

**A PANDEMIA DE COVID-19 E SUA REPRESENTAÇÃO EM ARTIGOS
PUBLICADOS EM PORTUGUÊS NA ÁREA DE SAÚDE PÚBLICA**
um olhar da indexação automática

Fatima Cristina Lopes dos Santos¹
Fundação Oswaldo Cruz
fatimaclsantos@gmail.com

Cícera Henrique da Silva²
Fundação Oswaldo Cruz
cicera.henrique@icict.fiocruz.br

Rosane Abdala Lins³
Fundação Oswaldo Cruz
rosane.abdala@icict.fiocruz.br

Resumo

Compreendendo a indexação como representação de conteúdo e seu importante papel no sistema de comunicação na ciência, buscou-se refletir sobre a informação em saúde representada em artigos de periódicos sobre a COVID-19, entendendo a relevância da visibilidade e recuperação dessa produção. Vale ressaltar que o periódico científico - e por conseguinte seus artigos - além de outros papéis estratégicos na comunicação da ciência, é registro da memória, chancela de qualidade e difusor do conhecimento. Com o objetivo de avaliar a aplicação da indexação automática em artigos publicados em português na área de saúde pública, este estudo exploratório baseou-se em 15 artigos publicados em 2020 e distribuídos em 4 periódicos integrantes da Coleção Saúde Pública na SciELO Brasil. Fez-se a análise entre o conteúdo dos trabalhos e sua representação (palavras-chave e descritores DeCS), observando o uso da indexação automática e da linguagem controlada na representação temática dessa informação em saúde. A metodologia dividiu-se nas seguintes etapas: 1. Seleção da fonte de informação; 2. Coleta da informação; 3. Tratamento e análise e dos dados coletados; 4. Mapeamento da atribuição de termos nos artigos científicos estudados, observando o papel da indexação automática nesse processo. O quadro avaliativo com a aplicação da indexação automática nos artigos analisados verificou que esse processo tem potencial para contribuir com a coerência na representação do conteúdo dos artigos estudados. Foi possível verificar que diversos termos foram identificados em ambos os processos, incluindo vários reconhecidos como descritores autorizados pelo DeCS.

Palavras-chave: Indexação Automática. Mineração de Textos. Artigos Científicos. Saúde Pública. COVID-19.

**THE COVID-19 PANDEMIC AND THE REPRESENTATION IN ARTICLES
PUBLISHED IN PORTUGUESE IN THE PUBLIC HEALTH AREA**
a look at automatic indexing

¹Doutoranda do Programa de Pós-Graduação em Informação e Comunicação em Saúde (PPGICS/ICICT/Fiocruz). Tecnologista em Saúde-Bibliotecária da Escola Nacional de Saúde Pública Sergio Arouca (ENSP/Fiocruz). Orcid: <https://orcid.org/0000-0003-0737-565X>.

²Doutora em Ciência da Informação e da Comunicação, Université Aix-Marseille III. Professora do Programa de Pós-Graduação em Informação e Comunicação em Saúde (PPGICS/ICICT/Fiocruz). Pesquisadora colaboradora do Laboratório de Informação Científica e Tecnológica em Saúde do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (LICTS/ICICT/Fiocruz). Orcid: <https://orcid.org/0000-0002-3298-795X>.

³Doutora em Informação e Comunicação em Saúde (PPGICS/ICICT/Fiocruz). Professora do Programa de Pós-Graduação em Informação e Comunicação em Saúde (PPGICS/ICICT/Fiocruz). Pesquisadora do Laboratório de Informação Científica e Tecnológica em Saúde do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (LICTS/ICICT/Fiocruz). Orcid: <https://orcid.org/0000-0002-7518-3265>.



Abstract

Understanding indexing as a representation of content and its important role in the communication system in science, we sought to reflect on the health information represented in journal articles about COVID-19, understanding the relevance of the visibility and recovery of this production. It is worth mentioning that the scientific journal - and therefore its articles - in addition to other strategic roles in the communication of science, is a record of memory, seal of quality and disseminator of knowledge. With the objective of evaluating the application of automatic indexing in articles published in Portuguese in the area of public health, this exploratory study was based on 15 articles published in 2020 and distributed in 4 journals that are part of the Public Health Collection at SciELO Brazil. An analysis was carried out between the content of the works and their representation (key words and DeCS descriptors), observing the use of automatic indexing and controlled language in the thematic representation of this health information. The methodology was divided into the following steps: 1. Selection of the source of information; 2. Collection of information; 3. Treatment and analysis of collected data; 4. Mapping of the assignment of terms in the scientific articles studied, observing the role of automatic indexing in this process. The evaluation framework with the application of automatic indexing in the analyzed articles verified that this process has the potential to contribute to the coherence in the representation of the content of the articles studied. It was possible to verify that several terms were identified in both processes, including several recognized as descriptors authorized by DeCS.

Keywords: Automatic Indexing. Text Mining. Scientific Articles. Public Health. COVID-19.

1 INTRODUÇÃO

Para obtenção de alto grau de precisão na recuperação da informação na *web* - além da especificidade da linguagem de indexação - faz-se necessária a coerência na representação da informação, visando a busca orientada em ambientes virtuais organizados para tal. Assim, podemos associar o grau de coerência da indexação - extensão com que há concordância quanto aos termos a serem usados para indexar um documento (LANCASTER, 2004) - com o grau de precisão de um sistema de recuperação da informação, isto é, o desempenho desse sistema quando a maioria dos itens recuperados é considerada útil ou inútil por quem realiza ou solicita a busca nesse ambiente (ARAÚJO JÚNIOR, 2007).

Quando esses meios virtuais tratam de informação em saúde e do processo de comunicação científica, a representação da informação envolve o complexo modelo de negócio de periódicos científicos - ainda o principal canal de comunicação e divulgação entre pares. Os procedimentos para a publicação de informação científica foram estabelecidos e consolidados há mais de três séculos, beneficiando editores de periódicos científicos com as assinaturas de periódicos (KURAMOTO, 2006). Os resultados obtidos por um pesquisador e publicados em artigos de periódicos são frequentemente retomados por outros cientistas, que dão continuidade ao estudo, fazendo avançar a ciência, produzindo tecnologias e produtos baseados nesses resultados (CAMPELLO; CENDÓN; KREMER, 2007).

Dado que uma das funções do periódico científico é ser um canal formal de comunicação na ciência, a coerência na representação de artigos científicos contribui para a organização dessa informação, unificando padrões em sua representação temática. A criação e circulação de terminologias em distintos cenários comunicativos são testemunhos de que

essas cumprem, prioritariamente, a dupla função de fixar o conhecimento técnico-científico e de promover sua transferência de modo pontual (KRIEGER; FINATTO, 2004).

Sendo assim, vale destacar que a informação científica e tecnológica em saúde vem se constituindo em uma área interdisciplinar por excelência, tornando fundamental a busca por propostas teóricas e metodológicas que incentivem a articulação entre diversas áreas do conhecimento. Esse processo de representação da informação em saúde e sua visibilidade tem sido objeto de reflexão e de aprofundamento: em estudos já publicados, verificou-se maior adequação das palavras-chave dos artigos estudados com o vocabulário da área de saúde - Descritores em Ciências da Saúde (DeCS) - do que com o conteúdo dos mesmos, o que pode acarretar num distanciamento da representatividade do conteúdo desses textos (SANTOS, 2017; SANTOS; MOLLICA; GUEDES, 2019; SANTOS; SILVA; LINS, 2019). Sobre isso, Lancaster (2004) recomenda que a etapa de tradução - que envolve a conversão de conceitos do texto em termos de indexação - não deve ser influenciada pelas características do vocabulário a ser usado.

