

UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE COMUNICAÇÕES E ARTES  
DEPARTAMENTO DE INFORMAÇÃO E CULTURA

AILA RENATA OKADA GOULART

**Uso do Processamento de Linguagem Natural no desenvolvimento de um sistema de organização do conhecimento: um estudo de caso de documentação na área da saúde**

SÃO PAULO  
2024

AILA RENATA OKADA GOULART

**Uso do Processamento de Linguagem Natural no desenvolvimento de um sistema de organização do conhecimento: um estudo de caso de documentação na área da saúde**

Trabalho de conclusão de curso de graduação em Biblioteconomia, apresentado ao Departamento de Informação e Cultura como requisito para a obtenção do título de Bacharel em Biblioteconomia.

Orientação: Prof.<sup>a</sup> Dr<sup>a</sup>. Cibele Araújo Camargo Marques dos Santos

SÃO PAULO  
2024

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catalogação na Publicação  
Serviço de Biblioteca e Documentação  
Escola de Comunicações e Artes da Universidade de São Paulo  
Dados inseridos pelo(a) autor(a)

---

Goulart, Aila Renata Okada  
Uso do Processamento de Linguagem Natural na  
construção de sistemas de organização do conhecimento: um  
estudo de caso para a elaboração de documentação na área  
da saúde / Aila Renata Okada Goulart; orientadora, Cibele  
de Araújo Camargo Marques dos Santos. - São Paulo, 2024.  
91 p.: il.

Trabalho de Conclusão de Curso (Graduação) -  
Departamento de Informação e Cultura / Escola de  
Comunicações e Artes / Universidade de São Paulo.  
Bibliografia

1. Terminologia. 2. Linguística de corpus. 3.  
Processamento de linguagem natural. 4. Doenças  
negligenciadas. I. Santos, Cibele de Araújo Camargo  
Marques dos. II. Título.

CDD 21.ed. - 020

## FOLHA DE APROVAÇÃO

Nome: Goulart, Aila Renata Okada

Título: Uso do Processamento de Linguagem Natural no desenvolvimento de um sistema de organização do conhecimento: um estudo de caso na área da saúde

Aprovado em: \_\_\_\_ / \_\_\_\_ / \_\_\_\_\_

Banca:

Nome:

Instituição:

Nome:

Instituição:

Nome:

Instituição:

## AGRADECIMENTOS

Agradeço e dedico este trabalho ao meu companheiro e melhor amigo, meu marido Júlio César Goulart, por complementar meu lado humano e sentimental com seu lado exato e racional.

À minha família, em especial à minha falecida *Batchan* Kasuko, que sempre soube da importância do conhecimento e dos estudos; à minha mãe, Lidia, que sempre fez de tudo para as filhas; para minha irmã, Karina, que me proporcionou essa experiência única que é ser irmã mais velha; ao meu tio, Augusto, que é como um pai para mim; às minhas tias Edna, Vera, Bete e Marly, que também cuidaram de mim e incentivavam os estudos e o hábito da leitura desde criança; ao meu primo, William, que me mostrou o mundo dos livros, quadrinhos, videogames e RPG.

À minha orientadora, a querida Profª. Cibele, por ser uma pessoa tão inspiradora.

Também agradeço todos os professores e funcionários do CBD, da ECA e da Universidade, que contribuem para que todos os anos tantos alunos iniciem, percorram e concluam sua jornada acadêmica.

Às minhas amigas Aline, Natália e Silvia, por compartilharem seu tempo, suas palavras e seus conhecimentos nos momentos que precisei de alguma ajuda, orientação ou distração.

Aos meus queridos amigos Daiane e Renato, pelas conversas, viagens e apoio imensurável e incomparável.

À minha chefe, Simone, às minhas colegas de trabalho, Kátia, Luciane, Mayara, Milena, Valéria e especialmente à Rita, por me fazerem pertencer a um lugar que tanto gosto: a Biblioteca.

“We can only see a short distance ahead,  
but we can see plenty there that needs to be done.”  
(Turing, 1950, p. 460)<sup>1</sup>.

---

<sup>1</sup> “Podemos avistar só um pequeno trecho do caminho à nossa frente, mas ali já vemos muito do que precisa ser feito.”

## RESUMO

**GOULART, A. R. O. Uso do Processamento de Linguagem Natural na construção de sistemas de organização do conhecimento:** um estudo de caso para a elaboração de documentação na área da saúde. 2024. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) — Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2024.

O presente trabalho tem como proposta averiguar a possibilidade de utilizar técnicas de Processamento de Linguagem Natural (PLN) e de análise de Linguística de *Corpus* na construção de uma Terminologia na área da saúde. O objetivo geral é averiguar como a recuperação e o acesso à informação podem ser beneficiados pelo PLN; como objetivos específicos foram estabelecidos a definição de Processamento de Linguagem Natural, e em quais áreas relacionadas à Biblioteconomia ele poderia ser aplicado, no caso a Linguística Documentária, Terminologia e Linguística de *Corpus*. A metodologia utilizada foi a de revisão bibliográfica seguida de uma parte prática por meio de levantamento de termos relacionados às denominadas Doenças Negligenciadas utilizando as ferramentas de pesquisa e análise de *corpus* BootCat e AntConc. Os termos utilizados para a pesquisa do *corpus* foram: doenças negligenciadas, doenças esquecidas, doenças tropicais negligenciadas, elefantíase filarial, helmintíase, hanseníase, oncocercose, esquistossomose, tracoma e popular. Os resultados obtidos dessas buscas foram: 126 termos recuperados no total, dos quais foram selecionados 32 substantivos para a confecção da fichas terminológicas de cada uma dessas palavras: Brasil, casos, causa, cegueira, Chagas, controle, cura, dengue, dia, diagnóstico, doença, eliminação, esquistossomose, filariose, hanseníase, helmintíase, hepatites, iniciativa, leishmaniose, malária, oncocercose, pessoa, problema, raiva, rios, saúde, sintomas, tracoma, transmissão, tratamento e tuberculose. Os termos levantados foram utilizados para a elaboração de uma árvore de domínio terminológico. Com a pesquisa, observou-se que os termos com maior número de ocorrências são “esquistossomose” e “hanseníase”, com 17 e 21 termos candidatos, respectivamente, e que “doenças esquecidas”, “elefantíase filarial”, “helmintíase” e “popular”, foram os termos pesquisados que não apresentaram termos candidatos. Discute-se a possibilidade da seleção do português como filtro de busca ter influenciado nos resultados obtidos. Também foi observado que muitos documentos e informações recuperados são provenientes de sites da área da saúde. Embora este resultado tenha sido inesperado, o uso das técnicas de PLN e de Linguística de *Corpus* mostrou-se pertinente e satisfatório nos estudos terminológicos.

**Palavras-chave:** Terminologia. Linguística de *corpus*. Processamento de linguagem natural. Doenças negligenciadas.

## ABSTRACT

GOULART, A. R. O. **Use of Natural Language Processing for elaborating knowledge organization systems: a case study in medical documentation.** 2024. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2024.

The present work aims to investigate the possibility of using Natural Language Processing (NLP) techniques and Corpus Linguistics analysis in constructing a Terminology in the healthcare field. The general objective of this work is to examine how information retrieval and access can benefit from NLP. Specific objectives include defining Natural Language Processing and exploring its potential applications in related areas of Library Science, specifically Documentary Linguistics, Terminology, and Corpus Linguistics. The methodology used in this work involved a literature review followed by a practical part through a case study of term retrieval related to so-called Neglected Diseases, using the search and corpus analysis tools BootCat and AntConc, also the DeCS/MeSH thesaurus was used as a main source of information for the medical terms. The terms used to search for the corpus are: neglected diseases, forgotten diseases, neglected tropical diseases, elephantiasis filarial, helminthiasis, leprosy, onchocerciasis, schistosomiasis, trachoma and popular. The results obtained included a total of 126 recovered terms, from which 32 nouns were selected for the creation of terminological cards for each noun: Brazil, cases, cause, blindness, Chagas, control, cure, dengue, day, diagnosis, disease, elimination, schistosomiasis, filariasis, leprosy, helminthiasis, hepatitis, initiative, leishmaniasis, malaria, onchocerciasis, person, problem, rabies, rivers, health, symptoms, trachoma, transmission, treatment and tuberculosis. The identified terms were used to develop a domain tree of terminology. The research revealed that "schistosomiasis" and "leprosy" were the terms with the highest number of occurrences, with 17 and 21 candidate terms, respectively. Terms such as "neglected diseases", "filariasis", "helminthiasis", and "popular" did not show candidate terms. The study discusses the possibility that using Portuguese as a search filter may have influenced the results. Additionally, it was noted that many retrieved documents and information originated from healthcare websites. Although the results were unexpected, the use of NLP and Corpus Linguistics techniques proved to be relevant and satisfactory in terminological studies.

Keywords: Terminology. Corpus linguistics. Natural language processing. Neglected diseases.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	10
1.1	<b>Justificativa.....</b>	12
1.2	<b>Objetivos .....</b>	13
1.3	<b>Metodologia .....</b>	13
<b>2</b>	<b>PROCESSAMENTO DE LINGUAGEM NATURAL.....</b>	15
2.1	<b>Marcos históricos e paradigmas do PLN .....</b>	16
2.2	<b>Aplicações e desafios das técnicas de PLN.....</b>	17
<b>3</b>	<b>LINGUÍSTICA DOCUMENTÁRIA E TERMINOLOGIA.....</b>	20
<b>4</b>	<b>LINGUÍSTICA DE <i>CORPUS</i>.....</b>	24
4.1	<b>Aplicações do PLN na LC .....</b>	24
4.2	<b>Programas de coleta e análise de <i>corpus</i> .....</b>	25
<b>5</b>	<b>PROCEDIMENTOS METODOLÓGICOS .....</b>	26
5.1	<b>BootCat .....</b>	26
5.2	<b>AntConc .....</b>	34
<b>6</b>	<b>TERMINOLOGIA DE DOENÇAS NEGLIGENCIADAS .....</b>	49
<b>7</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	55
	<b>REFERÊNCIAS .....</b>	57
	<b>APÊNDICES .....</b>	60

## 1 INTRODUÇÃO

Desde os anos 2000 e o fortalecimento do Google como principal fonte de pesquisa na web, o desenvolvimento de programas e ferramentas que facilitassem a troca de informações teve avanços promissores no que diz respeito à identificação e análise de texto – tanto no mesmo idioma quanto em versões e traduções para o inglês e outras línguas.

Para que a troca de informações seja bem-sucedida, além da “mineração de dados” (*data mining*), a análise de grande quantidade de dados (*big data*) e a Inteligência Artificial, é preciso estudar e desenvolver mais a área do Processamento de Linguagem Natural (PLN). Sem essa análise linguística não seria possível criar os programas que interagem com as pessoas por voz, como a Siri da Apple e a Alexa da Amazon, ou por texto, como o Chat GPT, da empresa OpenAI, e o Gemini (antigo BARD), da Google, que viraram um sucesso entre o final de 2022 e início de 2024. Tais programas mostraram-se recursos importantes para facilitar a vida das pessoas em geral, e também podem ser muito úteis para pesquisadores de todas as áreas tanto como ferramenta quanto como objeto de estudo.

O foco do PLN é sua aplicação na interação humano computador e a geração automática de conteúdo (sejam textos, imagens, áudios ou vídeos), mas também pode ser aplicado em diversas áreas. Uma dessas áreas é a medicina, onde uma vez que nos atendimentos médicos é necessário que haja um diálogo entre os termos técnicos e especializados com a linguagem coloquial, simples e com regionalismos, utilizada pelos pacientes, que, em sua maioria, são pessoas comuns, sem conhecimento dos nomes de doenças, procedimentos e medicamentos que os profissionais da saúde empregam nos atendimentos, com colegas de profissão e nos estudos e pesquisas.

O advento da internet facilitou a vida dos pacientes, possibilitando que estes tivessem um meio mais acessível de procurar sintomas e possíveis tratamentos - tanto financeiramente, por não precisarem pagar por um atendimento médico; de distância, por meio da telemedicina; e também de entendimento, no caso de não compreenderem alguma instrução ou qual medicamento foi receitado. Como exemplo disso, temos o ano de 2020 e pandemia de Covid-19, que já tornou a década de 2020 um marco histórico não apenas na área da saúde (devido à pandemia, às pesquisas focadas nesta nova doença, no desenvolvimento de tratamentos e de vacinas), mas também por todo o avanço tecnológico impulsionado pelo isolamento social e a urgência de se manter as atividades educacionais, sociais e econômicas.

Os computadores, celulares e outros eletrônicos, junto da internet, passaram a ser ferramentas básicas de trabalho e estudo, ajudando a democratizar o acesso à informação e

contribuindo para que se formasse um grande compêndio de dados que necessitam de tratamento. Com isso, a sociedade como um todo identificou uma nova necessidade no que diz respeito à produção, organização, armazenamento, recuperação e disseminação de informação. Houve um maior interesse por parte da população para saber o que era a Covid-19, o coronavírus e os termos técnicos dos procedimentos médico-hospitalares e houve também grande incentivo dos meios de comunicação para informar as pessoas sobre a pandemia e termos específicos das doenças que viriam a se tornar parte cotidiana do nosso vocabulário.

No entanto, ainda que tenha ocorrido um aumento da busca por um entendimento dos termos relacionados à medicina, existem doenças que sempre ficaram - e que continuam - à margem do interesse tanto por parte da população quanto das pesquisas científicas. As denominadas Doenças Negligenciadas vão na contramão do *boom* informacional que vivemos nessa década de 2020. Mesmo com um número elevado de casos, e até mesmo serem consideradas epidemias, as Doenças Negligenciadas não recebem a devida atenção por questões de interesse econômico.

É nesse contexto que o presente trabalho se desenvolve. Seu principal objetivo é averiguar como o PLN pode beneficiar a recuperação e o acesso à informação. Foram estabelecidos como objetivos específicos descrever o que é Processamento de Linguagem Natural, pesquisar sobre a aplicação do PLN na Linguística de *corpus* e na Terminologia, e investigar se o Processamento de Linguagem Natural pode servir como suporte no desenvolvimento de novos sistemas de organização do conhecimento.

Para atender aos objetivos propostos, este trabalho é composto por quatro partes. A primeira fala brevemente sobre o PLN, seus marcos históricos, paradigmas, aplicações e desafios para as técnicas existentes. A segunda seção aborda como o PLN pode dialogar com a Biblioteconomia, por isso entramos nos campos da Linguística Documentária e da Terminologia. Após análise dessas áreas, verificou-se que um dos caminhos utilizados para se construir uma terminologia especializada é através da Linguística de *Corpus* (LC). A terceira parte do trabalho trata da LC, com as aplicações do PLN para este campo de estudo e os programas mais conhecidos utilizados na análise de *corpus* – que serão utilizados neste trabalho: BootCat e AntConc. Por último, após consulta ao tesauro DeCS/MeSH da BIREME, faremos a aplicação das técnicas de PLN e dos programas mencionados para verificar a viabilidade da construção de um princípio de terminologia na área da saúde sobre doenças negligenciadas.

A construção da terminologia prosseguiu da seguinte maneira: o *software* BootCat foi utilizado para reunir sites com conteúdo relacionado às Doenças Negligenciadas e formar o

*corpus* a ser analisado pelo programa AntConc. Com este segundo *software* realizou-se o levantamento dos termos com mais ocorrências, sendo selecionados substantivos com uma frequência de mais de 10 ocorrências. A partir da seleção das palavras que atendiam a estes requisitos, efetuou-se o preenchimento de fichas terminológicas para cada um dos termos candidatos. Através da análise do contexto em que essas palavras estavam localizadas nos textos foi possível criar uma proposta de árvore de domínio de termos referentes ao assunto “Doenças Negligenciadas”.

Discute-se a possibilidade de se utilizar a expressão "também conhecido(a) como" para ampliar busca e conseguir localizar termos efetivamente populares (utilizados pela população em geral); a influência da seleção de um idioma específico (neste caso, o português) sob os resultados recuperados estarem relacionados com as Doenças Negligenciadas mais comuns no Brasil; a complexidade da construção de um *corpus* e de uma terminologia não baseados em fontes específicas da área da saúde recuperou um resultado não esperado, trazendo muitos sites e artigos médicos e sobre saúde, porém abre possibilidades para estudos futuros sobre o assunto.

### 1.1 Justificativa

O Processamento de Linguagem Natural realiza a análise de informações, seja da linguagem escrita ou falada, de acordo com os parâmetros específicos de cada idioma. Esse domínio relacionado à computação foi se desenvolvendo de acordo com as necessidades comunicacionais da evolução científica e tecnológica.

Como os maiores avanços tecnológicos ocorrem quase sempre, pioneiramente, no campo da medicina, optou-se por pesquisar como o PLN é utilizado na área da saúde. Por ser uma campo que utiliza e gera grande quantidade de dados/informações, é um objeto de estudo relevante para a Biblioteconomia e a Ciência da Informação.

A área da saúde também se mostra pertinente como objeto de estudo terminológico, pois é uma área fundamental para a sociedade e seria muito importante que se conseguisse trabalhar com um sistema mais dinâmico de coleta, organização e recuperação de informação cuja atualização pudesse ocorrer de forma a acompanhar a velocidade de produção de informação e conteúdo e fosse uma fonte de busca confiável para o pesquisador realizar as buscas para seus trabalhos e também para que as pessoas, os usuários/pacientes do sistema de saúde, possam, por meio do acesso a esse conhecimento mais especializado, conseguirem compreender melhor o que estão sentindo e explicar seus sintomas nas consultas, bem como entender as orientações passadas pelos médicos e profissionais da saúde e seguir seus tratamentos corretamente.

## 1.2 Objetivos

Este trabalho tem como objetivo geral averiguar como o PLN pode beneficiar a recuperação e o acesso à informação. Para isso, foram estabelecidos como objetivos específicos descrever o que é PLN e em qual(is) área(s) de estudo ele pode ser utilizado na Linguística de *Corpus* e na Terminologia para, futuramente, servir como suporte no desenvolvimento de novos sistemas de organização do conhecimento por meio da identificação de termos mais relevantes de um assunto específico, neste caso as Doenças Negligenciadas.

## 1.3 Metodologia

O presente trabalho utilizou o método de pesquisa exploratória e revisão bibliográfica sobre a temática a ser estudada (Gil, 2002; Silveira, 2009), seguindo-se de um estudo de caso e um levantamento seletivo de informações para a construção de um *corpus* terminológico de documentação em saúde.

Como ponto de partida para a pesquisa, primeiramente, buscou-se contextualizar o que é a Programação de Linguagem Natural e como ela pode ser inserida no campo da Biblioteconomia e Ciência da Informação. Os autores base deste trabalho são Jurafsky e Martin (2008), ambos da Universidade de Stanford, nos Estados Unidos) e Nunes (2008), do ICMC-USP São Carlos.

A partir do entendimento das técnicas de PLN, buscou-se uma área da Biblioteconomia em que essas técnicas pudessem ser aplicadas, elegendo a Terminologia como campo de estudo. Os autores adotados como referência para esta parte são Lara e Tálamo (2007), Maimone e Tálamo (2011), Lima (2004), da ECA-USP, e Krieger (2004), docente da Unisinos de Porto Alegre.

Após consultar a literatura da área, verificou-se que o caminho a ser seguido para se adotar as técnicas de PLN na Terminologia seria através da Linguística de *Corpus*. Nesta área, a referência são Sardinha (2000), da PUC-SP, e Biderman (2001), professora emérita da UNESP.

Passamos então para os procedimentos metodológicos. Optou-se, a princípio, por se fazer um estudo de caso para realizar uma análise sobre a viabilidade da construção de uma terminologia da área da saúde com o uso de ferramentas de coleta e análise de *corpus* baseadas técnicas de Processamento de Linguagem Natural. Este processo mostrou-se como um levantamento seletivo de dados e informações.

Dentro da área da saúde, o campo escolhido foi o das Doenças Negligenciadas, pois elas possuem um grande potencial de pesquisa, mas ainda não são o foco de investimentos de grandes centros de estudo. O vocabulário dos Descritores em Ciências da Saúde/Medical Subject Headings (DeCS/MeSH) foi utilizado para a definição de “Doenças Negligenciadas” e para a coleta de palavras-chave de pesquisa de acordo com os termos relacionados. Após estabelecer as palavras para pesquisa, partiu-se para a etapa de obtenção do *corpus* para análise. As ferramentas utilizadas no presente trabalho foram os softwares BootCaT, para o levantamento de conteúdos de *websites* que irão compor o *corpus*, e o AntConc, para a análise dos dados obtidos.

Na construção das fichas terminológicas, foram consultados, além do DeCS/MeSH, os dicionários online Priberam, Dicio, Michaelis e Sinônimos para a coleta das definições dos termos selecionados, sinônimos e área do conhecimento. O contexto de uso dos termos foi obtido através do recurso *File View* do software AntConc. O conteúdo e a análise das fichas terminológicas serviram de base para a elaboração da proposta de árvore de domínio.

## 2 PROCESSAMENTO DE LINGUAGEM NATURAL

O Processamento de Linguagem Natural, PLN, em português, ou *Natural Language Processing* (NLP, sigla em inglês), é a área da Ciência da Computação que estuda o desenvolvimento da capacidade dos programas de computadores reconhecerem, analisarem, interpretarem e até mesmo gerarem textos em linguagens humanas.

Por Linguagem Natural, faz-se necessário esclarecer a diferença entre Linguagem e Língua. De acordo com Benveniste (1991), “Linguagem” abrange todas as formas que os indivíduos utilizam para se comunicar, incluindo sons, gestos, desenhos, pinturas, fala e escrita. O termo mais adequado para essa área de estudo, portanto, seria Processamento de Língua Natural, uma vez que a Ciência da Computação trabalha com os textos e os sons da língua falada. Uma das poucas pesquisadoras que utiliza o termo Língua e não Linguagem, é a Profª. Maria das Graças Volpe Nunes, do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (ICMC – USP), em São Carlos.

