



17 A 20 DE NOVEMBRO  
SÃO PAULO - SP

## Especialização de inteligência artificial para indexação automática em Ciências da Saúde: integração da ferramenta DeCS Finder IA com Annif

**Palavras-chave:** Indexação e Redação de Resumos. Algoritmos de Classificação. Processamento de Linguagem Natural. Aprendizado de Máquina. Vocabulário Controlado.





**Abstract:** The lack of specialized solutions for automatic indexing in health motivated this study, which reports the integration of the DeCS Finder IA tool with Annif to improve automatic indexing in health, based on the LILACS Methodology and the DeCS controlled vocabulary. Models were trained using 5,000 LILACS records with Omikuji Bonsai and DeCS Highlighter and evaluated both quantitatively and qualitatively. Results indicated good precision and thematic alignment. The interface enables supervised curation, and future developments include integration with Fi-Admin and expansion of the corpus, aiming to enhance subject retrieval and strengthen the application of DeCS in real-world settings.

**Keywords:** Abstracting and Indexing. Classification Algorithms. Natural Language Processing. Machine Learning. Vocabulary, Controlled.

## 1 INTRODUÇÃO

O DeCS Finder IA é uma ferramenta desenvolvida pelo Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde - BIREME/OPAS/OMS para apoiar a organização e a recuperação da informação em saúde, por meio da descoberta automática de descritores do tesouro DeCS – Descritores em Ciências da Saúde.

O DeCS, desenvolvido pela BIREME desde 1986, é um vocabulário controlado que amplia o MeSH (*Medical Subject Headings*) da NLM (*National Library of Medicine*, dos Estados Unidos), incorporando termos relevantes para a realidade dos países da América Latina e Caribe (LAC), especialmente em suas cinco categorias exclusivas: Saúde Pública; Vigilância em Saúde; Homeopatia; Ciência e Saúde e Medicinas Tradicionais, Complementares e Integrativas, e os traduzindo aos idiomas falados na região LAC, português, espanhol e francês - além do inglês do MeSH, de modo a promover a padronização semântica e a interoperabilidade entre sistemas de informação<sup>1</sup>.

A indexação é um processo de análise documental cujo objetivo é representar o conteúdo de documentos por termos padronizados provenientes de uma linguagem de indexação, facilitando a recuperação da informação (RI). Nas bibliotecas, ela funciona como uma “tradução” do conteúdo documental em descritores que são como pistas para RI. Essa atividade visa mediar a relação entre o usuário e a informação, sendo essencial para a eficácia dos sistemas de RI e para o desempenho do próprio sistema de informação. Segundo Fujita (2003, p. 61) a indexação “[...] é uma combinação

---

<sup>1</sup> Sobre o DeCS – Descritores em Ciências da Saúde. São Paulo: BIREME/OPAS/OMS, 2025. Disponível em: <https://decs.bvsalud.org/sobre-o-decs/>. Acesso em: 29 jun. 2025.



metodológica altamente estratégica entre o tratamento do conteúdo de documentos e sua recuperação por um usuário” evidenciando a relação entre processo e finalidade da indexação.

A indexação exige mais do que a simples “extração” de termos, sendo uma atividade cognitiva que requer compreensão do conteúdo, estratégias cognitivas e atenção aos objetivos da representação temática. Na sequência, é necessário que os conceitos identificados sejam compatibilizados com os termos da linguagem utilizada, de modo a garantir a coerência semântica da base de dados (Redígolo; Fujita, 2015).

O esforço em reduzir a carga de trabalho e o tempo despendido na indexação constitui preocupação presente no campo da informação desde a década de 1950 (Silva; Correa; Gil-Leiva, 2020). Porém, a automatização da indexação se tornou uma alternativa viável diante do crescimento da capacidade computacional disponível e os recentes avanços tecnológicos especialmente no campo da Inteligência Artificial, como integração com Sistemas Gerenciadores de Banco de Dados (SGBD); Processamento de Linguagem Natural (PLN); Aprendizado de Máquina (*Machine Learning*) e a combinação de métodos que integram abordagens estatísticas, linguísticas e associativas.