Sob esse ponto de vista, independentemente do processo de indexação realizado - em linguagem natural ou controlada - o autor ao construir seu texto, seleciona da terminologia de seu campo de conhecimento, termos/conceitos que ele julga adequados para explicitação e comunicação de suas ideias e reflexões. Ele molda seu discurso de acordo com seus propósitos comunicativos, e o representa tematicamente por palavras-chave, com o objetivo de chamar a atenção do leitor para os conceitos que ele considera mais importantes em sua produção (FÓRIS, 2013).

A partir deste cenário, o estudo buscou evidenciar a importância da coerência na indexação da informação em saúde. A expectativa com o desenvolvimento desta pesquisa é contribuir com estudos sobre a indexação de artigos científicos na área de saúde pública, particularmente na representação temática de artigos sobre o novo coronavírus indexados na base SciELO Saúde Pública (SciELO SAÚDE PÚBLICA, 2022).

2 JUSTIFICATIVA

O cientista legitima suas atividades perante a sociedade apresentando soluções e benefícios para a comunidade - local ou mundial - onde seu estudo é desenvolvido. Essa atitude não teve início nas instituições acadêmicas, mas em práticas realizadas individual ou coletivamente, com a troca de experiências feita em diversos ambientes, incluindo salões e cafés. O cientista profissional do século XIX surge a partir de uma tradição semiprofissional,

e para continuar suas pesquisas e receber estímulos - incluindo os financeiros - as instituições acadêmicas e de fomento foram essenciais, como ainda são (VELHO, 1985; BURKE, 2003).

E por buscar compreender os fenômenos da natureza e os problemas que surgem perante a humanidade, o pesquisador chega a respostas que tendem a ser provisórias e sujeitas a transformações ao longo do tempo. Esses resultados estimulam e orientam a evolução humana e interferem na identidade dos povos e das nações, estabelecendo verdades fundamentais de cada época (TARGINO, 2000). Dessa forma, a ciência é uma atividade que está em constante movimento, pois suas verdades e valores dependem da sociedade - da qual o cientista é integrante - que a elabora.

Quando um texto científico é publicado, seu autor visa ser identificado em sua comunidade como produto da ciência certificada, que ao ser avaliado por seus pares, passa a incorporar o conhecimento produzido naquela área do conhecimento. Ao ser publicado num periódico científico, a possibilidade de recuperação desse texto aumenta quanto melhor seu conteúdo estiver representado. McGarry (1999, p.130) diz que:

A função da linguagem escrita como extensão da memória depende em muito de sua capacidade de representar informações de forma sistemática e conveniente [...] deve representar uma mensagem coerente e compreensível

36

Para Araújo Júnior (2007) a representação do conteúdo de um documento está na base do conceito de indexação, bem como seus conceitos relacionados: descritor, linguagem de indexação e termo de indexação. A indexação busca identificar os assuntos contidos no texto de um documento, traduzindo-os para uma linguagem de indexação natural - palavras que ocorrem no texto - ou linguagem de indexação controlada - termos/descriptores autorizados em vocabulário controlado (LANCASTER, 2004). Quanto mais específica a representação das informações identificadas no texto em análise, maior a probabilidade de recuperação de informações relevantes pelo sistema.

Na indexação, termos e descritores podem ser vistos como elementos de representação e disseminação do conhecimento, mas distinguem-se pela sua natureza e características básicas. O termo é uma unidade léxica que assume valor semântico próprio numa área de conhecimento, pela razão de integrar uma comunicação especializada. O descritor é um componente de uma linguagem constituída por gestores de informação, sendo, portanto, artificial (KRIEGER; FINATTO, 2004). A etapa de tradução de um documento, que segundo Lancaster (2004) envolve a conversão da análise conceitual num conjunto de termos de

indexação, é de suma importância para a representação e está relacionada com a escolha de palavras-chave e de termos autorizados em um vocabulário controlado.

Sales (2007) afirma que as linguagens artificiais auxiliam as atividades de representação de conteúdos informacionais, viabilizando um armazenamento de informações que poderão ser facilmente recuperáveis. Bocatto e Torquetti (2012) defendem que o uso de uma linguagem de indexação controlada permite a representação na indexação e a pesquisa por assunto com maior exatidão e equidade com as necessidades informacionais dos usuários.

Lancaster (2004) acrescenta que a coerência na indexação se refere à extensão com que há concordância quanto aos termos a serem usados para indexar o documento. O autor assegura que quanto mais termos retirados do próprio conteúdo do texto, para indexação do documento, maior será a probabilidade de uma representação de qualidade. Essa qualidade influi diretamente na tarefa de recuperação da informação: se os descritores selecionados para representar o conteúdo de um documento não forem coerentes, certamente não será recuperado com facilidade, comprometendo o processo como um todo (ARAÚJO JÚNIOR, 2007). Isto é, a recuperação do documento mais pertinente à questão de busca é aquela cuja indexação proporciona a identificação de conceitos mais pertinentes ao seu conteúdo, produzindo uma correspondência com o assunto pesquisado (FUJITA, 2003).

A pesquisa científica necessita de comunicação para reivindicar sua importância no campo de estudo, e a divulgação dos resultados de pesquisa à comunidade científica e sua posterior aprovação por essa comunidade é o que assegura credibilidade e visibilidade aos resultados (LATOURE; WOOLGAR, 1997; MEADOWS, 1999). Para alcançar esse crédito na comunidade científica, o trabalho intelectual do pesquisador depende do processo de comunicação científica, que possui canais caracterizados como formais e informais. Ambos são indispensáveis para a operacionalização da construção das pesquisas, mas possuem cronologias distintas e são utilizados em momentos diferentes nesse processo de comunicação (CHRISTOVÃO, 1979; TARGINO, 2000).

Na comunicação formal, a informação passa pelo crivo de uma comissão de avaliação e é divulgada por meio de livros, artigos em periódicos, resumos ou trabalhos completos publicados em anais de eventos, refletindo maior rigidez e controle. Uma das grandes vantagens dos canais formais é a garantia de armazenamento e recuperação, tendo em vista que geralmente esses canais passam por controles bibliográficos (CHRISTOVÃO, 1979; VANZ; SILVA FILHO, 2019). A comunicação informal possui um caráter mais pessoal e não passa pelo crivo de uma comissão de avaliação, a troca de informação ocorre por meio de contatos interpessoais, com a vantagem de maior atualização e menor custo. São os encontros

presenciais entre pesquisadores (congressos, seminários, etc.), e-mails, blogs, mídias acadêmicas (Mendeley, Research Gate, etc.), mídias sociais (Facebook, Twitter, Instagram, You Tube, etc.). Os trabalhos de congressos já não são considerados estritamente informais, sendo conceituados por Christovão (1979) como semiformais.

E diante do volume de informação disponível hoje, o processo de comunicação científica torna-se cada vez mais intrincado. A internet modificou profundamente o modo de elaborar e acessar informação, disponibilizando uma quantidade significativa de produção científica, permitindo uma comunicação dinâmica e abrangente, multiplicando o número de periódicos científicos *on-line* publicados em diversas áreas do conhecimento. Essas inovações viabilizam a disseminação da pesquisa científica, provocando mudanças no modo como o conhecimento científico é difundido e acessado (MIRANDA; CARVALHO; COSTA, 2018). Tanto a comunicação informal quanto a formal são amplamente divulgadas por esses meios, e seus conceitos começam a apresentar dificuldades pela falta de clareza em suas características (CHRISTOVÃO, 1979; MUELLER, 2007).

