

PUBLICAÇÃO DE DADOS DE PESQUISA EM ENFERMAGEM: NOVAS PRÁTICAS EM CIÊNCIA ABERTA

Publication of research data in nursing: new practices in open science

Fabiana de Melo Amaral Gonçalves Pinto

Universidade Federal Fluminense,

Niterói, RJ, Brasil

fabianamelo@id.uff.br

<https://orcid.org/0000-0002-8118-3165> 

A lista completa com informações dos autores está no final do artigo 

RESUMO

Objetivo: Apresenta dados referente ao mapeamento de publicação de dados de pesquisa pela área da Enfermagem via *datapaper*, em periódicos científicos, e *dataset*, em repositório de dados.

Método: Os dados foram extraídos a partir da interface das seguintes bases de dados via Portal de Periódicos da Capes: *Web of science*, *Science direct*, *Scopus*, *Lilacs*, *Medline* e *Cinahl*. Utilizou-se também como fonte para extração de dados os repositórios multidisciplinares: *Figshare*, *Dryade* e *Zenodo*.

Resultado: A partir de buscas sistematizadas diretamente nas interfaces das fontes citadas, estruturou-se as tipologias documentais apontadas no artigo (*datapapers* e *datasets*) em arquivo tipo .csv disponibilizados no repositório *Figshare* sob o DOI (<https://doi.org/10.6084/m9.figshare.19790425.v3>).

Conclusões: Os dados da pesquisa foram organizados com base no padrão de metadados *Dublin Core*, além de metadados autorais criados para atender o objetivo da pesquisa. Em formato textual e numérico, consistem em dados secundários extraídos das interfaces das fontes citadas sobre a área da Enfermagem.

PALAVRAS-CHAVE: Publicação de dados de pesquisa. Enfermagem. Dados de pesquisa. Artigo de dados. Conjunto de dados de pesquisa.

ABSTRACT

Objective: It presents data regarding the mapping of publication of research data by the Nursing area via *datapaper*, in scientific journals, and *dataset*, in a data repository.

Method: Data were extracted from the interface of the following databases via the Capes Journal Portal: *Web of science*, *Science direct*, *Scopus*, *Lilacs*, *Medline* and *Cinahl*. The multidisciplinary repositories: *Figshare*, *Dryade* and *Zenodo* were also used as a source for data extraction.

Result: Based on systematized searches directly in the interfaces of the cited sources, the document typologies pointed out in the article (*datapapers* and *datasets*) were structured in a .csv type file available in the *Figshare* repository under the DOI (<https://doi.org/10.6084/m9.figshare.19790425.v3>).

Conclusions: The research data was organized based on the *Dublin Core* metadata standards, in addition to authorial metadata created to meet the research objective. In textual and numerical format, they consist of secondary data extracted from the interfaces of the cited sources on the field of Nursing.

KEYWORDS: Publication of research data. Nursing. Search data. Data article. Research Data set.

1 APRESENTAÇÃO

Os dados apresentados e descritos neste artigo foram produzidos a partir da pesquisa de doutorado de Pinto (2022) realizada no âmbito do Programa de Pós-Graduação em Ciência da Informação do Instituto Brasileiro de Informação em Ciência e Tecnologia (PPGCI/IBICT), em convênio com a Universidade Federal do Rio de Janeiro (UFRJ), desenvolvida entre os anos de 2018 à 2022.

O cenário de produção e recuperação dos dados da pesquisa situa-se no campo da *e-science*, onde novas práticas científicas se estabelecem e, ao mesmo tempo, se fundem

às práticas já sedimentadas, aprimorando, oportunizando e ampliando novas relações e redes de trabalhos colaborativos em ambientes digitais.

A coleta dos dados da pesquisa nas bases e repositórios explorados foi conduzida pela seguinte questão norteadora: “Os dados produzidos em pesquisas da área da Enfermagem estão sendo publicados em *datapapers* e/ou *datasets*?”

