



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO

THEORETICAL FRAMEWORK CONSTRUCTION PROCESS THROUGH DATA EXTRACTION FROM THE SCOPUS ENVIRONMENT AND PRODUCTION INDICATORS

PROCESO DE CONSTRUCCIÓN DEL MARCO TEÓRICO MEDIANTE EXTRACCIÓN DE DATOS DEL ENTORNO SCOPUS E INDICADORES DE PRODUCCIÓN

Edson Melo de Souza¹, Gustavo Pessarelo Smulkowski Souza²

e412538

<https://doi.org/10.47820/recima21.v4i1.2538>

PUBLICADO: 01/2023

RESUMO

Contexto: A pesquisa e elaboração de referencial teórico de relevância em produções científicas é um processo árduo e demorado para o autor iniciante ou com pouca experiência. Objetivo: Este artigo apresenta um processo de pesquisa bibliográfica para a construção inicial do referencial teórico com o uso da Tecnologia da Informação (TI) para acesso, recuperação e gerenciamento da bibliografia. Materiais e métodos: Os experimentos executados empregaram metodologias qualitativas e indutivas, visando a realização de pesquisas *online* na base de periódicos eletrônicos Scopus, por meio de buscas por palavras-chave e parâmetros de filtragem. Para selecionar as melhores referências, foram utilizados o índice "h" e a classificação dos periódicos pelo Qualis-CAPES. Além disso, o programa gerenciador de referências Mendeley foi a ferramenta de TI usada para o gerenciamento da bibliografia. Resultados: As estratégias de pesquisa propostas para a seleção de palavras-chave, em conjunto com a análise da qualidade das publicações, forneceram condições para reduzir de 18.087 para 22 artigos, totalizando 2.263 citações relevantes para a construção do referencial teórico. O programa Mendeley permitiu a organização e exploração de detalhes dos artigos com facilidade e praticidade. Conclusão: Por fim, infere-se que as estratégias de pesquisa utilizadas em conjunto com a metodologia proposta e com a ferramenta de TI selecionada mostraram-se eficientes como suporte ao processo de busca, recuperação e catalogação de artigos científicos para a elaboração de referencial teórico relevante.

PALAVRAS-CHAVE: Referencial Teórico. Artigo Científico. Processo de Pesquisa.

ABSTRACT

Context: The research and elaboration of a theoretical framework of relevance in scientific productions is an arduous and time-consuming process for a novice author or one with little experience. Objective: This manuscript presents a bibliographical research process for the initial construction of the theoretical framework with the use of Information Technology (IT) for access, recovery, and management of the bibliography. Materials and methods: The experiments carried out used qualitative and inductive methodologies, aiming to carry out online searches in the Scopus electronic journal base, through searches by keywords and filtering parameters. To select the best references, the h-index, and journals ranked by Qualis-CAPES were used. In addition, the Mendeley reference manager program was the IT tool used for managing the bibliography. Results: The research strategies proposed for the selection of keywords, together with the analysis of the quality of publications, provided conditions to reduce from 18,087 to 22 articles, with 2,263 relevant citations to construction of the theoretical framework. The Mendeley program allowed the organization and exploration of the details of the articles with ease and practicality. Conclusion: Finally, it is inferred that the research strategies used in conjunction with the proposed methodology and the selected IT tool proved to be

¹ Doutor em Informática e Gestão do Conhecimento. Mestre em Engenharia de Produção com ênfase em Tecnologias de Gestão de Processos de Usinagem. Especialização em Gestão Estratégica Empresarial. Extensão em Práticas de Docência para o Ensino Superior e Bacharel em Ciências da Computação, ambos pela Universidade Nove de Julho (UNINOVE). Professor na Universidade Nove de Julho (UNINOVE). Membro do grupo de pesquisa GMOCOMP - Grupo Modelagem Computacional de Processos.

² Bacharelado em Ciências e Tecnologia na Universidade Federal do ABC – UFABC. Aluno de Iniciação Científica na área de Neurociência. Bolsista CNPq.



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE
DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

efficient in supporting the process of searching, retrieving, and cataloging scientific articles for the elaboration of a reference relevant theory.

KEYWORDS: *Theoretical Bases. Scientific Manuscript. Research Process.*

RESUMEN

Contexto: La investigación y elaboración de un marco teórico de relevancia en las producciones científicas es un proceso arduo y lento para un autor novel o con poca experiencia. Objetivo: Este manuscrito presenta un proceso de investigación bibliográfica para la construcción inicial del marco teórico con el uso de Tecnologías de la Información (TI) para el acceso, recuperación y gestión de la bibliografía. Materiales y métodos: Los experimentos realizados utilizaron metodologías cualitativas e inductivas, con el objetivo de realizar búsquedas en línea en la base de revistas electrónicas Scopus, a través de búsquedas por palabras clave y parámetros de filtrado. Para seleccionar las mejores referencias se utilizó el índice h y revistas clasificadas por Qualis-CAPES. Además, el programa gestor de referencias de Mendeley fue la herramienta informática utilizada para la gestión de la bibliografía. Resultados: Las estrategias de investigación propuestas para la selección de palabras clave, junto con el análisis de la calidad de las publicaciones, proporcionaron condiciones para reducir de 18.087 a 22 artículos, con 2.263 citas relevantes para la construcción del marco teórico. El programa Mendeley permitió la organización y exploración de los detalles de los artículos con facilidad y practicidad. Conclusión: Finalmente, se infiere que las estrategias de investigación utilizadas en conjunto con la metodología propuesta y la herramienta informática seleccionada demostraron ser eficientes para apoyar el proceso de búsqueda, recuperación y catalogación de artículos científicos para la elaboración de una teoría relevante de referencia.

PALABRAS CLAVE: *Bases Teóricas. Manuscrito científico. Proceso de investigación.*

INTRODUÇÃO

O desenvolvimento de um artigo científico requer a elaboração de um referencial teórico relevante e consistente (AZEVEDO, 2016). Essa etapa demanda grande esforço e tempo do pesquisador, principalmente para aqueles que não possuem experiência, sólidos conhecimentos no campo da ciência ou amplo domínio sobre o assunto em que está desenvolvendo sua pesquisa (EVANS; CVITANOVIC, 2018).

A publicação dos achados científicos no formato de artigo surgiu no ano de 1665 e teve como foco a divulgação de trabalhos sobre a Filosofia Natural, proporcionando uma melhora na divulgação das pesquisas, que até então ocorria pela troca de cartas entre os pesquisadores. Essa prática tornava o processo lento e custoso, problema o qual foi solucionado com a criação dos periódicos científicos, possibilitando padronizar e acelerar o debate de ideias sobre os assuntos mais relevantes no mundo científico europeu da época (ARAÚJO *et al.*, 2017).

Com o advento da internet pública em meados dos anos 1980, a difusão do conhecimento científico alcançou patamares anteriormente inimagináveis na visão de alguns autores (LORDÉLO, 2011). Essa difusão foi impulsionada pela Tecnologia da Informação (TI), que proporcionou o surgimento dos periódicos científicos eletrônicos na década de 2000. Desde então, os resultados de pesquisas científicas têm sido divulgados quase que de forma imediata (SANDES-GUIMARÃES; DINIZ, 2014).



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Ainda nos anos 2000, houve o surgimento de plataformas indexadoras de periódicos como a *Web of Science*, *Scopus* e o *IEEE Xplore Digital Library*. Essas plataformas oferecem ferramentas digitais que permitem realizar buscas personalizadas, reunindo quantidades expressivas de artigos em um mesmo lugar (SINGH P; SINGH K, 2017), característica que mitiga parte do esforço para construção do referencial teórico.

Apesar da existência das ferramentas digitais oferecidas pelas plataformas e periódicos, ainda há grande esforço para compilar a bibliografia obtida. Desta forma, pesquisadores iniciantes enfrentam dificuldades por desconhecerem métodos e ferramentas tecnológicas, dispendendo muito tempo para realizarem a construção inicial do referencial teórico para seus projetos e artigos, mesmo com disciplinas ofertadas sobre o assunto dentro dos programas de pós-graduação *Lato Sensu* e *Stricto Sensu* (TEIXEIRA *et al.*, 2019; SILVA *et al.*, 2019).

Em vista deste cenário, surge a seguinte questão de pesquisa: A utilização de estratégias para extração de dados de artigos, nas plataformas indexadoras e de periódicos com o suporte de ferramentas de TI, pode apoiar o processo inicial de construção de um referencial teórico de qualidade e minimizar a falta de experiência de pesquisadores iniciantes?

Com base no exposto, o objetivo principal deste trabalho é apresentar ao pesquisador iniciante os principais termos e conceitos inerentes à pesquisa científica. Um segundo objetivo é mostrar como realizar uma pesquisa bibliográfica em bases de artigos de periódicos e plataformas indexadoras com o uso de recursos de TI. Por fim, apresenta uma ferramenta de TI que apoia a organização e o manuseio dos artigos obtidos para a construção de um referencial teórico para suas produções acadêmicas (NILASHI, 2019).

