26º ENANCIB
Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação
A Amazônia na sociedade global em rede: assimetrias e potencialidades informacionais, tecnológicas e ambientais
Belém, Pará, 09 a 13 de novembro de 2026
Universidade Federal do Pará

ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
ISSN 2177-3688

GT 7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação

DADOS DE PESQUISA EM ENFERMAGEM: PRIMEIRAS ANÁLISES TAXONÔMICAS

NURSING RESEARCH DATA: *INITIAL TAXONOMIC ANALYSES*

Josimara Dias Brumatti – Universidade Federal Fluminense (UFF)
Paloma Rodrigues Moreira – Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Regiane Vanessa de Souza Baía – Universidade Federal Rural da Amazônia (UFRA)
Luana Farias Sales Marques – Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

Modalidade: Trabalho Completo

Resumo: Os dados de pesquisa em saúde apresentam elevada complexidade, diversidade e sensibilidade, demandando práticas estruturadas de gestão, curadoria e compartilhamento. Este estudo teve como objetivo analisar dados de pesquisa em Enfermagem disponíveis em repositórios nacionais, à luz da taxonomia proposta por Sales e Sayão (2019), e discutir suas implicações para a gestão de dados e a operacionalização dos princípios FAIR. Trata-se de uma pesquisa exploratória, descritiva, com abordagem quali-quantitativa. Os resultados indicaram predominância de dados observacionais e experimentais, majoritariamente em estágio final de processamento, organizados sobretudo em formatos de visualização e textual, e amplamente disponibilizados em acesso aberto. Conclui-se que a taxonomia constitui instrumento relevante para apoiar a gestão de dados e a implementação dos princípios FAIR, ao favorecer a identificação de requisitos de metadados, estratégias de preservação e condições de compartilhamento. Recomenda-se o desenvolvimento de uma taxonomia para cada área do conhecimento, com metadados mais específicos para domínios do conhecimento, maior autonomia do pesquisador na escolha de licenças e ações formativas sobre implicações éticas e jurídicas do reuso.
Palavras-chave: dados de pesquisa; repositório de dados; taxonomia; enfermagem; ciência aberta.

Abstract: Health research data is highly complex, diverse, and sensitive, requiring structured practices for management, curation, and sharing. This study aimed to analyze nursing research

data available in national repositories, based on the taxonomy proposed by Sales and Sayão (2019), and discuss its implications for data management and the implementation of the FAIR principles. This is an exploratory, descriptive study with a qualitative-quantitative approach. The results showed a predominance of observational and experimental data, mostly in the final stage of processing, organized mainly in visualization and textual formats, and widely available in open access. It is concluded that the taxonomy is a relevant tool to support data management and the implementation of the FAIR principles, as it favors the identification of metadata requirements, preservation strategies, and sharing conditions. It is recommended to develop a taxonomy for each area of knowledge, with more specific metadata for knowledge domains, greater researcher autonomy in choosing licenses, and training actions on the ethical and legal implications of reuse.
Keywords: research data; data repository; taxonomy; nursing; open science.

1 INTRODUÇÃO

Os dados de pesquisa nas ciências da saúde são essenciais para aprimorar diagnósticos, orientar tratamentos e prevenir doenças, de forma a construir uma medicina forte, na qual o paciente recebe um tratamento adequado e toda a sociedade se beneficia. Ao ter acesso a informações relevantes, é possível desenvolver vacinas e medicamentos que favorecem a vida.

A ciência aberta possibilita o acesso livre ao conhecimento científico, disponibilizando os dados abertos para a sociedade, segundo Albagli (2017), é um movimento que propõe aumentar a visibilidade, a produtividade e a circulação do conhecimento científico. Já os repositórios de dados de pesquisa são infraestruturas necessárias para que ocorra a gestão de dados. Sayão e Sales (2016) apontam que os repositórios de dados são infraestruturas desenvolvidas para dar suporte ao ciclo de gestão da informação, incluindo processos de curadoria de dados.

A Enfermagem brasileira tem desempenhado um papel importante na adoção das práticas FAIR e na promoção da Ciência Aberta, estudos apontam que a área de Enfermagem tem se destacado internacionalmente na discussão sobre gestão de dados científicos e interoperabilidade da informação, buscando alinhar suas práticas de pesquisa aos princípios FAIR, que orientam a organização, a abertura e o reuso de dados científicos (Alencar; Henning; Paula, 2021). Em 2020, a abertura do escritório Go FAIR Brasil Enfermagem reforçou esse protagonismo, indicando um movimento institucional de fortalecimento da gestão, da abertura e do reuso de dados de pesquisa. A criação da Rede GO FAIR Brasil Saúde Enfermagem promove o compartilhamento de dados de pesquisa produzidos na área, incentivando práticas de gestão de dados alinhadas aos princípios FAIR.

Apesar desses avanços, permanecem necessários estudos que analisem como os dados de pesquisa da área são disponibilizados e se atendem a padrões consistentes de descrição, documentação e preservação. “O reuso de dados em enfermagem ainda é pouco explorado, apesar de seu potencial para fortalecer a pesquisa, otimizar recursos e ampliar a visibilidade científica da área” (Guimarães *et al.*, 2025, p. 12).