Para Nunes (2007, p. 3), “[...] em PLN buscam-se soluções para problemas computacionais, ou seja, tarefas, sistemas, aplicações ou programas, que requerem o tratamento computacional de uma língua natural (português, inglês, etc.), que seja escrita (texto) ou falada (fala)”. A professora explica também que “PLN é quase sinônimo de processamento de língua escrita” e um dos obstáculos enfrentados é “não poder contar com recursos importantes para a solução de problemas, como a entonação, o volume da voz, entre outros.”. Ela ainda complementa que os trabalhos que estudam a fala são tratados pela área de Processamento de sinais dentro da Engenharia Elétrica, por questões tecnológicas; e os trabalhos envolvendo línguas naturais alternativas (como as línguas de sinais) são objetos de estudo que envolvem alguma forma de automatização e processamento/análise de imagens e/ou vídeos.

O que diferencia o processamento de uma linguagem do processamento de outros dados é o conhecimento da língua e a capacidade de extrair informações que estão incorporadas no contexto ou até mesmo que levam em conta a inferência de outros textos e/ou outras fontes de informação. Por exemplo, “para que um programa consiga contabilizar a quantidade de palavras em um arquivo, é preciso conhecimento sobre a definição de palavra e assim ele se torna um sistema de processamento de língua” (Jurafsky; Martin, 2008, p. 2, tradução nossa).<sup>2</sup>

---

<sup>2</sup> No original: “[...] when it is used to count the words in a file, it requires knowledge about what it means to be a word and thus becomes a language processing system.” (Jurafsky; Martin, 2008, p. 2).

## 2.1 Marcos históricos e paradigmas do PLN

Historicamente, o PLN pode ser dividido em grandes marcos, ou períodos, de acordo com o passar das décadas. Após a Segunda Guerra Mundial e o desenvolvimento da “Máquina de Turing” para decifrar (e traduzir) os códigos das mensagens utilizadas pelos exército nazista, um dos impulsos para o desenvolvimento do processamento das línguas naturais foi a automatização da tradução de textos.

Nas década de 1950, surgiram os primeiros tradutores automáticos, abrindo caminho para que, na década 1960, fosse criado o primeiro de inteligência artificial da História capaz de responder perguntas e gerar textos: o ELIZA. Este *chatbot*<sup>3</sup> foi desenvolvido no MIT por Joseph Weizenbaum e tinha como objetivo ser um programa de conversa entre um psicólogo e seu paciente.

Nos anos 1970, os estudos linguísticos se concentraram mais em definir regras, gramáticas e, a partir daí, surgiram 4 paradigmas: o simbólico, o estatístico, o neural e o híbrido (Caseli; Nunes; Pagano, 2023).

O paradigma simbólico foi o padrão que existiu até a década de 1980, e seguia a regra de que todo conhecimento linguístico era expresso por regras gramaticais e lógicas comprehensíveis ao ser humano. Na década de 1990 e o aumento da memória e da capacidade de processamento dos computadores trouxe o paradigma estatístico, no qual grandes quantidades de texto (*corpus*) passaram a ser utilizadas para “ensinar” as máquinas a identificarem determinadas sequências (regras) e representar probabilidades de ocorrências. Foi nessa época, até a década de 2010, que o PLN se destacou para a tradução automática.

O final dos anos 1990 e o início do novo milênio nos anos 2000 trouxeram a popularização dos computadores pessoais e da Internet, ocasionando a demanda da automatização de tarefas voltadas também para o cotidiano das pessoas. O aprendizado de máquina (ou *machine learning*, em inglês) ganhou espaço, e, na década de 2010, o paradigma neural tornou-se o foco das pesquisas da área da computação. Este é o paradigma mais utilizado nas tarefas de PLN, pois consegue trabalhar com enormes quantidades de dados (valores numéricos, não somente texto, símbolos ou regras) e em vários níveis de complexidade. Agora, na década de 2020, o que predomina é o paradigma híbrido, que associa o simbólico com qualquer outro paradigma para tentar suprir a necessidade de explicações dos algoritmos e se obter o resultado desejado.

---

<sup>3</sup> Chatbot – programa que simula uma conversa com pessoas.

Com a análise deste breve histórico, nota-se que os eventos mais marcantes do Processamento de Linguagem Natural ocorreram quando os objetivos de realizar tarefas envolvendo textos e respostas complexas (aqueles que se caracterizam por seguir uma estrutura gramatical e que façam sentido) foram alcançados, sempre com o intuito de promover a interação do ser humano com a máquina e na realização/automatização de tarefas. No entanto, cada grande avanço no processamento de dados linguísticos necessita de etapas que realizam tarefas mais básicas, mas não menos importantes para o resultado final. Todas as aplicações computacionais que processam texto, mesmo as mais básicas, podem se beneficiar das técnicas de PLN.

## 2.2 Aplicações e desafios das técnicas de PLN

Para se conseguir atender a demanda do processamento linguístico, foi necessário desenvolver técnicas específicas para analisar cada aspecto das línguas naturais. Tais aspectos separam-se entre a língua falada e as estruturas fonéticas e fonológicas, e a língua escrita, com a análise das classes gramaticais das palavras – estudadas pela morfologia e a sintaxe e o significado das palavras e expressões compostas pelas palavras, área da semântica e da pragmática/o discurso. Pode-se dizer que a complexidade aumenta de acordo com as estruturas analisadas.

Antes de um computador iniciar o processamento da linguagem, é preciso definir o que os programas precisam identificar – palavras, expressões, frases e etc. Esta etapa é denominada de pré-processamento, e consiste em padronizar o texto (*i.e.*: diferenciação de letras maiúsculas e minúsculas, pontuação e caracteres especiais), assim como detectar e eliminar palavras comuns e sem relevância semântica para a análise estatística dos textos (*i.e.*: artigos e preposições, também denominadas de *stopwords*). O pré-processamento possui três níveis de reconhecimento: o morfológico, que analisa a forma da palavra, como raízes e afixos para entender sua estrutura e seu significado; o sintático, que verifica a estrutura e a função gramatical das sentenças e como as palavras se relacionam uma com as outras; e o semântico, que verifica o significado das palavras e como elas devem ser combinadas para formar frases e sentenças com sentido.

No nível morfológico destacamos a técnica de *bag of words* (Saco de Palavras em português), que representa o texto como uma contagem das palavras presentes, sem levar em consideração a ordem da ocorrência. Para o nível semântico, temos a *word embedding* (incorporação de Palavras), que permite representar palavras como vetores numéricos, capturando a semântica; em resumo, é estabelecida uma sequência de palavras que ocorre

sempre da mesma maneira, mostrando que há um padrão que precisa ser seguido para haver sentido na frase.

Existem duas técnicas que podem ser aplicadas em todos os níveis de reconhecimento: os *n-grams*, que são sequências de um determinado número de itens (*n*) de uma amostra de texto ou fala. Estes itens podem ser letras (caracteres), morfemas/fonemas, sílabas ou palavras. Já a técnica de *Tokenização* seria o oposto dos *n-grams*, pois consiste em dividir um texto em unidades menores, chamadas *tokens*; os *tokens* podem ser palavras, frases ou até mesmo caracteres, tudo depende do nível de especificidade desejado. Enquanto nos *n-grams* se estabelece um padrão de busca partindo do um padrão, na *tokenização* observa-se o texto primeiro e depois se define qual o parâmetro de pesquisa.

De acordo com Caseli, Nunes e Pagano (2023), além do uso do Processamento de Linguagem Natural na Tradução Automática de textos, outras aplicações das técnicas merecem destaque: corretor ortográfico ou grammatical (sistemas de auxílio à escrita), assistentes virtuais (*chatbots*), sistemas de sumarização automática, sistemas de recomendação de produtos, músicas e/ou vídeos em sites de *e-commerce* ou de entretenimento, sistemas de classificação textual, sistemas de detecção de *fake news* e sistemas de recuperação de informação. A Professora Nunes reforça a importância do PLN para a Recuperação da Informação (RI):

[...] esses sistemas se destinam a recuperar documentos relevantes a uma dada consulta, entre uma coleção de documentos. Com a Web, essa aplicação ganhou muito destaque, tendo o sistema de consulta Google como seu mais conhecido representante. Se antes a consulta deveria obedecer a linguagens rígidas de busca (como o SQL), hoje em dia, o usuário é livre para formalizar sua consulta em LN, podendo até parametrizá-la de modo a aperfeiçoar a busca. **Quando a RI trata a consulta como um padrão qualquer, dizemos que não se trata de PLN. Mas quando a consulta é vista e processada como um fenômeno linguístico, requerendo qualquer tipo de processamento básico, então dizemos que se trata de uma RI linguisticamente motivada, e, portanto, de uma aplicação do PLN** (Nunes, 2008, p. 5-6).

Para tanto, são necessários **recursos linguísticos-computacionais**, cujo planejamento e construção constituem tarefas por si só. Exemplos desses recursos são corpora, léxicos ou dicionários, ontologias e gramáticas. A Professora também afirma que é preciso pensar nos problemas de PLN, bem como a busca por soluções, da mesma forma que a Linguagem Natural: de acordo com o contexto da aplicação, o usuário alvo e a tecnologia existente. Assim como na Biblioteconomia e CI e os estudos de usuário, pode acontecer de o usuário ser subestimado, criando algo mais simples e depois descobrir que o que foi desenvolvido não atende às necessidades reais, ou superestimar o usuário, criando algo muito mais complexo do que se precisa, fazendo com que o usuário não se adapte ao sistema e não havendo interação humano-computador.

Embora muitos desafios tenham sido superados, o PLN ainda necessita avançar nos quesitos de precisão, relevância e interpretação. A ambiguidade é uma das barreiras para o aprendizado das línguas naturais pelas máquinas, bem como os neologismos, expressões idiomáticas — que variam de região para região dentro de um mesmo país, de um mesmo estado e até mesmo dentro de uma cidade — mostram que as línguas naturais estão vivas e em constante mudança. Essa dinâmica das relações humanas traz à tona a necessidade de que os bancos de dados dos computadores sejam atualizados constantemente para que possam reconhecer as interações dos seres humanos.

### 3 LINGUÍSTICA DOCUMENTÁRIA E TERMINOLOGIA

No campo de estudo/domínio da Biblioteconomia e da Ciência da Informação também é necessário trabalhar com línguas, conceitos e termos, que são estudados na Linguística Documentária (LD) e na Terminologia. Lara e Tálamo (2007, p.2) definem a Linguística Documentária como “um subcampo no interior da Ciência da Informação que visa o estabelecimento do corpo de princípios teórico-metodológicos da linguagem documentária entendida como subconjunto da língua geral.” A LD é uma linguagem artificial que usa a linguagem natural para a determinação de termos técnicos de especialidade para cada área do conhecimento.

Cunha (2008, p 227), em seu Dicionário de Biblioteconomia e Documentação, define a Linguagem Documentária como “[...] 2. Conjunto organizado de termos normalizados, empregados para representar o conteúdo dos documentos com a finalidade de memorizá-los para pesquisas posteriores [...]”.

Para Maimone e Tálamo,

[...] tanto a Terminologia como a Linguística contribuem de modo significativo para organização e representação da informação documentária, uma vez que a primeira contribui metodologicamente nas relações *conceito-termo*, *termo-conceito* e *sistema de conceitos* e a segunda auxilia nos modos de ‘estruturação’ das linguagens documentárias em relação às definições e organização dos elementos linguísticos na referida estrutura (Maimone, Tálamo, 2011, p. 8).

Por trabalhar com dados (como sinônimo de informação), a Biblioteconomia e a CI podem se beneficiar do uso dos programas de computador e outras técnicas, como o Processamento de Linguagem Natural. Como já foi visto na seção 2.2, as tarefas possíveis de serem automatizadas dentro da Ciência da Informação são a criação de resumos, a sumarização e a indexação. No entanto, para que a sumarização e a indexação possam ocorrer, é preciso que sejam estabelecidos termos para vocabulários controlados e tesouros.

Cunha (2008, p. 360) descreve que um termo para a Linguística é uma “[...] palavra (ou expressão) empregada para denotar um conceito. O termo, não obstante indicar um assunto simples (conceito simples), pode ser composto por mais de uma palavra” e para o contexto da Biblioteconomia, é uma “[...] palavra (ou expressão) empregada para a inclusão temática de um item em sistemas de informação e para a recuperação posterior [...]”.

Compreendendo então o conceito de termo, parte-se para a definição de Terminologia:

LING 1. “Área do conhecimento, inter e transdisciplinar, que trata dos conceitos e de suas representações (termos e símbolos). 2. Disciplina que trata da formação e atribuição de nomes (nominação) aos conceitos, quer nas áreas temáticas

específicas, quer no conjunto de todas as áreas do conhecimento” (WN, 11-21). 3. Conjunto de termos que representa o sistema de conceitos de uma área específica do conhecimento. 4. Conjunto organizado de termos técnicos ou científicos, em uma ou mais línguas. Os significados desses termos são explicados ou definidos de modo apropriado” [...] 5. Publicação, na qual o sistema de conceitos de uma área específica do conhecimento é representado por termos.” (Cunha, 2008 p. 360).

Para Lima (2004), a Terminologia é uma disciplina interdisciplinar que se ocupa do estudo dos termos, sua origem, formação, significado e uso. Ela também se preocupa com a padronização dos termos, para garantir a sua aceitação e compreensão por parte de todos os especialistas de uma determinada área, sendo assim importante para a sociedade como um todo, pois contribui para o desenvolvimento do conhecimento científico e tecnológico. Além disso, os estudos terminológicos são fundamentais para a comunicação e a disseminação da informação, pois permitem que os especialistas compartilhem seus conhecimentos com o público em geral.

Por fim, Krieger (2004, p. 58) reforçam que “[...] os estudos teóricos e aplicados de Terminologia validam-se na medida em que descrevem e recolhem os termos em seus reais contextos de ocorrência [...].” e diz que o termo

reveia sua funcionalidade no processo comunicacional, configurando-se como elemento de representação e divulgação do conhecimento. [...] [e que] é uma unidade léxica que assume um valor semântico próprio de uma área de conhecimento, pela razão de integrar uma comunicação especializada. A despeito desse valor especializado, é um elemento natural das línguas naturais comportando variação e sinônimia, bem como outros efeitos da linguagem em funcionamento (Krieger, 2004, p. 61).

Maimone e Tálamo (2011) complementam que

Ao proceder com o funcionamento social de uma área especializada do conhecimento, a Terminologia trabalha com o estudo científico dos conceitos e respectivos termos, que constituem um conjunto expressivo e comunicativo, possibilitando a transferência do conhecimento especializado (Maimone; Tálamo, 2011, p. 5).

Com essas definições, temos que a Terminologia é interdisciplinar e é obtida dentro de contextos reais de uso da linguagem natural. A partir dos contextos de uso é que se verifica quais os termos devem ser estabelecidos como Termo Geral e quais serão os Termos Relacionados.

Portanto, em uma sociedade informacional cada vez mais conectada, onde a velocidade é ritmo predominante, faz-se necessário que as organizações que trabalham com o processamento e a recuperação da informação acompanhem este ritmo acelerado e possam entregar os resultados de forma imediata. O PLN tem grande potencial para auxiliar a tarefa da Terminologia, uma vez que é capaz de extrair informações relevantes por meio da análise automática dos textos, permitindo sua rápida elaboração e constante atualização.

Um uso pertinente do PLN na construção de uma terminologia especializada seria na área da saúde, pois é fundamental para a sociedade e a produção de conteúdo e de pesquisa são constantes, além de ser uma área que precisa ser colaborativa e dinâmica tendo em vista que o número de interessados que não possuem formação médica tem aumentado de forma bastante expressiva depois do ano de 2020. Apesar de a área da saúde possuir uma terminologia muito bem estabelecida, faz-se necessário o diálogo entre os termos característicos de determinada área do conhecimento e os termos que a população em geral utiliza no dia a dia, especialmente se for levado em consideração que a internet tem se tornando cada vez mais acessível às pessoas, possibilitando que a sociedade como um todo realize pesquisas sobre diversos assunto e, para isso, serão usadas as palavras, os termos, que são de conhecimento popular, e não os termos inerentes ao vocabulário médico.

O próprio DeCS/MeSH tem exemplos de como um termo técnico especializado também tem um termo alternativo, embora, muitas vezes, os termos alternativos também sejam técnicos, não cumprindo plenamente a função de disponibilizar todos os termos que podem estar relacionados ao termo preferido. Considerando que na relação profissional-paciente é fundamental passar informações de forma clara, é preciso considerar que os médicos, enfermeiros, entre outros profissionais da saúde, tenham uma fonte de informação precisa, confiável e de fácil acesso também para os termos conhecidos pela maior parte da população – ou ao menos para aquela parcela da população que ele(a) está atendendo. O DeCS/MeSH precisa acompanhar a velocidade com que as palavras, frases e construções de linguagem são constituídas nos atendimentos médicos-ambulatoriais dos dispositivos de saúde e por seus trabalhadores.

De acordo com Nunes (2008) os desenvolvedores de programas de PLN focados na Tradução Automática (TA) se depararam com uma tarefa muito mais difícil do que se esperava. Verificou-se que o processo de se traduzir um texto envolve muitas etapas e cada uma deve ser tratada separadamente, pois todas possuem características específicas. Esse problema da TA trouxe uma forma de solucionar essa dificuldade: fragmentando o problema determinando sua complexidade, avaliando os recursos disponíveis e a contribuição de cada etapa e combinando as soluções de cada parte. A Professora ainda menciona que mesmo se buscando pela automatização dos processos, ainda cabe a um falante humano avaliar os resultados dos sistemas de PLN, e isso abre espaço para um campo de pesquisa específico para esta atividade. Tal necessidade mostra que o ser humano ainda é — e nos resta esperar que continue sendo — necessário no processo de avaliação da automatização de qualquer tarefa, uma vez que é o

usuário final. A avaliação dos sistemas de PLN já é um campo de estudo por si só devido a sua complexidade.

Para tentar compreender um pouco da importância e da necessidade da Terminologia, existem normas internacionais para se estabelecer termos. Neste trabalho será usada a norma ISO 704:2009, referente aos princípios e métodos do trabalho terminológico. A norma aborda os passos para o trabalho terminológico, como a identificação de objetos, conceitos, suas relações e representação desses aspectos por meio de designações e definições. O documento ainda enfatiza a importância de conhecimento especializado da área e indica um passo a passo para o fazer terminológico, que vai desde a identificação do contexto/assunto, passando pelas propriedades dos objetos para determinar as características que podem ser extraídas e combinadas para formar um conceito e analisar as relações entre os conceitos para organizá-los em um sistema.

Embora a ISO 704:2009 descreva as definições e os processos que envolvem o desenvolvimento do trabalho terminológico, a norma não indica **como** se fazer a coleta dos termos e nem recomenda maneiras de como se chegar até eles.

Mesmo com a descrição da norma ISO, ainda existe uma lacuna entre a necessidade, a norma e o trabalho prático, e foi pensando nisso que a Profª Marilda Lopes Ginez de Lara e a Prof.<sup>a</sup> Maria de Fátima Gonçalves Moreira Tálamo, enquanto docentes do Departamento de Informação e Cultura da Escola de Comunicações e Artes da Universidade de São Paulo, propuseram uma metodologia didática e prática para o ensino de Terminologia (Lara; Tálamo, 2007). As Professoras Lara e Tálamo mostram uma sugestão de como criar uma terminologia a partir de termos já selecionados e ensinam como fazer fichas terminológicas, nas quais são descritos termos preferidos, termos específicos, termos relacionados, definições, descrições, contextos de uso e a estrutura hierárquica, as Professoras também demonstram a criação de uma árvore de domínio baseada nos termos analisados.

Tendo em vista a importância de se criar terminologias de acordo com as necessidades que vão surgindo, o uso da *web* mostra-se uma alternativa viável para a coleta de informações em grande quantidade e até mesmo sendo possível realizá-la de forma colaborativa e com um retorno de respostas em curto período de tempo. Além da internet, *softwares* de coleta e análise de *corpus* também são recursos possíveis de serem utilizados na recuperação e organização de conteúdo. Na seção a seguir serão exploradas algumas técnicas utilizadas na Linguística de *Corpus*.

## 4 LINGUÍSTICA DE *CORPUS*

A Linguística de *Corpus* (LC) é definida por Sardinha (2000) da seguinte forma:

A Linguística de *Corpus* ocupa-se da coleta e exploração de corpora, ou conjunto de dados linguísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem através de evidências empíricas extraídas por meio de computador (Sardinha, 2000, p. 325).

Essa abordagem empirista da LC pressupõe que a linguagem é um sistema probabilístico, o que significa que os traços linguísticos, embora sejam possíveis inúmeras combinações na teoria, elas não ocorrem com a mesma frequência e algumas não são usadas, não são consideradas corretas. O autor também diz que o *corpus* é uma fonte de informação primordial — uma vez que faz o registro da língua que de fato é utilizada por falantes e escritores — e que a investigação da frequência da ocorrência dos traços linguísticos não deve ser abordada como trivial, pois é através da comprovação da frequência dos traços lexicais, sintáticos, semânticos, discursivos e etc., que se torna possível fazer uma estimativa da probabilidade da teoria a ser estudada.

O uso do computador para auxiliar a tarefa da coleta e análise de *corpus* foi essencial para que se conseguisse coletar, organizar e analisar uma grande quantidade de corpora. Essa abordagem empirista da LC se encaixa no Paradigma Estatístico do PLN, como visto na seção 2.1 deste trabalho.

### 4.1 Aplicações do PLN na LC

A Recuperação da Informação trabalha com a premissa de que precisa ser eficaz e eficiente. Os softwares de PLN conseguem auxiliar na eficácia, pois são capazes de reunir todas as pesquisas desejadas em uma única busca/tarefa, o que, consequentemente, torna a recuperação da informação mais eficiente, uma vez que irá poupar o tempo do pesquisador, e, ainda para mostrar a importância do uso do PLN, a Professora Maria das Graças Volpe Nunes complementa que:

[...] a qualidade do conhecimento extraído depende diretamente da preparação ou pré-processamento do texto em estado bruto. E aí entra em cena o PLN. Informações morfossintáticas e semânticas associadas ao texto incrementam consideravelmente a extração do conhecimento, em número e qualidade (Nunes, 2008, p. 10).

