publicações por meio de fontes de informação como, por exemplo, [Unpaywall](#)⁴, [Scopus](#)⁵ e [Dimensions](#), faz-se válido considerar o estudo das hipóteses de OACA e OAAA de modo mais específico. Ou seja, a partir das comparações estabelecidas com os indicadores bibliométricos e altmétricos, pode-se verificar quais as variações observáveis entre os subtipos de acesso aberto supramencionados, ampliando a lógica de classificação binária (aberto vs. fechado) dos tipos acesso dos documentos.

Na condução de revisão sistemática de literatura sobre a temática de OACA, a qual compreende o recorte temporal de 2001 a 2020, [Langham-Putrow, Bakker e Riegelman \(2021\)](#) observaram, dentre a amostra de 134 documentos verificados, a elevada preponderância dos artigos como tipo de publicação mais recorrente (107/134), seguidos pelas apresentações em eventos (15/134), *preprints* (8/134) e teses e dissertações (4/134). O trabalho de [Langham-](#)

⁴ Disponível em: <https://unpaywall.org/>. Acesso em: 12 ago. 2021.

⁵ Disponível em: <https://www.scopus.com/>. Acesso em: 12 ago. 2021.

[Putrow, Bakker e Riegelman \(2021\)](#) também pontua que fatores, além do tipo de acesso, podem influenciar a questão da confirmação ou negação de OACA, a exemplo das bases de dados selecionadas, o tipo de publicação analisado e as variações disciplinares consideradas. Diante de tal colocação e tendo em mente o destaque recebido pelos *preprints* enquanto canais de comunicação científica relevantes para a disseminação otimizada das descobertas associadas às pesquisas sobre COVID-19 durante o período de enfrentamento da pandemia, o presente estudo considera como significativa a observação da variável tipo de publicação, dando foco aos artigos e *preprints*, com o intuito de comparar os possíveis desdobramentos desse aspecto nos indicadores altmétricos e de citação.

Na concepção de Zhang *et al.* (2020), o desenvolvimento de estudos por meio da aplicação de abordagens bibliométricas revelam colaborações significativas para resolução de crises de saúde, uma vez que propiciam a visualização dos esforços empenhados no entendimento dos surtos causados por doenças⁶. Dada a alta demanda por informações e a dinamicidade de produção científica vinculada ao contexto oriundo dos desdobramentos causados pela COVID-19, os primeiros estudos bibliométricos sobre essa doença foram desenvolvidos logo após a declaração do estado de ESPII. Portanto, essas pesquisas iniciais consideraram recortes temporais mais limitados, com menos de um ano de cobertura (BERMÚDEZ-RODRÍGUEZ *et al.*, 2020; NASAB; RAHIM, 2020; ORTIZ NÚÑEZ; STABLE RODRÍGUEZ, 2021; TORRES PASCUAL; TORRELL-VALLESPIN, 2020; ZHAI, *et al.*, 2020). De modo a contribuir com as investigações já desenvolvidas, este trabalho propõe ampliar o recorte temporal da pesquisa bibliométrica sobre COVID-19 a fim de observar novos desdobramentos e possibilidades eclodidas após o período inicial de disseminação da doença.

Na perspectiva de [Boetto et al.](#) (2021, p. 1190, tradução nossa) “[...] pela primeira vez na história da humanidade, estamos enfrentando uma pandemia que é descrita, debatida e investigada em tempo real pela comunidade científica por meio dos canais convencionais de pesquisa (e.g. artigos de periódicos) em conjunto com as mídias sociais *online*”. Tal colocação demonstra que a elaboração e disseminação da informação científica sobre COVID-19 consistiu em um esforço coletivo dos pesquisadores que utilizaram todas as possibilidades disponíveis para o avanço da ciência. Em virtude da dinamicidade e facilidade das redes sociais, essas foram utilizadas pela comunidade para estabelecer trocas, discussões e validações de *preprints* e

⁶ “Using bibliometric approaches to analyse scientific output relating to recent public health emergencies helps reveal the specific contributions academia makes to resolve health crises. Several studies stand out as good examples of where bibliometrics has provided a useful international view of medical and multidisciplinary research efforts” (ZHANG *et al.*, 2020, p. 748)

publicações sobre a temática. Essas atividades geram dados relacionados com a atenção dada às publicações científicas fora do contexto acadêmico formal, que podem ser analisados por meio da Altmétrie.

Cientes do contexto delineado, [Torres-Salinas, Robinson-Garcia e Castillo-Valdivieso \(2020\)](#) desenvolveram um estudo altmétrico sobre a produção científica de COVID-19 disponível até abril de 2020, estabelecendo comparações entre artigos de periódicos e *preprints*, tanto de acesso aberto quanto fechado. Em seus resultados, os autores apresentam que publicações analisadas em acesso aberto receberam maior atenção nas redes sociais quando contrastadas com as fechadas. [Bermúdez-Rodríguez et al. \(2020\)](#) e [Boetto \(2021\)](#), adotando perspectivas, métodos e fontes de informação diferentes, também desenvolveram análises altmétricas sobre a produção científica na temática de COVID-19, demonstrando que os métodos bibliométricos e altmétricos funcionam de modo complementar para o estabelecimento de compreensão sobre as dinâmicas de disseminação e impacto das publicações científicas.

Dadas as considerações apresentadas como principais pontos de justificativa para o desenvolvimento da presente pesquisa, a seguir, encontra-se desenvolvido o referencial teórico que apoia a reflexão sobre os conceitos essenciais para realização da investigação proposta.

2 REFERENCIAL TEÓRICO

Neste capítulo serão abordados aspectos e assuntos fundamentais para o embasamento teórico da pesquisa. De início, apresenta-se a comunicação e a produção científica de forma mais ampla, a fim de contextualizar o tópico que discute esses conceitos em tempos de COVID-19, destacando de que modo o período pandêmico impactou nas dinâmicas tradicionais de comunicação e publicação do conhecimento científico e apresentando o *preprint* como um canal que recebeu bastante destaque para veicular novas descobertas na circunstância considerada. Posteriormente, trata-se da origem e evolução do acesso aberto abordando, também, as possibilidades e subtipos atuais de disponibilização da produção científica abertamente e os argumentos favoráveis ao MAA. Por fim, são apresentadas questões sobre as vantagens de citação e altmétricas de acesso aberto, colocando sob perspectiva os indicadores bibliométricos de citação e altmétricos como possibilidades para viabilizar as investigações sobre as hipóteses de OACA e OAAA.

2.1 A COMUNICAÇÃO E PRODUÇÃO CIENTÍFICA

No processo de pesquisa e construção consensual do conhecimento científico, a comunicação ganha protagonismo como atividade fundamental para a continuidade do fazer científico. A importância da comunicação para a ciência fez com que diversos pesquisadores, ao longo da história, se debruçassem sobre essa temática buscando compreendê-la, por exemplo, com relação à definição, às características, à estrutura, aos atores envolvidos e aos principais canais utilizados. Segundo Pinheiro (2012), no campo da Ciência da Informação (CI), a comunicação científica está entre as subáreas mais estudadas. O interesse por essa disciplina na CI, de acordo com Mueller e Passos (2000), é observado na literatura da área desde o período que antecede a década de 1950, mas ganhou impulso a partir da publicação do trabalho de Derek Solla Price, em 1963.

Targino e Torres (2014) questionam: o que haveria acontecido se as descobertas realizadas não tivessem sido comunicadas, compartilhadas e disseminadas? Para Ziman (1984), a publicação dos resultados de pesquisa é um princípio básico da ciência. Uma vez publicadas, as descobertas podem ser analisadas, avaliadas, validadas e/ou questionadas pela comunidade científica. Isto posto, a comunicação dos resultados de pesquisa se mostra como fundamental para garantir o intercâmbio de informações entre os pesquisadores e para manter a dinâmica de produção do conhecimento. Assim,

[...] pode ser que novas descobertas venham demonstrar que aquela teoria não era afinal um instrumento de predição tão bom quanto se acreditava. E adaptações ou novas teorias serão propostas, de novo testadas, aceitas ou rejeitadas. O papel central da comunicação científica na produção do conhecimento científico e na obtenção do consenso que o avaliza se torna claro nesse processo (MUELLER; PASSOS, 2000, p. 14)

Mas afinal, o que se compreende como comunicação científica? [Garvey \(1979\)](#) atribui a essa um papel fundamental para o desenvolvimento das pesquisas e para o avanço do conhecimento científico, entendendo-a como a essência da ciência, aspecto que se torna evidente a partir da observação do título que identifica o livro do autor: *Communication: the essence of Science*. Ainda segundo Garvey (1979, p. ix, tradução nossa), a definição de comunicação científica integra um “[...] conjunto completo de atividades associadas à produção, disseminação e uso da informação desde o momento em que o cientista tem a ideia de sua pesquisa até que a informação sobre os resultados dessa pesquisa seja aceita como

constituente do conhecimento científico”⁷. Já Meadows (1999, p. vii) expõe que, dado seu grau de relevância, a comunicação científica encontra-se situada no coração da ciência e “[...] é tão vital quanto a própria pesquisa”. Logo, nota-se que as interpretações de ambos autores atribuem *status* de protagonismo para tal conceito no contexto da ciência.

Quando se trata das questões terminológicas sobre comunicação científica, Pinheiro (2012) aponta que a expressão está associada a dois termos em inglês: *scholarly communication* e *scientific communication*. Sendo o segundo cunhado pela primeira vez em 1939 por John Desmond Bernal a partir da publicação da obra *The social function of science* (CARIBÉ, 2015).

De acordo com Meadows (1999), as necessidades dos cientistas de coletar dados, testar hipóteses, desenvolver teorias e modificar ideias faz com que esses se envolvam com a comunicação e busquem aprimorá-la, uma vez que essa se mostra essencial para a evolução das pesquisas e a construção dos conhecimentos. Nesse sentido, Borko (1968, p. 4, tradução nossa) assinala que “se os procedimentos de troca entre a comunicação e a informação não forem melhorados, todos os outros trabalhos científicos serão impedidos; a falta de comunicação resultará na duplicação de esforço e num lento progresso” .

O interesse pelo entendimento sobre as dinâmicas e processos de comunicação estabelecidas entre os membros da comunidade acadêmica impulsionaram, ao longo do tempo, a proposição de diversos modelos conceituais que buscam descrever e representar os fluxos de informação científica. Conforme Fernandes e Vilan Filho (2021), os modelos, quando validados, são ferramentas que auxiliam na compreensão dos eventos associados à comunicação científica. Diante do entendimento da potencialidade das representações como fontes relevantes para compreensão do sistema de comunicação, os autores apresentam uma análise de vinte e sete modelos identificados na literatura da CI.

O mais antigo e considerado como um dos mais representativos entre os modelos existentes, é o de [Garvey e Griffith \(1972\)](#), que demonstra, por meio da dinâmica de evolução do tempo, a sequência de eventos ocorridos no processo de comunicação científica, desde a concepção do projeto de pesquisa até o registro da primeira citação do trabalho. A representação proposta pelos autores exemplifica que todo esse processo durava, em média, sete anos. O referido modelo foi proposto a partir da lógica de disseminação da produção científica no

⁷ “[...] our definition includes the full spectrum of activities associated with the production, dissemination, and use of information from the time the scientist gets the idea for his research until information about the results of this research is accepted as a constituent of scientific knowledge” (GARVEY, 1979, p. ix).

âmbito da Psicologia, todavia, ao longo do tempo, esse demonstrou características aplicáveis e de interesse de outras áreas do conhecimento.

A partir da concepção do fluxo supramencionado, outras propostas surgiram. Algumas delas com o intuito de aprimorar o trabalho de Garvey e Griffith (1972), a exemplo das proposições feitas por Hurd (1996) e Costa (2000), e outras adotando perspectivas distintas, como é o caso do modelo Unisist, publicado pela United Nations Educational, Scientific and Cultural Organization (Unesco) (FERNANDES; VILAN FILHO, 2021; LEITE, 2011). Segundo Mueller e Passos (2000), na representação do fluxo de informação científica proposta por Garvey e Griffith (1972), é possível notar tanto a comunicação formal quanto a informal. Já no modelo Unisist, por exemplo, são apresentados três tipos de comunicação: formal, informal e tabular, que é utilizada para disseminar documentos entendidos como complementares, a exemplo de dados, tabelas e mídias.

Além das possibilidades de classificação dos tipos de comunicação existentes, os canais utilizados para a transmissão da informação científica também podem ser de diferentes tipologias, sendo que as principais categorias citadas pela literatura são: formal e informal. Meadows (1999) esclarece que as práticas iniciais vinculadas aos processos de comunicação científica remetem às trocas de manuscritos e cartas efetuadas entre os acadêmicos gregos dos séculos V e IV a.C. A partir desse marco, o autor descreve uma evolução histórica ocidental do surgimento dos canais utilizados para comunicação entre integrantes da comunidade científica, iniciando nas cartas e abordando até a origem dos periódicos científicos na segunda metade do século XVII.

A descrição da evolução dos canais de comunicação apresentada por Meadows (1999) demonstra que a comunicação científica traçou um caminho progressivo de formalização de seus canais. Tais mudanças seguiram uma lógica evolutiva da comunicação temporária e restrita a um grupo pequeno de cientistas (informal), para o estabelecimento da comunicação formal, que tende a ser mais durável e acessível para os indivíduos, uma vez que encontra-se publicada e pode ser preservada (MEADOWS, 1999; TARGINO, 2000).

Dentre os tipos de canais de comunicação científica existentes, estão situados artigos científicos publicados em periódicos. Consoante Mueller e Passos (2000), os artigos científicos são caracterizados como o principal veículo formal de circulação do conhecimento científico. É por meio deles que os resultados de pesquisa publicados podem ser conhecidos, avaliados, questionados e citados pelos pares. Assim, os artigos científicos permitem que a comunidade

acadêmica coopere entre si para o compartilhamento de novas descobertas e, consequentemente, para a construção da ciência.

Mesmo que os artigos científicos assumam o papel de protagonismo como canal formal mais usado para divulgação de resultados de pesquisa, faz-se válido destacar a existência e a importância de outras possibilidades de comunicação do conhecimento, visto que, no entendimento de Targino (2000, p. 21, grifo da autora), “os cientistas, para difusão de suas pesquisas, sobretudo os resultados parciais, não escolhem de imediato os meios convencionais. São cada vez mais comuns as pré-edições (*preprints*), as versões provisórias (*prepapers*) e as comunicações em congressos ou outros encontros científicos, publicadas ou não”. Como exemplos de canais formais, podemos enumerar os livros, os relatórios de pesquisa, as teses e dissertações. Já no que se refere à comunicação informal, existem os eventos científicos, os *e-mails*, os *preprints*, e outras formas de comunicação verbal não registradas (TARGINO, 2000).

Nota-se que comunicação científica possui caráter dinâmico e apresenta, no decorrer do tempo, transformações vinculadas principalmente aos canais adotados e aos fluxos de informação propostos. Além disso, em situações como as de emergências de saúde pública, a lógica formal de comunicação científica é pressionada a superar desafios em busca de conhecimentos que auxiliem na superação da crise e garantam a manutenção do direito à saúde. [Jorge \(2018\)](#) aponta que

[...] a pesquisa no contexto de uma emergência é essencial. Novas respostas, em tempos diferentes das pesquisas rotineiras, são extremamente necessárias para frear e controlar a propagação de doenças e todos os riscos globais que elas representam. Desta forma, a pesquisa pode fornecer respostas imediatas para garantir a contenção, o tratamento e a prevenção dos eventos, auxiliando na redução do intervalo de tempo entre a declaração de emergência de saúde pública e a disponibilidade de tecnologias médicas eficazes ([JORGE, 2018, p. 110](#)).

A partir das colocações apresentadas, entende-se a relevância de se buscar uma compreensão mais específica sobre as alterações nos fluxos de comunicação e produção científica em períodos de crise, com o foco no contexto mudanças provocadas pela pandemia de COVID-19, que constitui uma das frentes temáticas abordadas pelo presente trabalho.

2.1.1 Comunicação e produção científica em tempos de COVID-19

A caracterização feita pela OMS de uma doença como ESPII tende a aumentar imediatamente o número de publicações sobre o assunto (ZHANG *et al.*, 2020). Os autores

observaram esse fenômeno a partir da análise da produção científica sobre *Severe Acute Respiratory Syndrome* (SARS), influenza A (H1N1), Ebola, Zika e COVID-19. Ao colocar em risco a saúde da população mundial, a pandemia de COVID-19 impulsionou uma maratona sem precedentes em busca de respostas e soluções para a contenção da crise e, como decorrência disso, houve o incremento na quantidade de investigações sobre o assunto.

Além do aumento no volume de publicações, consoante [Bermúdez-Rodríguez et al. \(2020\)](#), a pandemia enfatizou a necessidade de velocidade e de eficiência nos resultados de pesquisa, sendo esses aspectos fundamentais para os processos de elaboração e disseminação do conhecimento científico. Todavia, características existentes no sistema de comunicação científica vigente tendem a dificultar a agilidade no processo de elaboração e disseminação do conhecimento, a exemplo da manutenção dos padrões de ciência fechada e das barreiras linguísticas em decorrência da ênfase dada à publicação no idioma inglês (LARIVIÈRE; SHU; SUGIMOTO, 2020).

Jorge, Albagli e Souza (2020) pontuam que um dos desafios vividos pelos cientistas que trabalharam com a temática da COVID-19 durante o surto da doença tem relação com o acesso às informações e aos dados de forma ágil para subsidiar os questionamentos, as análises, as investigações e, conseqüentemente, otimizar o tempo para alcançar resultados e obter respostas. Paliativamente e com vistas a minimizar a problemática de disponibilização e acesso aos dados e às informações sobre COVID-19, instituições como a *Wellcome Trust*⁸ e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)⁹ elaboraram recomendações solicitando o compartilhamento de novos conhecimentos em acesso aberto.

Publicada em 30 de janeiro de 2020, a declaração da *Wellcome Trust*¹⁰ direciona o apelo pelo compartilhamento rápido e aberto dos resultados e dados aos pesquisadores, periódicos e financiadores, justificando que a concessão de tais informações contribui para a formulação de respostas à crise e para salvar vidas. Essa declaração, assinada por mais de 150 signatários (incluindo instituições de pesquisa, editoras, servidores de *preprints*, universidades e outros), estabelece compromissos direcionados à disponibilização e compartilhamento imediato,

⁸ Disponível em: <https://wellcome.org/>. Acesso em: 15 mar. 2021.

⁹ Disponível em: <http://mailer.periodicos.capes.gov.br/?m=136&p=view&pi=ViewBrowserPlugin&uid=69c6496339f687bc1e7ff4ea4624e28f>

¹⁰ Disponível em: <https://wellcome.org/press-release/sharing-research-data-and-findings-relevant-novel-coronavirus-ncov-outbreak>

gratuito e aberto das publicações, dos *preprints* e dos dados, ao menos pelo período de duração do estado de emergência.

No cenário brasileiro, a [carta aberta](#)¹¹ divulgada em conjunto pela Academia Brasileira de Ciências (ABC), a Associação Brasileira de Imprensa (ABI) e a Sociedade Brasileira para o Progresso da Ciência (SBPC) em 17 de março de 2020, solicitou a liberação de acesso irrestrito a todo conteúdo, seja escrito ou audiovisual, que tratasse da temática da pandemia ou do coronavírus. A carta ainda ressalta que “para os cientistas, pesquisadores, médicos e pessoal da área da saúde, é essencial o acesso aberto a todos os dados e à produção científica e técnica sobre a pandemia e aquela relacionada ao coronavírus” ([SOCIEDADE BRASILEIRA PARA O PROGRESSO DA CIÊNCIA, 2020, online](#))

Com intuito de reforçar o pedido feito pela carta supramencionada, por meio de ofício datado de 20 de março de 2020, a CAPES emitiu aos editores internacionais uma solicitação de disponibilização de artigos e outros documentos relevantes para o estudo da temática em questão de forma irrestrita e ágil ([COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR, 2020](#)). Editoras como Clarivate Analytics, Elsevier, Emerald, Royal Society, Wiley e outras liberaram o acesso aos conteúdos relacionados ao novo coronavírus por tempo limitado, sendo esses intervalos de disponibilização variáveis e definidos por cada editor.

As solicitações e apelos supramencionados demonstram que a remoção de barreiras de acesso às publicações é fundamental para o avanço do conhecimento científico. Entretanto, segundo [Larivière, Shu e Sugimoto \(2020\)](#), por mais que as ações de abertura temporária de acesso aos resultados e dados de pesquisa tenham sido importantes para o momento emergencial, elas não são suficientes, visto que a proporção de literatura disponibilizada por meio de tais medidas não reflete o todo existente, mantendo limitações de acesso à produção científica essencial para subsidiar o trabalho dos pesquisadores. Nesse sentido, os mesmos autores expõem, com base em dados de registros indexados na Web of Science (WoS), que mais da metade (51,5%) dos artigos publicados sobre coronavírus desde 1960 estão com acesso fechado, ou seja, apresentam alguma barreira (cobrança de taxas ou assinatura) para a disponibilização do conteúdo completo da publicação gratuitamente e com a declaração de direitos de propriedade aos autores.

¹¹ Disponível em: <http://portal.sbpnet.org.br/noticias/sbpc-abc-e-abi-pedem-que-meios-de-comunicacao-liberem-acesso-a-todas-as-informacoes-sobre-coronavirus/>

Dado que “cada cientista enxerga não só com seus próprios olhos, mas também com os olhos de seus predecessores e de seus colegas” (MUELLER; PASSOS, 2000, p. 14), compreender e analisar o que já foi publicado é fundamental para que as pesquisas avancem e gerem resultados. Desse modo, a fim de que os pesquisadores tenham acesso às descobertas prévias feitas por seus pares, faz-se necessário que essas estejam registradas e disponibilizadas publicamente.

Além dos aspectos associados aos tipos de acesso às publicações, a elevada demanda por esforços de pesquisa sobre COVID-19 gerou outras consequências a serem observadas a partir da perspectiva da comunicação científica. [Santos-d'Amorim \(2021\)](#) elenca duas principais: *aceleração de processos de revisão por pares* e *aumento da quantidade de preprints*. O primeiro ponto trata da redução observada, durante o período de pandemia, no tempo utilizado pelas editoras para avaliar as versões submetidas aos periódicos científicos, sugerir melhorias e adaptações aos trabalhos analisados e publicar a versão final da pesquisa. Sendo a revisão por pares considerada como um mecanismo relevante para certificar “a qualidade, confiabilidade, integridade e consistência da literatura acadêmica” (NASSI-CALÒ, 2015, *online*), levanta-se a preocupação de que a aceleração nos processos de avaliação impactem negativamente na qualidade da produção científica publicada sobre COVID-19 e pandemia.

Assim como houve o incremento na quantidade de artigos publicados durante a eclosão da crise causada pelo novo coronavírus, autores como [Vlasschaert, Topf e Hiremath \(2020\)](#), Torres-Salinas, Robinson-Garcia e Castillo-Valdivieso (2020) e [Fraser et al. \(2021\)](#) abordam em seus estudos a questão da elevação no volume de *preprints* sobre COVID-19. Esse tipo de publicação, apesar de não ser nova no cenário da comunicação científica, ganhou espaço recentemente em virtude das exigências por velocidade e abertura nos processos de disseminação de novas descobertas científicas colocadas pela ameaça à saúde pública global. Dada a conjuntura apresentada, a seguir, busca-se caracterizar o *preprint*, de modo a corroborar com o entendimento sobre o papel desse tipo de canal de comunicação científica no contexto de interesse do presente estudo.

2.1.2 Preprint

Os *preprints* são caracterizados como manuscritos que não passaram pelo processo formal de revisão por pares e que são disponibilizados *online* e gratuitamente com o intuito de disseminar novas descobertas para a comunidade acadêmica de modo mais rápido. Ou então, segundo [Kling \(2005\)](#), entende-se que os *preprints*, também reconhecidos como *e-preints*, são

manuscritos ainda não publicados, mas que já podem ter sido revisados por pares; submetidos e/ou aceitos para publicação; bem como, destinados para efetiva publicação. Logo, para o referido autor, o *status* de “não publicado” em periódicos formais é o que confere ao manuscrito a qualidade de *preprint*, havendo ou não o processo de revisão por pares. Kling (2005) também menciona que o *preprint* representa a relação existente entre dois documentos: a primeira versão de uma publicação (a qual será submetida aos procedimentos de avaliação) e a publicação definitiva.

Sarabipour *et al.* (2019) e [Vlasschaert, Topf e Hiremath \(2020\)](#) destacam que os *preprints* possuem características vantajosas, a exemplo da disponibilização em acesso aberto, das possibilidades de *feedbacks* e da rápida disseminação de resultados. Entretanto, assim como existe o lado positivo do uso de *preprints*, a ausência de validação por meio do processo de revisão por pares é uma desvantagem destacada por autores como Zhang *et al.* (2020), Vlasschaert *et al.* (2020) e Majumder e Mandl (2020), pois gera questionamentos relacionados à credibilidade, qualidade e desinformação associada aos conteúdos disponibilizados. Logo, mesmo que os *preprints* configurem canais relevantes para a comunicação das descobertas científicas e tenham demonstrado grande importância para o avanço da ciência em períodos emergenciais, como o surto de COVID-19, o uso desse canal de comunicação informal demanda cautela e criticidade dos pesquisadores, visto que, ao analisar um *preprint*, o cientista ocupa simultaneamente o papel de leitor e revisor do conteúdo acessado.

A reflexão em torno do conceito e das características do *preprint* perpassa pelo entendimento sobre a maneira como esse tipo de publicação se assemelha ou se distancia dos atributos dos artigos científicos revisados por pares, uma vez que, como já mencionado anteriormente e destacado por Kling (2005), o *preprint*, no processo de produção do conhecimento científico, ocupa uma posição mediana entre os extremos do primeiro registro de pesquisa até a sua publicação final em canais formais de comunicação científica após a avaliação pelos pares. Assim, no **Quadro 1** abaixo, encontra-se expressa uma comparação entre as características dos *preprints* e dos artigos científicos tradicionais.

Quadro 1 - Diferenças entre *preprints* e artigos tradicionais

Características	Preprints	Artigos
Velocidade de publicação* *período de tempo transcorrido entre a submissão e publicação efetiva	Instantaneamente Costumam ficar disponíveis <i>online</i> até 48h após a submissão.	Geralmente de semanas a meses

Acesso	Gratuito, aberto	Variável Podem ser publicados tanto abertamente, quanto em periódicos que cobram assinatura e/ou acesso pago. Muitos periódicos publicaram artigos temporariamente gratuitos durante o período de pandemia da COVID-19.
Revisão por pares	Normalmente, não há revisão por pares externos antes de submissão em repositórios. Adota mecanismo para revisão por pares aberta (comentários, mídias sociais, troca de e-mails).	Revisado às cegas por pares.
Credibilidade	Muito variável A verificação prévia da qualidade depende dos autores do manuscrito.	Variável São submetidos a procedimentos editoriais formais e revisão por pares. Todavia, também podem apresentar falhas e erros.

Fonte: [Vlasschaert, Topf e Hiremath \(2020, p. 421, tradução nossa\)](#) com adaptações.

[Rodríguez \(2019\)](#) esclarece que a evolução na direção da aceitação do *preprint* é resultado da combinação entre quatro pontos principais:

- a) Entendimento do *preprint* como uma etapa natural do processo de produção e publicação de um artigo científico;
- b) Melhorias relacionadas aos padrões de publicação, as quais passam a considerar a relevância do momento inicial da pesquisa para a sua posterior consolidação e registro através da publicação;
- c) Defesa do acesso à informação e do arquivamento no contexto do MAA;
- d) Em resposta a atrasos relacionados ao tempo para publicação, demanda pela aceleração e imediatismo no contexto da comunicação da ciência.

O último tópico supramencionado se aproxima das ponderações feitas por [Vlasschaert, Topf e Hiremath \(2020\)](#) a respeito da expansão do volume de *preprints*, nas temáticas de saúde, depositados em repositórios específicos para esse tipo de publicação durante a pandemia de COVID-19. Segundo os autores, o cenário de necessidade de aceleração da disseminação dos dados e dos resultados de pesquisa na área médica, com foco na nefrologia, impulsionou um aumento da quantidade de *preprints* produzidos e disponibilizados, fazendo com que esses passassem a receber maior atenção dos pesquisadores dessas áreas. Os autores ainda colocam que, antes da pandemia de COVID-19, a questão da aceitação e utilização dos *preprints* nas

áreas clínicas e da saúde era baixa, principalmente em virtude da ausência do processo de revisão pelos pares.

Faz-se válido ressaltar que o grau de utilização e de aceitação dos *preprints* como possibilidades de disseminação e acesso aos resultados iniciais de investigações científicas é um comportamento variável entre as áreas e as disciplinas, de modo que não é possível considerar uma lógica geral de produção e disseminação dessa tipologia de publicação para todas as ramificações do conhecimento científico. Por mais que, em áreas como a saúde, por exemplo, a questão do aumento da visibilidade dos *preprints* seja um acontecimento mais novo ([VLASSCHAERT, TOPF E HIREMATH, 2020](#)), em outras, como na Física e na Matemática, a prática de uso dos *preprints* encontra-se estabelecida há bastante tempo ([FRASER et al., 2020](#); [RODRÍGUEZ, 2019](#)).

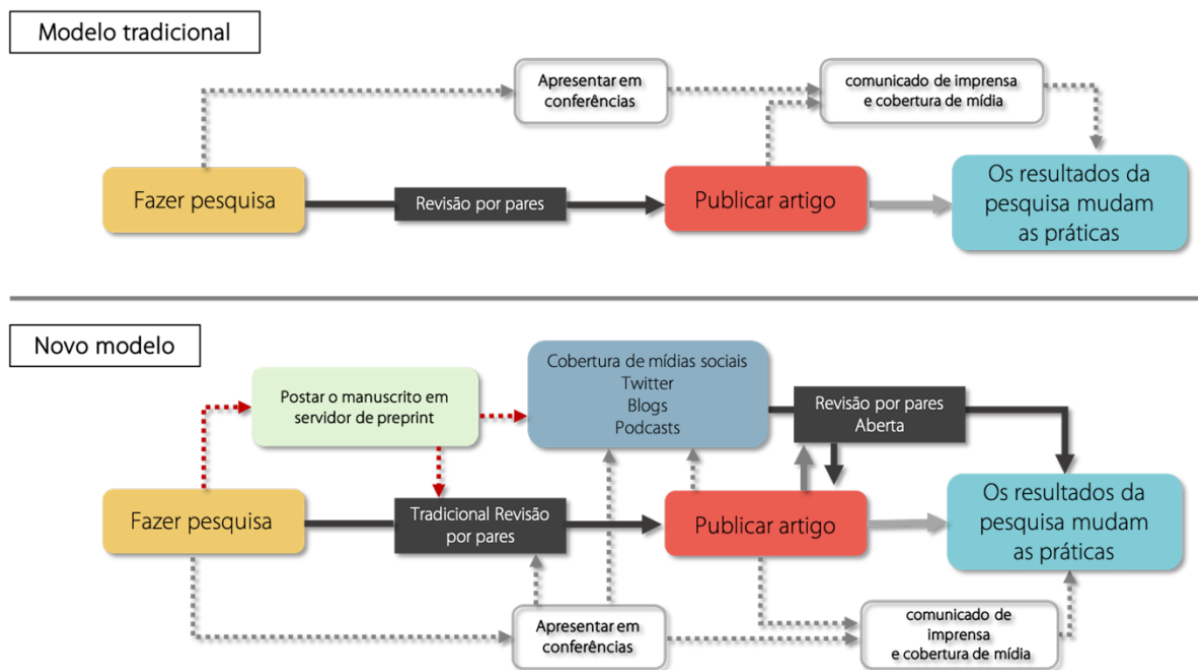
Mesmo que, em determinados contextos, a popularização do *preprint* enquanto canal de comunicação científica seja algo novo, o *preprint*, em si, não é um acontecimento recente. Um exemplo disso é que, no modelo de comunicação científica considerado como o mais tradicional e clássico, o de Garvey e Griffith (1972), os *preprints* já eram mencionados na etapa anterior à publicação de um artigo no periódico científico. Até chegar a versão atual de um documento digital submetido e disponibilizado abertamente em servidores/repositórios específicos, o modelo de *preprint* percorreu um caminho evolutivo. Inicialmente, “eles circulavam como cartas, memorandos ou cópias de rascunhos por e-mail entre pesquisadores, para comentários antes da submissão do periódico” ([RODRÍGUEZ, 2019, p. 6](#)). Nesse contexto, observa-se que a circulação das versões iniciais das publicações ocorria de modo mais restrito e direcionado a um grupo pequeno de pesquisadores e, além disso, nota-se que a distribuição desses documentos também ocorria no formato impresso. Nesse momento, os *preprints* cumpriam com as funções de serem mecanismos para que os proponentes das pesquisas recebessem *feedbacks* e sugestões informais de seus pares e, de certa forma, garantissem a autoria pela investigação em andamento, uma vez que parte da comunidade científica ficava ciente dos estudos propostos ([TARGINO, 2000](#)).

Com a evolução das possibilidades tecnológicas inauguradas no ambiente digital, as versões ora impressas passaram a ser disponibilizadas de forma eletrônica, fazendo com que *preprints* recebessem denominações como *e-prints* ou *e-preprints* ([RODRÍGUEZ, 2019](#)). A partir de então, começaram a surgir os servidores de *preprints* (que serão detalhados mais à frente), onde as versões digitais prévias das publicações podem ser disponibilizadas

gratuitamente para toda comunidade acadêmica e sociedade em geral, que desejem ter acesso aos resultados preliminares de pesquisas realizadas.

As possibilidades de comunicação e disseminação das descobertas científicas influenciam e são influenciadas pelo contexto social e tecnológico em que estão inseridas. Logo, a contemporaneidade, marcada pela dinâmica de difusão de informações por meio das mídias sociais, promove a apresentação dessas mídias como novos canais utilizados para comunicar ciência. Conscientes de tal cenário e, ao estudarem a popularização dos *preprints* em virtude das alterações causadas pela pandemia de COVID-19 no fluxo tradicional de comunicação científica, [Vlasschaert, Topf e Hiremath \(2020, online\)](#) propuseram uma sistematização comparativa (**Figura 1**) entre o que os autores entendem como o tradicional e o novo modelo de comunicação científica para a nefrologia, subcampo da área da saúde.

Figura 1 - Comparação entre as diferenças do modelo tradicional e o modelo atual de comunicação científica na área da saúde



Fonte: [Vlasschaert, Topf e Hiremath \(2020\)](#) traduzido por [Santos-d'Amorim \(2021, p. 12\)](#)

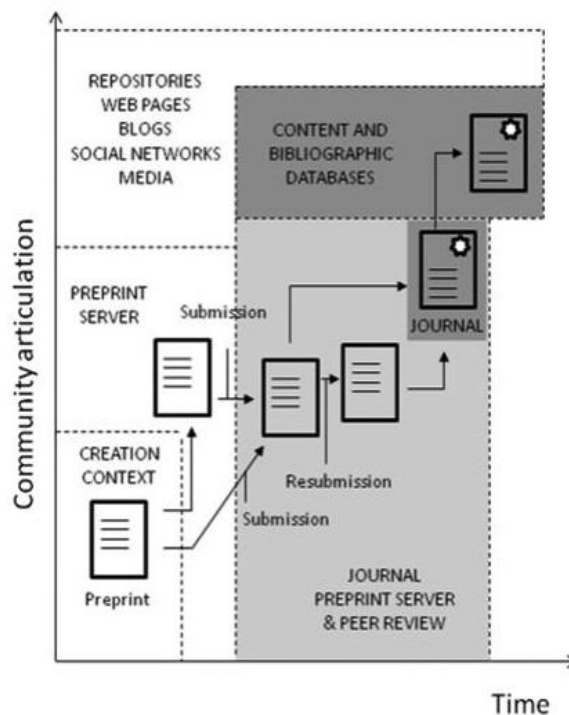
Em relação ao novo modelo proposto por [Vlasschaert, Topf e Hiremath \(2020\)](#), na **Figura 1**, são válidos alguns apontamentos:

- a) Fica evidente o papel de destaque que os autores deram para os *preprints* e para as mídias sociais no fluxo de publicação de novos artigos;

- b) Além da revisão tradicional por pares, é apresentada a revisão aberta, demonstrando o interesse dos autores em revelar as possibilidades de inovação e adaptação nas etapas de publicação de uma pesquisa a partir da observação das influências atuais da dinamicidade do ambiente digital e das mídias e redes sociais;
- c) Os autores demonstram, de forma explícita, somente a etapa de postagem do manuscrito em um servidor de *preprint*. Entretanto, entende-se que a representação do novo modelo ficaria mais completa se a etapa de submissão do artigo a um periódico científico fosse claramente indicada na **Figura 1**, por mais que seja possível compreender a existência dessa etapa a partir da interpretação do contexto;
- d) Na atualidade, os *preprints* são utilizados tanto para a viabilização da ciência, pois proporcionam o acesso rápido aos novos conhecimentos gerados, quanto para a consolidação da ciência, uma vez que viabilizam etapas de avaliação e certificação de descobertas recentes.

Ao tratar dos passos de articulação da comunidade acadêmica para disseminar a informação científica em detrimento do tempo, [Rodríguez \(2019\)](#) propõe uma representação visual, apresentada na **Figura 2**, em que a submissão do *preprint* em um servidor configura o primeiro marco de arquivamento de um manuscrito como documento científico.

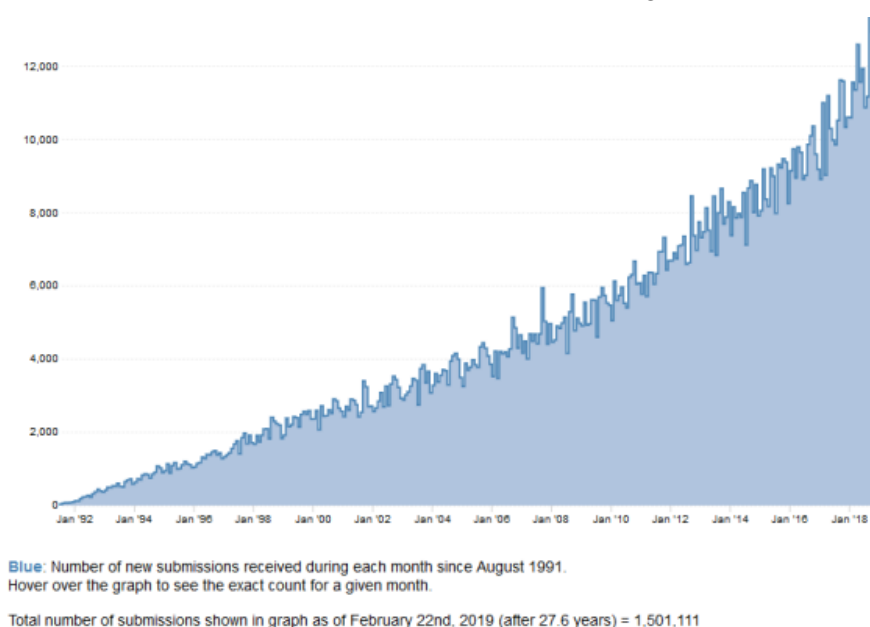
Figura 2 - Fluxo de surgimento do *preprint* no contexto de desenvolvimento da pesquisa até a publicação dos resultados em periódicos posteriormente disseminados.



Fonte: [Rodríguez \(2019, p. 5\)](#)

Os servidores de *preprints* são plataformas que arquivam, recuperam e disponibilizam os documentos/manuscritos digitais, promovendo a articulação entre as comunidades científicas *online* (RODRÍGUEZ, 2019). Como alternativa proposta para a disponibilização de informação científica de forma acessível, foi criado, 1991, o primeiro servidor de *preprint*: arXiv¹². Desde sua criação, o repositório registra números crescentes de submissão, conforme expressa o **Gráfico 1**.