Neste trabalho, indexação automática refere-se a “qualquer procedimento que permita identificar e selecionar os termos que representam o conteúdo dos documentos, sem a intervenção direta do indexador” (Robredo, 2005, p. 170). Porém, este conceito pode ser decomposto em três mais específicos, conforme Narukawa (2011) com base em Gil Leiva (1999): a indexação assistida por computador, a indexação semiautomática e a indexação automática.

A distinção entre as três reside no nível de independência dos sistemas, sendo a indexação assistida por computador em que o indexador realiza toda a análise do conteúdo, usando o sistema apenas para armazenar a representação. Já na indexação semiautomática o sistema analisa o conteúdo e o indexador avalia os termos propostos para indexação. E na indexação automática não há a presença do indexador humano, o documento é analisado por um sistema de indexação que determina a indexação. Neste modelo não há tampouco validação humana dos termos definidos pelo sistema (Narukawa, 2011).

Neste trabalho apresenta-se um relato de experiência de desenvolvimento e implementação do DeCS Finder IA e sua integração com a ferramenta Annif, destacando



o uso do algoritmo DeCS Highlighter no aprimoramento da indexação semiautomática baseada no vocabulário controlado DeCS.

### **1.1 Premissas técnicas da integração do DeCS Finder IA e do algoritmo DeCS Highlighter com o Annif**

A ferramenta DeCS Finder<sup>2</sup> foi lançada em 2021, utilizando inicialmente apenas um algoritmo léxico de Processamento de Linguagem Natural (PLN) denominado DeCS Highlighter. Em 2024, foi publicada a versão beta do DeCS Finder com Inteligência Artificial (IA), incorporando modelos mais avançados de análise de texto para a recomendação de descritores embasada pela Metodologia LILACS.

Para integrar métodos mais avançados de PLN ao DeCS Finder IA optou-se pelo Framework Annif de indexação. Annif é uma ferramenta de código aberto desenvolvida pela Biblioteca Nacional da Finlândia, que combina PLN e aprendizado de máquina. A ferramenta dispõe de algoritmos de análise como TF-IDF, Omikuji, FastText e redes neurais. Seu objetivo é executar a indexação automática e a classificação a partir de diferentes coleções de documentos (Brito; Martins, 2023).

Outra possibilidade existente no Annif é a realização de *ensembles*, a combinação de dois ou mais algoritmos visando aprimorar o resultado da indexação, de acordo com Martins e Brito (2023) muitas ferramentas usam um único algoritmo de classificação, enquanto o Annif permite configurar projetos com diversos backends.

A escolha do Annif deve-se à possibilidade de treinamento com vocabulários controlados, permitindo a personalização dos sistemas de indexação automática e ao uso de *ensembles* que proporciona maior precisão ao combinar diferentes modelos, reduzindo limitações inerentes a métodos isolados. Tais características são relevantes para a organização do conhecimento em Ciências da Saúde em que é essencial o uso da terminologia especializada que está em constante atualização. Entretanto é necessário destacar que em sua documentação explicita-se que o Annif não considera a estrutura interna do vocabulário, exigindo apenas a compreensão das URIs e dos rótulos de preferência de cada assunto ou conceito (Brito; Martins, 2022).

---

<sup>2</sup> DeCS Finder IA. São Paulo: BIREME/OPAS/OMS, 2025. Disponível em: <https://decsfinder.bvsalud.org>. Acesso em: 29 jun. 2025.



Entre as premissas fundamentais para o desenvolvimento da integração entre DeCS Finder IA e Annif está a definição clara sobre o nível de intervenção humana na indexação automática. Consideram-se distintos níveis dessa intervenção, como: indexação manual assistida, o indexador analisa todo o conteúdo e o sistema apenas armazena a representação e Indexação Automática, a indexação é realizada integralmente por sistema, sem intervenção humana na análise ou edição dos termos. Nesse cenário, o papel do humano se volta para o projeto, desenvolvimento, aprimoramento e atualização do sistema para garantir que os resultados gerados sejam confiáveis.