Apenas o uso de palavras-chave não são o suficiente para que se obtenha resultados satisfatórios. O que os Bibliotecários aprenderam com sua prática diária da profissão, também

é descrito por Nunes em pesquisas que mostram que é preciso verificar o contexto em que as palavras estão inseridas, ou seja, como os termos ocorrem dentro do texto.

Para Biderman (2001), ao avaliar a qualidade e a relevância de *corpus*, além de considerar qual a função o termo utilizado desempenha, é preciso ele seja equilibrado e represente os diferentes usos da língua, sendo, portanto, necessário que haja uma proporção balanceada dos diferentes tipos de texto e/ou de temas nele incluídos, bem como dos diferentes gêneros e variedades dos usos linguísticos. Além disso, o *corpus* deve ser atualizado constantemente e ter uma configuração que permita que ele seja reutilizável.

A autora defende que o desenho de um *corpus* deve ser cuidadoso para garantir que ele seja uma representação fiel da língua. Isso é importante para que o *corpus* possa ser usado para extrair generalizações sobre a língua e para melhorar a qualidade de produtos linguísticos, como dicionários e vocabulários, e finaliza dizendo que “A Lexicografia e a Terminologia são indubitavelmente as áreas do conhecimento lingüístico que mais se beneficiaram com as informações propiciadas pelos *corpora* de língua escrita e falada pelas ferramentas computacionais. [...]” (Biderman, 2001, p. 92).

#### **4.2 Programas de coleta e análise de *corpus***

A Linguística de *Corpus* foi evoluindo de acordo com o avanço da tecnologia, portanto, sua história se desenvolve a partir das tecnologias disponíveis para exploração, armazenamento e análise de *corpus* – tanto em tamanho quanto em diversidade e rapidez (Sardinha, 2000).

O Processamento de Linguagem Natural propiciou a otimização de programas de computador específicos para a análise de *corpora* textuais. No entanto, esses programas possuem a versão somente em língua inglesa, fazendo com que o pesquisador/usuário tenha algum conhecimento do Inglês ou precise de um treinamento para aprender os recursos oferecidos.

A próxima sessão será dedicada aos procedimentos metodológicos e aos programas mais conhecidos na Linguística de *Corpus*: o BootCat e o AntConc.

## 5 PROCEDIMENTOS METODOLÓGICOS

A busca no DeCS/MeSH ocorreu no dia 25 de fevereiro de 2024 e a pesquisa dos sites com o BootCaT foi feita entre os dias 25 e 29 de fevereiro de 2024.

Após este procedimento, o programa AntConc foi utilizado para realizar a análise do conteúdo dos sites coletados pelo *software* BootCat.

Com uma lista de termos identificados pelo BootCaT e pelo AntConc, foi feita uma proposta de terminologia “popular” de doenças negligenciadas, considerando que o *corpus* foi coletado em páginas na internet às quais todas as pessoas têm acesso, e não apenas com acesso especial. A proposta seguiu os passos descritos por Lara e Tálamo (2007) em seu modelo didático por meio de fichas terminológicas adaptadas e pela construção de uma árvore de domínio.

### 5.1 BootCat

O BootCaT “é um programa que automatiza o processo de encontrar textos de referência na web com a finalidade de compor um *corpus* único” (Baroni; Bernardini, 2004, tradução nossa).<sup>4</sup> Os desenvolvedores do programa o descrevem como “uma ferramenta muito útil para trabalhar com tradução e construção de **terminologias**, e informa que o BootCat é usado em cursos/aulas para a construção de *corpora* de diversos tamanhos e especificações. Além de permitir aos usuários diferentes níveis de controle dos parâmetros de busca, o programa possibilita a análise das URLs recuperadas e transforma as páginas em arquivos de texto no formato “.txt”, já limpos dos códigos Html da página”<sup>5</sup>.

Ainda na descrição da ferramenta, os desenvolvedores destacam que os *corpora* são recursos essenciais para aqueles que trabalham com linguagens especializadas e mais frequentemente em áreas onde neologismos e novos termos surgem num ritmo acelerado e uma referência padronizada deve ser composta por conjuntos de textos que sejam direcionados e fáceis de se construir/coletar.

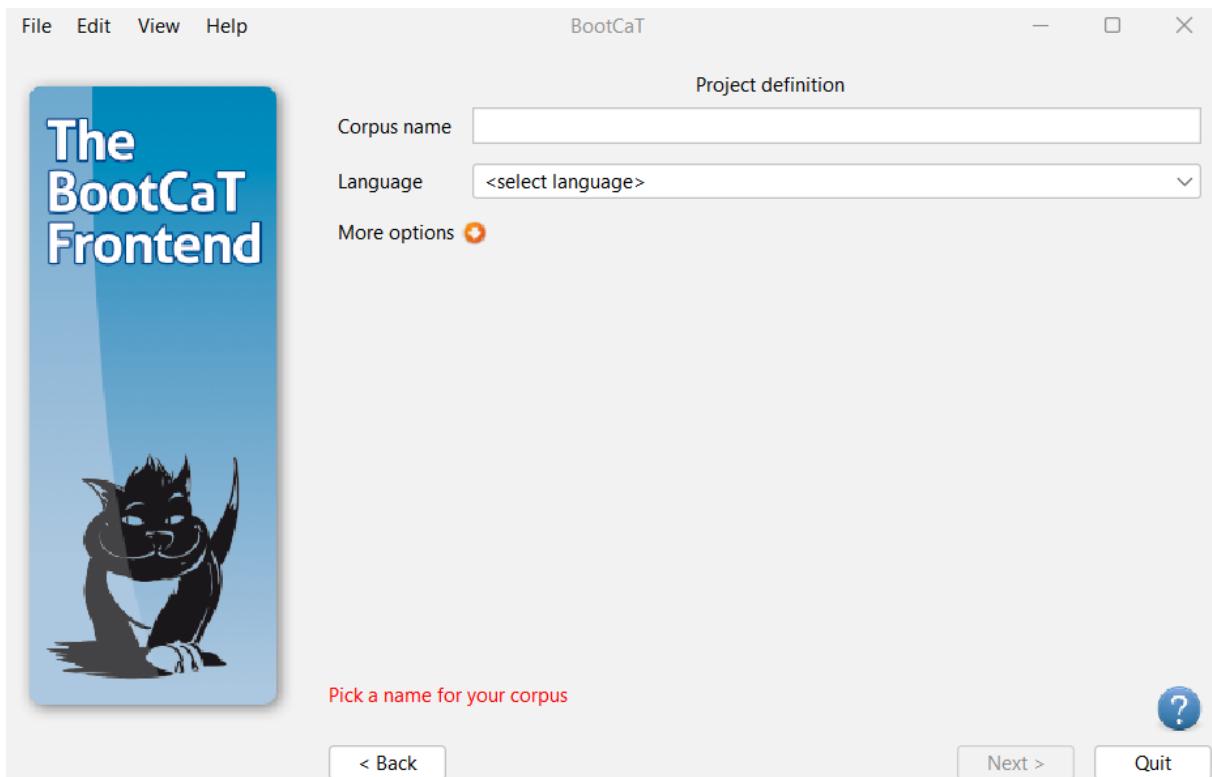
---

<sup>4</sup> No original: “BootCaT automates the process of finding reference texts on the web and collating them in a single *corpus*”.

<sup>5</sup> No original: “The pipeline allows varying levels of control. In the first step, users provide a list of single- or multi-word terms to be used as seeds for text collection. These are then combined into “tuples” of varying length and sent as queries to a search engine, which returns a list of potentially relevant URLs. At this point the user has the option of inspecting the URLs and trimming them; the actual web pages are then retrieved, converted to plain text and saved in “txt” format.

Na tela inicial (Figura 1), é solicitado que o usuário indique um nome para o *corpus* a ser gerado e também traz como primeiro filtro de pesquisa a escolha do idioma a ser utilizado na busca – o programa disponibiliza a busca em mais de 150 línguas, dentre elas o português.

Figura 1 - Tela inicial de parametrização do BootCat

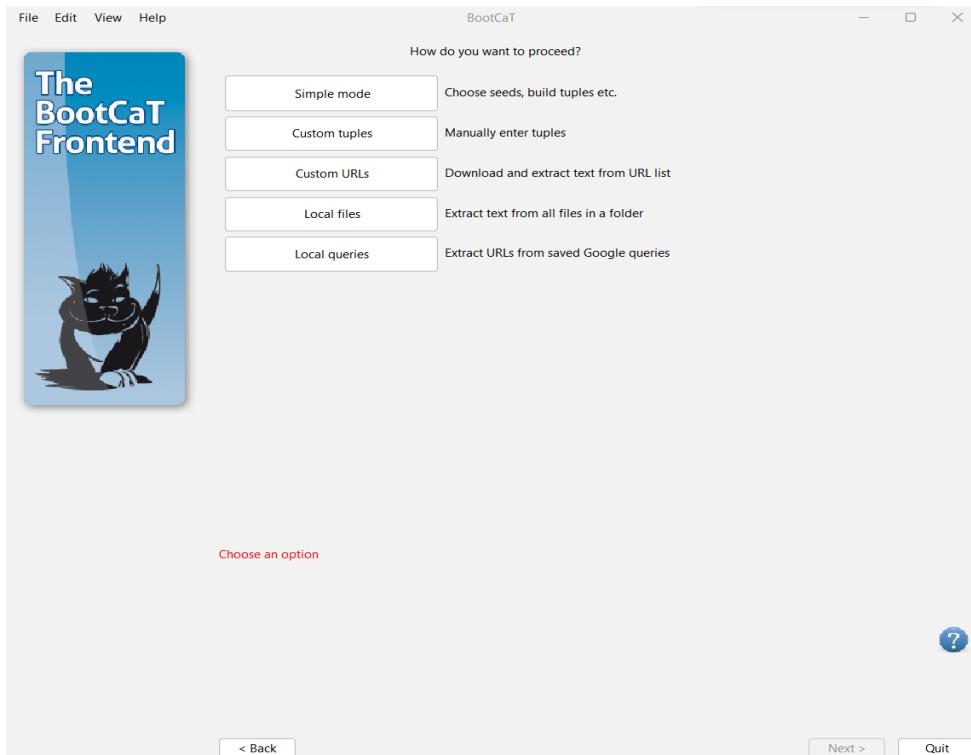


Fonte: a autora

No menu “*More options*” é possível adicionar listas de palavras, em arquivo “.txt.”, para que sejam adicionados filtros nas pesquisas. O BootCat denomina essas listas como “*Blacklists*”, para que não sejam selecionados os resultados com as palavras contidas no arquivo anexado nesta indicação, e também as “*Whitelists*”, fazendo com que as buscas só recuperem páginas que contenham um número mínimo e/ou máximo de palavras daquelas especificadas.

Na tela seguinte (Figura 2), o programa permite a personalização de alguns parâmetros, como, por exemplo, a customização manual das tuplas e em quais tipos de URLs a busca deve ser feita. Este refinamento não foi utilizado, pois o intuito deste trabalho é fazer com o que o *corpus* seja o mais abrangente possível de acordo com o que está disponível na internet. A opção “*Simple mode*” (Modo simples) foi a escolhida.

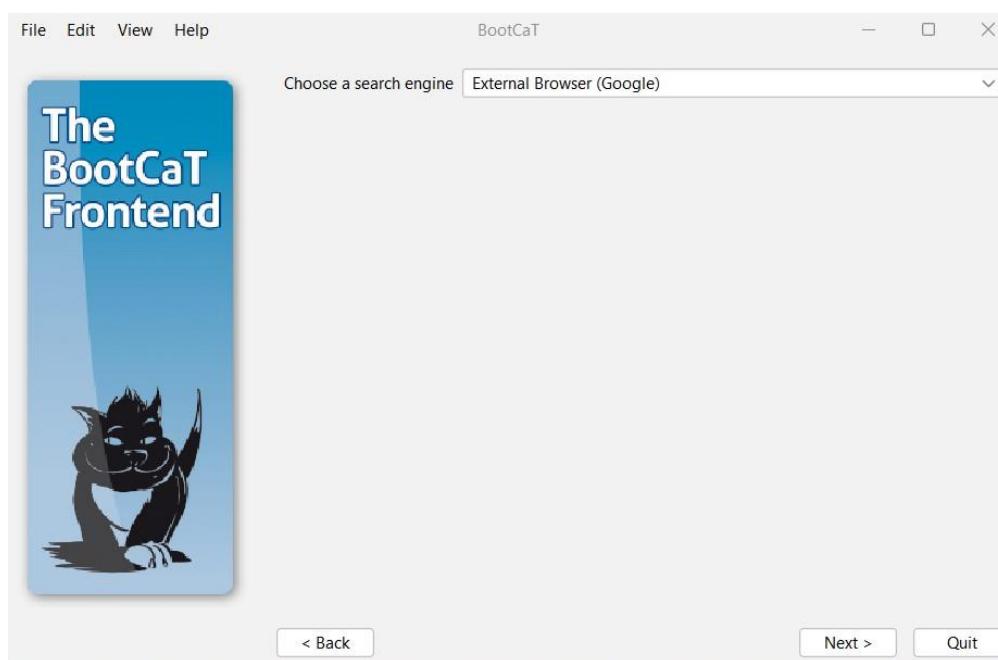
Figura 2 - Tela 2 de parametrização para a busca



Fonte: a autora

Ao selecionar o modo simples, é exibida a tela para a seleção da ferramenta de busca, que neste modo só o Google aparece como opção. (Figura 3):

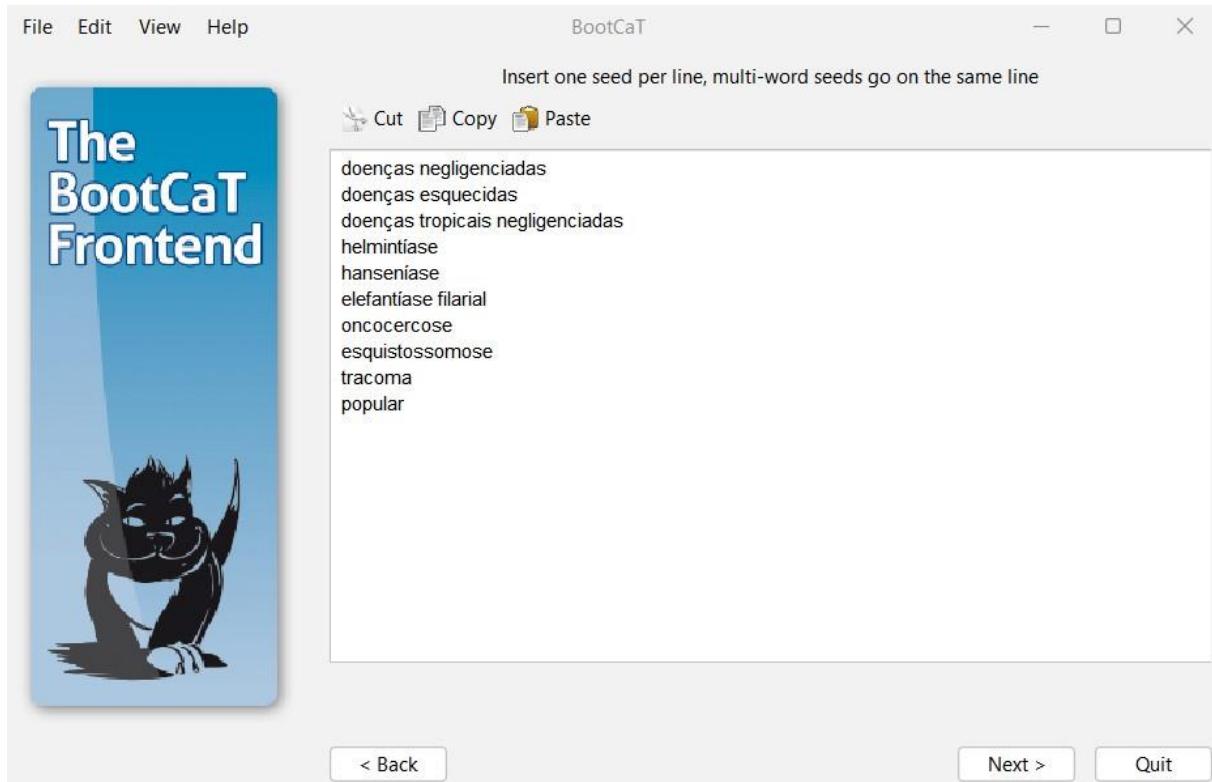
Figura 3 – Ferramenta de busca



Fonte: a autora

A próxima tela do programa (Figura 4) solicita que sejam inseridos os termos a serem buscados. Cada linha corresponde a um termo (*seed*), portanto, se houver mais de uma palavra em uma linha, o próprio BootCaT identifica que é uma expressão ou um termo composto. Neste caso, foram utilizadas dez “sementes” retiradas do DeCS/MeSH a partir do conceito principal de Doenças Negligenciadas: o próprio termo “doenças negligenciadas”, “doenças esquecidas”, “doenças tropicais negligenciadas”, “elefantíase filarial”, “helmintíase”, “hanseníase”, “oncocercose”, “esquistossomose”, “tracoma” e “popular”. O processo de pesquisa dos termos e a definição de Doenças Negligenciadas será abordado no item 5 Procedimentos metodológicos.

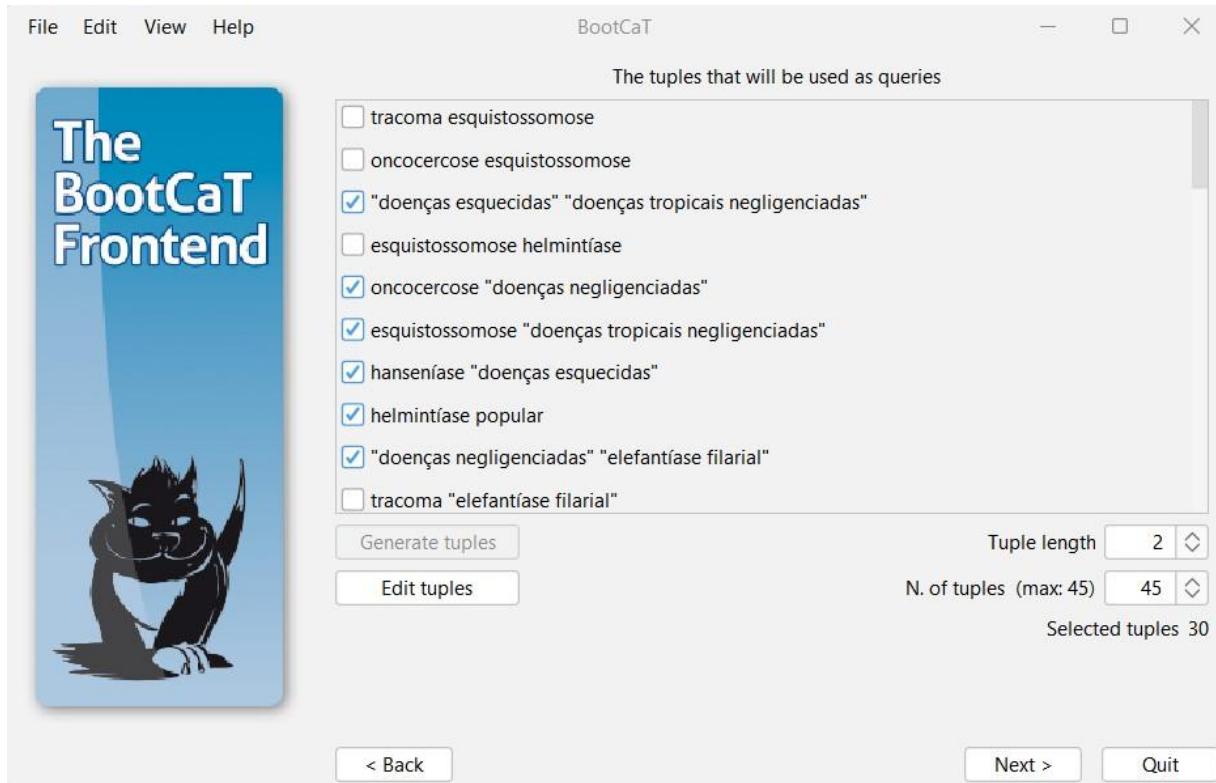
Figura 4 – Tela de identificação dos termos a serem pesquisados



Fonte: a autora

Após informar os termos de busca, o próximo passo é indicar as combinações entre os termos para que haja uma maior abrangência para a pesquisa dos sites. Essas combinações possíveis são chamadas de tuplas (em inglês, *tuple*). Ao indicar a quantidade de termos que irão compor as tuplas, ou seja, o comprimento das tuplas (*tuple length*), o BootCaT já fornece o número máximo de tuplas que poderão ser formadas (Figura 5).