De caráter exploratório, a pesquisa buscou identificar e recolher dados que revelassem a prática de publicação de dados de pesquisa pela área da Enfermagem no contexto da Ciência Aberta. Os dados produzidos demonstram-se valiosos para a comunidade científica, na medida em que apresentam mais um formato de publicação e comunicação entre pares, além de ampliar as possibilidades de interlocução, bem como a transparência dos dados, promoção de novas pesquisas através do reuso de dados e colaboração em pesquisas.

A produção dos dados está amparada pela relevância e potencial que a prática pode exercer na engrenagem da ciência como um todo. A descrição do conjunto de dados via *datapaper* bem como sua disponibilização em repositórios (*datasets*) são ações distintas que viabilizam a transparência dos processos e métodos científicos, e que, conseqüentemente, potencializa e fomenta o avanço da ciência. Configuram-se como uma prática recente que contribui para o avanço da gestão de dados de pesquisa, além de fortalecer o movimento da Ciência Aberta de maneira geral.

A importância e utilidade dos dados produzidos pelo estudo para a comunicação e comunidade científica devem-se a possibilidade de sua reutilização em novas pesquisas, uma vez que estes, estão depositados em repositório de dados e disponíveis para reutilização, compactuando com uma das prerrogativas do movimento da Ciência Aberta.

O conjunto de dados da pesquisa poderá beneficiar especialmente pesquisas no campo da Enfermagem, na medida em que revelam dados sobre o comportamento da comunicação entre pares na área investigada. Os dados encontram-se em formato aberto, acessíveis e passíveis de reutilização.

O modelo de curadoria dos dados presente no conjunto de dados produzido pode servir de base para outras investigações de mesma natureza em outras áreas do conhecimento.

Novos vieses de pesquisa são oportunizados a partir da publicação de dados, seja através de sua descrição em artigos e/ou sua disponibilização em repositórios. O aumento da visibilidade científica, ampliação da capacidade de respostas mais rápidas à sociedade, conforme foi observado na pandemia do COVID-19, otimização de tempo gasto em alguma

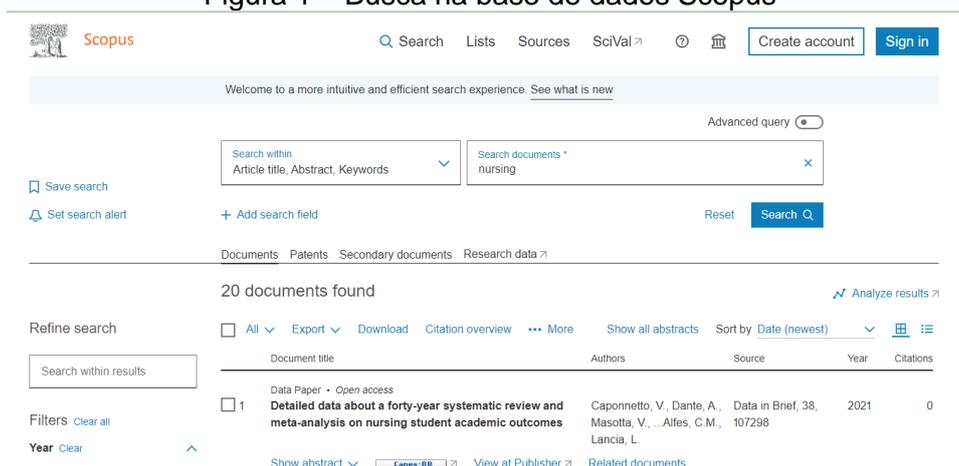
das etapas da pesquisa, colaboração, compartilhamento e reuso de dados entre os pares, estímulo à inovação e à descoberta científica, possibilidade de melhoria das práticas, métodos investigativos e experimentais de pesquisa, dentre outros.