Gráfico 1 - Crescimento do número de submissões mensais registradas no arXiv entre 1992 e 2019



Fonte: [Appel \(2019, p. 45\)](#)

Ao considerar os servidores de *preprint* arXiv, bioRxiv¹³ e medRxiv¹⁴ e por meio da ferramenta COVID-19 Primer¹⁵, [Santos-d'Amorim \(2021\)](#) identificou que, até março de 2021, foram indexados mais de 16 mil *preprints* sobre a temática COVID-19 em diversas disciplinas e áreas. Nesse sentido, Majumder e Mandl (2020, p. 629, tradução nossa) afirmam que “o impacto dos preprints no discurso e nas tomadas de decisão relativas ao surto de COVID-19 em andamento sugere que devemos repensar como recompensamos e reconhecemos as contribuições da comunidade durante as crises de saúde pública presentes e futuras” e tal

¹² Disponível em: <https://arxiv.org/>

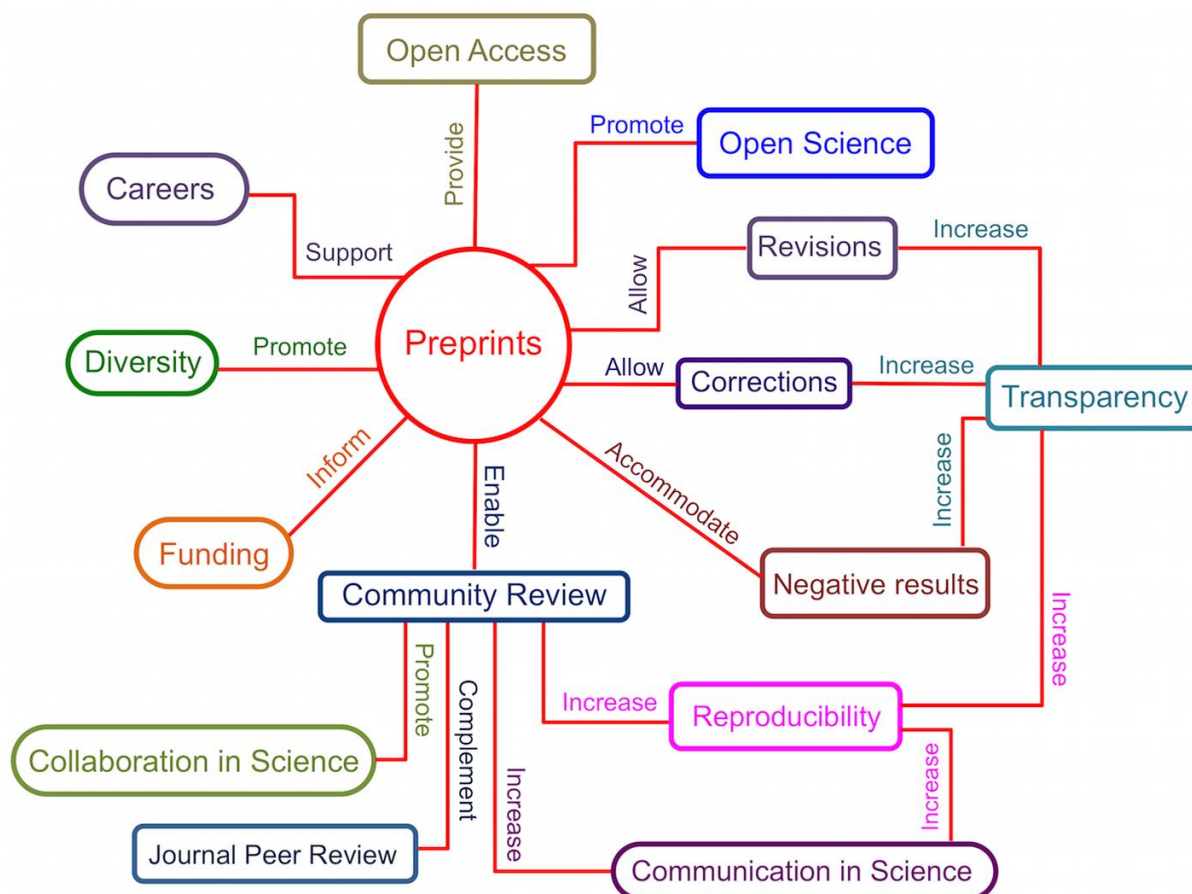
¹³ Disponível em: <https://www.biorxiv.org/>

¹⁴ Disponível em: <https://www.medrxiv.org/>

¹⁵ Disponível em: <https://covid19primer.com/>

pensamento reforça a necessidade de estudos que busquem investigar as contribuições e consequências da popularização de uso dos *preprints* para a comunicação científica de modo geral. E, de forma mais específica, conforme recomenda Zhang *et al.* (2020), desenvolver pesquisas que tentem compreender melhor de quais maneiras os *preprints* podem afetar ou beneficiar as respostas dadas pela academia às emergências de saúde pública.

Figura 3 - Mapeamento das influências dos *preprints* no contexto acadêmico



Fonte: [Sarabipour et al. \(2019, p. 4\)](#).

Conforme sumariza a **Figura 3**, [Sarabipour et al. \(2019\)](#) propõem uma análise a respeito dos pontos positivos diretamente relacionados ao uso dos *preprints*, levando em consideração as contribuições desse tipo de canal de comunicação para a carreira de pesquisadores iniciantes da área de Ciências da Vida. Além das questões já mencionadas, como a otimização de tempo para disponibilização de resultados e a colaboração que os *preprints* fornecem para fomentar as ideias defendidas pelo MAA e a Ciência Aberta, Sarabipour *et al.* (2019) argumentam que o tipo de publicação em questão propicia o estabelecimento de contato e colaboração entre os pesquisadores e gera visibilidade às investigações, uma vez que algumas discussões e interações relativas aos *preprints* ocorrem no contexto das redes sociais. Nesse sentido, os autores apontam

que a publicação de um *preprint* implica na elevação de indicadores altmétricos (menções em mídias sociais e notícias, por exemplo) e indicadores de citação dos artigos científicos, frutos da mesma pesquisa do manuscrito, que são publicados *a posteriori*.

Com ênfase na área da Medicina, [Serghiou e Ioannidis \(2018\)](#) ponderam que os processos de debates e avaliações criadas em torno dos *preprints* geram dados sobre uso, visualização, *downloads*, índices altmétricos e de citação, que podem ser analisados. Tendo em vista tal colocação, os autores investigaram os dados oriundos da PubMed, Altmetric.com e CrossRef vinculados a 7760 *preprints* disponibilizados entre 2013 a 2017, a partir dos resultados calculados para as medianas de *downloads*, visualizações de resumos das publicações arquivadas, citações da CrossRef e Altmetric *score*.

Um dos aspectos mencionados na reflexão supradelineada sobre as possibilidades e tipos de disponibilização do conhecimento científico foi o acesso aberto. Para a comunicação científica, o MAA levanta questionamentos e novas possibilidades que contrastam com o modelo de comunicação tradicional da ciência, baseado na publicação de novas descobertas em periódicos científicos que adotam modelos de negócios voltados à cobrança de taxas de assinatura e acesso às informações. Considerando a contextualização posta, a próxima seção busca aprimorar e detalhar a questão do acesso aberto evidenciando novos pontos relevantes para o embasamento deste estudo.

2.2 ACESSO ABERTO

Gratuito, *online*, irrestrito, livre de custos de assinaturas e de restrições de direitos autorais, e com possibilidades de compartilhamento e de reuso dos conteúdos publicados. Essas são características que qualificam e orientam a compreensão sobre em que consiste o acesso aberto. Diferentemente de um acontecimento repentino e isolado, o acesso aberto surge como fruto de uma construção progressiva baseada em ações, iniciativas e documentos múltiplos. De acordo com a linha do tempo do MAA traçada por [Suber \(2009\)](#), a partir da década de 60 do século XX, começaram a ser observados acontecimentos relevantes para estimular o delineamento de um cenário propício para a lapidação da ideia que passou a ser definida posteriormente como acesso aberto ([SUBER, 2009](#)).

Mesmo havendo sinais que indicassem necessidades relacionadas ao que é defendido pelo MAA desde a década de 60, [Appel \(2019, p. 43\)](#) assinala que “os anos 1990 podem ser considerados como um marco para o surgimento de crescentes movimentos de reação aos

cercamentos da informação e do conhecimento científico e pela proposição de práticas alternativas de publicação e disseminação científica”. Foi também no referido período que, segundo [Cintra \(2017\)](#), ocorreu o advento e a expansão das Tecnologias de Informação e Comunicação (TICs), acompanhadas dos primeiros indícios de uso dos periódicos eletrônicos.

Dentre as novas possibilidades de disponibilização da informação científica em meio eletrônico inauguradas no fim do século XX, encontra-se uma iniciativa caracterizada como pioneira e significativa para o contexto do acesso aberto: a criação do repositório de arquivos de *preprints* arXiv.org¹⁶, em 1991. Voltado para o armazenamento de texto completo das publicações das áreas de Física, Matemática, Ciência da Computação, Biologia, Estatística e Finanças, Engenharia elétrica, Economia e outras áreas afins. Esse repositório, atualmente hospedado e gerenciado pela Cornell University Library, apresentou uma nova forma de promover a circulação dos *preprints* entre os pares ([APPEL, 2019](#); [COSTA; LEITE, 2017](#)).

[Rodrigues \(2004\)](#) destaca que a criação da Scholarly Publishing and Academic Resources Coalition (SPARC), em 1998, pela Association of Research Libraries (ARL) foi um outro fato importante para a consolidação do MAA. A SPARC tornou-se uma organização internacional científica atuante na promoção do acesso aberto e publicou, juntamente com a Triangle Research Libraries Network (TRLN), a [Declaring Independence](#) (Declaração de Independência). Escrita por Michael L. Rosenzweig (2001), o texto dessa declaração faz um apelo aos pesquisadores pela mobilização em busca de alternativas de disponibilização dos resultados de pesquisas de forma independente das editoras e dos periódicos científicos comerciais, utilizando alternativas que atendam às necessidades da comunidade acadêmica de acesso à informação científica para dar prosseguimento às investigações e produção do conhecimento. Tal declaração expressa que o perfil curioso e explorador dos cientistas é um aspecto a ser levado em consideração e que pode auxiliar no descobrimento de caminhos alternativos para disponibilizar os resultados de investigação científica por meio do uso da gestão cautelosa aliada às tecnologias.

Ainda com relação às ações que antecederam a formalização do MAA e forneceram suporte à concepção desse movimento, [Sarmiento et al. \(2005\)](#) enumeram a [Declaração de Santo Domingo](#) (1999) e a [Declaração sobre a Ciência e o Uso do Conhecimento Científico](#) (UNESCO, 1999), documentos elaborados no contexto da Conferência Mundial sobre a Ciência para o século XXI, que colocou em discussão a relevância da ideia de acesso igualitário e

¹⁶ Disponível em: <https://arxiv.org/>. Acesso em: 27 nov. 2021.

universal ao conhecimento científico e a necessidade da remoção de barreiras dos sistemas educacional e de pesquisa.

O ano de 1999 também foi marcado pelo estabelecimento da [Open Archives Initiative](#) (OAI), em português, Iniciativa de Arquivos Abertos, que se firmou a partir das discussões elaboradas na Convenção de Santa Fé. A proposta da OAI, em essência, envolve o desenvolvimento e a promoção de padrões de interoperabilidade com o intuito de usufruir das possibilidades trazidas pelo emergente cenário de disponibilização das publicações científicas em meio eletrônico. Para Costa e Leite (2017), os desdobramentos da OAI, como o protocolo de interoperabilidade *Open Archives Initiative Protocol for Harvesting* (OAI-PMH), mostraram-se fundamentais para a prosperidade dos sistemas de acesso aberto.

A combinação dos acontecimentos mencionados com outros eventos, documentos e iniciativas, compilados cronologicamente no **Quadro 2**, gerou expectativas voltadas para a ampliação das possibilidades de produção, distribuição e acesso às publicações científicas antes distribuídas principalmente no formato impresso e por meio da mediação de editoras científicas e comerciais cujo modelo de negócio encontrava-se baseado na cobrança de assinaturas para disponibilizar o acesso aos conteúdos publicados ([APPEL, 2019](#)).

Quadro 2 - Marcos relevantes para o MAA

Ano	Marco/Evento/Documento
1991	Lançamento do repositório de preprints arXiv.org
1992	Reunião da Association of Research Libraries (ARL)
1993	Lançamento do Projeto MUSE, pioneiro na distribuição de documentos <i>online</i>
1994	Proposta de autoarquivamento por Stevan Harnad, conhecida como “ A subversive proposal ”
1997	Lançamento do PubMed com incorporação do conteúdo da Medline
	Início do desenvolvimento do projeto Scientific Electronic Library Online (SciELO)
	A base Medline torna-se de acesso aberto e é incorporada pelo PubMed.
1998	Lançamento da SPARC (Scholarly Publishing and Academic Resources Coalition)
	SciELO começa a operar publicamente.

1999	Definição da Open Archives Initiative (OAI) na ocasião da Convenção de Santa Fé
	Declaração sobre a Ciência e Uso do Conhecimento Científico (<i>Declaration on Science and the Use of Scientific Knowledge</i>), Unesco
2000	Lançamento da PubMed Central (PMC) com artigos de texto completo de livre acesso
	<i>Public Library of Science</i> (PLOS)
2001	Lançado o <i>Wayback Machine</i> pelo <i>Internet Archive</i>
2002	Budapest Open Access Initiative (BOAI)
	Projetos SHERPA e <i>Rights Metadata for Open archiving</i> (RoMEO) pelo JISC-FAIR
	Lançamento do software DSpace pelo Massachusetts Institute of Technology (MIT)
2003	Declaração de Bethesda
	Declaração de Berlim
	Tradução do <i>Open Journal System</i> (OJS) é feita pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e é lançado o Sistema Eletrônico de Editoração de Revistas (SEER)
2004	Declaração da Organization for Economic Cooperation and Development (OECD) sobre o acesso aos dados da pesquisa financiada por fundos públicos
	Princípios de Washington D. C. para o Livre Acesso à Ciência
2005	Declaração de São Paulo em apoio ao acesso aberto pelo IBICT. Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica
2006	IFLA e UNESCO lançam o Manifesto Diretrizes sobre a Internet da IFLA / UNESCO
	Lançamento da revista <i>PLoS ONE</i> .
2007	<i>Open Access Research</i> emite chamada de trabalhos e torna-se a primeira revista em acesso aberto revisada por pares dedicada ao próprio acesso aberto
2008	Publicação do livro <i>Open access to knowledge and information: scholarly literature and digital library initiatives: the South Asian scenario</i> , de Anup Kumar Das e edição da Unesco
	Peter Suber e Robin Peek lançam o Open Access Directory
2010	Primeira Conferência Luso-Brasileiro sobre Acesso Aberto (CONFOA)
	10 anos de publicação da BOAI e Boicote à Elsevier

2012	Lançado o projeto SciELO Books
------	--------------------------------

Fonte: elaborado com base em [Fausto \(2013, online\)](#), Cintra (2017, p. 45) e [Costa e Leite \(2017, p. 45\)](#).

Consoante [Rodrigues \(2004, p. 26\)](#), a origem do acesso aberto está estabelecida nos “[...] problemas, limitações e contradições do sistema de comunicação da ciência, em particular os relacionados com as revistas científicas”. O modelo tradicional de publicação e distribuição dos periódicos custeado principalmente pelas taxas de assinatura e de acesso aos artigos oriundos de investigações científicas gera aos atores envolvidos no contexto da comunicação científica, como universidades, institutos de pesquisa e agências de fomento, uma espécie de duplicação dos custos e investimentos relacionados ao financiamento dos estudos (SARMENTO *et al.*, 2005), visto que são investidos recursos para o custeio das pesquisas e, posteriormente, para ter acesso às publicações provenientes das descobertas efetuadas. Assim, uma vez que os resultados das pesquisas desenvolvidas por meio do financiamento público são cedidos às editoras comerciais de forma gratuita, juntamente com os direitos de propriedade pelo conteúdo, surge uma barreira para a disponibilização da informação científica como um bem público.

O pagamento duplo, tanto pelo financiamento das investigações quanto pelo acesso aos resultados produzidos, gera o que [Kuramoto \(2006, p. 92\)](#) considera como uma “situação paradoxal”. Na perspectiva desse autor, a lógica de circulação da informação científica descrita anteriormente contraria o ponto de vista ético, pois considerando um contexto em que boa parte das pesquisas científicas é financiada com recursos públicos, “[...] o Estado, para promover o acesso àquilo que produz, é obrigado a arcar com os custos de manutenção das coleções das revistas em que são publicados os resultados de sua produção científica” ([KURAMOTO, 2006, p. 92](#)).

Um exemplo prático da problemática do duplo custeio e dos impactos negativos na manutenção do acesso à informação científica é conhecido na literatura como a crise dos periódicos científicos ([MUELLER, 2006](#); [ARBELÁEZ, 2014](#); [COSTA; LEITE, 2017](#)). Deflagrada no contexto das unidades de informação, tal crise representou o desafio enfrentado pelas bibliotecas universitárias de diversos países, incluindo os Estados Unidos e Canadá, para arcar com os custos elevados de manutenção das assinaturas dos periódicos científicos que compunham suas coleções no período que compreende as décadas de 1970 e 1980. Os orçamentos disponíveis para as bibliotecas tornaram-se incompatíveis com as taxas crescentes cobradas pelas editoras científicas, gerando descontinuidade nos acervos e dificuldades aos

pesquisadores que precisavam acessar às publicações científicas para dar andamento e embasamento aos seus estudos. A contextualização, as causas, as ações imediatas, bem como a busca por soluções para contenção da crise e as consequências decorrentes dessa conjuntura foram registradas em relatório da [Association of Research Libraries \(1989\)](#).

Diante da problemática mencionada e das possibilidades potenciais que os avanços tecnológicos e a ampliação da Internet demonstraram, soluções alternativas que auxiliassem na disseminação da produção acadêmica de forma livre e gratuita mostraram-se mais viáveis, impulsionando então a formalização do MAA. Logo, os acontecimentos do final do século XX modelaram a ideia de acesso aberto, entretanto, a consolidação do conceito ocorreu efetivamente por meio da publicação da tríade de declarações onde estão firmadas as bases do acesso aberto: [BOAI \(2002\)](#), [Bethesda Statement on Open Access publishing \(2003\)](#) e [Berlin Declaration on Open Access to knowledge in the Sciences and Humanities \(2003\)](#). A primeira, foi responsável por cunhar oficialmente o termo *open access* e registrar a definição mais significativa e relevante de acesso aberto (PIWOWAR *et al.*, 2018), entendido como:

a [...] disponibilidade gratuita [da literatura científica] na Internet, permitindo que qualquer usuário possa ler, baixar, copiar, distribuir, imprimir, buscar ou criar links para os textos completos destes artigos, rastreá-los para indexação, transferi-los como dados para software, ou usá-los para qualquer outra finalidade legal, sem barreiras financeiras, legais ou técnicas que não sejam inseparáveis do acesso à própria Internet. A única restrição à reprodução e à distribuição e o único papel dos direitos autorais neste domínio deve ser o de conferir aos autores o controle sobre a integridade do seu trabalho e o direito de serem devidamente reconhecidos e citados (BOAI, 2002, *online*, tradução nossa)¹⁷

Os aspectos evidenciados nessa definição sinalizam as características básicas do acesso aberto, às possibilidades de uso e reúso da literatura científica, a necessidade de garantia e reconhecimento dos direitos dos autores, bem como o uso da Internet como um fator de reversão na lógica de comunicação e distribuição da informação. Tão grande é a importância da BOAI para o MAA, visto que essa representa o compromisso da comunidade científica com o acesso aberto, que [Costa e Leite \(2017, p. 46\)](#) afirmam que a organização histórica do movimento em questão pode ser observada a partir da consideração de três períodos: “pré-BOAI, BOAI, e pós-

¹⁷By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited (BOAI, 2002, *online*).

BOAI”. De acordo com a BOAI, a remoção de barreiras de acesso corrobora com a aceleração da pesquisa, o enriquecimento da educação, a utilização da literatura e o compartilhamento de aprendizados, revelando assim um apelo à igualdade de acesso ao conhecimento.

Em 2003, foram publicadas as outras duas declarações componentes da tríade anteriormente mencionada. Com o intuito de ampliar e acelerar a disponibilização da literatura científica em acesso aberto, a Declaração de Bethesda fornece recomendações aos atores envolvidos no contexto de produção e disseminação do conhecimento científico (por exemplo, cientistas, bibliotecários, editores e agências de fomento à pesquisa) e fixa condições formais para se considerar uma publicação como de acesso aberto. Já a Declaração de Berlim, destaca que não disponibilizar a informação científica de forma ampla e imediata para a sociedade significa alcançar apenas metade da missão de disseminação do conhecimento. Nesse sentido, o documento enfatiza o papel fundamental da Internet para modificar a comunicação e distribuição da informação, bem como destaca questões que apontam o acesso aberto como um movimento capaz de proporcionar benefícios para a ciência e para a sociedade.

Mesmo que com objetivos distintos, as três declarações se complementam e influenciaram diretamente na consolidação e ênfase do que se compreende como acesso aberto. Assim, levando em consideração os pontos em comum entre as declarações de Budapeste, Bethesda e Berlim, Suber (2015, *online*, tradução nossa) propôs a *BBB definition*, que considera que o “acesso aberto remove as barreiras de preços (assinaturas, taxas de licenciamento, taxas de visualização pagas) e as barreiras de permissão (a maioria das restrições de *copyright* e licenciamento)”. A visualização comparada dos principais aspectos abordados pelas declarações está disposta no **Quadro 3**.

Quadro 3 - Principais aspectos das declarações de Budapeste, Bethesda e Berlim

	Declaração de Budapeste (2002)	Declaração de Bethesda (2003)	Declaração de Berlim (2003)
Área do conhecimento (origem)	Todas	Biomedicina	Ciências e Humanidades
Ênfases/Focos	-Definição formal para o conceito de acesso aberto; -Proposição de duas estratégias principais para alcance do acesso aberto.	-Fornece recomendações para atores envolvidos no contexto da publicação científica; - Estabelece condições para classificar uma publicação como de acesso aberto	- Destaca a Internet como instrumento funcional ao serviço de uma base de conhecimento científico global e do pensamento humano; -Paradigma do acesso

			aberto eletrônico; -Benefícios do acesso aberto para o contexto científico e social.
Qualidade (revisão por pares)	- Artigos de periódicos revisados; -Preprints sem revisão.	- Artigos revisados;	- Aprovação pela comunidade científica (revisado)
Direitos autorais	- Controle sobre a integridade do trabalho - Direito de propriedade intelectual assegurado aos autores - Atribuição da autoria por meio de citação	- Atribuição da autoria por meio de citação	- Atribuição da autoria por meio de citação
Utilização prevista para as publicações	-Diversas: leitura, download, cópia, distribuição, impressão, links, rastreamento, indexação; -Uso para finalidades legais	-Diversos, respeitando as finalidades legais	-Diversos, respeitando as finalidades legais
Permite obras derivadas?	Não mencionado	Sim	Sim
Acesso	-Gratuito; -Público (mundial); -Online	-Gratuito -Mundial -Irrevogável -Perpétuo (demanda preservação)	-Gratuito -Mundial -Irrevogável

Fonte: elaborado pela autora com base em [Sarmiento et al. \(2005, p. 6\)](#) e [Rios, Lucas e Amorim \(2019, p. 166\)](#).

Em consonância com o escopo do presente estudo, a compreensão da base firmada a partir das três declarações mencionadas anteriormente mostra-se suficiente para contextualizar a origem e o desenvolvimento do acesso aberto. Entretanto, vale destacar que o MAA recebeu influência de diversas outras declarações e manifestos, tanto de abrangência global, quanto nacional. A enumeração e análise das declarações que, de alguma forma, contribuíram para evolução do acesso aberto é tratada de forma mais detalhada em [Rios, Lucas e Amorim \(2019\)](#) e [Costa e Leite \(2017\)](#).

De posse de uma definição explícita de acesso aberto e suas características substanciais, era o momento do MAA dedicar esforços para ampliar e disseminar as ideias firmadas. Então, levando em consideração duas estratégias principais de ação para alcance do acesso aberto propostas pela BOAI, o autoarquivamento e os periódicos de acesso aberto, o movimento se dedicou em avançar na implementação de soluções e sistemas que permitissem a disponibilização da literatura científica livre de barreiras. Nesse processo, os modelos de negócios adotados pelas editoras passaram por adaptações e modificações, fazendo surgir

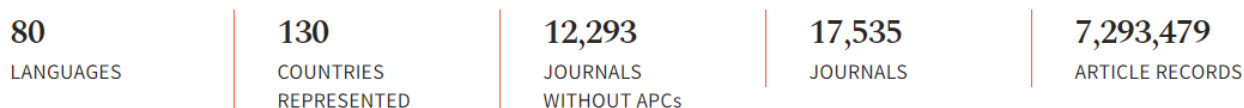
variados subtipos e categorias de acesso aberto (APPEL, 2019; PIWOWAR *et al.*, 2018). Análise e detalhamento dessas novas tipologias encontra-se delineado na próxima subseção.

2.2.1 Os subtipos de acesso aberto

Aproximadamente duas décadas após o marco de formalização do MAA por meio da BOAI, observa-se uma tendência de ampliação e adesão ao movimento. Nesse período, avanços e possibilidades surgiram. Em termos quantitativos, o movimento apresenta expressivos volumes. O [Directory of Open Access Journals \(DOAJ\)](#), base de dados de referência quando se trata da curadoria, manutenção e fornecimento de informações confiáveis relativas aos periódicos científicos de acesso aberto, registrou, em março de 2022, valores superiores a 17 mil periódicos e 7 milhões de artigos em acesso aberto (**Figura 4**). Dentre as publicações periódicas indexadas no DOAJ, mais de 12 mil são caracterizadas como periódicos científicos que não cobram *Article Processing Charge* (APC), ou seja, taxas relacionadas ao processamento e publicação dos artigos.

A tendência de crescimento percebida no DOAJ é também confirmada no âmbito dos repositórios de acesso aberto, por meio da consulta ao [Directory of Open Access Repositories \(OpenDOAR\)](#)¹⁸, serviço responsável por hospedar e revisar registros de repositórios que disponibilizam recursos e resultados científicos gratuitamente e em acesso aberto. No mês de março de 2022, as estatísticas do referido diretório explicitam um total de 5841 repositórios em acesso aberto registrados. A evidência de crescimento é percebida quando esse valor é comparado aos 78 repositórios presentes no OpenDOAR em 2005, ano de seu lançamento (**Gráfico 2**).

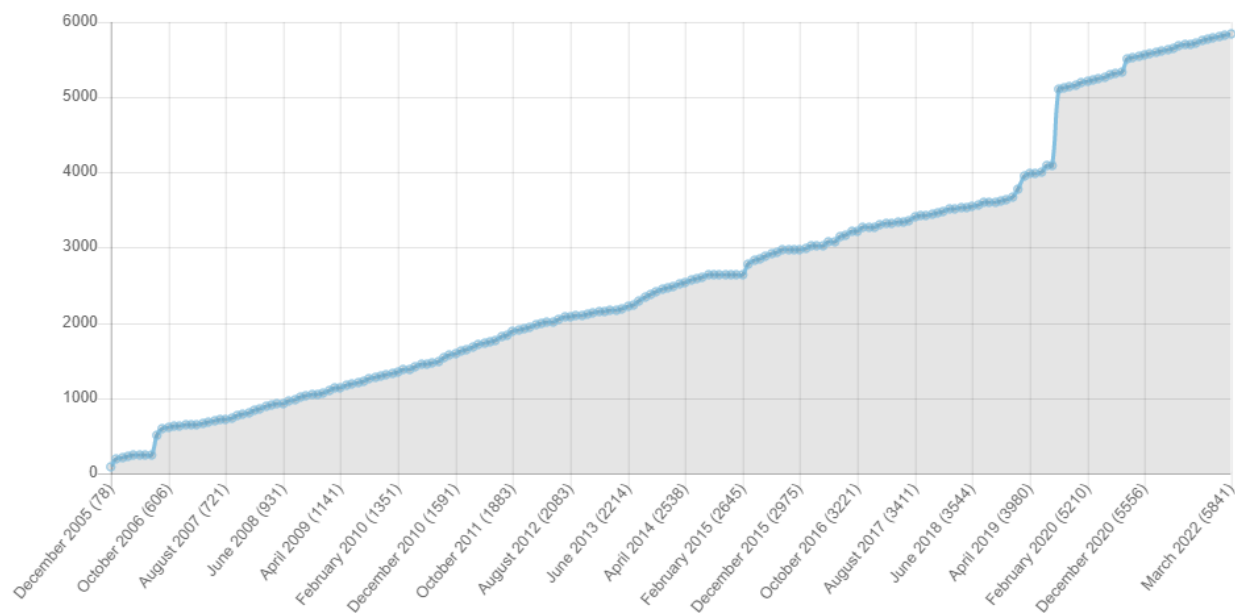
Figura 4 - Quantitativo de periódicos científicos e artigos em acesso aberto registrados no DOAJ (março 2022)



Fonte: Dados disponibilizados por <https://doaj.org/> em março de 2022.

¹⁸ Disponível em: <https://v2.sherpa.ac.uk/opensoar/about.html>. Acesso em: 17 ago. 2021.

Gráfico 2 - Crescimento do número de repositórios de acesso aberto registrados no OpenDOAR (dezembro de 2005 a março de 2022)



Fonte: Dados do [OpenDOAR](#) (2022).

O *autoarquivamento*, por meio do depósito em repositórios digitais de acesso aberto, e os *periódicos científicos de acesso aberto* foram as duas principais estratégias complementares propostas pela BOAI (2002) para o alcance do objetivo de disponibilização da literatura científica sem barreiras de acesso e custos. Tais estratégias foram, posteriormente, intituladas por [Harnad et al. \(2008\)](#) como *green road* (via verde) e *gold road* (via dourada). A primeira, representando a escolha de utilizar um repositório de acesso aberto para realizar o autoarquivamento, enquanto que a segunda via destaca a possibilidade de publicar artigos em periódicos científicos de acesso aberto. A terminologia proposta por [Harnad et al. \(2008\)](#) foi amplamente aceita pela literatura da área e constituiu o primeiro passo para o desenvolvimento do código de cores que simboliza as ramificações e os subtipos de acesso que eclodiram com o desenvolvimento do MAA.

O texto da BOAI (2002) destaca que

Enquanto sugerimos as duas estratégias já mencionadas, também encorajamos a experimentação de outras maneiras para realizar a transição dos métodos atuais de disseminação para o acesso aberto. **Flexibilidade, experimentação e adaptação** para as circunstâncias locais são os melhores caminhos para garantir que o progresso em diversos ambientes seja rápido, seguro e duradouro (BOAI, 2002, *online*, tradução nossa, grifo nosso).

E o trecho supracitado corrobora para o entendimento de que o acesso aberto tem como característica a fluidez e favorece o surgimento de novas estratégias em busca da forma mais viável para assegurar a literatura científica livre para leitura e para reuso. Nesse sentido, [White et al. \(2021\)](#) apontam três aspectos, ocorridos no início dos anos 2000, que influenciaram o estabelecimento do cenário contemporâneo do acesso aberto marcado pela pluralidade de subtipos, são eles:

- O crescimento de periódicos da via dourada, financiados por APC em vez de assinaturas;
- A adoção de opções de acesso aberto híbrido por periódicos de assinatura, tornando os artigos individuais abertamente acessíveis através do pagamento de APC.
- O crescimento de repositórios institucionais e temáticos, proporcionando uma via alternativa de acesso aberto verde aos artigos individuais sem taxas de publicação ([WHITE et al., 2021, p. 2](#))

Consoante [Appel e Albagli \(2019\)](#), as APC surgem como uma alternativa oriunda das adaptações nos fluxos editoriais e dos modelos de negócio adotados pelas editoras científicas para a disponibilização de publicações periódicas em acesso aberto. Tennant *et al.* (2020) colocam que o processamento de artigos envolve etapas de edição, revisão, composição, infraestrutura de serviços, hospedagem do periódico, e divulgação das publicações, e esses procedimentos influenciam os custos de produção do periódico, que tendem a variar a depender da editora e se desdobram nos valores de APC cobrados.

As informações relativas aos custos de produção e processamento de artigos não costumam ser explicitadas pelos editores, colocando sob questionamento o aspecto de transparência e clareza quanto aos valores definidos como APC (TENNANT *et al.*, 2020). Com o passar do tempo, a elevação dos preços cobrados como APC chamaram atenção da comunidade acadêmica, que deu início às reflexões associadas aos possíveis empecilhos gerados pelas APC para a livre circulação da informação científica. Na perspectiva de [White et al. \(2021\)](#), a cobrança de APC adicionou um novo tipo de despesa ao contexto do acesso aberto, dificultando a redução de custos almejada pelo movimento.

Além de compreendidas como modelo de negócio criado pelas editoras comerciais na tentativa de se adequarem às prerrogativas defendidas pelo MAA, [Appel e Albagli \(2019\)](#) também esclarecem que as APC também podem ser analisadas como uma nova barreira que dificulta o desenvolvimento de periódicos de acesso aberto, visto que as taxas acabam por privilegiar regiões geográficas com mais recursos financeiros destinados à pesquisa e à

produção científica. Ao levar esse ponto em consideração, nota-se que cobrança de APC, de certa forma, transfere as dificuldades enfrentadas pelos leitores para acessarem a literatura científica já publicada, para problemáticas encaradas pelos autores na etapa de publicação dos artigos, fase em que as APC são cobradas para disponibilizar os artigos em acesso aberto. Isto posto, a depender, por exemplo, da região geográfica, os elevados valores de APC podem representar um gatilho para uma crise de publicação de resultados científicos em revistas que adotam APC como modelo de negócio.

Piwowar *et al.* (2018) argumentam que o modelo de negócio de acesso aberto híbrido compartilha características com o dourado, visto que, em ambos subtipos, os artigos encontram-se abertos para leitura e licenciados de forma a permitir o reuso. Todavia, esses subtipos podem se diferenciar, por exemplo, quanto ao local de publicação, visto que os artigos dourados são publicados em periódicos considerados totalmente como acesso aberto, enquanto que os artigos híbridos podem ser publicados em periódicos comerciais que cobram taxas para disponibilizar o conteúdo do artigo gratuitamente para os leitores da revista. Os artigos do tipo dourado são publicados necessariamente em periódicos científicos totalmente de acesso aberto, ou seja, revistas que, além de disponibilizarem os textos completos de todas as publicações *online*, também deixam claras aos usuários as licenças de acesso e uso do conteúdo disponível. Além disso, os periódicos classificados como dourados podem ser validados pelo DOAJ, enquanto os de acesso híbrido não, visto que o DOAJ só indexa periódicos que apresentem características de acesso aberto livre em detrimento do acesso aberto entendido como gratuito.

Para [Suber \(2008\)](#) e [DOAJ \(2020\)](#), o avanço e a aceitação do conceito de acesso aberto fez com que esse passasse a compreender duas ramificações: o acesso aberto livre (*Libre Open Access*) e o acesso aberto gratuito (*Gratis Open Access*). O primeiro remove barreiras de preço e algumas barreiras de permissão, visto que se caracteriza por manter o conteúdo digital completo da publicação acessível gratuitamente para os usuários e define, de forma explícita, as questões associadas aos direitos autorais e as licenças de uso e reuso. Já o segundo, não apresenta esclarecimentos explícitos quanto às licenças, rompendo apenas com barreiras de preço ao disponibilizar os conteúdos das publicações gratuitamente para leitura.

Outra característica dos periódicos híbridos é que os artigos por eles publicados podem não ficar disponíveis para acesso no ato da publicação, tendo que aguardar o pagamento de taxa, pelo autor ou instituição vinculada a ele, para liberação do conteúdo gratuitamente ao leitor (PIWOWAR *et al.*, 2018). Nesse sentido, a análise dos prós e contras vinculados ao modelo de acesso híbrido levanta na literatura de acesso aberto a discussão sobre a problemática do *double*

dipping, caracterizado como o pagamento duplo por um mesmo artigo, em que, primeiro os autores ou instituições de financiamento pagam taxas para publicar (APC) e, depois, as universidades e/ou instituições de pesquisa custeiam as taxas para acessar a publicação (MITTERMAIER, 2015). Pelas características, Björk (2012) posiciona o acesso híbrido entre as publicações tradicionais com o modelo de negócio baseado na cobrança de assinaturas e os periódicos científicos de acesso aberto do tipo dourado. Tal caráter intermediário desse tipo de acesso gera questionamentos e dúvidas associadas ao entendimento do acesso híbrido como representante viável do acesso aberto (MITTERMAIER, 2015). Entretanto, faz-se válido ressaltar a ausência de um consenso quanto à aceitação ou rejeição do acesso híbrido enquanto acesso aberto. Para os fins da presente pesquisa e tendo em vista o entendimento de Piwowar *et al.* (2018), considera-se o acesso aberto híbrido como um subtipo de acesso aberto.

O último ponto enumerado por White *et al.* (2021) e mencionado anteriormente como influente para definir o contexto atual do MAA é o aumento no número de repositórios temáticos e institucionais. Tal crescimento é também demonstrado por meio do **Gráfico 2** e representa a expansão da via verde como uma possibilidade de disponibilização das publicações em acesso aberto sem cobrança de taxas como as APC.

Harnad (2010) enfatiza a necessidade de maior atenção ao acesso aberto verde em detrimento do dourado, argumentando que, diferentemente da estratégia baseada nos periódicos que se vincula às expectativas das editoras, a via verde se relaciona diretamente com os interesses dos fornecedores dos resultados das investigações científicas, os pesquisadores, que almejam a divulgação de seus trabalhos. O autor também afirma que a via verde pode ser adotada e expandida por meio de instrumentos mandatórios provenientes de instituições de fomento à pesquisa, universidades e sociedades científicas, por exemplo. Enquanto que o mesmo não se mostra viável para as publicações periódicas, que atendem às determinações dos editores.

Os anos subsequentes à proposição inicial das duas principais estratégias para evolução do acesso aberto, representadas pelo autoarquivamento e os periódicos de acesso aberto, foram marcados pela busca pela viabilização dos objetivos do MAA. Nesse processo adaptativo, surgiram alguns subtipos complementares de acesso aberto, ampliando o espectro de possibilidades ora concentrado no acesso verde e dourado. Archambault *et al.* (2014) coloca que a disponibilização aberta de artigos em periódicos científicos envolve grande diversidade de

[...] formatos (e.g. *preprint*, HTML, PDF), difundidos por meio de uma ampla gama de mídias (e.g. repositórios institucionais, sites de editoras, agregadores de conteúdo)

e vários níveis de disponibilidade (imediate, atrasada, transitória), que fez com que as definições básicas do acesso aberto, verde e dourado, se tornassem insuficientes ([ARCHAMBAULT et al., 2014, p. 2, tradução nossa](#)).

Dentre as alternativas de classificação propostas, Piwowar *et al.* (2018) e Langham-Putrow, Bakker e Riegelman (2021) esclarecem que nem todas foram aceitas de modo consensual no contexto do MAA, fazendo com que algumas categorizações de subtipos fossem mais disseminadas que outras.

Conforme [Taga \(2016\)](#), a classificação dos subtipos de acesso aberto pode ser estipulada com relação a três pontos: (1) o método ou canal adotado; (2) o modelo de negócio definido pelo periódico; (3) as licenças declaradas para uso. O autor ainda exemplifica possibilidades para cada um dos pontos elencados. No que se refere ao canal, estão os subtipos verde e dourado. Vinculado ao segundo aspecto, [Taga \(2016\)](#) menciona o acesso aberto imediato ao conteúdo completo do periódico (*full journal immediate OA*); o híbrido (*hybrid OA*); e o atrasado (*delayed OA*) por meio do cumprimento do período de embargo. Por fim, quanto às licenças, tem-se o acesso aberto livre (*libre open access*) e o gratuito (*gratis OA*) ([SUBER, 2008](#)). De modo complementar, Piwowar *et al.* (2018) afirmam que o subtipo livre encontra-se mais próximo da definição de acesso aberto apresentada pela BOAI (2002), pois rompe com as barreiras de preço e permissão.