Figura 5 – Tela de seleção de tuplas

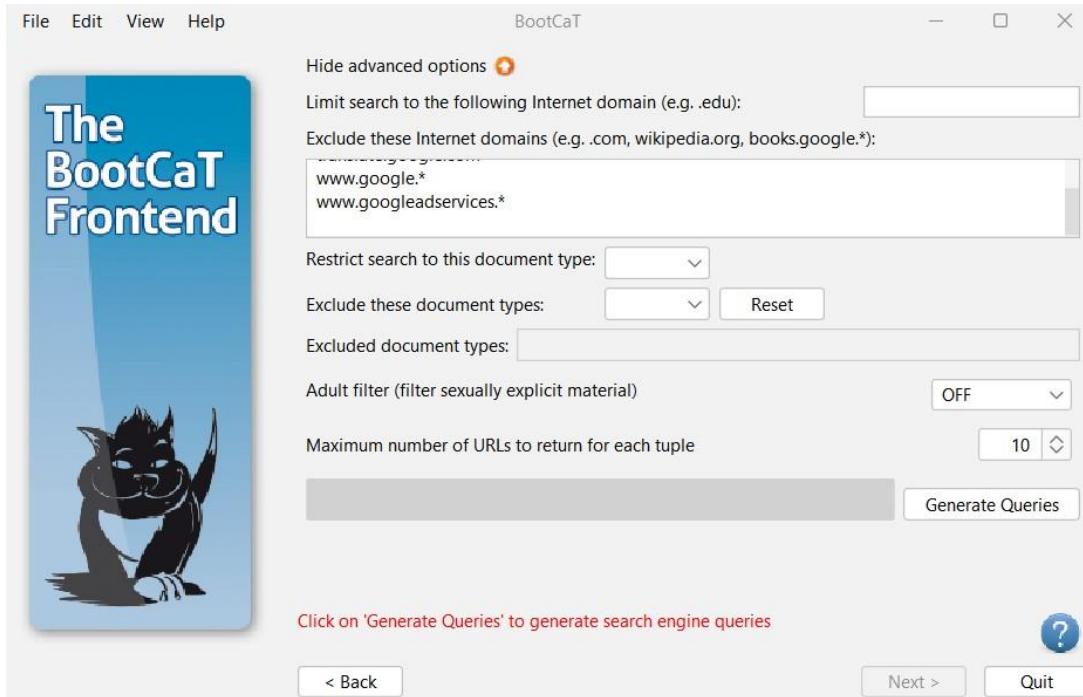


Fonte: a autora

Para esta pesquisa foram utilizados os seguintes parâmetros: tuplas de dois termos e um máximo de 45 tuplas a serem geradas, das quais 30 foram selecionadas (Figura 5). Os critérios de seleção utilizado foi descartar as tuplas compostas por duas denominações de doenças (ex.: tracoma e “elefantíase filarial”).

Outra etapa de parametrização para a coleta do *corpus* são as opções de busca avançada que a ferramenta oferece (Figura 6). É possível limitar o tipo de domínio de sites nos quais será realizada a busca, bem como restringir o tipo de documento existentes e se há conteúdo sinalizado como conteúdo adulto nas páginas da *web* e o número de URLs recuperadas para cada tupla.

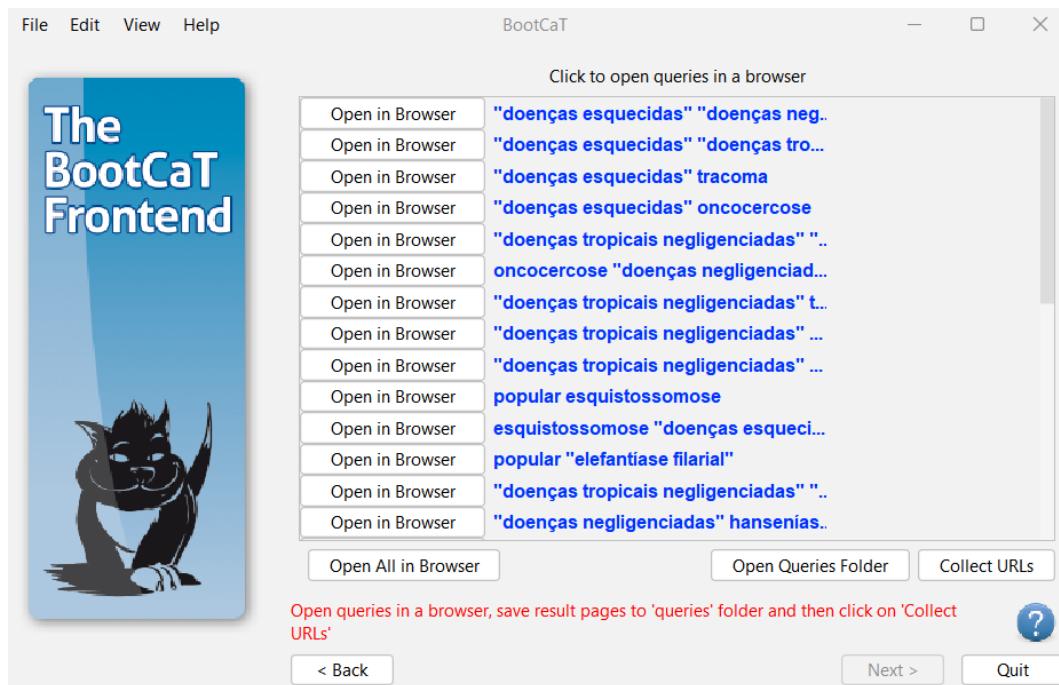
Figura 6 – Tela de refinamento da busca



Fonte: a autora

Para realizar a pesquisa, é só clicar no botão “*Generate Queries*” (Figura 6) e aguardar. A próxima tela traz a lista com os resultados da pesquisa.

Figura 7 – Lista com os resultados da pesquisa

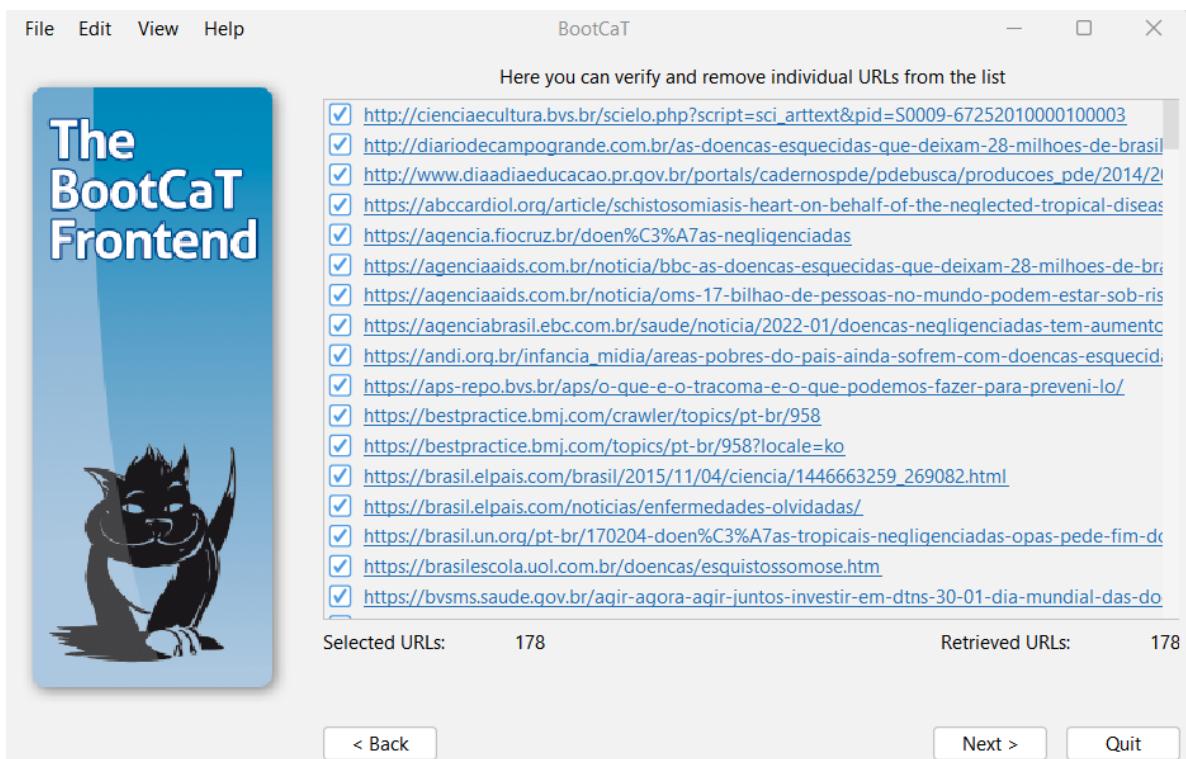


Fonte: a autora

A etapa seguinte exige que o pesquisador abra todas as páginas de resultados no navegador de internet (*Open in browser*), e é possível escolher qual site será baixado ou não. Somente após este processo, o BootCaT permite que as URLs sejam salvas. O próprio programa cria automaticamente uma pasta para realizar a coleta dos links e a conversão dos conteúdos em arquivo .txt.

A Figura 8 mostra que foram recuperadas 178 URLs de acordo com os parâmetros de busca informados. Para este trabalho, todos os resultados foram utilizados.

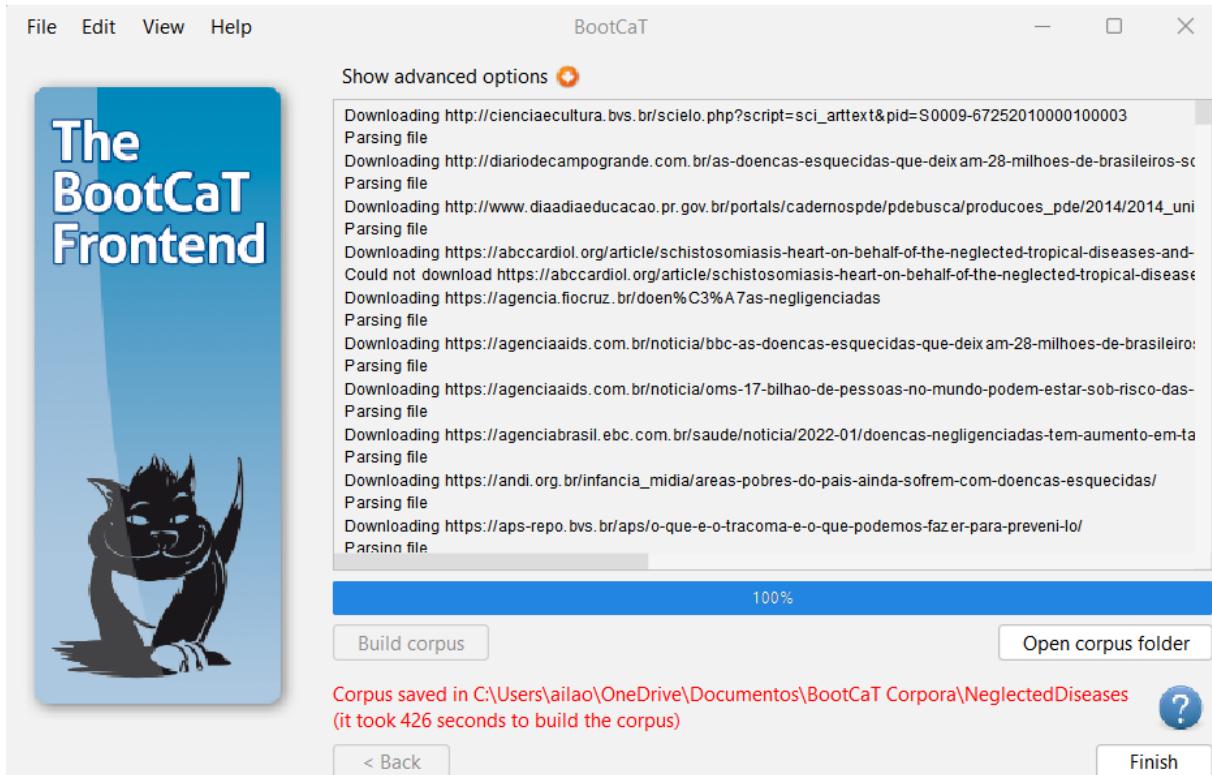
Figura 8 – Tela de URLs recuperadas na pesquisa



Fonte: a autora

A última etapa para a criação do *corpus* é fazer o download do conteúdo das páginas obtidas (Figura 9), o qual o BootCat salva automaticamente no formato .txt.

Figura 9 – Etapa final de elaboração do *corpus* através do BootCaT



Fonte: a autora

Após salvar as URLs do *corpus* levantado, o BootCaT também fornece um relatório (*report*) em Excel com algumas informações das páginas pesquisadas. Ele organiza o arquivo com seguintes itens: *Downloaded file* (Nome do arquivo baixado), *Extracted plain file* (Nome do arquivo extraído), *Extracted XML file* (Nome do arquivo em XML), *Approximate token count* (Número aproximado de tokens), *Character count* (Número de caracteres), URL (o site recuperado), *Redirected from URL* (Endereço de redirecionamento de site), *Downloader* (tipo de download), *Detected languages* (Idiomas detectados nos sites), *Mime type* (tipo de conteúdo, se é texto em HTML ou se é um aplicativo de arquivo PDF), Status (se foi feito o download ou não), *Skipped sentences* (Frases não contabilizadas), *Downloaded file size* (Tamanho do arquivo baixado), *Extracted file size* (Tamanho do arquivo extraído), *HTML Extraction mode* (Tipo de HTML extraído), *Creation date* (Data de criação), e *Download date* (Data de download).

Neste *corpus*, dos 178 resultados obtidos, dez não puderam ser baixados, sendo um deles redirecionado de outro link e outras 13 URLs foram identificadas como “língua errada”, isto é, quando o conteúdo não era totalmente em português, trazendo outros idiomas (espanhol e inglês), e um *link* categorizado como não sendo possível determinar o idioma. Isso mostra que

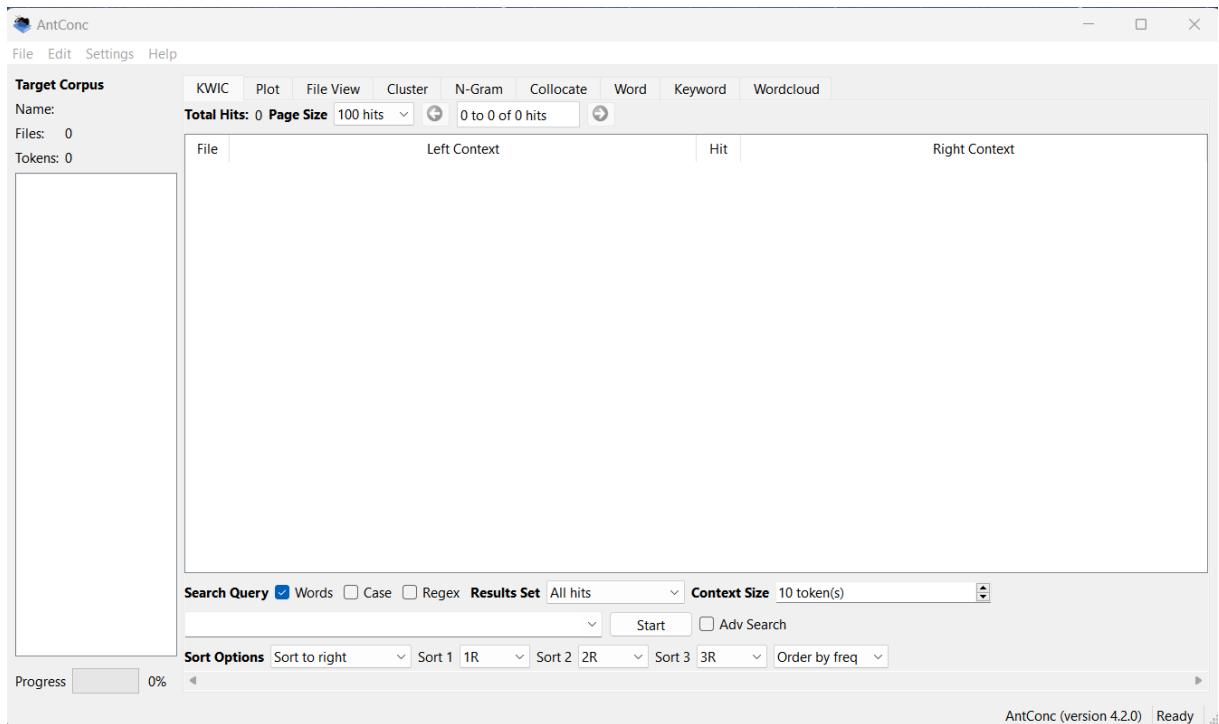
a ferramenta informa tudo o que foi localizado na busca, porém não permite que o usuário faça o download dos arquivos que não seguem os parâmetros estabelecidos.

## 5.2 AntConc

O AntConc é um software gratuito de análise de *corpus*. Além de ser disponibilizado de forma gratuita, um dos atrativos do AntConc é que o programa possui uma interface amigável para a interação humano-computador, não sendo necessário o conhecimento de linguagem de programação para utilizá-lo (Figura 10).

A versão do software utilizada neste trabalho foi a 4.2.0. Dentre as ferramentas disponíveis no programa, é possível verificar a frequência de ocorrência, concordâncias e conjuntos de palavras.

Figura 10 – Tela inicial do software AntConc



Fonte: a autora

Embora o software tenha uma versão somente em inglês, Julia de Souza Borba Gonçalves<sup>6</sup>, na época em que era estudante do curso de Relações Internacionais da Universidade Estadual Paulista “Julio de Mesquita Filho” (UNESP) e participou do Laboratório de Novas Tecnologias de Pesquisa em Relações Internacionais (LANTRI) na mesma

---

<sup>6</sup> Atualmente é doutoranda do curso de Relações Internacionais da Universidade de Brasília (UnB).

Universidade, criou um manual para utilização do AntConc, que está disponível no site do desenvolvedor do programa. Gonçalves o descreve da seguinte forma:

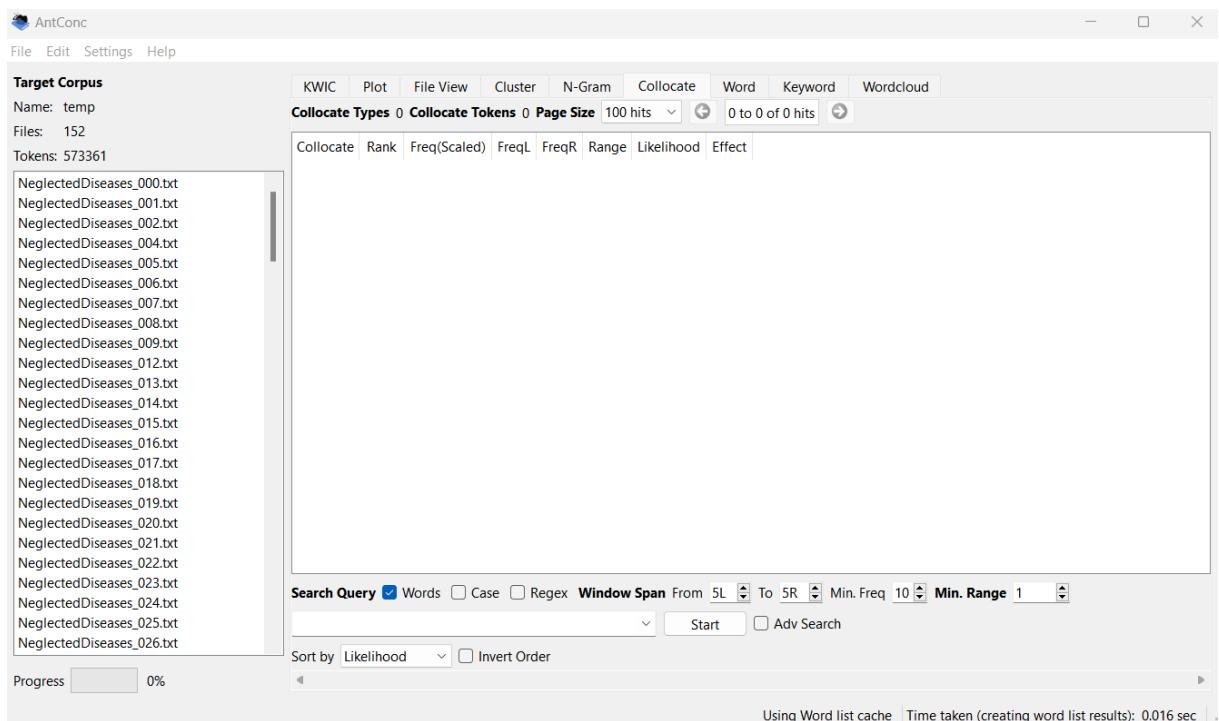
Software gratuito para realizar análise de documentos escritos, podendo ajudar os pesquisadores que fazem análise de discurso. O software oferece opções diversas que permite o pesquisador a fazer uma análise apurada envolvendo grandes quantidades de documentos; dentre essas opções, destacam-se a localização de palavras-chaves em determinados contextos, índice de frequência de palavras e de palavras semelhantes, lista de palavras-chaves e lista de palavras. (Gonçalves, 2016, p. 4)

A leitura do manual descrito acima e as instruções disponíveis em vídeos no canal do criador do *software* no YouTube permitem que até um usuário pouco experiente consiga utilizar o programa sem grandes dificuldades.

## Funcionalidades do AntConc

Para utilizar o AntConc é preciso que o *corpus* a ser analisado esteja em formato “.txt” e inserir os arquivos no comando “*Corpus Manager*”. Para o *corpus* de Doenças Negligenciadas temos 152 arquivos e um total de 573.361 tokens (Figura 11).

Figura 11 – Arquivos do *corpus* no AntConc



Fonte: a autora

A seguir, são destacadas as principais funcionalidades e ferramentas do AntConc:

**KWIC (KeyWord In Context)** – permite que o pesquisador veja as palavras e frases que são mais comumente usadas em um *corpus* textual.

**Plot tool** – mostra os resultados como se fosse um código de barras, mostrando as posições (em qual parte do texto) os resultados da busca ocorrem.

**File view** – nesta aba é possível verificar a localização da palavra ou expressão pesquisada dentro do texto, ou seja, no contexto de uso.

**Clusters/N-grams** – busca por condições/regras específicas, podendo ser em conjuntos (clusters) e os *N-grams* determina o número de palavras (n) que deverá ter no conjunto, isso permite que sejam identificadas as expressões mais comuns num *corpus*.

**Collocates** – é o oposto do que faz os *clusters*, os *collocates* identificam padrões não sequenciais da língua, verificando os termos mais recorrentes e que estejam próximos da palavra indicada, mas que não sejam necessariamente os termos sequencialmente anteriores ou posteriores.

**Keywords** – esta aba possui uma ferramenta que mostra quais palavras são menos frequentes, e possibilita que sejam identificadas palavras características de determinado *corpus*.

**Wordcloud** – com base nos resultados das análises das outras abas, é possível criar uma nuvem de palavras com os termos com mais ocorrências.

A ferramenta *Collocate* se mostrou mais pertinente para este trabalho, uma vez que o objetivo da pesquisa é identificar termos, palavras e expressões que tenham alguma relação com as “doenças negligenciadas”, e não exatamente quais são as palavras antecedem ou vêm depois do termo em questão.