Nesse sentido, questiona-se: como os dados de pesquisa em Enfermagem, disponíveis em repositórios nacionais, são classificados à luz da taxonomia proposta por Sales e Sayão (2019) e de que forma essa classificação pode apoiar a gestão de dados de pesquisa na área? Este estudo justifica-se pela possibilidade de contribuir para o fortalecimento das práticas FAIR e da gestão de dados no contexto brasileiro.

O objetivo é analisar os dados de pesquisa da área da saúde, utilizando a taxonomia para sua classificação. Pretende-se a) identificar repositórios institucionais da área da Saúde em âmbito nacional para fins de seleção dos dados de pesquisa; b) analisar a classificação dos dados de pesquisa; c) discutir as implicações dos resultados para a gestão de dados de pesquisa em Enfermagem.

A opção por analisar repositórios nacionais justifica-se pela necessidade de compreender como os dados de pesquisa em Enfermagem vêm sendo organizados e disponibilizados no contexto brasileiro, contribuindo para o fortalecimento das práticas de gestão e compartilhamento de dados. Além disso, a análise busca oferecer subsídios para iniciativas de promoção dos princípios FAIR, especialmente as desenvolvidas pelo GO FAIR Brasil Enfermagem, voltadas à adoção de padrões e boas práticas para dados de pesquisa na área.

2 CIÊNCIA ABERTA E PRINCÍPIOS FAIR

A Ciência Aberta (*Open Science*) busca tornar a pesquisa científica disponível a todos por meio de diversas abordagens, que incluem desde o acesso aberto às publicações, abertura de software e hardware, bem como compartilhamento de dados, códigos e outros objetos de pesquisa e ainda Ciência Cidadã. A presente pesquisa está focada no compartilhamento de dados, partindo do pressuposto de que tornar acessível o conhecimento científico possibilita a reutilização dos dados, maior colaboração entre a comunidade científica, acelerando a disseminação do conhecimento e os benefícios para a sociedade.

No entanto, para que o compartilhamento de dados seja efetivo e este pressuposto

seja verdadeiro, faz-se necessário que os dados compartilhados passem, antes de sua abertura, por um processo minucioso de gestão, que permita que possam ser encontráveis, acessíveis e interoperáveis para que finalmente sejam reusados na construção de novos conhecimentos. Assim, torna-se fundamental que os princípios FAIR de *Findable* (Localizáveis), *Accessible* (Acessíveis), *Interoperable* (Interoperáveis) e *Reusable* (Reutilizáveis), sejam considerados na gestão desses dados de pesquisa, tornando-as mais fáceis de descobrir, acessar, integrar e usar por máquinas e pessoas e, proporcionando o desenvolvimento da ciência aberta e a pesquisa reprodutível.

Os princípios FAIR podem ser definidos como:

> Um conjunto de princípios destinados a melhorar a utilidade dos dados, marcando um refinamento significativo dos conceitos essenciais para aumentar o valor dos dados e facilitar sua reutilização por seres humanos e máquinas. Isso promove uma interação mais eficiente entre os usuários interessados nos dados e aqueles que os disponibilizam (European Commission, 2018, não paginado).

A adoção dos princípios FAIR amplia a gestão dos dados de pesquisa para além do armazenamento, considerando também seu uso futuro. Segundo Wilkinson *et al*. (2016), esses princípios orientam a organização e a descrição dos dados para que sejam encontrados, acessados, compreendidos e reutilizados por pessoas e sistemas.

O GO FAIR (2018) promove a implantação de práticas e serviços pautados na adoção integral dos princípios FAIR, apresentam os Princípios FAIR e destacam características vinculadas a cada um deles, apresentando elementos a serem considerados na publicação de dados. No contexto da Ciência Aberta, os princípios FAIR reconhecem os dados de pesquisa como recursos que podem ser reutilizados em diferentes etapas e contextos. Segundo Henning, Ribeiro, Santos e Santos (2019), sua aplicação exige uma gestão padronizada e qualificada, com informações sobre a origem, o contexto, a estrutura e as condições de uso dos dados. Portanto, sua simples disponibilização não garante a reutilização.

Cabe destacar que os princípios FAIR não pressupõem necessariamente a abertura irrestrita dos dados, seu propósito é promover melhores práticas de gestão, descrição, preservação e compartilhamento de dados e outros objetos digitais de pesquisa. A consolidação dos princípios FAIR também tem ampliado a discussão sobre os mecanismos e serviços necessários para sua implementação. Nesse cenário, a aplicação dos princípios não se restringe à definição de características desejáveis para os dados, mas envolve práticas de gestão, infraestrutura, serviços e instrumentos capazes de apoiar sua adoção ao longo do ciclo

de vida dos dados de pesquisa (Veiga; Henning; Príncipe; Santos, 2025).

Nesse contexto, o princípio *Accessible* refere-se à possibilidade de acesso por meio de protocolos e condições claramente definidas, e não necessariamente ao acesso aberto.

2.1 REPOSITÓRIO DE DADOS DE PESQUISA

Uma das formas de armazenar e disseminar dados de pesquisa se dá por meio da utilização de repositórios de dados de pesquisa, que aumentam a visibilidade e consolidam a partilha de dados em diversos formatos (Sanchez; Vechiato, 2019).