A expansão de subtipos de acesso aberto adicionou à interpretação sobre as possibilidades existentes certo grau de complexidade. Enquanto alguns subtipos compartilham fortes similaridades entre si, outros demonstram consideráveis diferenças, de modo que a comparação entre as definições de cada um auxilia na identificação das distinções e aproximações. Isto posto, a seguir, encontra-se exposto o **Quadro 4**, que sistematiza informações significativas das principais categorizações de subtipos relacionados ao acesso aberto.

Quadro 4 - Categorização de subtipos relacionados ao acesso aberto

Subtipos	Definição	Cobrança de APC?	Licença explícita?	Fontes consultadas
Verde <i>Green</i>	Publicação acessível legalmente e de forma gratuita em um repositório de acesso aberto .	✗	✓	Harnad (2010)
Diamante/ Platina <i>Diamond/ Platinum</i>	Conteúdo publicado imediatamente em um periódico de acesso aberto que não cobra APC , mantido por meio de esforço e financiamento oriundos, por exemplo, de organizações sem fins lucrativos e não comerciais, como universidades, institutos de pesquisa e sociedades científicas.	✗	✓	Appel (2019) ; Fuchs e Sandoval (2013) ; Langham-Putrow; Bakker; Riegelman (2021) White et al. (2021)
Dourado <i>Gold</i>	Artigos imediatamente publicados em periódicos disponibilizados <i>online</i> e completamente em acesso aberto . O reconhecimento de um periódico científico como de acesso aberto dourado tende a ocorrer mediante indexação desse no DOAJ. Neste subtipo, a cobrança de APC representa uma forma de viabilizar os processos editoriais e de publicação.	✓	✓	Harnad (2010) Piwowar et al. (2018) White et al. (2021)
Híbrido <i>Hybrid</i>	Após o pagamento de APC, o artigo é disponibilizado individualmente em acesso aberto em periódicos originalmente caracterizados pela cobrança de assinatura. No acesso híbrido, o pagamento de APC representa a opção fornecida pela editora comercial para publicar um artigo específico em acesso aberto.	✓ (artigo individual)	✓	Björk (2012) Taga (2016) White et al. (2021)
Bronze <i>Bronze</i>	Publicações gratuitamente acessíveis através do <i>site</i> do periódico/editora, mas sem a definição explícita sobre as licenças de compartilhamento e uso de conteúdo.	?	✗	Piwowar et al. (2018)

Bronze imediato	Publicações do subtipo bronze, mas disponibilizadas imediatamente. Não há período de embargo.	?	×	Santos-d'Amorim (2021)
Bronze atrasado	Publicações do subtipo bronze, mas disponibilizadas após o vencimento do período de embargo.	?	×	Santos-d'Amorim (2021)
Cinza <i>Gray</i>	Disponibilização de publicações e conteúdos, pelos autores, em perfis de redes sociais acadêmicas (e.g. ResearchGate, Academia.edu e Mendeley) ou sites pessoais , sem a observação atenta de questões relacionadas às licenças e às transferências de direitos autorais.	×	×	Björk (2017) Langham-Putrow; Bakker; Riegelman, (2021)
Preto <i>Black</i>	Publicações compartilhadas ilegalmente em sites piratas , como Sci-Hub e LibGen. Segundo Piwowar <i>et al.</i> (2018), sites como os mencionados destoam da proposta de acesso aberto defendida pela BOAI.	×	×	Björk (2017) Langham-Putrow; Bakker; Riegelman, (2021) Piwowar et al. (2018)

Fonte: elaborado pela autora

Legenda: ✓ Sim; × Não; ? Informação não identificada

Com base nas fontes consultadas e de modo comparativo, o **Quadro 4** evidencia e sumariza informações essenciais para a caracterização dos subtipos de acesso aberto mencionados de modo mais evidente na literatura sobre a temática em questão. Além das tipologias supramencionadas e descritas, ainda é possível identificar outras ramificações associadas aos subtipos exemplificados. Taga (2016, p. 51, grifo do autor), por exemplo, aponta que “o acesso aberto dourado pode ser particularizado em cinco tipos: *full journal immediate open access*, *platinum open access*, *hybrid open access*, *delayed open access* e *promotional open access*”. Faz-se válido ressaltar que a interpretação sobre os vínculos de subordinação e/ou relação entre os subtipos não é um consenso entre os estudos sobre acesso aberto. As variadas possibilidades foram observadas, ao longo do tempo, a partir de diferentes lentes. Nesse sentido, no contexto do presente trabalho, são consideradas como definições norteadoras para a abordagem dos subtipos de acesso aberto as destacadas no **Quadro 4**.

O subtipo bronze demonstra um desafio latente no contexto do acesso aberto: a identificação simples e explícita das licenças de compartilhamento e uso de conteúdos publicados. Diante de tal problemática, foram desenvolvidos serviços e ferramentas *online* que funcionam como diretórios de políticas editoriais de periódicos de acesso aberto. Esses, a partir dos dados fornecidos por meio dos cadastros feitos pelos editores, buscam agregar, organizar e disseminar informações sobre as permissões de uso e autoarquivamento dos artigos. Dentre os serviços existentes, o [Sherpa Romeo](https://v2.sherpa.ac.uk/romeo/)¹⁹ é o pioneiro e com maior abrangência quando comparado a outras ferramentas equivalentes, como o [Diadorim](https://diadorim.ibict.br/)²⁰ e o [Dulcinea](https://www.accesoabierto.net/dulcinea/)²¹.

Em contraponto ao acesso aberto e seus subtipos, está o acesso fechado (*closed*), o qual engloba publicações que não se adequam em nenhuma possibilidade de acesso aberto, como os artigos publicados em periódicos que adotam o modelo de negócio tradicional, o qual fornece o conteúdo científico apenas mediante pagamento de taxas de assinatura (PIWOWAR *et al.*, 2018). [Archambault et al. \(2014\)](#) definem também o acesso aberto restrito (*restricted OA*), entendido como gratuito, mas com restrições para *download* e reuso. A questão da restrição também pode estar associada a liberação de acesso apenas para membros de grupos específicos.

Do ponto de vista de ferramentas e bases de dados que indexam publicações científicas e atribuem aos documentos uma classificação quanto aos tipos de acesso, percebe-se a adoção

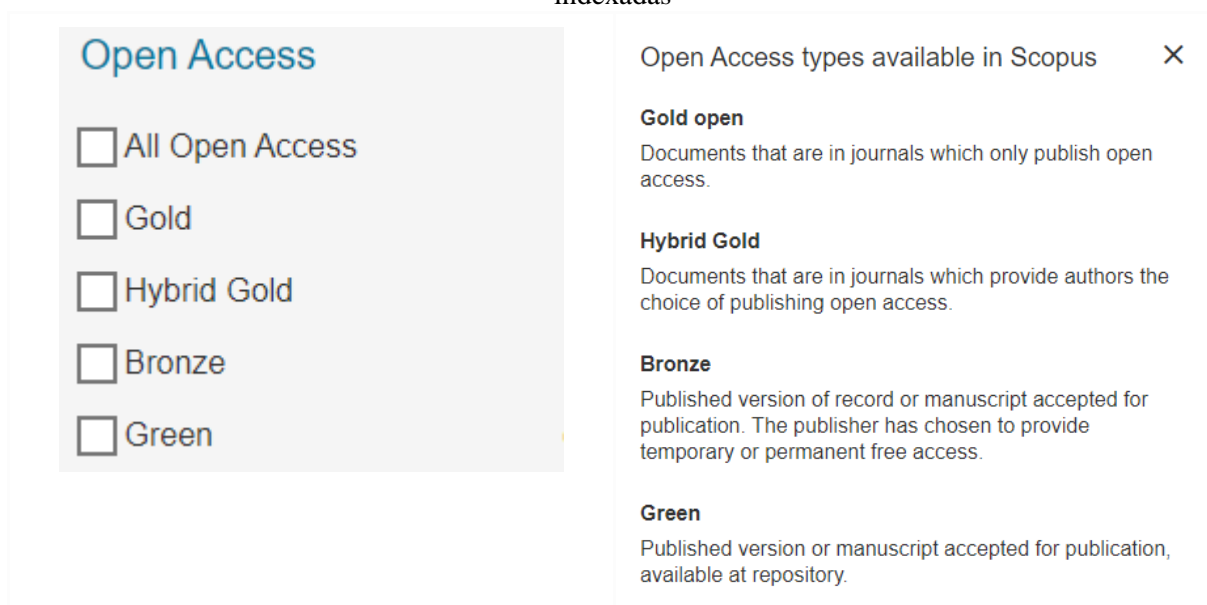
¹⁹ Disponível em: <https://v2.sherpa.ac.uk/romeo/>. Acesso em: 12 out. 2021.

²⁰ Disponível em: <https://diadorim.ibict.br/>. Acesso em: 12 out. 2021.

²¹ Disponível em: <https://www.accesoabierto.net/dulcinea/>. Acesso em: 12 out. 2021.

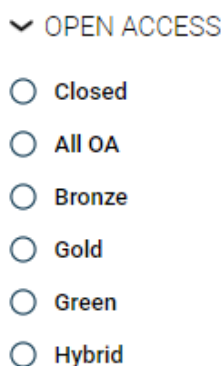
de um espectro de subtipos mais resumido do que o apresentado pelo **Quadro 4**. Ao definirem o tipo de acesso de uma publicação, [Unpaywall](https://unpaywall.org/)²², [Scopus](https://www.scopus.com/)²³ e [Dimensions](https://www.dimensions.ai/)²⁴ consideram como possibilidades os seguintes subtipos: dourado, híbrido, bronze e verde, conforme demonstram as **Figuras 5 e 6**, que representam o filtro “*open access*” disponível em ambas bases com as respectivas opções para seleção.

Figura 5 - Subtipos utilizados pela Scopus para classificação de acesso aberto em publicações indexadas



Fonte: <https://www.scopus.com/> em outubro de 2021.

Figura 6 - Subtipos utilizados pela Dimensions para classificação de acesso aberto em publicações indexadas



Fonte: <https://www.dimensions.ai/> em outubro de 2021.

²² Disponível em: <https://unpaywall.org/>. Acesso em: 12 ago. 2021.

²³ Disponível em: <https://www.scopus.com/>. Acesso em: 12 ago. 2021.

²⁴ Disponível em: <https://www.dimensions.ai/>. Acesso em: 12 ago. 2021.

Diante de tais possibilidades de classificação de subtipos de acesso aberto, estudos como Van Leeuwen, Costas e Robinson-Garcia (2019), Robinson-Garcia, Costas e Van Leeuwen (2020) e Piwowar *et al.* (2018) utilizaram a ferramenta Unpaywall para coletar dados relativos à classificação dos tipos de acesso registrados pelas publicações. Diante de tais pesquisas, nota-se a possibilidade de análise da variável acesso nas publicações para além do sistema binário baseado nas categorias acesso aberto ou fechado. Entretanto, tais autores destacam limitações e desafios associados ao entendimento dos processos de classificação dos tipos de acesso utilizados pela ferramenta adotada para coleta dos dados, com ênfase nos subtipos bronze e híbrido.

Na percepção de Van Leeuwen, Costas e Robinson-Garcia (2019), o Unpaywall efetua a classificação automática de um artigo de acesso aberto como bronze quando não é rastreada evidência clara sobre a licença de uso declarada pela editora responsável pela publicação. Logo, mesmo que uma publicação tenha uma licença registrada, mas por empecilhos técnicos essa não possa ser identificada, poderá ocorrer uma classificação inadequada do tipo de acesso da publicação, fazendo com que um artigo originalmente do subtipo dourado, por exemplo, seja classificado como bronze. Os fluxos que representam a lógica utilizada pelo Unpaywall para atribuir automaticamente os subtipos de acesso às publicações encontram-se detalhados nas **Figuras 13 e 14** deste trabalho, as quais estão localizadas na seção **4.2.2**, que expõe os resultados e as discussões associadas ao **OE2** da presente pesquisa.

A contextualização construída por meio das informações apresentadas anteriormente evidenciam o desafio e a amplitude de possibilidades associadas ao estudo dos subtipos e alternativas de acesso aberto. As definições colocadas na presente seção tem por intuito orientar as investigações e análises elaboradas por esta pesquisa no que tange ao acesso aberto em publicações na temática da COVID-19. A seguir, a fim de complementar o entendimento sobre o MAA, encontram-se enumerados argumentos presentes na literatura como justificativas que advogam a favor do acesso aberto

2.2.2 Argumentos a favor do acesso aberto

Inicialmente, as publicações de acesso aberto eram questionadas sob a perspectiva da autenticidade e qualidade. Após esforços e iniciativas coletivas, incluindo algumas das pontuadas no **Quadro 2**, a visão em relação ao MAA se transformou e passou a demonstrar elementos de contribuição quanto a promoção de visibilidade; presença global; acessibilidade; aumento do impacto em citações e aplicações; elevação na colaboração; facilidade de *feedbacks*, comentários e reflexões críticas sobre publicações disponibilizadas em acesso aberto ([KANJILAL; KUMAR DAS, 2015](#)).

[Furnival e Hubbard \(2011\)](#) evidenciam o papel de importância das declarações, manifestos e políticas de acesso aberto como formas de conscientização e instrumentos para disseminação de apoio e incremento da credibilidade das ideias levantadas pelo MAA de modo oficial e internacional. Nesse sentido, retoma-se ao texto da BOAI (2002) para destacar um trecho que sinaliza argumentos favoráveis sobre a viabilidade e desdobramentos positivos para adoção ao movimento: “[...] o acesso aberto é economicamente viável, dá aos leitores um poder extraordinário para encontrar e fazer *uso* da literatura relevante e fornece aos autores e seus trabalhos vastas e *mensuráveis* novas formas de *visibilidade, leitura e impacto*” (BOAI, 2002, *online*, tradução nossa, grifo nosso).²⁵

Ao tratar sobre os benefícios do acesso aberto e considerando as características destacadas pela BOAI (2002), Swan (2012) interpreta os pontos positivos grifados anteriormente de modo interligado e concatenado, expondo que ao maximizar a visibilidade das publicações científicas, o acesso aberto tende a ampliar e permitir novas possibilidades de uso dos resultados de investigações e, conseqüentemente, influencia no impacto promovido pela pesquisa tanto no contexto acadêmico e quanto além das fronteiras da academia.

Mostra-se evidente que o acesso aberto não consiste na única solução para todas as questões vinculadas à acessibilidade de publicações científicas, havendo, por exemplo, barreiras linguísticas, de conectividade, entre outras. Todavia, é válido o reconhecimento da potencialidade do acesso aberto para contribuir com a melhoria nas formas de se ter contato com os resultados científicos, corroborando também com o surgimento de inovações voltadas para a remoção de barreiras que impedem a comunicação e disseminação da ciência (TENNANT *et al.*, 2020).

²⁵ “[...] open access is economically feasible, that it gives readers extraordinary power to find and make use of relevant literature, and that it gives authors and their works vast and measurable new visibility, readership, and impact” (BOAI, 2002, *online*).

[Hrynaszkiewicz \(2016\)](#), ao investigar questões voltadas à sustentabilidade dos modelos de negócio de periódicos de acesso da área de ciências sociais e políticas, estruturou sete exemplos de pontos positivos, abordados como benefícios do acesso aberto que estão especificados no **Quadro 5** a seguir.

Quadro 5 - Benefícios do acesso aberto

	Benefício	Exemplo/contexto
1	Mais rápido e eficiente progresso na pesquisa	Os pesquisadores podem se concentrar na pesquisa sem perder tempo e se preocupar em obter acesso ao conhecimento. As plataformas <i>online</i> de acesso aberto geralmente podem coordenar a revisão por pares e o processo de publicação mais rapidamente, sem restrições de impressão.
2	Aumento na visibilidade e compreensão da pesquisa também fora da comunidade acadêmica.	Há um forte incentivo para que a pesquisa em ciências políticas e sociais - pesquisas que devem afetar a política governamental baseada em evidências - seja acessível fora da academia.
3	Melhoria no retorno do investimento em pesquisa	Permitir visibilidade máxima e reutilização perpétua por meio do acesso aberto oferece melhor retorno do investimento em pesquisa. Para cada \$1 investido no Projeto Genoma Humano - um exemplo de acesso aberto às informações de pesquisa - \$16 foram devolvidos (GRUEBER; AND TRIPP, 2011 apud HRYNASZKIEWICZ, 2016 , p. 192).
4	Aumento do impacto da pesquisa, mensurado por meio da bibliometria	Numerosos estudos mostraram um aumento no número de leitores e citações de artigos de acesso aberto em comparação com artigos semelhantes por trás das barreiras de assinatura.
5	Facilitação de pesquisa e colaboração interdisciplinar	Perguntas de pesquisa entre disciplinas, por exemplo, incluindo elementos de economia, sociologia e saúde pública, são cada vez mais necessárias. As informações de pesquisa podem ser acessadas independentemente do campo usual de investigação do pesquisador, estimulando novas hipóteses interdisciplinares.
6	A literatura de pesquisa torna-se um recurso para pesquisas futuras	A pesquisa em mineração de texto e dados é habilitada quando o acesso aberto a artigos de periódicos permite a reutilização irrestrita da literatura, por humanos e máquinas (ROBERTS, 2001 apud HRYNASZKIEWICZ, 2016 , p. 192).
7	Condução/Impulsão da inovação na comunicação científica	A publicação em acesso aberto nasceu da web e está intimamente ligada à inovação, por exemplo, em bibliometria, publicação e integração de dados e esforços para melhorar a reprodutibilidade científica.

Fonte: [Hrynaszkiewicz](#) (2016, p. 192, tradução nossa) com adaptações.

As informações destacadas pelo **Quadro 5** demonstram aspectos oportunizados por meio do acesso aberto. Argumentos como progresso para a ciência, aceleração da pesquisa, melhor aproveitamento de recursos investidos em estudos e investigações são trazidos de forma complementar aos tópicos inicialmente mencionados nesta seção. Hrynaszkiewicz (2016)

reforça a visibilidade e o impacto como benefícios do acesso aberto, apontando o uso dos métodos bibliométricos como opção para viabilizar a mensuração de tais aspectos.

No que tange a observação dos impactos provocados pelo acesso aberto, observa-se que esse gera efeitos no contexto acadêmico que se refletem de duas principais maneiras: “(i) através da associação com um maior impacto registrado para artigos científicos, como resultado da disponibilidade e reutilização, e (ii) por meio da possibilidade de permitir, de forma não restritiva, que os pesquisadores usem ferramentas para explorar a literatura acadêmica” (TENNANT *et al.*, 2020, p. 5, tradução nossa). Para além da academia, Tennant *et al.* (2020) consideram que o acesso aberto também produz impacto social, visto que corrobora para que as descobertas se tornem acessíveis potencialmente por qualquer pessoa e ou instituição com conexão à Internet e que detenha habilidades para encontrar e ler as informações publicadas. Nesse sentido, os autores ainda afirmam que o acesso aberto estimula a manutenção do processo de aprendizagem contínua e favorece a inovação e a criatividade.

Os argumentos a favor do acesso aberto apresentados anteriormente estimulam a reflexão a respeito da existência de diferenças de impacto e uso associados às publicações classificadas como de acesso aberto em comparação com as de acesso fechado. Tal questão impulsionou, na literatura, o surgimento da hipótese que sinaliza factíveis vantagens de citação de acesso aberto, conhecida como OACA. A contextualização sobre essas vantagens, seus desdobramentos e tipos de mensuração e análise por meio de indicadores bibliométricos são tratados a seguir.

2.3 INDICADORES PARA MENSURAR AS VANTAGENS DO ACESSO ABERTO

“A disponibilidade online e gratuita aumenta substancialmente o impacto de um periódico” (em inglês, *free online availability substantially increases a paper's impact*) é a afirmação que intitula o artigo de [Lawrence \(2001, tradução nossa\)](#), estudo publicado na área de Ciência da Computação e considerado como a precursor da ideia do que atualmente são tratadas pela literatura como vantagens de citação de acesso aberto. Diante da análise das taxas de citação de 119.924 artigos, a investigação identificou que a disponibilização *online* e gratuita influenciava na maximização do impacto (interpretado pelo autor como número de citações acumuladas) das publicações. De posse desses resultados, [Lawrence \(2001\)](#) recomenda aos autores e editores que facilitem o acesso aos resultados de pesquisa, mencionando como justificativa a geração de benefícios para a ciência e a sociedade.

À época da publicação do artigo de [Lawrence \(2001\)](#), o conceito de acesso aberto ainda não havia sido firmado, fato que ocorreu pouco tempo depois pela publicação da BOAI em fevereiro de 2002. Com a eclosão do MAA e a disseminação dos argumentos a favor do movimento (abordados na seção **2.2.2**), com ênfase na lógica de que “[...] dá visibilidade deriva o uso, e do uso deriva o impacto” (SWAN, 2012, p. 28, tradução nossa), ocorreu a consolidação da hipótese OACA, que considera que publicações em acesso aberto tendem a registrar um número total de citações mais elevado do que as não disponibilizadas *online* e gratuitamente sem barreiras de acesso.

[Langham-Putrow, Bakker e Riegelman \(2021\)](#) esclarecem que a nomenclatura “vantagens de citação de acesso aberto” é consistente e efetivamente utilizada na literatura, sendo empregada inclusive em estudos em que os resultados sinalizam possíveis “desvantagens” ao reportarem evidências negativas ou inconclusivas quanto às investigações da hipótese de OACA.

O termo “vantagem” conduz a significação de algo que encontra-se em posição de privilégio ou superioridade em relação ao referencial a que é comparado ([MICHAELIS, c2015](#)). Consoante Cunha e Cavalcanti (2008, p. 375), vantagem cumulativa considera a “ideia contida nas leis de Bradford e Lotka de que o sucesso alimenta mais sucesso. P. ex., um autor muito citado ou que venda muitos livros torna-se mais conhecido e, muitas vezes, dominante em relação a outros”. Ambas leis mencionadas são exemplos de leis bibliométricas, importantes para firmar as bases de uma das abordagens mais tradicionais dos estudos métricos, a

Bibliometria. Por conseguinte, nota-se, de certa forma, a aproximação dos métodos bibliométricos para avaliação de questões de vantagem.

Noronha e Maricato (2008) colocam que, além das investigações baseadas nas proposições das leis bibliométricas, os estudos bibliométricos envolvem a análise de citação, que considera dados os quais permitem investigações associadas ao impacto, tipos de documentos citados, análise da literatura científica (vida média, obsolescência), entre outras.

Segundo [Kanjlal e Kumar Das \(2015\)](#), existem quatro dimensões de avaliação da pesquisa: produtividade, reputação, visibilidade e impacto. O aumento das duas últimas pode ser visto como um tipo de desdobramento causado pela abertura das publicações científicas. Por se relacionarem com a avaliação dos produtos das investigações, observa-se que a mensuração de tais aspectos pode ser viabilizada por meio da aplicação de métodos existentes no âmbito dos estudos métricos da informação, que visam avaliar, utilizando uma abordagem quantitativa, os insumos e as questões vinculadas à comunicação e produção científica. Em decorrência desse tipo de estudos, faz-se possível a geração de indicadores que sinalizam as tendências de pesquisas (NORONHA; MARICATO, 2008).

Assim, com um olhar mais quantitativo, os indicadores corroboram com o entendimento e a observação de dinâmicas e aspectos associados à produção científica. A partir da realização de revisão sistemática de literatura com foco nos documentos que tratam da hipótese de OACA, [Langham-Putrow, Bakker e Riegelman \(2021\)](#) identificaram os principais indicadores considerados para a investigação das vantagens de citação de acesso aberto, são eles: média de citações por artigo, mediana de citações por artigos e citações cumulativas. Dessa forma, sem a intenção de esgotar o assunto e buscando contextualizar o uso de indicadores como alternativas para investigar as hipóteses de vantagens de acesso aberto, as próximas seções tratam sobre os indicadores bibliométricos e alométricos.

Após a abordagem dos tópicos relacionados às OACA e os indicadores bibliométricos, reflete-se sobre a hipótese de vantagens alométricas de acesso aberto (OAAA), oriunda na literatura seguindo a mesma lógica considerada para as vantagens de citação, de que a disponibilização das publicações científicas em acesso aberto favorecem o aumento da atenção recebida por elas em ambientes digitais consolidados no âmbito da Web 2.0, como as mídias e redes sociais.

2.3.1 Indicadores bibliométricos de citação

“Não há ciência ou tecnologia sem medidas. E principalmente sem medidas exatas”
(LE COADIC, 2004, p. 53).

Enquanto fenômeno sociocultural, a ciência envolve aspectos complexos, a exemplo da produtividade, institucionalização, publicação, comunicação e qualidade. Os desdobramentos das atividades e resultados científicos produzem inovações e influenciam as estruturas da sociedade. Dado tal contexto, Maltrás Barba (2003) busca expressar as relações existentes entre os indicadores bibliométricos e os conceitos e elementos inseridos no contexto do sistema de publicações científicas. Assim, o autor define indicadores bibliométricos como as “[...] medidas obtidas a partir das análises estatísticas das características quantificáveis da literatura científica” (MALTRÁS BARBA, 2003, p. 13, tradução nossa) e considera o documento científico como objeto de estudo da Bibliometria, campo do conhecimento partícipe da Cientometria.

A Bibliometria, como campo interdisciplinar, considera componentes advindos da Matemática, das Ciências Sociais e Naturais, das Engenharias e das Ciências da vida, afirma Glänzel (2003). Poderia então esse campo contribuir de alguma forma com todas essas áreas? Interpretada também como metodologia que aplica a análise quantitativa para descrever questões associadas à literatura acadêmica e outros meios de comunicação (ARAÚJO, 2006), a Bibliometria pode corroborar com as mais diversas áreas ao mensurar e lidar com os processos informacionais dessas a partir da análise da produção científica, identificando lacunas e possíveis melhorias a serem reportadas às instituições e grupos de pesquisa, aos cientistas, aos países e aos periódicos, por exemplo (KANJILAL; KUMAR DAS, 2015).

Tanto Mueller e Passos (2000) quanto Glänzel (2003) evidenciam a importância e o pioneirismo do trabalho de Derek de Solla Price (1963) - *Little Science, Big Science* - para a Bibliometria e a comunicação científica, visto que esse afirma o entendimento da ciência enquanto fenômeno mensurável e considera a investigação do aumento do volume da literatura científica como um indicador relevante para o mapeamento da evolução do conhecimento. Por meio da obra de Price (1963), os estudos bibliométricos se expandiram e passaram a medir e considerar questões sobre produtividade dos pesquisadores e distribuição de periódicos (MUELLER; PASSOS, 2000).

Antes mesmo da popularização do termo “Bibliometria”, fato que, de acordo com Vanti (2002) e Momesso e Noronha (2017), ocorreu após a publicação do artigo de Pritchard (1969), o qual aborda as diferenças entre os termos “bibliografia estatística” e bibliometria, estudos

dedicados à realização de análises estatísticas voltadas ao entendimento da literatura científica já existiam. Dentre eles está o trabalho de [Gross e Gross \(1927\)](#) apontado por Glänzel (2003) como um dos primeiros a realizar uma investigação baseada na análise de citações com a finalidade de embasar a aquisição de periódicos de Química por uma biblioteca.

[Hoppen \(2014\)](#) considera que indicadores, de modo geral, são parâmetros usados durante a avaliação de atividades. Conforme perspectiva de Maltrás Barba (2003), os indicadores bibliométricos estabelecem diretamente relações com as atividades científicas e seus resultados e podem ser agrupados em três tipos: qualidade, colaboração e produção. Além desses, existem também os indicadores de impacto, que buscam analisar a repercussão obtida por uma publicação científica perante o público interessado, levando em consideração essencialmente, a análise de citações ([HOPPEN, 2014](#)).

Ao abordar a variedade de teorias bibliométricas existentes, [Araújo \(2006, p. 18\)](#) enfatiza a análise de citações como sendo “a área mais importante da bibliometria”, uma vez que viabiliza investigações centradas em identificar e descrever padrões observáveis na produção científica. As citações representam e registram as interações e elos estabelecidos entre as publicações, os pesquisadores, as instituições e as áreas de pesquisa, uma vez que a partir de um conjunto de referências, torna-se possível analisar vínculos existentes entre as pesquisas e identificar redes de cientistas participantes de uma mesma comunidade ([ARAÚJO, 2006; HOPPEN, 2014](#)). Assim, nota-se que as citações corroboram com a evolução da ciência e destacam a relevância do uso da literatura científica para o andamento das pesquisas.

A contagem de citações é um indicador de bibliométrico popular quando se trata de avaliação científica. De acordo com [Mattedi e Spiess \(2017\)](#), é papel da avaliação da atividade científica mensurar o volume, bem como o impacto da produção acadêmica a fim de mapear a dinâmica e a evolução do conhecimento de variadas áreas do conhecimento. Nesse sentido, o indicador de contagem de citações detém potencial para demonstrar o grau de utilização de uma publicação a partir da análise do quantitativo de citações recebidas.

Além do indicador de contagem, existem outros índices e indicadores bibliométricos baseados em dados de citação. Alguns deles, são elencados por [Kanjlilal e Kumar Das \(2015\)](#) e Andrés (2009):

- a) Fator de Impacto do periódico (*Journal Impact Factor*): criado com o intuito de identificar os periódicos mais relevantes em cada área de estudo, o cálculo deste indicador se baseia na média de citações recebidas por um periódico no ano corrente em razão dos artigos publicados em anos anteriores, podendo haver

variação na quantidade de anos considerados. Normalmente, o intervalo considerado é de dois anos (ANDRÉS, 2009; NASCIMENTO, 2017);

- b) Índice de imediatez do periódico (*Journal Immediacy Index*): indica a velocidade com a qual publicações são incorporadas às referências de outros trabalhos;
- c) Índice h (*H-index*): índice do pesquisador e é calculado a partir do número de artigos publicados pelo autor que registram citações maiores ou iguais a esse número;
- d) Meia-vida (*Cited Half-Life/Citing Half-Life*): trata do número de anos, tendo em vista o ano corrente, em que se pode observar 50% do total de citações recebidas pelo periódico;
- e) Autocitação (*Self-Citation*);
- f) Relação entre artigos citados vs. não citados (*Cited vs. Uncited papers ratio*);
- g) Índice de influência do artigo (*Article Influence score*);
- h) Citações por documento (*Cited per document*);
- i) Colaboração internacional (*International Collaboration*).

No contexto da comunicação científica, a afirmação *publish or perish* (publicar ou perecer) evidencia a associação existente entre a questão do produtivismo e o reconhecimento, sendo o segundo uma forma de recompensa que estimula o desenvolvimento da carreira acadêmica. Maltrás Barba (2003, p. 26) interpreta que “o conhecimento é pago com reconhecimento” e a partir deste, tem-se a evolução do prestígio e, posteriormente, a construção da reputação do pesquisador frente aos seus pares. Segundo [Pacheco et al. \(2018\)](#) a citação caracteriza-se como a maneira mais comum de representação do reconhecimento entre os membros da comunidade científica, fazendo com que essa seja utilizada como base para o cálculo de diversos indicadores, com ênfase na mensuração de impacto acadêmico.

[Harnad et al. \(2008\)](#) entende que o impacto representa uma medida de produtividade de progresso associada a uma investigação, de modo que mensura o grau em que o registro da pesquisa foi lido, utilizado, aplicado e citado. Conforme interpretação de Nascimento (2017), ao se tratar da temática de avaliação da produção científica, deve-se considerar a questão do impacto, estudado principalmente a partir das citações. Logo, percebe-se que

Em busca de medidas quantitativas de impacto científico, administradores e formuladores de políticas frequentemente recorrem aos dados de citações. Uma

variedade de medidas de impacto pode ser derivada de dados brutos de citações. No entanto, é muito comum avaliar o impacto científico em termos de taxas médias de citação de periódicos ([BOLLEN *et al.*, 2009, p. e6022](#), tradução nossa)

Uma vez que as citações se mostram relevantes para análise da produção científica e para embasar o cálculo de diversos indicadores bibliométricos, pergunta-se: quais as motivações que influenciam na ocorrência de uma citação? A fim de compreender melhor os motivos que influenciam no efetivo registro de uma citação, recorre-se aos trabalhos de [Spinak \(1996\)](#) e Glänzel (2003), os quais listam possíveis razões para que um autor cite uma publicação com base em motivos apresentados por Weinstock (1971). Algumas das motivações identificadas são destacadas a seguir:

Motivos sérios ou funcionais:

- a) Homenagem a um pioneiro no assunto estudado;
- b) Dar crédito a um trabalho relacionado;
- c) Identificar metodologias e procedimentos;
- d) Introduzir ou apresentar histórico sobre um assunto;
- e) Efetuar correções (no próprio trabalho e no trabalho de terceiros);
- f) Críticas a um trabalho anterior;
- g) Concordância e validação do conteúdo publicado;
- h) Alertar o leitor sobre trabalhos futuros;
- i) Divulgação de trabalhos anteriores que não foram citados ou ou divulgados corretamente;
- j) Autenticação de dados publicados como resultados;
- k) Identificar publicações originais nas quais as ideias ou conceitos já tenham sido discutidos;
- l) Rejeição ou questionamento sobre o trabalho ou as idéias de outros autores (reclamação negativa);

Para além dos motivos classificados como funcionais, [Spinak \(1996\)](#) considera que as citações podem ocorrer, também, por motivações estratégicas, político-partidárias e ideológicas, e persuasivas. Por outro lado, a partir da investigação das dinâmicas de citação em artigos científicos, [Erikson e Erlandson \(2014\)](#) propõem uma taxonomia de motivos para realizar uma citação com base em quatro categorias principais, responsáveis por agrupar as motivações, são elas: argumentação, alinhamento social, alinhamento mercantil e dados. Em

cada uma delas, os autores posicionam motivos para citar, representados como subcategorias. A estrutura de distribuição de motivações proposta por [Erikson e Erlandson \(2014\)](#) encontra-se disposta abaixo:

- a) Argumentação: delimitação; suporte ativo; crítica ativa; suporte passivo; leitura adicional;
- b) Alinhamento social: tradição científica; autoimagem científica; compensação pelo esforço;
- c) Alinhamento mercantil: crédito; credenciais próprias; material de troca; autopromoção; empenho;
- d) Dados: revisão; meta-análise; texto de estudo.

Segundo [Erikson e Erlandson \(2014\)](#), as três primeiras categorias de motivos são aplicáveis a todas as tipologias de artigos científicos, enquanto que a categoria de fatores para citações baseada em dados melhor se adequa às publicações que visam a investigação de estudos de terceiros, a exemplo de artigos de revisão de literatura. Dentre as categorias listadas, a de Argumentação é apontada pelos autores como a mais tradicional, visto que destaca o uso de citações para embasar a defesa de um ponto de vista. Qual o posicionamento ou conceito adotado pelo autor no âmbito de um campo científico específico? Esse é o ponto essencial considerado pela categoria de Alinhamento social. Já a terceira categoria, Alinhamento mercantil, trata do uso das citações como instrumento para a obtenção de diversos tipos de crédito, envolvendo inclusive a questão do reconhecimento já mencionada nesta sessão. Por fim, a categoria de dados considera a literatura citada como dados relevantes para desenvolver explicações sobre temas singulares.

Os exemplos supramencionados demonstram a heterogeneidade de fatores com potencial para influenciar a realização de uma citação. Logo, o aumento no número de citações registradas em uma publicação pode ser modificado conforme a variável considerada. Os estudos que se dedicam à investigação da OACA consideram o tipo de acesso como a variável elementar a ser avaliada, podendo essa ser relacionada ou comparada a outros aspectos de análise.

[Furnival e Hubbard \(2011, p. 163\)](#) alegam que “na medida em que o fator de impacto e as citações ainda têm prestígio entre a comunidade científica, a OACA constitui um argumento forte a favor de AA [acesso aberto], seja via publicação em revistas AA, seja via submissão de

artigos num repositório institucional”. Assim, confirma-se que os estudos sobre OACA se associam às metodologias bibliométricas para averiguar as relações e os desdobramentos que a disponibilização da literatura em acesso aberto causam nos aspectos mensuráveis de análise da produção científica, a exemplo das citações.

2.3.1.1 Estudos sobre vantagens de citação de acesso aberto

A percepção de que a disponibilização *online* e gratuita de publicações científicas poderia gerar desdobramentos positivos em relação ao incremento no volume de citações recebidas e, conseqüentemente, no impacto dessas, tem a autoria atribuída a [Lawrence \(2001\)](#). Mesmo sem nomear tal hipótese com o termo atualmente utilizado na literatura, OACA, o trabalho de [Lawrence \(2001\)](#) inaugura a reflexão a respeito da possível existência de relações de influência entre duas variáveis associadas à uma publicação: o acesso e as citações.

Em função da data de publicação do artigo de [Lawrence](#), maio de 2001, faz-se relevante a consideração de dois fatores: a) o início dos anos 2000 foi marcado por diversas mudanças e evoluções no que tange às TICs e nessa conjuntura, os periódicos eletrônicos começaram a ganhar maior visibilidade, fazendo com que muitas editoras produzissem tanto a versão impressa como eletrônica das publicações, logo, ao se referir a “disponibilidade *online* da literatura científica”, entende-se que o autor se refere aos artigos de conferência com arquivos localizados na web ([CINTRA, 2017](#)). b) Conforme já mencionado na seção 2.4, nos termos do artigo de [Lawrence \(2001\)](#), o acesso recomendado era tratado como “gratuito e *online*”, não há menção ao termo acesso aberto como compreendido atualmente, visto que esse foi firmado posteriormente.

A reflexão iniciada por [Lawrence \(2001\)](#) recebeu continuidade na literatura por meio de outros estudos desenvolvidos em variadas áreas do conhecimento, sendo alguns deles selecionados para compor uma bibliografia temática intitulada [OACA list](#)²⁶ e elaborada pela SPARC Europe (2015). Composta por 70 trabalhos publicados entre 2001 e 2015, a bibliografia expõe a variação de resultados positivos, negativos e inconclusivos obtidos pelas pesquisas mediante testes da hipótese de OACA em diversos contextos e adotando abordagens e metodologias diferentes. Dentre as referências enumeradas pela lista da SPARC Europe (2015), nota-se que 46 (65,7%) identificaram/confirmaram a hipótese de OACA; 17 (24,3%) não

²⁶ Disponível em: <https://sparceurope.org/what-we-do/open-access/sparc-europe-open-access-resources/open-access-citation-advantage-service-oaca/oaca-list/>. Acesso em: 14 out. 2021.

encontram evidências que confirmassem OACA; e 7 (10%) estudos mostraram-se inconclusivos, com presença de dados não significativos ou que fugiam do escopo de mensuração de existência ou não de vantagens. Assim, no contexto da *OACA list*, percebe-se a prevalência de investigações que confirmaram a existência de vantagens de citação em publicações disponibilizadas em acesso aberto.

Ao cumprir o papel de divulgação de estudos sobre a temática de OACA e em virtude da ampliação da produção científica relacionada ao assunto, em 2016, a *SPARC Europe* encerrou a atualização da *OACA list*, fazendo com que essa passasse a não incorporar referências de trabalhos publicados após esse período. Considerando um recorte temporal mais amplo que a *SPARC Europe* (2015), 2001 a 2020, [Langham-Putrow, Bakker; Riegelman \(2021\)](#) conduziram uma revisão sistemática de literatura sobre OACA com base em levantamento realizado em 17 bases de dados. Inicialmente, foram recuperados 5744 itens. Após etapas de tratamento dos dados, remoção de duplicatas e aplicação de critérios de inclusão e exclusão, as autoras obtiveram uma amostra de 134 estudos analisados integralmente.