Para cada, o termo “doenças negligenciadas” arquivo foi pesquisado manualmente na ferramenta *KWIC*, e no comando *Collocates* foi selecionado o parâmetro *Sort by “likelihood”* (Ordenar por similaridade). Em três termos, “doenças esquecidas”, “elefantíase filarial” e “helmintíase” foi preciso diminuir a frequência mínima de ocorrências de 10 para 1, pois não havia resultados no primeiro parâmetro. Também foi realizada uma busca apenas com a palavra “elefantíase”, uma vez que a expressão “elefantíase filarial” não recuperou muitos resultados.

A funcionalidade *File view* foi útil para se conseguir o contexto de ocorrência das palavras selecionadas, utilizado nas fichas terminológicas.

Após a verificação do AntConc, foi feita a seleção dos dados para análise e estabelecer um vocabulário prévio, para depois prosseguir para uma proposta de terminologia com as palavras obtidas. Os dados para análise de cada termo estão nas tabelas 1 a 12.

Para esta contagem final, foram retiradas as *stopwords* e outras palavras que não são relevantes semanticamente (normalmente artigos). Na proposta de terminologia deste trabalho, optou-se por utilizar os substantivos, que foram destacados em azul para uma melhor visualização das ocorrências de cada *Collocate* nos documentos (*range*). Depois, as palavras selecionadas foram divididas de acordo com sua classe gramatical, para que fosse possível trabalhar apenas com substantivos.

Substantivo é definido pelo dicionário Michaelis (2024) como “Classe de palavras com que nomeamos os seres em geral, ações, estados e noções”. Optou-se por utilizar essa classe de palavras para facilitar a identificação de relações, uma vez que as ações e outros qualificadores (adjetivos, advérbios, etc.) poderia tornar a terminologia muito ampla e complexa. Para a identificação desses substantivos, foram utilizados os dicionários online Dicio, Michaelis e Priberam.

Após a identificação dos substantivos, seguiu-se com a seleção dos termos candidatos a compor esta pequena terminologia. Para isso, levou-se em consideração a coluna *Range* (alcance), e escolheu-se os termos com um *range* igual ou superior a 10.

## **Doenças esquecidas**

Um dos termos com menos ocorrências nos documentos e as palavras relacionadas também não se mostraram muito recorrentes, aparecendo, em sua maioria, em apenas um documento como mostra a Tabela 1. Foram 21 ocorrências, 56 tokens; sendo 17 selecionadas, das quais 6 são substantivos:

Tabela 1 – Doenças esquecidas

<i>Collocate</i>	<i>Range</i>
Atacam	1
Deixam	2
Ainda	4
Livrar	1
Ajudar	2
desigualdade	2
Angola	1

Agenda	2
Luxo	1
pródigas	1
desprezadas	1
controlar	2
brasileiros	2
indústria	1
condomínios	1
adverte	1
farmacêutica	1

Fonte: a autora

Termos candidatos: 0

### Doenças Negligenciadas

Este termo apareceu 510 vezes nos documentos levantados e trouxe um total de 64 tokens na ferramenta *Collocate*. Dentre esses tokens, foram selecionados 24 termos, sendo 18 substantivos, como é possível ver na Tabela 2.

Tabela 2 - Doenças negligenciadas

<i>Collocate</i>	<i>Range</i>
medicinais	1
plantas	1
derivados	1
federais	1
gastos	1
bibliométrica	1
discussão	2
texto	1
epidemiologia	1
brasil	25
escola	1
medicamentos	14
análise	2
prevenção	3
abordadas	1
combate	9
iniciativa	10
doenças	29
pesquisas	6
problema	12
tuberculose	5
saúde	22

negligenciadas	15
hanseníase	8

Fonte: a autora

Termos candidatos: 6

Brasil, medicamentos, iniciativa, doenças, problema, saúde.

### Doenças Tropicais Negligenciadas

O sinônimo de Doenças negligenciadas (Tabela 3) recuperou 444 ocorrências, 33 tokens, sendo 23 selecionados e 11 substantivos. Aqui ressaltamos o aparecimento de três siglas que foram destacadas em amarelo: OMS, sigla para Organização Mundial da Saúde; DTN, abreviação de doenças tropicais negligenciadas; e DTNS abreviação de doenças tropicais negligenciadas subnotificadas.

Tabela 3 – Doenças tropicais negligenciadas

Collocate	Range
oms	20
Relatório	1
Primeiro	5
Sobre	13
Hoje	4
Mundo	5
Parte	1
Mundial	25
Dia	23
Dtn	8
DTNS	15
Global	3
Superar	1
Tropicais	9
Combate	5
negligenciadas	14
Plano	2
Impacto	1
Doenças	22
organização	9
Grupo	9
Contra	10
Principais	9

Fonte: a autora

Termos candidatos: 2

Dia, doenças.

### **Elefantíase**

Este termo obteve 71 *hits*, 10 *tokens*, 7 termos selecionados, sendo 4 substantivos. Os resultados na Tabela 4 mostraram-se mais específicos dos termos médicos.

Tabela 4 - Elefantíase

Collocate	Range
filariose	19
hidrocele	3
linfática	16
erisipela	1
urina	1
leitosa	1
conhecida	10

Fonte: a autora

Termos candidatos: 1

Filariose.

### **Elefantíase filarial**

“Elefantíase filarial” foi o termo que trouxe palavras menos conhecidas (Tabela 5), e mais específicos da terminologia da área da saúde. Foram apenas 6 correspondências, 11 tokens e todos selecionados para análise. Dos 11 tokens, 8 são substantivos, porém todos possuem um *range* muito baixo, por isso nenhum termos foi escolhido como possível termo candidato.

Tabela 5 – Elefantíase filarial

Collocate	Range
helmintíase	2
wuchereria	1
filariose	2
inquéritos	1
bancrofti	1
anatômicos	1
tricostrongiloidíase	1
Mycetoma	1

elephantiasis	1
cromoblastomicose	1
Micetoma	1

Fonte: a autora

Termos candidatos: 0

### Esquistossomose

O termo que mais recuperou resultados foi esquistossomose, com 1023 correspondências e 72 tokens, dentre os quais foram 57 selecionados. Na Tabela 6 vemos em destaque os termos candidatos. Pontos a serem observados: houve um retorno do próprio termo “esquistossomose” em 16 documentos; este foi o termo que se relacionou com todas as outras seis denominadas doenças negligenciadas analisadas neste trabalho.

Tabela 6 – Esquistossomose

<i>Collocate</i>	<i>Range</i>
Mansoni	16
Diretrizes	3
Técnicas	3
mansônica	10
helmintíases	13
Controle	13
vigilância	4
Filariose	32
Linfática	25
Aguda	9
transmitidas	5
Chagas	33
idiopática	5
Solo	5
leishmaniose	29
associada	9
esquistossomose	16
cardiovascular	6
tratamento	17
hanseníase	37
secundária	5
oncocercose	21
Causada	16
complicação	6
induzida	5

transmissão	10
schistosoma	13
hepatoesplênica	2
malária	20
doença	47
relacionada	9
edição	1
diagnóstico	14
epidemiologia	7
quimioterapia	6
dengue	21
pacientes	9
manejo	5
usados	5
transmitida	13
geográfica	5
tabela	8
raiva	13
introdução	3
envolvimento	6
parasitária	10
pulmonar	5
endêmica	10
pode	14
saúde	16
crônica	10
conhecida	9
cura	8
taxa	7
tracoma	23
resultados	8
Distribuição	8

Fonte: a autora

Termos candidatos: 17

Helmintíases, controle, filariose, chagas, leishmaniose, esquistossomose, tratamento, hanseníase, oncocercose, transmissão, malária, doença, diagnóstico, dengue, raiva, saúde, tracoma.

## Hanseníase

Nesta pesquisa, a hanseníase obteve a segunda maior quantidade de correspondências, com 993 *hits*. Foram identificados 53 tokens, dos quais 39 foram selecionados e 35 são substantivos. Este termo também se recuperou a ele mesmo e também trouxe muitos termos de outras doenças, como mostra a Tabela 7:

Tabela 7 – Hanseníase

<i>Collocate</i>	<i>Range</i>
Guia	2
Cadernos	1
Básica	2
Atenção	5
Controle	12
tuberculose	18
Hanseníase	15
leishmanioses	8
Doença	47
Chagas	25
Malária	12
Casos	28
leishmaniose	26
Dengue	22
esquistossomose	36
Novos	15
Hepatites	10
diagnóstico	15
Virais	8
Cura	11
Pessoas	16
Filariose	22
Raiva	15
História	3
Doentes	3
Sintomas	10
Lepra	8
Contatos	5
Caso	10
Brasil	20
Problema	9
referências	5
Tratamento	19
Sinais	5
eliminação	12

detecção	3
linfática	12
eliminar	5
pessoa	6

Fonte: a autora

Termos candidatos: 21

Controle, tuberculose, hanseníase, doença, chagas, malária, casos, leishmaniose, dengue, esquistossomose, hepatites, diagnóstico, cura, pessoas, filariose, raiva, sintomas, caso, brasil, tratamento, eliminação.

## Helmintíase

A busca inicial por este termo não trouxe muito resultado. Foi um dos casos em que foi necessário diminuir o *range* para um. Com isso, obteve-se 22 *hits*; 18 *tokens*, dentre os quais 17 selecionados e chegando-se a 12 substantivos, elencados na Tabela 8:

Tabela 8 – Helmintíase

Collocate	Range
helmintos	2
enteropatias	2
seleccione	1
filariose	3
obscura	1
deixando	1
animal	1
restrita	1
helmintíase	1
categoria	1
lugares	1
infestação	1
oncocercose	4
parasitas	1
linfática	2
verminoses	2
elefantíase	2

Fonte: a autora

Termos candidatos: 0

## Oncocerose

A palavra “oncocerose” teve 519 correspondências, com 28 dos 41 *tokens* selecionados. Na Tabela 9, temos que desses 28, 18 são substantivos, sendo duas siglas: OCP, abreviação do inglês *Onchocerciasis Control Programme* (Programa de Controle da Oncocerose), e OEPA, *Onchocerciasis Elimination Program in the Americas* (Programa de Eliminação da Oncocerose nas Américas). Por não serem em língua portuguesa, essas duas siglas serão desconsideradas.

Tabela 9 – Oncocerose

<i>Collocate</i>	<i>Range</i>
eliminação	10
programa	7
Africano	3
Cegueira	17
Filariose	25
Linfática	18
Controle	4
Américas	8
Rios	14
Ocp	3
Oepa	5
Foco	6
África	5
Ocidental	3
esquistossomose	21
Brasil	6
oncocerose	7
Tracoma	17
Indígena	4
yanomami	4
parasitária	7
Eliminar	8
transmitidas	5
Focos	3
Humana	5
Onde	6
leishmaniose	9
distribuição	5

Fonte: a autora

Termos candidatos: 6

Eliminação, cegueira, filariose, rios, esquistossomose, tracoma.

## Popular

Foi o termo que recuperou menos resultados, mesmo aparecendo 99 vezes nos documentos, só foram identificados 5 tokens, dentro os quais 4 foram selecionados e todos são substantivos (Tabela 10).

Tabela 10 – Popular

<i>Collocate</i>	<i>Range</i>
participação	3
nome	7
educação	2
controle	2

Fonte: a autora

Termos candidatos: 0

## Tracoma

O último termo, Tracoma, recebeu 843 correspondências, 68 tokens, dos quais 48 foram selecionados e 38 são substantivos. O termo também recuperou a ele mesmo e outras doenças, como é possível ver na Tabela 11:

Tabela 11 – Tracoma

<i>Collocate</i>	<i>Range</i>
eliminação	10
problema	12
tracoma	13
cegueira	17
nacionais	1
folicular	10
inflamatório	8
programas	3
focos	3
mapeamento	3
cegante	1
Ativo	5
recentes	1
países	6
brasil	5
validação	5

conjuntivite	9
causa	13
helmintíases	11
oncocercose	17
fronteira	1
endêmico	6
intenso	6
escolares	3
chagas	23
filariose	17
bactéria	9
global	7
eliminar	6
<b>pesquisa</b>	2
ocular	9
formas	7
epidemiológico	7
situação	5
prevalência	7
Raiva	9
contra	2
prevenção	8
esquistossomose	23
pesquisas	3
Risco	8
pública	10
ações	2
méxico	1
estratégia	3
importante	7
presença	6
busca	2

Fonte: a autora

Termos candidatos: 10

Eliminação, problema, tracoma, cegueira, causa, helmintíases, oncocercose, chagas, filariose, esquistossomose.

Após o levantamento dos termos candidatos, tendo sido obtidos 63 termos, foi feita a organização das palavras por ordem alfabética e a verificação daqueles que se repetiram. Com esta análise, chegou-se a um total de 32 termos: Brasil, caso, causa, cegueira, chagas, controle,

cura, dengue, dia, diagnóstico, doença, eliminação, esquistossomose, filariose, hanseníase, helmintíase, hepatites, iniciativa, leishmaniose, malária, medicamentos, oncocercose, pessoas, problema, raiva, rios, saúde, sintomas, tracoma, transmissão, tratamento, tuberculose.

A organização *corpus* obtido é importante para identificar as características dos termos recuperados e, a partir dessas características, torna-se possível fazer a análise de domínio que vai viabilizar o aprimoramento das ferramentas de construção do conhecimento como as terminologias e ontologias.

## 6 TERMINOLOGIA DE DOENÇAS NEGLIGENCIADAS

Em oposição às doenças mais populares e às doenças mais pesquisadas, como, por exemplo, Coronavírus (termo mais pesquisado no Brasil e no mundo inteiro no ano de 2020, de acordo com o Google Trends), temos também as doenças que, por serem tratadas deliberadamente como menos importantes, são denominadas Doenças Negligenciadas. Esse termo é usado para chamar um grupo de doenças que são conhecidas – algumas sendo até consideradas enfermidades consideradas graves – mas que não despertam interesse para a pesquisa médica.

As Doenças Negligenciadas foram selecionadas como assunto principal para o levantamento do *corpus* a ser analisado pelas ferramentas de análise de *corpus*.

Usando a principal referência de terminologia na área da saúde, o DECS/MESH, ao se realizar a busca pelo termo: “doenças negligenciadas”, obtém-se a seguinte nota de escopo:

Doenças que são subfinanciadas e possuem baixo reconhecimento, mas são as principais ameaças em países pouco desenvolvidos. A Organização Mundial da Saúde apontou seis doenças infecciosas tropicais como negligenciadas em países industrializados que são endêmicas em muitos países em desenvolvimento (HELMINTÍASE, HANSENÍASE, ELEFANTÍASE FILARIAL, ONCOCERCOSE, ESQUISTOSSOMOSE E TRACOMA). (Descritores em Ciências da Saúde: DeCS, 2023)

Os termos alternativos informados/fornecidos no DeCS são “Doença Negligenciada”, “Doenças Esquecidas” e “Doenças Tropicais Negligenciadas”. Sua estrutura hierárquica segue o seguinte esquema: Termo Geral “Condições Patológicas, Sinais e Sintomas”; Termos Subordinados “Processos Patológicos” → “Atributos de Doença” → “Doenças Negligenciadas”.

A pesquisa no DeCS/MeSH forneceu as palavras-chave para a busca com a ferramenta BootCaT, bem como parte das informações contidas nas fichas terminológicas confeccionadas a partir dos termos levantados como termos candidatos.

Após a busca feita com o BootCaT, detalhada no subitem 4.2.1, e a recuperação e análise inicial feita com o software AntConc, foram obtidos um total de 183 substantivos. O próximo passo foi selecionar os substantivos que se repetiram para utilizá-los como termos candidatos a compor a terminologia aqui proposta. Foram encontrados 32 termos com essa característica. A tabela 12, a seguir, mostra a relação dos termos candidatos:

Tabela 12 – Seleção dos termos candidatos

<b>Termo candidato</b>	<b>Termo pesquisado relacionado</b>	<b>Range</b>	<b>Repetições</b>	<b>Singular / Plural</b>	<b>Range total</b>
Brasil	Doenças negligenciadas	25	5	-	56
casos	Hanseníase	28	-	10 singular 28 plural	38
causa	Tracoma	13	-	-	13
cegueira	Oncocercose	17	2	-	34
Chagas	Esquistossomose	33	3	-	57
controle	Esquistossomose	13	4	-	31
cura	Esquistossomose	8	2	-	19
dengue	Esquistossomose	21	2	-	43
dia	Doenças tropicais negligenciadas	23	-	-	23
diagnóstico	Esquistossomose	14	2	-	29
doença	Esquistossomose	47	4	2 singular / 2 plural	145
eliminação	Hanseníase	12	3	-	32
esquistossomose	Esquistossomose	16	4	-	96
filariose	Elefantíase	19	7	-	120
hanseníase	Doenças negligenciadas	8	3	-	60
helmintíase	Elefantíase filarial	2	4	2 singular/2plural	27
hepatites	Hanseníase	10	-	-	-
iniciativa	Doenças negligenciadas	10	-	-	-
leishmaniose	Esquistossomose	29	4	3singular/1plural	72
malária	Esquistossomose	20	2	-	32
oncocercose	Esquistossomose	21	4	-	49
pessoa	Hanseníase	6	2	1singular/1plural	22
problema	Doenças negligenciadas	12	3	-	33
raiva	Esquistossomose	13	3	-	37
rios	Oncocercose	14	-	-	-
saúde	Doenças negligenciadas	22	2	-	38
sintomas	Hanseníase	10			
tracoma	Esquistossomose	23	3	-	40
transmissão	Esquistossomose	10			
tratamento	Esquistossomose	17	2	-	36
tuberculose	Doenças negligenciadas	5	2	-	5

Fonte: a autora

Na primeira coluna da tabela temos o Termo Candidato, selecionado a partir do número de ocorrências. A segunda coluna corresponde a qual termo pesquisado o termo candidato apareceu mais vezes. A coluna *Range* informa quantas ocorrências o termo candidato teve na análise do termo pesquisado. Já as Repetições indicam o máximo de vezes que o termo candidato ocorreu em um documento e a coluna Singular/Plural mostra se houve distinção entre a ocorrência do termo no singular ou no plural. A última coluna informa a soma de ocorrências do termo em todos os arquivos do *corpus*.

### Fichas terminológicas

Seguindo os exemplos de fichas terminológicas como descritos por Lara e Tálamo (2007) e pela Prof.<sup>a</sup> Dr.<sup>a</sup> Vânia Mara Alves Lima na disciplina de Linguística Documentária do curso de Biblioteconomia da ECA-USP no primeiro semestre de 2022, foram selecionados os alguns campos, que originaram o seguinte modelo:

Quadro 1- Modelo de ficha terminológica

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	<i>Para organizar e identificar mais rapidamente os termos que compõem o conjunto terminológico</i>
<b>TERMO</b>	<i>Palavra selecionada do corpus</i>
<b>SINÔNIMOS</b>	<i>Palavras de significado igual ou semelhante ao termo escolhido</i>
<b>CONTEXTO</b>	<i>Trecho ou parágrafo em que o termo selecionado é utilizado; este campo tem a finalidade de auxiliar a visualização de possíveis relações entre os termos</i>
<b>DEFINIÇÃO</b>	<i>Descrição do termo</i>

Fonte: a autora

Para este estudo, foram produzidas as fichas terminológicas dos 32 termos selecionados para a confecção dessa proposta de terminologia de doenças negligenciadas. A seguir, temos um exemplo de uma das fichas. O conjunto de todas as fichas podem ser conferidas no item Apêndice A, ao final deste trabalho.

O preenchimento dos campos “sinônimos” e “definição” foi feito por meio de pesquisa dos termos em dicionários online (Priberam, Dicio, Sinônimos e Michaelis). A “área do conhecimento”, além das opções anteriores, poderá contar com o auxílio do campo “contexto”.

A ferramenta *File view* do AntConc fornece todas as ocorrências do termo pesquisado em cada documento, por isso foi utilizada para compor o item “contexto” das fichas. Devido à quantidade de arquivos e ocorrências que foram compiladas, optou-se por selecionar quatro definições dos arquivos do *corpus* de doenças negligenciadas reunido na presente pesquisa. A

lista com os respectivos links dos arquivos utilizados nas fichas confeccionadas neste trabalho pode ser conferida no Apêndice B.