BASE TEÓRICA

O ARTIGO CIENTÍFICO

De acordo com a ABNT - Associação Brasileira de Normas Técnicas (NBR 6022, 2018, p. 2) o artigo científico é definido como “[...] parte de uma publicação, com autoria declarada, de natureza técnica e/ou científica.”. De outra forma, um artigo científico pode ser descrito como um documento que tem como objetivo relatar, de forma lógica e concisa, os resultados obtidos por um ou mais pesquisadores sobre uma pesquisa ou parte dela (HUGUET; GAYA; PALOU, 2018). É a materialização dos resultados de uma pesquisa em um documento que possa ser distribuído de forma física ou digital.

A estrutura de um artigo científico, ainda segundo a ABNT (NBR 6022, 2018, p. 4) é constituída de “[...] elementos pré-textuais, textuais e pós-textuais [...]”. Cada um dos elementos possui itens obrigatórios e opcionais, dos quais os obrigatórios são adotados em publicações mundiais. Esses elementos são: título, autor(es), resumo, palavras-chave, data de submissão/aprovação, introdução, desenvolvimento, resultados, considerações finais e referências.



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Entre os itens obrigatórios está a nomeação dos autores que participaram da elaboração do artigo, conhecido como coautoria ou *byline*. A literatura não é muito clara quanto à ordem em que estes devem ser inseridos, uma vez que isso depende de cada área do conhecimento e do periódico em que será publicado o artigo (YANG; WOLFRAM; WANG, 2017). A ordem de aparição dos nomes dos pesquisadores ou posição autoral indica o grau de relevância e de contribuição, o que deve ser considerado em relação à experiência acadêmica, técnica e de senioridade do pesquisador. No Brasil, o primeiro autor é considerado como principal.

Segundo Ferreira e Serra (2015), há três grandes orientações para identificar a posição autoral: "(1) o iniciador do projeto de pesquisa, (2) a contribuição relativa de cada um dos autores, (3) a importância da publicação para a progressão do pesquisador.". Ainda, segundo os mesmos autores, há outras formas que utilizam a experiência acadêmica, percepção de contribuição individual, entre outros, principalmente quando há alunos de pós-graduação envolvidos.

Um artigo científico tem seu registro realizado mediante a publicação em uma revista científica especializada, denominada de periódico. Entretanto, antes que seja publicado, há um processo de avaliação por pesquisadores e/ou especialistas da área que tenham amplo conhecimento sobre o assunto proposto. Esse processo é conhecido como avaliação por pares ou *peer review*, definido por Patrus, Dantas e Shigaki (2016) como "Um dos procedimentos quase universalmente aceitos pela comunidade científica na indicação da qualidade de trabalhos científicos [...]". A revisão por pares atesta não só a qualidade de um artigo, mas também a relevância do assunto para a comunidade científica. Ademais, o periódico que fará a publicação deve estar alinhado à área de pesquisa e possuir qualificação de acordo com a comunidade científica (MATTSSON; SUNDBERG; LAGET, 2011).

A reunião de diversos artigos científicos fornece condições para um pesquisador buscar respostas às suas perguntas e/ou fundamentar uma proposição, no devido respeito ao que já foi publicado em sua área de investigação. Portanto, constitui a base do referencial teórico de uma pesquisa.

PERIÓDICOS CIENTÍFICOS

A concepção dos periódicos científicos data de aproximadamente 350 anos com o surgimento do francês *Journal des Sçavans* e o inglês *Philosophical Transactions*, ambos em 1665, com foco na divulgação de trabalhos sobre a Filosofia Natural. Até então o trabalho de divulgação das pesquisas ocorria por meio do envio de cartas entre os pesquisadores, o que tornava o processo altamente custoso e moroso. A chegada desses periódicos trouxe enorme inovação, possibilitando padronizar e acelerar o debate de ideias acerca dos assuntos mais relevantes no mundo científico europeu da época (MUELLER, 2000).

A partir dessas referências, no século XVIII houve uma evolução para a publicação mais ampla, em formato de boletins para membros da *Royal Society*, instituição fundada em 1660 com o



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

objetivo de disseminar o conhecimento científico (SPRAT, 1734). Nesse período houve um crescimento expressivo de periódicos, principalmente pela necessidade do registro de propriedade intelectual e expansão para as sociedades ocidentais. Nessa mesma época também foi implantado o sistema de revisão por pares (PATRUS; DANTAS; SHIGAKI, 2016).

Um periódico científico tem como funções a comunicação formal de resultados de pesquisas, manter e perpetuar o conhecimento registrado ao longo do tempo e fomentar a reflexão e o debate de ideias entre pesquisadores (MUELLER, 2000). Neste cenário de vasto conhecimento registrado, os periódicos são divididos em áreas de conhecimento e, dentro dessas áreas, ainda há subdivisões que visam apresentar especificidades sobre o conhecimento abordado.

Segundo Oliveira (2008), “A utilização do meio eletrônico na comunicação entre pesquisadores já pode ser considerada um fato comum, já que desde seu surgimento, a Internet foi utilizada por eles para a troca de informações e para a comunicação.” Para Lordêlo (2011), “[...] a internet surgiu como uma ferramenta para ajudar a disseminação da ciência, propagando informações e auxiliando no processo instrucional e educacional das sociedades.”. Desta forma, a internet tornou-se um ambiente fértil para o crescimento e propagação do conhecimento científico. Em termos da divulgação científica, a TI proporcionou grande impulso para o surgimento de um elevado número de periódicos eletrônicos na década de 2000, os quais passaram a disponibilizar quase que imediatamente os resultados de pesquisas científicas (SANDES-GUIMARÃES; DINIZ, 2014).

Devido ao volume de informações geradas, surgiram obstáculos para a localização de publicações, uma vez que é necessário conhecer especificamente a área de atuação do periódico para então realizar a busca de artigos. Há também dificuldades em relação a transparência do periódico, quantidade e qualidade das publicações, grau de importância para a comunidade científica, entre outros aspectos (ARAÚJO *et al.*, 2017). Assim, os periódicos possuem uma classificação denominada estrato indicativo de qualidade ou somente estrato, a qual elenca a produção científica nele publicada.

No Brasil, a CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior é a responsável por realizar esta classificação por intermédio do Qualis-Periódicos ou Qualis-Capes. Esta classificação é realizada periodicamente e distribuída de forma quadrienal através da Plataforma Sucupira (*Online*), vigente para (2013-2016) no momento da escrita deste trabalho, onde os periódicos são classificados em um sistema de letras e pesos relativos que indicam a qualidade, sendo, em ordem decrescente de pesos, as classificações A1 - o mais elevado, A2, B1, B2, B3, B4, B5 e C - peso zero (SANTOS; FRANÇA, 2022). Internacionalmente, as publicações recebem diferentes métodos de classificação, gerando alguns impasses dentro da comunidade científica. Há diferentes classificadores como o *SCImago Journal & Country Rank* e o *Google Scholar*, ambos com critérios diferentes de ranqueamento (FARIAS, 2017; SCIMAGO, 2021). Assim, não apenas



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE
DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

periódicos nacionais são classificados, mas todos os periódicos coletados mundialmente, seguindo critérios quantitativos adotados pela CAPES (FRIGERI; MONTEIRO, 2014).

O processo de seleção de um periódico para submissão ou recuperação de um artigo está diretamente relacionado ao *ranking* atribuído pelos atores classificadores. Portanto, há necessidade da compreensão sobre os indicadores de uma produção científica, a fim de obter publicações que sejam relevantes e com impacto representativo dentro da comunidade e reconhecidos pelas universidades (PIMENTA *et al.*, 2017).

INDICADORES SOBRE ARTIGOS E PERIÓDICOS

O conhecimento qualitativo é baseado na subjetividade de uma narrativa, enquanto o quantitativo está relacionado a números, os quais podem ser analisados com o auxílio da estatística (CRESWELL; POTH, 2017). Na visão de Lundeborg (2006), a geração de conhecimento é materializada por intermédio da produção acadêmica. Assim, Alan Pritchard propôs no final da década de 1960 o termo “bibliometria”, que pode ser definido, segundo Vouga e Amatucci (2015), como “[...] a aplicação de métodos estatísticos e matemáticos na análise de obras literárias.”. Em outras palavras, a bibliometria visa estudar e identificar características comuns presentes nos artigos científicos, permitindo que sejam atribuídos indicadores mediante a quantificação do conhecimento.

De acordo com o contexto, a seleção de um artigo deve ser realizada com base em indicadores e/ou métricas que possam medir a qualidade da produção científica publicada a partir da avaliação de seus conteúdos (PIMENTA *et al.*, 2017). Entre os indicadores utilizados, estão o fator de impacto (FI) e o índice “h” ou (*h-index*) proposto por Jorge E. Hirsch em 2005, que fornecem condições de medir a qualidade de um artigo ou periódico (DROESHER; SILVA, 2014).