Mesmo na dificuldade de caracterizar esses meios, o canal formal ainda é o mais utilizado para legitimar o conhecimento científico. O veículo mais utilizado no sistema formal de comunicação científica é o periódico, que surge na Europa na segunda metade do século XVII pela necessidade de comunicação dos novos conhecimentos e invenções entre os cientistas da época. Entre outras razões para o surgimento do periódico científico, temos a necessidade de um debate coletivo de questões da ciência, a urgência de comunicação com um público crescente e a expectativa de lucro por parte de editores. Para o pesquisador-autor, o artigo publicado em um periódico é uma informação relativamente rápida de produzir, fazendo com que ideias e/ou resultados de pesquisa sejam conhecidos com agilidade (MEADOWS, 1999; STUMPF, 2005).

Assim, neste trabalho foi realizado um exercício de indexação na expectativa de poder contribuir para a representação de artigos publicados nos periódicos brasileiros em português em saúde pública. A necessidade de uma investigação mais aprofundada, principalmente se tratando de artigos publicados em português na área de saúde e sua importância na comunicação científica nacional (SANTOS; MOLLICA; GUEDES, 2019) foi a principal motivação para o desenvolvimento deste estudo. Mesmo com mais de 300 anos e todas as questões apresentadas no texto - além do surgimento da *web 2.0* (também conhecida como *web social* ou *web interativa*) e a introdução de novos veículos na comunicação científica - a relevância do periódico científico e seu formato de publicação em artigos não foram alterados.

Até mesmo alguns problemas se tornaram comuns aos impressos e eletrônicos, como a preservação, representação e recuperação de seus conteúdos.

Dentre as ações que influenciam no acesso aos artigos de periódicos, a indexação é fundamental, pois identifica e atribui conceitos pertinentes ao conteúdo do documento, facilitando a precisão na recuperação. A indexação aponta os assuntos contidos num documento, traduzindo-os para uma linguagem de indexação natural - palavras que ocorrem no texto - ou linguagem de indexação controlada - termos/descriptores autorizados em vocabulário controlado (ABNT, 1992; LANCASTER, 2004). A linguagem de indexação viabiliza o controle de vocabulário num sistema de informação. Sem a mediação dessas linguagens a recuperação da informação pode ficar inviável, com um sistema inconsistente devido à falta de controle de um vocabulário (CRUZ; FUJITA, 2021).

Dessa forma, as bases de indexação são importantes colaboradoras na recuperação dos conteúdos, aumentando o alcance da divulgação e exposição dos artigos publicados nos periódicos aceitos em seus ambientes (SOUZA-SANTOS, 2010). Quanto mais específica a representação dos conteúdos identificados no texto em análise, maior a probabilidade de recuperação de temas relevantes pelo sistema. Nos periódicos científicos, refere-se ao processo de indexação de seus títulos em bases de dados, nacionais ou internacionais, que tem por função disseminar e divulgar a produção científica de uma determinada área, aumentando a visibilidade dos textos publicados e influenciando a comunicação científica.

Esse processo é visto como um reflexo da excelência de um periódico, que é considerado de maior qualidade científica em comparação aos periódicos que não são indexados. Isso faz com que as editoras busquem se adequar aos critérios dessas bases indexadoras para ingressarem nesse grupo seletivo. O trabalho publicado em periódicos indexados permite o acesso quase que imediato - no caso de publicações em acesso aberto - ao seu conteúdo, e ainda é visto como a forma mais legítima de reconhecimento e prestígio acadêmico e institucional (VETTORE; COIMBRA JR, 2010).

É importante ressaltar o que Lancaster (2004) escreve sobre a indexação de um documento: não há um conjunto correto de termos, isso depende para que tipo de público esse documento será representado/recuperado. Mas é de suma importância a pesquisa dos itens e expressões utilizadas, tanto para a construção de um vocabulário controlado quanto para a escolha dos termos que irão representar o conteúdo de um texto científico, pois supõe-se que o autor de um texto científico ao publicar seus achados de pesquisa utiliza-se primeiro de linguagem natural, representando assim sua área de conhecimento. Esse autor molda seu discurso de acordo com seus propósitos comunicativos, com o objetivo de chamar a atenção

do leitor - seu par na comunidade científica - para os conceitos que ele considera importantes no texto.

Quando o texto científico - principalmente o artigo de periódico que possui grandes restrições de formato - se constrói de forma coerente, numa atribuição de termos que representarão devidamente os conceitos inseridos no texto, a escolha dos termos para sua representação será mais eficaz. Guedes (2010) defende que o artigo de periódico permanece no centro de uma teia formada por gêneros textuais de publicações de processo de pesquisa e que as interrelações são bem conhecidas e reconhecidas no ambiente acadêmico. Para Fóris (2013), o texto é uma rede de conceitos que devem estabelecer coerência; se essa rede não estiver bem construída por termos representativos desses conceitos, ela poderá se tornar inconsistente (FÓRIS, 2013). Fortalecendo a importância da indexação, Migués et al. (2013) escrevem que:

O uso das palavras-chave potencializa o acesso ao conteúdo dos documentos, para além da informação que é representada pelo título e resumo; traduz o pensamento dos autores, e mantém o contato com a realidade da prática cotidiana, acompanhando a evolução científica e tecnológica, que é refletida pelos documentos. (MIGUÉS et al., 2013, p.3).

40

A unidade terminológica, isto é, o termo, é simultaneamente, tanto elemento constitutivo da produção do saber quanto componente linguístico, cujas propriedades favorecem a univocidade da comunicação especializada (KRIEGER; FINATTO, 2004), e tal processo pode ser concretizado por meio da atribuição de palavras-chave ou descritores. Segundo Krieger e Finatto (2004) a existência e a circulação de terminologias em distintos cenários comunicativos são testemunhos de que essas cumprem, prioritariamente, a dupla função de fixar o conhecimento técnico-científico e de promover sua transferência de modo pontual.

Na saúde pública o vocabulário mais indicado pelos periódicos da área é o Descritores em Ciências da Saúde (DeCS), produzido pelo Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde (Bireme), criado em 1967 pela Organização Pan-Americana da Saúde/Organização Mundial da Saúde (OPAS/OMS) para cooperação técnica em informação científica em saúde. O DeCS foi desenvolvido a partir do Medical Subject Headings (MeSH) da United States National Library of Medicine (NLM) e é um vocabulário trilingue que serve como linguagem única de indexação de artigos de revistas científicas, livros, anais de congressos, relatórios técnicos e outros tipos de materiais (DECS, 2022).

Assim, o profissional de informação, que irá indexar o texto publicado pelos cientistas, muitas vezes terá a recomendação de utilizar descritores autorizados pelo vocabulário controlado da área. Um estudo preliminar analisou orientações de uso de vocabulários controlados em 78 periódicos científicos brasileiros na área de saúde pública, e verificou que 44 periódicos recomendam ou informam a obrigatoriedade de uso de vocabulário controlado, prevalecendo a indicação do DeCS, o que aponta a preocupação desses periódicos com a padronização dos termos utilizados na representação temática de seus artigos (SANTOS; SILVA; LINS, 2019).

Cabe então aqui discutir sobre uma indexação automática que auxilie tanto o autor na representação de seus trabalhos quanto o profissional de informação no momento de indexar esses documentos, no intuito de colaborar com a recuperação e visibilidade dessa produção. Esse tipo de indexação utiliza programas que analisam o conteúdo dos textos e apresentam termos que podem ser os descritores desses textos (GIL LEIVA, 1999). Para Lancaster (2004) existem dois tipos de indexação automática: por extração, que extrai e seleciona de forma automática termos do texto; e por atribuição, onde termos são atribuídos automaticamente de um vocabulário controlado. Na indexação automática por atribuição, o vocabulário controlado atua no processo de análise automática do texto do documento e na sua representação, isto é, o resultado dessa análise é condicionado à atribuição dos descritores desse vocabulário (NARUKAWA; GIL LEIVA; FUJITA, 2009).