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	01
<b>TERMO</b>	Brasil
<b>SINÔNIMOS</b>	Não possui
<b>CONTEXTO</b>	<p>“ [...] Espalhadas por todo o planeta, mais de um bilhão de pessoas vivem com menos de US\$ 2 por dia: seja no <b>Brasil</b> ou outros países da América Latina e do Caribe, na África, na Ásia e, também, nos Estados Unidos e em alguns países da Europa[...].” (arquivo 000).</p> <p>“ [...] Embora não apareçam na lista de doenças negligenciadas do <b>Brasil</b>, o calor também representa uma boa notícia para o mosquito por trás de dengue, zika e chikungunya — o Aedes aegypti[...].” (arquivo 005).</p> <p>“ [...] É considerada um problema de saúde pública no <b>Brasil</b>, o país mais afetado das Américas, onde estima-se que 1,5 milhão de indivíduos estejam infectados pelo Schistosoma mansoni[...].” (arquivo 017).</p> <p>“ [...] Parasitos das mais diversas espécies que infectam o ser humano são responsáveis pelas doenças endêmicas de grande importância no <b>Brasil</b> e o seu diagnóstico exige uma boa capacitação do profissional [...].” (arquivo 040).</p>
<b>DEFINIÇÃO</b>	<p>“ Maior país da América do Sul cuja extensão territorial pode chegar aos 8.5 milhões de quilômetros quadrados, sendo formado pela união de 26 estados e pelo distrito Federal.” (Dicio, 2024).</p> <p>“(1) País localizado na costa leste da América do Sul, entre a Colômbia e o Peru, que faz fronteira com o Oceano Atlântico. É delimitado ao norte pela Venezuela, Guiana, Suriname e Guiana Francesa; ao sul pelo Uruguai; e a oeste pela Argentina. Sua capital é Brasília. (2) A República Federativa do Brasil é formada por 5 regiões (norte, nordeste, centro-oeste, sudeste e sul), 26 Estados e o Distrito Federal (Brasília). É o quinto país do mundo em área total e o maior da América do Sul.” (DeCS/MeSH, 2024).</p>

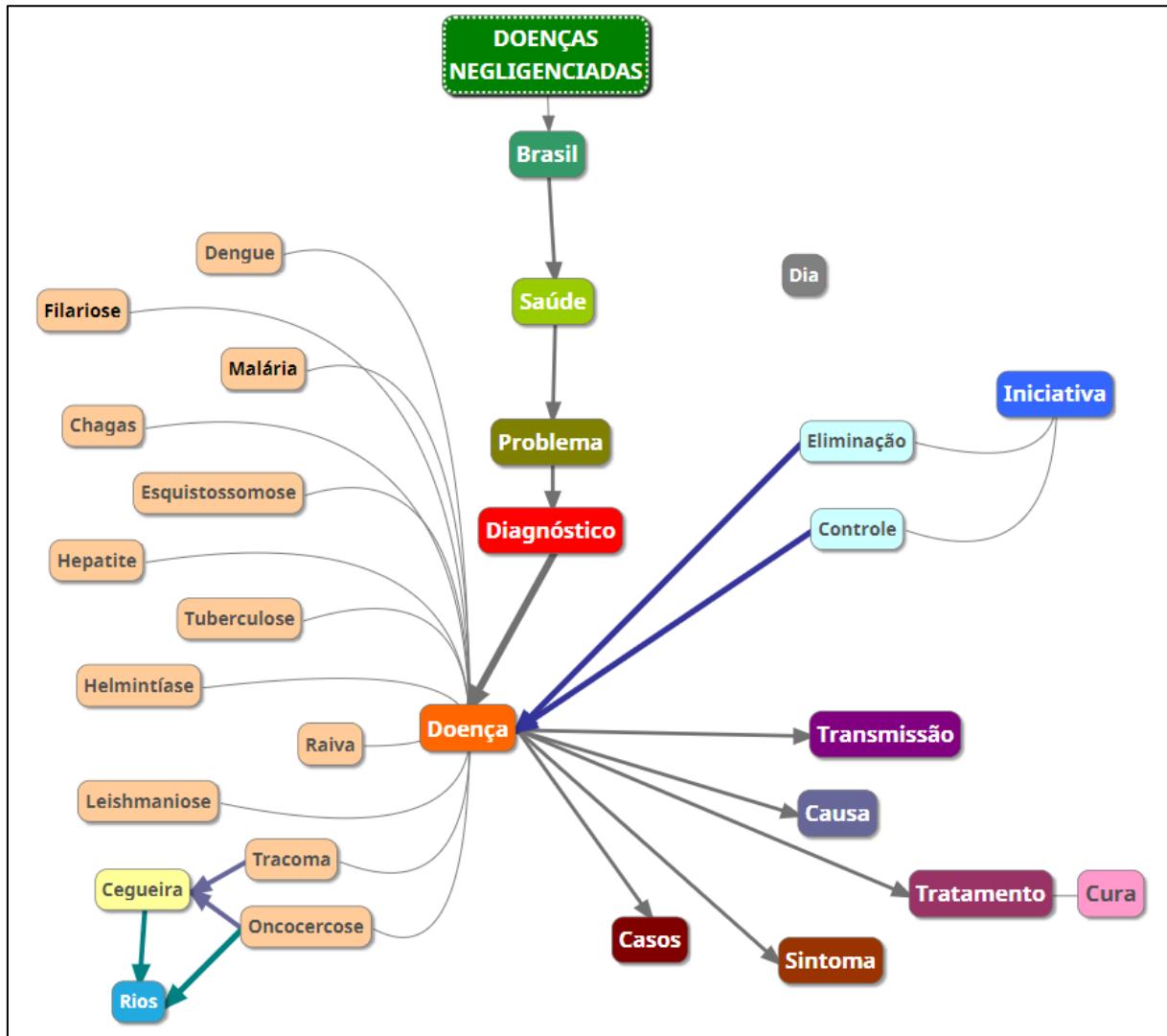
Com as fichas terminológicas organizadas, seguiu-se para uma última sistematização dos termos levantados. Os termos levantados foram distribuídos em categorias de alta generalização, por meio da noção de conjunto, que gerou a árvore de domínio utilizada como recurso de visualização das relações entre as palavras do *corpus* (Lara; Tálamo, 2007, p. 6).

A distribuição dos termos foi feita de acordo com as seguintes noções relacionais observadas nas definições dos dicionários, no DeCS/MeSH e nas ocorrências no *corpus*:

- Foi estabelecido como domínio principal o termo "Brasil", uma vez que a pesquisa focou em sites e documentos em língua portuguesa, obtendo os resultados referentes ao Brasil;
- Como assunto mais geral, temos o termo "saúde", de onde podemos relacionar "problema", que tem "diagnóstico" como termo subordinado;
- "Diagnóstico" deriva "doenças", e aqui temos um grupo grande de termos relacionados: "Chagas", "dengue", "esquistossomose", "filariose", "hanseníase", "helmintíase", "hepatite", "leishmaniose", "malária", "oncocercose", "raiva", "tracoma" e "tuberculose".
- "Doenças" é um termo que, além de trazer as doenças negligenciadas, também se ramifica em "casos", "causa", "sintoma", "tratamento" e "transmissão";
- "Pessoa" mostrou relação com "sintoma";
- Para o termo "tratamento" temos a relação com "cura";
- "Iniciativa", no sentido de 'realizar um ato' ou 'ação' tem dois termos relacionados "controle" e "eliminação";
- "Oncocercose" e "tracoma" têm uma relação em comum: "cegueira";
- "Rios" tem relação com "oncocercose" e "cegueira";
- O termo "dia" ficou deslocado, uma vez que os contextos de uso não mostraram relação com os outros substantivos relacionados.

A partir das relações estabelecidas, foi possível a construção da seguinte árvore de domínio:

Figura 12 - Proposta de Árvore de Domínio do *corpus* de Doenças Negligenciadas



Fonte: a autora

A árvore de domínio, que segue a ideia de *mind maps* (mapas mentais), também chamados de *topic maps* (mapas de tópicos), permite a visualização geral dos termos selecionados e de suas relações.

## 7 CONSIDERAÇÕES FINAIS

A área da Ciência da Computação, assim como a Ciência da Informação são campos multi e transdisciplinares, permitindo que os estudos dialoguem entre si e se beneficiem dos conhecimentos um do outro. Por meio deste trabalho foi possível constatar a existência de um diálogo entre a Biblioteconomia e o Processamento de Linguagem Natural com o uso da Linguística de *Corpus*. – uma vez que são áreas multi e interdisciplinares.

Os estudos da Terminologia mostram-se ser um campo que necessita de constante revisão e avaliação de atualização e até mesmo a criação de terminologias para novos domínios de especialidade. A aplicação das técnicas de PLN apresentam-se como relevantes para determinar quais as palavras que podem formar uma terminologia, bem como quais serão os termos preferenciais e os termos relacionados. Para que uma terminologia de especialidade seja criada, é preciso buscar um *corpus* amplo e composto, primeiramente, de diversas áreas do conhecimento, para que se verifique os contextos reais de uso das palavras e/ou expressões que são candidatas a termo preferido e que irá representar o conceito para determinado conjunto de descritores. Além disso, quando se faz uma terminologia, também se verifica uma tendência com a qual os pesquisadores estão trabalhando.

Após este trabalho de pesquisa, verificou-se que a área de PLN carece de fontes teóricas em língua portuguesa, tendo como referência no assunto a Professora Maria das Graças Volpe Nunes, docente no Instituto de Ciências Matemáticas e Computação da Universidade de São Paulo, campus São Carlos.

O programa BootCaT se mostrou bastante eficaz na recuperação de *websites* relacionados ao termo desejado, detalhando em seu relatório final algumas informações que podem ser relevantes para o pesquisador tanto para a análise linguística quanto em relação a qual o tipo de *website* o assunto pesquisado tem mais conteúdo. A funcionalidade de salvar as pesquisas já em formato .txt é um ponto que também facilita muito a tarefa do usuário na coleta e organização dos *corpora*.

O programa AntConc, embora tenha ferramentas que sejam mais relevantes para os estudos de aspectos mais específicos da área da Linguística, pode auxiliar a tarefa dos Bibliotecários que trabalham com Terminologia e desenvolvimento de tesouros e vocabulários controlados.

Com esta pesquisa, observou-se que “esquistossomose” e “hanseníase” são as Doenças Negligenciadas mais recorrentes, pois obtiveram os maiores números de correspondências – respectivamente 17 e 21 termos candidatos. Em contrapartida, os termos “doenças esquecidas”,

“elefantíase filarial”, “helmintíase” e “popular” foram os que não tiveram substantivos com ocorrência para atender os critérios de palavras selecionadas como termos candidatos, o que pode significar que as Doenças Negligenciadas não possuem tanto conteúdo publicado e difundido para a população comum, dificultando, assim, o acesso às informações pelas pessoas que não possuem conhecimento na área da saúde. Uma possibilidade seria utilizar na busca a expressão “também conhecido(a) como”, para que outros termos das doenças sejam recuperados.

Outro ponto para reflexão é que mesmo pesquisando as DN de acordo com os termos fornecidos pelo DeCS/MeSH, devido ao filtro de busca do idioma ter sido definido como português, os termos encontrados passaram a abranger também as Doenças Negligenciadas mais comuns no Brasil.

A árvore de domínio que foi proposta nessa pesquisa seguiu a distribuição sugerida por Lara e Tálamo (2007), e permite uma melhor visualização das relações entre os termos selecionados e descritos nas fichas terminológicas.

O princípio de construção de um *corpus* sobre Doenças Negligenciadas mostrou-se bastante complexo e acabou não retornando um resultado esperado de mais palavras que indicassem o uso popular para se referir às Doenças Negligenciadas. Além disso, a ideia inicial de realizar a busca em fontes que não fossem específicas da área da saúde mostrou-se insatisfatória, uma vez que as pesquisas retornaram muitos resultados de sites referentes à área da saúde. No entanto, o *corpus* reunido representa um possível caminho para se desenvolver uma terminologia especializada baseado em fontes que não sejam exclusivamente de artigos acadêmicos e trabalhos acessíveis somente à comunidade da área da saúde.

Recomenda-se estudos futuros para a obtenção de *corpora* terminológicos por meio das técnicas de PLN e dos programas de busca e análise de *corpus*, podendo ser aplicados a todas as áreas do conhecimento, não somente às linguagens documentárias e os *topic maps*, mas até mesmo em glossários e ferramentas de organização do conhecimento - em especial as que precisam ser ainda necessitam ser desenvolvidas e as que necessitam de constante atualização para acompanhar a rápida evolução da criação de conteúdo e de informações. Para que isso aconteça, é preciso que os Linguistas e os Bibliotecários trabalhem em conjunto e realizem mais estudos sobre Inteligência Artificial e olhem para as questões que envolvem as terminologias e a organização do conhecimento.

## REFERÊNCIAS<sup>7</sup>

- AGUIAR, D. A.; SANTANA, W. J. G. Inteligência artificial, as limitações da linguagem natural e seus impactos: estado da arte. In: **Revista Americana de Empreendedorismo e Inovação**. v. 5, n. 1, p. 7-17, mar. 2023. Disponível em: <https://periodicos.unespar.edu.br/index.php/raei/article/view/7423/5557>. Acesso em: 01 out. 2023.
- ANTHONY, L. (2023). **AntConc** (Version 4.2.4) [Computer Software]. Tokyo, Japan: Waseda University. Disponível em: <https://www.laurenceanthony.net/software>. Acesso em: 23 out 2023.
- BARONI, M; BERNARDINI, S. 2004. **BootCaT**: Bootstrapping corpora and terms from the web. Proceedings of LREC 2004. Disponível em: [https://sslmit.unibo.it/~baroni/publications/lrec2004/bootcat\\_lrec\\_2004.pdf](https://sslmit.unibo.it/~baroni/publications/lrec2004/bootcat_lrec_2004.pdf). Acesso em: 23 out. 2023.
- BENVENISTE, E. Vista d'olhos sobre o desenvolvimento da linguística. In: BENVENISTE, E. **Problemas de linguística geral I**. São Paulo : Ed. Nacional ; EDUSP, 1991. p.19-33
- BIDERMAN, M. T. C. A língua e o computador. In: BIDERMAN, M. T. C. **Teoria linguística**: teoria lexical e linguística computacional. São Paulo: Martins Fontes. 2001. p.75-93
- CASELI, H. M; NUNES, M. G. V; PAGANO, A. O que é PLN? In: CASELI, H. M; NUNES, M. G. V. (org.). **Processamento de linguagem natural**: conceitos, técnicas e aplicações em português. 2023. Disponível em: <https://brasileiraspln.com/livro-pln/1a-edicao/> . Acesso em: 06 nov. 2023.
- CLARO, D. B; SANTOS, J; SOUZA, M; VIEIRA, R; PINHEIRO, V. Extração de informação. In: CASELI, H. M; NUNES, M. G. V. (org.). **Processamento de linguagem natural**: conceitos, técnicas e aplicações em português. 2023. Disponível em: <https://brasileiraspln.com/livro-pln/1a-edicao/parte8/cap17/cap17.html>. Acesso em: 06 nov. 2023.
- CUNHA, M. B.; CAVALCANTI, C. R. O. **Dicionário de biblioteconomia e arquivologia**. Brasília: Briquet de Lemos, 2008.
- DESCRITORES em Ciências da Saúde: DeCS. 2023. São Paulo: BIREME / OPAS / OMS, 2017. Disponível em: <http://decs.bvsalud.org>. Acesso em: 25 fev. 2024.
- DICIO Dicionário Online de Português. Disponível em: <https://www.dicio.com.br/>. Acesso em: 07 mar. 2024.

---

<sup>7</sup> De acordo com a ABNT 6023 (2018).

DICIONÁRIO Priberam da Língua Portuguesa. Disponível em:  
<https://dicionario.priberam.org/>. Acesso em: 07 mar. 2024.

GIL, A. C. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2002.

GONÇALVES, J. S. B. Tutorial AntConc. Software para a realização de análises qualitativas. **Documentos LANTRI**, n.1, p. 1-18, set. 2016. Disponível em:  
[https://www.laurenceanthony.net/software/antconc/resources/help\\_AntConc344\\_portuguese.pdf](https://www.laurenceanthony.net/software/antconc/resources/help_AntConc344_portuguese.pdf). Acesso em 29 out 2023.

GOOGLE TRENDS. Veja o que esteve em alta em 2020 - Global. Disponível em:  
<https://trends.google.com.br/trends/yis/2020/GLOBAL/?hl=pt>. Acesso em: 21 nov. 2023.

ISO 704. **Terminology work – Principles and methods**. Geneva: ISO, 2009.

ISO 1087. **Terminology work and terminology science – Vocabulary**. Geneva: ISO, 2019. Disponível em: <https://www.iso.org/obp/ui/en/#iso:std:iso:1087:ed-2:v1:en>. Acesso em: 25 nov. 2023.

JURAFSKY, D.; MARTIN, J. H. **Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition**. 2 ed. New Jersey: Pearson, 2008. p. 1-16.

KRIEGER, M. G. Documentação. In: FINATTO, M. J. B. **Introdução à terminologia: teoria e prática**. São Paulo: Contexto, 2004. p. 58-73.

LARA, M.L.G.; TÁLAMO, M.F.G.M. Uma experiência didática na interface Linguística Documentária e Terminologia. **Datagramazero**, v. 8, n.5, Artigo 1, p. 1-11, 2007. Disponível em: <https://cip.brapci.inf.br/download/176079>. Acesso em: 26 out. 2023.

LIMA, J. L. O.; ALVARES, L. Organização e representação da informação e do conhecimento. In: ALVARES, L. **Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações**. São Paulo: B4 Editores, 2012. p 21-48.

LIMA, V. M. A. **Da classificação do conhecimento científico aos sistemas de recuperação de informação: enunciação de decodificação e decodificação da informação documentária**. 2004. Tese (Doutorado em Ciências da Comunicação) - Escola de Comunicações e Artes - Universidade de São Paulo, São Paulo, 2004. Disponível em: <https://www.teses.usp.br/teses/disponiveis/27/27143/tde-06032006-150120/publico/Daclassificacaodoconhecimentocientificoaossistemasderecuperaçaoadeinformacao.pdf>. Acesso em: 26 nov. 2023.

LOPES, L.; VIEIRA, R. FINATTO, M. J.; MARTINS, D.; ZANETTE, A.; RIBEIRO JR, L.C. Extração automática de termos compostos para construção de ontologias: um experimento na área da saúde. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, [S.]

- l.], v.3, n. 1, p. 76-88, mar. 2009. Disponível em:  
<https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/821>. Acesso em: 23 jan. 2023.*
- MAIMONE, G. D; TÁLAMO, M. F. G. M. Linguística e terminologia: contribuições para a elaboração de tesouros em ciência da informação. **DataGramZero – Revista de Ciência da Informação**. v. 12, n. 2, abr. 2011.
- MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 9 ed. São Paulo: Atlas, 2022.
- MOREIRA, V. P. Recuperação de informação. In: CASELI, H. M; NUNES, M. G. V. (org.). **Processamento de linguagem natural: conceitos, técnicas e aplicações em português**. 2023. Disponível em: <https://brasileiraspln.com/livro-pln/1a-edicao/>. Acesso em: 06 nov. 2023.
- NUNES, M. G. V. O processamento de línguas naturais: para quê e para quem? **Notas didáticas do ICMC**. Instituto de Ciências Matemáticas e de Computação. São Carlos. n. 73, p. 1-12, maio 2008. Disponível em:  
[https://web.icmc.usp.br/SCATUSU/RT/Notas\\_Didaticas/nd\\_73.pdf](https://web.icmc.usp.br/SCATUSU/RT/Notas_Didaticas/nd_73.pdf). Acesso em: 20 set. 2023.
- PAGANO, A. et al. PLN na saúde. In: CASELI, H. M; NUNES, M. G. V. (org.). **Processamento de linguagem natural: conceitos, técnicas e aplicações em português**. 2023. Disponível em: <https://brasileiraspln.com/livro-pln/1a-edicao/>. Acesso em: 06 nov. 2023.
- SARDINHA, T. B. Linguística de corpus: histórico e problemática. **D.E.L.T.A.**, v. 16, n. 2, p. 323-367, 2000. Disponível em: <https://revistas.pucsp.br/index.php/delta/article/view/39903>. Acesso em: 16 jan. 2023.
- SINÔNIMOS. Dicionário online. Disponível em: <https://www.sinonimos.com.br/>. Acesso em: 07 mar. 2024.
- SILVEIRA, D. T; CÓRDOVA, F. P. A pesquisa científica. In: GERHARDT, T. E; SILVEIRA, D. T. (org.). **Métodos de pesquisa**. Porto Alegre: Editora UFRGS, 2009. p. 31-42.
- TURING, A. M. Computing machinery and intelligence. **Mind**, v. LIX, n. 236, p. 433-460, out. 1950. DOI <https://doi.org/10.1093/mind/LIX.236.433>. Disponível em:  
<https://academic.oup.com/mind/article/LIX/236/433/986238>. Acesso em: 22 out. 2023.
- COMPUTADORES e inteligência. Disponível em:  
<https://turingtextos.blogspot.com/2017/02/computadores-e-inteligencia-alan.html>. Acesso em: 22 out. 2023.
- VALVERDE, R. Doenças negligenciadas. Agência Fiocruz de notícias – Saúde e ciência para todos. Especiais. Disponível em: <https://agencia.fiocruz.br/doen%C3%A7as-negligenciadas>. Acesso em: 10 dez. 2023.
- ZANCHETA, E.; GASPARI, F. BootCaT tutorial. Disponível em:  
<https://docs.sslmit.unibo.it/doku.php?id=bootcat:start>. Acesso em: 03 mar. 2024.