De acordo com Sayão e Sales (2016), para apoiar o desempenho dos processos de gestão, é importante um arcabouço tecnológico e gerencial que aborde, em toda a sua completude, o ciclo de vida dos dados. Na parte central desse arcabouço estão os repositórios digitais de dados de pesquisa que, por muitas razões e demandas, de forma veloz, tornam-se parte essencial da infraestrutura mundial de pesquisa.

Os pesquisadores precisam de repositórios para depositar os dados de pesquisa e, por conta disso, as bibliotecas universitárias e os centros de pesquisa começaram uma avaliação de forma a realizar a qualificação de repositórios confiáveis (Curty, 2018). Essa avaliação é importante, visto que aborda a confiabilidade do repositório de dados e suas qualificações. Na seção que se segue, serão abordados os dados de pesquisa no domínio da enfermagem.

2.2 DADOS DE PESQUISA NO DOMÍNIO DA ENFERMAGEM

A Enfermagem produz diversos dados que, muitas vezes, são considerados sensíveis por envolverem informações pessoais de patologias em indivíduos. Contudo, esses dados podem contribuir de forma significativa para que pesquisadores os utilizem no aprimoramento de seus estudos e, consequentemente, no avanço da pesquisa científica na área da saúde.

Brumatti *et al.* (2025) afirmam, em seu estudo, que 95% dos dados de pesquisa da área de Enfermagem publicados em repositórios de dados encontram-se em acesso aberto. Já em relação aos dados de pesquisa oriundos da publicação de periódicos, o acesso ainda é fechado. Segundo os autores, dados abertos em enfermagem representam não apenas uma oportunidade de ampliar a transparência e a colaboração científica, mas também um campo de disputa política, econômica e ética.

Diante da interdisciplinaridade, da natureza heterogênea e aplicações dos dados, Sales

e Sayão (2019), partindo dessa premissa, observaram a necessidade de caracterização dos dados de pesquisa, uma vez que há uma inexistência de um instrumento terminológico e classificatório que visa criar uma taxonomia para classificar os tipos de dados através de uma abordagem teórico-metodológica da organização do conhecimento.

No âmbito da Ciência da Informação, a taxonomia tem como função a organização de conteúdo, alguns princípios básicos de categorização, além de cânones e princípios originados de teorias como a formulada pelo matemático indiano Shiyali Ramamrita Ranganathan (1937). O processo elaborativo de uma taxonomia se utiliza das seguintes etapas: captura do conhecimento, análise dos documentos e informações, elaboração da estrutura classificatória e validação (Campos; Gomes, 2007). Dentro desses princípios, destaca-se a Categorização que busca analisar o domínio conceitual, de forma dedutiva, formando uma cadeia sistemática (Campos; Gomes, 2007). Embora tradicionalmente, o papel da taxonomia seja ordenar sistematicamente conteúdos, Sales e Sayão (2019) propõem uma taxonomia cuja função foi não apenas reunir as formas de classificação de dados de pesquisa encontrada na literatura, mas também ajudar na caracterização de dados gerados em um domínio para fins de tomada de decisão em relação à própria gestão de dados, bem como para a construção de políticas e infraestruturas adequadas ao tipo de dado gerado naquele domínio. Essa taxonomia de dados de pesquisa contempla oito características de divisão que permite que o dado seja analisado sob diversas facetas, a saber: quanto à origem, ao processamento, abordagem, natureza, sensibilidade, materialidade, perenidade e abertura dos dados.