De acordo com Bornmann e Daniel (2005), o índice “h” é calculado com base no número de artigos publicados de um autor e a relação do número de citações para os artigos publicados. Como exemplo, se um autor publica 8 artigos que possuem 8 citações, então o seu *h-index* é 8. Por outro lado, se houver publicações, mas não houver ao menos uma (1) citação para os artigos, o *h-index* será zero (0), não significando que o autor não produziu. Esse indicador é utilizado não apenas para avaliar autores, mas também periódicos e universidades, devido sua simplicidade (RUIZ; GRECO; BRAILE, 2009). Uma forma de obter tais índices é por meio do SCiVal da Elsevier, VantagePoint, InCites e o Google Scholar Metrics, ferramentas computacionais analíticas que fornecem dados sobre produções (LEYDESDORFF; WOUTERS; BORNMANN, 2016).

Já o FI de um periódico é um indicador numérico calculado sobre as citações das publicações em um determinado período (GIUFFRIDA *et al.*, 2019). Apesar de ser um balizador de qualidade das publicações, é importante destacar que seu valor é válido apenas dentro de uma mesma área ou subárea do conhecimento, de modo que comparações entre áreas podem causar distorções devido ao volume de publicações entre elas e taxas de obsolescência diferenciadas (STREHL, 2005). De forma geral, quanto maior for o FI do periódico dentro de uma mesma área, melhor será a sua



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

classificação. Segundo Mugnaini (2012), o FI de um artigo “[...] se tornou num tipo de símbolo de status para periódicos [...]”. Ainda, segundo o mesmo autor, apesar de ser amplamente utilizado, existem divergências sobre sua aplicação, uma vez que a quantidade de documentos é muito grande, gerando inconsistências de índices sobre citações e bibliografias.

Para contornar esse problema, os indicadores do *Journal Citation Reports* (JCR) do *Institute for Scientific Information* (ISI) passaram a ser utilizados e, segundo Droeshner e Silva (2014), “[...] apresenta anualmente três indicadores por título de periódico: o índice de citação imediata (*immediacy index*), a meia-vida das citações (*cited Half-Life*) e o FI (*impact fator*)”.

Portanto, a utilização dos indicadores citados permite realizar uma seleção mais criteriosa e confiável da produção científica publicada.

FERRAMENTAS DE TI PARA GESTÃO DA BIBLIOGRAFIA

A evolução da TI têm proporcionado cada vez mais as possibilidades de acelerar o processo produtivo e criativo em diversas áreas. Com a criação de bases de periódicos *online*, a pesquisa e recuperação digital de artigos vêm se consolidando a cada ano. Assim, surgiram programas de computador para auxiliar no gerenciamento da bibliografia e referências (YAMAKAWA *et al.*, 2014).

Na internet é possível localizar esses programas de maneira gratuita ou comercializada. Essas ferramentas podem realizar a reunião e o gerenciamento da bibliografia em um computador *desktop* ou *SmartPhone* tanto de forma local como online. Esses programas permitem realizar anotações, destacar trechos em textos, importar artigos e referências, entre outras funcionalidades (KRATOCHVÍL, 2017).

Entre as ferramentas para gestão de bibliografia se destacam os programas Zotero (<https://www.zotero.org/>), EndNote (<https://access.clarivate.com/login?app=endnote>), RefWorks (<http://refworks.proquest.com/>) e Mendeley (<http://www.mendeley.com/>) (IVEY; CRUM, 2018), este último mantido pela Elsevier-Scopus.

De forma geral, os programas citados oferecem condições de armazenar artigos, criar anotações, gerar citações e lista de referências diretamente em editores de texto como o *MS Word* e *Open Office*, através da instalação de *plugins*. De forma individual, podem oferecer a recuperação automática de arquivos através do DOI (*Digital Object Identifier System* ou Identificador Digital de Documentos) ou de um endereço na internet, facilitando a organização e acesso aos documentos originais.

Entre os programas citados, o EndNote é distribuído de forma comercial e o RefWorks através de vínculo institucional com uma IES. Já o Zotero e o Mendeley são gratuitos, diferenciando na quantidade de espaço fornecido para armazenamento *online*, não apresentando limites para isso de forma local.

Além das opções apresentadas, ainda pode ser utilizada a plataforma digital da Universidade Federal de São Carlos (UFSC) que disponibiliza o Mecanismo Online para Referências (MORE).



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Essa ferramenta fornece funcionalidades para realização de citações diretas, indiretas e textuais para o referencial teórico (MORE, 2019), permitindo ainda o armazenamento das referências geradas. É importante destacar que esse recurso se limita a geração de citações e referências, não permitindo realizar a gestão completa da bibliografia.

Portanto, o gerenciamento de artigos recuperados de bases eletrônicas pode ser realizado mediante as ferramentas digitais citadas, colaborando para que haja uma otimização no processo da construção de um documento científico.

MATERIAIS E MÉTODOS

Para o início de uma pesquisa científica é necessário que se tenha definido o que será investigado, ou seja, a questão ou pergunta inicial. A partir dessa definição, é necessário realizar uma pesquisa abrangente nas bases de dados de periódicos, a fim de obter artigos que já trataram sobre o assunto. Para o processo de pesquisa de artigos é necessário definir um conjunto de critérios como palavras-chave, intervalo de publicação e número de citações, os quais permitirão realizar as buscas e a seleção de referências com mais facilidade, acurácia e qualidade.

Segundo (PIMENTA *et al.*, 2017), os seguintes critérios são necessários para iniciar o processo de pesquisa de artigos:

- Escolha de plataformas indexadoras de periódicos;
- Intervalo de publicação dos artigos – últimos cinco anos completos;
- Relevância dos indicadores de qualidade dos periódicos (FI);
- Publicações realizadas somente no formato de artigo;
- Seleção das palavras-chave ou *keywords*.

Inicialmente, a escolha de uma plataforma indexadora de periódicos pode tornar o processo de pesquisa exaustivo. Para isso, é indicado o acesso ao Portal de Periódicos CAPES (<https://www.periodicos.capes.gov.br>), o qual é um ponto de partida adequado quando não há conhecimento sobre um periódico ou tema específico, oferecendo opções de busca por assunto, periódico, livro ou base de periódicos, além possibilitar a personalização da busca (PACHECO RAMOS; MAGALHAES DE OLIVEIRA; SANTOS, 2018).

Quando já há conhecimento sobre o tema, o mais indicado é a utilização de bases de periódicos e ferramentas de busca com grande relevância como o *Scopus*, *Web of Science*, *IEEE Xplore*, *Google Scholar* e a Biblioteca Eletrônica Scielo (OLIVEIRA *et al.*, 2015, PIMENTA *et al.*, 2017). O *Google Scholar* não é uma base de periódicos, mas sim um agregador de resultados de publicações (PACKER, 2014) que permite realizar buscas para documentos indexados.

Além de periódicos, é importante realizar pesquisas em Anais de Congressos ou *Conference Proceedings*, os quais são uma coleção de trabalhos científicos publicados em congressos, seminários, encontros, *workshops*, simpósios, entre outros, tanto nacionais como internacionais sobre um determinado tema (WEB of SCIENCE, 2012).



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

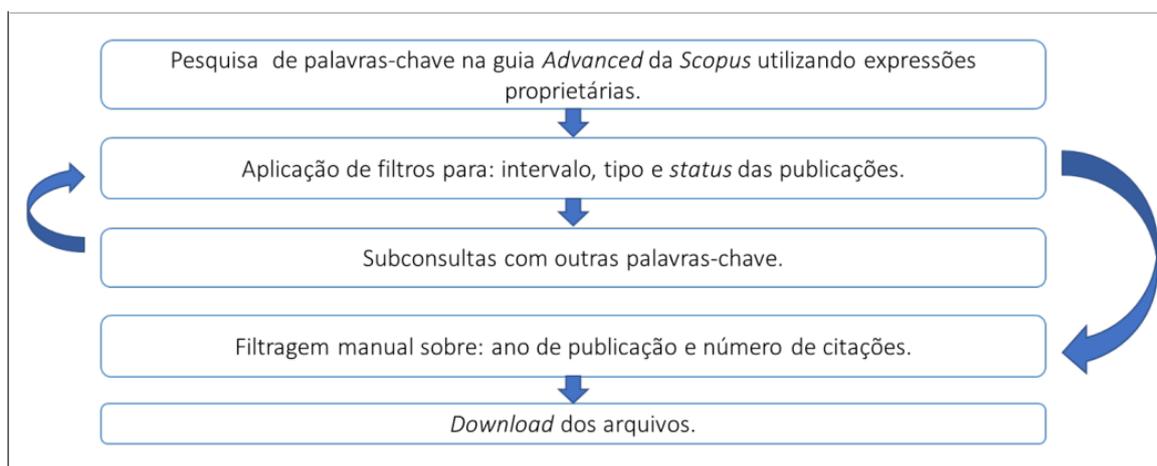
mesma sequência nas pesquisas, uma vez que combinações diferentes impactam nos resultados, o que pode gerar um número elevado de resultados com pouca ou nenhuma relevância, inviabilizando o processo inicial do levantamento dos dados sobre as publicações. Isso ocorre devido aos diferentes tipos de algoritmos presentes nas plataformas para realização das consultas (HABIBI; POPESCU-BELIS, 2015).