## **2 METODOLOGIA**

O trabalho aqui apresentado caracteriza-se como um relato de experiência, tendo como foco a descrição do processo de implementação da integração entre o DeCS Finder IA e Annif, enfatizando a documentação dos procedimentos, decisões metodológicas e os resultados alcançados. Assim, a literatura citada cumpre função de sustentação teórica e contextualização, mas não corresponde a uma revisão de literatura compreensiva.

### **2.1 Seleção do corpus**

O processo de seleção do corpus levou em consideração a qualidade da indexação manual existente, de modo a especializar os modelos treinados conforme os critérios da Metodologia LILACS<sup>3</sup>.

O corpus utilizado para treinamento dos modelos foi composto por cerca de 5.000 registros bibliográficos da base LILACS<sup>4</sup>, previamente indexados manualmente com DeCS. Foi selecionada uma amostra de registros indexados por 10 indexadores proficientes na Metodologia LILACS e no uso do vocabulário DeCS. Não houve recorte temporal ou de idioma, sendo incluídos artigos escritos em inglês, espanhol e português. Para a realização dos treinamentos o corpus foi particionado seguindo a proporção

---

<sup>3</sup> Metodologia LILACS. São Paulo: BIREME/OPAS/OMS, 2025. Disponível em: <https://lilacs.bvsalud.org/metodologia-lilacs/>. Acesso em: 29 jun. 2025.

<sup>4</sup> LILACS - Literatura Latino-Americana e do Caribe em Ciências da Saúde. São Paulo: BIREME/OPAS/OMS, 2025. Disponível em: <https://lilacs.bvsalud.org/>. Acesso em: 29 jun. 2025.



80/10/10 para conjuntos de treinamento, teste e validação. O conjunto de treinamento (80% do corpus) é utilizado para ajustar os parâmetros internos do modelo, permitindo que ele aprenda padrões presentes nos dados. O conjunto de validação (10%) é empregado durante o treinamento para acompanhar o desempenho do modelo. Já o conjunto de teste (10%) é destinado à etapa final de avaliação, funcionando como uma amostra independente que possibilita verificar a capacidade do modelo de aplicar o aprendizado a dados não acessados anteriormente. Essa estratégia é amplamente adotada em experimentos de aprendizado de máquina por garantir equilíbrio entre aprendizado e avaliação crítica dos resultados.

## **2.2 Treinamento dos modelos e configuração do pipeline**

Para a composição do projeto no Annif foram incluídos como parte do corpus títulos, resumos e descritores de assunto, descritores pré-codificados e descritores de tipos de publicação assinalados aos respectivos artigos.

O algoritmo Omikuji Bonsai foi priorizado por sua adequação a classificações hierárquicas extensas, como o vocabulário DeCS. Testes comparativos prévios mostraram que o Omikuji Bonsai superou o TF-IDF na hierarquização dos descritores, especialmente por sua capacidade de representar a estrutura semântica do vocabulário DeCS. Essa abordagem favorece tanto a escalabilidade quanto a precisão dos resultados. A escolha do Omikuji Bonsai deve-se à estrutura em árvores otimizadas, que organiza grandes volumes de descritores e responde com agilidade a classificações complexas.

O vocabulário utilizado no treinamento foi estruturado no formato Turtle (.ttl) e modelado em SKOS. Cada descritor do DeCS foi representado como um skos:Concept com as propriedades skos:prefLabel (termo preferido), skos:altLabel (termos alternativos/não preferidos), skos:scopeNote (nota de escopo DeCS), skos:notation (código hierárquico DeCS) e skos:broader (termos genéricos).

## **2.3 Avaliação**

A avaliação quantitativa dos modelos utilizou métricas clássicas da recuperação da informação e já disponíveis no ambiente Annif: precisão (proporção de verdadeiros positivos entre os resultados sugeridos), revocação (proporção de verdadeiros positivos entre os esperados) e F1-score como medida harmônica entre ambas.



Complementarmente, foi realizada uma validação qualitativa com especialistas em indexação da BIREME, que analisaram a relevância e adequação dos termos sugeridos, comparando-os com os descritores atribuídos manualmente nos registros do corpus de treinamento.