1 MÉTODOS E INSTRUMENTOS

A metodologia utilizada para a localização e extração dos dados para a produção do conjunto de dados da pesquisa consistiu na busca direta na interface das bases de dados e repositórios de dados via Portal de Periódicos da Capes: *Web of science*, *Science direct*, *Scopus*, *Lilacs*, *Medline* e *Cinahl*. Utilizou-se também como fonte para extração de dados os repositórios de dados de pesquisa multidisciplinares de acesso livre: *Figshare*, *Dryade* e *Zenodo*. A partir de buscas sistematizadas diretamente nas interfaces das fontes citadas, estruturou-se os dados sobre as tipologias documentais apontadas no artigo (*datapapers* e *datasets*) em arquivo tipo .csv, respectivamente: *Dataper_Enfermagem.csv* e *Datasets_Enfermagem.csv*, que estão disponibilizados no repositório *Figshare* (<https://figshare.com/>). A curadoria e organização dos dados da pesquisa foram planejadas com base no padrão de metadados *Dublin Core*, além de metadados autorais criados para atender o objetivo da pesquisa. Em formato textual e numérico, consistem em dados secundários extraídos diretamente das interfaces das fontes citadas sobre a área da Enfermagem.

Para identificação e seleção de dados sobre publicação de dados via *data paper*, utilizamos as bases de dados: *Web of science*, *Science direct*, *Scopus*, *Lilacs*, *Medline* e *Cinahl*. O critério aplicado para seleção dos dados considerou os seguintes filtros, respectivamente: “*Article title, Abstract, Keyword*”: *Nursing*, associado ao filtro “*Document Type*”: *data paper*, sem delimitação de data, em todas as bases eleitas pelo estudo, conforme exemplificado pela Figura 1:

Figura 1 – Busca na base de dados Scopus



Fonte: Autora, 2023.

O mesmo método para seleção dos dados foi aplicado nas demais bases de dados exploradas. Para extração dos dados utilizou-se a ferramenta de exportação de dados disponibilizado pela própria interface. Os dados foram estruturados em planilha .csv e disponibilizado no repositório *Figshare* sob o DOI <https://doi.org/10.6084/m9.figshare.19790425.v4>.

Na tabela 1 apresentamos o resumo dos dados extraídos para a identificação do tipo de publicação *data paper* nas bases de dados pesquisadas no período de 08/11/2021.

Tabela 1 - Publicação de dados de pesquisa de Enfermagem em publicações periódicas via *data paper*¹

Base de dados	Data paper	Período
Web of Science	75	2015-2021
Science Direct	86	2015-2021
Scopus	16	2015-2021
Lilacs	0	-
Medline	0	-
Cinahl	0	-
Total	177	2015-2021

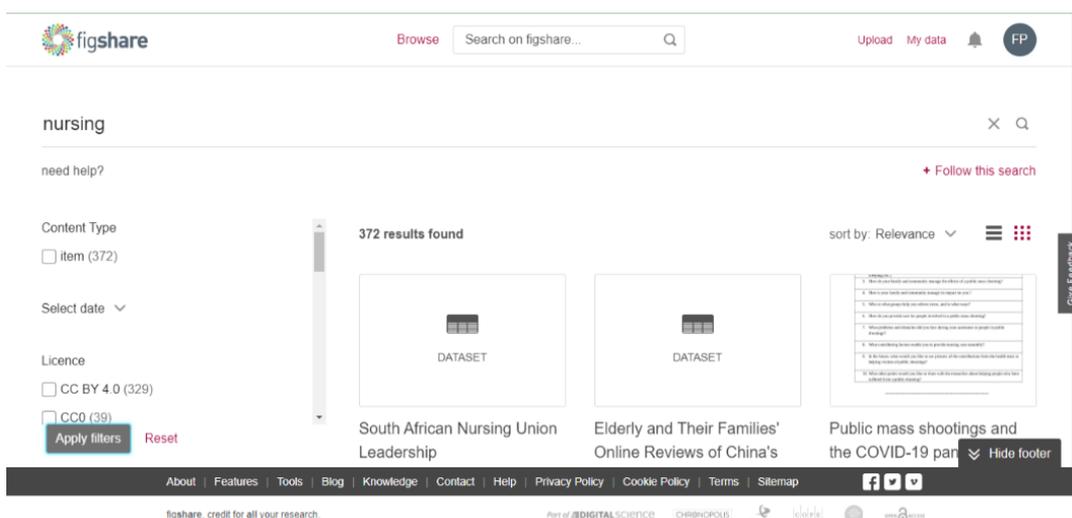
Fonte: Autora, 2021

O recorte temporal considerado na coleta de dados e representado na Tabela 1, de 2015 a 2021, refere-se à totalidade dos *datasets* e *data papers* da área de Enfermagem disponíveis nas fontes coletadas.