A base de periódicos *Scopus* foi selecionada para a realização das pesquisas por ser a maior base de dados bibliográficos sobre publicações científicas, disponibilizando resumos, citações e textos completos de periódicos acadêmicos com revisão por pares (MAÑANA-RODRÍGUEZ, 2015), além de fornecer métodos de buscas sofisticados, incluindo a possibilidade de uso de expressões proprietárias de fácil entendimento.

MODELO DE PESQUISAS NA PLATAFORMA SCOPUS

Utilizando as palavras-chave descritas anteriormente e com autenticação realizada via uma IES, iniciou-se o processo de pesquisa aplicando-se a metodologia proposta neste trabalho, mostrada na Figura 2.

Figura 2: Processo de pesquisa proposto para pesquisa de artigos.



Fonte: Os Autores (2022).

As pesquisas iniciais foram realizadas utilizando o procedimento mostrado na Figura 2 para todos os termos da lista previamente selecionada, aplicando-se um filtro para o intervalo entre os anos de 2017 e 2021 (últimos cinco anos), e somente para publicações do tipo artigo, formato *All Open Access (Golden Open)*, em língua inglesa e marcados com o *status* de Final, por meio da expressão de busca no modo avançado da Scopus, mostrada na Figura 3, resultando em 18.087 documentos.

A justificativa para seleção de artigos *All Open access (Golden Open)* é o fato de que a recuperação dos arquivos digitais não necessita de acesso por meio de assinatura ou por uma IES, garantindo a reprodutibilidade desta pesquisa no que diz respeito à aquisição dos arquivos originais



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE
DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

das publicações. Os procedimentos a seguir foram realizados na data de 10 nov. 2022, podendo apresentar diferenças numéricas em reproduções posteriores devido a alterações na plataforma Scopus.

Figura 3: Expressão de busca da primeira pesquisa.

Expressão Scopus	<i>ALL ("word cloud") OR ("text" AND "mining") OR ("text mining") AND (LIMIT-TO (OA , "publisherfullgold")) AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2021))</i>
------------------	---

Fonte: Os Autores (2022).

O resultado obtido na pesquisa inicial mostrou a existência de um elevado número de títulos (18.087) que não se relacionavam diretamente ao assunto *Word Cloud*, embora contivessem as palavras-chave com outras relações nas publicações. Para melhorar a qualidade dos resultados e diminuir o número de documentos, decidiu-se utilizar somente a palavra-chave “*word cloud*”, mantendo-se os parâmetros anteriores para o refinamento da busca, conforme mostrado na Figura 4, resultando em 298 documentos.

Figura 4: Expressão de busca da segunda pesquisa.

Expressão Scopus	<i>ALL("word cloud") AND (LIMIT-TO (PUBSTAGE,"final")) AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO (PUBYEAR,2017) OR LIMIT-TO (PUBYEAR,2018) OR LIMIT-TO (PUBYEAR,2019) OR LIMIT-TO (PUBYEAR,2020) OR LIMIT-TO (PUBYEAR,2021)) AND (LIMIT-TO (OA," publisherfullgold"))</i>
------------------	--

Fonte: Os Autores (2022).

Apesar da diminuição de exemplares, foi constatado que ainda existiam documentos onde a palavra-chave não estava presente. Essa descoberta ocorreu após a leitura de alguns resumos (*abstracts*), evidenciando a necessidade da aplicação de um novo filtro. Assim, foi aplicado um novo filtro Figura 5, realizando a pesquisa da palavra-chave contida apenas no *abstract* e mantendo-se os parâmetros anteriores, resultando em 109 documentos.

Figura 5: Expressão de busca da segunda pesquisa.

Expressão Scopus	<i>ABS ("word cloud") AND (LIMIT-TO (OA , "publisherfullgold")) AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2021))</i>
------------------	---

Fonte: Os Autores (2022).

Nessa última pesquisa, o número de exemplares obtidos mostrou que o assunto apresenta relevância de forma geral, principalmente nas áreas da medicina (33), ciências sociais (32), ciências da computação (28), engenharia (12) e ciências ambientais (10). Esses números são fornecidos em um painel especial no Scopus denominado *Subject Area* para cada pesquisa realizada, cuja divisão é



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

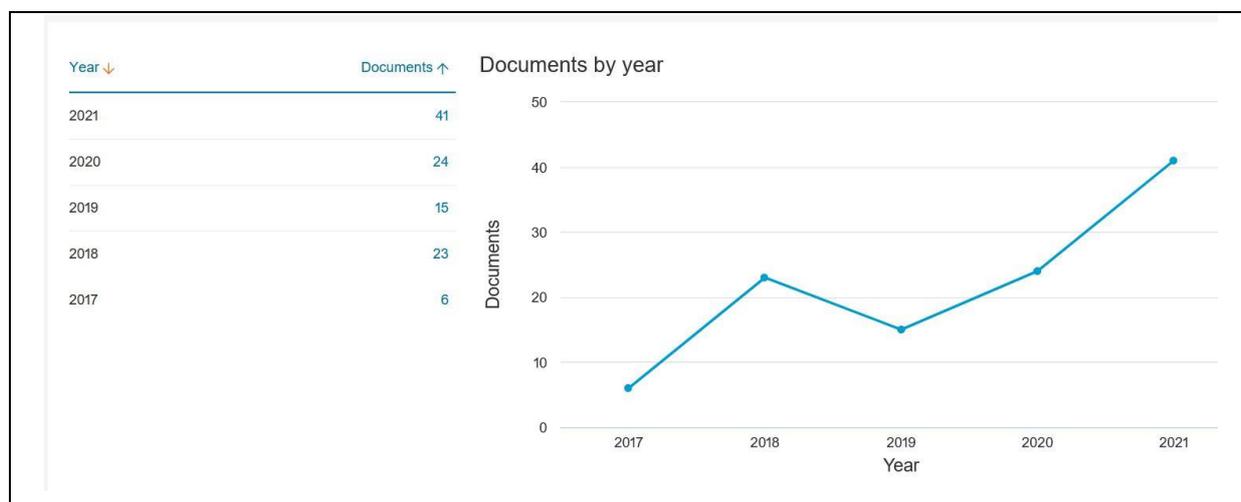
PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

mantida pelo Scimago - portal que fornece indicadores de produções científicas contidas no banco de dados do Scopus (VILLASEÑOR-ALMARAZ *et al.*, 2019).

Na Figura 6 é mostrada a evolução das publicações ao longo do período de cinco anos, totalizando 109 artigos, mostrando que o uso da técnica de *Word Cloud* está em crescimento nos últimos anos para apresentar resultados de aglomeração de palavras de forma rápida e objetiva.

Na Figura 7 são apresentados os percentuais de aparição da palavra-chave nas áreas de pesquisa, mostrando que a técnica de *Word Cloud* é utilizada em todas as áreas do conhecimento, considerando aquelas abordadas pela Scopus. Essas áreas podem sofrer alterações dependendo da plataforma em que os dados forem pesquisados.

Figura 6: Evolução das publicações entre os anos de 2017 e 2021.



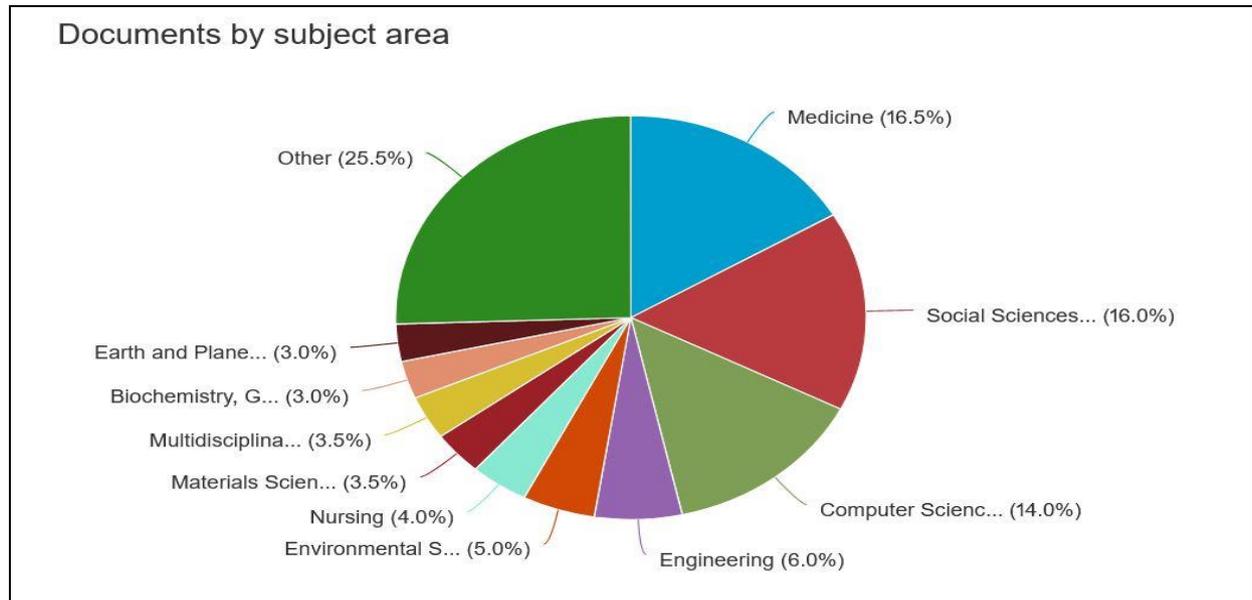
Fonte: Os Autores (2022).