Do conjunto de trabalhos verificados por [Langham-Putrow, Bakker e Riegelman \(2021\)](#), 47,8% confirmaram a ocorrência de OACA, 27,6% não encontraram indícios de vantagens de citação, 23,9% encontraram confirmação de OACA em apenas alguns subconjunto da amostra estudada e, 0,8% (equivalente a um estudo) mostrou resultados inconclusivos com relação à testagem da hipótese de OACA. Essas autoras destacam que, para além do tipo de acesso, fatores como as bases de dados definidas para coleta de dados, o tipo de publicação selecionado e diferenças disciplinares podem influenciar na questão da confirmação ou negação de OACA. Isto posto, durante o exame das publicações sobre vantagens de citação, [Langham-Putrow, Bakker e Riegelman \(2021\)](#) consideraram as seguintes variáveis de análise: tipo de acesso aberto, tipos de publicação e disciplinas. Com relação às tipologias de publicação, as autoras afirmam o predomínio dos artigos de periódicos, sendo esses 107 de 134 publicações.

No que se refere ao tipo de acesso, [Langham-Putrow, Bakker e Riegelman \(2021\)](#) notaram que: 72% dos estudos deram maior atenção ao acesso aberto verde, 39,6% enfatizaram a análise do acesso aberto dourado; 39,4% investigaram tanto o acesso verde quanto o dourado; e em 52,2% das publicações não ficou evidente a especificação sobre os subtipos de acesso aberto, sendo esse analisado de modo geral em comparação com o acesso fechado. A análise do tipo de acesso das publicações também é efetuada por [Archambault et al. \(2013\)](#), que examina os seguintes subtipos de acesso aberto nas observações realizadas: dourado, verde e

híbrido. De modo complementar, [Piwowar et al. \(2018\)](#), além de levar em consideração os subtipos já mencionados, insere no contexto dos estudos sobre OACA a categoria bronze.

A investigação da hipótese de OACA considerando diferenças disciplinares e os tipos de publicação (*preprint* e artigos em acesso aberto) também é realizada por [Antelman \(2004\)](#). A autora avalia se a variação das taxas de adoção de acesso aberto entre quatro disciplinas (Matemática, Engenharia Elétrica e Eletrônica, Ciências Políticas e Filosofia) demonstra desdobramentos no impacto das publicações quando comparadas a outras não disponibilizadas livremente e *online*. [Antelman \(2004\)](#) justifica que a escolha das disciplinas mencionadas se deu em virtude da expectativa de considerar diferentes pontos de vista de adoção do acesso aberto e por serem disciplinas que, tradicionalmente, publicam ativamente resultados de pesquisa em *preprints*.

Ao pesquisar as OACA, [Eysenbach \(2006\)](#) desenvolveu, escolhendo abordagem em nível de artigo, um estudo bibliométrico comparando artigos de acesso aberto e fechado publicados em um mesmo periódico. [Antelman \(2004\)](#), [Eysenbach \(2006\)](#) e [Cintra \(2017\)](#) são exemplos de alguns dos estudos que calculam em suas investigações o indicador de média de citações. Já [Archambault et al. \(2013\)](#) examina a vantagem de citação, em artigos publicados entre 2004 a 2011 em 22 áreas do conhecimento, por meio da análise dos valores obtidos para a média de citações relativa (*average of relative citation*), que considera as taxas de citação normalizadas a partir das diferenças existentes entre as disciplinas para utilizar referências e receber citações.

Consoante [Costa \(2008, p. 223\)](#), “sabe-se que o impacto da pesquisa é, especialmente nos dias atuais, a moeda mais valiosa no mercado de ações da pesquisa, e o acesso aberto aumenta de modo exponencial esses indicadores”. Nesse sentido, estudos relacionados à OACA buscam mensurar e averiguar contextos e possibilidades em que a relação diretamente proporcional entre acesso aberto e impacto pode ser validada ou negada. Como destaca [Piwowar et al. \(2018\)](#), a hipótese de OACA não é universalmente apoiada, dadas as variadas metodologias adotadas e as diferentes conjunturas observadas.

Considerando a heterogeneidade existente entre os estudos associados a OACA e com o objetivo de propor possíveis aprimoramentos às investigações sobre o tema, [Langham-Putrow, Bakker; Riegelman \(2021\)](#) recomendam boas práticas a serem adotadas por pesquisas bibliométricas realizadas no contexto discutido, são elas:

[...] relatar claramente sobre os tipos de acesso aberto, o número de artigos de acesso aberto ou não, a comparabilidade dos dados usados, o período de tempo (janela de

citação) e uma justificativa para essa escolha, uma métrica comum (e.g., média e mediana de citações por artigo), e fatores de conflito e suas respectivas abordagens ([LANGHAM-PUTROW, BAKKER; RIEGELMAN, 2021](#), p. 9, tradução nossa).

Tendo em vista tais recomendações e ciente das possibilidades de investigação da hipótese de OACA exemplificadas anteriormente, a presente pesquisa busca aproximar as discussões sobre as vantagens do acesso aberto da temática da produção científica sobre COVID-19 e para tal fim, recorre também, às perspectivas complementares de investigação fornecidas pela Altmétria. Dessa forma, a próxima seção trata dos indicadores alométricos.

2.3.2 Indicadores alométricos e vantagens alométricas de acesso aberto

Seriam as métricas tradicionais baseadas em citação suficientes para promover a avaliação da produção científica? Ou, os processos de avaliação podem ser aprimorados por meio do uso de novas possibilidades de interação e disseminação da informação científica inauguradas a partir da popularização da Web 2.0? Para [Araújo \(2014\)](#), a confluência entre a comunicação científica e as tecnologias digitais como mídias sociais, wikis e blogs sinaliza rotas alternativas e complementares para as investigações desenvolvidas no âmbito dos estudos métricos da informação.

A necessidade de melhorias e de diversificação dos critérios utilizados para avaliação da produção científica são questões defendidas pela [San Francisco Declaration on Research Assessment](#) (Declaração de São Francisco sobre Avaliação da Pesquisa - DORA) (2012), que recomenda o reconhecimento e aplicação das oportunidades fornecidas pelas publicações digitais e o uso de métricas que avaliem o impacto da pesquisa em nível de artigo e não do periódico como um todo.

Diante do cenário que busca a expansão da percepção sobre o que é e o que gera o impacto, é publicado o marco da Altmétria, intitulado: *Altmetrics: a manifesto* ([PRIEM et al., 2010](#)). Tendo em vista o argumento da impossibilidade de se ler todo volume de literatura científica existente e considerando os indicadores alométricos como filtros que auxiliam na seleção de fontes e pesquisas relevantes utilizando mecanismos diferentes dos indicadores baseados em citação, [Priem et al.](#) (2010, *online*, tradução nossa) afirmam que “as medidas de contagem de citações são úteis, mas não suficientes”. Logo, os indicadores alométricos são apresentados como caminhos para mensurar o impacto no ecossistema diversificado e digital estabelecido a partir dos avanços das tecnologias e recursos web.

[Costas, Zahedi e Wouters \(2015\)](#) afirmam que a Altmétria demonstra potencial de contribuição no que tange à expansão do entendimento sobre impacto científico,

complementando o conceito a partir de abordagens consideradas por outros tipos de impacto, como por exemplo: social, educacional e cultural. Estudos como os desenvolvidos por [Gontijo e Araújo \(2019\)](#) e [Wei \(2020\)](#) vinculam o uso de indicadores altmétricos à percepção do impacto social obtido pelas publicações científicas em mídias e redes sociais.

Altmétrie, enquanto área, pode ser compreendida como “[...] o estudo e uso de medidas de impacto acadêmico baseadas na atividade em ferramentas e ambientes online” ([PRIEM, 2014, p. 266, tradução nossa](#)). Ao tratarem sobre aspectos relevantes para o entendimento sobre Altmétrie, [Maricato e Martins \(2017\)](#) evidenciam as relações estabelecidas entre essa área e a Bibliometria, Cientometria e Webometria, apontando que o principal ponto de conexão entre essas é a comunicação científica, a qual envolve a investigação e avaliação das dinâmicas de elaboração, disseminação e uso da produção científica. Além disso, os autores registram o entendimento da Altmétrie enquanto:

[...] uma área emergente do campo de Ciência da Informação, que se ocupa do estudo, produção e uso de indicadores científicos e tecnológicos. Baseada em ferramentas que captam informações de múltiplas fontes, que ao serem socializadas em mídias e redes sociais, geram ações e interações de uma grande variedade de atores sociais e novas formas de relacionamento com objetos digitais e entre pessoas ([MARICATO; MARTINS, 2017, p. 55](#)).

A definição supramencionada sintetiza elementos essenciais para compreensão básica da Altmétrie: indicadores, mídias e redes sociais, multiplicidade, dinamicidade e ferramentas *online*. De acordo com [Souza \(2014, p. 261\)](#), outra característica da Altmétrie é a variedade, que possibilita o acompanhamento de “[...] produtos diversos além do artigo (dados brutos, códigos de programação, experimentos, blogs, microblogs, comentários, anotações, etc), plataformas diversas além do periódico tradicional (repositórios institucionais, bibliotecas digitais, comunidades online, etc), e públicos diversos além do acadêmico”.

Os indicadores altmétricos mensuram atividades desempenhadas no âmbito dos ambientes digitais anteriormente citados, registrando atividades como: visualização (*websites* dos editores e repositórios); *downloads* e *uploads*; reuso/adaptação (Github); compartilhamento (Facebook, Twitter); capturas e armazenamento (Mendeley); comentários e menções (Twitter, Facebook, blogs, etc.) ([COSTAS; ZAHEDI; WOUTERS, 2015](#); [MARICATO; MARTINS, 2017](#); [PRIEM, 2014](#); [SOUZA, 2014](#)). Tal diversidade de indicadores é um dos quatro pontos enumerados por [Piwowar \(2013\)](#) como potenciais vantagens da Altmétrie. As outras três vantagens são:

- Compreensão mais apurada sobre impacto, mostrando-nos quais produtos acadêmicos são lidos, discutidos, salvos e recomendados, bem como citados;
- Dados mais oportunos, mostrando evidências de impacto em dias ao invés de anos;
- Uma vitrine sobre o impacto de produtos acadêmicos nativos da web, como conjuntos de dados, software, postagens de blogs, vídeos e outros ([PIWOWAR, 2013](#), p. 9, tradução nossa).

Tais colocações demonstram que o interesse por questões voltadas para as vantagens mensuradas por meio de indicadores também é um aspecto levado em consideração no âmbito da Almetria. Desse modo, observa-se a possibilidade de ampliação das reflexões ora desenvolvidas sobre OACA (vide seção 2.4.1.1), que consiste em uma das diversas vantagens associadas ao acesso aberto ([FURNIVAL; HUBBARD, 2011](#)).

Para além da OACA, é possível considerar também a hipótese de vantagens alométricas de acesso aberto (OAAA), a qual supõe que as publicações disponibilizadas em acesso aberto apresentam maiores chances de atrair mais atenção em ambientes *online*. Diante dessa suposição, [Holmberg et al. \(2019\)](#) desenvolvem estudo para investigar se artigos científicos publicados por universidades da Finlândia, entre 2012 a 2014, demonstravam algum tipo de OAAA. De acordo com os autores, a hipótese de OAAA surge também pela inferência de que, uma vez acessíveis abertamente para o público além do acadêmico, os artigos ficam mais suscetíveis a demonstrar maior impacto social, mensurado em razão dos indicadores alométricos. Assim, para nortear a pesquisa, [Holmberg et al. \(2019, p. 645\)](#) levantaram o seguinte questionamento: “artigos em periódicos de acesso aberto apresentam atividade alométrica mais frequente que os artigos presentes em periódicos baseados em assinatura?”.

O trabalho de [Holmberg et al. \(2019\)](#) analisou dados alométricos vinculados ao Mendeley, Twitter, Facebook, notícias, blogs e Wikipedia. Durante a coleta de dados, os autores utilizaram os *Digital Object Identifiers* (DOIs) de 38.819 publicações para realizar consultas nas *Application Programming Interface* (APIs) do Mendeley e Altmetric.com, bem como coletar ocorrências de citações na base Web of Science. Na análise de dados, os autores buscaram observar diferenças entre campos do conhecimento e as plataformas de onde os dados foram coletados. Dentre os resultados, [Holmberg et al. \(2019\)](#) observaram que existem consideráveis diferenças entre plataformas, principalmente no que diz respeito ao público-alvo de cada uma delas. Os autores afirmam que o Mendeley é enfaticamente usado por acadêmicos, enquanto que o Twitter demonstra maior heterogeneidade de usuários.

Mesmo não utilizando a expressão OAAA para designar a investigação de vantagens alométricas, estudos como [Adie \(2014\)](#) e [Cintra, Furnival e Milanez \(2018\)](#) buscam analisar,

por meio de indicadores alométricos, a influência do acesso aberto na atenção *online* recebida pelas publicações na web social. Na amostra de dados verificada por [Adie \(2014\)](#), foi observada maior atenção recebida pelos artigos no Twitter e no Mendeley, enquanto que nos blogs e portais de notícias, não foi notado indício de vantagens alométricas. Já [Cintra, Furnival e Milanez \(2018\)](#), investigam vantagens de citação e menções na web social em relação ao acesso aberto em periódicos científicos da área de CI que adotam como modelo de negócio a cobrança de APC. Para o desenvolvimento da pesquisa, os autores coletaram dados da Web of Science, Google Acadêmico, Altmetric.com e Mendeley. Os resultados obtidos por [Cintra, Furnival e Milanez \(2018\)](#) demonstraram vantagens de citação e de menção nos artigos de acesso aberto híbrido publicados no *Journal of the American Society for Information Science and Technology* (JASIST) e *Scientometrics*.

O trabalho de [Tennant et al. \(2020, p. 10\)](#), ao abordar questões sobre os impactos acadêmico, econômico e social do acesso aberto, comenta que, de modo amplo, artigos em acesso aberto indicam vantagens de atenção nas mídias sociais e a investigação de tal fato pode contribuir para melhoria nas formas de mensuração de impacto social.

As considerações e estudos apresentados anteriormente sinalizam os indicadores alométricos como possibilidades viáveis para complementar investigações até então focadas na análise da hipótese de OACA. Nessa conjuntura, é apresentada também a OAAA, hipótese que adota lógica semelhante ao que é entendido como vantagens de citação, mas com ênfase no estudo da atenção recebida pelas publicações de acesso aberto na web 2.0. Assim, dadas as interpretações isoladas sobre os indicadores de citação e alométricos, a seguir, busca-se abordar, brevemente, a questão das correlações entre eles.

2.3.3 Correlação entre indicadores de citação e alométricos

Além de observadas de modo complementar aos indicadores bibliométricos, estariam as métricas utilizadas no contexto da Almetria correlacionadas às citações? Essa é a pergunta base do trabalho de [Costas, Zahedi e Wouters \(2015\)](#). Os autores empenharam esforços para mensurar a relação entre o número de citações de uma publicação e o impacto médio de documentos publicados em periódicos. Com base nos DOIs, foram coletados dados da Web of Science e Altmetric.com. Como resultado principal, foram evidenciadas relações positivas, porém fracas (com valores de coeficiente de correlação de Pearson entre 0,21 e 0,4), entre os indicadores estudados.

A literatura científica registra a existência de alguns trabalhos dedicados, em algum grau, à análise das relações estabelecidas entre indicadores bibliométricos e altmétricos. Para tal fim, são adotadas diferentes fontes de informação e métodos, que variam conforme os objetivos firmados pelas pesquisas. Com o foco nos dados bibliométricos do Google Acadêmico e altmétricos fornecidos diretamente pelas APIs do Twitter e Facebook, de um único periódico da CI, a revista DataGramZero, Araújo, Caran e Souza (2016) observaram baixa correlação entre o número de menções e o número de citações por artigo. Os autores notaram também, no conjunto de dados considerados, correlações entre a variável de tempo de publicação com número de citações e menções.

Diante de dados de apenas um periódico, Journal of Medical Internet Research (JMIR), [Eysenbach \(2011\)](#), utilizando as correlações de Pearson, observou aproximações classificadas entre moderadas e fortes entre citações da Scopus e Google Scholar e dados altmétricos de menção no Twitter. Por outro lado, [Thelwall \(2018b\)](#) estudou correlações entre citações e contagem de leitores no Mendeley associadas ao tempo decorrido desde a publicação do trabalho. Destaca-se que, ao invés de utilizar o coeficiente de correlação de Pearson, como observado em outros trabalhos mencionados anteriormente, [Thelwall \(2018b\)](#) adota o coeficiente de correlação de Spearman. As análises realizadas pelo autor demonstram que os indicadores calculados a partir dos dados da ferramenta de gerenciamento de referências são mais efetivos para observar publicações recentes que as citações.

Os estudos de correlação são relevantes para compreender as aproximações e possibilidades de complemento entre dados bibliométricos e altmétricos. Assim como as análises altmétricas, são heterogêneos no que tange às abordagens adotadas, fornecendo possibilidades variadas de compreensão entre aspectos da Bibliometria e Almetria. Defronte às informações estabelecidas nesta e nas seções anteriores deste trabalho, a continuidade apresentada a seguir trata da metodologia do estudo.

3 METODOLOGIA

O presente capítulo encarrega-se de descrever a metodologia definida para alicerçar o desenvolvimento da investigação. Primeiramente, apresenta-se a caracterização da pesquisa, levando em consideração a concepção filosófica, bem como a classificação da abordagem e método do estudo. Em seguida, encontram-se elucidadas as etapas previstas e as fontes consideradas para realização da coleta de dados. Por fim, são enumerados fatores associados aos procedimentos metodológicos, que estão organizados de acordo com os objetivos específicos deste trabalho.

3.1 CARACTERIZAÇÃO DA PESQUISA

Consoante Creswell e Creswell (2021), o planejamento e a definição dos procedimentos de pesquisa devem deliberar sobre três principais componentes: a perspectiva filosófica, a abordagem de pesquisa e o estabelecimento do tipo de método. Tais elementos corroboram com a caracterização da pesquisa.

A perspectiva ou concepção filosófica diz respeito às crenças básicas e ideias mais abrangentes que orientam as ações desenvolvidas na investigação, influenciando a prática. Além disso, Creswell e Creswell (2021) explicam que a identificação da perspectiva filosófica apoia a detecção da abordagem do estudo (qualitativa, quantitativa ou mista), visto que fornece uma orientação geral quanto à natureza e ao universo com os quais a pesquisa científica se relaciona. Diante disso, afirma-se que o presente estudo adota como pressuposto filosófico o *Pós-Positivismo*, caracterizado pelo desenvolvimento de observações e mensurações empíricas, em que o pesquisador parte de uma teoria, coleta dados com a intenção de verificá-la e realiza análises de modo a compreender se a hipótese inicial pode ser defendida ou refutada. Outro traço presente da pesquisa pós-positivista é a objetividade, usada na aplicação dos métodos e relevante para estudos do tipo quantitativo.

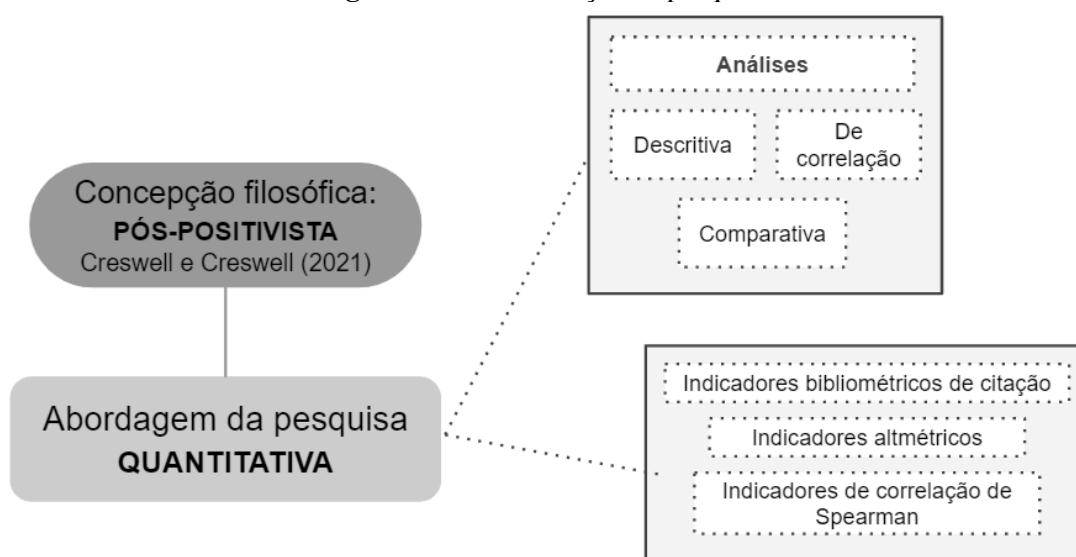
A definição de um posicionamento mais amplo que englobe os outros elementos de caracterização da pesquisa também é orientada por Bryman (2012). Como sinônimo de concepção filosófica, o autor utiliza a expressão “perspectiva epistemológica”.

A perspectiva filosófica supramencionada relaciona-se diretamente com a abordagem de pesquisa adotada pelo presente estudo, a *quantitativa*, a qual é estruturada e baseada em análises numéricas que possibilitam o exame e observação de relações entre variáveis. Por variáveis, Creswell e Creswell (2021, p. 3) entendem “[...] medidas, geralmente, com

instrumentos para que os dados numéricos possam ser analisados com procedimentos estatísticos”.

Na conjuntura da presente pesquisa, diretamente associados à abordagem quantitativa estão os indicadores bibliométricos, alométricos e de correlação, já discutidos nas seções anteriores deste relatório (2.4.1, 2.4.2 e 2.4.3). Assim, metodologicamente, este estudo caracteriza-se a partir dos elementos e componentes anteriormente destacados e a lógica de caracterização da pesquisa está exposta de forma sistemática na **Figura 7** abaixo.

Figura 7 - Caracterização da pesquisa



Fonte: elaborada pela autora com base em Bryman (2012) e Creswell e Creswell (2021)

3.2 FONTES DE DADOS

Albrecht e Ohira (2000, p. 133) afirmam que “o objetivo de uma base de dados é fornecer informação atualizada (recursos estruturais), precisa e confiável (não dar a informação pela metade) e de acordo com a demanda (oferecer o que o usuário necessita)”. Assim, as autoras atestam as bases de dados como fontes de informação importantes e qualificadas. Isto posto, na sequência, apresenta-se as fontes de dados utilizadas pela presente pesquisa, acompanhadas de suas respectivas descrições.

Quadro 6 - Fontes de dados utilizadas na pesquisa

Fonte	Descrição	Dados gerados
Dimensions.ai	Desenvolvida em 2018 e mantida pela empresa Digital Science, a Dimensions é	<ul style="list-style-type: none"> Dados de identificação da publicação: DOI, título,

	<p>uma base de dados multidisciplinar que reúne dados vinculados tanto às publicações tradicionais, como aos registros de ensaios clínicos, patentes e políticas. A proposta da base é estruturar dados associados a todo o ciclo de vida da pesquisa, considerando desde o financiamento até o mapeamento do impacto. A base disponibiliza um módulo gratuito para realização de buscas, todavia, para exportação de dados e outras funcionalidades mais completas, é necessário possuir autorização de acesso. Pesquisadores podem realizar solicitação de acesso temporário para fins de pesquisa por meio de formulário disponível no site da base. O pedido de solicitação é analisado pela empresa Digital Science.</p>	<p>data, editora, periódico);</p> <ul style="list-style-type: none"> • Dados de citação; • Tipos de acesso: aberto e fechado; • Subtipos de acesso aberto: bronze, dourado, híbrido e verde; • Tipo de publicação: artigo ou <i>preprint</i>.
<p>Plataforma Altmetric.com (API)</p>	<p>Ferramenta altmétrica paga que monitora e rastreia atividades <i>online</i> relacionadas à produção científica em ambientes como Facebook, Twitter, Mendeley, LinkedIn, Wikipédia, plataformas como Publons, portais de apontamentos de pesquisas (F1000), entre outros. O acesso à API da ferramenta depende de autorização prévia. O pedido de solicitação é analisado pela empresa Digital Science. Por um período do estado pandêmico, o acesso ficou liberado temporariamente.</p>	<ul style="list-style-type: none"> • Dados altmétricos das publicações no Twitter.
<p>Mendeley (API)</p>	<p>O Mendeley caracteriza-se como uma ferramenta gratuita para gerenciamento de citações e referências e é entendido, também, como uma rede social acadêmica.</p>	<ul style="list-style-type: none"> • Dados altmétricos de leitores/capturas (extraídos diretamente do Mendeley).

Fonte: elaborado pela autora

A consulta às APIs da plataforma Altmetric.com e do Mendeley foi efetuada por meio da ferramenta [Odisseia Métricas](#)²⁷, que foi desenvolvida no Google Colab com o intuito de possibilitar o acesso mais intuitivo as APIs mencionadas, viabilizando a extração de dados (RAMOS; MARICATO, 2020).

As informações evidenciadas pelo **Quadro 6** descrevem as fontes de dados selecionadas pela presente investigação. Destaca-se que o uso combinado da Dimensions e Altmetric.com é

²⁷ Disponível em:

https://colab.research.google.com/drive/1L75mJbqteVE0kMoTMK1_kLVXWFZpQumi#scrollTo=ohRqrEb7RXlm. Acesso em: 4 fev. 2022.

adotado pela pesquisa bibliométrica e altmétrica de Gontijo e Araújo (2019), que analisa o impacto acadêmico e social em artigos científicos de acesso aberto. Ainda é válido apontar estudos como o de [Thelwall \(2018a\)](#) e [Hook, Porter e Herzog \(2018\)](#), que analisam a Dimensions de acordo com aspectos estruturais, limitadores e de qualidade dos dados. Por fim, o uso da Plataforma Altmetric.com e do Mendeley com interesses próximos ao desta pesquisa é observado no trabalho de [Cintra, Furnival e Milanez \(2017\)](#). Assim, definidas as fontes de informação, o próximo passo da pesquisa consistiu no uso dessas para efetuar as etapas de coleta e tratamento dos dados, conforme detalhado na seção a seguir.

3.3 COLETA E TRATAMENTO DE DADOS

Realizada em fevereiro e março de 2022, a coleta ocorreu com o intuito de reunir os dados necessários para embasar a execução dos procedimentos metodológicos associados aos objetivos específicos da pesquisa, bem como viabilizar as análises e interpretações que serão apresentadas como resultados do estudo desenvolvido. Dessa forma, a presente pesquisa considerou a realização da coleta e do tratamento dos dados por etapas, as quais são enumeradas e detalhadas a seguir.

Etapa 1 - Coleta de dados do conjunto de artigos e preprints sobre COVID-19 na base de dados Dimensions.ia

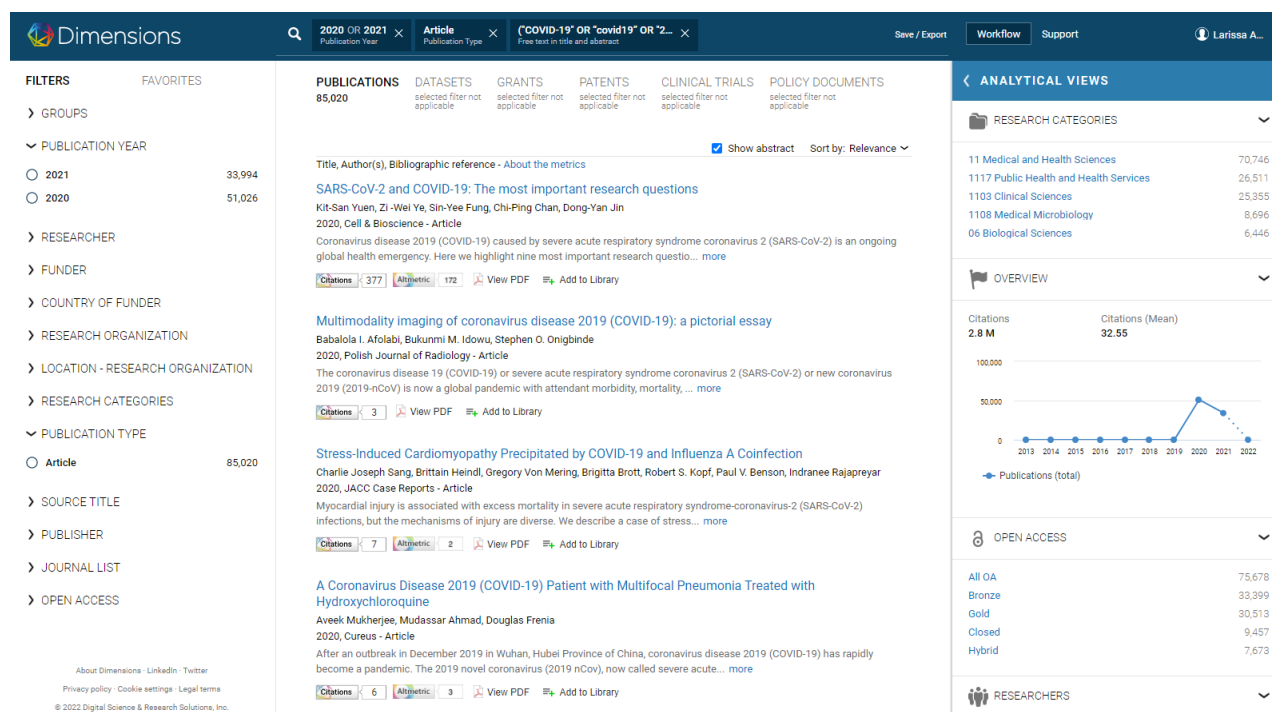
No contexto deste estudo, a escolha da Dimensions para coleta de dados se justifica em razão das potencialidades que a referida base possui para fornecer dados relevantes ao embasamento das análises voltadas para identificação da existência de relação entre o tipo de acesso e tipo de publicação com quantidade de citações. A partir da referida fonte de informação, foram coletados dados: bibliométricos; de identificação das publicações; de tipos de acesso: aberto e fechado; de subtipos de acesso aberto: bronze, dourado, híbrido e verde; e de tipo de publicação.

No que tange aos dados sobre os tipos de acesso, a Dimensions demonstrou um espectro diversificado de classificação, englobando os subtipos de acesso considerados por [Piwowar et al. \(2018\)](#): fechado, dourado, verde, híbrido e bronze. Vale destacar que a classificação do tipo de acesso apresentada pela Dimensions se baseia na integração de dados provenientes de duas fontes relacionadas ao acesso aberto: Unpaywall e o DOAJ ([BODE et al., 2019](#)). Ainda sobre a origem dos dados obtidos a partir da Dimensions, registra-se, conforme [Bode et al. \(2019\)](#) e

[Dimensions \(2019\)](#), que a base extrai as citações a partir de fontes como [Crossref](#) , [PubMed Central](#) e/ou [OpenCitations](#).

Para execução desta etapa 1 da coleta, foi utilizado o módulo restrito da Dimensions.ai, denominado Dimensions Analytics (**Figura 8**), o qual pôde ser acessado mediante *login* temporário concedido pela empresa [Digital Science](#), mantenedora e responsável pela base. Na Dimensions, é possível escolher entre algumas opções de busca. As consultas podem ser efetuadas em todos os metadados do registro bibliográfico (*full data*), somente nos campos de título e resumo (*title and abstract*), ou por meio do *Digital Object Identifier* (DOI). Um vez selecionada uma dessas opções, faz-se necessário o preenchimento do campo de pesquisa com uma expressão de busca criada a partir da combinação de termos e operadores booleanos e lógicos aceitos pela base (e.g. AND, OR, NOT, (), ?, *, ~n).

Figura 8 - Interface do módulo Dimensions Analytics



Fonte: elaborada pela autora

Durante a consulta realizada pela presente pesquisa, foi selecionada a opção de busca por *título e resumo*. Para a elaboração da expressão de busca, foram escolhidos termos sobre a temática de COVID-19 a partir da análise das *strings* de busca utilizadas por [Bermúdez-Rodríguez et al. \(2020\)](#), [Torres-Salinas, Robinson-Garcia e Castillo-Valdivieso \(2020\)](#), [Zhai, et al. \(2020\)](#) e [Zhang et al., 2020](#) em suas investigações, bem como, pelo estudo dos termos

descritores sobre o assunto²⁸ elencados pelo vocabulário estruturado e multilíngue de Descritores em Ciências da Saúde (DeCS). Como resultado, a *expressão de busca* obtida e utilizada para realizar as consultas e a coleta dos dados de registros bibliográficos na base Dimensions foi:

("COVID-19" OR "covid19" OR "2019-nCoV" OR "coronavirus disease 2019") AND ("coronavirus" OR "corona virus" OR "coronavírus") AND ("SARS-CoV-2" OR "severe acute respiratory syndrome coronavirus 2" OR "síndrome respiratória aguda grave 2")

Além da consulta por meio da expressão de busca, a Dimensions disponibiliza filtros para refinamento dos registros recuperados. Durante a realização da presente etapa, foram utilizados dois filtros específicos. A definição deles, bem como os respectivos valores de preenchimento considerados encontram-se enumerados abaixo:

- a) *Publication Year* (ano de publicação): 2020 e 2021;
- b) *Publication Type* (tipo de publicação): Article e *Preprint*.

Faz-se válido destacar que o recorte temporal proposto para a coleta de dados em questão, 2020 e 2021, compreende o período de anos completos desde o início da vigência do estado pandêmico de COVID-19 declarado oficialmente pela OMS em março de 2020. Com relação ao filtro de tipo de publicação, este foi aplicado com o intuito de delimitar o universo da pesquisa apenas a artigos e *preprints*, fazendo com que outras tipologias documentais indexadas pela Dimensions fossem desconsideradas, a exemplo de livros, capítulos de livros, monografias e trabalhos publicados em eventos.

²⁸ Descritores sobre COVID-19 indexados no DeCS. Disponível em: https://decs.bvsalud.org/thi?filter=thi_termall&q=COVID . Acesso em: 12 fev. 2022.

Figura 9 - Expressão de busca e filtros utilizados para pesquisar artigos e *preprints* na Dimensions.

The screenshot shows a search interface with two filter buttons at the top: "2020 OR 2021" (Publication Year) and "Article OR Preprint" (Publication Type). Below these is a text box containing the following search query: ("COVID-19" OR "covid19" OR "2019-nCoV" OR "coronavirus disease 2019") AND ("coronavirus" OR "corona virus" OR "coronavírus") AND ("SARS-CoV-2" OR "severe acute respiratory syndrome coronavirus 2" OR "síndrome respiratória aguda grave 2").

Below the query box is a "Hide operator info" link and a list of Boolean operators with their definitions:

- AND Requires both terms on either side of the Boolean operator to be present for a match
- OR Requires that either term (or both terms) be present for a match
- NOT Requires that the following term not be present
- () Use parentheses to control the Boolean logic for a query
- ? Single character wildcard. Cannot be used as a leading wildcard or inside of quotes.
- * Multiple characters wildcard. Cannot be used as a leading wildcard or inside of quotes.
- ~n Proximity search, e.g. "ambient noise"~4

At the bottom of the interface, there is a search bar with the following options: "Search in: Full data Title and abstract (Applied if no other field is specified)

Fonte: elaborada pela autora

Dessa forma, diante da aplicação da expressão de busca e dos filtros supramencionados, conforme demonstra a **Figura 9**, a base de dados recuperou um total de 91.807 registros, sendo 85.020 artigos e 6.787 *preprints*. Os dados relacionados aos registros recuperados pela Dimensions foram exportados diretamente da base em lotes de até 50 mil registros e em arquivos de planilhas no formato XLSX. Após o *download* de todos os lotes, os dados foram reunidos em uma única planilha, gerando o conjunto de dados (*dataset*) inicial da pesquisa.

Etapa 2 - Análise do dataset inicial e seleção dos dados de identificação e de citação

De posse do conjunto de dados coletados na **Etapa 1**, deu-se início à análise da estrutura das planilhas e das características dos dados exportados da Dimensions, de modo que fosse possível realizar a primeira filtragem das informações obtidas. Assim, foi elaborado um quadro de análise de dados e metadados, que encontra-se disponível na íntegra no **APÊNDICE A** desta dissertação.

Após observação das características e da descrição dos dados, ficou evidente que os *datasets* exportados a partir da Dimensions apresentam um conjunto de metadados padrão, que reúne dados que podem ser agrupados em algumas categorias de análise, são elas:

- a) Identificação bibliográfica das publicações: DOI, códigos identificadores, título, resumo, fonte, agradecimentos, data, editor, volume, paginação;
- b) Tipo de acesso e tipo de publicação: acesso fechado e aberto (levando em consideração os subtipos bronze, dourado, híbrido e verde), classificação dos tipos de documentos;
- c) Autoria e afiliação: autores, autores correspondentes da publicação, afiliação institucional;
- d) Financiamento e localização: instituições de pesquisa; cidade, estado e país da instituição de pesquisa; dados de financiamento; país financiador; grupo financiador;
- e) Métricas: quantidade de citações; citações recentes; indicadores calculados pela própria Dimensions (*Relative Citation Ratio* e *Field Citation Ratio*); Altmetric Score;
- f) Fontes: links e indicação das fontes originais das publicações;
- g) Áreas, assuntos e temas: grandes áreas temáticas e categorias de classificação de assuntos.

Diante da organização e agrupamento dos metadados nas categorias de análise pontuadas acima e, tendo em vista o escopo da presente pesquisa, fez-se a seleção de quais metadados seriam mantidos e quais seriam ignorados por não demonstrarem relação direta com os objetivos deste trabalho. Como resultado desta etapa, apresenta-se o **Quadro 7**, que lista os metadados de identificação e citação selecionados inicialmente para dar continuidade à coleta de dados.

Quadro 7 - Seleção de dados de identificação e citação das publicações coletados a partir da Dimensions

Categoria de análise	Metadado	Descrição
Identificação bibliográfica	<i>DOI (Digital Object Identifier)</i> <i>Digital Object Identifier</i>	DOI - Código identificador único da publicação
	<i>Title</i>	Título da publicação
	<i>Source title</i>	Fonte da publicação

	<i>Publisher</i>	Editora
	<i>PubYear</i>	Ano da publicação
	<i>Publication Date (online)</i>	Data da publicação digital. Formato: ano ou ano-mês-dia
Acesso aberto e tipo de publicação	<i>Open Access</i>	Classificação do tipo e subtipo de acesso da publicação. Disponível tanto para artigo quanto <i>preprint</i>
	<i>Publication Type</i>	Classificação do tipo de publicação. Neste estudo são analisadas apenas duas: artigos e preprints
Métricas	<i>Times cited</i>	Quantidade de citações recebidas. Disponível tanto para artigos quanto para preprints
	<i>Altmetric</i>	Valor correspondente ao <i>Altmetric Attention Score</i>
Fontes	<i>Source Linkout</i>	Link da fonte
	<i>Dimensions URL</i>	Link do registro da publicação indexada na dimensions

Fonte: elaborado pela autora

Etapa 3 - Organização, tratamento dos dados e elaboração de listagem com os DOIs do respectivo conjunto de publicações

Com o objetivo de gerar a listagem de DOIs dos artigos e *preprints* sobre COVID-19 a ser utilizada na coleta de dados altmétricos de leitores/capturas e menções no Twitter, a presente etapa de organização e tratamento inicial dos dados compreendeu as seguintes atividades:

- a) *Remoção de linhas em branco*: a planilha resultante da **Etapa 2** apresentava linhas sem nenhuma informação registrada. Assim, essas foram removidas do conjunto de dados para evitar inconsistências futuras;
- b) *Verificação e remoção de linhas de registros sem DOIs*: para o presente estudo, os DOIs das publicações caracterizam-se como dados indispensáveis, visto que as fontes de informação selecionadas para a coleta utilizam os DOIs como as chaves primárias das consultas. Situação semelhante é descrita nos trabalhos de [Holmberg et al. \(2019\)](#) e

[Piwowar et al. \(2018\)](#), em que os DOIs são utilizados como chave de consulta às APIs de fontes de dados. Logo, sem os identificadores não é possível recuperar outros dados da publicação. Portanto, as linhas da planilha que demonstravam o campo de DOI em branco foram removidas no *software* Microsoft Office Excel. Para detalhes sobre o quantitativo de linhas excluídas, vide **Quadro 8**;

- c) *Tratamento de inconsistências de caracteres e verificação de duplicatas de DOIs*: a execução desta atividade foi desempenhada na ferramenta de tratamento de dados complexos [OpenRefine](#)²⁹, que possibilita a padronização e agrupamento de valores semanticamente parecidos, mesmo que sejam identificadas variações e inconsistências de caracteres maiúsculos e minúsculos, espaçamentos e pontuações. Após a padronização de valores, os DOIs duplicados foram selecionados e removidos;
- d) *Remoção de DOIs duplicados*: o OpenRefine não identificou nenhuma duplicata para DOI em *preprints* (os DOIs duplicados e removidos eram de publicações do tipo artigo). No todo, foram removidos 152 DOIs, como demonstra o **Quadro 8** a seguir

Quadro 8 - Detalhamento do quantitativo de registros, por tipo de publicação, componentes da amostra da pesquisa

Tipo de publicação	Etapa 1 Registros de publicações exportados da Dimensions	Etapa 3 Registros sem DOI (Removidos)	Etapa 3 Registros com DOI (Mantidos)	Etapa 3 DOIs duplicados (Removidos)	Etapa 3 <i>Amostra final</i> de registros com DOI, sem duplicatas
Artigos	85.020	1.391	83.629	152	83.477
Preprints	6.787	264	6.523	0	6.523
Total	91.807	1.655	90.152	152	90.000

Os valores dispostos no **Quadro 8** sistematizam as etapas e as atividades de tratamento e organização de dados efetuadas até a obtenção da amostra final de 90.000 DOIs relativos a publicações na temática de COVID-19, incluindo *preprints* e artigos. Diante dessa listagem de identificadores foi possível dar início a próxima e última etapa da coleta de dados, a qual encontra-se descrita abaixo.