## 2.4 Desenvolvimento da interface

A interface gráfica do DeCS Finder foi adaptada para apresentar os descritores recomendados pelo Annif em paralelo aos resultados do algoritmo DeCS Highlighter. A organização permite comparar os conjuntos de sugestões e selecionar os termos mais adequados, caracterizando a indexação semiautomática supervisionada. As sugestões podem ser exportadas em formato .txt, facilitando a curadoria humana. A interface oferece acesso ao Serviço de Atenção ao Usuário para envio de sugestões e comentários à equipe do projeto.

**Figura 1** – Interface gráfica DeCS Finder IA – versão beta

DeCS/MeSH  
Descritores em Ciências da Saúde  
Finder

English | Español | Português | Français

Idioma do seu texto: Português  
Idioma dos termos: O mesmo localizado no texto  
Tipos de termos: Descritores, Qualificadores

Cole abaixo o seu texto:

O diabetes situa-se entre as dez principais causas de morte nos países ocidentais, sendo uma das condições mais prevalentes da atualidade e alvo de muitos estudos. Nas últimas décadas a prevalência de DM2 vem aumentando, o que se caracteriza como uma epidemia mundial e um problema de saúde pública. Objetiva-se neste estudo, discutir os benefícios dos principais compostos bioativos presentes em alimentos funcionais na prevenção do diabetes mellitus tipo 2, com base em uma revisão de literatura. Trata-se de uma revisão integrativa da literatura, com abordagem qualitativa e exploratória, realizada durante o período de novembro de 2021 a fevereiro de 2022, no qual foram consultadas bases de dados eletrônicas como Scielo, Lilacs, PubMed e Google Acadêmico. Os compostos bioativos mencionados na tabela 1 e 2 apresentam vários efeitos fisiológicos e benéficos ao organismo de uma pessoa diabética, como a redução da absorção da glicose e resistência insulínica contribuído para o controle da glicemia e melhoria na qualidade de vida das indivíduos portadoras da DM2. Contribuições para:

Termos extraídos:

- Qualidade de Vida
- Prevenção de Doenças
- Revisão
- Alimento Funcional
- Prevalência
- Saúde Pública
- Pessoas
- Diabetes Mellitus Tipo 2
- Literatura
- LILACS
- Exercício Físico

Termos identificados por IA:

- Humanos
- Revisão
- Idoso
- Revisão Sistemática
- Diabetes Mellitus Tipo 2
- Qualidade de Vida
- Feminino
- Adolescente
- Masculino
- Adulto
- Pé Diabético

Fonte: <https://decsfinder.bvsalud.org/dmf>.

Descrição: Interface gráfica do usuário DeCS Finder IA versão beta de abril de 2025.

## 3 RESULTADOS E DISCUSSÕES

Os testes realizados demonstraram que a integração do Annif ao DeCS Finder IA foi capaz de produzir sugestões de descritores tematicamente relevantes, alinhadas à estrutura do vocabulário DeCS e à prática de indexação adotada pela Metodologia LILACS. Os modelos treinados com Omikui e vocabulário em SKOS apresentaram desempenho satisfatório nas métricas de precisão, revocação e F1-score, especialmente

quando comparados a abordagens estatísticas isoladas. A validação qualitativa também indicou boa aderência dos descritores sugeridos, com contribuições significativas para a curadoria semiautomática.

A inclusão dos tipos de termos “descritores pré-codificados” e “descritores de tipos de publicação” no corpus de treinamento dos modelos contribuiu para o aumento das métricas quantitativas, que ampliaram a capacidade de predição dos termos. No entanto, observou-se que esses termos, em alguns casos, influenciaram excessivamente a seleção, sendo propostos em detrimento de descritores de assunto alinhados ao conteúdo principal.