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Figura 7: Publicações por área entre os anos de 2017 e 2021.



Fonte: Os Autores (2022).

Além das opções apresentadas, a plataforma Scopus ainda oferece outras ferramentas gráficas para auxiliar no processo de análise dos resultados da pesquisa. Esses recursos podem ser acessados diretamente nos resultados das pesquisas.

Em continuidade ao processo, foi realizada a exportação dos dados recuperados da plataforma Scopus. Essa exportação resultou em um arquivo no formato tabular CSV (*Comma-separated values*) contendo dados sobre as 109 publicações. Na Figura 8 é mostrado um fragmento do arquivo recuperado.

Figura 8: Publicações por área entre os anos de 2017 e 2021.

Author(s) ID	Titles	Year	Source title	DOI	Cited by	Open Access	Source	EID
55626073900; 5721	Social representati	2018	Movimento	10.22456/1982-8918.65543	2	All Open Access; Gt Scopus	2-s2.0-85049092897	
57211874433; 368	Analysis of scientifi	2018	Innovation and Ma	10.1108/INMR-02-2018-006	4	All Open Access; Gt Scopus	2-s2.0-85061662475	
55737826100; 109	A new indicator to	2021	Applied Sciences (S	10.3390/app11041894	1	All Open Access; Gt Scopus	2-s2.0-85101925365	
57224542645; 368	Developing adaptiv	2021	JMIR Serious Game	10.2196/25997	4	All Open Access; Gt Scopus	2-s2.0-85107849831	
36552939700; 572	Using topic modelli	2021	PLoS ONE	10.1371/journal.pone.026040	0	All Open Access; Gt Scopus	2-s2.0-85121053575	
56327328700; 555	Revealing New Tec	2021	International Journ	10.15837/ijcc.2021.2.4101	0	All Open Access; Gt Scopus	2-s2.0-85104036158	
57192176905; 506	Identifying urban ft	2018	ISPRS International	10.3390/ijgi7120459	36	All Open Access; Gt Scopus	2-s2.0-85061453906	
57216782966; 561	Impact model of tc	2020	Sustainability (Swit	10.3390/SU12083430	3	All Open Access; Gt Scopus	2-s2.0-85084647993	
57193608906; 571	The Influence of CC	2021	Frontiers in Psycho	10.3389/fpsyg.2021.681091	6	All Open Access; Gt Scopus	2-s2.0-85109032132	
55737826100; 353	Public responses to	2021	ISPRS International	10.3390/ijgi10030126	3	All Open Access; Gt Scopus	2-s2.0-85106471802	
55845931300; 859	Empirical advances	2017	BMC Medical Infor	10.1186/s12911-017-0519-0	16	All Open Access; Gt Scopus	2-s2.0-85027959746	
57188721939; 572	Let's talk numbers:	2018	BMJ Open	10.1136/bmjopen-2018-0230	2	All Open Access; Gt Scopus	2-s2.0-85054777963	
57208328701; 572	A text mining applic	2018	Central Bank Revie	10.1016/j.cbrev.2018.05.001	6	All Open Access; Gt Scopus	2-s2.0-85064514536	
57203872029; 571	Prospecting Inform	2018	IEEE Access	10.1109/ACCESS.2018.28702	36	All Open Access; Gt Scopus	2-s2.0-85053321373	
57200417246; 572	Bibliometric analys	2020	Revista Baiana de I	10.18471/rbe.v34.37107	1	All Open Access; Gt Scopus	2-s2.0-85110210926	
56464173700; 650	Trends in HIV termi	2018	JMIR Public Health	10.2196/publichealth.8552	4	All Open Access; Gt Scopus	2-s2.0-85047518723	
43261733900; 575	Does Online Educa	2021	Forum Geografic	10.5775/fg.2021.056.d	0	All Open Access; Gt Scopus	2-s2.0-85126761856	
55695244600; 572	Status of weed con	2021	Advances in Weed	10.51694/AdvWeedSci/2021;:	4	All Open Access; Gt Scopus	2-s2.0-85106737922	
57216763062	Examining the rural	2020	GeoJournal of Tour	10.30892/jgtg.29212-487	6	All Open Access; Gt Scopus	2-s2.0-85084551677	
24768324100; 571	Capturing twitter n	2021	Sustainability (Swit	10.3390/su132313356	3	All Open Access; Gt Scopus	2-s2.0-85120742703	
26323728300; 551	Research priorities	2018	Journal of Global H	10.7189/jogh.08.010501	16	All Open Access; Gt Scopus	2-s2.0-85045373935	
36110358700; 572	Keyword analysis o	2020	Journal of People, I	10.11628/kspp.2020.23.2.23	3	All Open Access; Gt Scopus	2-s2.0-85101279685	
57212769265	Even when no one	2019	Education Sciences	10.3390/educsci9030233	1	All Open Access; Gt Scopus	2-s2.0-85073352936	

Fonte: Os Autores (2022).



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Uma vez publicado um artigo, as citações são uma métrica que evidencia a qualidade da produção. Embora citações não ocorram de forma imediata após a publicação de um artigo, podendo até mesmo nunca serem citados, vários fatores podem influenciar para que um artigo possua nenhuma ou poucas citações, como a recência da publicação, assunto muito específico, periódico com baixo FI, entre outros. Assim, um novo filtro foi aplicado para determinar qualitativamente quais artigos seriam lidos na íntegra para a formação do referencial teórico inicial. O critério utilizado foi de artigos que possuíam dez ou mais citações no período de cinco anos. Esse novo filtro resultou em 22 documentos, somando 2.263 citações entre eles.

O Scopus e o Qualis-CAPES não fornecem as métricas FI, SJR e o *h-index* dos periódicos na exportação dos dados, exigindo uma consulta manual no site *Resurchify* (<https://www.resurchify.com/ranking>), que fornece esses dados. O arquivo contendo informações sobre as publicações obtido anteriormente foi atualizado com os novos dados.

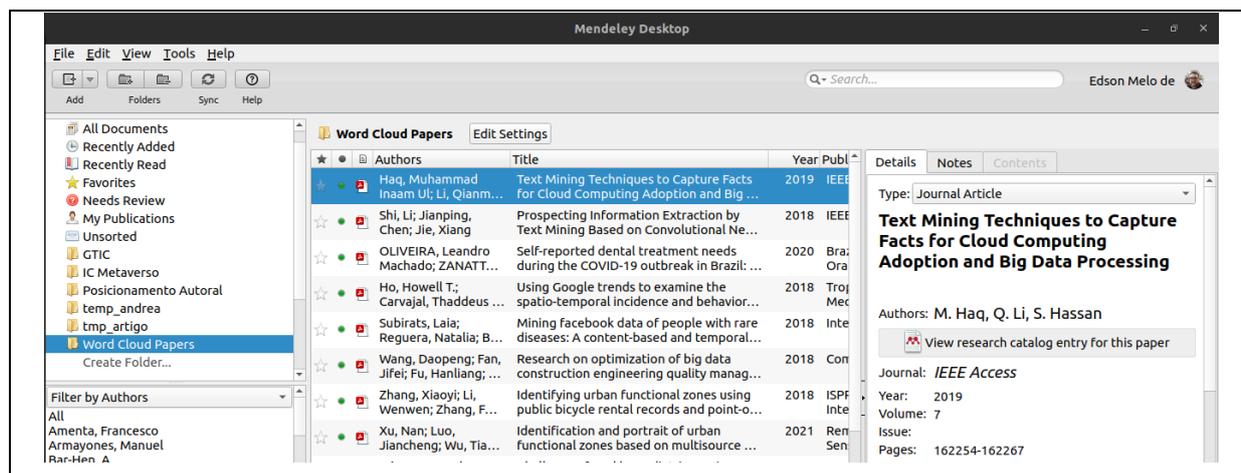
Por fim, utilizando o DOI, os 22 artigos em formato digital *Portable Document Format* (PDF) foram obtidos através do Resolve a *DOI Name* (<https://dx.doi.org/>), que é o mantenedor dos registros eletrônicos dos documentos digitais nesse formato.

CATALOGAÇÃO DOS ARTIGOS E MANUTENÇÃO DAS INFORMAÇÕES

A construção do referencial teórico não é algo estanque e, até que o documento não esteja finalizado, principalmente para dissertações e teses, é importante revisitar a literatura, realizando novas buscas, de modo a detectar se não há algo novo que foi publicado no intervalo da última pesquisa. Isso reforça que o trabalho que está sendo desenvolvido ainda não foi publicado e traz novidades dentro do tema estudado.