¹ Pesquisa rodada em 08/11/2021.

Para identificação e extração de dados via repositório utilizamos as seguintes fontes: *Figshare*, *Dryade* e *Zenodo*. O critério aplicado para seleção dos dados considerou os seguintes filtros, respectivamente: “Assunto”: *Nursing*, (no campo de busca principal) associado ao filtro “*Item type*”: *dataset*, sem delimitação de data, em todos os repositórios verificados pelo estudo, conforme exemplificado pela Figura 2:

Figura 2 – Busca no repositório de dados *Figshare*



Fonte: Autora, 2023.

O mesmo método de extração dos dados foi aplicado aos repositórios. Utilizou-se a ferramenta de exportação de dados disponibilizado pela própria interface. Os dados foram estruturados em planilha .csv e disponibilizado no repositório *Figshare* sob o DOI <https://doi.org/10.6084/m9.figshare.19790425.v4>.

Na tabela 2 apresentamos o resumo dos dados extraídos das interfaces dos repositórios explorados na pesquisa para a identificação do tipo de publicação *dataset* no período de 08/11/2021.

Tabela 2– Disponibilização de dados de pesquisa de Enfermagem em repositórios de dados via *dataset*

Repositórios de dados	Datasets	Período de depósito
Figshare	224	2013-2021
Dryad	134	2017-2021
Zenodo	103	2017-2021
Total	461	2013-2021

Fonte: Autora, 2021.

O recorte temporal considerado na coleta de dados em repositórios e representado na Tabela 2, de 2013 a 2021, refere-se à totalidade dos *datasets* da área de Enfermagem disponíveis nas fontes exploradas. Não houve um recorte temporal pré-estabelecido, tendo em vista a juventude deste tipo de publicação.

A pesquisa na íntegra que originou os dados encontra-se no repositório institucional da Universidade Federal Fluminense - RIUFF, no endereço eletrônico: <https://app.uff.br/riuff/handle/1/26650>.

2 TABELA DE ESPECIFICAÇÕES

Tabela 3 – Tabela de especificações

Área de Conhecimento	B. <i>Information use and sociology of information</i>
Área de assunto específica	Enfermagem
Idioma	Português
Tipo de dados	Número Texto
Como os dados foram adquiridos	Os dados foram extraídos diretamente das interfaces de busca das bases e repositórios de dados. Não foi utilizada ferramenta de coleta ou software para extração de dados.
Estado dos dados	Brutos
Parâmetros para coleta de dados	Foram considerados como parâmetros para a coleta de dados as seguintes tipologias documentais: conjuntos de dados (<i>dataset</i>) e artigos de dados (<i>data paper</i>) produzidos em pesquisas na área da Enfermagem.
Descrição da coleta de dados	<p>Para identificação e seleção de dados sobre publicação de dados <i>via data paper</i>, utilizamos as bases de dados: <i>Web of science</i>, <i>Science direct</i>, <i>Scopus</i>, <i>Lilacs</i>, <i>Medline</i>, e <i>Cinahl</i>, considerando os seguintes filtros, respectivamente: “<i>Article title</i>, <i>Abstract</i>, <i>Keyword</i>”: <i>Nursing</i>, associado ao filtro “<i>Document Type</i>”: <i>data paper</i>, sem delimitação de data;</p> <p>Para identificação e extração de dados <i>via dataset</i>, utilizamos os repositórios: <i>Figshare</i>, <i>Dryade</i> e <i>Zenodo</i>. O critério aplicado para seleção dos dados considerou os seguintes filtros, respectivamente: “<i>Assunto</i>”: <i>Nursing</i>, (no campo de busca principal) associado ao filtro “<i>Item type</i>”: <i>dataset</i>, sem delimitação de data.</p> <p>Os resultados das buscas foram extraídos via ferramenta de exportação das próprias interfaces de busca em formato .csv e deram</p>