Etapa 4 - Consulta e coleta dos dados altmétricos

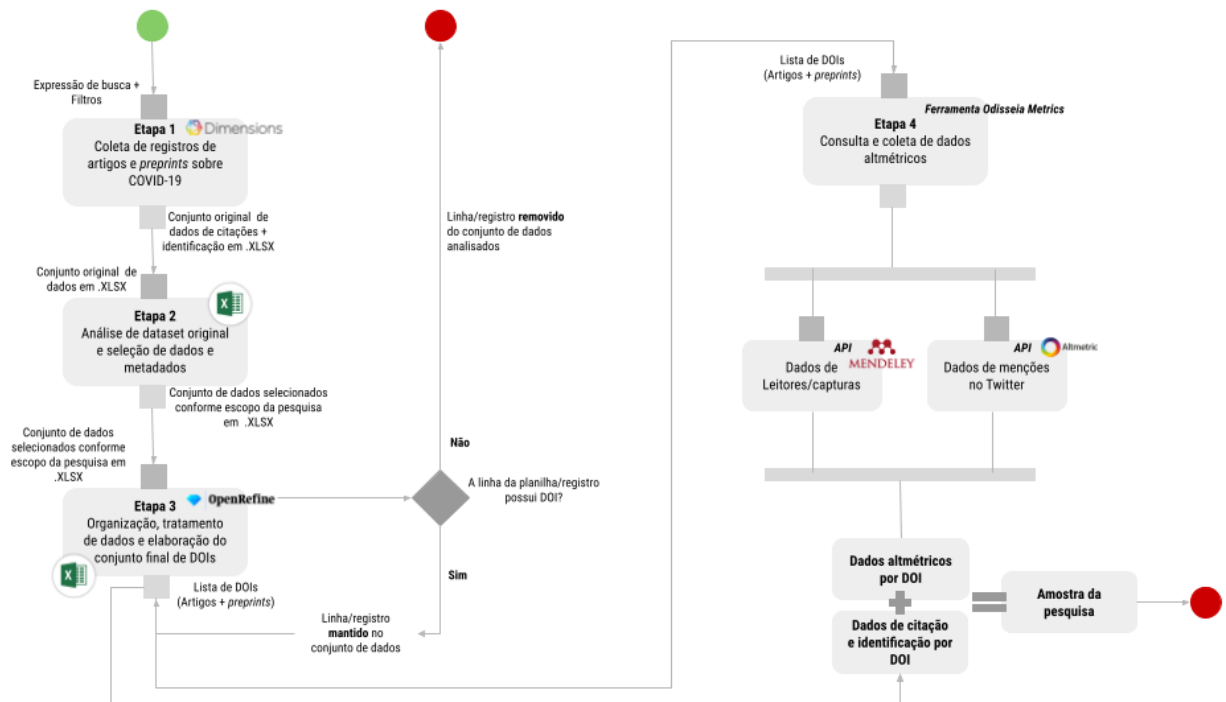
²⁹ Disponível em: <https://openrefine.org/>. Acesso em: 19 fev. 2022.

Conforme já apresentado na seção **3.2 fontes de dados** e esquematizado por meio do **Quadro 6**, a coleta dos dados altmétricos de leitores/capturas no Mendeley e menções no Twitter ocorreu através da extração automática de dados altmétricos realizada pela ferramenta Odisseia Métrics (RAMOS; MARICATO, 2020) a partir da lista de DOIs obtida na Etapa 3. A referida ferramenta se conecta às APIs tanto do Mendeley quanto da Plataforma Altmetric.com e, após realizar as consultas solicitadas através dos DOIs, gera planilhas com os dados extraídos no formato XLSX.

Os dados altmétricos obtidos nesta etapa foram compilados e incorporados ao conjunto principal de dados da pesquisa, elaborado a partir das atividades e procedimentos de coleta, tratamento e organização dos dados realizados nas **Etapas 1, 2 e 3** descritas anteriormente. Dessa forma, ao fim da **Etapa 4**, todos os dados necessários para efetuar os cálculos dos indicadores bibliométricos e altmétricos, bem como dar início a interpretação dos resultados da pesquisa, haviam sido coletados e tratados.

A fim de concluir esta parte da pesquisa e representar o passo a passo das etapas de coleta e tratamento de dados por meio de uma visão sistematizada dos procedimentos desempenhados, apresenta-se a **Figura 10** abaixo, que sintetiza o conteúdo descrito na seção 3.3 do presente trabalho

Figura 10 - Etapas e procedimentos de coleta e tratamento de dados



Fonte: elaborada pela autora.

3.4 ANÁLISE DE DADOS

As etapas e atividades de coleta e tratamento de dados expressas anteriormente forneceram os dados necessários para a realização das análises responsáveis por fomentar a obtenção dos resultados da presente pesquisa. Dessa forma, o conteúdo expresso nesta seção e nas subseções abaixo descrevem aspectos relevantes aplicados para desempenhar a análise dos dados da amostra deste estudo, composta por 90.000 publicações, sendo 83.477 artigos e 6.523 *preprints*, conforme já assinalado pelo **Quadro 8**.

De modo a orientar as decisões tomadas no decorrer da análise dos dados, fez-se necessário, primeiramente, executar duas atividades associadas ao detalhamento das variáveis da pesquisa: definição e caracterização, uma que, “o tipo de variável é crucial no cálculo de estatísticas descritivas e na representação gráfica de resultados” (FAVERO; BELFIORE, 2017, p. 18). As informações resultantes de tais atividades encontram-se indicadas na subseção a seguir.

3.4.1 Definição e caracterização das variáveis da pesquisa

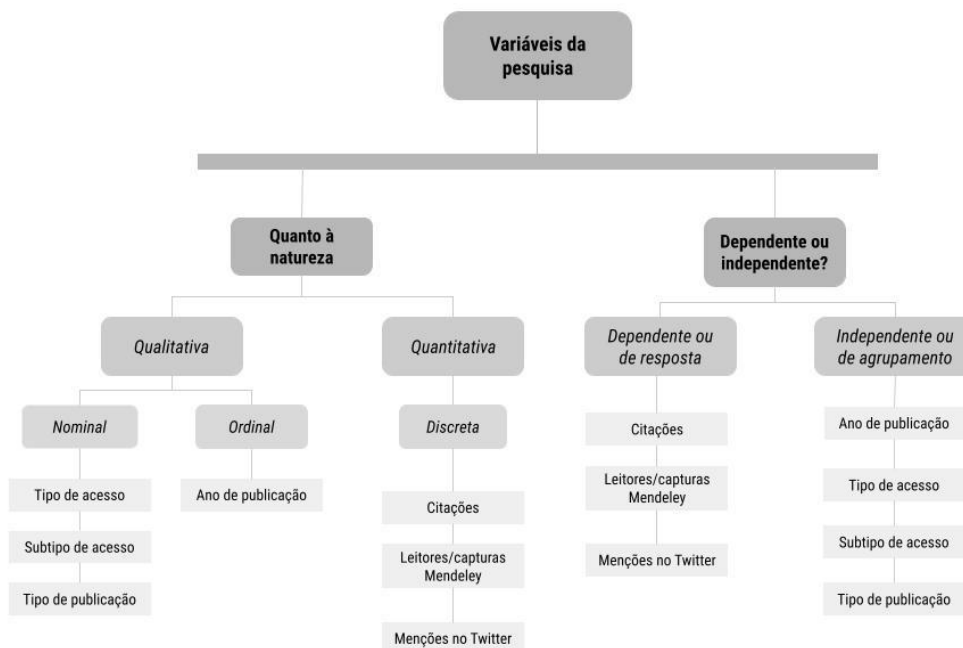
Compreendidas como características, atributos ou propriedades passíveis de variação e que podem ser observadas, mensuradas e/ou classificadas em mais de uma categoria, as variáveis desempenham papel relevante no direcionamento das análises realizadas pela pesquisa (CRESWELL; CRESWELL; GIL, 2019). Isto posto, no âmbito do presente estudo, as variáveis de pesquisa são definidas da seguinte forma:

- a) *Citações*: representam a quantidade de vezes em que uma publicação foi citada formalmente em outros documentos. Os valores de citação foram coletados a partir da base de dados Dimensions. De acordo com as informações fornecidas pela base, os dados de citação da [Dimensions \(2019\)](#) são extraídos e agregados a partir das seguintes fontes: Crossref, PubMed Central, OpenCitations e diretamente de registros de textos completos disponibilizados por editoras;
- b) *Leitores/capturas Mendeley*: originalmente tratados somente como “readers/leitores” pelo Mendeley, representam a quantidade de usuários dessa ferramenta de gerenciamento de referências que capturaram a publicação para inclusão em sua biblioteca de bibliografias. [Silva Filho e Vanz \(2019\)](#), nomearam este indicador como “capturas”. Em consonância com o que propuseram esses autores e com a literatura

- internacional, este estudo adota o termo "leitores/capturas" em detrimento de "leitores" ou "capturas" individualmente, uma vez que a seleção de uma publicação para a biblioteca Mendeley não garante, necessariamente, a leitura dessa pelo usuário;
- c) *Menções no Twitter*: simbolizam o número de contas do Twitter que twittaram/mencionaram uma publicação, por meio do DOI, em postagens realizadas na referida mídia social. Os valores associados a esta variável foram coletados a partir da consulta a API da ferramenta Altmetric.com;
 - d) *Subtipos de acesso*: refere-se à classificação mais específica do *status* de acesso da publicação. Em publicações identificadas com o tipo de acesso aberto, os valores possíveis de subtipos de acesso são: dourado, verde, bronze e híbrido. Para maiores detalhes sobre cada um desses, vide **Quadro 4** - Categorização de subtipos relacionados ao acesso aberto. Para fins de pesquisa, em publicações do tipo fechado, o único valor de subtipo de acesso considerado é: fechado;
 - e) *Tipo de acesso*: classificação mais ampla atribuída para a situação de acesso em que se encontra a publicação. No contexto deste trabalho, são considerados somente dois valores possíveis: aberto ou fechado;
 - f) *Tipo de publicação*: classificação da tipologia dos documentos levando em consideração características específicas. Neste estudo são analisados apenas dois tipos de publicação, os artigos e os *preprints*;

Além das definições supramencionadas, para entender as variáveis da pesquisa é válido apresentar a caracterização do tipo de cada uma delas. A classificação das variáveis orienta a tomada de decisão no contexto da investigação.

A partir das possibilidades de classificação de variáveis de pesquisa expostas por autores como Barbetta (2012) e Gil (2019), as variáveis do presente estudo foram caracterizadas quanto à natureza (qualitativa nominal ou ordinal; quantitativa discreta ou contínua) e de acordo com o aspecto de dependência que demonstram entre si. As variáveis dependentes podem ser compreendidas como os focos de mensuração de um estudo, podendo receber influência das variáveis independentes (GIL, 2019). Isto posto, o resultado da categorização efetuada para o presente trabalho encontra-se expresso na **Figura 11**.

Figura 11 - Categorização das variáveis da pesquisa quanto à natureza e dependência

Fonte: elaborada pela autora.

Uma vez esclarecida a classificação das variáveis demonstradas na **Figura 11**, o andamento da pesquisa deu-se por meio do desenvolvimento dos procedimentos de análises relatados na subseção seguinte.

3.4.2 Procedimentos de análise

Assim como os procedimentos de tratamento e organização dos dados (vide seção 3.3) demandaram a adoção de ferramentas de apoio, para a análise dos dados, foram utilizados *softwares* estatísticos como o Microsoft Office Excel e o [Jamovi](https://www.jamovi.org/)³⁰(versão sólida 2.2.5). O segundo caracteriza-se como uma ferramenta estatística gratuita, de código aberto, desenvolvida por membros da comunidade científica, a partir da linguagem estatística R, com o intuito de fornecer uma alternativa intuitiva e descomplicada para os pesquisadores efetuarem análises estatísticas diversas, entre elas: testes de normalidade, correlação e regressão, testes paramétricos e não paramétricos, testes de confiabilidade e outros.

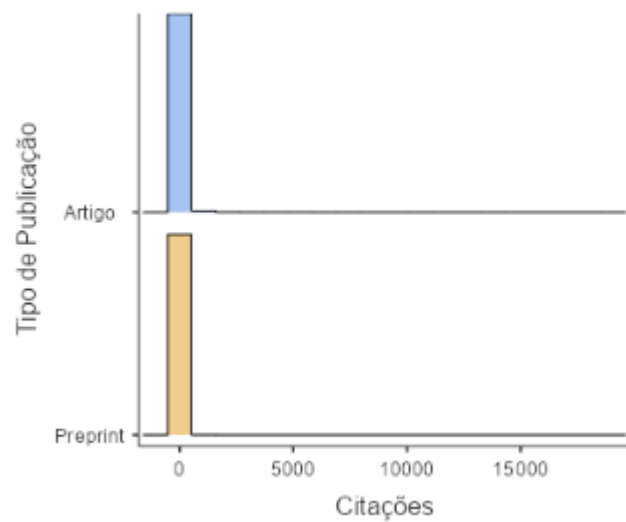
A princípio, fez-se uso do Microsoft Office Excel para iniciar as análises descritivas dos dados coletados e tratados, gerar tabelas e as primeiras visualizações dos dados a partir das distribuições de frequências e porcentagens por variável de pesquisa. Já o Jamovi foi utilizado para ampliar a análise descritiva do estudo; calcular as medidas de tendência central das

³⁰ Disponível em: <https://www.jamovi.org/> . Acesso em: 25 mar. 2022.

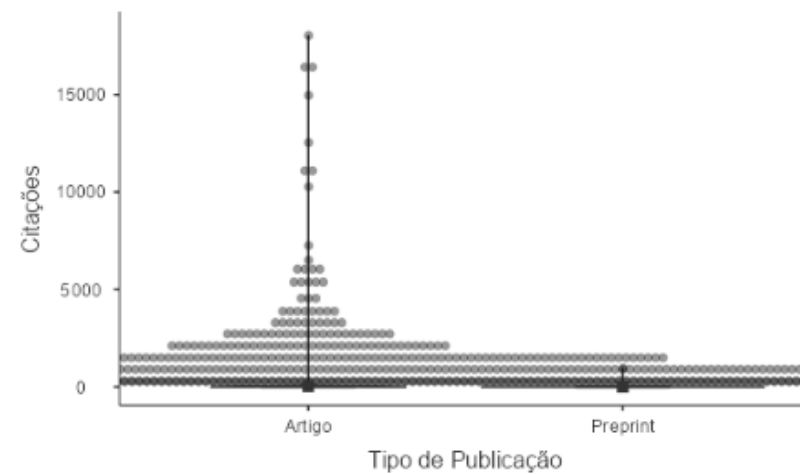
variáveis (média, moda, mediana); gerar gráficos para observação dos tipos de distribuição; realizar testes estatísticos e cálculo de correlação entre variáveis.

Para Gil (2019), a mensuração da correlação entre as variáveis de pesquisa deve ser precedida pela verificação do tipo de distribuição (normal/simétrica ou não normal/assimétrica) dos dados quantitativos, de modo que seja possível identificar qual a análise mais adequada para ser aplicada ao estudo (paramétrica: contextos em que se observa distribuições normais; não-paramétrica: indicada para distribuições assimétricas). A verificação da normalidade das variáveis foi realizada por meio da comparação entre os valores de tendência central e da observação de histogramas e gráficos de dispersão (*scatterplot*) gerados no Jamovi. As médias, modas e medianas foram comparadas (**Quadro 9**), uma vez que, em distribuições normais, essas apresentam valores iguais ($\bar{X}=Mo=Md$) (BARBETTA, 2012, p. 108), enquanto que “para uma curva assimétrica, a média distancia-se da moda, e a mediana situa-se em uma posição intermediária”, podendo ser classificada como: assimétrica positiva ou à direita; assimétrica negativa ou a esquerda (FAVERO; BELFIORE, 2017, p. 91). As médias, medianas e modas apresentadas no quadro **Quadro 9** se mostram diferentes, sendo mais um indicativo de assimetria na distribuição dos valores da amostra.

O histograma, gráfico que representa a distribuição de frequência de uma variável quantitativa (tanto discreta como contínua) por meio de barras verticais, permite a observação do formato da curva da distribuição, que apresenta, quando normal, a forma de um sino. Em casos de distribuições assimétricas dos valores da variável, a curva apresenta formato mais irregular. Os histogramas representados pelos **Gráficos 3, 4 e 5** evidenciam distribuições assimétricas à esquerda para as variáveis quantitativas de citação, leitores/capturas e menções no Twitter. Já os gráficos de dispersão (**Gráficos 6, 7, 8**) indicam elevada concentração de valores entre os intervalos de 0 a 5.000 para as variáveis quantitativas. Esses gráficos ainda demonstram a alta taxa de publicações com 0 citações, leitores/capturas e menções.

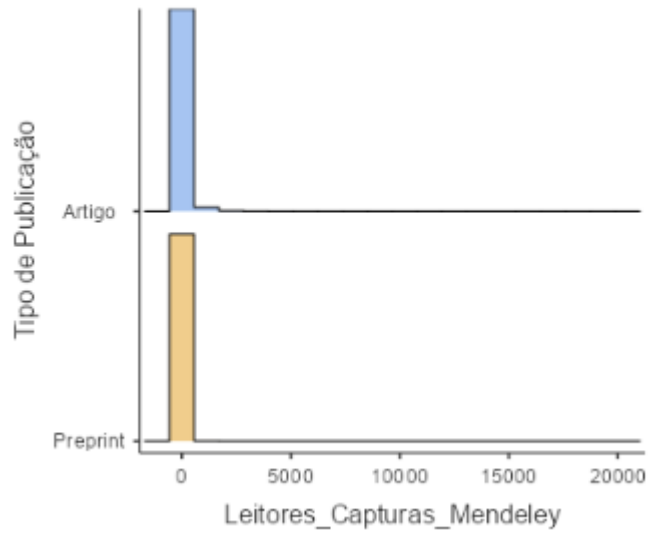
Gráfico 3 - Histograma de citações por tipo de publicação (2020-2021)

Fonte: dados da pesquisa

Gráfico 6- Dispersão dos dados de citação por tipo de publicação

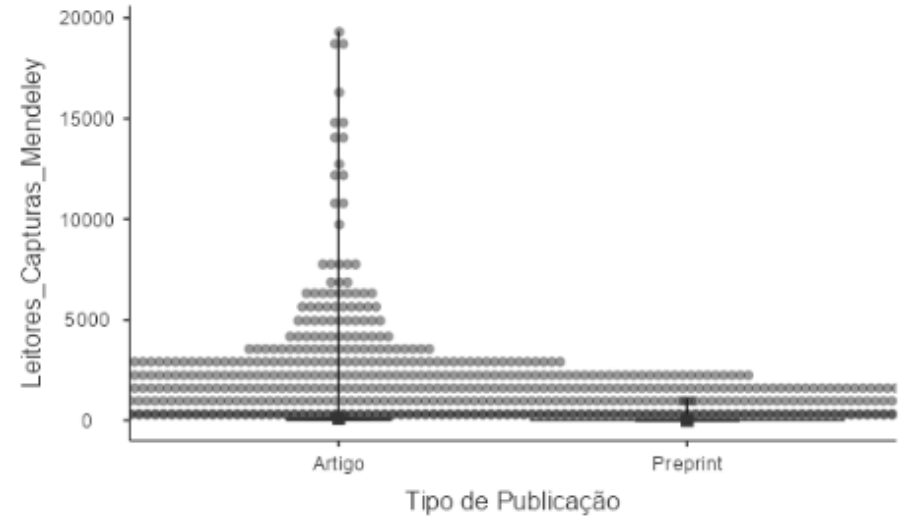
Fonte: dados da pesquisa

Gráfico 4 - Histograma de leitores/capturas por tipo de publicação (2020-2021)



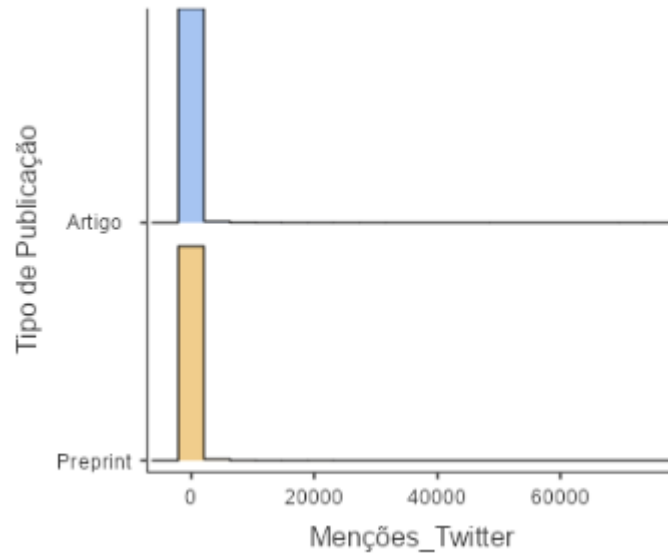
Fonte: dados da pesquisa

Gráfico 7- Dispersão dos dados de leitores/capturas por tipo de publicação



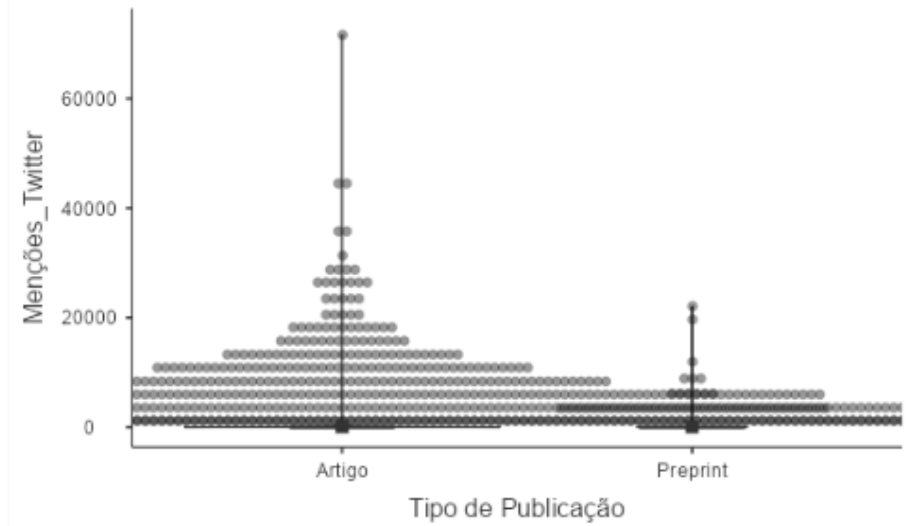
Fonte: dados da pesquisa

Gráfico 5 - Histograma menções no Twitter por tipo de publicação (2020-2021)



Fonte: dados da pesquisa

Gráfico 8- Dispersão dos dados de menções no Twitter por tipo de publicação



Fonte: dados da pesquisa

Quadro 9 - Medidas de tendência central (variáveis quantitativas por tipo de publicação)

Variáveis quantitativas	Tipo de Publicação	Média	Mediana	Moda
Citações	Artigo	33.01	5	0.0
	<i>Preprint</i>	7.98	1	0.0
Leitores/Capturas	Artigo	109.15	53	0.0
	<i>Preprint</i>	4.44	0	0.0
Menções Twitter	Artigo	93.10	3	0.0
	<i>Preprint</i>	71.78	4	0.0

Fonte: dados da pesquisa

Quadro 10 - Medidas de forma (variáveis quantitativas por tipo de publicação)

Variáveis quantitativas	Tipo de Publicação	Assimetria		Curtose	
		Assimetria	SE	Curtose	SE
Citações	Artigo	43.2	0.00848	2935	0.0170
	<i>Preprint</i>	12.5	0.03032	261	0.0606
Leitores/Capturas	Artigo	25.4	0.00848	1121	0.0170
	<i>Preprint</i>	18.8	0.03032	529	0.0606
Menções Twitter	Artigo	27.3	0.00848	1251	0.0170
	<i>Preprint</i>	22.5	0.03032	712	0.0606

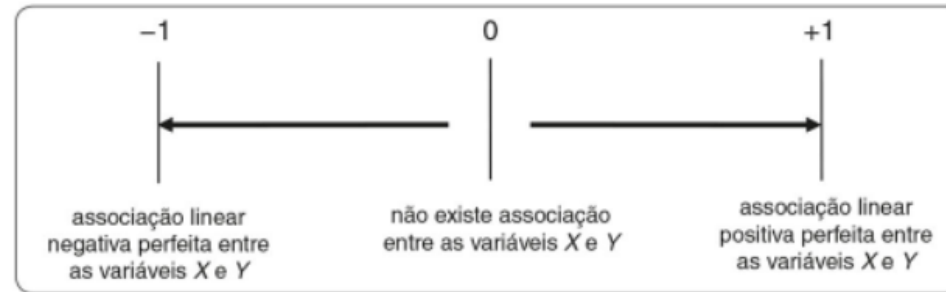
Fonte: dados da pesquisa

Ainda para analisar a distribuição das variáveis, no Jamovi, foram calculadas as medidas de forma das variáveis: assimetria (*skewness*) e curtose (*kurtosis*), as quais viabilizam a identificação da forma da distribuição em torno da média (**Quadro 10**). Como curtose, entende-se o grau de achatamento da curva de uma distribuição de frequências (FAVERO; BELFIORE, 2017).

As análises de distribuição (**Quadro 9, Gráficos 3, 4 e 5**) permitiram determinar que as variáveis quantitativas da pesquisa apresentam distribuições assimétricas, fato que indica a aplicação de análises do tipo não-paramétricas ao estudo, de modo que o coeficiente de correlação mais adequado para ser calculado é o de Spearman (r_s). Esse é caracterizado como uma medida de associação entre duas variáveis quantitativas e é calculado a partir da equação expressa abaixo, cujos valores obtidos como resultados podem variar entre -1 e 1 (**Figura 12**):

$$r_{sp} = 1 - \frac{6 \sum_{k=1}^n d_k^2}{n \cdot (n^2 - 1)}$$

Figura 12 - Interpretação do intervalo de valores possíveis para o coeficiente de correlação de Spearman



Fonte: Favero e Belfiore (2017, p. 119)

A análise e a interpretação mais detalhada dos valores obtidos para o coeficiente de correlação podem ser realizada com base nas informações dispostas no **Quadro 11** abaixo, que relaciona o valor de r_s com a respectiva intensidade da correlação

Quadro 11 - Escala de interpretação para os valores do coeficiente de correlação de Spearman (r_s)

Valor do coeficiente de correlação de Spearman (r_s)	Força/ Intensidade da correlação	Direção
-1,0 a -0,9	Muito forte ou perfeita	Negativa
-0,9 a -0,7	Forte	Negativa
-0,7 a -0,4	Moderada	Negativa
-0,4 a -0,2	Fraca	Negativa
-0,2 a 0	Insignificante	Negativa
0 a 0,2	Insignificante	Positiva

0,2 a 0,4	Fraca	Positiva
0,4 a 0,7	Moderada	Positiva
0,7 a 0,9	Forte	Positiva
0,9 a 1,0	Muito forte ou perfeita	Positiva

Fonte: elaborado pela autora com base em [Navarro e Foxcroft \(2019\)](#) e Favero e Belfiore (2017).

Quadro 12 - Síntese dos procedimentos de análise conforme objetivos específicos da pesquisa

Objetivos específicos	Fontes de informação para coleta	Ferramenta de análise	Tipo de análise
OE1 Comparar as variações entre o acesso aberto e os impactos bibliométrico de citação e altmétricos (menções e leitores/capturas) em artigos e preprints sobre COVID-19;	Dimensions Mendeley Altmetric.com	Excel Jamovi 2.2.5	Análise descritiva Observação de distribuição de frequência das variáveis; Geração de tabelas e gráficos nas ferramentas de análise utilizadas
OE2 Identificar a relação entre os indicadores de citações, menções e leitores/capturas e os subtipos de acesso (verde, bronze, dourado, híbrido e fechado);	Dimensions Mendeley Altmetric.com	Excel Jamovi 2.2.5	Análise descritiva e comparativa Observação de distribuição de frequência das variáveis; Geração de tabelas e gráficos nas ferramentas de análise utilizadas
OE3 Mensurar correlações entre indicadores bibliométricos de citação e altmétricos em artigos e <i>preprints</i> ;	Dimensions Mendeley Altmetric.com	Excel Jamovi 2.2.5	Análise de correlação Cálculo de correlação do coeficiente de Spearman; Matrizes de correlação

Fonte: elaborado pela autora.

3.5 LIMITAÇÕES DA PESQUISA

Em virtude das decisões tomadas no desenvolvimento do estudo, bem como em decorrência dos recortes e delimitações realizados nas etapas de coleta, tratamento e análise dos dados descritas anteriormente (seções 3.3 e 3.4, respectivamente), a pesquisa apresentou limitações, as quais encontram-se pontuadas a seguir:

- a) Fontes de informações utilizadas para coleta de dados: a escolha pelas fontes de informação utilizadas implica na definição de um conjunto restrito de informações a serem analisadas. Mesmo que o presente estudo tenha optado por investigar dados oriundos de diferentes fontes (vide **Quadro 6**) e, tendo em vista que a Dimensions tem características de uma base de dados abrangente, é sabido que nenhuma fonte possui a cobertura total de publicações sobre uma temática;
- b) Recorte temporal definido: trata-se de um período de tempo sucinto, que compreende apenas dois anos, 2020 e 2021, entretanto foi delimitado em virtude das características associadas à temática do estudo. Entende-se, também, que o intervalo temporal definido pode influenciar nos valores de acúmulo de impacto bibliométrico de citação e altmétrico, visto que, em outras datas as mesmas publicações podem registrar valores diferentes, sejam esses mais baixos ou mais altos;
- c) Uso do DOI como chave principal para a coleta de dados altmétricos: a necessidade de adotar um identificador único a exemplo do DOI como a chave principal para coleta de dados implica na automática exclusão de publicações que não possuem tal identificador. Assim, observa-se uma limitação quanto à diminuição da quantidade de publicações com potencial de análise pela pesquisa. Entretanto, é reconhecida a dificuldade atual em se coletar, de forma automatizada, dados, principalmente altmétricos, de publicações que não possuem nenhum tipo de link direto (URL, DOI, PMID) ([HAUSTEIN et al., 2014](#));
- d) Classificação dos artigos a partir dos dados da Dimensions, que utiliza as informações do Unpaywall e DOAJ: a categorização adotada pelas fontes mencionadas decide por um subtipo de acesso em detrimento de outro ([ORR, 2021](#)) e não engloba toda a gama de subtipos de acesso aberto existentes, gerando limitações relacionadas ao estudo dos valores identificados para as variáveis de tipo e subtipo de acesso (para mais detalhes, vide as informações apresentadas na seção **4.2**, na subseção que trata do OE2).

4 RESULTADOS E DISCUSSÕES

O presente capítulo tem por intuito a construção do relato e o detalhamento dos resultados de pesquisa fruto das análises dos dados efetuadas de acordo com as informações e procedimentos declarados na metodologia deste estudo. Inicialmente, é construída uma descrição mais ampla sobre os dados e, em um segundo momento, apresenta-se os resultados estruturados em três subseções, as quais foram definidas com base nos objetivos da dissertação.

4.1 DESCRIÇÃO GERAL DOS DADOS

Na busca pela identificação do impacto bibliométrico e altmétrico em publicações de acesso aberto na temática COVID-19, a presente pesquisa analisou dados quantitativos de citações, menções e capturas/leitores relacionados a 83.477 (92,7%) artigos e 6.523 (7,3%) *preprints*, que compõem a amostra total de 90.000 publicações. Essas encontram-se distribuídas entre os anos de 2020 e 2021, período que representa os dois primeiros anos completos de vigência do estado pandêmico declarado em decorrência do alastramento do novo coronavírus e suas respectivas variantes pelo mundo.

A observação dos dados revelou que 60,3% das publicações, somando os artigos e *preprints*, foram publicizados em 2020, enquanto que os 39,7% restantes da amostra foram em 2021 (**Quadro 13**). Nota-se que o primeiro ano do intervalo temporal considerado, registrou um maior quantitativo de publicações que o segundo. De acordo com Zhang (2020), o incremento acentuado no volume da produção científica, logo após a declaração de uma ESPII pela OMS, é um comportamento que também ficou evidente em surtos de doenças causadas por outros vírus, como Ebola, H1N1, Zika e SARS.

Além da possibilidade de desaceleração natural da produção científica sobre temáticas relacionadas às emergências de saúde pública com o passar do tempo (ZHANG, 2020), é relevante considerar que o valor mais baixo de artigos e *preprints* identificados em 2021 (39,7%) pode se caracterizar como um desdobramento associado à data em que foi realizada a coleta de dados do presente estudo, fevereiro e março de 2022. Tendo em vista que os procedimentos de registro e indexação de publicações em bases de dados podem sofrer atrasos decorrentes das logísticas e dos fluxos de processamento adotados pelas fontes de informação, no momento da coleta. Assim, publicações de 2021 ainda não registradas na base consultada (Dimensions) não puderam ser recuperadas.

Quadro 13 - Quantitativo de artigos e preprints sobre COVID-19 recuperados na base Dimensions (2020-2021)

Ano de publicação	Artigos	Preprints	Total
2020	49.872	4.390	54.262 (60,3%)
2021	33.605	2.133	35.738 (39,7%)
Total	83.477 (92,7%)	6.523 (7,3%)	90.000

Fonte: dados da pesquisa

Autores como [Fraser et al. \(2021\)](#) e [Santos-d'Amorim \(2021\)](#) mencionam em seus trabalhos o destaque dado ao *preprint* no contexto de disseminação e comunicação dos resultados científicos sobre COVID-19. Segundo os resultados obtidos por [Fraser et al. \(2021\)](#), quando comparados os quantitativos totais de *preprints* sobre Ebola, Zika e COVID-19 presentes nos repositórios bioRxiv e medRxiv, os *preprints* relacionados à COVID-19 demonstraram uma proporção aproximadamente duas vezes maior de registros recuperados. Mesmo que a utilização de *preprints* para viabilizar a comunicação científica tenha ganhado mais espaço no contexto da temática investigada pela presente pesquisa, na amostra analisada, o tipo de publicação predominante são os artigos, que representam 92,7% do total (**Quadro 13**). Os dados de [Fraser et al. \(2021\)](#) também demonstram uma composição de amostra com a maior parte da produção científica sobre COVID-19 sendo de artigos (65%) em detrimento de *preprints* (35%).

As descrições iniciais apresentadas levam em conta as distribuições de frequência por ano e por tipo de publicação e subsidiam a interpretação dos resultados que serão expressos adiante, os quais encontram-se organizados com base nos objetivos específicos da dissertação.

4.2 RESULTADOS ORIENTADOS AOS OBJETIVOS ESPECÍFICOS

O conteúdo apresentado a seguir estrutura a apresentação de resultados alcançados com a investigação desenvolvida, bem como as análises associadas a eles, a partir dos três objetivos específicos da pesquisa.

4.2.1 OE1 Comparar as variações entre o acesso aberto e os impactos bibliométrico de citação e altmétricos (menções e leitores/capturas) em artigos e preprints sobre COVID-19

As hipóteses de OACA e OAAA existentes na literatura de acesso aberto consideram que as variações do tipo de acesso podem implicar em valores mais elevados para indicadores de citação, menção e leitores/capturas de publicações. Conforme já mencionado na seção **3.4.1 Definição e caracterização das variáveis da pesquisa**, para fins do presente estudo, o tipo de

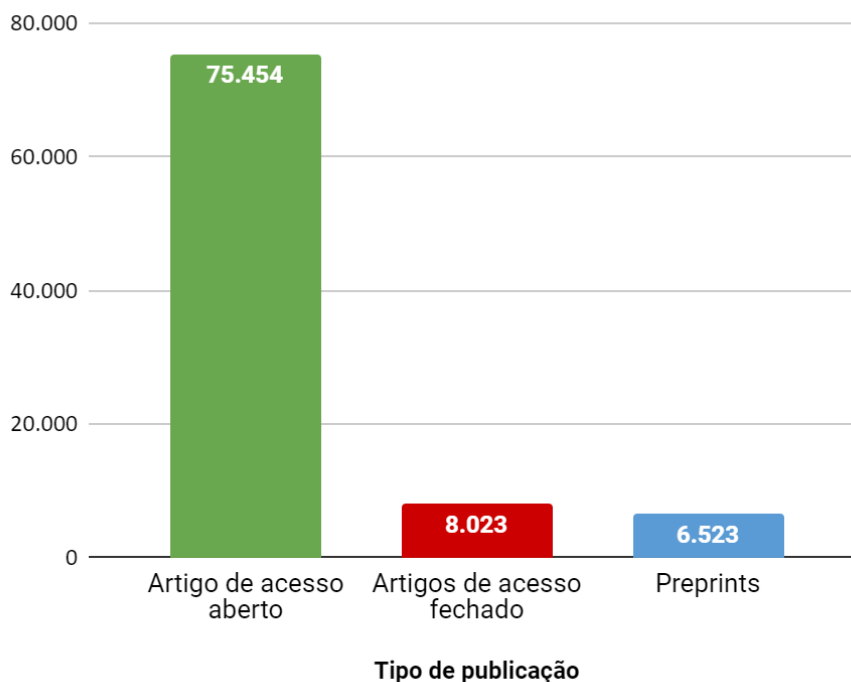
acesso é compreendido como o *status* mais amplo de disponibilização em que se encontra a publicação, podendo ser classificado como **aberto** ou **fechado**. Vale lembrar que a classificação do acesso das 90.000 publicações analisadas foi obtida a partir dos dados coletados pela base Dimensions, que utiliza o [Unpaywall](#) e o DOAJ como fontes para indicar se uma publicação é aberta ou fechada ([BODE et al., 2019](#)).