Nesta etapa foram importados os 22 arquivos para o programa gerenciador de bibliografias Mendeley. Na Figura 9 é mostrado um fragmento da tela do Mendeley com os artigos importados ao centro.

Figura 9: Fragmento da tela do programa Mendeley com dados sobre artigos importados.



Fonte: Captura de tela do programa Mendeley (2022).



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Na Figura 9 é mostrado um painel localizado na lateral esquerda, onde é possível criar pastas para organizar arquivos e referências sobre publicações. Em destaque está a pasta denominada *Word Cloud Papers* com os arquivos da pesquisa. Essa flexibilidade permite um elevado grau de organização sobre temas. Ainda na mesma lateral, parte inferior, visualiza-se parcialmente uma lista com os autores dos artigos importados, em ordem alfabética.

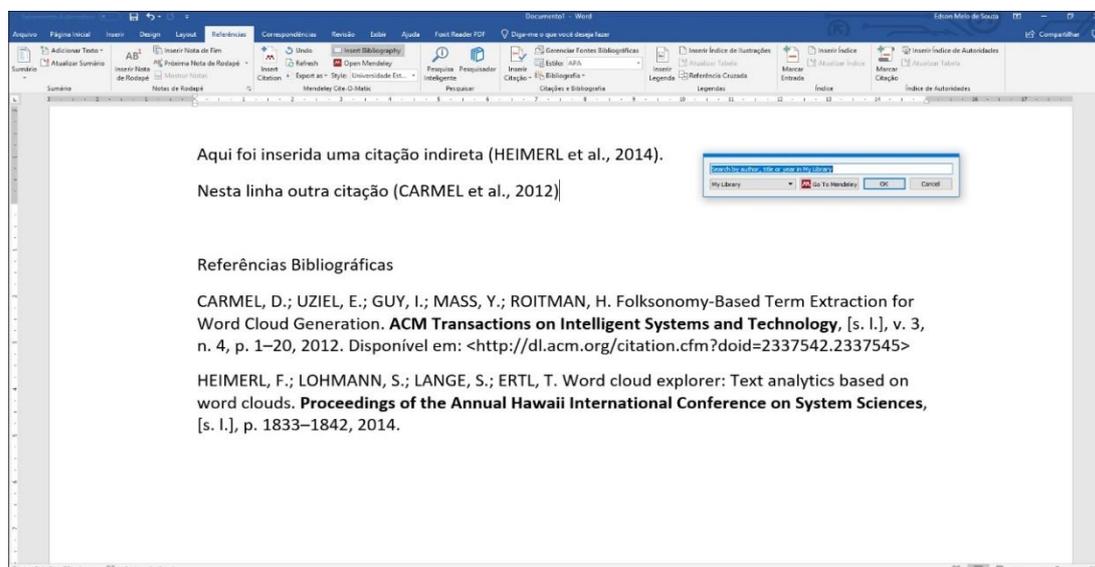
Na lateral direita encontram-se alguns detalhes sobre o documento que está selecionado (em destaque), recuperados diretamente dos arquivos importados. Também é possível recuperar informações de um artigo utilizando o DOI atribuído ao documento de forma *online*, diretamente das fontes mantenedoras.

Por fim, na parte superior da tela é possível realizar pesquisas por palavras-chave diretamente nos textos dos arquivos e em anotações realizadas pelo utilizador. Ademais, o programa permite que os arquivos sejam armazenados e sincronizados na internet, o que facilita a visualização e recuperação sem a necessidade de possuir o programa instalado em um dispositivo computacional.

OUTROS RECURSOS DA FERRAMENTA MENDELEY

Os recursos de gerenciamento de bibliografia são complexos e, neste aspecto, o programa Mendeley fornece recursos que podem facilitar a inserção de citações e a criação da lista de referências de forma automática em processadores de texto, como o Microsoft Word (MS Word), aplicando inclusive estilos de formatação de bibliografia como o ABNT, APA, IEEE, entre outros. Tais estilos podem ser incluídos através do próprio programa em suas configurações. O recurso de citação e inclusão automática da lista de referências utilizado no processador de textos Microsoft Word, no padrão ABNT, é mostrado em um fragmento simples de texto na Figura 10.

Figura 10: Citações e lista de referências inseridas no MS Word pelo programa Mendeley.



Fonte: Os Autores (2022).



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE
DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Na Figura 10 é possível visualizar um exemplo de inserção para duas citações e a lista de referências gerada em ordem alfabética e no padrão ABNT, recurso utilizado para criação deste trabalho. No alto da imagem visualiza-se a ferramenta de inserção de citações.

Outro ponto a ser destacado é sobre a inserção da lista de referências que, quando inserida no processador de textos, pode precisar de ajustes dependendo de como os dados foram importados ou incluídos nos detalhes do artigo dentro do programa Mendeley.

ANÁLISE DOS RESULTADOS

A metodologia proposta proporcionou condições de reduzir de 18.087 para 22 o número de documentos a serem utilizados para a construção do referencial teórico inicial sobre o tema *Word Cloud* ou Nuvem de Palavras. Já a utilização do programa Mendeley permitiu organizar de forma fácil e estruturada os detalhes de cada artigo, mantendo a qualidade da pesquisa. Na Tabela 1 são mostrados os 22 artigos utilizados para a construção do referencial teórico. Os arquivos das publicações estão disponíveis no repositório público (https://osf.io/yjs5v/?view_only=d3dfb98ffd854313a3250939d443f42c).



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Figura 11: Artigos selecionados para a construção do referencial teórico.

Título	Year	FI	h-index	SJR	Qualis	Citações
<i>MiRBase: From microRNA sequences to function</i>	2019	19,33	569	8,241	A1	1603
<i>Analysis of patient narratives in disease blogs on the internet: An exploratory study of social pharmacovigilance</i>	2017	10,88	41	1,893	B2	188
<i>Tweet topics and sentiments relating to COVID-19 vaccination among Australian twitter users: Machine learning analysis</i>	2021	7,52	158	1,736	A1	55
<i>"What kind of a science is sustainability science?" An evidence-based reexamination</i>	2018	4,17	109	0,664	B2	47
<i>Big data analysis for personalized health activities: Machine learning processing for automatic keyword extraction approach</i>	2018	3,28	64	0,54	B2	40
<i>Text mining with sentiment analysis on seafarers' medical documents</i>	2021	21,35	132	4,584	A1	38
<i>Prospecting Information Extraction by Text Mining Based on Convolutional Neural Networks-A Case Study of the Lala Copper Deposit, China</i>	2018	4,34	158	0,927	B3	36
<i>Identifying urban functional zones using public bicycle rental records and point-of-interest data</i>	2018	3,51	52	0,721	B3	36
<i>Research on optimization of big data construction engineering quality management based on RnN-LSTM</i>	2018	1,12	38	1,788	B1	30
<i>Understanding smart city—a data-driven literature review</i>	2020	4,17	109	0,664	B2	24
<i>Role of artificial intelligence in shaping consumer demand in e-commerce</i>	2020	4,00	38	0,793	n/c	23
<i>Empirical advances with text mining of electronic health records</i>	2017	3,64	79	0,833	A2	16
<i>Mining facebook data of people with rare diseases: A content-based and temporal analysis</i>	2018	4,54	138	0,814	A2	16
<i>Research priorities for adolescent health in low and middle-income countries: A mixed-methods synthesis of two separate exercises</i>	2018	4,18	44	1,349	A1	16
<i>Using Google trends to examine the spatio-temporal incidence and behavioral patterns of dengue disease: A case study in metropolitan Manila, Philippines</i>	2018	3,59	24	0,924	n/c	15
<i>Challenges faced by pediatric nursing workers in the face of the covid-19 pandemic</i>	2020	1,56	39	0,434	B1	14
<i>Global sentiments surrounding the COVID-19 pandemic on Twitter: Analysis of Twitter trends</i>	2020	10,88	41	1,893	B2	13
<i>Determining public opinion of the COVID-19 pandemic in South Korea and Japan: Social network mining on Twitter</i>	2020	3,01	33	0,614	B2	12
<i>Dashboard of Sentiment in Austrian Social Media During COVID-19</i>	2020	3,61	10	0,735	B1	11
<i>Self-reported dental treatment needs during the COVID-19 outbreak in Brazil: an infodemiological study</i>	2020	2,57	50	0,736	A2	10
<i>Text Mining Techniques to Capture Facts for Cloud Computing Adoption and Big Data Processing</i>	2019	4,34	158	0,927	B3	10
<i>Identification and portrait of urban functional zones based on multisource heterogeneous data and ensemble learning</i>	2021	13,63	303	3,862	C	10

Fonte: Os Autores (2022).