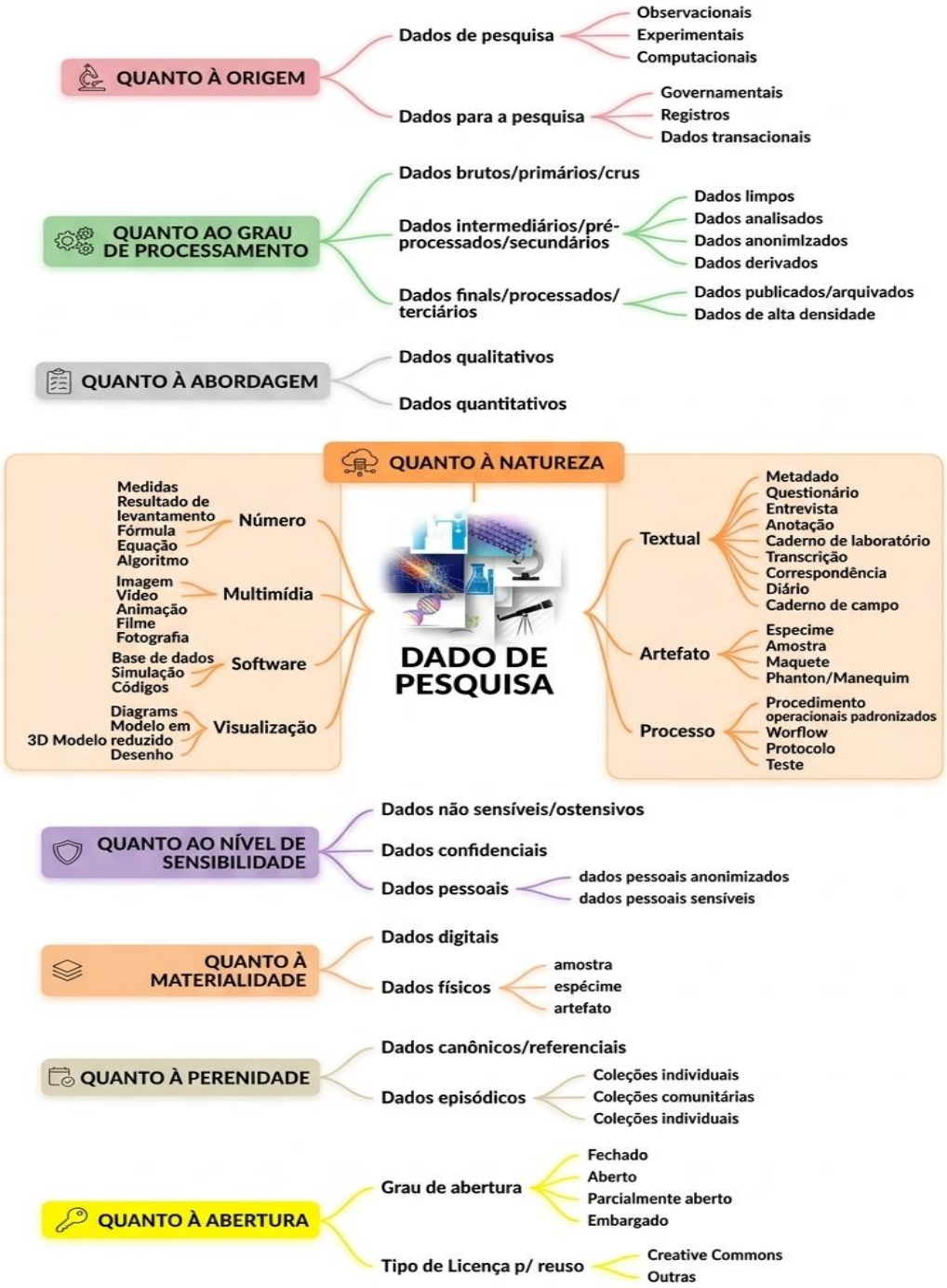
3 PROCEDIMENTOS METODOLÓGICOS

A pesquisa[1] constitui-se de estudo de caso de natureza quali-quantitativa de abordagem descritiva, por meio de uma população definida, pois identifica e classifica os conjuntos de dados de pesquisa em Enfermagem, a partir da taxonomia proposta por Sales e Sayão (2019) que classifica os dados segundo oito classes fundamentais para a gestão e a curadoria de dados de pesquisa, a figura 1 abaixo mostra 8 classes e suas subdivisões. O conjunto de dados analisados nesta pesquisa será depositado em um repositório institucional, garantindo armazenamento e preservação, assim como a confiabilidade na pesquisa e no método empregado.

[1] Para esta pesquisa, o Plano de Gestão de Dados na Plataforma FioDMP, que pode ser acessado através do link: https://fiodmp.fiocruz.br/S5H0NST.

A pesquisa[2] foi realizada em duas etapas: 1ª: identificação dos repositórios nacionais foi realizada por meio da plataforma *re3data* (*Registry of Research Data Repositories*). Na página inicial da plataforma, selecionou-se a opção "Navegar por país", optando-se pelo esquema de visualização em formato de texto. Em seguida, foi selecionado o país Brasil, resultando na recuperação de 23 repositórios.

Figura 1 - Taxonomia para dados de pesquisa



Fonte: elaboração própria (2026) baseado em Sales e Sayão (2019).

[2] A pesquisa foi realizada em dezembro de 2025.

Utilizou-se os seguintes critérios: tipo de repositório institucional, o que resultou em 14 registros, e software Dataverse, reduzindo o conjunto a 9 repositórios, a saber: Repositório de Dados de Pesquisa da Universidade de Campinas - Redu; Repositório de Dados de Pesquisa da Rede Cariniana; Repositório de Dados de Pesquisa do SciELO - SciELO Data; Repositório Comum de Dados de Pesquisa - Deposita Dados; Repositório de Dados de Pesquisa do Instituto Brasileiro de Informação em Ciência e Tecnologia - Aleia; Repositório de Dados de Pesquisa do Instituto Federal Goiano - Dados Abertos de Pesquisas; Repositório de Dados de Pesquisa da Universidade Federal de São Paulo - Domus Dados; Repositório de Dados de Pesquisa da Fundação Oswaldo Cruz - Arca Dados e Repositório de Dados de Pesquisa da Empresa Brasileira de Pesquisa Agropecuária - Recape.

A 2ª etapa da pesquisa, foi realizada a análise das classes, nos repositórios de dados selecionados foram analisados individualmente. A identificação dos conjuntos de dados de pesquisa foi realizada por meio do campo de busca de cada repositório, que recupera ocorrências dos termos nos metadados disponíveis. Foram utilizados os descritores “enfermagem”, “enfermería” e “nursing”, com o objetivo de delimitar uma população de análise vinculada ao domínio da Enfermagem. Dessa forma, o estudo concentrou-se nos conjuntos recuperados a partir desses descritores, não tendo como finalidade o mapeamento exaustivo da produção da área. Foram considerados para análise apenas os conjuntos de dados publicados no ano de 2025.