	origem ao conjunto de dados da pesquisa (https://doi.org/10.6084/m9.figshare.19790425.v4).
Localização da fonte de dados	Bases de dados via Portal Capes (https://www.periodicos.capes.gov.br): <i>Web of Science</i> (https://www-webofscience.ez24.periodicos.capes.gov.br/wos/woscc/basic-search), <i>Science Direct</i> (https://www-sciencedirect.ez24.periodicos.capes.gov.br/), <i>Scopus</i> (https://www-scopus.ez24.periodicos.capes.gov.br/search/form.uri?display=basic), <i>Lilacs</i> (https://lilacs.bvsalud.org/), <i>Medline</i> (https://pubmed.ncbi.nlm.nih.gov/) e <i>Cinahl</i> (https://web-s-ebscohost.ez24.periodicos.capes.gov.br/ehost/search/basic?vid=1&sid=f825920a-6355-4133-a698-5f7978e41a38%40redis). Repositórios de dados: <i>Figshare</i> , <i>Dryade</i> e <i>Zenodo</i> .
Acessibilidade de dados	Nome do repositório: [FIGSHARE] Número de identificação de dados: https://doi.org/10.6084/m9.figshare.19790425.v4 . Endereço: https://figshare.com/

2.1 Descrição do conjunto de dados

Os conjuntos de dados produzidos pela pesquisa encontram-se disponibilizados no repositório de dados *Figshare* (<https://figshare.com/>), DOI <https://doi.org/10.6084/m9.figshare.19790425.v4>, sob o título: Publicação de dados de pesquisa_Enfermagem. Os dados foram organizados em dois arquivos no formato .csv, nomeados respectivamente como: *Datasets_Enfermagem.csv* e *Datapaper_Enfermagem.csv*.

Ambos os arquivos foram estruturados a partir do padrão de metadados Dublin Core² para descrever e representar os dados produzidos pela pesquisa. Foram necessários além destes, criar alguns metadados específicos para descrever dados que não foram adequadamente identificados pelo padrão utilizado, conforme representação nos quadros a seguir:

Quadro 1 - Metadados utilizados no conjunto de dados: *Datasets_Enfermagem*

DC - Dublin Core	Descrição / Equivalência
dc.description.source	Nome do Repositório de dados
dc.creator	Autor(s)
dc.title	Título
dc.language	Idioma
dc.subject keyword	Palavra-chave
dc.description.abstract	Resumo
dc.date.created	Data de publicação
dc.year.created	Ano de publicação
dc.identifier.doi	DOI

² <http://dublincore.org>

dc.right	Licença
dc.format	Tipo de arquivo

Fonte: Autora, 2021.

A curadoria utilizada para o gerenciamento dos dados da pesquisa do arquivo *Datasets_Enfermagem.csv* considerou os seguintes metadados: *dc.description.source* (nome do repositório onde o *dataset* foi localizado); *dc.creator* (nome do(s) *autor(s)* do *dataset*); *dc.title* (nome atribuído ao *dataset*); *dc.language* (idioma dos dados que compõem o *dataset*); *dc.subject.keywords* (assuntos atribuídos pelo autor para representar a temática dos dados); *dc.description.abstract* (resumo produzido pelo autor para apresentar o contexto de produção do *dataset*); *visualization* (quantidade de vezes que o *dataset* foi visualizado no repositório); *downloads* (número de vezes que o *dataset* foi “baixado” do repositório); *cited reference* (número de vezes que o conjunto de dados foi citado a partir do repositório³); *dc.date.created* (mês do depósito); *dc.year.created* (ano do depósito); DOI (Identificador persistente do *dataset*); *open access designation* (indica o tipo de acesso: *open*, *closed*, *restrict or embargoed* para o *dataset*); *dc.right* (tipo de licença atribuída ao *dataset*); outros repositórios depositados (identificação do depósito do *dataset* em outro repositório); Indexador (base indexadora do repositório onde o *dataset* foi depositado); *Altimetric* (serviço de métricas alternativas para identificação de citação do *dataset* em blogs, redes sociais, entre outras fontes); *dc.format* (extensão do arquivo); Arquivo complementar (conteúdo complementar ao *dataset* que explica e/ou descreve a natureza ou condição dos dados).