Até onde se sabe, alguns protótipos com base na indexação automática da informação foram implementados em outras áreas do conhecimento, como o estudo realizado em textos sobre habitação (ARAÚJO JUNIOR; TARAPANOFF, 2006; ARAÚJO JUNIOR, 2007). Na área das ciências da saúde conhece-se o estudo de Narukawa, Gil Leiva e Fujita, (2009) e a pesquisa de Marcondes, Costa e Martins (2016) com experiências baseadas na web semântica.

É neste contexto de pesquisa aplicada na área que este estudo investigou a representação temática de artigos publicados em português na área de saúde pública, por meio da análise da coerência semântica entre o conteúdo dos trabalhos e sua representação, isto é, de que forma o conteúdo desses textos se reflete em suas palavras-chave/descriptores e nas palavras/termos identificados via mineração de textos. Nesse sentido, Strehl afirma que:

O conceito de coerência é um dos quesitos para uma indexação de qualidade, já que todos os documentos que tratam sobre o mesmo assunto devem estar representados da mesma forma, não havendo perda de informações no momento da recuperação (STREHL, 1998, p.330).

Assim, compreendendo a indexação como representação de conteúdo e seu importante papel no sistema de comunicação ciência, refletiu-se sobre como a pandemia de COVID-19 está representada em textos da área de saúde pública, especificamente em artigos, entendendo a relevância da visibilidade e recuperação dessa produção. Cabe aqui lembrar Guanaes e Guimarães (2012) para quem o periódico científico, e por conseguinte seus artigos, além de outros papéis estratégicos na comunicação da ciência, são registros da memória, chancela de qualidade e difusores do conhecimento.

Esses papéis se atualizam neste momento com o surgimento da COVID-19 (do inglês *Coronavirus Disease* 2019), doença infecciosa causada pelo coronavírus da Síndrome Respiratória Aguda Grave 2 (SARS-CoV-2) e identificada pela primeira vez em dezembro de 2019, em Wuhan, na China. A Organização Mundial da Saúde (OMS) declarou, em 30 de janeiro de 2020, que o surto dessa nova doença constituiu uma Emergência de Saúde Pública de Importância Internacional - o mais alto nível de alerta da Organização - e em 11 de março de 2020 a COVID-19 foi caracterizada pela OMS como uma pandemia. Até 28 de julho deste ano, foram confirmados no mundo 571.198.904 casos, com 6.387.863 mortes. Na Região das Américas, 169.649.287 pessoas foram infectadas pelo novo coronavírus, no Brasil já são 33.659.879 casos confirmados e 677.494 mortes, conforme dados da Organização Pan-Americana da Saúde (OMS, 2022; OPAS, 2022).

Numa emergência de saúde como essa, a comunicação na ciência é imprescindível para a atualização de dados sobre pesquisas, tratamentos, protocolos e processos relacionados ao novo coronavírus e à COVID-19. Torna-se essencial compreender de forma aprofundada o significado e as repercussões sociais, econômicas e humanitárias de uma crise complexa e de múltiplas dimensões, que afeta de forma significativa o futuro das nações (CARVALHO; LIMA; COELI, 2020). Cientistas de todo o mundo se mobilizam na busca de alternativas que diminuam o número de vidas perdidas e o sofrimento de milhões de pessoas, fazendo com que o volume de artigos produzidos aumente exponencialmente com contribuições de diversas áreas do conhecimento. Essa mobilização resultou - entre outras ações - em diversas vacinas que desde o início de suas aplicações já passam de 12 bilhões de doses administradas globalmente (OMS, 2022).

Nessas situações de urgência, a comunicação científica mantém processos de avaliação que influenciam no tempo de publicação de um artigo. Dependendo do periódico, após o aceite do texto esperam-se meses para que o artigo seja publicado. Na área da saúde - nos periódicos mais concorridos nacionais e internacionais - o tempo entre a submissão e a aceitação e/ou publicação pode chegar, em média, a 12/18 meses (FERRARI, 2018).

O sistema de *fast-track* acelera esse processo de avaliação. Nesse tipo de procedimento, os textos sobre o assunto em destaque, atualmente a pandemia de COVID-19, recebem tratamento prioritário dos editores e seguem para a avaliação dos pares com mais agilidade. Os periódicos podem adotar fluxos diferentes, alguns organizam ações internas para que os estudos sejam priorizados, outros criam áreas específicas em seus sites, permitindo o acesso aos textos e indicando que os trabalhos estão sob análise, sendo chamados de *preprints* nesses casos. O período de aceitação de artigos neste fluxo também fica a critério do periódico, podendo ser aberto e encerrado de acordo com o interesse das editoras.

Buscando elaborar um recorte sobre a representação dessa produção sobre COVID-19, foi avaliada a aplicação da indexação automática em artigos publicados em português na área de saúde pública. Para isso, a pesquisa realizou-se em quatro etapas:

1. Seleção da fonte de informação;
2. Coleta da informação;
3. Tratamento e análise e dos dados coletados;
4. Mapeamento da atribuição de termos nos artigos científicos estudados, observando o papel da indexação automática e do grau de coerência semântica nesse processo.

Com o estudo, espera-se contribuir com as pesquisas sobre a indexação de artigos científicos na área de Saúde Pública, particularmente na coerência da representação temática de artigos sobre o novo coronavírus.

3 METODOLOGIA

Para atingir o objetivo proposto, esta pesquisa quanti-qualitativa seguiu o processo de mineração de textos, que consiste no uso de técnicas de análise e extração de dados a partir de textos, frases ou apenas palavras em grandes volumes de produções textuais (ARAÚJO JUNIOR, 2007; FERREIRA; CORREA, 2020). A seguir descrevem-se as etapas já citadas:

3.1 SELEÇÃO DA FONTE DE INFORMAÇÃO

Na área de ciências da saúde algumas fontes de informação especializadas são utilizadas com mais frequência: as bases de dados e os vocabulários controlados (OLIVEIRA; ALMEIDA; SOUZA, 2015). Como orientação para escolha de uma base de dados para pesquisa, Lancaster (2004) indica os seguintes critérios para avaliação, de acordo com a utilidade ao responder as necessidades de informação: 1. Cobertura: relacionado ao volume do conteúdo sobre determinado assunto disponível na base de dados; 2. Recuperação: os dados devem ser recuperados por meio de uma estratégia de busca não muito complexa; 3.

Previsibilidade: o documento deve ser reconhecido como relevante a partir das informações contidas na base de dados, isto é, com que eficiência o pesquisador pode aferir se aquele documento será útil ou não para seu estudo; 4. Atualidade: velocidade com que novas publicações são incluídas na base.

Sendo assim, esta pesquisa utilizou a fonte SciELO Saúde Pública, por se tratar de um acervo consolidado sobre a área que será estudada e congrega periódicos de renome em seus países de origem, em sua maioria indexados em bases de dados internacionais como Medline, Web of Science e Scopus (SciELO SAÚDE PÚBLICA, 2022).

Lançado em 1998, a Scientific Electronic Library Online (SciELO) é um programa internacional, implementado por meio de uma rede descentralizada de coleções nacionais de periódicos revisados por pares de diversos países que disponibilizam via acesso aberto mais de mil periódicos e cerca de 50 mil artigos por ano (SciELO SAÚDE PÚBLICA, 2022). A coleção Saúde Pública faz parte desse programa e possui 18 títulos de periódicos, divididos entre as temáticas Ciências da Saúde e Humanidades que, até o momento, já divulgou cerca de 2.400 números e 47.000 artigos.