## APÊNDICES

### **APÊNDICE A – Fichas Terminológicas dos termos coletados do *corpus* de Doenças Negligenciadas**

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	01
<b>TERMO</b>	Brasil
<b>SINÔNIMOS</b>	Não possui
<b>CONTEXTO</b>	<p>“ [...] Espalhadas por todo o planeta, mais de um bilhão de pessoas vivem com menos de US\$ 2 por dia: seja no <b>Brasil</b> ou outros países da América Latina e do Caribe, na África, na Ásia e, também, nos Estados Unidos e em alguns países da Europa[...].” (arquivo 000).</p> <p>“ [...] Embora não apareçam na lista de doenças negligenciadas do <b>Brasil</b>, o calor também representa uma boa notícia para o mosquito por trás de dengue, zika e chikungunya — o <i>Aedes aegypti</i>[...].” (arquivo 005).</p> <p>“ [...] É considerada um problema de saúde pública no <b>Brasil</b>, o país mais afetado das Américas, onde estima-se que 1,5 milhão de indivíduos estejam infectados pelo <i>Schistosoma mansoni</i>[...].” (arquivo 017).</p> <p>“ [...] Parasitos das mais diversas espécies que infectam o ser humano são responsáveis pelas doenças endêmicas de grande importância no <b>Brasil</b> e o seu diagnóstico exige uma boa capacitação do profissional [...].” (arquivo 040).</p>
<b>DEFINIÇÃO</b>	<p>“ Maior país da América do Sul cuja extensão territorial pode chegar aos 8.5 milhões de quilômetros quadrados, sendo formado pela união de 26 estados e pelo distrito Federal” (Dicio, 2024).</p> <p>“(1) País localizado na costa leste da América do Sul, entre a Colômbia e o Peru, que faz fronteira com o Oceano Atlântico. É delimitado ao norte pela Venezuela, Guiana, Suriname e Guiana Francesa; ao sul pelo Uruguai; e a oeste pela Argentina. Sua capital é Brasília. (2) A República Federativa do Brasil é formada por 5 regiões (norte, nordeste, centro-oeste, sudeste e sul), 26 Estados e o Distrito Federal (Brasília). É o quinto país do mundo em área total e o maior da América do Sul” (DeCS/MeSH, 2024).</p>

FICHA TERMINOLÓGICA	
NÚMERO	02
TERMO	Casos
SINÔNIMOS	Acontecimento, ocorrência, episódio, evento
CONTEXTO	<p>“[...] Entre 2016 e 2020 foram identificados 583.960 <b>casos</b> novos das DTNs, com uma média anual de 116.792 <b>casos</b>. Em todo o período, as maiores taxas de detecção de <b>casos</b> foram verificadas nas Regiões Norte e Nordeste[...].” (arquivo 006).</p> <p>“[...] Mais do que estabelecer o diagnóstico diferencial, porém, eles são importantes para orientar o tratamento e, quando necessário, para implementar as medidas epidemiológicas na comunidade para evitar que novos <b>casos</b> da doença ocorram [...]”(arquivo 035).</p> <p>“[...] Mais de 90% dos <b>casos</b> de malária ocorrem na Região Norte e há surtos de filariose linfática e oncocercose [...]” (arquivo 054).</p> <p>“[...] No <b>caso</b> de Patrícia, a doença evoluiu para um quadro clínico agravado principalmente em razão do diagnóstico tardio[...].” (arquivo 95).</p>
DEFINIÇÃO	<p>“8 Manifestação individual de uma doença; ocorrência” (Michaelis, 2024).</p> <p><b>TERMO ALTERNATIVO PARA “RELATOS DE CASOS”</b></p> <p>“Nota de escopo: Apresentações clínicas que podem ser seguidas pelos estudos avaliativos que conduzem eventualmente a um diagnóstico (DeCS/MeSH, 2024).</p>

FICHA TERMINOLÓGICA	
NÚMERO	03
TERMO	Causa
SINÔNIMOS	Fator, agente, fonte, origem
CONTEXTO	<p>“[...] a leishmaniose cutânea <b>causa</b> ulceração e resulta na destruição parcial ou completa das mucosas do nariz, boca e garganta[...].” (arquivo 014).</p> <p>“[...] A alteração cardíaca é a forma mais importante de limitação do portador da doença e a principal <b>causa</b> de morte[...].” (arquivo 063).</p>

	<p>“[...] O tracoma é a principal <i>causa</i> de cegueira infecciosa e é responsável por prejuízos visuais em 1,9 milhões de pessoas, das quais 450 mil apresentam cegueira irreversível[...].” (arquivo 114).</p> <p>“[...] Admite-se que este declínio teve como <i>causa</i> principal a melhoria das condições sócio-econômicas experimentadas pelos povos europeus ao longo das Idades Moderna e Contemporânea[...].” (arquivo 163).</p>
<b>DEFINIÇÃO</b>	<p>“2 Aquilo que provoca o início ou determina a origem de algo; agente, origem, princípio” (Michaelis, 2024).</p> <p><b>TERMO ALTERNATIVO PARA “CAUSALIDADE”</b></p> <p>“Nota de escopo: Processo de relacionar causas com os efeitos que elas produzem. Causas são qualificadas como necessárias sempre que devem preceder um efeito e suficientes quando elas iniciam ou produzem um efeito. (Tradução livre do original: Last, 2001)” (Decs/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	04
<b>TERMO</b>	Cegueira
<b>SINÔNIMOS</b>	Ablepsia
<b>CONTEXTO</b>	<p>“[...] O tracoma é a causa mais frequente de <i>cegueira</i> evitável, representando 3% da cegueira do mundo[...].” (arquivo 009).</p> <p>“[...] Tais cuidados são indispensáveis por dois motivos: 1) evitar as recorrências que podem levar à <i>cegueira</i> e 2) controlar/interromper a cadeia de transmissão da bactéria[...].” (arquivo 035).</p> <p>“[...] O tracoma é a segunda causa de <i>cegueira</i> no mundo, depois da catarata[...].” (arquivo 067).</p> <p>“[...] Já a oncocercose, também chamada de <i>cegueira</i> dos rios ou mal do garimpeiro, foi eliminada na Colômbia, no Equador, na Guatemala e no México[...].” (arquivo 096).</p>
<b>DEFINIÇÃO</b>	<p>“1. [Medicina]: Estado de quem está privado do sentido da visão ou tem visão muito reduzida.” (Priberam, 2024)</p> <p><b>CONCEITO PREFERIDO</b></p> <p>“Nota de escopo: Incapacidade de enxergar ou ausência da percepção visual. Esta afecção pode ser o resultado de DOENÇAS OCULARES, DOENÇAS DO NERVO ÓPTICO, doenças do</p>

	QUIASMA ÓPTICO ou DOENÇAS CEREBRAIS que afetam as VIAS VISUAIS ou LOBO OCCIPITAL” (DeCS/Mesh, 2024).
--	--

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	05
<b>TERMO</b>	Chagas
<b>SINÔNIMOS</b>	Doença de Chagas – termo preferido para desambiguação da palavra “chaga”, que quer dizer ferimento
<b>CONTEXTO</b>	<p>“[...] O Ministério da Saúde vai destinar R\$ 20 milhões para a pesquisa das chamadas doenças negligenciadas, como a dengue, a doença de <i>Chagas</i>, a hanseníase, as leishmanioses, a malária e a tuberculose, que atingem sobretudo os países pobres e, por isso, raramente são alvo da curiosidade e dos investimentos dos laboratórios farmacêuticos[....]” (arquivo 080).</p> <p>“[...] As DTNs mais comuns no país são: doença de <i>Chagas</i>, esquistossomose, filariose linfática, hanseníase, leishmaniose visceral, leishmaniose tegumentar, oncocercose, raiva humana, tracoma e acidente ofídico[...].” (arquivo 006).</p> <p>“[...] O Programa Sanar tem como objetivo reduzir ou eliminar enquanto problema de saúde pública as seguintes doenças transmissíveis negligenciadas: tuberculose, hanseníase, esquistossomose, doença de <i>Chagas</i>, leishmaniose, filariose, geohelmintíases e tracoma[...].” (arquivo 063).</p> <p>“[...] O Ministério da Saúde vai destinar R\$ 20 milhões para a pesquisa das chamadas doenças negligenciadas, como a dengue, a doença de <i>Chagas</i>, a hanseníase, as leishmanioses, a malária e a tuberculose, que atingem sobretudo os países pobres e, por isso, raramente são alvo da curiosidade e dos investimentos dos laboratórios farmacêuticos[...].” (arquivo 080).</p>
<b>DEFINIÇÃO</b>	<p>“Busca por ‘doença de Chagas’</p> <p>[Medicina] Doença infecciosa aguda ou crônica, causada pelo protozoário <i>Trypanosoma cruzi</i> e transmitida ao homem por alguns insetos hematófagos, endêmica da América Central e do Sul e estudada inicialmente pelo médico brasileiro Carlos Chagas. = TRIPANOSOMÍASE AMERICANA” (Priberam, 2024).</p> <p>“CONCEITO PREFERIDO – DOENÇA DE CHAGAS</p> <p>Nota de escopo: Infecção com parasita protozoário TRYPANOSOMA CRUZI, uma forma de TRYPANOSSOMOSE endêmica nas Américas Central e do Sul. Foi denominada pelo médico brasileiro Carlos Chagas, que descobriu o parasita. A infecção pelo parasita (somente com resultado sorológico positivo)</p>

	se diferencia das manifestações clínicas que se desenvolvem após alguns anos, como destruição dos GÂNGLIOS PARASSIMPÁTICOS, CARDIOMIOPATIA CHAGÁSICA e disfunção do ESOFÂGO ou CÓLON” (DeCS/MeSH, 2024).
--	--

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	06
<b>TERMO</b>	Controle
<b>SINÔNIMOS</b>	Vigilância, inspeção, fiscalização, domínio
<b>CONTEXTO</b>	<p>“[...] Segundo Neves (2005, p. 259), algumas medidas têm de ser tomadas para o <i>controle</i> desta e de outras helmintases intestinais como: a construção de redes de esgoto ou fossas sépticas com tratamento das fezes humana[...]" (arquivo 002).</p> <p>“[...] Este eBook aborda os principais parasitos intestinais, apresentando 19 espécies cujas informações estão divididas da seguinte forma: doença causada, morfologia, habitat, vias de transmissão, ciclo de vida, sintomas e patogenia, distribuição geográfica, diagnóstico parasitológico, <i>controle</i> e prevenção[...]" (arquivo 040).</p> <p>“[...] Na avaliação de Júlio Croda, o Brasil retrocedeu de dez a 20 anos no combate a essas doenças. Segundo ele, será necessário reconstruir os serviços de saúde já que todos os programas nacionais de <i>controle</i> para essas doenças sofreram algum impacto[...]" (arquivo 054).</p> <p>“[...] A análise da produção bibliográfica tem papel relevante para fortalecimento do desenvolvimento de pesquisas e planejamento estratégico de programas para o <i>controle</i> de tracoma e doenças tropicais negligenciadas em geral[...]" (arquivo 161).</p>
<b>DEFINIÇÃO</b>	“Ação de controlar, de possuir domínio sobre algo ou alguém” (Dicio, 2024).

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	07
<b>TERMO</b>	Cura
<b>SINÔNIMOS</b>	Melhora

<b>CONTEXTO</b>	<p>“[...] Ele deve ser mantido durante a fase ativa da doença e depois periodicamente até que seja constatada a <i>cura</i> definitiva do tracoma[...].” (arquivo 041).</p> <p>“[...] O objetivo do tratamento é a <i>cura</i> da infecção, interrupção da cadeia de transmissão da bactéria e diminuição da circulação do agente etiológico na comunidade, o que leva à redução da frequência das reinfecções e da gravidade dos casos[...].” (arquivo 114).</p> <p>“[...] Ademais, a crença na <i>cura</i> para a infecção esquistossomótica está bem sedimentada no grupo, e, sobretudo, a infecção não é considerada algo que coloque a vida sob ameaça de morte[...].” (arquivo 155).</p> <p>“[...] Na perspectiva biomédica da saúde, a hanseníase é uma doença como outra qualquer, que manifesta sinais, sintomas, necessita de exames para haver confirmação do diagnóstico, e, em seguida, um tratamento adequado que evolua até a <i>cura</i> do indivíduo, tudo gratuito e disponível pelo SUS[...].” (arquivo 162).</p>
<b>DEFINIÇÃO</b>	<p>“1. Ato ou efeito de curar ou de se curar. 2. Período em que se segue um regímen ou tratamento contra uma doença. 3. Recuperação da saúde” (Priberam, 2024).</p>

FICHA TERMINOLÓGICA	
<b>NÚMERO</b>	08
<b>TERMO</b>	Dengue
<b>SINÔNIMOS</b>	febre dengue
<b>CONTEXTO</b>	<p>“[...] Por estarem associadas a contextos tropicais empobrecidos, não se espalham para países distantes e só raramente afetam viajantes como, por exemplo, durante surtos de <i>dengue</i>[...].” (arquivo 020).</p> <p>“[...] No ano passado, foram registrados perto de 1 milhão de casos de <i>dengue</i> no Brasil, com pouco mais de 500 mortes[...].” (arquivo 081).</p> <p>“[...] A <i>dengue</i>, transmitida por mosquitos, não ocorre apenas na África, no Sudeste Asiático e na América do Sul, mas também já atingiu a região do Mediterrâneo e o sul da Europa, lembra May[...].” (arquivo 168).</p> <p>“[...] Façamos um contraponto com a <i>dengue</i> que é observada em 129 países e que tem potencial para infectar mais de 3,9 bilhões de pessoas[...].” (arquivo 177).</p>
<b>DEFINIÇÃO</b>	<p>“MED Doença febril infecciosa cujos sintomas sobrevêm repentinamente, caracterizada por fortes dores de cabeça, nos olhos,</p>

	<p>nos músculos e nas articulações, inflamação na garganta, manifestações catarrais e, às vezes, erupções cutâneas, além de inchaços doloridos, transmitida ao homem pela picada do mosquito <i>Aedes aegypti</i>; febre dengue” (Michaelis, 2004).</p> <p><b>“TERMO PREFERIDO</b></p> <p>Nota de escopo: Doença febril aguda transmitida por picada de mosquitos AEDES infectados com o VÍRUS DA DENGUE. É autolimitada e caracterizada por febre, mialgia, cefaleia e exantema. A DENGUE GRAVE é uma forma mais virulenta da dengue” (DeCS/MeSH, 2024).</p>
--	---

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	09
<b>TERMO</b>	Dia
<b>SINÔNIMOS</b>	Data, tempo, presente, momento
<b>CONTEXTO</b>	<p>“No período reprodutivo, uma fêmea de Ascaris quando fecundada, pode colocar mais de 200 mil ovos por <i>dia</i> [...]” (arquivo 002).</p> <p>“Desde 2021, todo dia 30 de janeiro é marcado como <b><i>Dia</i></b> Mundial das Doenças Tropicais Negligenciadas[...]” (arquivo 043).</p> <p>“[...]As fêmeas fecundadas produzem uma quantidade enorme de embriões, com produção estimada de 500 a 3.800 microfilárias por <i>dia</i> [...]” (arquivo 085).</p> <p>“[...]Se algum <i>dia</i> a medicina conseguir controlar o sistema imunológico do ser humano da mesma maneira que este organismo controla, poderíamos resolver problemas como a rejeição de órgãos transplantados ou de qualquer tipo de implantes, só para citar alguns exemplos”, comenta De Marco[...]” (arquivo 172).</p>
<b>DEFINIÇÃO</b>	“3 CRON Período de tempo em que a Terra descreve uma volta completa em torno de seu eixo de rotação, geralmente contado desde a meia-noite até vinte e quatro horas depois; dia civil” (Michaelis, 2024).

FICHA TERMINOLÓGICA	
NÚMERO	10
TERMO	Diagnóstico
SINÔNIMOS	Diagnose, análise
CONTEXTO	<p>“[...] Entre as condições que aparecem na lista, algumas possuem <b>diagnóstico</b> e tratamento disponíveis há anos na rede pública, como a hanseníase e o tracoma[...].” (arquivo 005).</p> <p>“[...] Para atingir a meta de eliminação, atualmente são promovidas medidas preventivas, além de ações de caracterização epidemiológica do foco em área Yanomami. "Promover o acesso da população ao <b>diagnóstico</b> precoce e tratamento oportuno no âmbito da atenção primária da saúde é um importante desafio", afirma Marilza[...].” (arquivo 061).</p> <p>“[...] Ou seja, indivíduos portadores de manifestações crônicas, geralmente, são amicrofilarêmicos, o que muitas vezes dificulta o <b>diagnóstico</b> etiológico da condição clínica apresentada[...].” (arquivo 097).</p> <p>“[...] A hanseníase tem tratamento e cura. Porém, se no momento do <b>diagnóstico</b> o paciente já apresentar alguma deformidade física instalada, esta pode ficar como seqüela permanente no momento da alta[...].” (arquivo 163).</p>
DEFINIÇÃO	“Qualificação dada por um médico a uma enfermidade ou estado fisiológico, com base nos sinais que observa” (Michaelis, 2024).

FICHA TERMINOLÓGICA	
NÚMERO	11
TERMO	Doença
SINÔNIMOS	Enfermidade, distúrbio, mazela, moléstia, patologia
CONTEXTO	<p>“[...] Outras doenças transmissíveis com tendência declinante são a difteria, a coqueluche e o tétano acidental, que têm em comum o fato de serem imunopreveníveis; a <b>doença</b> de Chagas, endêmica há várias décadas em nosso país; a febre tifóide, associada a condições sanitárias precárias; e a oncocercose, a filariose e a peste, todas com áreas de ocorrência restritas[...].” (arquivo 018).</p> <p>“[...] A hanseníase é uma <b>doença</b> crônica, cujo agente é a bactéria <i>Mycobacterium leprae</i>, pode acometer todas as pessoas[...].” (arquivo 38).</p>

	<p>“[...] A tuberculose é a <b>doença</b> negligenciada responsável pelo maior número de mortes entre as populações mais vulneráveis, segundo Croda[...]" (arquivo 054).</p> <p>“[...] Quem tem uma <b>doença</b> jamais vai se enxergar como um ponto na estatística, e este é o principal desafio dos processos de tratamento de saúde ao redor do mundo —do desenvolvimento de novos medicamentos aos tratamentos em si, passando por diagnósticos, exames e análises, cada pessoa precisa ser tratada com humanidade[...]" (arquivo 169).</p>
<b>DEFINIÇÃO</b>	<p>“1 MED Processo de alteração biológica, com sintomas característicos, muitas vezes imperceptíveis, que pode afetar o corpo todo ou uma ou, ainda, várias de suas partes, resultando na deterioração ou enfraquecimento do estado de saúde de um ser (homem ou animal); enfermidade, moléstia, mal, falta de saúde” (Michaelis, 2024).</p> <p><b>“CONCEITO PREFERIDO</b></p> <p>Nota de escopo: Processo patológico definido com um quadro característico de sinais e sintomas. Pode afetar o corpo inteiro ou quaisquer de suas partes. Sua etiologia, patologia e prognóstico podem ser conhecidos ou desconhecidos” (DeCS/Mesh, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	12
<b>TERMO</b>	Eliminação
<b>SINÔNIMOS</b>	Corte, supressão, expulsão
<b>CONTEXTO</b>	<p>“[...] o tratamento dos indivíduos infestados assume uma importância fundamental, pois a <b>eliminação</b> dos vermes adultos, através de medicamentos específicos, interfere diretamente no ciclo biológico do parasita[...]" (arquivo 002).</p> <p>“[...] Em 2003, a Organização Mundial da Saúde (OMS) começou a mudar o paradigma no controle e na <b>eliminação</b> de um grupo de doenças tropicais negligenciadas (DTN)[...]" (arquivo 020).</p> <p>“[...] De acordo com Marilza, o conceito de <b>eliminação</b> consiste na manutenção da interrupção da transmissão local da oncocercose [...]" (arquivo 061).</p>

	“[...] Até 2020, nove países no mundo mantinham validada a <b>eliminação</b> do tracoma como problema de saúde pública[...]" (arquivo 161).
<b>DEFINIÇÃO</b>	“Ato ou efeito de eliminar (-se)” (Michaelis, 2024).

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	13
<b>TERMO</b>	Esquistossomose
<b>SINÔNIMOS</b>	Xistose, barriga d'água, doença do caramujo, bilharzíaze
<b>CONTEXTO</b>	<p>“[...]O Brasil Saudável visa erradicar até 2030 doenças como malária, Chagas, tracoma, filariose, <b>esquistossomose</b>, oncocercose e algumas verminoses intestinais[...]" (arquivo 005).</p> <p>“[...]A <b>esquistossomose</b>, também conhecida como barriga d'água, xistose ou doença do caramujo, é uma doença causada por um platelminto[...]" (arquivo 015).</p> <p>“[...] A complicação mais importante da <b>esquistossomose</b> crônica é a hipertensão arterial pulmonar (HAP), com sinais e sintomas não diferentes dos causados por outras etiologias[...]" (arquivo 065).</p> <p>“[...]De acordo com a Organização Mundial da Saúde (OMS), atualmente 17 doenças parasitárias crônicas são classificadas como DTNs, dentre elas raiva, tracoma, doença de Chagas (trypanossomíase americana), leishmaniose, filariose linfática, oncocercose (cegueira dos rios), <b>esquistossomose</b> (bilharzíaze), helmintíases transmitidas pelo solo e água", atesta o Boletim Epidemiológico[...]" (arquivo 092).</p>
<b>DEFINIÇÃO</b>	“MED Afecção grave no ser humano causada por três espécies de vermes do gênero Schistosoma: Schistosoma haemotobium, Schistosoma japonicum e Schistosoma mansoni, sendo este o mais comum no Brasil, causador de danos ao fígado, ao intestino e ao baço; esquistosomíase, esquistossomíase, xistosa, xistossomíase, xistosomo, xistossomose” (Michaelis, 2024).