Os modelos testados apresentaram variações significativas nas métricas de avaliação. Em termos de desempenho médio por documento (doc avg), o modelo ensemble denominado internamente como nn3 obteve os melhores resultados: precisão de 0,5217, revocação de 0,5283 e F1-score de 0,5051. Já o modelo Omikuji Bonsai, com configuração padrão, apresentou precisão de 0,4809, revocação de 0,4882 e F1-score de 0,4663, indicando um desempenho inferior, porém consistente. O modelo híbrido nn3 com Highlighter teve desempenho intermediário, com F1-score de 0,4975. Esses resultados evidenciam o potencial de combinação de abordagens neurais com técnicas lexicais para melhoria da indexação automática.

**Quadro 1** – Métricas de avaliação dos modelos comparados

|                      | Omikuji - DTA | Omikuji Bonsai | Omikuji Bonsai DTEA | nn2(bonsai peso 2) | nn3    | nn3 - Highlighter peso 2 |
|----------------------|---------------|----------------|---------------------|--------------------|--------|--------------------------|
| Precision (doc avg): | 0.4725        | 0.4809         | 0.4781              | 0.5137             | 0.5217 | 0.5130                   |
| Recall (doc avg):    | 0.4814        | 0.4882         | 0.4836              | 0.5174             | 0.5283 | 0.5224                   |
| F1 score (doc avg):  | 0.4588        | 0.4663         | 0.4630              | 0.4963             | 0.5051 | 0.4975                   |
| F1@5:                | 0.4194        | 0.4279         | 0.4240              | 0.4467             | 0.4499 | 0.4577                   |
| NDCG:                | 0.5688        | 0.5742         | 0.5696              | 0.5990             | 0.6069 | 0.6036                   |
| NDCG@5:              | 0.6982        | 0.7037         | 0.6999              | 0.7248             | 0.7283 | 0.7354                   |
| NDCG@10:             | 0.6225        | 0.6284         | 0.6239              | 0.6556             | 0.6637 | 0.6600                   |
| Precision@1:         | 0.9074        | 0.9043         | 0.8951              | 0.9105             | 0.9043 | 0.8920                   |
| Precision@3:         | 0.7490        | 0.7438         | 0.7490              | 0.7654             | 0.7675 | 0.7788                   |
| Precision@5:         | 0.6340        | 0.6432         | 0.6401              | 0.6692             | 0.6745 | 0.6852                   |
| True positives:      | 1531          | 1558           | 1549                | 1646               | 1675   | 1653                     |
| False positives:     | 1709          | 1682           | 1691                | 1526               | 1510   | 1553                     |
| False negatives:     | 1735          | 1708           | 1717                | 1620               | 1591   | 1613                     |



|                      |                                                                         |     |     |     |      |     |
|----------------------|-------------------------------------------------------------------------|-----|-----|-----|------|-----|
| Documents evaluated: | 324                                                                     | 324 | 324 | 324 | 324  | 324 |
| Duração época        |                                                                         |     |     | 97s | 180s |     |
| nn2                  | por padrão é esemble de Omikuji Bonsai x DeCS Highlighter               |     |     |     |      |     |
| nn3                  | por padrão é esemble de Omikuji Bonsai x Bonsai DTEA x DeCS Highlighter |     |     |     |      |     |

Fonte: elaborado pelos autores a partir de métricas fornecidas por Annif.

Descrição: Tabela comparativa com métricas de desempenho de seis modelos de atribuição de descritores.

#### 4 CONSIDERAÇÕES FINAIS

A experiência demonstrou que, mesmo com avanços técnicos, a atuação humana continua sendo central na validação e no refinamento dos resultados sugeridos pelos modelos. O uso supervisionado da ferramenta mostrou-se eficaz para potencializar o trabalho dos indexadores, sem substituí-los preservando a interpretação crítica e o julgamento especializado inerentes ao processo de indexação.

Como limitação, destaca-se que os resultados dos modelos são sensíveis à composição do corpus de treinamento o que reforça a importância de bases representativas e criteriosamente selecionadas para o sucesso da aprendizagem automática.

Como desdobramento futuro da iniciativa, prevê-se a integração dos resultados de Annif e DeCS Highlighter em lista única com pontuação combinada, integrando aprendizado de máquina e técnicas lexicais consolidadas na saúde. Essa integração ampliar a consistência semântica das sugestões e oferecer um conjunto mais confiável de descritores.