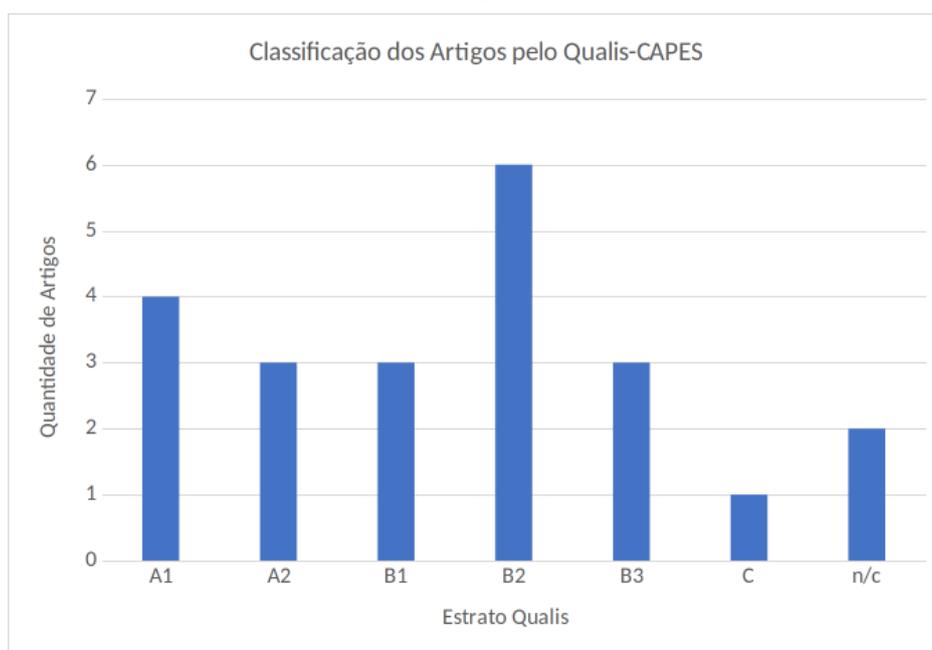


RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Na Tabela 1 são mostrados o nome dos periódicos, títulos dos artigos, ano de publicação, os indicadores de qualidade FI, *h-index*, SJR, Qualis-CAPES e o número de citações. Na Figura 12 é mostrada a distribuição quantitativa do estrato Qualis-CAPES para os 22 artigos selecionados, representando 18,8% (A1), 13,64% (A2), 13,64% (B1), 18,8% (B2), 13,64% (B3) e 4,55% (C) reforçando a qualidade das publicações. Dois artigos (9,09%) não possuíam classificação no Qualis-CAPES, embora possuam FI elevado com 4,0 e 3,59. É importante destacar que o indicador *h-index* sofre muitas variações devido aos artigos terem sido publicados em áreas diferentes. Entretanto, apresentam informações relevantes acerca do tema pesquisado.

Figura 12 – Distribuição dos artigos pelo estrato Qualis-CAPES.



Fonte: Os Autores (2022).

Considerando o objetivo deste trabalho, os artigos apresentam indicadores que permitem utilizá-los para a criação do referencial teórico inicial. Outro indicador importante para verificação da qualidade e relevância das publicações são as citações. Os 22 artigos selecionados anteriormente totalizaram 2.263 citações, onde os menos citados receberam dez (10) e o maior 1.603 citações. O artigo com maior citação foi retirado da amostra por gerar distorções na apresentação gráfica. Na Figura 13 são mostrados os dados sobre os 21 artigos selecionados.



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE
DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

Figura 13: Número de citações dos artigos.



Fonte: Os Autores (2022)

Por fim, os resultados analisados evidenciam que as publicações selecionadas para o tema *Word Cloud* pesquisado são relevantes e apresentam qualidade científica, além de serem usados como ferramenta visual para apresentar resultados.

CONSIDERAÇÕES

A construção inicial do referencial teórico para um trabalho científico é uma etapa de grande importância, uma vez que toda pesquisa deve ser fundamentada a partir de referências validadas pela comunidade científica. Assim, este trabalho buscou apresentar um processo de como alunos iniciantes ou com pouca experiência em pesquisas acadêmicas podem construir um referencial teórico inicial, mostrando os principais termos e conceitos inerentes à pesquisa científica. Ademais, apresentou um processo de como realizar uma pesquisa bibliográfica em bases de artigos de periódicos e plataformas indexadoras com o uso da TI como pilar de sustentação tecnológica.

Utilizando a ferramenta *Mendeley* para gerenciamento de referências, foram apresentados procedimentos de como realizar a catalogação, inserção de anotações e aplicação de destaques nos artigos obtidos nas pesquisas para destacar os pontos importantes, permitindo assim otimizar as buscas nas referências e elaborar o referencial teórico inicial.

Como resultado, obteve-se 22 artigos relevantes e de qualidade por meio de filtragem pelos indicadores bibliométricos FI, SJR, *h-index* e do estrato Qualis-CAPES, que podem direcionar o aprofundamento da pesquisa sobre o tema *Word Cloud*, objeto de aplicação neste trabalho. Ademais, mostrou como facilitar o gerenciamento das referências por meio da inserção automática de citações e criação da lista de referências no processador de textos Microsoft Word.

Por fim, este trabalho mostrou como utilizar estratégias para extração de dados de artigos nas plataformas indexadoras e de periódicos com o suporte de ferramentas de TI, apoiando o



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

processo inicial de construção de um referencial teórico de qualidade, minimizando a falta de experiência de pesquisadores iniciantes para a construção de um referencial teórico inicial.

REFERÊNCIAS

ARAÚJO, R. M. *et al.* Gestão de periódicos: um estudo na área de Administração, Ciências Contábeis e Turismo. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 22, n. 49, p. 42, 2017. DOI: 10.5007/1518-2924.2017v22n49p42

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: Informação e documentação: Referências. Rio de Janeiro: ABNT, 2018.

AZEVEDO, D. Revisão de Literatura, Referencial Teórico, Fundamentação Teórica e Framework Conceitual em Pesquisa - diferenças e propósitos. **Working Paper**, 2016. Disponível em: <https://unisinus.academia.edu/DeborazAzevedo/Papers>. Acesso em: 21 nov. 2022.

BORNMANN, L.; DANIEL, H. D. Does the h-index for ranking of scientists really work? **Scientometrics**, v. 65, n. 3, p. 391-392, 2005. DOI: 10.1007/s11192-005-0281-4

BORNMANN, L.; DANIEL, H. D. The state of h index research. Is the h index the ideal way to measure research performance? **EMBO reports**, v. 10, n. 1, p. 2-6, 2009. DOI: 10.1038/embor.2008.233

COPPERSMITH, G.; KELLY, E. Dynamic Wordclouds and Vennclouds for exploratory data analysis. **Association for Computational Linguistics Workshop on Interactive Language Learning and Visualization**, p. 22-29, 2014.

CRESWELL, J. W.; POT, C. N. **Qualitative inquiry and research design**: Choosing among five approaches. 4. ed. Washington: Sage Publications, 2017. 448 p.

DROESCHER, F. D.; SILVA, E. L. Da. O pesquisador e a produção científica. **Perspectivas em Ciência da Informação**, v. 19, n. 1, p. 170-189, 2014. DOI: 10.1590/s1413-99362014000100011

EVANS, M. C.; CVITANOVIC, C. An introduction to achieving policy impact for early career researchers. **Palgrave Communications**, 2018. DOI: 10.1057/s41599-018-0144-2

FARIAS, S. A. DE. Internacionalização dos Periódicos Brasileiros. **Revista de Administração de Empresas**, v. 57, n. 4, p. 401-404, 2017. DOI: 10.1590/s0034-759020170409

FERREIRA, M. A. S. P. V.; SERRA, F. R. A. Coautoria em Artigos Científicos de Administração: Perspectivas de Pesquisadores Internacionais. **Administração: Ensino e Pesquisa**, v. 16, n. 4, p. 663, 2015. DOI: 10.13058/raep.2015.v16n4.381.

FRIGERI, M.; MONTEIRO, M. S. A. Qualis Periódicos: indicador da Política Científica no Brasil? **Estudos de Sociologia**, v. 19, n. 37, p. 299-315, 2014. Disponível em: <http://seer.fclar.unesp.br/estudos/article/download/6266/5227>. Acesso em: 21 nov. 2022.

GIUFFRIDA, C. *et al.* Are all citations worth the same? Valuing citations by the value of the citing items. **Journal of Informetrics**, v. 13, n. 2, p. 500-514, 2019. DOI: 10.1016/j.joi.2019.02.008

HABIBI, M.; POPESCU-BELIS, A. Keyword Extraction and Clustering for Document Recommendation in Conversations. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, v. 23, n. 4, p. 746-759, 2015. DOI: 10.1109/TASLP.2015.2405482



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE
 DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
 Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

HUGUET, J.; GAYA, J. M.; PALOU, A. B. J. Actas Urológicas Españolas: El estilo de la comunicación científica. **Actas Urológicas Españolas** (English Edition), v. 42, Issue 9, p. 551-556 2018. DOI: 10.1016/j.acuro.2018.02.013

IVEY, C.; CRUM, J. Choosing the right citation management tool: EndNote, Mendeley, RefWorks, or Zotero. **Journal of the Medical Library Association**, v. 106, n. 3, p. 399–403, 2 jul. 2018.