Tendo em vista o escopo da pesquisa, a variável tipo de publicação poderia assumir somente dois valores: artigos ou *preprints*. Os artigos, dependendo do lugar em que são disponibilizados e das licenças adotadas, podem receber a classificação de acesso tanto aberto como fechado. Todavia, os *preprints*, por se caracterizarem como documentos acessíveis em plataformas que os arquivam, recuperam e disponibilizam gratuitamente e diretamente *online*, são classificados somente como publicações de acesso aberto. Isto posto, a distribuição da amostra do estudo por tipo de acesso demonstra que 91,1% das publicações analisadas, reunindo tanto artigos como *preprints*, são de acesso aberto, enquanto que o acesso fechado representa 8,9% do total de artigos na temática de COVID-19 (**Gráfico 9**).

A composição da amostra presente no **Gráfico 9** revela expressiva presença de publicações em acesso aberto sobre COVID-19 (91,1%). [Araújo et al. \(2017\)](#), ao estudarem a produção científica sobre Zika vírus a partir de dados oriundos da base Scopus, observaram uma elevação no quantitativo de publicações disponibilizadas em acesso aberto sobre o tema a partir de 2016, mesmo ano em que a Organização Mundial da Saúde (OMS) fez um apelo a comunidade científica para divulgar os resultados de pesquisa relacionados ao vírus da Zika em acesso aberto.

Como mencionado na seção **2.1.1 Comunicação e produção científica em tempos de COVID-19**, após a declaração do estado pandêmico de COVID-19 pela OMS no início de 2020, instituições de fomento, sociedades acadêmicas e outros integrantes do contexto da comunicação científica se mobilizaram e divulgaram pedidos formais apelando aos cientistas pela divulgação dos resultados de pesquisa abertamente. Tais ações e recomendações podem ser um fator a ser levado em consideração quando se observa a variável tipo de acesso em produções científicas associadas aos anos iniciais de temáticas voltadas às ESPII, visto que podem influenciar a porcentagem de publicações em acesso aberto identificadas na amostra investigada.

Gráfico 9 - Quantidade de publicações sobre COVID-19 recuperados na base Dimensions (2020-2021) por tipo de acesso (aberto e fechado)



Fonte: dados da pesquisa

Tendo em vista o perfil de tipo de acesso acima exposto para as 90.000 publicações examinadas neste trabalho, adiante, serão elaboradas análises bibliométricas e altmétricas dos dados levando em consideração a classificação de acesso dos artigos e *preprints*.

4.2.1.1 Análise bibliométrica de citação por tipo de acesso

A concretização da citação ocorre quando, em seu conteúdo, um documento se refere a ou menciona uma outra publicação já existente ([SPINAK, 1996](#)). Assim, a frequência de citações registradas para uma publicação assinala a quantidade de vezes em que essa foi formalmente indicada em outros documentos. Os artigos sobre COVID-19 dos anos de 2020 e 2021 componentes da amostra verificada receberam, ao todo, 2.755.622 citações, enquanto que os *preprints* foram citados 52.063 vezes.

Tabela 1 - Indicadores de citação de artigos e preprints por tipo de acesso (aberto e fechado)

Tipo de publicação	Tipo de acesso	Publicações totais (PT)	Publicações citadas (PC) Citações > 0	%PC Citações > 0	Citações (C)	Média C/PC	Desvio padrão C	Mediana
Artigo	Aberto	75.454	62.105	82,3%	2.679.894	43,15	208.4	6
	Fechado	8.023	5.011	62,5%	75.728	15,11	53.6	1
Preprint	Aberto	6.523	3.862	59,2%	52.063	13,48	28.9	1

Fonte: dados da pesquisa

Legenda: PT= quantidade total de artigos ou preprints identificados na amostra; PC= quantidade de publicações que receberam, ao menos, uma citação, ou seja, registraram valores de citação maiores que 0; %PC= porcentagem de publicações que receberam, ao menos, uma citação em relação ao total de publicações ($PC/PT*100$); C= somatória das citações recebidas; **Média C/PC** = média de citações por publicações citadas (são consideradas somente as publicações que registraram citações > 0); **Desvio padrão C**= valor que mensura a dispersão dos dados em relação a média. **Mediana** = mediana de citações.

Os dados presentes na **Tabela 1** demonstram que nem todas as publicações do conjunto investigado possuem citações associadas aos DOIs correspondentes. Entre os 6.523 *preprints*, 3.862 foram citados ao menos uma vez. E, dos 83.477 artigos da amostra, 67.116 receberam ao menos uma citação. Quando examinados sob o ponto de vista do tipo de acesso, os artigos classificados como de acesso aberto exibiram uma maior porcentagem de publicações citadas (82,3%) que os acesso fechado (62,5%).

Autores como [Wang, Liu, Mao e Fang \(2015\)](#) interpretam que o número de publicações com ao menos uma citação dividido pelo conjunto total de publicações analisadas gera um indicador que representa o potencial de citação de um documento. No presente estudo, o indicador que calcula tal potencial é o %PC. A partir da **Tabela 1**, nota-se que, dentre as publicações de acesso aberto, os artigos demonstraram maior potencial de citação (82,3%) que os *preprints* (59,2%). Já, ao comparar os valores de %PC de publicações de acesso aberto *versus* fechado, os artigos de acesso aberto também obtiveram maior destaque.

A verificação dos valores resultantes para o indicador média de citações (**média C/PC**) expressa que o artigo de acesso aberto foi a tipologia de publicação e acesso que recebeu, em média, mais citações (43,15). Por outro lado, os artigos de acesso fechado e os *preprints* demonstraram resultados de média de citações muito inferiores, 15,11 e 13,48 respectivamente (**Tabela 1**). Sendo a média a soma de todas as citações registradas dividida pelo número de

publicações citadas, essa medida de tendência central pode sofrer forte influência de valores atípicos, ou seja, pode ser influenciada pela grande variação existente entre os valores mínimo e máximo existentes na amostra. No conjunto de publicações na temática de COVID-19 há, por exemplo, artigos de acesso aberto que não receberam citações, enquanto outros foram citados 18.029 vezes. Nesse sentido, para complementar a análise bibliométrica de citação neste trabalho, além dos indicadores de tendência central (média e mediana de citações), foi calculado também um indicador que se caracteriza como uma medida de dispersão de dados, o desvio padrão de citações. Esse indica o grau de dispersão dos dados a partir da média e, quanto maior o valor do desvio obtido, maior é a dispersão existente entre a quantidade de vezes em que as publicações foram citadas.

A análise da **Tabela 1** evidencia que, assim como ocorreu nos resultados dos indicadores já mencionados, os artigos de acesso aberto possuem o maior desvio padrão de citações entre os valores encontrados (208.4) e esse expressa considerável diferença em relação aos números identificados para artigos de acesso fechado (53.6) e *preprints* (28.9). Logo, nota-se que, além de terem recebido em média mais citações, os artigos de acesso aberto da amostra estudada, também apresentaram o maior desvio entre o quantitativo de citações identificadas a partir da Dimensions.

Diferentemente da média, a mediana é um indicador que, nos resultados calculados, recebe menos influência de valores discrepantes. Agresti e Finlay (2012, p. 61) afirmam que “a média pode ser fortemente afetada pelos valores atípicos, enquanto a mediana não”. Dessa forma, tendo em vista a característica assimétrica de distribuição dos dados da amostra (vide **3.4.2 Procedimentos de análise**), foram calculadas, no módulo de análise descritiva do Jamovi, as medianas de citação dos artigos de acesso aberto e fechado e dos *preprints*. Os resultados encontrados para o indicador em questão demonstraram, mais uma vez, destaque das citações em artigos de acesso aberto (6) em detrimento dos outros tipos de publicação e acesso. Os artigos de acesso fechado e os *preprints* apresentaram equivalência na localização do ponto médio da quantidade de citações, ambos possuem mediana igual a 1 (**Tabela 1**).

A análise bibliométrica de citações, a partir do conjunto de valores apresentados como resultados associados aos indicadores evidenciados na **Tabela 1**, demonstra que os artigos de acesso aberto são a combinação entre as variáveis tipo de publicação e tipo de acesso que manifestou os maiores valores para todos os indicadores calculados.

4.2.1.2 Análise altmétrica de Leitores/capturas por tipo de acesso

Caracterizado como uma ferramenta gratuita de gerenciamento de referência, o Mendeley é um serviço coordenado pela empresa Elsevier que oferece possibilidades de armazenamento, organização, compartilhamento, destaque e citação de referências de forma *online* (ELSEVIER, 2022). Os usuários cadastrados no Mendeley são chamados de *readers*/leitores e, por meio da ferramenta, podem encontrar e salvar publicações de interesse em suas bibliografias pessoais. A quantidade de vezes em que as publicações são salvas pelos usuários no sistema gera os dados relativos ao que, no contexto da pesquisa, consideramos como quantitativo de leitores/capturas por publicação. Desse modo, após a coleta e tratamento desses dados altmétricos, foram calculados outros indicadores associados aos leitores/capturas, cujos resultados encontram-se sumarizados pela **Tabela 2**.

Tabela 2 - Indicadores de leitores/capturas de artigos e preprints por tipo de acesso (aberto e fechado)

Tipo de publicação	Tipo de acesso	Publicações totais (PT)	Publicações lidas/capturadas (PLC)	%PLC Leitores/capturas > 0	Leitores/capturas (LC)	Média LC/PLC	Desvio padrão LC	Mediana
Artigo	Aberto	75.454	72.989	96,7%	8.670.305	118,79	314,4	57
	Fechado	8.023	7.081	88,3%	441.424	62,33	129,3	27
Preprint	Aberto	6.523	838	12,8%	28.956	34,56	28,4	0

Fonte: dados da pesquisa

Legenda: PT= quantidade total de artigos ou preprints identificados na amostra; PLC= quantidade de publicações que registraram, ao menos, um leitor/captura. Ou seja, registraram valores de leitores/capturas maiores que 0; %PLC= porcentagem de publicações que registraram, ao menos, um leitor/captura em relação ao total de publicações; LC= somatória dos leitor/captura identificados; **Média LC/PLC** = média de leitores/capturas por publicações lidas/capturadas (são consideradas somente as publicações que registraram leitores/capturas > 0); **Desvio padrão LC**= valor que mensura a dispersão dos dados em relação a média de leitores/captura. **Mediana** = mediana de leitores/capturas.

Os primeiros indicadores integrantes da análise altmétrica, quantidade de publicações lidas/capturadas (PLC) e porcentagem de publicações com, ao menos, um leitor/captura (%PLC), buscam averiguar a questão de cobertura da ferramenta dentre a amostra do estudo. Ou seja, entender se o quantitativo de artigos e *preprints* na temática sobre COVID-19 encontrados no Mendeley é significativo. Os valores presentes na **Tabela 2** demonstram que a cobertura no gerenciador de referência em questão é positivamente representativa para artigos,

uma vez que 96,7% dos artigos de acesso aberto e 88,3% dos de acesso fechado foram capturados pelos usuários da ferramenta (leitores), no mínimo, uma vez. Por outro lado, a presença dos *preprints* no Mendeley se mostrou baixa, somente 12,8% dos 6.523 analisados registraram leitores/capturas associados. [Fraser et al. \(2020\)](#), ao avaliarem os vínculos existentes entre a variação dos tipos de publicações e o comportamento dos dados de citação e altmétricos provenientes da plataforma Altmetric.com, também identificaram elevada cobertura para artigos no Mendeley, de modo que 92% dos artigos do conjunto analisado pelos autores registraram atenção na referida ferramenta. Dessa forma, observa-se que o tipo de publicação artigo expressou maior potencial de obtenção de capturas (% PLC) a partir do Mendeley no presente estudo, havendo uma cobertura levemente superior para os publicados em acesso aberto.

Os resultados encontrados para média de publicações lidas/capturadas e para a mediana de leitores/capturas revelaram o destaque dos artigos de acesso aberto em relação às outras tipologias. Esses registraram, em média, 118,79 leitores/capturas por artigo, com considerável diferença para os artigos de acesso fechado que foram lidos/capturados, pela média LCP/PLC, 62,33 vezes.

A menor mediana obtida foi a de *preprints*, que com o resultado igual a 0 (**Tabela 2**), revela o elevado quantitativo de publicações desse tipo componentes da amostra que não registraram nenhum leitor/captura associados. Em contrapartida, os artigos de acesso aberto demonstraram elevado valor de mediana, 57, seguidos dos artigos de acesso fechado, em que o ponto médio, ou seja, a quantidade de leitores/capturas situada no centro dos valores da amostra, se localiza na posição 27. A mesma lógica se manteve para os resultados do desvio padrão LC, sendo a maior dispersão observada em relação à média registrada para os artigos de acesso aberto (314,4) e a menor para os *preprints* (28,4).

Diante dos aspectos indicados anteriormente, percebe-se que, assim como ocorreu na análise bibliométrica de citações, no contexto dos dados de leitores/capturas, os artigos de acesso aberto demonstraram vantagem significativa em relação aos artigos de acesso fechado e os *preprints*. Faz-se válido ressaltar, também, que a análise altmétrica dos dados provenientes do Mendeley revelou ampla cobertura da ferramenta com relação aos artigos, e baixa presença de *preprints* da amostra com leitores/capturas identificadas. A observação da alta cobertura de dados do Mendeley entre artigos também foi evidenciada em estudos como o de [Zahedi, Costas e Wouters \(2014\)](#), de modo que os autores destacam o potencial da referida ferramenta para o desenvolvimento de estudos altmétricos vinculados a publicações científicas do tipo artigo.

Na interpretação de [Maflahi e Thelwall \(2018\)](#) o Mendeley é uma ferramenta estruturada de modo que os usuários possam deixar a indicação de que leram ou que tem a intenção de fazer a leitura de artigos salvos nas bibliotecas de referências, fazendo com que os dados provenientes da ferramenta possam revelar um impacto altmétrico inicial da versão publicada de um artigo. Entretanto, tendo em vista que uma publicação pode apresentar mais de uma versão (e.g. preprints e versões de visualização antecipada disponibilizadas pelos editores no site oficial do periódico), [Maflahi e Thelwall \(2018, p. 164, tradução nossa\)](#) apontaram a dificuldade de controle de versionamento da publicação como uma limitação da ferramenta, destacando que o Mendeley “[...] não deixa claro como transfere os leitores entre diferentes versões de artigos para associar preprints às versões publicadas do mesmo artigo e se atualiza automaticamente os metadados para alguns editores”³¹. Assim, depreende-se que tal característica da ferramenta pode ter contribuído com a baixa cobertura de leitores/capturas entre os *preprints* analisados pelo presente estudo.

4.2.1.3 Análise altmétrica de menções no Twitter por tipo de acesso

Estudos como o de [Thelwall et al. \(2013\)](#) interpretam dados de menções no Twitter como um indicador altmétrico promissor para verificar aspectos de popularidade das publicações, indicando-as como um dos indicadores altmétricos com maior cobertura entre artigos indexados na PubMed. Tal entendimento também se aplica ao contexto da presente pesquisa, visto que, para todos os tipos de publicação e acesso da amostra analisada foram identificados valores superiores a 50% para o indicador de porcentagem de publicações mencionadas (PMT), sendo 76,4% em artigos de acesso aberto, 55,6% em artigos de acesso fechado e 66,6% para *preprints*. Esses resultados indicam a porcentagem de publicações que registram quantidade de menções no Twitter diferente de 0 (**Tabela 3**).

A questão da alta cobertura do Twitter também pode ser observada a partir dos valores da soma da quantidade de menções identificadas na amostra da pesquisa. Uma vez comparados os quantitativos de citações (C) (**Tabela 1**) e leitores/capturas (LC) (**Tabela 2**), com os de menções no Twitter, observa-se, pela **Tabela 3**, que nesta foram encontrados valores maiores tanto para artigos de acesso fechado (585.623) quanto para *preprints* (468.241). Para artigos de acesso aberto, foi registrado um valor expressivo de 7.185.771 menções, que não superam o

³¹ “[...] it is not clear how it transfers readers between different versions of articles to associate preprints with the published versions of the same article and if it automatically updates the metadata for some publishers” ([MAFLAHI; THELWALL, 2018](#), p. 164)

valor total de leitores/capturas (8.670.305 - **Tabela 2**) registrado para a mesma tipologia de publicação e acesso.

Tabela 3 - Indicadores de menções no Twitter de artigos e preprints por tipo de acesso (aberto e fechado)

Tipo de publicação	Tipo de acesso	Publicações totais (PT)	Publicações mencionadas (PMT)	%PMT Menções > 0	Menções no Twitter (MT)	Média MT/PMT	Desvio padrão MT	Mediana
Artigo	Aberto	75.454	57.637	76,4%	7.185.771	124,7	849	3
	Fechado	8.023	4.464	55,6%	585.623	131,2	704	1
Preprint	Aberto	6.523	4.342	66,6%	468.241	107,8	552	4

Fonte: dados da pesquisa

Legenda: **PT**= quantidade total de artigos ou preprints identificados na amostra; **PMT**= quantidade de publicações que registraram, ao menos, uma menção no Twitter. Ou seja, registraram valores de leitores/capturas maiores que 0; **%PMT**= porcentagem de publicações que registraram, ao menos, uma menção no Twitter em relação ao total de publicações; **MT**= somatória das menções no Twitter; **MT/PMT** = média de menções no Twitter por publicações mencionadas (são consideradas somente as publicações que registraram menções > 0); **Desvio padrão MT**= valor que mensura a dispersão dos dados em relação a média de menções no Twitter; **Mediana** = mediana de menções no Twitter.

Diferentemente do que foi evidenciado pelas análises bibliométrica de citações e altmétrica de leitores/capturas apresentadas anteriormente, os artigos de acesso aberto não demonstraram vantagens para todos os indicadores de menção no Twitter calculados a partir dos dados da pesquisa. A verificação da **Tabela 3** expressa que, dentre as publicações mencionadas do Twitter, as com a maior média MT/PMT são os artigos de acesso fechado (131,2), seguidos pelos de acesso aberto (124,7) e, por fim, os *preprints* (107,8). Dessa forma, observa-se, em média, uma vantagem dos artigos de acesso fechado em relação aos de acesso aberto quanto aos dados de atenção das publicações no Twitter.

Ao investigarem o uso do Facebook e do Twitter por 25 periódicos científicos revisados pelos pares na área de medicina geral, [Kamel Boulos e Anderson \(2014\)](#) constataram que as revistas da amostra com as melhores estratégias de uso das duas mídias sociais mencionadas para a divulgação das publicações eram de acesso fechado e exigiam o pagamento dos leitores para disponibilização do texto completo. Destarte, os resultados encontrados pelo presente estudo e por [Kamel Boulos e Anderson \(2014\)](#) indicam uma possibilidade de existência de

estruturas mais robustas de divulgação das publicações em mídias sociais entre as editoras de acesso fechado.

Pela comparação entre os resultados de média, mediana e desvio padrão das três categorias de publicação analisadas (**Tabela 3**), nota-se uma maior homogeneidade entre os valores que os observados entre os dados de citações e leitores/capturas. Mesmo sendo distintos, os resultados dos indicadores demonstram diferenças pequenas entre eles. A maior mediana registrada foi para os *preprints* (4). Esse tipo de publicação também recebeu destaque com relação ao potencial altmétrico expresso pelo indicador %PMT, visto que, dos 6.523 *preprints* componentes da amostra, 66,6% receberam alguma menção no Twitter, enquanto que apenas 12,8% foram lidos/capturados no Mendeley (**Tabela 2**) e 59,2% receberam ao menos uma citação (**Tabela 1**).

Diante da verificação das relações entre as variáveis tipo de acesso (aberto e fechado), tipo de publicação (artigos e *preprints*), citações, leitores/capturas e menções no Twitter por meio da análise dos resultados obtidos para os indicadores bibliométricos e altmétricos calculados e apresentados nas **Tabelas de 1 a 3**, a próxima subseção dará enfoque às diferenças e aproximações existentes entre os indicadores métricos levando em consideração o detalhamento subtipos de acesso.

4.2.2 OE2 Identificar a relação entre os indicadores de citações, menções e leitores/capturas e os subtipos de acesso (verde, bronze, dourado, híbrido e fechado)

Abordados com maior detalhamento na seção **2.2.1** do presente trabalho, os subtipos de acesso constituem uma classificação das publicações dos diferentes aspectos associados à disponibilização de uma publicação, como local de acesso (periódicos, sites de editoras, repositórios), características do periódico em que se encontra publicado o artigo (é de acesso aberto ou não?) e identificação de licenças de uso e compartilhamento do conteúdo. A relevância de tais questões para delimitação dos subtipos de acesso fica evidente na definição de acesso aberto destacada por Piwowar *et al.* (2018, p. 4, tradução nossa), a qual evidencia que “os artigos de acesso aberto são gratuitos para leitura online, seja no site do editor ou em um repositório de acesso aberto”.

A combinação entre os fatores mencionados anteriormente fez com que, ao longo da evolução do MAA, surgissem na literatura variados subtipos de acesso, sendo os principais deles enumerados e detalhados no **Quadro 4** da subseção **2.2.1**. Atualmente, por meio da

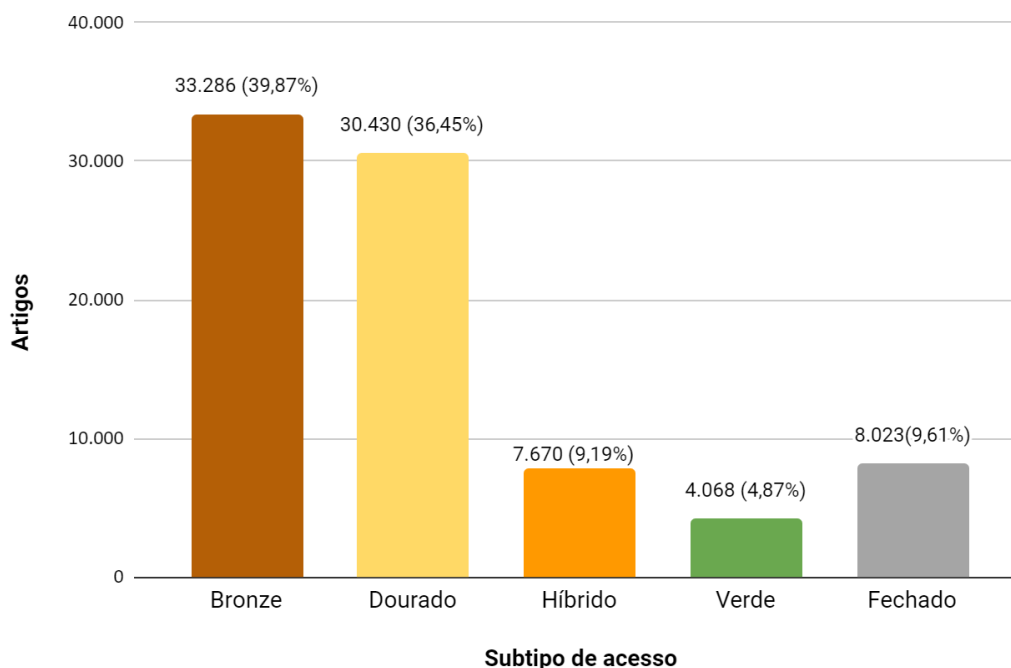
análise de metadados das publicações, algumas ferramentas conseguem, consideradas as devidas limitações, identificar evidências que possibilitam a classificação automática do subtipo de acesso de publicações com DOI, sendo a principal delas o Unpaywall ([ROBINSON-GARCIA; COSTAS, R.; VAN LEEUWEN, 2020](#)). Essa caracteriza-se como uma base de dados aberta, que realiza coleta de informações de acesso aberto das publicações em mais de 50 mil fontes, incluindo editores e repositórios, e é composta por mais de 32 milhões de artigos científicos ([UNPAYWALL, 2022](#)).

O Unpaywall realiza integração com diversas outras ferramentas, tendo como objetivo o fornecimento de dados sobre acesso aberto. Dentre as fontes com as quais o Unpaywall se integra estão as seguintes bases de dados: Dimensions, Lens, Scopus e Web of Science ([UNPAYWALL, 2022](#)). Ou seja, como já mencionado na metodologia deste estudo, os dados de tipo e subtipo de acesso obtidos para amostra da pesquisa foram coletados pela Dimensions com base na conexão existente com o Unpaywall.

Faz-se válido ressaltar que o mecanismo de identificação automática dos subtipos de acesso adotado pelo Unpaywall não consegue englobar toda a gama de subcategorias de acesso aberto apresentadas pelo **Quadro 4**. Tal ferramenta permite a classificação de um artigo nas seguintes categorias de acesso: bronze, dourado, híbrido, verde e fechado. Logo, subtipos como diamante/platina, cinza e preto não são identificados pela ferramenta, sendo essa uma das limitações do Unpaywall. Em decorrência de tal fato, a análise da relação entre os subtipos de acesso e os indicadores bibliométricos de citação e altmétricos será realizada somente para os artigos dos cinco subtipos identificados pelo Unpaywall e, conseqüentemente, pela Dimensions.

Outro ponto relevante de esclarecimento sobre a interpretação dos resultados alcançados para o segundo OE da pesquisa envolve a variável tipo de publicação. Tendo em vista que as publicações do tipo *preprint* são classificadas somente como acesso aberto verde, em decorrência das características de disponibilização e acesso atreladas a elas, a análise comparativa entre os valores de indicadores métricos com relação às aproximações e distanciamentos existentes pela variação dos subtipos de acesso foi desenvolvida apenas para os 83.477 artigos componentes da amostra.

Gráfico 10 - Quantitativo de artigos sobre COVID-19 recuperados na base Dimensions (2020-2021) distribuído por subtipo de acesso (bronze, dourado, híbrido, verde, fechado)



Fonte: dados da pesquisa

A distribuição de frequências representada pelo **Gráfico 10** revela a quantidade de artigos, e a respectiva porcentagem, dividida entre os cinco subtipos de acesso examinados pela presente pesquisa. Verifica-se que os artigos do subtipo bronze configuram a maior parte da amostra de artigos analisados (39,87%), seguidos dos dourados, com 36,45%. O subtipo que demonstrou menor evidência foi o verde (4,87%) e a parcela identificada como acesso híbrido expôs um tímido percentual em comparação com o todo (9,19%).

Uma elevada porcentagem de publicações do subtipo bronze entre o conjunto de artigos também foi verificada por [Price e Ozakan \(2021\)](#) ao estudarem a produção científica sobre COVID-19 elaborada pelo Imperial College entre janeiro e setembro de 2020. Nesse estudo, 30,5% da amostra era composta por artigos bronze, enquanto que os subtipos dourado, verde e híbrido somaram 66,1% das publicações. Os resultados de [Price e Ozakan \(2021\)](#) corroboram com a reflexão de um aspecto ora mencionado na seção **2.1.1** desta dissertação, de que: como uma forma de colaboração com os esforços globais empenhados no combate à pandemia, uma variedade de editores passaram a disponibilizar de forma gratuita e temporária o acesso aos resultados de pesquisa sobre a COVID-19. A disponibilização da publicação para leitura gratuita por um período de tempo sem a definição clara de uma licença de uso e reuso do conteúdo acessado são características do subtipo de acesso bronze. Nesse sentido, entende-se

que os apelos pela abertura temporária do acesso às descobertas científicas sobre o vírus causador da pandemia refletiram na grande porção de artigos classificados como bronze nessa temática nos primeiros anos do surto da doença (2020 e 2021).

O destaque do quantitativo de artigos bronze em contraste com o baixo percentual de publicações do subtipo verde também foi observado em Piwowar *et al.* (2019) durante a análise do *status* de acesso, a partir dos dados do Unpaywall, de uma amostra composta por 300.000 artigos. Na interpretação dos autores, a quantidade reduzida de publicações classificadas como verde se dá, pois “artigos autoarquivados só são contabilizados como verdes quando não há nenhuma outra versão hospedada e disponibilizada por uma editora, ou seja, o acesso aberto verde é às vezes é ‘sombreado’ por artigos dourados, bronze ou híbridos” (PIWOWAR *et al.*, 2019, p. 11, tradução nossa).

Caso, durante a classificação do subtipo de acesso da publicação, o mecanismo do Unpaywall identifique que um artigo encontra-se em mais de um local (representado pelo metadado *oa_location*), a ferramenta busca comparar e tratar os atributos do artigo de modo a definir o que é chamado de *best_oa_location* (em português: melhor localização de acesso aberto). Ou seja, por meio de uma lógica de classificação predefinida, o Unpaywall decide por um subtipo de acesso em detrimento de outro (ORR, 2021), de modo que uma mesma publicação não pode ser classificada ao mesmo tempo, por exemplo, como verde e híbrida, mesmo que esteja disponível em um repositório e em um periódico mediante a declaração de alguma licença aberta. Nesse caso hipotético, o Unpaywall classificaria o referido artigo como híbrido.

Segundo Orr (2021, online), para determinar a “melhor localização de acesso aberto” e classificar um artigo em um subtipo de acesso e não em outro, o Unpaywall atribui uma pontuação à publicação levando em consideração os seguintes critérios:

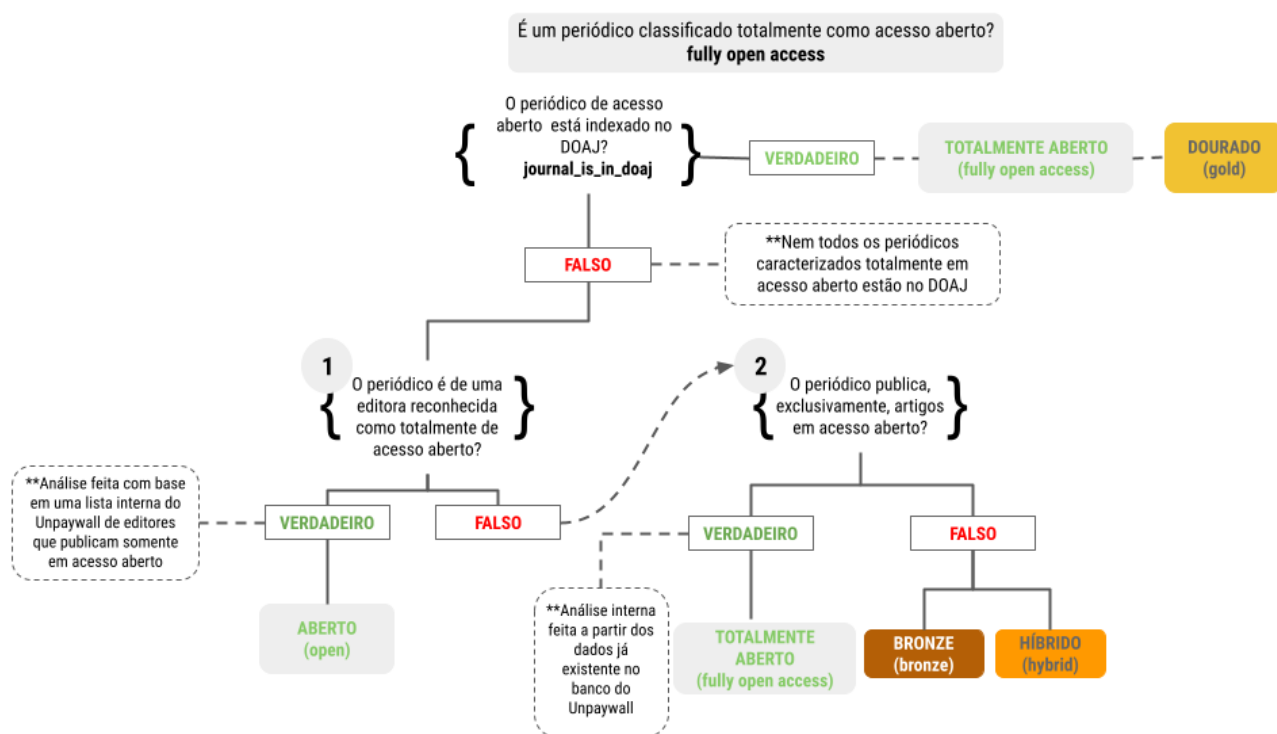
- a) Tipo de hospedagem da publicação (*host_type*): a presença de um artigo vinculado a uma editora é considerado melhor que repositório;
- b) Versão (*version*): a ordem decrescente de pontuação para as versões é: versão publicada > versão aceita > versão submetida;
- c) PDF identificado: se a publicação analisada apresenta link relacionado para o PDF, essa é melhor classificada do que os artigos sem arquivo de PDF relacionado;
- d) Em repositórios: registros identificados a partir do DOI são melhor classificados do que aqueles em que a identificação ocorre por meio do título;

e) Classificação de repositórios: aparentemente, existe uma hierarquia de classificação entre repositórios existentes, em que os mais tradicionais, a exemplo do arXiv e PubMed Central, ocupam melhores posições.

A verificação do conjunto de critérios acima, segundo [Orr \(2021, online\)](#), resulta na definição do que a ferramenta considera como a “melhor localização de acesso aberto” e isso, conseqüentemente, auxilia na atribuição do subtipo de acesso que será indicado ao artigo analisado. De acordo com [Priem \(2020, online\)](#), diferentemente do DOAJ, que tem por objetivo verificar aspectos de credibilidade e de legitimidade dos periódicos de acesso aberto, o foco do trabalho do Unpaywall está no monitoramento das possibilidades de acesso associadas às publicações.

Os apontamentos supramencionados demonstram que os resultados exibidos pelo **Gráfico 10** recebem influência direta da metodologia de classificação de acesso adotada pelo mecanismo de categorização automática da fonte de dados utilizada por este estudo. Dessa forma, a seguir, com o apoio das sistematizações expressas pelas **Figuras 13 e 14**, busca-se esclarecer o fluxo de identificação dos subtipos de acesso utilizado pelo Unpaywall ao verificar os DOIs das publicações.

Figura 13 - Fluxo de identificação do tipo de acesso de um periódico (completamente aberto/dourado ou fechado) pelo Unpaywall

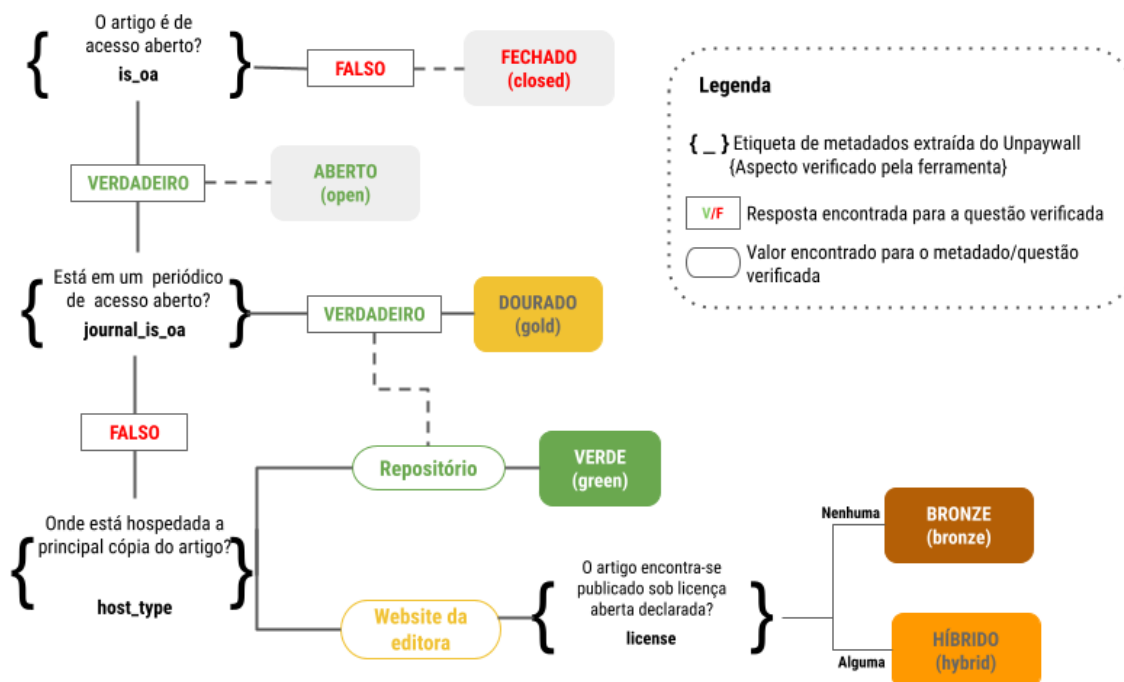


Fonte: elaborada pela autora com base em [Priem \(2020, online\)](#)

A coleta e a verificação automática de dados vinculados ao status de acesso de uma publicação a partir de um DOI demanda uma padronização dos metadados e dados a serem rastreados pelo mecanismo de consulta do Unpaywall. Assim, essa ferramenta possui um formato de dados e metadados³² próprio. Diversas são as possibilidades de consulta aos metadados, entretanto, para o contexto do presente estudo, buscou-se compreender o fluxo de classificação do tipo de acesso dos periódicos (**Figura 13**) e a lógica de definição dos subtipos de acesso para os artigos (**Figura 14**). As análises das **Figuras 13** e **14** demonstram que dois processos de classificação possuem relações entre si. O tipo de status de acesso atribuído a um periódico influencia na delimitação do subtipo de acesso que encontra-se publicado na referida revista científica.

³² Formato de dados e metadados do Unpaywall. Disponível em: <http://unpaywall.org/data-format#oa-location-object>. Acesso em: 13 maio 2022.

Figura 14 - Fluxo de identificação dos subtipos de acesso de um artigo (completamente aberto/dourado ou fechado) pelo Unpaywall



Fonte: elaborada pela autora com base em [Priem \(2021, online\)](#) e [Robinson-Garcia, Costas e Van Leeuwen \(2020, p. 4\)](#).

As informações da **Figura 14** evidenciam o caráter prioritário estabelecido pela ferramenta Unpaywall para os subtipos de acesso diretamente relacionados à publicação dos artigos em periódicos que fornecem, de alguma forma, a possibilidade de disponibilização de artigos em acesso aberto (bronze, dourado e híbrido) em detrimento dos repositórios. Tal aspecto justifica o perfil de subtipos de acesso da amostra de artigos da presente pesquisa representado pelo **Gráfico 8** e, além disso, caracteriza-se como uma das limitações deste estudo, que busca investigar a relação entre os indicadores de citações, menções e leitores/capturas e os subtipos de acesso (verde, bronze, dourado, híbrido e fechado) levando em consideração os dados fornecidos pela Dimensions que utiliza o Unpaywall como fonte de coleta.

A partir da caracterização supradelineada dos dados examinados para a variável subtipos de acesso, busca-se construir, a seguir, uma interpretação dos indicadores bibliométricos e altmétricos por meio da observação comparativa entre os valores registrados para cada categoria de acesso (**Tabela 4**).