A seguir, através do Quadro 2, foram apresentados os metadados autorais criados para descrever os aspectos dos dados que não foram contemplados pelo *Dublin Core*.

Quadro 2 – Metadados autorais criados pela pesquisa para o conjunto de dados:
Datasets_Enfermagem

Metadados criados	Descrição
<i>Visualization</i>	Número de visualização do arquivo
<i>Download</i>	Número de downloads do arquivo
<i>Cited reference</i>	Número de citação do arquivo
<i>Open Access Designation</i>	Tipo de acesso ao arquivo
Volume	Volume do periódico
<i>Start page</i>	Página inicial do artigo
<i>End page</i>	Página final do artigo
<i>Outros repositórios depositados</i>	Mais de um depósito
<i>Indexador</i>	Base indexadora
<i>Altimetric</i>	Métrica alternativa
Arquivo complementar	Arquivo explicativo de dados
<i>Repository</i>	Nome do repositório

³ Medição realizada pela plataforma Dimension (<https://www.dimensions.ai/>).

Fonte: Autora, 2021.

Para a curadoria de dados aplicada ao arquivo *Datapaper_Enfermagem.csv* considerou-se os seguintes metadados representados através do Quadro 3:

Quadro 3 - Metadados utilizados no conjunto de dados: *Datapaper_Enfermagem*

Dublin Core	Descrição/Equivalência
<i>dc.description.source</i>	Nome da Base de dados (Fonte)
<i>dc.creator</i>	Autor(s)
<i>dc.creator.full name</i>	Sobrenome do autor
<i>dc.title</i>	Título do artigo
<i>dc.source title</i>	Título do periódico
<i>dc.language</i>	Idioma
<i>dc.subject keyword</i>	Palavra-chave do autor
<i>dc.description.abstract</i>	Resumo
<i>dc.date.created</i>	Data de publicação
<i>dc.date.year</i>	Ano de publicação
<i>dc.publisher</i>	Editora
<i>dc.identifier.doi</i>	DOI

Fonte: Autora, 2021

Onde, *dc.description.source* descreve o nome da base de dados onde o artigo foi localizado; *dc.creator* descreve o primeiro nome do(s) autor(s) do *datapaper*; *dc.creator.full name* refere-se ao sobrenome dos autores do *datapaper*; *dc.title* descreve o título do *datapaper*; *dc.source title* (título do periódico); *dc.language* (idioma do *datapaper*); *dc.subject keyword* (assuntos/palavra-chave atribuídos pelo autor para representar a temática do *datapaper*); *dc.description.abstract* (resumo produzido pelo autor para apresentar o contexto de produção do *datapaper*); *cited reference* (número de vezes que o conjunto de dados foi citado a partir da base de dados); *dc.publisher* (Editora); *dc.date.created* (data da publicação); *dc.date.year* (ano de publicação); volume (indicação de volume do periódico); *start page* (página inicial do artigo); *end page* (página final do artigo); DOI (Identificador persistente do *datapaper*); *Open Access Designation* (indica o tipo de acesso: *open*, *closed*, *restrict or embargoed* para o *datapaper*).

REFERÊNCIAS

CINAHL. Massachusetts: EBSCO, 2021. Disponível em: <https://web-p-ebscobhost.ez24.periodicos.capes.gov.br/ehost/search/basic?vid=1&sid=b61a978e-a73d-4d66-b155-2d21d3f3c330%40redis>. Acesso em: 08 nov. 2021.