Dado que a pesquisa almejou investigar a representação temática de artigos sobre a pandemia de COVID-19 publicados em português, foram incluídos no estudo somente periódicos brasileiros indexados na SciELO Saúde Pública que aceitam artigos em português e indicam o uso do vocabulário DeCS. Portanto, a partir dos critérios elaborados para a seleção dos periódicos do estudo, apresentam-se a seguir os títulos em que os artigos foram coletados:

Quadro 1 - Periódicos brasileiros da Coleção SciELO Saúde Pública que aceitam artigos em português e usam o DeCS

Título/Área	Editor	ISSN	Idiomas aceitos
1. Cadernos de Saúde Pública Ciências da Saúde	Fundação Oswaldo Cruz/ Escola Nacional de Saúde Pública Sergio Arouca	0102-311X 1678-4464	Português; Inglês; Espanhol
2. Ciência & Saúde Coletiva Ciências da Saúde	Associação Brasileira de Saúde Coletiva (ABRASCO)	1413-8123 1678-4561	Português; Inglês; Espanhol; Francês
3. Epidemiologia e Serviços de Saúde Ciências da Saúde	Ministério da Saúde/ Secretaria de Vigilância em Saúde	1679-4974 2237-9622	Português; Espanhol
4. Revista Brasileira de Epidemiologia Ciências da Saúde	Associação Brasileira de Saúde Coletiva (ABRASCO)	1415-790X	Português; Inglês; Espanhol
5. Revista de Saúde Pública Ciências da Saúde	Universidade de São Paulo/ Faculdade de Saúde Pública	0034-8910 1518-8787	Português; Inglês; Espanhol
6. Saúde em Debate Ciências da Saúde	Centro Brasileiro de Estudos de Saúde (CEBES)	0103-1104 2358-2898	Português; Inglês; Espanhol

Fonte: Elaboração das autoras

O vocabulário DeCS também é visto como fonte de informação, já que seus descritores foram objetos do estudo. Após seleção dos artigos e coleta dos termos mais representativos desses textos - via palavras-chave e mineração de texto - esses termos foram pesquisados no DeCS para verificar seu reconhecimento, ou não, pelo vocabulário.

3.2 COLETA DA INFORMAÇÃO

O principal objeto de análise deste estudo é o artigo de periódico sobre a pandemia de COVID-19 publicado em português na coleção SciELO Saúde Pública no ano de 2020. Após testes com estratégias de busca, a pesquisa sistematizada não levou em consideração as tipologias: comunicação rápida, relato breve, editorial, resenha, relato de caso e carta (tipos de literatura identificados na coleção pesquisada).

Os artigos fazem parte de um universo de 917 textos coletados em periódicos da SciELO Saúde Pública (quadro 1) para uma pesquisa em que, por meio de uma amostragem aleatória, chegou-se à 106 textos. Aplicando o recorte temático sobre COVID-19 nessa amostra, identificaram-se 15 artigos para este estudo (figura 1).

Figura 1 - Universo e amostra do estudo



Fonte: Elaboração das autoras

A partir disso, apresentam-se os artigos incluídos nesta pesquisa:

Quadro 2 - Artigos selecionados para o estudo

1	ALMEIDA, Celia; CAMPOS, Rodrigo Pires. Multilateralismo, ordem mundial e Covid-19: questões atuais e desafios futuros para a OMS. <i>Saúde Debate</i> , v.44, n.esp, p.13-39, 2020. Disponível em: <10.1590/0103-11042020E401>.
2	ALPINO, Tais de Moura Ariza et al. COVID-19 e (in)segurança alimentar e nutricional: ações do Governo Federal brasileiro na pandemia frente aos desmontes orçamentários e institucionais. <i>Cad. Saúde Pública</i> , v.36, n.8, e00161320, 2020. Disponível em: <10.1590/0102-311X00161320>.
3	BEZERRA, Évilly Carine Dias et al. Análise espacial das condições de enfrentamento à COVID-19: uma proposta de Índice da Infraestrutura da Saúde do Brasil. <i>Ciência & Saúde Coletiva</i> , v.25, n.12, p.4957-4967, 2020. Disponível em: <10.1590/1413-812320202512.34472020>.
4	CASTRO, Jorge Abrahão. Proteção social em tempos de Covid-19. <i>Saúde Debate</i> , v.44, n.esp, p.88-99, 2020. Disponível em: <10.1590/0103-11042020E405>.
5	CAVALCANTE, João Roberto et al. COVID-19 no Brasil: evolução da epidemia até a semana epidemiológica 20 de 2020. <i>Epidemiol. Serv. Saude</i> , v.29, n.4, e2020376, 2020. Disponível em: <10.5123/S1679-49742020000400010>.
6	ESTRELA, Fernanda Matheus et al. Pandemia da Covid 19: refletindo as vulnerabilidades a luz do gênero, raça e classe. <i>Ciência & Saúde Coletiva</i> , v.25, n.9, p.3431-3436, 2020. Disponível em: <10.1590/1413-81232020259.14052020>.
7	MINAYO, Maria Cecília de Souza; FREIRE, Neyson Pinheiro. Pandemia exacerba desigualdades na Saúde. <i>Ciência & Saúde Coletiva</i> , v.25, n.9, p.3555-3556, 2020. Disponível em: <10.1590/1413-81232020259.13742020>.
8	NORONHA, Kenya Valeria M. de Souza et al. Pandemia por COVID-19 no Brasil: análise da demanda e da oferta de leitos hospitalares e equipamentos de ventilação assistida segundo diferentes cenários. <i>Cad. Saúde Pública</i> , v.36, n.6, e00115320, 2020. Disponível em: <10.1590/0102-311X00115320>.
9	NUNES, Bruno Pereira et al. Multimorbidade e população em risco para COVID-19 grave no Estudo Longitudinal da Saúde dos Idosos Brasileiros. <i>Cad. Saúde Pública</i> , v.36, n.12, e00129620, 2020. Disponível em: <10.1590/0102-311X00129620>.
10	OLIVEIRA, Wanderson Kleber et al. Como o Brasil pode deter a COVID-19. <i>Epidemiol. Serv. Saude</i> , v.29, n.2, e2020044, 2020. Disponível em: <10.5123/S1679-49742020000200023>.
11	PASSOS, Taciana Silveira; ALMEIDA-SANTOS, Marcos Antonio. Trabalho sexual em período de pandemia por COVID-19 no contexto ibero-americano: análise de anúncios em websites. <i>Ciência & Saúde Coletiva</i> , v.25, n.11, p.237-4248, 2020. Disponível em: <10.1590/1413-812320202511.26622020>.
12	RIBEIRO-SILVA, Rita de Cássia et al. Implicações da pandemia COVID-19 para a segurança alimentar e nutricional no Brasil. <i>Ciência & Saúde Coletiva</i> , v.25, n.9, p.3421-3430, 2020. Disponível em: <10.1590/1413-81232020259.22152020>.
13	SANTOS, Ronaldo Teodoro; GUIMARÃES, Juarez Rocha. Democracia sem sentimento de república: o SUS nos tempos da Covid-19. <i>Saúde Debate</i> , v.44, n.esp, p.73-87, 2020. Disponível em: <10.1590/0103-11042020E404>.
14	SILVA, Andrey Ferreira et al. Elementos precipitadores/intensificadores da violência conjugal em tempo da Covid-19. <i>Ciência & Saúde Coletiva</i> , v.25, n.9, p.3475-3480, 2020. Disponível em: <10.1590/1413-81232020259.16132020>.
15	SOUZA, Wayner Vieira de et al. Cem dias de COVID-19 em Pernambuco, Brasil: a epidemiologia em contexto histórico. <i>Cad. Saúde Pública</i> , v.36, n.11, e00228220, 2020. Disponível em: <10.1590/0102-311X00228220>.