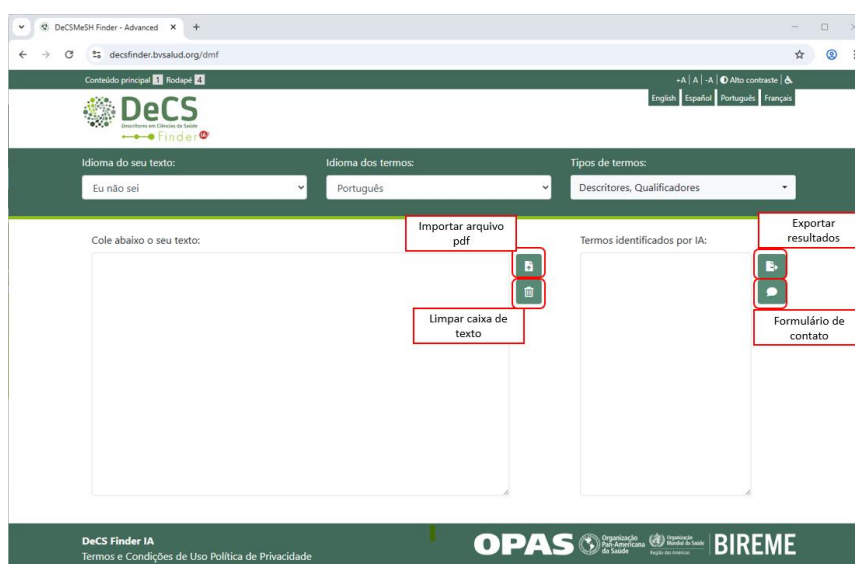
Também está prevista a ampliação do corpus de treinamento para cerca de 40 mil registros, todos provenientes da base de dados LILACS, mantendo assim o "padrão ouro" de indexação como parâmetro de treinamento, conforme diretrizes estabelecidas na Metodologia LILACS (BIREME), que orienta a qualidade e consistência na indexação temática de registros em saúde.

Assim como, a incorporação da ferramenta ao ambiente Fi-Admin, sistema de administração de fontes de informação bibliográfica mantido pela BIREME/OPAS/OMS, que visa possibilitar sua integração direta ao fluxo de entrada de dados do acervo da

BVS, otimizando o processo de sugestão de descritores no momento da descrição bibliográfica.

Prevê-se também ações centradas na renovação da interface do DeCS Finder IA, com foco em usabilidade e clareza semântica, incluindo a disponibilização de instruções mais claras sobre sua utilização; a melhoria das interfaces assistida e direta para maior clareza de suas funcionalidades; a possibilidade de anexar arquivos de texto simples (.txt) e textos disponibilizados online (via url) para análise; links diretos nos resultados para o descritor no DeCS e ampliação da seção de perguntas frequentes. As melhorias buscam facilitar a indexação semiautomática e aprimorar a interação dos diferentes perfis de usuários com a ferramenta.

**Figura 2** – Interface gráfica DeCS Finder IA



Fonte: <https://decsfinder.bvsalud.org/dmf>

Descrição: Interface do DeCS Finder IA, versão de produção de agosto de 2025.

Mesmo sendo automatizada, a qualidade dos modelos do Annif depende diretamente do vocabulário controlado utilizado e da qualidade do corpus de treinamento. A adoção de um conjunto de dados indexado por profissionais experientes e o uso estruturado do vocabulário DeCS foram fatores essenciais para a obtenção de resultados alinhados à prática da indexação em saúde.

Desta forma, os desenvolvimentos futuros do DeCS Finder IA encontram desafios na necessidade de atualização contínua do vocabulário incorporado ao modelo e na integração completa com o sistema de gestão de fontes de informação. Primeiro, considerando que o DeCS é atualizado anualmente, um modelo com vocabulário



desatualizado comprometeria a qualidade da indexação. Segundo, na integração completa com o FI-Admin, que consiste em um requisito estratégico para viabilizar uma indexação automática completa incluída no fluxo de processamento dos dados bibliográficos.