DRYAD. Dryad. [S. l.], 2021. Digital Repository. Disponível em: <https://datadryad.org/stash>. Acesso em: 06 maio 2021.

DUBLIN CORE METADATA INITIATIVE (DCMI). 2016. Disponível em: <http://dublincore.org/>. Acesso em: 19 fev. 2019.



FIGSHARE. Figshare. [S. I.], 2021. Digital Repository. Disponível em: <https://figshare.com>. Acesso em: 06 maio 2021.

LILACS. São Paulo: BIREME, 2021. Disponível em: <https://lilacs.bvsalud.org/>. Acesso em: 08 nov. 2021.

MEDLINE. Bethesda: National Library of Medicine, 2021. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/>. Acesso em: 08 nov. 2021.

PINTO, Fabiana de Melo Amaral Gonçalves. **Publicação de dados de pesquisa em Enfermagem**: novas práticas em Ciência Aberta. 2022. Tese (Doutorado em Ciência da Informação) - Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022. Disponível em: <https://app.uff.br/riuff/handle/1/26650>. Acesso em: 16 dez. 2022.

PINTO, Fabiana de Melo Amaral Gonçalves. **Publicação de dados de pesquisa_Enfermagem**. Figshare. [dataset], 2022. Disponível em: <https://doi.org/10.6084/m9.figshare.19790425.v3>. Acesso em: 28 maio 2023.

PINTO, Fabiana de Melo Amaral Gonçalves. **Plano de gestão de dados de pesquisa (PDG)_Dados de pesquisa_Enfermagem**. Figshare. [dataset], 2023. Disponível em: <https://doi.org/10.6084/m9.figshare.23098667.v1> Acesso em: 28 maio 2023.

SCIENCE DIRECT. Amsterdã: Elsevier, 2021. Disponível em: <https://www-sciencedirect.ez24.periodicos.capes.gov.br/>. Acesso em: 08 nov. 2021.

SCOPUS. Amsterdã: Elsevier, 2021. Disponível em: <https://www-scopus.ez24.periodicos.capes.gov.br/search/form.uri?display=basic#basic>. Acesso em: 08 nov. 2021.

WEB OF SCIENCE. Londres: Clarivate Analytics, 2021. Disponível em: <https://www-webofscience.ez24.periodicos.capes.gov.br/wos/woscc/basic-search>. Acesso em: 08 nov. 2021.

ZENODO. Zenodo. [S. I.], 2021. Digital Repository. Disponível em: <https://zenodo.org>. Acesso em: 06 maio 2021.

NOTAS

CONTRIBUIÇÃO DE AUTORIA

Concepção e elaboração do manuscrito: F. M. A. G. Pinto

Coleta de dados: F. M. A. G. Pinto

Análise de dados: F. M. A. G. Pinto

Discussão dos resultados: F. M. A. G. Pinto

Revisão e aprovação: F. M. A. G. Pinto

CONJUNTO DE DADOS DE PESQUISA

FINANCIAMENTO

Não se aplica.

CONSENTIMENTO DE USO DE IMAGEM

Não se aplica.

APROVAÇÃO DE COMITÊ DE ÉTICA EM PESQUISA

Não se aplica.



CONFLITO DE INTERESSES

Não se aplica.

LICENÇA DE USO – uso exclusivo da revista

Os autores cedem à **Encontros Bibli** os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a [Licença Creative Commons Attribution](#) (CC BY) 4.0 International. Esta licença permite que **terceiros** remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico. Os **autores** têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

PUBLISHER

Universidade Federal de Santa Catarina. Programa de Pós-graduação em Ciência da Informação. Publicação no [Portal de Periódicos UFSC](#). As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

EDITORES

Edgar Bisset Alvarez, Ana Clara Cândido, Patrícia Neubert, Genilson Geraldo, Jônatas Edison da Silva, Mayara Madeira Trevisol.

HISTÓRICO

Recebido em: 08-03-2023 – Aprovado em: 02-08-2023 – Publicado em: 28-09-2023