Fonte: Elaboração das autoras

46

3.3 TRATAMENTO E ANÁLISE DOS DADOS COLETADOS

Os artigos foram estudados com o auxílio da mineração de textos, utilizando o MAXQDA, um *software* para análise de dados qualitativos, quantitativos e métodos mistos de

pesquisa (MAXQDA, 2022). Antes do uso do *software*, os textos foram tratados considerando-se os seguintes critérios (GUEDES, 2010):

- Exclusão de autores, afiliações e títulos dos autores, resumos, palavras-chave, sumários, referências, cabeçalhos, notas de rodapé e agradecimentos;
- Consideração das citações como partes dos textos e formas flexionadas das palavras como palavras distintas.

Ao escrever sobre a indexação por extração automática, Lancaster (2004) afirmou que é possível escrever programas simples para contar as palavras num texto, desde que este tenha sido cotejado com uma lista de palavras proibidas, a fim de eliminar palavras não-significativas (artigos, preposições, conjunções e assemelhados) e, em seguida, ordenar essas palavras segundo a frequência de sua ocorrência. Essa lista de palavras proibidas, ou *stop words*, também foi aplicada na mineração dos textos.

Finalizando o processo, os termos compilados na extração automática via mineração de textos foram comparados com os termos localizados nos indicadores de representação (título, palavras-chave, descritores DeCS), verificando como o conteúdo dos artigos se reflete nesses indicadores.

47

3.4 MAPEAMENTO DA ATRIBUIÇÃO DE TERMOS

A atribuição dos termos nos artigos estudados foi mapeada, observando o papel da indexação automática e do grau de coerência semântica nesse processo. A coerência na representação temática dos artigos foi investigada a partir dos 3 resultados coletados nos textos: nas palavras-chave, nos descritores DeCS e na mineração de textos.

Dessa forma buscou-se verificar se o uso da indexação automática pode colaborar com a representação desses textos apresentados na SciELO Saúde Pública.

4 RESULTADOS

Dos seis periódicos selecionados para o estudo, somente quatro apresentaram textos de acordo com os critérios já citados: Ciência & Saúde Coletiva, com seis artigos; Cadernos de Saúde Pública, com quatro artigos; Saúde em Debate, com três artigos e Epidemiologia e Serviços de Saúde, com dois artigos (quadro 2).

Verificou-se que a maioria das palavras-chave dos artigos faz parte do vocabulário controlado da área (DeCS): em 13 dos 15 textos analisados, todas as palavras-chave utilizadas

são descritores autorizados, significando que os autores seguiram a recomendação dos periódicos.

Sobre a coerência entre o conteúdo dos artigos e sua representação temática, isto é, se o conteúdo dos textos foi contemplado nas palavras-chave, títulos e nos termos de alta frequência (via mineração de texto), identificou-se alta coerência em sete artigos (textos 2, 3, 6, 7, 8, 11 e 12), média coerência em cinco artigos (textos 1, 4, 9, 13 e 14) e baixa coerência em três artigos (textos 5, 10 e 15). Em alguns casos, termos de alta frequência não foram contemplados nos títulos e nas palavras-chave, enquanto termos de baixa frequência foram contemplados nas palavras-chave e títulos.

Chamou atenção a palavra-chave Infecções por Coronavirus, utilizada em cinco artigos, e que mesmo sendo um descritor autorizado pelo DeCS não teve representatividade no conteúdo dos textos, ao contrário do termo COVID-19 que, mesmo com alta frequência nesses textos não foi contemplado como palavra-chave. Essa situação pode estar refletindo o momento da publicação dos artigos - início de 2020 - praticamente no mesmo período de análise, revisão e publicação do descritor COVID-19 no vocabulário DeCS (COVID-19, 2022), ou seja, não houve tempo suficiente para incorporação pelos autores para o novo descritor autorizado. O tempo da nomeação da doença foi mais rápido.

As palavras-chave mais utilizadas - que também são descritores autorizados - foram COVID-19 (9 artigos), Pandemia(s) (9 artigos) e Infecções por Coronavírus (5 artigos). O uso das outras palavras-chave identificadas ficou entre 1 e 2 artigos.

Os termos de maior relação com a pandemia da COVID-19 foram Brasil, Epidemiologia, Política(s) Pública(s), Segurança Alimentar e Nutricional, Sistema Único de Saúde e Vigilância em Saúde. Com o uso da indexação automática, observou-se que a representação dos artigos poderia se completar com os diversos termos elaborados nesse processo, incluindo os que também são descritores autorizados pelo vocabulário recomendado (quadro 3).

Quadro 3 - Mapeamento da atribuição de termos

Palavras-chave	Indexação Automática	Descritores autorizados pelo DeCS
	Abastecimento Alimentar	Abastecimento de Alimentos
	Abuso de Álcool	Alcoolismo
	Acesso à Saúde	
	Acesso aos Cuidados de Saúde	Acesso aos Serviços de Saúde
Acesso aos Serviços de Saúde	Acesso aos Serviços de Saúde	Acesso aos Serviços de Saúde
	Agricultura Familiar	Agricultura
	Alimentação Adequada	Direito Humano à Alimentação Adequada
	Alimentos	Alimentos

	América Latina	América Latina
	Análise Espacial	Análise Espacial
	Análise de Conteúdo	
Anúncios		Anúncio
	Anúncios de Prostituição Anúncios em Websites	Anúncio Profissionais do Sexo Redes de Comunicação de Computadores Trabalho Sexual
	Autoridade Sanitária	Autoridades de Saúde
	Auxiliares e Técnicos de Saúde	Pessoal Técnico de Saúde
Brasil	Brasil	Brasil
Classe Social		Classe Social
Condições Sociais		Condições Sociais
	Consumo de Álcool	Consumo de Bebidas Alcoólicas
COVID-19	COVID-19 Combate à COVID-19	COVID-19
	Coronavírus	Infecções por Coronavírus
Democracia	Democracia Brasileira	Democracia
	Desigualdade de Renda	Fatores Socioeconômicos
Desigualdades		Disparidades nos Níveis de Saúde
	Desigualdades em Saúde	Disparidades nos Níveis de Saúde
	Determinantes Sociais	Determinantes Sociais da Saúde
	Direitos Humanos	Direitos Humanos
	Distanciamento Social	Distanciamento Físico
	Distribuição Espacial	Características de Residência Demografia
	Efeito do Álcool	
	Emergência em Saúde	Emergências
Epidemiologia		Epidemiologia
	Estudo Longitudinal da Saúde dos Idosos Brasileiros (ELSI-Brasil)	
Estudos de Séries Temporais		Estudos de Séries Temporais
	Falta de Recursos	Pobreza
Fatores Epidemiológicos		Fatores Epidemiológicos
Gênero		Identidade de Gênero
	Governo Federal	Governo Federal
Idoso	Idoso(s) Idosos Brasileiros	Idoso
	Indicadores de Saúde	Indicadores Básicos de Saúde
	Iniquidades Sociais	Iniquidades em Saúde
	Insegurança Alimentar	Insegurança Alimentar
	Institucionalização do SUS	Institucionalização Sistema Único de Saúde
	Internação por COVID-19	COVID-19 Hospitalização
	Isolamento Social	Isolamento Social
	Indústria do Sexo	
	Informações Sanitárias	
Infecções por Coronavírus		Infecções por Coronavírus
Infraestrutura do Sistema de Saúde	Infraestrutura de Saúde	Infraestrutura Sistemas de Saúde
	Leitos	Número de Leitos em Hospital
	Leitos Hospitalares	Número de Leitos em Hospital
	Marcadores Sociais de Gênero, Raça e Classe	Classe Social Grupos Raciais