A análise gerou o seguinte critério de exclusão: *datasets* com mesmo título, indexados nos idiomas pesquisados.

4 RESULTADOS E DISCUSSÃO

Cada conjunto de dados identificado foi analisado individualmente e classificado nas oito categorias da taxonomia proposta por Sales e Sayão (2019): origem, grau de processamento, abordagem, natureza dos dados, nível de sensibilidade, materialidade, perenidade e abertura. Com vistas à transparência metodológica, reprodutibilidade e reuso dos resultados, os dados produzidos pela pesquisa serão depositados no repositório Aleia[3].

Esta pesquisa revelou os seguintes resultados: Dos 9 repositórios analisados, 4 não foram passíveis de análise por não conterem dados de pesquisa em Enfermagem. Dos cinco

[3] Disponível em: https://aleia.ibict.br/.

repositórios incluídos na pesquisa a saber: Redu, Deposita Dados, Arca Dados, SciELO Data e Domus, totalizou-se 80 conjuntos de dados. Após eliminar as duplicidades por idioma das palavras-chave, chegamos ao número final de 66 conjuntos de dados.

4.1 QUANTO À ORIGEM

Observou-se predominância de dados de pesquisa (69,7%), entendidos como aqueles gerados diretamente no contexto da investigação, enquanto 30,3% corresponderam a dados para pesquisa, originalmente produzidos com outras finalidades e posteriormente utilizados em estudos.

Entre os tipos de dados segundo a origem, os observacionais foram os mais frequentes (48,5%), seguidos pelos experimentais (31,8%), dados provenientes de registros (15,2%) e dados computacionais (4,5%).

A presença expressiva de dados observacionais é compatível com características da pesquisa em Enfermagem. Salmon (2015) destaca que a observação permite acompanhar atividades, comportamentos e acontecimentos nos contextos em que ocorrem, podendo ser empregada em pesquisas qualitativas ou quantitativas.

No conjunto analisado nesta pesquisa, esse resultado mostra a importância dos dados produzidos em situações concretas de cuidado, assistência e investigação. Para a gestão desses dados, a identificação de sua origem é relevante porque auxilia na definição de requisitos de documentação, preservação e descrição, conforme a proposta de Sales e Sayão (2019), que relaciona a caracterização dos dados às decisões sobre gestão, políticas e infraestrutura.

4.2 QUANTO AO GRAU DE PROCESSAMENTO

Predominaram os dados finais ou processados (65,2%), enquanto os dados brutos representaram 22,7% e os intermediários 12,1%.

A predominância de dados já processados chama atenção para a necessidade de documentação das etapas pelas quais esses dados passaram. A disponibilização de um conjunto final não permite, por si só, compreender as operações realizadas sobre os dados originais. Henning, Ribeiro, Santos e Santos (2019) ressaltam que o reuso depende da existência de informações sobre origem, contexto, estrutura e condições de uso, e não apenas da disponibilização do arquivo.

Nesse sentido, informações sobre limpeza, transformação, anonimização, exclusão de registros e demais procedimentos realizados durante o tratamento dos dados são importantes para sua interpretação e para a possibilidade de reprodução da pesquisa.

4.3 QUANTO À ABORDAGEM

Os dados provenientes de pesquisas qualitativas representaram 60,6% dos conjuntos analisados, enquanto 22,7% estavam relacionados a pesquisas quantitativas e 16,7% a estudos de abordagem mista.

A presença expressiva de dados qualitativos encontra apoio na própria tradição de pesquisa da Enfermagem. Oliveira, Presado e Baixinho (2025) destacam a importância da pesquisa qualitativa para compreender vivências, necessidades, comportamentos e contextos relacionados ao cuidado, elementos que nem sempre podem ser adequadamente apreendidos apenas por medidas quantitativas.

A predominância encontrada, portanto, ajuda a compreender outras características dos dados identificados no estudo, especialmente a presença de materiais textuais provenientes de entrevistas, questionários, relatos e registros.

4.4 QUANTO À NATUREZA DOS DADOS

Os dados classificados como de visualização foram os mais frequentes (53%), incluindo tabelas, diagramas, gráficos e modelos. Os dados textuais representaram 30,3%, seguidos pelos números (12.1%), multimídia (3,0%) e dados de processo (1,5%). Não foram identificados conjuntos classificados como software ou artefatos físicos.

A presença de dados textuais guarda relação com o predomínio das abordagens qualitativas encontradas na pesquisa. Entrevistas, narrativas, respostas abertas e registros são materiais frequentes em estudos que procuram compreender experiências e contextos de cuidado, como discutem Oliveira, Presado e Baixinho (2025).

Para o reuso, porém, o formato do arquivo não é suficiente. A compreensão desses dados depende de informações sobre o contexto em que foram produzidos, instrumentos utilizados, métodos de coleta e tratamento realizado. Essa preocupação se aproxima do princípio Reuso do FAIR, segundo o qual os dados devem ser suficientemente descritos para que possam ser compreendidos e utilizados posteriormente. O próprio referencial adotado no estudo ressalta que a organização e a descrição são condições importantes para que os

dados sejam encontrados, acessados, compreendidos e reutilizados.