Apesar da existência de ferramentas de indexação automática, o conhecimento e o uso prático dessa abordagem ainda são limitados entre os profissionais da informação. A experiência aqui relatada busca contribuir para ampliar esse repertório e apoiar sua adoção crítica e contextualizada em ambientes reais de trabalho.

Com base nos avanços relatados e nas direções propostas para a continuidade do projeto, o DeCS Finder IA mostra-se como uma iniciativa promissora para a modernização dos processos de indexação em saúde. Sua integração com o Annif, aliada ao uso do vocabulário DeCS, abre caminho para soluções mais interoperáveis, responsivas e adaptáveis às necessidades das bibliotecas e repositórios científicos da região.

O DeCS Finder IA, ao integrar tecnologias de inteligência artificial com vocabulário controlado especializado, mostra-se também como um recurso estratégico para a modernização dos fluxos operacionais da Biblioteca Virtual em Saúde (BVS). Sua aplicação contribui diretamente para o fortalecimento da interoperabilidade entre bases de dados e sistemas de informação regionais, além de ampliar o apoio à gestão temática da informação em saúde. Ao automatizar etapas críticas do processo de indexação e facilitar a curadoria de descritores, a ferramenta busca viabilizar maior escalabilidade e padronização nas atividades de organização do conhecimento.

Adicionalmente, as palavras-chave atribuídas a este trabalho foram inicialmente sugeridas pela ferramenta DeCS Finder IA e, em seguida, analisadas e validadas pelos autores com base em critérios semânticos e de representatividade temática. Todas as palavras-chave selecionadas correspondem a descritores controlados do vocabulário DeCS.

Consideramos que a metodologia desenvolvida possui potencial de replicação em diferentes sistemas de informação que adotem o vocabulário DeCS, favorecendo a padronização semântica e o fortalecimento da recuperação temática da informação em saúde em contextos diversos.



## REFERÊNCIAS

BRITO, Jean Carlos Borges; MARTINS, Dalton Lopes. Framework genérico para geração automática de assuntos e indexação em repositório digital. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 28, fluxo contínuo, 2023. Disponível em: <https://doi.org/10.1590/1981-5344/46629>. Acesso em: 08 jun. 2025.

BRITO, Jean Carlos Borges; MARTINS, Dalton Lopes. Geração automática de metadados: estudo de caso utilizando a técnica de indexação automática estatística com a ferramenta ANNIF. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais eletrônicos** [...]. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxii/enancib/paper/view/777>. Acesso em: 28 jun. 2025.

FUJITA, Mariângela Spotti Lopes. A identificação de conceitos no processo de análise de assunto para indexação. **RDBCi: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.1, n.1, p. 60-90, 2003. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2089>. Acesso em: 15 ago. 2025.

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

NARUKAWA, Cristina Miyuki. **Estudo de vocabulário controlado na indexação automática**: aplicação no processo de indexação do Sistema de Indización SemiAutomática (SISA). 2011. 222f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2011. Disponível em: <http://hdl.handle.net/11449/93677>. Acesso em: 16 ago. 2025.

REDÍGOLO, Franciele Marques; FUJITA, Mariângela Spotti Lopes. A leitura profissional do catalogador e seu papel como mediadora da informação. **Informação & Informação**, Londrina, v. 20, n. 3, p. 356–376, 2015. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/21683>. Acesso em: 29 jun. 2025.

ROBREDO, J. **Documentação de hoje e de amanhã**: uma abordagem revisitada e contemporânea da Ciência da Informação e de suas aplicações biblioteconômicas, documentárias, arquivísticas e museológicas. 4. ed. rev. e ampl. Brasília DF: Edição de autor, 2005.

SILVA, Sâmela Rouse de Brito; CORREA, Renato Fernandes; GIL-LEIVA, Isidoro. Avaliação direta e conjunta de Sistemas de Indexação Automática por Atribuição. **Informação & Sociedade: Estudos**, João Pessoa, v. 30, n. 4, p. 1-27, 2020. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/57259/32624>. Acesso em: 29 jun. 2025.