		Identidade de Gênero
	Mercado de Trabalho	Mercado de Trabalho
Mercantilização		Mercantilização
Mídias Sociais		Mídias Sociais
	Ministério da Saúde	Sistemas Nacionais de Saúde
	Morbidades	Morbidade
	Mulheres	Mulheres
Multimorbidade		Multimorbidade
	Óbitos por COVID-19	COVID-19 Mortalidade
OMS	Organização Mundial da Saúde	Organização Mundial da Saúde
Ordem Mundial	Ordem Mundial	
	Países em Desenvolvimento	Países em Desenvolvimento
Pandemia(s)	Pandemia	Pandemias
	Pandemia da COVID-19	COVID-19 Pandemias
Política		Política
	Política Brasileira	Brasil Política
Política(s) Pública(s)		Política Pública
	População Brasileira	Brasil População
	População em Situação de Rua	Pessoas em Situação de Rua
	População Vulnerabilizada	Populações Vulneráveis
	Prevenção da COVID-19	Prevenção de Doenças
	Profissionais de Enfermagem	Profissionais de Enfermagem
	Profissionais de Saúde	Pessoal de Saúde
	Profissionais do Sexo	Profissionais do Sexo
Proteção Social em Saúde	Proteção Social	Proteção Social em Saúde
	Quarentena	Quarentena
Raça e Saúde		Grupos Raciais Saúde
	Redes de Apoio às Mulheres	Apoio Social Mulheres
	Renda	Renda
	Risco	Risco
	Risco de COVID-19 Grave	COVID-19 Risco
	SARS-CoV-2	SARS-CoV-2
Saúde	Saúde	Saúde
	Saúde dos Idosos	Saúde do Idoso
Saúde Pública	Saúde Pública	Saúde Pública
Saúde Suplementar		Saúde Suplementar
Segurança Alimentar e Nutricional	Segurança Alimentar	Segurança Alimentar e Nutricional
	Serviços de Saúde	Serviços de Saúde
	Sistema de Saúde	Sistemas de Saúde
Sistema Único de Saúde	Sistema Único de Saúde	Sistema Único de Saúde
	Situação de Saúde	Diagnóstico da Situação de Saúde
	Soberania Popular	
	Suporte Institucional	
	Suporte Social	
	Taxas de Incidência e Mortalidade	Incidência Mortalidade
	Trabalhadores do Sexo	
	Trabalho e Renda	Renda Trabalho

Trabalho Sexual	Trabalho Sexual	Trabalho Sexual
	Unidade de Saúde	Centros de Saúde
	Uso de Álcool	Transtornos Relacionados ao Uso de Álcool
Vigilância em Saúde Pública	Vigilância em Saúde	Vigilância em Saúde Pública
	Violência Conjugal	
	Violência Doméstica	Violência Doméstica
Violência por Parceiro Íntimo		Violência por Parceiro Íntimo
	Websites de Prostituição	Redes de Comunicação de Computadores Trabalho Sexual

Fonte: Elaboração das autoras

5 CONSIDERAÇÕES FINAIS

O estudo, de caráter exploratório, visou avaliar a aplicação da indexação automática em artigos publicados em português na área de saúde pública, em artigos publicados em 2020 em periódicos integrantes da Coleção Saúde Pública na SciELO Brasil.

Observou-se que ocorreu alta coerência na metade dos artigos estudados. Nos outros textos a coerência ficou entre baixa e média: em alguns casos os termos de alta frequência não foram utilizados, e em outros casos as palavras-chave foram utilizadas mesmo não estando na lista dos termos mais frequentes do artigo. Esse resultado pode refletir uma escolha consciente dessas palavras-chave pelos autores, levando em consideração a indicação do vocabulário controlado.

Mesmo com a obrigatoriedade de uso, outras palavras-chave (que também são descritores) poderiam representar de forma mais assertiva o conteúdo de alguns artigos. É razoável supor que há desconhecimento em como usar todos os recursos que o vocabulário oferece. Reconhece-se a importância do uso da terminologia em áreas específicas, mas essa questão não pode interferir na visibilidade da inovação que esteja sendo desenvolvida em pesquisas dessas áreas.

A partir deste exercício exploratório, pode-se considerar que a mineração de texto e a indexação automática têm potencial para contribuir com a coerência na representação de artigos científicos, pois foi possível verificar que diversos termos coerentes ao conteúdo dos textos foram identificados nesse processo, incluindo os que também são descritores autorizados pelo DeCS (quadro 3).

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). **NBR 12676**: Métodos para análise de documentos - Determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro: ABNT, 1992.

ARAÚJO JÚNIOR, Rogério H. **Precisão no processo de busca e recuperação da informação**. Brasília: Thesaurus, 2007.

ARAÚJO JÚNIOR, Rogério H; TARAPANOFF, Kira. Precisão no processo de busca e recuperação da informação: uso da mineração de textos. **Ciência da Informação**, v.35, n.3, p.236-247, 2006.

BOCCATO, Vera R. Casari; TORQUETTI, Melissa Camargo. Interoperabilidade entre linguagens de indexação como recurso de modelagem de repertório terminológico de coordenadorias de comunicação social em ambientes universitários: uma proposta metodológica. **Informação & Informação**, v. 17, n. 3, p. 76-101, 2012. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/10800/pdf>>. Acesso em: 8 ago 2013.

BURKE, Peter. A consolidação do conhecimento: antigas e novas instituições. In: **Uma história social do conhecimento**: de Gutemberg a Diderot: Rio de Janeiro: Jorge Zahar, 2003. v.1

CAMPELLO, Bernadete Santos; CENDÓN, Beatriz Valadares; KREMER, Jeannette Marguerite (orgs.). **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: UFMG, 2007.

CARVALHO, Marília Sá; LIMA, Luciana Dias de; COELI, Cláudia Medina. Fast-track COVID-19 em CSP. **Cad. Saúde Pública**, v. 36, n. 8, e00204820, 2020.

CHRISTOVÃO, Heloisa T. **Da comunicação informal à comunicação formal**: identificação da frente de pesquisa através de filtros de qualidade. Dissertação (Mestrado em Ciência da Informação) - CNPQ/IBICT/UFRJ, Rio de Janeiro, 1979. p. 6-21.

CRUZ, Maria C. Andrade; FUJITA, Mariângela S. Lopes. O uso de linguagem de indexação por bibliotecas universitárias brasileiras. **Informação & Informação**, v.26, n.1, p.574-600, 2021. Disponível em: <10.5433/1981-8920.2021v26n1p574 >.

COVID-19. Disponível em:

<https://decs.bvsalud.org/ths/resource/?id=59585&filter=ths_termall&q=covid#Details>. Acesso em 16 ago. 2022

DeCS: Descritores em Ciências da Saúde. Disponível em: <<https://decs.bvsalud.org/>>. Acesso em 20 jul. 2022.

FERRARI, Carlos K.B. Um guia para publicar artigos em ciências da saúde. **Pleiade**, v.12, n.26, p.05-13, 2018.

FERREIRA, Márcio H.W.; CORREA, Renato F. Mineração de textos científicos: análise de artigos de periódicos científicos brasileiros da área de Ciência da Informação. **Em Questão**, v.27, n.1, p.237-262, 2020. Disponível em: <<https://doi.org/10.19132/1808-5245271.237-262>>.

FÓRIS, Ágota. Network theory and terminology. **Knowledge Organization**, v.40, n.6, p.422-429, 2013.

FUJITA, Mariângela S. Lopes. A identificação de conceitos no processo de análise de assunto para indexação. **Rev Digital de Biblioteconomia e Ciência da Informação**, v.1, n.1, p.60-90, 2003. Disponível em: <<https://doi.org/10.20396/rdbci.v1i1.2089>>.

GIL LEIVA, Isidoro. **La automatización de la indización de documentos**. Gijón: Trea, 1999. 221 p.