KRATOCHVÍL, J. Comparison of the Accuracy of Bibliographical References Generated for Medical Citation Styles by EndNote, Mendeley, RefWorks and Zotero. **Journal of Academic Librarianship**, v. 43, n. 1, p. 57-66, 2017. DOI: 10.1016/j.acalib.2016.09.001

LEYDESDORFF, L.; WOUTERS, P.; BORNMANN, L. Professional and citizen bibliometrics: complementarities and ambivalences in the development and use of indicators-a state-of-the-art report. **Scientometrics**, v. 109, n. 3, p. 2129-2150, 2016. DOI: 10.1007/s11192-016-2150-8

LORDÊLO, F. S. A internet como ferramenta de divulgação científica sobre energias renováveis. **Comunicologia - Revista de Comunicação da Universidade Católica de Brasília**, v. 4, n. 1, p. 313-332, 2011. DOI: 10.17921/1415-5141.2011

LUNDBERG, J. *et al.* Is it better or just the same? Article identification strategies impact bibliometric assessments. **Scientometrics**. v. 66, p. 183-197, 2006. DOI: 10.1007/s11192-006-0013-4

MAÑANA-RODRÍGUEZ, J. A critical review of SCImago Journal & Country Rank. **Research Evaluation**, v. 24, n. 4, p. 343–354, out. 2015. DOI: 10.1093/reseval/rvu008

MATTSSON, P.; SUNDBERG, C. J.; LAGET, P. Is correspondence reflected in the author position? A bibliometric study of the relation between corresponding author and byline position, **Scientometrics**, p. 99-105, 2011. DOI: 10.1007/s11192-010-0310-9

MORE: Mecanismo online para referências, versão 2.0. **UFSC Rexlab**, 2013. Disponível em: <http://www.more.ufsc.br/>. Acesso em: 20 nov. 2022.

MUELLER, S. P. M. O periódico científico. **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: UFMG. 2000. Disponível em: <https://acervo.enap.gov.br/cgi-bin/koha/opac-detail.pl?biblionumber=6348>. Acesso em: 22 nov. 2022.

MUGNAINI, R. Ciclo Avaliativo De Periódicos No Brasil: Caminho Virtuoso ou Colcha de Retalhos? *In: XVI Encontro Nacional de Pesquisa em Ciência da Informação (XVI ENANCIB)*. João Pessoa, PB, 2012. Disponível em: <http://200.20.0.78/repositorios/handle/123456789/2947> Acesso em: 20 nov. 2022.

NILASHI, M. *et al.* An interpretive structural modelling of the features influencing researchers' selection of reference management software. **Journal of Librarianship and Information Science**, v. 51, n. 1, p. 34-46, 2019. DOI: 10.1177/0961000616668961

OLIVEIRA, A. B. *et al.* Comparação entre o Qualis/Capes e os índices H e G: o caso do portal de periódicos UFSC. **Informação & Informação**, v. 20, n. 1, p. 70, 2015. DOI: 10.5433/1981-8920.2015v20n1p70

OLIVEIRA, É. B. P. M. Periódicos científicos eletrônicos: definições e histórico. **Informação & Sociedade: Estudos**, v. 18, n. 2, p. 69-77, 2008.

PACHECO RAMOS, M. D.; MAGALHAES DE OLIVEIRA, R. D. C.; SANTOS, M. R. Estado da arte da pesquisa (auto)biográfica: uma análise do portal de periódicos CAPES. **Revista Brasileira de Pesquisa (Auto)biográfica**, v. 2, n. 5, p. 449, 2018. DOI: 10.31892/rbpab2525-



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
 Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

426X.2017.v2.n5.p449-469

PACKER, A. L. A eclosão dos periódicos do Brasil e cenários para o seu porvir. **Educação e Pesquisa**, v. 40, n. 2, p. 301–323, 2014. DOI: 10.1590/s1517-97022014061860

PATRUS, R.; DANTAS, D. C.; SHIGAKI, H. B. Pesquisar é preciso. Publicar não é preciso: história e controvérsias sobre a avaliação por pares. **Revista da Avaliação da Educação Superior** (Campinas), v. 21, n. 3, p. 799-820, 2016. DOI: 10.1590/S1414-40772016000300008

PEREIRA, M. G.; GALVÃO, T. F. Etapas de busca e seleção de artigos em revisões sistemáticas da literatura. **Epidemiologia e Serviços de Saúde**, v. 23, n. 2, p. 369-371, 2014. DOI: 10.5123/s1679-49742014000200019

PIMENTA, A. A. *et al.* A Bibliometria nas Pesquisas Acadêmicas. **SCIENTIA - Revista de Ensino, Pesquisa e Extensão**, v. 4, n. 7, p. 1-13, 2017.

RUIZ, M. A.; GRECO, O. T.; BRAILE, D. M. Fator de impacto, importância e influência no meio editorial, acadêmico e científico. **Revista Brasileira de Hematologia e Hemoterapia**, São Paulo, v. 31, n. 5, p. 355-360, 2009. DOI: 10.1590/S1516-84842009005000080

SANDES-GUIMARÃES, L. V. de; DINIZ, E. H. Gestão de periódicos científicos: estudo de casos em revistas da área de Administração. **Revista de Administração**, v. 49, n. 3, p. 449–461, 2014. DOI: 10.5700/rausp1160

SANTOS, Ynara Victoria Borges; FRANÇA, Eliacir Neves. O processo de formulação dos indicadores de excelência da avaliação realizada nos cursos de pós-graduação stricto sensu elaborados pela CAPES. **Brazilian Journal of Development**, v. 8, n. 1, p. 3549-3561, 2022.

SCIMAGO. SJR - SCImago Journal & Country Rank, 2021. Disponível em: <https://www.scimagojr.com/>. Acesso em: 21 nov. 2022.

SILVA, C. F. DA *et al.* Formação Docente na Área Contábil: Contribuições da Disciplina de Metodologia do Ensino oferecida na Pós-Graduação Stricto Sensu. **Sociedade, Contabilidade e Gestão**, v. 14, n. 3, p. 144–162, 12 ago. 2019.

SINGH, P.; SINGH, K. Exploring Automatic Search in Digital Libraries. *In: Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 15 jun. 2017. p. 236–241. DOI: 10.1145/3084226.3084275

SPRAT, T. **The History of the Royal Society of London, for the Improving of Natural Knowledge**. 4. ed. London: J. Knapton, 1734. Disponível em: https://play.google.com/books/reader?id=rtnO8tI4_hcC&lr=&printsec=frontcover&pg=GBS.PP1. Acesso em: 25 nov. 2022.

STREHL, L. O fator de impacto do ISI e a avaliação da produção científica: aspectos conceituais e metodológicos. **Ciência da informação**, v. 34, n. 1, p. 19-27, 2005. DOI: 10.1590/S0100-19652005000100003

TAVARES, F.; RODRIGUES, J.; FILHO, F. Metodologia de pesquisa bibliográfica com a utilização de método multicritério de apoio à decisão. **Production**, v. 24, n. 3, p. 508–520, 2014.

TEIXEIRA, F. C. *et al.* Formação de professores universitários na pós-graduação stricto sensu em Educação Física no Brasil: uma revisão sistemática. **Motrivivência**, v. 31, n. 59, p. 1–20, 30 jul. 2019.

TREINTA, F. T. *et al.* Metodologia de pesquisa bibliográfica com a utilização de método multicritério



RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

PROCESSO DE CONSTRUÇÃO DE REFERENCIAL TEÓRICO VIA EXTRAÇÃO DE
 DADOS DA PLATAFORMA SCOPUS E INDICADORES DE PRODUÇÃO
 Edson Melo de Souza, Gustavo Pessarelo Smulkowski Souza

de apoio à decisão. **Production**, v. 24, n. 3, p.508-520, 2014. DOI: 10.1590/S0103-65132013005000078

VILLASEÑOR-ALMARAZ, M. *et al.* Impact factor correlations with Scimago Journal Rank, Source Normalized Impact per Paper, Eigenfactor Score, and the CiteScore in Radiology, Nuclear Medicine & Medical Imaging journals. **La radiología medica**, v. 124, n. 6, p. 495–504, 6 jun. 2019. DOI: 10.1007/s11547-019-00996-z

VOUGA, G.; AMATUCCI, M. O que é bibliometria? Uma introdução ao Fórum. **Revista Eletrônica de Negócios Internacionais**, v. 10, p. 1-5, 2015. DOI: 10.18568/1980-4865.1021-5

WEB OF SCIENCE (Ed.). **CPCI Essay - IP & Science**, 2012. Disponível em: http://wokinfo.com/products_tools/multidisciplinary/webofscience/cpci/cpciessay. Acesso em: 21 nov. 2022.

YAMAKAWA, E. K. *et al.* Comparativo dos softwares de gerenciamento de referências bibliográficas: Mendeley, EndNote e Zotero. **Transinformação [online]**, n. 2, p. 167, 2014. DOI: 10.1590/0103-37862014000200006

YANG, S.; WOLFRAM, D.; WANG, F. The relationship between the author byline and contribution lists: a comparison of three general medical journals. **Scientometrics**, v. 110, n. 3, p. 1273-1296, 2017. DOI: 10.1007/s11192-016-2239-0