Tabela 4 - Indicadores bibliométricos de citação e alométricos (leitores/capturas e menções no Twitter) de artigos por subtipo de acesso (bronze, dourado, híbrido, verde, fechado)

	Subtipo de acesso	Artigos totais (AT)	Artigos com 0 (C, LC ou MT) Citações= 0 Leitores/capturas=0 Menções=0	Artigos > 0 (C, LC ou MT) Citações> 0 Leitores/capturas>0 Menções>0	%Artigos (A) Citações> 0 Leitores/capturas> 0 Menções>0	Soma (C, LC ou MT)	Média (C, LC ou MT) Soma/Artigos citados, lidos ou mencionados	Desvio padrão (C, LC ou MT)	Mediana (C, LC ou MT)
Citações (C)	Bronze	33.286	4.117	29.169	87,6%	1.599.489	54,8	280,3	9
	Dourado	30.430	7.258	23.172	76,1%	584.862	25,2	74,7	4
	Híbrido	7.670	1.389	6.281	81,9%	359.560	57,2	232,1	7
	Verde	4.068	585	3.483	85,6%	135.983	39	126,2	8
	Fechado	8.023	3.012	5.011	62,5%	75.728	15,1	53,6	1
Leitores/ Capturas (LC)	Bronze	33.286	490	32.796	98,5%	4.671.822	142,4	391,9	70
	Dourado	30.430	1.570	28.860	94,8%	2.584.057	89,5	199,7	42
	Híbrido	7.670	371	7.299	95,2%	964.057	132,1	345,7	55

	Verde	4.068	34	4.034	99,2%	450.369	111,6	204,4	63
	Fechado	8.023	942	7.081	88,3%	441.424	62,3	129,3	27
Menções no Twitter (MT)	Bronze	33.286	6410	26.876	80,7%	3.905.153	145,3	965	4
	Dourado	30.430	9040	21.390	70,3%	1.597.481	74,7	540	2
	Híbrido	7.670	1557	6.113	79,7%	1.301.517	212,9	1.177	6
	Verde	4.068	810	3.258	80,1%	381.620	117,1	958	4
	Fechado	8.023	3.559	4.464	55,6%	585.623	131,2	704	1

Fonte: dados da pesquisa

Legenda: **AT**= quantidade total de artigos identificados na amostra por subtipo de acesso; **Artigos com citações, leitores/capturas ou menções=0** representam a quantidade de artigos que não registraram nenhuma atenção bibliométrica ou alométrica a partir do DOI; **Artigos com citações, leitores/capturas ou menções>0** = quantidade de artigos que registraram, ao menos, uma atenção bibliométrica ou alométrica. Ou seja, registraram valores maiores que 0; **%A**= porcentagem de artigos que receberam alguma atenção em relação ao total; **Soma**= somatória de todas as citações, leitores/capturas ou menções no Twitter identificadas entre o conjunto de artigos da amostra; **Média** = média de de citações, leitores/capturas ou menções por publicações com alguma atenção registrada (são consideradas somente as publicações que registraram atenção > 0); **Desvio padrão**= valor que mensura a dispersão dos dados em relação a média. **Mediana** = citações ou leitores/capturas ou menções.

A questão da cobertura é um dos primeiros aspectos que podem ser averiguados a partir dos dados da **Tabela 4**. Consideradas as citações coletadas por meio da Dimensions, os leitores/capturas registrados no Mendeley e as menções no Twitter, qual delas demonstrou maior homogeneidade de cobertura em todos os subtipos de acesso? Os valores resultantes para o indicador %A realçam a baixa variação entre os subtipos da porcentagem de artigos que foram lidos/capturados, no mínimo uma vez, na ferramenta de gerenciamento de referências. Por outro lado, o conjunto mais heterogêneo quanto à cobertura é o de dados altmétricos provenientes do Twitter. Enquanto 80,7% do total de artigos do subtipo bronze receberam alguma menção na referida mídia social, pouco mais da metade (55,6%) dos 8.023 artigos de acesso fechado registraram alguma menção. No que se refere às citações, foram identificados quantitativos de artigos com valores de citação maiores que zero superiores a 75% para todos os subtipos de acesso.

Diante da **Tabela 4**, nota-se que, entre os cinco subtipos de acesso estudados por esta pesquisa, o bronze foi o que registrou os maiores potenciais de citação e menções no Twitter, o que significa que, dos 33.286 artigos bronze componentes na amostra examinada, 87,6% receberam alguma citação e 80,7% registraram menção em pelo menos um tweet publicado no Twitter. Esses expressivos valores de cobertura de citações e menções no Twitter recebidas pelos artigos do tipo bronze podem, de alguma maneira, ser consequência dos periódicos e das editoras responsáveis pelas publicações desses artigos. Ou seja, levanta-se como uma possibilidade que os artigos sobre COVID-19 do subtipo bronze tenham sido publicados em periódicos com maior reconhecimento entre os pares e, no período de 2020 e 2021, tenham sido temporariamente disponibilizados pelas editoras comerciais para a leitura gratuita da comunidade acadêmica, de modo que tais fatos possam ter influenciado nos elevados valores encontrados para o indicador %A em citações e menções. Entretanto, tendo em vista que as variáveis de editora e periódicos não fazem parte do foco de análise da presente pesquisa, os pontos anteriormente levantados funcionam somente como tópicos hipotéticos da discussão, fazendo com que, para a confirmação de tais possibilidades, sejam necessários estudos futuros sobre os artigos da categoria de acesso bronze associados às editoras e revistas científicas em que são publicados.

Para os dados de %A em leitores/capturas, o subtipo de acesso com o maior potencial de ser lido/capturado por usuários do Mendeley foi o verde. Já os artigos classificados como fechados expuseram os menores percentuais de publicações efetivamente citadas (62,5%),

lidas/capturadas (88,3%) e mencionadas (55,6%). Logo, quanto à cobertura, os subtipos de acesso aberto demonstraram vantagens em relação ao acesso fechado.

Os indicadores de média foram calculados levando-se em consideração a divisão dos valores dispostos na coluna de soma da **Tabela 4** pelos dados de artigos (AT). Os resultados demonstram que os subtipos bronze e híbrido evidenciaram ser, em média, os dois mais citados, lidos/capturados e mencionados. As posições de destaque entre eles se alteram dependendo da métrica analisada, mas os artigos classificados como híbridos e bronze se revezam nas posições de primeiro e segundo lugar. Para média de citações e leitores/capturas, os artigos de acesso fechado, mais uma vez, registraram os menores valores para os indicadores analisados. Por outro lado, no âmbito dos dados altmétricos de menções no Twitter, o subtipo de acesso com a média menos expressiva foi o dourado.

Os resultados para o indicador desvio padrão, responsável por mensurar a dispersão dos dados bibliométricos e altmétricos em relação aos valores médios identificados, revelaram que os mesmos subtipos de acesso (híbrido e bronze) que registraram as maiores médias também evidenciaram os resultados mais significativos de dispersão. Ou seja, apesar de terem sido, em média mais citados, capturados e mencionados, os artigos bronze e híbridos são também, dentre as cinco subtipos, os que possuem a maior variabilidade entre os valores de citação, leitores/captura e menções identificados para a amostra de artigos na temática de COVID-19. Os dados da **Tabela 4** demonstram que os menores desvios padrão registrados para as citações e leitores/capturas ocorrem entre os artigos fechados (53,6 e 129,3 respectivamente). O comportamento de dispersão entre os subtipos de acesso se modifica entre os dados de menções do Twitter, em que os artigos dourados são os que registraram o menor desvio padrão (540), seguidos pelos artigos fechados, que demonstraram o segundo menor desvio de menções no Twitter (704) (**Tabela 4**).

Por fim, a apuração das medianas, indicador que recebe menos influência dos valores discrepantes na distribuição dos dados (também conhecidos como *outliers*), identificou resultados muito próximos entre as medianas de citação para os artigos bronze (9), verde (8) e híbrido (7). O destaque para os dois primeiros subtipos mencionados (bronze e verde) também foi registrado entre os valores de mediana de leitores/capturas. A **Tabela 4** mostra que os resultados para esse indicador foram elevados em comparação com as outras métricas: bronze (70), verde (63), híbrido (55), dourado (42), fechado (27), fato que permite reafirmar uma maior cobertura do Mendeley em detrimento do Twitter e Dimensions para os artigos da amostra analisada. Mais uma vez, os artigos de acesso fechado registraram os menores valores para os

cálculos de mediana de citações, leitor/capturas e menções no Twitter. Assim, é possível afirmar, a partir da análise comparativa entre indicadores bibliométricos e altmétricos, que os subtipos bronze, dourado, híbrido e verde demonstraram vantagens, na maioria das situações de comparação simuladas pela pesquisa, em relação ao acesso fechado.

4.2.3 OE3 Mensurar correlações entre indicadores bibliométricos de citação e altmétricos em artigos e preprints

O foco do terceiro OE desta dissertação está situado na verificação da existência, bem como a força e a intensidade, de relações entre as variáveis bibliométricas (citações) e altmétricas (leitores/capturas e menções no Twitter) investigadas pela presente pesquisa.

Diante da classificação da natureza não paramétrica, com distribuição assimétrica, dos dados coletados (vide seção **3.4.2 Procedimentos de análise**), definiu-se o uso do coeficiente de correlação de Spearman (r_s) para efetuar a análise de relação entre as variáveis quantitativas do estudo. O coeficiente foi calculado por meio do *software* estatístico Jamovi, que também analisou a significância das correlações a partir do fornecimento do valor de p, responsável por indicar se os resultados de r_s são significativamente diferentes de 0.

Após a aplicação automática da fórmula de Spearman ao conjunto de dados bibliométricos e altmétricos dos artigos e *preprints* na temática de COVID-19, foram obtidos os resultados dispostos nas **Tabelas 5 e 6**. Faz-se válido ressaltar que, no contexto da presente pesquisa, a análise dos valores identificados para r_s teve como referencial interpretativo as informações apresentadas no **Quadro 11**, que encontra-se disposto nos **3.4.2 Procedimentos de análise** do presente trabalho.

Antes mesmo de avaliar a intensidade das correlações, é necessário verificar, por meio do valor de p, se essas são significativas. Valores de p menores que 0,05 ($p < 0,05$) indicam que o nível de significância estatística da amostra é superior a 95%. Para os dados da pesquisa (vide **Tabelas 5 e 6**), identificamos correlações significativas em todos os casos analisados, tanto para artigos como para *preprints* ($p < 0,001$).

Tabela 5 -Matriz de correlação das variáveis citações, leitores/captura e menções no Twitter para artigos

	Citações	Leitores/capturas no Mendeley	Menções no Twitter
Citações	—		
Leitores/capturas no Mendeley	0,836 ($p < 0,001$)	—	
Menções no Twitter	0,527 ($p < 0,001$)	0,454 ($p < 0,001$)	—

Fonte: dados da pesquisa

A matriz de correlação gerada pelo Jamovi para os dados de artigos (acesso aberto e fechado) está apresentada abaixo pela **Tabela 5**. Nota-se que os coeficientes obtidos entre as citações, as interações de usuários do Mendeley e do Twitter revelaram correlações positivas em todas as aproximações. A correlação mais forte identificada para artigos na temática de COVID-19 publicados entre os anos de 2020 e 2021 se deu entre os dados de citação e leitores/capturas ($r_s = 0,836$). As outras correlações analisadas entre citações e menções no Twitter ($r_s = 0,527$); e menções e leitores/capturas ($r_s = 0,454$) evidenciaram correlações positivas moderadas. A forte correlação encontrada entre dados de citação e alométricos de leitores/capturas ratifica uma tendência já identificada em estudos como os de [Maflahi e Thelwall \(2018\)](#), [Thelwall e Wilson \(2016\)](#) e [Zahedi, Costas e Wouters \(2014\)](#), que apontam a as contagens de leitores/capturas no Mendeley como os dados alométricos que demonstram o maior potencial de correlação com os indicadores de bibliométricos de citação, de modo que os primeiros, na perspectiva de [Maflahi e Thelwall \(2018\)](#), podem ser interpretados como relevantes indicadores de impacto inicial das publicações.

Outro aspecto importante a ser ressaltado no que tange à forte correlação observada entre as citações e leitores/capturas entre os artigos integrantes da amostra de artigos analisados pelo presente estudo é que a correlação não implica na afirmação de causalidade. Ou seja, um alto valor obtido para r_s demonstra que “[...] uma correlação positiva significativa indicaria que artigos altamente citados tendem a ser altamente lidos e vice-versa, mas não provaria que um causa o outro, nem que ambos medem de alguma forma a qualidade ou valor científico de um artigo” ([THELWALL; WILSON, 2016, p. 1966, tradução nossa](#)).

Tabela 6 -Matriz de correlação das variáveis citações, leitores/captura e menções no Twitter para *preprints*

	Citações	Leitores/capturas no Mendeley	Menções no Twitter
Citações	—		
Leitores/capturas no Mendeley	0,126 (p < 0,001)	—	
Menções no Twitter	0,544 (p < 0,001)	-0,079 (p < 0,001)	—

Fonte: dados da pesquisa.

Os resultados da análise de correlação para as publicações do tipo *preprint* (**Tabela 6**) demonstram que a maior correlação identificada, caracterizada como positiva moderada, se dá entre os dados bibliométricos de citação e os altmétricos de menção no Twitter ($r_s = 0,544$), esse resultado se diferencia da tendência, mencionada anteriormente, existente entre artigos, em que os dados altmétricos oriundos do Mendeley são os que demonstram correlações mais fortes com as citações. Assim, a mudança no tipo de publicação demonstrou diferenças no que tange a intensidade das correlações entre os indicadores de citação e os indicadores altmétricos, indicando que, entre os *preprints*, além de maior cobertura, as menções no Twitter demonstram uma correlação mais intensa com as citações do que os leitores/capturas.

[Wang, Glänzel e Chen \(2020\)](#), ao pesquisarem sobre impacto de citação e altmétrico em *preprints* da área de Biblioteconomia e Ciência da Informação, encontraram vantagens de impacto no Twitter para artigos que também possuíam a versão *preprint* disponibilizada no repositório arXiv. Assim como efetuado pelo presente estudo, [Wang, Glänzel e Chen \(2020\)](#) coletaram os dados de menções no Twitter a partir da ferramenta Altmetric.com. Os autores apontam que, aparentemente, os *links* das versões *preprint* das publicações tendem a ser mais mencionadas pelos usuários das mídias sociais do que os links das versões presentes nos sites das editoras, visto que as primeiras disponibilizam o texto completo do documento. Além disso, os mesmos autores pontuam que, pela maior disponibilidade, os *preprints* presentes no arXiv possuem maior probabilidade de serem identificados pelos mecanismos de rastreamento da plataforma Altmetric.com.

Faz-se válido pontuar que a Dimensions, principal fonte de dados deste estudo, coleta *preprints* de diferentes repositórios (medRxiv, bioRxiv, ChemRxiv, SciELO *preprints*, entre outros), não se restringindo somente ao arXiv. Logo, entende-se que as observações identificadas por [Wang, Glänzel e Chen \(2020\)](#) são apenas um indicativo interpretativo para os

dados de menções no Twitter relacionados aos *preprints*. Entretanto, faz-se válido apontar como possibilidade para análises futuras a consideração da variável de repositório de *preprints* para compreender melhor as variações da atenção alométrica sobre um mesmo tema em repositórios distintos.

As demais correlações verificadas, entre citações e leitores/capturas Mendeley ($r_s = 0,126$) e as interações registradas tanto no Twitter quanto na ferramenta de gerenciamento de referências ($r_s = -0,079$), demonstraram correlações caracterizadas como insignificantes, sendo a primeira na direção positiva e a segunda, negativa.

Diante do exposto, conclui-se que os resultados apresentados nas matrizes de correlação tanto de artigos como de *preprints* indicam que todas as correlações encontradas são significativas ($p < 0,05$); em nenhum dos casos calculados foram encontradas correlações fracas ou perfeitas; e, observou-se que a variação do tipo de publicação pode influenciar os valores obtidos para o coeficiente de correlação de Spearman.

5 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Períodos de emergências em saúde pública caracterizam-se como mais um dos grandes desafios impostos à ciência e aos cientistas e, com a COVID-19, não foi diferente. Os esforços no combate a pandemia, declarada pela OMS em 2020, demandaram mudanças e adaptações no fluxo formal de publicação dos resultados de pesquisa, fato que impulsionou o surgimento de novas questões e abordagens no âmbito da comunicação científica e seus temas correlatos, como o acesso aberto e os estudos métricos da informação. Nessa conjuntura, foi proposta uma investigação, materializada por esta dissertação, que buscou estudar aspectos inseridos ou associados ao arcabouço da comunicação científica, área que desperta o interesse da CI. Assim, com o intuito de contribuir com as discussões que aproximam a perspectiva informacional das investigações que tratam dos desdobramentos causados pela COVID-19, o estudo realizado buscou respostas para o seguinte questionamento: qual a relação entre as dinâmicas de impacto bibliométrico de citação e altmétrico em artigos e *preprints* na temática COVID-19?

Com abordagem quantitativa, a pesquisa coletou, tratou e analisou dados bibliométricos de citação e altmétricos de leitores/capturas e menções provenientes de fontes como a Dimensions, o Mendeley e a plataforma Altmetric.com. Quanto à cobertura expressa pelas ferramentas utilizadas, foi possível observar que, para artigos, a que demonstrou melhor abrangência foi o gerenciador de referências, com valores de cobertura superiores a 88% em todos os subtipos de acesso. Por outro lado, entre os *preprints*, o Twitter se destacou por cobrir a maior porcentagem desse tipo de publicação (66%), indicando que, entre as publicações sobre COVID-19, os *preprints* circularam mais em contextos com públicos mais amplos que o científico, que costuma concentrar as atividades em ambientes mais formais de comunicação (periódicos científicos, bases de dados e redes sociais acadêmicas, a exemplo do Mendeley).

A interpretação dos dados obtidos possibilitou a verificação das hipóteses H1 e H2 levantadas pelo estudo. Os resultados e discussões de pesquisa associadas aos OE 1 demonstraram a confirmação parcial de H1, que considera que os artigos de acesso aberto tendem a apresentar maiores valores para indicadores bibliométricos de citação e altmétricos quando comparados aos artigos de acesso fechado. Os dados da pesquisa evidenciaram vantagens de acesso aberto para artigos em relação aos indicadores de citação (média e mediana) e para indicadores de leitores/capturas (média e mediana). Entretanto, quando se trata dos dados de menções no Twitter, os artigos de acesso aberto não demonstraram a mesma

vantagem, visto que os artigos do tipo fechado expuseram valores mais elevados para os indicadores de média e desvio padrão.

A análise comparativa geral realizada para os resultados associados ao OE 2, que focou em identificar a relação entre os indicadores de citações, menções e leitores/capturas e os subtipos de acesso (verde, bronze, dourado, híbrido e fechado), evidenciou vantagens, em média, para os subtipos bronze e híbrido, para citações, leitores/capturas e menções no Twitter. Tais subtipos obtiveram os mais elevados valores de média na maioria dos contextos analisados. Faz-se válido ressaltar que tanto para os dados de citação como altmétricos, os subtipos com as maiores médias também registraram altas dispersões, mensuradas pelos indicadores de desvio padrão. Nesse sentido, faz-se relevante a análise dos valores de mediana, que, por receber menos influência de números extremos, indica vantagens mais estáveis entre os subtipos de acesso que a média.

Para citações, as maiores medianas registradas foram para as categorias bronze e verde. Em leitores/capturas, os destaques foram para artigos bronze e híbridos. E, as medianas de menções mais altas estão em publicações híbridas, bronze e verde, uma vez que os dois últimos subtipos apresentaram valores iguais. Em suma, os resultados da pesquisa expressam que as vantagens de acesso aberto dos subtipos bronze e híbrido em relação aos outros estudados pela pesquisa se faz evidente entre as publicações na temática de COVID-19. Tais indicativos retratam a relevância das análises que se empenham em identificar as nuances de variação existentes entre os subtipos de acesso com o intuito de buscar novas alternativas de compreensão sobre o acesso aberto que ampliem a discussão polarizada que enfatiza a dualidade: aberto *versus* fechado.

Investigar os subtipos de acesso amplia as possibilidades de descoberta sobre as maiores dificuldades de implementação do acesso aberto em determinadas temáticas ou áreas. No caso dos artigos sobre COVID-19, os elevados valores de percentual da amostra e do resultados dos indicadores para o subtipo bronze, apontam características de disponibilização temporária da produção científica no período de maior esforço para o combate à pandemia e um baixo comprometimento das editoras com o fortalecimento do acesso aberto por meio da declaração clara de licenças de uso e reuso dos conteúdos publicados. Assim, acredita-se que, em 2020 e 2021, as editoras que publicaram os artigos estudados por esta pesquisa optaram, em maior grau, por estratégias de disponibilização temporária do conteúdo para leitura gratuita em detrimento das estratégias iniciais e mais tradicionais de acesso aberto propostas pelo MAA a partir da BOAI, as vias verde e dourada.

A postura adotada pelas editoras para cumprirem com a demanda social pela abertura das informações imposta pelo estado pandêmico pode influenciar nos resultados de estudos bibliométricos e altmétricos futuros realizados sobre a produção científica na temática de COVID-19 que considerem um recorte temporal mais amplo que somente os anos iniciais da emergência em saúde pública. Pois, com o avançar do tempo, poderá ser notada uma porcentagem mais alta de artigos em acesso fechado do que encontrados pela presente pesquisa, visto que, passado o período de maior apelo pela abertura das publicações, as permissões de acesso temporário aos artigos poderão ser retiradas.

Por se tratar de um subtipo de acesso recente, proposto por [Piwowar et al. \(2018\)](#), entende-se que ainda há muito que se avançar no entendimento sobre as características e as implicações da adoção da estratégia de acesso bronze no contexto geral do MAA. Assim, são sugeridas novas investigações comparativas que levem em consideração as variações dos subtipos de acesso associadas às abordagens bibliométricas e altmétricas, bem como, outras temáticas e áreas do conhecimento.

Pelos resultados associados aos OE1 e OE3 do trabalho, a segunda hipótese da pesquisa (H2) foi refutada. Durante o desenho do estudo, foi levantada a possibilidade de que, em virtude das características de dinamicidade e aceleração da produção e comunicação científica durante o período pandêmico, os *preprints* na temática da COVID-19 poderiam registrar maiores impactos de citação, de menções no Twitter e leitores/captura em relação aos artigos de acesso aberto e fechado a serem observados a partir dos valores obtidos para os indicadores bibliométricos e altmétricos calculados. Entretanto, os valores de média, mediana e desvio padrão para *preprints* alcançados por este estudo foram menores que os resultados encontrados para artigos, tanto abertos como fechados.

Ainda sobre os *preprints* da amostra investigada, faz-se relevante ressaltar que esses obtiveram cobertura pouco representativa no Mendeley e maior abrangência no Twitter. Além disso, entre os *preprints*, os resultados obtidos para rs evidenciaram correlações mais fortes entre citações e menções no Twitter do que entre citações e leitores/capturas, invertendo a lógica de intensidade de correlação mais comum entre os artigos, em que os dados de leitores/capturas tendem a se correlacionar de forma mais intensa com os indicadores bibliométricos, fazendo com que alguns estudos, a exemplo de [Maflahi e Thelwall \(2018\)](#), [Thelwall e Wilson \(2016\)](#) e [Zahedi, Costas e Wouters \(2014\)](#), declarem os indicadores calculados a partir dos dados do Mendeley como indicativos iniciais de impacto em artigos.

Em face do exposto e apesar das limitações da pesquisa já expressas no decorrer da dissertação, admite-se que o estudo realizado corrobora com o avanço das discussões sobre as mudanças na comunicação científica em contextos emergenciais; amplia as reflexões sobre as possibilidades e os subtipos de acesso aberto, bem como sobre as hipóteses de OACA e OAAA na literatura nacional de CI. A partir disso e com o propósito de incorporar novos conhecimentos e ampliar as discussões expressas por este trabalho, são sugeridos estudos métricos complementares que busquem analisar outros subtipos de acesso aberto, como por exemplo, o diamante e o *black*, por meio da combinação e comparação entre dados provenientes de múltiplas fontes de informação, além das mais tradicionais quando se trata de acesso aberto, como DOAJ e Unpaywall.

REFERÊNCIAS

- ADIE, E. Attention! A study of open access vs non-open access articles. *Figshare*, 2014. Disponível em: <http://dx.doi.org/10.6084/m9.figshare.1213690>. Acesso em: 6 abr. 2021.
- AGRESTI, A.; FINLAY, B. Métodos estatísticos para as ciências sociais. 4. ed. Porto Alegre: Penso, 2012.
- ALBRECHT, R. F.; OHIRA, M. L. B. Bases de dados: metodologia para seleção e coleta de documentos. *Revista ACB: Biblioteconomia em Santa Catarina*, Santa Catarina, v. 5, n. 5, p. 131-144, 2000. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/71931> . Acesso em: 18 set. 2021.
- ANDRÉS, A. *Mesuring academic research: how to undertake a bibliometric study*. 1. ed. Oxford: Chandos Publishing, 2009.
- ANTELMAN, K. Do Open-Access articles have a greater research impact? *College & Research Libraries News*, Chicago, v. 65, n. 5, p. 372-382, 2004. Disponível em: <https://doi.org/10.5860/crl.65.5.372> . Acesso em: 25 ago. 2021.
- APPEL, A. L. *Dimensões tecnopolíticas e econômicas da comunicação científica em transformação*. Orientadora: Profa. Dra. Sarita Albagli. Rio de Janeiro, 2019. 172 f. Tese (Doutorado em Ciência da Informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro; Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, RJ, 2019. Disponível em: https://ridi.ibict.br/bitstream/123456789/1024/1/Pesquisa_Tese_AndreAppel_20190706.pdf. Acesso em: 20 fev. 2021.
- APPEL, A. L.; ALBAGLI, S. The adoption of Article Processing Charges as a business model by Brazilian Open Access journals. *Transinformação*, Campinas, v. 31, e180045, p. 1-12, abr. 2019. Disponível em: <https://doi.org/10.1590/2318-0889201931e180045> . Acesso em: 8 abr. 2019.
- ARAÚJO, C. A. Bibliometria: evolução histórica e questões atuais. *Em Questão*, Porto Alegre, v. 12, n. 1, p. 11-32, jan./jun. 2006. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/16>. Acesso em: 14 mar. 2021.
- ARAUJO, K.; SILVA, C.; GUIMARÃES, M.; LINS, R.; ASSEF NETO, R. A produção científica sobre zika em periódicos de acesso aberto. *Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*, v. 11, n. sup., p. 1-8, 2017. Disponível em: <https://doi.org/10.29397/reciis.v11i0.1391>. Acesso em: 12 fev. 2022.
- ARAÚJO, R. F. Ciência 2.0 e a presença online de pesquisadores: visibilidade e impacto. *Ciência da Informação em Revista*, Maceió, v. 1, n. 3, p. 32-40, set./dez. 2014. Disponível em: <https://www.seer.ufal.br/index.php/cir/article/view/1608/1192>. Acesso em: 5 nov. 2021.
- ARAÚJO, R. F.; CARAN, G. M.; SOUZA, I. V. P. Orientação temática e coeficiente de correlação para análise comparativa entre dados altmétricos e citações: uma análise da revista

datagramazero. *Em Questão*, Porto Alegre, v. 22, n. 3, p. 184-200, 2016. DOI 10.19132/1808-5245223.184-200. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/61912>. Acesso em: 19 out. 2021.

ARBELÁEZ, E. J. V. La política editorial de acceso abierto: horizontes de la democratización del conocimiento para el desarrollo. *Equidad y Desarrollo*, Bogotá, n. 21, p. 145-161, 2014. Disponível em: <https://ciencia.lasalle.edu.co/eq/vol1/iss21/7/>. Acesso em: 4 jun. 2021.

ARCHAMBAULT, É *et al.* *Proportion of open access papers published in peer-reviewed journals at the European and world levels: 1996–2013*. [S.l.]: European Commission, 2014. Disponível em: http://science-metrix.com/sites/default/files/science-metrix/publications/d_1.8_sm_ec_dg-rtd_proportion_oa_1996-2013_v11p.pdf. Acesso em: 18 set. 2021.

ASSOCIATION OF COLLEGE & RESEARCH LIBRARIES. *Principles and Strategies for the Reform of Scholarly Communication 1*. Chicago, EUA, 2003. Disponível em: <http://www.ala.org/acrl/publications/whitepapers/principlesstrategies>. Acesso em: 15 jul. 2021.

ASSOCIATION OF RESEARCH LIBRARIES. *Report of the ARL serials prices project: a compilation of reports examining the serials prices problem*. Washington, DC: ARL, maio 1989. Disponível em: <http://catalog.hathitrust.org/Record/001527850>. Acesso em: 3 abr. 2021.

BARBETTA, P. A. *Estatística aplicada às Ciências Sociais*. 8.ed. Florianópolis: Ed. da UFSC, 2012.

BERMÚDEZ-RODRÍGUEZ, T. *et al.* O impacto do acesso aberto na produção e difusão de conhecimento sobre a Covid-19. *Liinc em Revista*, Rio de Janeiro, v. 16, n. 2, p. e5296, 2020. DOI: 10.18617/liinc.v16i2.5296. Disponível em: <http://revista.ibict.br/liinc/article/view/5296>. Acesso em: 13 ago. 2021.

BERLIN Declaration on Open Access to knowledge in the Sciences and Humanities. Alemanha, 2003. Disponível em: <https://openaccess.mpg.de/Berlin-Declaration>. Acesso em: 19 abr. 2018.

BETHESDA Statement on Open Access Publishing. Bethesda, EUA, 2003. Disponível em: <http://legacy.earlham.edu/~peters/fos/bethesda.htm>. Acesso em: 19 abr. 2018.

BJÖRK, B. C. A lifecycle model of the scientific communication process. *Learned Publishing*, [s.l.], v. 18, n. 3, p. 165–176, Jul. 2005. DOI <https://doi.org/10.1087/0953151054636129>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1087/0953151054636129>. Acesso em: 7 fev. 2022.

BJÖRK, B. C. Gold, green, and black open access. *Learned Publishing*, [s.l.], v. 30, n. 2, p. 173–175, 2017. DOI 10.1002/leap.1096. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/leap.1096>. Acesso em: 14 set. 2021.

BJÖRK, B. C. The hybrid model for open access publication of scholarly articles: a failed

experiment? *Journal of the American Society for Information Science and Technology*, [s.l.], v. 63, n. 8, p. 1496–1504, ago. 2012. Disponível em: <https://doi.org/10.1002/asi.22709>. Acesso em: 25 maio 2021.

BODE, C. *et al.* *A guide to the Dimensions data approach: a collaborative approach to creating a modern infrastructure for data describing research: where we are and where we want to take it*. Cambridge, USA: Dimensions, 2019. DOI: 10.6084/m9.figshare.5783094 Disponível em: <https://www.dimensions.ai/resources/a-guide-to-the-dimensions-data-approach/>. Acesso em: 20 out. 2020.

BOETTO, E. *et al.* Using altmetrics for detecting impactful research in quasi-zero-day time-windows: the case of COVID-19. *Scientometrics*, [s.l.], v. 126, p. 1189–1215, 2021. Disponível em: <https://doi.org/10.1007/s11192-020-03809-7>.

BOLLEN, J. *et al.* A Principal Component Analysis of 39 Scientific Impact Measures. *PLOS One*, [s.l.], v. 4, n. 6, 2009. Disponível em: <https://doi.org/10.1371/journal.pone.0006022>. Acesso em: 10 nov. 2021.

BORKO, H. Information Science: What is it? *American Documentation*, [s.l.], v. 19, n.1, p.3-5, jan. 1968.

BRYMAN, A. *Social Research Methods*. 4. ed. Oxford: Oxford University Press, 2012.

BUDAPEST OPEN ACCESS INITIATIVE (BOAI). Budapest, Hungary, 14 Feb., 2002. Disponível em: <http://www.budapestopenaccessinitiative.org/read>. Acesso em: 19 abr. 2018.

CARIBÉ, R. C. V. Comunicação científica: reflexões sobre o conceito. *Informação & Sociedade: Estudos, Paraíba*, v. 25, n. 3, p. 89-104, 2015. Disponível em: <http://www.periodicos.ufpb.br/ojs/index.php/ies/article/view/23109>. Acesso em: 23 abr. 2021.

CINTRA, P. R. *Avaliação do impacto do acesso aberto em periódicos da área da Ciência da Informação: uma análise de indicadores bibliométricos e altmétricos*. 2017. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) – Universidade Federal de São Carlos, São Carlos, 2017. Disponível em: <https://repositorio.ufscar.br/bitstream/handle/ufscar/8552/DissPRC.pdf?sequence=1>. Acesso em: 19 fev. 2021.

CINTRA, P. R.; FURNIVAL, A. C.; MILANEZ, D. H. Vantagens de citação do acesso aberto em periódicos selecionados da ciência da informação: uma análise ampliada aos indicadores altmétricos. *Informação & Informação*, v. 22, n. 1, p. 129-149, 2017. DOI: 10.5433/1981-8920.2017v22n1p129. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/27146>. Acesso em: 15 jul. 2020.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). Editoras removem controle de acesso a conteúdos em atenção à pandemia. Brasília:

Capes, [2020]. Disponível em:

<http://mailer.periodicos.capes.gov.br/?m=136&p=view&pi=ViewBrowserPlugin&uid=69c6496339f687bc1e7ff4ea4624e28f> . Acesso em: 27 nov. 2021.

COSTA, M. P.; LEITE, F. C. L. *Repositórios institucionais da América Latina e o acesso aberto à informação científica*. Brasília: IBICT, 2017. Disponível em:

<https://repositorio.unb.br/handle/10482/23202>. Acesso em: 17 jan. 2021.

COSTA, S. Abordagens, estratégias e ferramentas para o acesso aberto via periódicos e repositórios institucionais em instituições acadêmicas brasileiras. *Liinc em Revista*, Rio de Janeiro, v. 4, n. 2, 2008. DOI 10.18617/liinc.v4i2.281. Disponível em:

<http://revista.ibict.br/liinc/article/view/3175> . Acesso em: 28 out. 2021.

COSTA, S. M. S. Mudanças no Processo de Comunicação Científica: O impacto do uso de novas tecnologias. In: MUELLER, S. P. M.; PASSOS, E. J. L.

(orgs.). *Comunicação Científica*. Brasília: Departamento de Ciência da Informação, Universidade de Brasília, 2000. p. 95-105.

COSTAS, R.; ZAHEDI, Z.; WOUTERS, P. Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *JASIST*, [s.l.], v. 66, n. 10, p. 2003-2019, jul. 2014. DOI:

<https://doi.org/10.1002/asi.23309>. Disponível em:

<https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.23309>. Acesso em: 8 abr. 2020.

CRESWELL, J. W.; CRESWELL, J. D. *Projeto de Pesquisa: métodos qualitativo, quantitativo e misto*. 5. ed. Porto Alegre: Penso, 2021.

CUNHA, M. B.; CAVALCANTI, C. R. O. *Dicionário de Biblioteconomia e Arquivologia*. Brasília: Briquet de Lemos, 2008.

DIMENSIONS. *Where are citations in Dimensions coming from?* 29 jul. 2019. Disponível em:<https://dimensions.freshdesk.com/support/solutions/articles/23000018849-where-are-citations-in-dimensions-coming-from-> .Acesso em: 8 nov. 2021.

DOAJ. What does DOAJ define as open access? *Blog DOAJ*, 11 nov. 2020. Disponível em: <https://blog.doaj.org/2020/11/17/what-does-doaj-define-as-open-access/>. Acesso em: 14 mar. 2022.

ELSEVIER. *About Mendeley*. 2022. Disponível em:

<https://www.elsevier.com/solutions/mendeley>. Acesso em: 13 mar. 2022.

ERIKSON, M. G.; ERLANDSON, P. A taxonomy of motives to cite. *Social Studies of Science*, [s.l.], v. 44, n. 4, p. 625-637, 2014. DOI 10.1177/0306312714522871. Disponível em: <https://journals.sagepub.com/doi/10.1177/0306312714522871>. Acesso em: 8 out. 2021.

EYSENBACH, G. Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. *Journal of Medical Internet Research*, [s.l.], v. 13, n. 4, 2011. Disponível em: <https://www.jmir.org/2011/4/e123/>. Acesso em: 18 set. 2021.

EYSENBACH, G. Citation Advantage of Open Access Articles. *PLoS Biology*, [s.l.], v. 4, n. 5, May 2006. Disponível em: <https://doi.org/10.1371/journal.pbio.0040157>. Acesso em: 28 jul. 2021.

FAUSTO, S. Evolução do Acesso Aberto: breve histórico. *SciELO em Perspectiva*, [s.l.], 2013. Disponível em: <https://blog.scielo.org/blog/2013/10/21/evolucao-do-acesso-aberto-breve-historico/>. Acesso em: 16 dez. 2020.

FAVERO, L. P.; BELFIORE, P. Manual de Análise de Dados: Estatística e Modelagem Multivariada com Excel®, SPSS® e Stata®. Rio de Janeiro: Elsevier, 2017.

FERNANDES, H. D. H; VILAN FILHO, J. L. Fluxo da informação científica: uma revisão dos modelos propostos na literatura em Ciência da Informação. *Em questão*, Porto Alegre, v. 27, n. 2, p. 138-163, abr./jun. 2021. Disponível em: <https://doi.org/10.19132/1808-5245272.138-163>. Acesso em: 19 mar. 2021.

FRASER, N. *et al.* The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS Biol*, [s.l.], v. 19, n. 4, e3000959, apr. 2021. Disponível em: <https://doi.org/10.1371/journal.pbio.3000959>. Acesso em: 13 jul. 2021.

FRASER, N.; MOMENI, F.; MAYR, P.; PETERS, I. The relationship between bioRxiv preprints, citations and altmetrics. *Quantitative Science Studies*, v. 1, n. 2, p. 618–638, 2020. Disponível em: https://doi.org/10.1162/qss_a_00043. Acesso em: 3 maio 2022.

FUCHS, C.; SANDOVAL, M. The diamond model of open access publishing: Why policy makers, scholars, universities, libraries, labour unions and the publishing world need to take non-commercial, nonprofit open access serious. *TipleC*, [s.l.], v. 11, n. 2, p. 428-443, 2013. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.426.6430&rep=rep1&type=pdf> . Acesso em: 19 ago. 2021.

FURNIVAL, A. C.; HUBBARD, B. Acesso aberto às publicações científicas: vantagens, políticas e *advocacy*. *InCID: R. Ci. Inf. e Doc.*, Ribeirão Preto, v. 2, n. 2, p. 160-177, jul./dez. 2011. Disponível em: <https://doi.org/10.11606/issn.2178-2075.v2i2p160-177>. Acesso em: 15 jun. 2021.

GARVEY, W. D. *Communication: the essence of science: facilitating information exchange among librarians, scientists, engineers, and students, librarians, scientists, engineers, and students*. Oxford: Pergamon Press, 1979. Disponível em: <https://books.google.com.br/books?id=3RijBQAAQBAJ&lpg=PA1&hl=pt-BR&pg=PR7#v=onepage&q&f=false>. Acesso em: 19 jan. 2022.

GARVEY, W. D.; GRIFFITH, B.C. Communication and information processing within scientific disciplines: Empirical findings for Psychology. *Information Storage and Retrieval*, Elmsford, v. 8, n. 3, p. 123-136, 1972. DOI [https://doi.org/10.1016/0020-0271\(72\)90041-1](https://doi.org/10.1016/0020-0271(72)90041-1). Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/0020027172900411>. Acesso em: 12 nov. 2021.

GIL, A. C. *Métodos e Técnicas de Pesquisa Social*. 7.ed. São Paulo: Atlas, 2019.

GLÄNZEL, W. Bibliometrics as a research field: a course on theory and application of bibliometric indicators. [S.l.]: Course handouts, 2003.

GONTIJO, M. C. A.; ARAÚJO, R. F. Dados bibliométricos e altmétricos de artigos científicos sobre inteligência artificial: análise do impacto acadêmico e social. *Múltiplos Olhares em Ciência da Informação*, v. 9, n. 2, 2019. Disponível em: <https://periodicos.ufmg.br/index.php/moci/article/view/1913>. Acesso em: 11 abr. 2020.

GROSS, P. L. K.; GROSS, E. M. College Libraries and Chemical Education. *Science*, [s.l.], v. 66, n. 1713, p. 385-389, 1927. Disponível em: <https://www.science.org/doi/10.1126/science.66.1713.385>. Acesso em: 23 ago. 2021.

HARNAD, S. Gold Open Access publishing must not be allowed to retard the progress of Green Open Access Self-Archiving. *Logos*, [s.l.], jan. 2010. DOI 10.1163/095796511X559972. Disponível em: https://brill.com/view/journals/logo/21/3-4/article-p86_10.xml. Acesso em: jul. 2021.

HARNAD, S. The access/impact problem and the green and gold roads to open access: an update. *Serials Review*, [s.l.], v. 34, n. 1, p. 36–40, 2008. DOI 10.1080/00987913.2008.10765150. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0098791307001530>. Acesso em: 20 maio 2021.

HAUSTEIN, S.; PETERS, I.; SUGIMOTO, C. R.; THELWALL, M.; LARIVIÈRE, V. Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *JASIST*, [s.l.], v. 65, n. 4, p. 656-669, Apr. 2014. Disponível em: <https://doi-org.ez54.periodicos.capes.gov.br/10.1002/asi.23101>. Acesso em: 12 maio 2022.