4.5 QUANTO À PERENIDADE

A maior parte dos conjuntos foi classificada como de perenidade intermediária (97,0%), enquanto 3,0% foram classificados como episódicos.

Esse resultado indica que os dados analisados possuem valor para além de seu uso imediato na pesquisa que os originou, embora não se caracterizem como grandes bases referenciais de caráter permanente. Nesse ponto, a classificação proposta por Sales e Sayão (2019) é particularmente útil porque permite considerar a permanência dos dados como um dos elementos para orientar decisões sobre preservação.

Sayão e Sales (2016) destacam que a gestão dos dados deve considerar todo o seu ciclo de vida e que os repositórios fazem parte da infraestrutura necessária para apoiar esse processo. Assim, a definição de estratégias de preservação não deve considerar apenas o formato digital, mas também o valor científico, a possibilidade de reuso e o contexto em que os dados foram produzidos.

4.6 QUANTO AO NÍVEL DE SENSIBILIDADE

Os dados não sensíveis corresponderam a 45,5% dos conjuntos analisados. Os dados pessoais representaram 37,9% e os confidenciais 16,7%. A proporção de dados pessoais e confidenciais é compatível com uma área que trabalha diretamente com indivíduos, experiências de saúde e contextos assistenciais. Esse aspecto exige cuidados específicos de gestão, principalmente no que se refere à anonimização, às condições de acesso e às responsabilidades envolvidas no compartilhamento.

Brumatti *et al.* (2025) chamam atenção para o fato de que a abertura de dados em Enfermagem envolve não apenas questões científicas, mas também aspectos éticos, econômicos e políticos. No mesmo sentido, os princípios FAIR não determinam que todos os dados sejam disponibilizados de forma irrestrita. O acesso pode ocorrer sob condições previamente estabelecidas, especialmente quando estão envolvidos dados pessoais ou sensíveis.

O resultado reforça, portanto, a necessidade de equilibrar compartilhamento e proteção, de acordo com a orientação de tornar os dados tão abertos quanto possível e tão fechados quanto necessário.

4.7 QUANTO À MATERIALIDADE

Os dados digitais representaram 78,8% dos conjuntos analisados, enquanto 21,2% estavam associados a elementos de natureza física. A predominância do formato digital é esperada em repositórios de dados, cuja função envolve armazenamento, preservação e compartilhamento de objetos digitais. Entretanto, a presença de elementos físicos mostra que alguns conjuntos estão relacionados a amostras, materiais coletados, entrevistas ou outros objetos que precisam ser representados digitalmente por meio de metadados, descrições, transcrições ou digitalização.

Esse resultado reforça a importância dos repositórios não apenas como locais de armazenamento, mas como parte de uma infraestrutura de gestão capaz de representar objetos de natureza diversa. Sayão e Sales (2016) situam os repositórios justamente como componentes centrais da infraestrutura de gestão dos dados de pesquisa.

4.8 QUANTO AO GRAU DE ABERTURA

Os conjuntos classificados como abertos representaram 87,9% do total analisado. Os dados fechados corresponderam a 7,6%, os parcialmente abertos a 3,0% e os embargados a 1,5%. O percentual encontrado se aproxima do observado por Brumatti *et al.* (2025), que identificaram elevada presença de dados de pesquisa em Enfermagem disponíveis em acesso aberto nos repositórios analisados pelos autores. Embora os universos das duas pesquisas sejam distintos, os resultados indicam presença significativa de práticas de compartilhamento aberto na área.

A existência de dados fechados, parcialmente abertos ou embargados não deve, contudo, ser interpretada automaticamente como ausência de conformidade com os princípios FAIR. Conforme discutido no referencial deste estudo, o princípio “Acessível” está relacionado à existência de condições claras de acesso, e não à obrigatoriedade de abertura irrestrita.

Quanto ao licenciamento, todos os conjuntos utilizavam licenças *Creative Commons*. A CC BY 4.0 foi predominante, com 68,2%, seguida pela CC BY-NC 4.0, com 30,3%, e pela CC BY-ND 4.0, com 1,5%. Essas diferenças mostram que o acesso aberto não elimina as condições estabelecidas para o reuso. A escolha da licença define possibilidades distintas de utilização, sobretudo em relação ao uso comercial e à produção de obras derivadas. A adoção de licenças *Creative Commons* estabelece condições para o compartilhamento e o reuso dos conjuntos

de dados, mas não substitui os requisitos éticos e legais envolvidos na pesquisa. No caso de dados oriundos de pesquisas com seres humanos, mesmo quando submetidos a processos de anonimização, o licenciamento deve ser considerado em conjunto com as condições de consentimento, proteção dos participantes e legislação aplicável, incluindo a Lei Geral de Proteção de Dados Pessoais (LGPD). Assim, a existência de uma licença aberta não significa, por si só, que os dados possam ser reutilizados sem a observância dessas condições.

A documentação também apresentou diferenças entre os conjuntos. Em 62,1% dos casos havia arquivo “Leia-me”, enquanto 37,9% não disponibilizavam esse tipo de documento. A presença do Leia-me foi mais frequente no SciELO Data, onde existem orientações específicas de depósito e revisão curatorial.