GUANAES, Paulo C. Vieira; GUIMARÃES, Maria Cristina S. Modelos de gestão de revistas científicas: uma discussão necessária. **Perspectivas em Ciência da Informação**, v.17, n.1, p.56-73, 2012. Disponível em: <<https://doi.org/10.1590/S1413-99362012000100004>>.

GUEDES, Vânia Lisbôa S. **Nominalizações deverbais em artigos científicos: uma contribuição para a análise e a indexação temática da informação**. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

KRIEGER, Maria da Graça; FINATTO, Maria José B. **Introdução à terminologia: teoria e prática**. São Paulo: Contexto, 2004.

KURAMOTO, Hélio. Informação científica: proposta de um novo modelo para o Brasil. **Ciência da Informação**, v.35, p.91-102, 2006.

LANCASTER, Brasília Frederick Wilfrid. **Indexação e resumos: teoria e prática** 2. ed. rev. atual.: Briquet de Lemos/Livros, 2004.

LATOUR, Bruno; WOOLGAR, Steve. A credibilidade científica. In: **A vida de laboratório**. Rio de Janeiro: Relume Dumará, 1997.

MACGARRY, Kevin. **O contexto dinâmico da informação: uma análise introdutória**. Brasília: Briquet de Lemos, 1999.

MARCONDES, Carlos Henrique; COSTA, Leonardo Cruz da; MARTINS, Sergio de Castro. Descoberta de conhecimento em artigos digitais em ciências biomédicas. **Informação & Informação**, v.21, n.2, p.170-216, 2016. Disponível em: <<http://dx.doi.org/10.5433/1981-8920.2016v21n2p170>>.

MAXQDA. Disponível em: <https://www.maxqda.com/brasil>. Acesso em: 14 jul. 2022.

MEADOWS, Arthur J. Canais da comunicação científica. In: **A comunicação científica**. Brasília,DF: Briquet de Lemos, 1999.

MIGUÉS, Ana; NEVES, Bruno; SILVA, Ana Luísa; TRINDADE, Álvaro; BERNARDES, José Augusto. A importância das palavras-chave dos artigos científicos da área das Ciências Farmacêuticas, depositados no Estudo Geral: estudo comparativo com os termos atribuídos na MEDLINE. **InCID: Revista de Ciência da Informação e Documentação**, v.4, n.2, p.112-125, 2013.

MIRANDA, Ana C. Carvalho; CARVALHO, Edirsana M. Ribeiro; COSTA, Maria Ilza. O impacto dos periódicos na comunicação científica. **BIBLOS - Revista do Instituto de Ciências Humanas e da Informação**, n.1, v.32, p.1-22, 2018.

MUELLER, Suzana P.M. A ciência, o sistema de comunicação científica e a literatura científica. In: CAMPELLO, Bernardete S.; CENDÓN, Beatriz V.; KREMER, Jeannette (orgs.). **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: UFMG, 2007.

NARUKAWA, Cristina Miyuki; GIL LEIVA, Isidoro; FUJITA, Mariângela S. Lopes. Indexação automatizada de artigos de periódicos científicos: análise da aplicação do software SISA com uso da terminologia DeCS na área de Odontologia. **Informação & Sociedade**:

ASKLEPION: Informação em Saúde, Rio de Janeiro, v. 2, 2022, Edição Especial, p. 33-55.

Estudos, v.19, n.2, p.99-118, 2009. Disponível em:

<<https://periodicos.ufpb.br/index.php/ies/article/view/2925>>.

OLIVEIRA, Jacqueline P.; ALMEIDA, Maurício B.; SOUZA, Renato R. Fontes de informação especializada em ciências da saúde: análise de características e proposta de critérios para avaliação. In: **Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), XVI, Anais...** 23 a 30 de outubro de 2015, João Pessoa, PB. Disponível em: <<http://www.ufpb.br/evento/index.php/enancib2015/enancib2015/paper/viewFile/3104/1273>>

ORGANIZAÇÃO MUNDIAL DA SAÚDE (OMS). **Painel da OMS sobre o coronavírus (COVID-19)**. Disponível em: <<https://covid19.who.int/>>. Acesso em 28 jul. 2022.

ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE (OPAS). **Folha informativa COVID-19: escritório da OPAS e da OMS no Brasil**. Disponível em: <<https://www.paho.org/pt/covid19>>. Acesso em 28 jul. 2022.

SALES, Rodrigo de. Suportes teóricos para pensar linguagens documentárias. **Rev. Digital de Biblioteconomia e Ciência da Informação**, v.5, n.1, p.96-114, 2007.

SANTOS, Fatima Cristina Lopes dos. **Coerência na representação temática de artigos científicos indexados no repositório de saúde pública da Fundação Oswaldo Cruz**. 2017. 258 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal do Rio de Janeiro/Escola de Comunicação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, 2017.

SANTOS, Fatima C.L.; MOLLICA, Maria Cecília M.; GUEDES, Vânia Lisboa S. Coerência na representação temática de artigos científicos na área de saúde pública. **Perspectivas em Ciência da Informação**, v.24, n.3, p.194-213, 2019. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/3727>>.

SANTOS, Fatima C.L.; SILVA, Cícera Henrique da; LINS, Rosane Abdala. Estudo sobre indicação de vocabulários nas orientações para autores de periódicos científicos na área da saúde. In: Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), XX, **Anais...**, 21 a 25 de outubro de 2019, Florianópolis, SC. Disponível em: <<https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/1147/807>>.

SciELOSaúde Pública. Disponível em:

<<https://www.scielo20.org/redescielo/wp-content/uploads/sites/2/2018/09/Informe-SciELO-Sa%C3%BAde-P%C3%BAblica.pdf>>. Acesso em: 14 abr. 2022.

SOUZA-SANTOS, Reinaldo. Busca e uso de fontes bibliográficas. In: HORTALE, Virginia Alonso; MOREIRA, Carlos O. Fiúza; BODSTEIN, Regina C. Andrade; RAMOS, Célia Leitão (Orgs.). **Pesquisa em saúde coletiva: fronteiras, objetos e métodos**. Rio de Janeiro: Fiocruz, 2010.

STREHL, Leticia. Avaliação da consistência da indexação realizada em uma biblioteca universitária da artes. **Ciência da Informação**, v.27, n.3, p.329-335, 1998.

STUMPF, Ida R. Chitto. Avaliação de originais nas revistas científicas: uma trajetória em busca do acerto. FERREIRA, Sueli M.S. Pinto; TARGINO, Maria das Graças (orgs.). **Preparação de revistas científicas: teoria e prática**. São Paulo: Reichmann & Autores, 2005.

TARGINO, Maria das Graças. Comunicação científica: uma revisão de seus elementos básicos. **Informação e Sociedade: Estudos**, v.10, n.2, p. 67-85, 2000.

VANZ, Samile A. de Souza; SILVA FILHO, Rubens da Costa. O protagonismo das revistas na comunicação científica: histórico e evolução. In: CARNEIRO, Felipe F. Barros;

ASKLEPION: Informação em Saúde, Rio de Janeiro, v. 2, 2022, Edição Especial, p. 33-55.

FERREIRA NETO, Amarílio; SANTOS, Wagner. **A comunicação científica em periódicos**. Curitiba: Appris, 2019.

VELHO, Léa M.L.S. Como medir a ciência? **Rev. Bras. Tecnologia**, v.16, n.1, p.35-41, 1985.

VETTORE, Mário Vianna; COIMBRA JÚNIOR, Carlos E. A publicação científica. In: HORTALE, Virgínia Alonso et al. **Pesquisa em saúde coletiva: fronteiras, objetos e métodos**. Rio de Janeiro: Fiocruz, 2010.