HOLMBERG, K. *et al.* Do articles in open access journals have more frequent altmetric activity than articles in subscription-based journals? An investigation of the research output of Finnish universities. *Scientometrics*, [s.l.], v. 122, n. 1, p. 645-659, 2019. DOI: <https://doi.org/10.1007/s11192-019-03301-x>. Disponível em: <https://link.springer.com/article/10.1007/s11192-019-03301-x> . Acesso em: 17 jan. 2021.

HOOK, D. W.; PORTER, S. J.; HERZOG, C. Dimensions: Building Context for Search and Evaluation. *Frontiers in Research Metrics and analytics*, [s.l.], v. 3, p. 1-11, ago. 2018. Disponível em: <https://doi.org/10.3389/frma.2018.00023>. Acesso em: 28 out. 2020.

HOPPEN, N. H. F. *A Neurociências no Brasil de 2006 a 2013, indexada na Web of Science: produção científica, colaboração e impacto*. 2014. 163 f. Dissertação (Mestrado em Comunicação e Informação) - Programa de Pós-graduação em Comunicação e Informação, Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014. Disponível em: <http://hdl.handle.net/10183/107958>. Acesso em: 7 set. 2021.

HRYNASZKIEWICZ, I. Open access journals: a sustainable and scalable solution in social and political sciences? *European Political Science*, [s.l.], v. 15, p. 191–200, 2016. Disponível em: <https://doi.org/10.1057/eps.2015.87> . Acesso em: 13 ago. 2021.

HURD, J. M. Models of Scientific Communications Systems. In: CROWFORD, S. Y.; HURD, J. M.; WELLER, A. C. (orgs.). *From Print to Electronic: the transformation of scientific communication*. Medford, MA: ASIS, 1996. p. 9-33.

JORGE, V. A. *Abertura e compartilhamento de dados para pesquisa nas situações de emergência em saúde pública: o caso do vírus Zika*. Rio de Janeiro, 2018. 263 f. Orientadora: Sarita Albagli. Tese (Doutorado em Ciência da Informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2018. Disponível em: <https://www.arca.fiocruz.br/handle/icict/32604>. Acesso em: 12 fev. 2021.

JORGE, V. A.; ALBAGLI, S.; SOUZA, A. R. Apresentação. *Liinc em Revista*, Rio de Janeiro, v. 16, n. 2, e5535, dez. 2020. DOI <https://doi.org/10.18617/liinc.v16i2.5341>. Disponível em: <http://revista.ibict.br/liinc/article/view/5535/5117>. Acesso em: 15 jul. 2021.

KAMEL BOULOS, M. N.; ANDERSON, P. F. Preliminary survey of leading general medicine journals' use of Facebook and Twitter. *Journal of the Canadian Health Libraries Association*, [s. l.], v. 33, n. 2, p. 38–47, Jul. 2014. DOI 10.5596/c2012-010. Disponível em: <https://journals.library.ualberta.ca/jchla/index.php/jchla/article/view/22579>. Acesso em: 10 maio. 2022.

KANJILAL, U.; KUMAR DAS, A. *Introduction to Open Access*. Paris: UNESCO, 2015. (Open Access for Library Schools). Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000231920?posInSet=1&queryId=9eeab845-0f22-49fa-a64b-b119afd1697a>. Acesso em: 21 maio 2021.

KLING, R. The internet and unrefereed scholarly publishing. *Annual Review of Information Science and Technology*, [s.l.], v. 38, n. 1, p. 591–631, 22 set. 2005. DOI <http://doi.wiley.com/10.1002/aris.1440380113>. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/aris.1440380113>. Acesso em: 4 jul. 2021.

KUPFERSCHMIDT, K. A completely new culture of doing research: coronavirus outbreak changes how scientists communicate. *Science*, [s.l.], 26 Feb. 2020. DOI 10.1126/science.abb4761. Disponível em: <https://www.sciencemag.org/news/2020/02/completely-new-culture-doing-research-coronavirus-outbreak-changes-how-scientists>. Acesso em: 19 mar. 2020.

KURAMOTO, Hélio. Informação Científica: proposta de um novo modelo para o Brasil. *Ciência da Informação*, Brasília, v. 35, p. 91-102, 2006. Disponível em: <https://doi.org/10.1590/S0100-19652006000200010>. Acesso em: 20 maio 2021.

LANGHAM-PUTROW, A; BAKKER, C; RIEGELMAN, A. Is the open access citation advantage real? A systematic review of the citation of open access and subscription-based articles. *PLoS ONE*, Barcelona, v. 16, n. 6, e0253129, Jun. 2021. Disponível em:

<https://doi.org/10.1371/journal.pone.0253129>. Acesso em: 3 ago. 2021.

LARIVIÈRE, V.; SHU, F.; SUGIMOTO, C. The Coronavirus (COVID-19) outbreak highlights serious deficiencies in scholarly communication. *LSE Impact Blog*. London, 5 mar. 2020. Disponível em: <https://blogs.lse.ac.uk/impactofsocialsciences/2020/03/05/the-coronavirus-covid-19-outbreak-highlights-serious-deficiencies-in-scholarly-communication/>. Acesso em: 26 nov. 2021.

LAWRENCE, S. Free online availability substantially increases a paper's impact. *Nature*, [s.l.], v. 411, n. 521, p. 521-522, May. 2001. Disponível em: <https://www.nature.com/articles/35079151>. Acesso em: 26 set. 2021.

LE COADIC, Yves-François. A ciência da informação. 2 ed. Brasília: Briquet de Lemos, 2004.

LEITE, F. C. L. Modelo genérico de gestão da informação científica para instituições de pesquisa na perspectiva da comunicação científica e do acesso aberto. 2011. 262 f. Tese (Doutorado em Ciência da Informação)—Universidade de Brasília, Brasília, 2011. Disponível em: <https://repositorio.unb.br/handle/10482/9753>. Acesso em: 7 fev. 2022.

MAFLAHI, N; THELWALL, M. How quickly do publications get read? The evolution of Mendeley reader counts for new articles, *JASIST*, [s.l.], v. 69, n. 1, p. 158-167, Jan. 2018. DOI <https://doi.org/10.1002/asi.23909>. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.23909>. Acesso em: 6 maio 2022.

MAJUMDER, M.S.; MANDL, K. D. Early in the epidemic: impact of preprints on global discourse about COVID-19 transmissibility. *The Lancet Glob Health*, [s.l.], v. 8, n. 5, e627-e630, 2020. DOI [https://doi.org/10.1016/S2214-109X\(20\)30113-3](https://doi.org/10.1016/S2214-109X(20)30113-3). Disponível em: <https://www.sciencedirect.com/science/article/pii/S2214109X20301133>. Acesso em: 15 ago. 2021.

MALTRÁS BARBA, B. *Los indicadores bibliométricos: fundamentos y aplicación al análisis de la ciencia*. Espanha: Ediciones Trea, 2003.

MARICATO, J. M.; MARTINS, D. L. Altmetria: complexidades, desafios e novas formas de mensuração e compreensão da comunicação científica na web social. *Biblios*, Pittsburgh, n. 68, p. 48-68, jul. 2017. Disponível em: <http://dx.doi.org/10.5195/biblios.2017.358>. Acesso em: 23 mar. 2021.

MATTEDI, M. A.; SPIESS, M. R. A avaliação da produtividade científica. *História, Ciências, Saúde: Manguinhos*, Rio de Janeiro, v. 24, n. 3, p. 623–643, 2017. Disponível em: <https://doi.org/10.1590/S0104-59702017000300005>. Acesso em: 24 out. 2021.

MEADOWS, A. J. A comunicação científica. Brasília: Briquet de Lemos, 1999.

MITTERMAIER, B. Double dipping in hybrid open access: chimera or reality? *ScienceOpen Research*, [s.l.], p. 1-12. DOI 10.14293/S2199-1006.1.SOR-SOCSCIAOWNTU.v1.

Disponível em: <https://www.scienceopen.com/hosted-document?doi=10.14293/S2199-1006.1.SOR-SOCSCI.AOWNTU.v1.2.abr.2022>.

MOMESSO, A. C.; NORONHA, D. P. Bibliométrie ou Bibliometrics: o que há por trás de um termo? *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 22, n. 2, p. 118-124, jun. 2017. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2831>. Acesso em: 8 nov. 2021.

MUELLER, S. P. M. A comunicação científica e o movimento de acesso livre ao conhecimento. *Ciência da Informação*, Brasília, v. 35, n. 2, p. 27-38, ago. 2006. Disponível em: <http://revista.ibict.br/ciinf/article/view/1138/1293>. Acesso em: 25 abr. 2018.

MUELLER, S. P. M.; PASSOS, E. J. L. As questões da comunicação científica e a ciência da informação. In: MUELLER, S. P. M.; PASSOS, E. J. L. (org.). *Comunicação científica*. Brasília: Ciência da Informação, 2000. p. 13-22.

NASAB F. R., RAHIM F. Bibliometric analysis of global scientific research on SARSCoV-2 (COVID-19). *MedRxiv*, [s.l.], 2020. DOI: <https://doi.org/2020.03.19.20038752>. Disponível em: <https://www.medrxiv.org/content/10.1101/2020.03.19.20038752v1>. Acesso em: 15 ago. 2020.

NASCIMENTO, A. G. *Altmetria para bibliotecários: guia prático de métricas alternativas para avaliação da produção científica*. São Paulo: Scortecci, 2017.

NASSI-CALÒ, L. Avaliação por pares: modalidades, prós e contras. *SciELO em Perspectiva*, [s.l.], 2015. Disponível em: <https://blog.scielo.org/blog/2015/03/27/avaliacao-por-pares-modalidades-pros-e-contras>. Acesso em: 29 jul. 2021.

NAVARRO, D. J.; FOXCROFT, D. R. *Learning statistics with jamovi: a tutorial for psychology students and other beginners. Version 0.75*. [S.l.]: Jamovi, 2022. DOI: 10.24384/hgc3-7p15. Disponível em: <https://www.learnstatswithjamovi.com/>. Acesso em: 1 junho 2022.

NORONHA, D. P.; MARICATO, J. M. Estudos métricos da informação: primeiras aproximações. *Encontros Bibli: Revista eletrônica de Biblioteconomia e Ciência da Informação*, [s.l.], v. 13, n. 1, p. 116-128, 2008. Disponível em: <https://doi.org/10.5007/1518-2924.2008v13nesp1p116>. Acesso em: 19 fev. 2021.

OPENDOAR. *OpenDOAR Statistics: an overview of the data held in OpenDOAR*. Jisc, Nov. 2021. Disponível em: https://v2.sherpa.ac.uk/view/repository_visualisations/1.html. Acesso em: 7 nov. 2021.

ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE. OMS afirma que COVID-19 é agora caracterizada como pandemia. OPAS, 11 mar. 2020. Disponível em: <https://www.paho.org/pt/news/11-3-2020-who-characterizes-covid-19-pandemic>. Acesso em: 14 maio 2021.

ORR, R. How is the best OA location determined? *Unpaywall support portal*, 19 jan. 2021. Disponível em: <https://support.unpaywall.org/support/solutions/articles/44001943223-how-is-the-best-oa-location-determined->. Acesso em: 17 maio 2022.

ORTIZ NÚÑEZ, R.; STABLE RODRÍGUEZ, Y. Análisis de la producción científica internacional sobre tratamientos contra la COVID-19. *Revista Cubana de Información en Ciencias de la Salud*, [s.l.], v. 32, n. 2, jun. 2021. ISSN 2307-2113. Disponível em: <http://www.rcics.sld.cu/index.php/acimed/article/view/1682>. Acesso em: 8 ago. 2021.

PACHECO, A. *et al.* Citações e métricas complementares: um estudo exploratório da sua correlação em artigos científicos em acesso aberto. *Cadernos BAD*, Portugal, n. 1, p. 125-138, 2018. Disponível em: <https://publicacoes.bad.pt/revistas/index.php/cadernos/article/view/1906/pdf>. Acesso em: 24 out. 2021.

PINHEIRO, L. V. R. Constituição epistemológica e social da comunicação científica no Brasil. In: PINHEIRO, L. V. R.; OLIVEIRA, E. C. P. (org.). *Múltiplas facetas da comunicação e divulgação científicas: transformações em cinco séculos*. Brasília, DF: IBICT, 2012. p. 115–149.

PIWOWAR, H. Introduction Altmetrics: What, why and where? *Bulletin of the American Society for Information Science and Technology*, [s.l.], v. 39, n. 4, p. 8-9, 2013. Disponível em: <https://doi.org/10.1002/bult.2013.1720390404>. Acesso em: 25 jul. 2014.

PIWOWAR, H. *et al.* The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, [s.l.], v. 6, p. e4375, 2018. Disponível em: <https://doi.org/10.7717/peerj.4375>. Acesso em: 4 mar. 2020.

PRICE, D. J. S. *Little science, big science*. New York: Columbia University, 1963.

PRICE, R.; OZKAN, Y. Characteristics of Imperial College London's COVID-19 research outputs. *Learned Publishing*, [s.l.], v. 34, n. 3, p. 358-369, jan. 2021. DOI: 10.1002/leap.1358. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/leap.1358>.

PRIEM, J. Altmetrics. In: CRONIN, B.; SUGIMOTO, C. R. (eds.). *Beyond bibliometrics: harnessing multidimensional indicators of performance*. Cambridge, MA: MIT Press, 2014. p. 263–287. Disponível em: <https://arxiv.org/abs/1507.01328>. Acesso em: 27 out. 2021.

PRIEM, J. How do we decide if a given journal is fully OA? *Unpaywall support portal*, 29 dez. 2020. Disponível em: <https://support.unpaywall.org/support/solutions/articles/44001792752-how-do-we-decide-if-a-given-journal-is-fully-oa-1777288-what-do-the-types-of-oa-status-green-gold-hybrid-and-bronze-mean->. Acesso em: 17 maio 2022.

PRIEM, J. What do the types of oa_status (green, gold, hybrid, and bronze) mean? *Unpaywall support portal*, 1 dez. 2021. Disponível em: <https://support.unpaywall.org/support/solutions/articles/44001777288-what-do-the-types-of-oa-status-green-gold-hybrid-and-bronze-mean->. Acesso em: 17 maio 2022.

PRIEM, J. *et al.* Altmetrics: a manifesto. 26 out. 2010. Disponível em: <http://altmetrics.org/manifesto/>. Acesso em: 27 jan. 2021.

PRITCHARD, A. Statistical bibliography or bibliometrics? *Journal of Documentation*, [s.l.], v. 25, n. 4, p. 348-349, 1969.

PULVERER, B. Preparing for Preprints. *EMBO J*, Heidelberg, Germany, v. 35, n. 24, p. 2617-2619, 2016. DOI: <https://doi.org/10.15252/embj.201670030>. Disponível em: <https://www.embopress.org/doi/full/10.15252/embj.201670030> . Acesso em: 25 out. 2021.

RAZUMOVA, I. K. COVID-19 Pandemic and Self-isolation: impact on bibliometrics and use. Part II. Citation and use. *Scholarly Research and Information*, [s.l.], v. 3, n. 2-3, p. 188-206, 2020. DOI <https://doi.org/10.24108/2658-3143-2020-3-2-3-188-206>. Disponível em: <https://www.neiconjournal.com/jour/article/view/95>. Acesso em: 20 jul. 2021.

RAMOS, T. B. C.; MARICATO, J. M. Odisseia Métrics: Ferramenta de Extração de Dados. Versão 1.0. 14 out. 2020. Disponível em: https://colab.research.google.com/drive/1L75mJbqteVE0kMoTMK1_kLVXWFZpQumi#scrollTo=ohRqrEb7RXIm. Acesso em: 27 out. 2021.

RIOS, F. P.; LUCAS, E. R. O.; AMORIM, I. S. Manifestos do movimento de acesso aberto: Análise de Domínio a partir de periódicos brasileiros. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, v. 15, n. 1, p. 148-169, jan. 2019. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/1152>. Acesso em: 17 out. 2021.

ROBINSON-GARCIA, N.; COSTAS, R.; VAN LEEUWEN, T. N. Open Access uptake by universities worldwide. *PeerJ*, [s.l.], v. 8, p. e9410, 8 jul. 2020. Disponível em: <http://doi.org/10.7717/peerj.9410>. Acesso em: 12 ago. 2020.

RODRIGUES, E. Acesso livre ao conhecimento: a mudança do sistema de comunicação da ciência e os profissionais de informação. *Cadernos BAD*, Portugal, n. 1, 2004. Disponível em: <https://publicacoes.bad.pt/revistas/index.php/cadernos/article/view/836>. Acesso em: 17 out. 2021.

RODRÍGUEZ, E. G. Preprints and preprint servers as academic communication tools. *Revista Cubana de Información en Ciencias de la Salud*, Cuba, v. 30, n. 1, 2019. Disponível em: <http://www.rcics.sld.cu/index.php/acimed/article/view/1324>. Acesso em: 10 set. 2021.

SAN Francisco Declaration on Research Assessment. San Francisco, Dec. 2012. Disponível em: <https://sfdora.org/read/> . Acesso em: 25 jul. 2021.

SARMENTO, F. *et al.* Algumas considerações sobre as principais declarações que suportam o movimento Acesso Livre. In: WORLD CONGRESS ON HEALTH INFORMATION AND LIBRARIES, 2005, Salvador. *Anais [...]*. Salvador: [s.n.], 2005. Disponível em: http://eprints.rclis.org/8512/1/Sarmento_Miranda_Baptista_Ramos_-_Vers%C3%A3o_Final.pdf.

SARABIPOUR, S. *et al.* On the value of preprints: An early career researcher perspective.

PLoS Biol, [s.l.], v. 17, n. 2, e3000151, 2019. Disponível em: <https://doi.org/10.1371/journal.pbio.3000151>. 13 set. 2021.

SERGHIOU, S.; IOANNIDIS, J. P. A. Altmetric Scores, Citations, and Publication of Studies Posted as Preprints. *JAMA Network*, [s.l.], v. 319, n. 4, p. 402-404, jan. 2018. DOI 10.1001/jama.2017.21168. Disponível em: <https://jamanetwork.com/journals/jama/fullarticle/2670247>. Acesso em: 21 set. 2021.

SILVA FILHO, R. C.; VANZ, S. A. S. Impacto de altmetrics sobre a visibilidade de artigos em acesso aberto da enfermagem brasileira: um estudo de caso. *Transinformação*, Campinas, v. 31, e190025, 2019. Disponível em: <http://dx.doi.org/10.1590/2318-0889201931e190025>. Acesso em: 19 abr. 2020.

SLOW SCIENCE ACADEMY. The slow science manifesto. Berlin, 2010. Disponível: <http://slow-science.org/>. Acesso em: 14 mar. 2021.

SOCIEDADE BRASILEIRA PARA O PROGRESSO DA CIÊNCIA. SBPC, ABC e ABI pedem que meios de comunicação liberem acesso a todas as informações sobre coronavírus. *SBPC*, São Paulo, 17 mar. 2020. Disponível em: <http://portal.sbpcnet.org.br/noticias/sbpc-abc-e-abi-pedem-que-meios-de-comunicacao-liberem-acesso-a-todas-as-informacoes-sobre-coronavirus/>. Acesso em: 27 nov. 2021.

SOUZA, I. V. P. Almetria: estado da arte. *Tendências da Pesquisa Brasileira em Ciência da Informação*, [s.l.], v. 7, n. 2, p. 251-268, jul./dez. 2014. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/315>. Acesso em: 14 set. 2021.

SUBER, P. Gratis and libre open access. *SPARC Open Access Newsletter*, [s.l.], n. 124, 2008. Disponível em: https://dash.harvard.edu/bitstream/handle/1/4322580/suber_oagratis.html?sequence=1&isAllowed=y. Acesso em: 12 jun. 2021.

SUBER, P. Open Access overview: focusing on open access to peer-reviewed research articles and their preprints. 1 ed. rev. *Earlham College*, 2015. Disponível em: <https://legacy.earlham.edu/~peters/fos/overview.htm>. Acesso em: 19 abr. 2019.

SUBER, P. *Timeline of the Open Access Movement*. 2009. Disponível em: <https://legacy.earlham.edu/~peters/fos/timeline.htm>. Acesso em: 18 jan. 2021.

SPARC EUROPE. *The open access citation advantage: list of studies until 2015*. Disponível em: <https://sparceurope.org/what-we-do/open-access/sparc-europe-open-access-resources/open-access-citation-advantage-service-oaca/>. Acesso em: 6 ago. 2020.

SPINAK, E. Dicionario enciclopédico de bibliometria, cienciometria e informetria. [S.l.]: UNESCO, 1996. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000243329>. Acesso em: 26 jul. 2021.

SWAN, A. *Policy guidelines for the development and promotion of open access*. Paris: UNESCO, 2012. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000215863>. Acesso em: 15 mar. 2020.

TAGA, V. *Acesso aberto: classificação dos artigos indexados na Scopus*. 2016. Dissertação (Mestrado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Santa Catarina, Santa Catarina, 2016. Disponível em: <https://repositorio.ufsc.br/xmlui/handle/123456789/175907>. Acesso em: 15 ago. 2021.

TARGINO, M. G. Comunicação científica: uma revisão de seus elementos básicos. *Informação & Sociedade: Estudos*, Paraíba, v. 10, n. 2, p. 37-85, 2000. Disponível em: <https://periodicos.ufpb.br/index.php/ies/article/view/326>. Acesso em: 6 mar. 2021.

TARGINO, M. G.; TORRES, N. H. Comunicação Científica além da Ciência. *Ação midiática*, Paraná, n. 7, 2014. Disponível em: <https://revistas.ufpr.br/acaomidiatica/article/view/36899>. Acesso em: 28 abr. 2021.

TENNANT, J. P *et al.* The academic, economic and societal impacts of Open Access: an evidence-based review. *F1000Research*, [s.l.], v. 5, n. 632, p. 1-55, 2020. (version 3; peer review: 4 approved, 1 approved with reservations). Disponível em: <https://doi.org/10.12688/f1000research.8460.3>. Acesso em: 27 jul. 2021.

THELWALL, M. Dimensions: A competitor to Scopus and the Web of Science? *Journal of Informetrics*, [s.l.], v. 12, n. 2, p. 430–435, mar. 2018a. Disponível em: <https://doi.org/10.1016/j.joi.2018.03.006>. Acesso em: 22 set. 2021.

THELWALL, M. Early Mendeley readers correlate with later citation counts. *Scientometrics*, [s.l.], v. 115, p. 1231–1240, 2018b. Disponível em: <https://doi.org/10.1007/s11192-018-2715-9>. Acesso em: 29 out. 2021.

THELWALL, M.; HAUSTEIN, S.; LARIVIÈRE, V.; SUGIMOTO, C. R. Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS ONE*, v. 8, n. 5, 2013, p. e64841. Disponível em: <https://doi.org/10.1371/journal.pone.0064841>. Acesso em: 17 abr. 2022.

THELWALL, M.; WILSON, P. Mendeley readership altmetrics for medical articles: an analysis of 45 fields. *JASIST*, v. 67, n. 8, p. 1962-1972, Aug. 2016. Disponível em: <https://doi.org/10.1002/asi.23501>. Acesso em: 20 jun. 2022.

TORRES PASCUAL, C.; TORRELL-VALLESPIN, S. Análisis bibliométrico de la producción científica latinoamericana y del Caribe sobre COVID-19 en PUBMED. *Revista Cubana de Información en Ciencias de la Salud*, La Habana, v. 31, n. 3, e1600, sept. 2020. Disponível em: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2307-21132020000300008&lng=es&nrm=iso. Acesso em: 10 jun. 2021.

TORRES-SALINAS, D.; ROBINSON-GARCIA, N.; CASTILLO-VALDIVIESO, P. Open Access and Altmetrics in the pandemic age: Forecast analysis on COVID-19 literature. *bioRxiv*, 2020. DOI <https://doi.org/10.1101/2020.04.23.057307>. Disponível em: <https://www.biorxiv.org/content/10.1101/2020.04.23.057307v1.full#ref-15>.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO). *Declaração sobre a ciência e o uso do conhecimento científico*. Budapeste, Hungria: UNESCO, 1999. Disponível em: <http://www.dominiopublico.gov.br/download/texto/ue000111.pdf>. Acesso em: 27 nov. 2021

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO). *Unisist: Study Report on the Feasibility of a World Science Information System*. Paris: Unesco, 1971.

UNPAYWALL. Portal da ferramenta Unpaywall. 2022. Disponível em: <https://unpaywall.org/>. Acesso em: 18 abr. 2022.

VAN LEEUWEN, T.; COSTAS, R.; ROBINSON-GARCIA, N. Indicators of open access publishing in the CWTS Leiden Ranking 2019. *CWTS Meaningful metrics*, Leiden, 15 may 2019. Disponível em: <https://www.cwts.nl/blog?article=n-r2w2a4>. Acesso em: 8 dez. 2020.

VANTAGEM. In: MICHAELIS: Dicionário Brasileiro da Língua Portuguesa. [S.l.]: Melhoramentos, c2015. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/vantagem/>. Acesso em: 26 nov. 2021.

VANTI, N. A. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ciência da Informação*, Brasília, v. 31, n. 2, p. 152-162, maio/ago. 2002. Disponível em: <https://doi.org/10.1590/S0100-19652002000200016>. Acesso em: 22 maio 2021.

VLASSCHAERT, C; TOPF, J. M.; HIREMATH, S. Proliferation of Papers and Preprints During the Coronavirus Disease 2019 Pandemic: Progress or Problems With Peer Review? *Advances in chronic kidney disease*, v. 27, n. 5, 2020, p. 418-426. DOI 10.1053/j.ackd.2020.08.003. Disponível em: [https://www.ackdjournal.org/article/S1548-5595\(20\)30119-1/pdf](https://www.ackdjournal.org/article/S1548-5595(20)30119-1/pdf). Acesso em: 23 dez. 2020.

WANG, X.; LIU, C.; MAO, W.; FANG, Z. The open access advantage considering citation, article usage and social media attention. *Scientometrics*, [s.l.], v. 103, n. 2, p. 555-564, mar. 2015. DOI: <https://doi.org/10.1007/s11192-015-1547-0> . Acesso em: 3 dez. 2021.

WANG, Z; GLÄNZEL, W.; CHEN, Y. The impact of preprints in Library and Information Science: an analysis of citations, usage and social attention indicators. *Scientometrics*, [s.l.], v. 125, p. 1403–1423, 2020. DOI <https://doi.org/10.1007/s11192-020-03612-4>. Disponível em: <https://link.springer.com/article/10.1007/s11192-020-03612-4>. Acesso em: 10 jun. 2022.

WEI, M. Research on Impact Evaluation of Open Access Journals. *Scientometrics*, n. 122, p. 1027–1049, fev. 2020. Disponível em: <https://doi.org/10.1007/s11192-019-03306-6>. Acesso em: 5 mar. 2020.

WEINSTOCK, M. Citation indexes. In: *Encyclopedia of Library and Information Science*. New York: Dekker, 1971. v. 5., p. 16-40.

WHITE, R. et al. Only two out of five articles by New Zealand researchers are free-to-access: a multiple API study of access, citations, cost of Article Processing Charges (APC), and the potential to increase the proportion of open access. *PeerJ*, [s.l.], v. 9, May 2021. Disponível em: <https://peerj.com/articles/11417/> . Acesso em: 8 ago. 2021.

WORLD HEALTH ORGANIZATION. *Emergencies*: International health regulations and emergency committees. 19 Dec. 2019. Disponível em: <https://www.who.int/news-room/q-a-detail/emergencies-international-health-regulations-and-emergency-committees>. Acesso em: 20 maio 2021.

YOUNG, J. S.; BRANDES, P. M. Green and gold open access citation and interdisciplinary advantage: a bibliometric study of two science journals. *The Journal of Academic Librarianship*, [s.l.], v. 46, n. 2, 2020. Disponível em: <https://doi.org/10.1016/j.acalib.2019.102105> . Acesso em: 19 jul. 2021.

ZAHEDI, Z.; COSTAS, R.; WOUTERS, P. How well developed are altmetrics? A cross-disciplinary analysis of the presence of ‘alternative metrics’ in scientific publications. *Scientometrics*, [s.l.], v. 101, p. 1491–1513, 2014. DOI <https://doi.org/10.1007/s11192-014-1264-0>. Disponível em: <https://link.springer.com/article/10.1007/s11192-014-1264-0>. Acesso em: 13 abr. 2022.

ZHAI, F. *et al.* Research progress of coronavirus based on bibliometric analysis. *International Journal of Environmental Research and Public Health*, [s.l.], v. 17, n. 11, p. 1-15, 2020. DOI [10.3390/ijerph17113766](https://doi.org/10.3390/ijerph17113766). Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7312058/>. Acesso em: 4 ago. 2021.

ZHANG, L. *et al.* How scientific research reacts to international public health emergencies: a global analysis of response patterns. *Scientometrics*, [s.l.], v. 124, p. 747–773, 2020. DOI <https://doi.org/10.1007/s11192-020-03531-4>. Disponível em: <https://link.springer.com/article/10.1007%2Fs11192-020-03531-4>. Acesso em: 18 abr. 2021.

ZIMAN, J. *An introduction to science studies*. Cambridge: Cambridge University Press, 1984.

APÊNDICE A

Quadro 14 - Análise dos metadados do *dataset* gerado pela Dimensions na Etapa 1 da coleta de dados

	Metadado	Descrição	Tipo de dado	Exemplo
Identificação bibliográfica	Rank	Número	Número	28696
	Publication ID	Código de identificação da publicação indexada na Dimensions	Código/Identificador	pub.1125681333
	DOI (Digital Object Identifier)	Digital Object Identifier	Código/Identificador	10.1016/j.jaccas.2020.05.068
	PMID (PubMed Indexing Number)	PubMed Indexing Number que é diferente do código utilizado pelo PubMed Central	Código/Identificador	32818146
	PMCID (Número de referência do PubMed Central)	Número de referência do PubMed Central. Pouco usado nos Preprints	Código/Identificador	PMC7074995
	Title	Título da publicação	Texto	Atrial Arrhythmias in a Patient Presenting With Coronavirus Disease-2019 (COVID-19) Infection

	Metadado	Descrição	Tipo de dado	Exemplo
Identificação bibliográfica	Abstract	Resumo	Texto	Utilizing 34 348 severe acute respiratory syndrome coronavirus 2 (SARS CoV-2) nucleic acid amplification test (NAAT) results from 2 health systems, we estimated the clinical sensitivity of a single SARS-CoV-2 NAAT. We found that SARS-CoV-2 NAAT has 82%-97% sensitivity for diagnosing coronavirus disease 2019 among symptomatic patients.
	Acknowledgements	Agradecimentos/créditos	Texto	We thank the clinical and environmental management services personnel at the Cleveland VA Medical Center for their participation in the study.
	Source title	Fonte da publicação. Para os artigos, é listado o periódico. Para os preprints, encontra-se listado o repositório de preprint em que a publicação foi indexada e disponibilizada (arXiv; bioRxiv; medRxiv; Preprint.org)	Texto	Artigos: Cell & Bioscience; Preprint: arXiv
	Anthology title	Campo vazio tanto para artigos como para preprints	Texto	Vazio

	Metadado	Descrição	Tipo de dado	Exemplo
Identificação bibliográfica	Publisher	Editora	Texto	Artigos: Elsevier, Springer, etc; Preprint: Cold Spring Harbor Laboratory
	MeSH terms	Termos descritores de assunto. Campo vazio para preprints	Texto. Palavras-chave	Artigos: Betacoronavirus; COVID-19; Clinical Laboratory Techniques; Coronavirus Infections; Humans; Pandemics; Pneumonia, Viral; RNA, Viral; Real-Time Polymerase Chain Reaction; SARS-CoV-2; Preprints: Vazio
	PubYear	Ano da publicação	Número	2020
	Publication Date (online)	Data da publicação digital. Formato: ano ou ano-mês-dia	Data (aaaa-mm-dd)	2021-10-13
	Publication Date (print) - Só para artigos	Data da publicação impressa. Não registrada para preprint. Formato: ano ou ano-mês ou ano-mês-dia. Variação entre os registros	Data (aaaa-mm) ou (aaaa-mm-dd)	2020-07

	Metadado	Descrição	Tipo de dado	Exemplo
	Volume - Só para artigos	Volume do periódico em que o artigo foi publicado. Dado disponibilizado somente para artigos	Número	9
	Issue - Só para artigos	Número do periódico em que o artigo foi publicado. Dado disponibilizado somente para artigos	Número	4
	Pagination	Paginação. Apresenta valores variados (página de início-página de fim; código; só uma página; DOI)	Número	Variável (27-36; e7386; 40; 10.1111/tbed.13620)
Acesso aberto e tipo de publicação	Open Access	Classificação do tipo e subtipo de acesso da publicação. Disponível tanto para artigo quanto preprint	Texto. Categorias	Closed ou All OA; Green
	Publication Type	Classificação do tipo de publicação. Neste estudo são analisadas apenas duas: artigos e preprints	Texto. Categorias	Article; Preprint
	Authors	Autores da publicação	Texto	Ju, Chun-Rong; Lian, Qiao-Yan; Zhang, Jian-Heng; Qiu, Tao; Cai, Zhi-Tao; Jiang, Wen-Yang; Zhang, Jing; Cheng, Qin; Chen, Gang; Li, Ning; Wang, Chun-Yan; He, Jian-Xing

	Metadado	Descrição	Tipo de dado	Exemplo
Autoria e afiliação	Authors (Raw Affiliation)	Autores da publicação com dados de afiliação institucional como consta na publicação, sem nenhum tratamento prévio	Texto	Shatri, Jeton (Department of Radiology, University Clinical Center of Kosovo, Prishtina, Kosovo, Albania,); Department of Anatomy Faculty of Medicine, University of Prishtina “Hasan Prishtina”, Kosovo, Albanian, .); Tafilaj, Lirim (Department of Infectious Disease, University Clinical Center of Kosovo, Prishtina, Kosovo, Albania,); Turkaj, Agon (Faculty of Medicine, University of Prishtina “Hasan Prishtina”, Kosovo, Albanian, .); Dedushi, Kreshnike (Department of Radiology, University Clinical Center of Kosovo, Prishtina, Kosovo, Albania,); Shatri, Metë (Faculty of Medicine, University of Prishtina “Hasan Prishtina”, Kosovo, Albanian, .); Bexheti, Sadi (Department of Anatomy Faculty of Medicine, University of Prishtina “Hasan Prishtina”, Kosovo, Albanian, .); Mucaj, Serbeze Kabashi (Department of Radiology, University Clinical Center of Kosovo, Prishtina, Kosovo, Albania,)
Autoria e afiliação				

	Metadado	Descrição	Tipo de dado	Exemplo
	Corresponding Authors	Autores da publicação destacados como correspondentes nas publicações	Texto	Tang, Shi-Lin (University of South China; First Affiliated Hospital of University of South China); Tang, Chao-Ke (University of South China; First Affiliated Hospital of University of South China)
	Authors Affiliations	Dados sobre afiliação institucional dos autores	Texto	Sharma, Madhu (Albert Einstein College of Medicine); Gorstein, Samuel (Albert Einstein College of Medicine); Aldrich, Margaret L. (Albert Einstein College of Medicine); Hsu, Daphne T. (Albert Einstein College of Medicine); Choueiter, Nadine F. (Albert Einstein College of Medicine)
Financiamento e	Research Organizations - standardized	Instituição de pesquisa com dados padronizados pela Dimensions	Texto	Texas Southern University
	GRID IDs	Identificadores registrados no GRID, banco de dados global gratuito e disponível abertamente de organizações relacionadas à pesquisa. Responsável por catalogar organizações relacionadas à pesquisa e fornecer a cada uma um identificador único e persistente	Código/Identificador	grid.231844.8; grid.6530.0; grid.416507.1; grid.17063.33; grid.7778.f; grid.7849.2; grid.19006.3e

	Metadado	Descrição	Tipo de dado	Exemplo
localização	City of Research organization	Cidade da instituição de pesquisa	Texto	San Diego; Philadelphia
	State of Research organization	Estado da instituição de pesquisa	Texto	California; California; California; California
	Country of Research organization	País da instituição de pesquisa	Texto	United States; United States
	Funder	Financiador da pesquisa	Texto	Los Alamos National Laboratory; National Institute of General Medical Sciences; Division of Mathematical Sciences
	Funder Group	Grupo financiador da pesquisa	Texto	DoE - United States Department of Energy; ICRP - International Cancer Research Partnership; NIH - National Institutes of Health; NSF - National Science Foundation
Financiamento e localização	Funder Country	País da instituição financiadora da pesquisa	Texto	United States; United States; United States

	Metadado	Descrição	Tipo de dado	Exemplo
	UIDs of supporting grants	Código relacionado a informações de financiamento	Código	grant.8349242; grant.8368371
	Supporting Grants		Número	Artigos: 81570408; 81770461; Preprints: FI2GM137804
Métricas	Times cited	Quantidade de citações recebidas. Disponível tanto para artigos quanto para preprints	Número	12
	Recent citations	Número de citações recebidas nos últimos dois anos. Disponível tanto para artigos quanto para preprints	Número	6
	RCR (Relative Citation Ratio)	É uma medida baseada na citação da influência científica de uma publicação. É calculado como as citações de um artigo, normalizadas para as citações recebidas por publicações financiadas pelo NIH na mesma área de pesquisa e ano. Explicação em: https://dimensions.freshdesk.com/support/solutions/articles/23000018841-what-is-the-rcr-how-is-the-rcr-score-calculated-	Número	0,93

	Metadado	Descrição	Tipo de dado	Exemplo
Métricas	FCR (Field Citation Ratio)	É uma medida baseada na citação da influência científica de um ou mais artigos. É calculado dividindo o número de citações que um artigo recebeu pelo número médio recebido por documentos publicados no mesmo ano e na mesma categoria de Fields of Research (FoR) utilizadas pela Dimensions. Explicação em: https://dimensions.freshdesk.com/support/solutions/articles/23000018848-what-is-the-fcr-how-is-it-calculated-	Número	5,11
	Altmetric	Valor correspondente ao <i>Altmetric Attention Score</i>	Número	24
Fontes	Source Linkout	Link da fonte	Link	https://academic.oup.com/jpids/article-pdf/9/3/370/33492964/piaa068.pdf
	Dimensions URL	Link do registro da publicação indexada na dimensions	Link	https://app.dimensions.ai/details/publication/pub.1128377612

	Metadado	Descrição	Tipo de dado	Exemplo
Áreas/Assuntos/ Temas	FOR (ANZSRC) Categories	Classificação utilizada pela Dimensions para categorizar as áreas do conhecimento. Baseia-se na Australian and New Zealand Standard Research Classification (ANZSRC). Lista completa disponível em: https://app.dimensions.ai/browse/categories/publication/for	Texto	11 Medical and Health Sciences; 1103 Clinical Sciences
	RCDC Categories	Classificação temática	Texto. Categorias	Prevention; Infectious Diseases; Pneumonia & Influenza; Biodefense; Pneumonia; Vaccine Related; Lung; Immunization; Emerging Infectious Diseases
	HRCS HC Categories	Classificação temática	Texto. Categorias	Infection; Respiratory
	HRCS RAC Categories	Classificação temática	Texto. Categorias	2.1 Biological and endogenous factors
	ICRP Cancer Types	Classificação para tipos de câncer	Texto. Categorias	Lung Cancer

	Metadado	Descrição	Tipo de dado	Exemplo
Áreas/Assuntos/ Temas	ICRP CSO Categories	Classificação temática	Texto. Categorias	5.3 Systemic Therapies - Discovery and Development
	Units of Assessment	Classificação temática	Texto	A01 Clinical Medicine
	Sustainable Development Goals	Classificação dos Objetivos de Desenvolvimento Sustentável. Dado observado em poucos registros. Lista completa disponível em: https://app.dimensions.ai/browse/categories/publication/sdg	Texto. Categorias	3 Good Health and Well Being