A ausência desse tipo de documentação pode limitar o reuso, mesmo quando os arquivos estão disponíveis em acesso aberto. Henning, Ribeiro, Santos e Santos (2019) chamam atenção justamente para a necessidade de registrar informações sobre origem, contexto, estrutura e condições de uso dos dados. Assim, documentação e abertura devem ser vistas como aspectos complementares: disponibilizar o arquivo amplia o acesso, mas é a documentação que permite compreender melhor o que está sendo reutilizado.

5 CONSIDERAÇÕES FINAIS

A aplicação da taxonomia de Sales e Sayão (2019) aos dados de pesquisa em Enfermagem permitiu identificar características recorrentes nos conjuntos analisados, com predominância de dados observacionais e experimentais, de abordagens qualitativas, organizados principalmente em formatos de visualização e textuais e, em sua maioria, disponibilizados em acesso aberto.

A presença de dados pessoais e confidenciais reforça a necessidade de cuidados específicos quanto à anonimização, às condições de acesso e ao compartilhamento. Nesses casos, a atribuição de uma licença de uso constitui apenas um dos elementos da gestão dos dados e deve estar articulada aos requisitos éticos, às condições estabelecidas no consentimento e às normas de proteção de dados aplicáveis. Esses resultados também mostram que a adoção dos princípios FAIR não depende necessariamente da abertura irrestrita dos dados, mas da existência de condições claras de acesso, descrição, documentação e reuso, aspecto especialmente importante em pesquisas que envolvem informações de saúde.

A classificação mostrou-se útil não apenas para descrever os conjuntos analisados, mas também para indicar necessidades relacionadas à gestão dos dados. A origem observacional, por exemplo, exige atenção à documentação do contexto em que os dados foram produzidos, aos instrumentos utilizados e à proveniência. A presença expressiva de dados qualitativos e textuais amplia a importância de metadados capazes de registrar métodos de coleta, características da população estudada, processos de tratamento e demais informações necessárias à interpretação futura dos dados. Já os dados pessoais e confidenciais demandam mecanismos que permitam diferentes níveis de acesso e o registro das condições de uso e compartilhamento.

Nesse sentido, a taxonomia pode apoiar decisões relacionadas à curadoria, preservação, descrição, licenciamento e configuração das infraestruturas de dados. Ao permitir reconhecer as características dos conjuntos produzidos em determinado domínio, a classificação oferece subsídios para que repositórios, instituições, pesquisadores e profissionais envolvidos na gestão de dados definam práticas mais adequadas às necessidades da área.

A análise também mostrou que a maior parte dos conjuntos encontra-se em acesso aberto e utiliza licenças *Creative Commons*, embora com diferenças quanto às possibilidades de reuso. Da mesma forma, a presença de arquivos leia-me em parte dos conjuntos indica avanços na documentação dos dados, mas sua ausência em parcela significativa da amostra mostra que ainda há espaço para aprimorar as informações necessárias à compreensão e reutilização dos conjuntos.

A taxonomia utilizada possui caráter geral e foi concebida para aplicação em diferentes áreas do conhecimento. Os resultados deste estudo sugerem que determinadas características da Enfermagem, especialmente a presença de dados qualitativos, observacionais, textuais e relacionados a indivíduos, podem demandar formas mais específicas de descrição. Estudos futuros podem avançar na definição de metadados e critérios de classificação mais próximos das particularidades da área, bem como na validação dessas categorias por pesquisadores e profissionais da Enfermagem.

Como limitação, reconhece-se que a classificação dos conjuntos envolve um componente interpretativo, ainda que os critérios tenham sido aplicados de forma padronizada com base na taxonomia adotada. O recorte da pesquisa, restrito aos conjuntos recuperados pelos termos definidos e publicados em 2025, também não permite representar

toda a produção de dados de pesquisa em Enfermagem disponível nos repositórios nacionais. A ampliação do período analisado, dos termos de busca e da participação de especialistas da área constitui uma possibilidade de continuidade deste estudo.

REFERÊNCIAS

ALBAGLI, Sarita. Ciência aberta como instrumento de democratização do saber. **Trabalho, Educação e Saúde**, Rio de Janeiro, v. 15, n. 3, p. 659-660, set./dez. 2017. Disponível em: https://www.scielo.br/j/tes/a/NjMd4mXhD43CKqXPcZKrmjj/. Acesso em: 12 mar. 2026.

ALENCAR, Maria Simone de Menezes; HENNING, Patricia; PAULA, Danielle Galdino de. A implementação da rede GO FAIR Brasil Saúde-Enfermagem. **Páginas a&b**: arquivos e bibliotecas, [*s. l.*], série 3, n. especial, p. 113-117, 2021. Disponível em: https://ojs.letras.up.pt/index.php/paginasaeb/article/view/10202. Acesso em: 12 mar. 2026.

BRUMATTI, Josimara Dias; BRUNONI, Jessica Silva; SANTOS, Marcia Rodrigues dos; ALENCAR, Maria Simone de Menezes *et al.* Dados de pesquisa abertos: panorama da Enfermagem a partir da base de dados OpenAlex. **Revista Ciência Aberta Lusófona**, [*s. l.*], v. 1, 2025. DOI: https://doi.org/10.82524/recal.2025.66. Disponível em: https://recal.rcaap.pt/recal/article/view/66. Acesso em: 7 jan. 2026.

CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Taxonomia e classificação: a categorização como princípio. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8., 2007, Salvador. **Anais** [...]. Salvador: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação, 2007. p. 1-14. Disponível em: https://enancib.ppgci.ufba.br/artigos/GT2--101.pdf. Acesso em: 26 jan. 2026.

CURTY, Renata. **Para onde os dados devem ir afinal?** Dados de pesquisa abertos. 2018. Disponível em: https://dadosdepesquisa.rnp.br/para-onde-os-dados-devem-ir-afinal. Acesso em: 8 jan. 2026.

EUROPEAN COMMISSION. **Communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions**. Europea Union: Bruxelas, 2018. Disponível em: https://shre.ink/3XhP. Acesso em: 5 jan. 2026.

GO FAIR INTERNATIONAL SUPPORT AND COORDINATION OFFICE. **GO FAIR initiative**. Leiden, 2018. Disponível em: https://www.go-fair.org/go-fair-initiative/. Acesso em: 18 jun. 2026.

GUIMARÃES, Clarissa Coelho Vieira; MIRANDA, Gabriela Cristina Oliveira de; CÂMARA, Lívia de Souza; BRUNONI, Jéssica Silva *et al.* Ciência aberta no âmbito da enfermagem: contribuições em estudos com reúso de dados. **Revista de Geopolítica**, São José dos Pinhais, v. 16, n. 4, e685, 2025. DOI: https://doi.org/10.56238/revgeov16n4-048.

Disponível em: https://revistageo.com.br/revgeo/article/view/685. Acesso em: 12 mar. 2026.

HENNING, Patricia Corrêa; RIBEIRO, Claudio José Silva; SANTOS, Luiz Olavo Bonino da Silva; SANTOS, Paula Xavier dos. GO FAIR e os princípios FAIR: o que representam para a expansão dos dados de pesquisa no âmbito da Ciência Aberta. **Em Questão**, Porto Alegre, v. 25, n. 2, p. 389-412, 2019. Disponível em: https://seer.ufrgs.br/index.php/EmQuestao/article/view/84753. Acesso em: 12 set. 2026.

OLIVEIRA, Ellen Synthia de; PRESADO, Maria Helena; BAIXINHO, Cristina Lavareda. The relevance of qualitative research in nursing: exploring experiences and context in care. **Revista Brasileira de Enfermagem**, [*s. l.*], v. 78, supl. 3, e78supl301, 2025. DOI: https://doi.org/10.1590/0034-7167.202578supl301. Disponível em: https://www.scielo.br/j/reben/a/P67RpB4XHgF8S9Hw3QfTDvQ/. Acesso em: set. 2026.

RANGANATHAN, Shiyali Ramamrita. **Prolegomena to library classification.** Madras: Madras Library Association, 1937.

SALES, Luana Farias; SAYÃO, Luís Fernando. Uma proposta de taxonomia para dados de pesquisa. **Conhecimento em Ação**, Rio de Janeiro, v. 4, n. 1, p. 31-48, jan./jun. 2019. DOI: https://doi.org/10.47681/rca.v4i1.26337. Disponível em: https://revistas.ufrj.br/index.php/rca/article/view/26337. Acesso em: 18 jan. 2026.

SALMON, Jenny. Using observational methods in nursing research. **Nursing Standard**, v. 29, n. 45, p. 36-41, 8 jul. 2015. DOI: https://doi.org/10.7748/ns.29.45.36.e8721. Disponível em: https://pubmed.ncbi.nlm.nih.gov/26153969/. Acesso em: set. 2026.

SANCHEZ, Fernanda Alves; VECHIATO, Fernando Luiz. Avaliação de repositórios de dados de pesquisa segundo critérios da encontrabilidade da informação. **Liinc em Revista**, Rio de Janeiro, v. 15, n. 2, p. 142-160, 2019. DOI: https://doi.org/10.18617/liinc.v15i2.4774. Disponível em: https://revista.ibict.br/liinc/article/view/4774. Acesso em: 6 jan. 2026.

SAYÃO, Luís Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, Londrina, v. 21, n. 2, p. 90-115, maio/ago. 2016. DOI: https://doi.org/10.5433/1981-8920.2016v21n2p90. Disponível em: https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/27939. Acesso em: 5 jan. 2026.

SCIELO. **Guia de depósito de dados de pesquisa**. Versão 1.1. [*S. l.*]: SciELO, maio 2025. Disponível em: https://wp.scielo.org/wp-content/uploads/Guia_deposito_pt.pdf. Acesso em: 15 jan. 2026.

VEIGA, Viviane; HENNING, Patricia; PRÍNCIPE, Pedro; SANTOS, Luiz Olavo Bonino da Silva. Um panorama dos princípios de dados FAIR: teoria, práticas e serviços. **Informação & Informação**, [*s. l.*], v. 29, n. 4, p. 267-282, 2025. Disponível em:

https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/51782. Acesso em: 12 set. 2026.

WILKINSON, Mark D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, [*s. l.*], v. 3, art. 160018, 2016. DOI: https://doi.org/10.1038/sdata.2016.18. Disponível em: https://www.nature.com/articles/sdata201618. Acesso em: 12 set. 2026.