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	14
<b>TERMO</b>	Filariose

<b>SINÔNIMOS</b>	Filaríase, elefantíase
<b>CONTEXTO</b>	<p>“[...] A <b>filariose</b> linfática, também conhecida como elefantíase, é uma doença crônica causada por um parasita[...].” (arquivo 017).</p> <p>“[...] A <b>filariose</b> linfática, que causa inchaço nas pernas, desapareceu nos estados da Bahia, de Alagoas e do Pará, mas ainda está em processo de certificação da eliminação em Pernambuco, onde havia focos ativos até poucos anos atrás[...].” (arquivo 078).</p> <p>“[...] Outras estão diretamente relacionadas às condições precárias de moradia ou à falta de saneamento básico e acesso à água potável, casos de esquistossomose e <b>filariose</b> linfática[...].” (arquivo 165).</p> <p>“[...] No texto, tracoma, esquistossomose, doença de Chagas, leishmaniose, <b>filariose</b>, oncocercose e hanseníase vêm fazer companhia às geohelmintoses. São moléstias ligadas ao meio ambiente, falta de saneamento básico e pobreza[...].” (arquivo 173).</p>
<b>DEFINIÇÃO</b>	<p>“MED Infestação com filárias ou doenças causadas pela presença de filárias no sangue e nos tecidos” (Michaelis, 2024).</p> <p><b>“CONCEITO PREFERIDO</b></p> <p>Nota de escopo: Infecção com vermes (trematódeos) do gênero SCHISTOSOMA. Três espécies produzem as doenças clínicas mais frequentes: SCHISTOSOMA HAEMATOBIUM (endêmico na África e Oriente Médio), SCHISTOSOMA MANSONI (no Egito, norte e sul do continente africano, em algumas ilhas das Índias Orientais, 2/3 do norte da América do Sul) e SCHISTOSOMA JAPONICUM (no Japão, China, Filipinas, Sulawesi, Tailândia, Laos). S. mansoni é detectado com frequência em porto-riquenhos que moram nos Estados Unidos” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	15
<b>TERMO</b>	Hanseníase
<b>SINÔNIMOS</b>	lepra, lazeira
<b>CONTEXTO</b>	<p>“[...] No Brasil, leishmaniose, tuberculose, doença de Chagas, malária, esquistossomose, hepatites, filariose linfática, dengue e <b>hanseníase</b> estão entre as principais doenças negligenciadas[...].” (arquivo 007).</p> <p>“[...] A <b>hanseníase</b> é uma doença infecciosa, contagiosa, que afeta os nervos e a pele e é causada por um bacilo chamado Mycobacterium</p>

	<p>leprae ou bacilo de Hansen, em homenagem à Gerhard Henrick Armauer Hansen, médico e bacteriologista norueguês, que descobriu a doença, em 1873[...]" (arquivo 022).</p> <p>“[...] Esquistossomose, <b><i>hanseníase</i></b>, raiva humana e outros problemas de saúde podem afetar quase 15% da população, mas recebem pouca atenção e investimento. Um novo plano nacional pretende eliminar ou controlar a maioria dessas condições[...]" (arquivo 052).</p> <p>“[...] O pesquisador já publicou mais de 200 artigos ligados ao tema, sobre pesquisas em novos tratamentos e novos medicamentos, determinantes sociais da <b><i>hanseníase</i></b> e determinação genética de transmissão da doença[...]" (arquivo 104).</p>
<b>DEFINIÇÃO</b>	<p>“MED Infecção crônica, contagiosa, causada pelo <i>Mycobacterium leprae</i>, ou bacilo de Hansen, que produz inicialmente manchas sem pigmentos na pele, tornando-a insensível e sem transpiração, evoluindo para lesões mais graves nas mucosas e nos nervos periféricos; elefantíase dos gregos, gafa2, gafeira, gafo, guarucaia, lazeira, lepra, leprose, macota, macutena, mal, mal-bruto, mal de cuia, mal de lázaro, mal de são lázaro, mal do sangue, mal-morfético, morfeia” (Michaelis, 2024).</p> <p>“CONCEITO PREFERIDO</p> <p>Nota de escopo: Infecção granulomatosa crônica causada pelo <i>MYCOBACTERIUM LEPRAE</i>. As lesões granulomatosas são manifestadas na pele, nas mucosas e nos nervos periféricos. Há dois tipos polares ou principais: a lepromatosa e a tuberculoide” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	16
<b>TERMO</b>	Helmintíase
<b>SINÔNIMOS</b>	Ancilostomíase, verminose
<b>CONTEXTO</b>	<p>“[...]As DTN incluem: esquistossomose, leishmaniose cutânea, leishmaniose visceral, envenenamento por picadas de animais, filariose, tracoma, escabiose e outras ectoparasitoses, fasciolíase, bouba, micetoma, oncocercose, <b><i>helmintíase</i></b>, teníase/cisticercose, doença de chagas, hanseníase, hidatidose/equinococose e raiva[...]" (arquivo 023).</p> <p>“[...]A Organização Mundial da Saúde apontou seis doenças infecciosas tropicais como negligenciadas em países industrializados que são endêmicas em muitos países em desenvolvimento</p>

	<p>(<b>HELMINTÍASE</b>, HANSENÍASE, ELEFANTÍASE FILARIAL, ONCOCERCOSE, ESQUISTOSSOMOSE e TRACOMA).[...]" (arquivo 047).</p> <p>"[...]agentes destrutivos para os vermes parasitários. São usados terapeuticamente para tratar a <b>HELMINTÍASE</b>, tanto no homem como nos animais[...]" (arquivo 048).</p> <p>"[...]A oncocerose passa a ocupar papel importante junto às outras grandes endemias somente em 1975 (WHO, 1975), deixando de ser considerada uma <i>helmintíase</i> obscura e restrita a lugares remotos [...]" (arquivo 085).</p>
<b>DEFINIÇÃO</b>	<p>"MED Infestação parasitária do organismo de homens e animais causada por helmintos, como a esquistossomose, a ascaridíase, a teníase, a ancilostomose e outras parasitoses" (Michaelis, 2024).</p> <p>"A infestação com vermes parasitas da classe dos helmintos" (DeCS/Mesh, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	17
<b>TERMO</b>	Hepatites
<b>SINÔNIMOS</b>	Figadeira,
<b>CONTEXTO</b>	<p>"Além disso, há metas para reduzir casos de tuberculose, HIV, hanseníase e <b>hepatites</b> virais — e acabar com a transmissão vertical (quando o agente infeccioso passa da mãe para o bebê na gestação, parto ou amamentação) de HIV, sífilis, hepatite B, Chagas e HTLV" (arquivo 005).</p> <p>"[...]Nesse grupo de doenças, destacam-se a tuberculose e as <b>hepatites</b> virais, especialmente as <b>hepatites</b> B e C, em função das altas prevalências, da ampla distribuição geográfica e do potencial evolutivo para formas graves que podem levar ao óbito[...]" (arquivo 018).</p> <p>"O agravamento da esquistossomose por associação com as <b>hepatites</b> virais B e C é controverso. Para alguns autores, a evolução das <b>hepatites</b> segue seu próprio curso natural, independente da esquistossomose, enquanto que, para outros, a associação das <b>hepatites</b> virais B e C pode agravar a esquistossomose, causando mais precocemente insuficiência hepática, ascite, icterícia, hipertensão portal e varizes do esôfago com sangramento" (arquivo 021).</p>

	<p>“A classificação da OMS atualmente inclui 20 doenças tropicais negligenciadas. Dentre as prevalentes no Brasil, podemos citar: hanseníase, doença de Chagas, esquistossomose, leishmaniose, <b>hepatites</b> e filariose linfática” (arquivo 130).</p>
<b>DEFINIÇÃO</b>	<p>“A forma hepatites é feminino plural de hepatite” (Priberam, 2024).</p> <p>“MED Inflamação aguda ou crônica do fígado, infecciosa (causada por vírus, bactérias, parasitas), sérica (transmitida pelo sangue) ou tóxica (causada por álcool, antibióticos etc.); figadeira”(Michaelis, 2024).</p> <p>“INFLAMAÇÃO do FÍGADO” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	18
<b>TERMO</b>	Iniciativa
<b>SINÔNIMOS</b>	Ação, projeto, resolução, atividade
<b>CONTEXTO</b>	<p>“[...]A Organização também possui uma <b>iniciativa</b> para eliminar cerca de 30 doenças infecciosas e condições relacionadas até 2030” (arquivo 006).</p> <p>“A <b>Iniciativa</b> Medicamentos para Doenças Negligenciadas (DNDi)<sup>29</sup> é uma organização sem fins lucrativos com o objetivo de corrigir o chamado “desequilíbrio fatal” entre o desenvolvimento insuficiente de novos medicamentos para doenças negligenciadas e sua elevada ocorrência, especialmente entre as populações mais pobres do mundo[...].” (arquivo 071).</p> <p>“Outra <b>iniciativa</b> desenvolvida no contexto de apoio ao Plano Brasil sem Miséria foi o projeto de “Expedições Fiocruz para um Brasil sem Miséria”<sup>17</sup>, com apoio da Capes, da Fiocruz e de diversos parceiros mobilizados a cada expedição[...].” (arquivo 119).</p> <p>“Sem desmerecer a importância de doações filantrópicas, há uma forte motivação econômica nessa <b>iniciativa</b>, que pode vir a ser a maior na área de saúde pública, em termos globais, até agora[...].” (arquivo 174).</p>
<b>DEFINIÇÃO</b>	“2 Ação ou efeito de ser o primeiro a propor e pôr em prática qualquer coisa: Tomou a iniciativa de terminar o relacionamento” (Michaelis, 2024).

FICHA TERMINOLÓGICA	
NÚMERO	19
TERMO	Leishmaniose
SINÔNIMOS	Leishmaníase
CONTEXTO	<p>“[...] Em 2020, a taxa de mortalidade para malária subiu 82,55%, apesar da queda de 29,3% nas internações. Doenças como a <i>leishmaniose</i> visceral e a leptospirose também registraram aumento de mortalidade de 32,64% e 38,98%, respectivamente [...]” (arquivo 007).</p> <p>“[...] A ocorrência de formas graves ou atípicas de doenças tropicais, como <i>leishmaniose</i> e doença de Chagas, tem sido observada no Brasil [...]” (arquivo 018).</p> <p>“[...] A <i>leishmaniose</i>, ou leishmaníase, é uma antropozoonose, ou seja, doença primária de animais que pode ser transmitida ao homem [...]” (arquivo 71).</p> <p>“[...] O aumento médio da temperatura do planeta favorece o espalhamento de insetos transmissores de doenças, como a <i>leishmaniose</i>[...].” (arquivo 170).</p>
DEFINIÇÃO	<p>“MED Designação das doenças causadas por protozoários do gênero Leishmania” (Michaelis, 2024).</p> <p>“Doença causada por qualquer uma das espécies de protozoários do gênero LEISHMANIA. Há quatro tipos clínicos principais dessa infecção: cutânea, (LEISHMANIOSE CUTÂNEA), cutânea difusa (LEISHMANIOSE CUTÂNEA DIFUSA), mucocutânea (LEISHMANIOSE MUCOCUTÂNEA) e visceral (LEISHMANIOSE VISCERAL)” (DeCS/MeSH, 2024).</p>

FICHA TERMINOLÓGICA	
NÚMERO	20
TERMO	Malária
SINÔNIMOS	paludismo, sezão
CONTEXTO	“[...] No Brasil, 99% dos casos de <i>malária</i> ocorrem na região amazônica nos estados do Acre, Amazonas, Amapá, Maranhão, Mato Grosso, Pará, Rondônia, Roraima e Tocantins. Nessa região, a <i>malária</i> causa consideráveis perdas sociais e econômicas na população sob risco [...]” (arquivo 017).

	<p>“[...]Embora o conceito seja um tanto impreciso — algumas dessas condições também aparecem nas zonas temperadas, por exemplo —, é tradicionalmente usado para reunir uma série de quadros diferentes, que vão desde a <b>malária</b> e a dengue até a dracunculíase e as micoses profundas[...].” (arquivo 052)</p> <p>“[...]A <b>malária</b> é um exemplo de doença com alta prevalência em países pobres que passou a ser vista como prioridade. Em 2018, a Fundação Bill &amp; Melinda Gates destinou US\$ 1 bilhão ao programa de combate à <b>malária</b> da OMS e encomendou um estudo que resultou em um roteiro de ações em prevenção e pesquisa para erradicá-la até 2050[...].” (arquivo 081).</p> <p>“[...]A Novartis está comprometida com a luta contra a <b>malária</b> nas últimas duas décadas. Em 1999, lançamos a primeira terapia combinada à base de artemisinina em dose fixa e, em 2009, a primeira terapia pediátrica dispersível desenvolvida em parceria com a Medicines for Malaria Venture (MMV) [...]” (arquivo 131).</p>
<b>DEFINIÇÃO</b>	<p>“1 MED Enfermidade aguda transmitida ao homem pela picada de mosquito do gênero <i>Anopheles</i>, quando infectado por protozoário do gênero <i>Plasmodium</i>, que causa acessos de febre intermitente e destrói grande parte das hemácias; febre palustre, maleita-brava, maligna, malina” (Michaelis, 2024).</p> <p>“Doença causada em humanos por protozoários de quatro espécies do gênero <i>PLASMODIUM</i>: <i>PLASMODIUM FALCIPARUM</i>, <i>PLASMODIUM VIVAX</i>, <i>PLASMODIUM OVALE</i> e <i>PLASMODIUM MALARIAE</i> e transmitida pela picada da fêmea infectada do mosquito do gênero <i>ANOPHELES</i>. A malária é endêmica em partes da Ásia, África, Américas Central e do Sul, Oceania e em certas ilhas Caribenhais. Caracteriza-se clinicamente por exaustão extrema associada com paroxismos de FEBRE alta, SUDORESE, CALAFRIOS e ANEMIA. Em ANIMAIS, a malária é causada por outras espécies de plasmódio” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	21
<b>TERMO</b>	Medicamentos
<b>SINÔNIMOS</b>	Drogas, preparados, ingredientes, substâncias, remédios, fármacos.
<b>CONTEXTO</b>	“[...] Os 13 maiores laboratórios farmacêuticos do mundo, que se engajaram na campanha, têm pleno conhecimento disso e não ignoram que, além de vender novos <b>medicamentos</b> que vierem a ser

	<p>desenvolvidos, terão também de subsidiar ou doar milhões de doses, uma vez que, se o objetivo é a erradicação de endemias, o esforço não pode ficar limitado pela baixa renda de boa parte dos povos que vivem em países tropicais” (arquivo 000).</p> <p>“A pandemia da COVID-19 interrompeu intervenções e programas de eliminação dessas doenças nas Américas, como campanhas de administração de <b>medicamentos</b> em massa, pesquisas e rastreamento ativo de casos. A maioria foi cancelada em 2020 e retomada gradualmente em 2021” (arquivo 014).</p> <p>“Essas doenças são consideradas negligenciadas devido à falta de investimento na descoberta de novos <b>medicamentos</b> e vacinas e pela baixa eficácia dos programas de controle” (arquivo 064).</p> <p>“[...] Apesar de afetar milhões e matar dezenas de milhares de pessoas por ano, as DTNs costumam receber pouca atenção de desenvolvedores de <b>medicamentos</b>, de formuladores de políticas ou da mídia hegemonicamente [...]” (arquivo 122).</p>
<b>DEFINIÇÃO</b>	<p>“1. Substância que, devidamente manipulada, se aplica ao organismo doente com o fim de o curar.” (Priberam, 2024).</p> <p>“Termo alternativo para ‘Preparações Farmacêuticas’ Drogas dirigidas para uso humano ou veterinário, apresentadas em sua formulação final. Estão incluídos aqui os materiais usados na preparação e/ou na formulação final” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	22
<b>TERMO</b>	Oncocercose
<b>SINÔNIMOS</b>	Cegueira dos rios
<b>CONTEXTO</b>	<p>“[...] Os fármacos para tratamento da <b>oncocercose</b> e FL interrompem a transmissão matando os vermes jovens, mas não matam os vermes adultos, tornando necessárias diversas rodadas de tratamento antes que os vermes adultos morram e a transmissão acabe[...].” (arquivo 041).</p> <p>“[...] A <b>oncocercose</b>, doença parasitária que pode causar cegueira, ainda tem um foco no país, em Roraima, na fronteira com a Venezuela, no território habitado pelos Ianomâmi[...].” (arquivo 078).</p> <p>“[...] Entre 1974 e 2002, a <b>oncocercose</b> foi controlada na África Ocidental utilizando principalmente o controle vetorial,</p>

	<p>suplementado pela administração da ivermectina em larga escala desde o ano de 1989[...]" (arquivo 113).</p> <p>"[...]A <b>oncocercose</b> é uma infecção produzida pelo parasita Onchocerca volvulus, transmitido aos seres humanos por picadas das moscas do gênero Simulium. A doença causa prurido, deformações dermatológicas, perda da visão e cegueira[...]" (arquivo 175).</p>
<b>DEFINIÇÃO</b>	<p>"MED, VET Doença causada pela infestação de vermes nematódeos do gênero Onchocerca, cuja espécie Onchocerca volvulus invade o corpo e os olhos das vítimas formando tumores fibrosos, podendo levar à cegueira" (Michaelis, 2024).</p> <p>"Infecção com nematoide do gênero ONCHOCERCA. Entre as características estão a presença de nódulos subcutâneos firmes preenchidos por vermes adultos, PRURIDO e lesões oculares". DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	23
<b>TERMO</b>	Pessoa
<b>SINÔNIMOS</b>	Indivíduo, ser humano, criatura
<b>CONTEXTO</b>	<p>"[...] A doença não escolhe sexo, raça ou idade, podendo ser contraída por qualquer <b>pessoa</b>[...]" (arquivo 015).</p> <p>"[...]Uma vez dentro de uma <b>pessoa</b>, os vermes criam larvas que fazem o seu caminho para fora da pele, onde podem infectar a próxima mosca preta que pica a <b>pessoa</b> [...]" (arquivo 068).</p> <p>"[...] O desenvolvimento de intervenções e métodos combinados da saúde pública nos diversos territórios do Brasil pode contribuir para o controle deste grupo de doenças e levar à melhoria da qualidade de vida da <b>pessoa</b> acometida por uma doença tropical negligenciada[...]" (arquivo 130).</p> <p>"[...] O diagnóstico da oncocercose deve ser feito pelo clínico geral, infectologista ou oftalmologista a partir da avaliação dos sinais e sintomas apresentados pela <b>pessoa</b>[...]" (arquivo 167).</p>
<b>DEFINIÇÃO</b>	"1 Criatura humana [...]. 2 Ser eminente ou importante. 3 Caráter peculiar que dá distinção a alguém" (Michaelis, 2024).

FICHA TERMINOLÓGICA	
NÚMERO	24
TERMO	Problema
SINÔNIMOS	adversidade, contratempo, obstáculo
CONTEXTO	<p>“[...]Um <b>problema</b> preocupante em relação às doenças negligenciadas é a coinfeção pelo vírus da imunodeficiência humana ou pela tuberculose, que favorece manifestações clínicas graves e falha terapêutica[...].” (arquivo 064).</p> <p>“[...]O <b>problema</b> é que o G-Finder considera como negligenciadas doenças que não são mais classificadas assim pela OMS, como malária, Aids e tuberculose, e foram elas o alvo da maioria dos recursos[...].” (arquivo 078).</p> <p>“[...]O México pôs fim à raiva humana transmitida por cães e ao tracoma como <b>problema</b> de saúde pública[...].” (arquivo 135).</p> <p>“[...]O <b>problema</b> fundamental é que a droga apenas elimina a forma adulta do parasita, mas o meio ambiente ainda continua contaminado[...].” (arquivo 172).</p>
DEFINIÇÃO	“1 Tema, em qualquer área do conhecimento, cuja solução ou resposta requer considerável pesquisa, estudo e reflexão. [...] 2 Questão levantada para inquirição, consideração, discussão, decisão ou solução [...]. 3 Dificuldade ou obstáculo que requer grande esforço para ser solucionado ou vencido. 4 Situação conflitante; dificuldade[...].” (Michaelis, 2024).

FICHA TERMINOLÓGICA	
NÚMERO	25
TERMO	Raiva
SINÔNIMOS	danação, hidrofobia, lissa, moléstia, rábia
CONTEXTO	<p>“[...] Há algumas doenças que se espalham praticamente por todo o território nacional, como os acidentes ofídicos (as mordeduras de serpentes), e outras que estão restritas a pouquíssimos municípios, como a <b>raiva</b> humana[...].” (arquivo 005).</p> <p>“[...]A redução na incidência e a concentração dos casos da <b>raiva</b> humana transmitida por animais domésticos também permitem prever sua próxima eliminação[...].” (arquivo 015).</p>

	<p>“[...]Hanseníase, dengue, leishmaniose, esquistossomose, <i>raiva</i> humana transmitida por cães, escabiose (sarna), doença de Chagas, parasitoses intestinais e tracoma são algumas das mais de 20 patologias presentes na região – onde também são conhecidas como doenças infecciosas negligenciadas[...]" (arquivo 083).</p> <p>“[...]Entre as principais Doenças Tropicais Negligenciadas com casos no Brasil, estão: hanseníase, febre chikungunya, esquistossomose, filariose linfática, geo-helmintíases, oncocerose, tracoma, doença de Chagas, leishmanioses, <i>raiva</i>, hidatidose, escabiose (sarna), micetoma e cromoblastomicose[...]" (arquivo139).</p>
<b>DEFINIÇÃO</b>	<p>“1 MED, VET Doença infecciosa, causada por vírus, transmitida ao homem por mordedura de animais infectados, principalmente cães, atacando o sistema nervoso central e provocando convulsão, tetania e paralisia respiratória[...]" (Michaelis, 2024).</p> <p>“Uma das VIROSES DO SISTEMA NERVOSO CENTRAL, aguda, que afeta mamíferos, inclusive o homem. É causada por VÍRUS DA RAIVA e é geralmente disseminada pela contaminação com saliva carregada de vírus através de mordidas infligidas por animais que carregam o vírus. Animais vetores importantes incluem o cão, gato, morcego, raposa, guaxinim, gambá e lobo” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	26
<b>TERMO</b>	Rios
<b>SINÔNIMOS</b>	Flumes, afluxos, cursos de água
<b>CONTEXTO</b>	<p>“As esquistossomoses originaram-se nas bacias dos <i>rios</i> Nilo, na África, e do Yangtze, na Ásia[...]" (arquivo 021).</p> <p>“Também conhecida como 'cegueira dos <i>rios</i>', a oncocerose é transmitida por simulídeos que, ao se alimentarem de um hospedeiro infectado, sugam microfilárias junto com o sanguem[...]" (arquivo 061).</p> <p>“A esquistossomose ainda depende de estratégias de vigilância, como o monitoramento da presença dos caramujos em <i>rios</i>[...]" (arquivo 078).</p> <p>“A oncocerose ou "cegueira dos <i>rios</i>" é uma doença parasitária que causa lesões na pele e nos olhos. Se não for tratada, pode causar</p>

	<p>cegueira. Na América Latina, o território onde a comunidade indígena Yanomami vive – entre Brasil e Venezuela – é o último foco da doença.</p> <p>Estima-se que, no final da década de 1980, cerca de 500 mil pessoas no Brasil, Colômbia, Equador, Guatemala, México e Venezuela corriam o risco de contrair esta doença, que é transmitida pela picada de mosquitos infectados que se reproduzem em <b>rios</b> e córregos[...]" (arquivo 134).</p>
<b>DEFINIÇÃO</b>	<p>“A forma rios é masculino plural de rio.</p> <p>1. Grande curso de água natural, quase sempre oriunda das montanhas, que recebe no trajeto águas de regatos e ribeiros, e deságua em outro curso de água, num lago ou no mar” (Priberam, 2024).</p> <p>“Grandes correntes naturais de ÁGUA DOCE formadas pela convergência de afluentes e que desembocam em um grande volume de água (lago ou oceano)” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	27
<b>TERMO</b>	Saúde
<b>SINÔNIMOS</b>	bem-estar físico, vigor, energia
<b>CONTEXTO</b>	<p>“[...] Com baixo poder aquisitivo e sem influência política, os doentes e sistemas de <b>saúde</b> de países pobres não conseguem gerar o retorno financeiro exigido pela maior parte das empresas[...]" (arquivo 000).</p> <p>“[...] Destaca-se a intensificação das ações de vigilância e controle da tuberculose e hanseníase, integradas à sífilis e focadas nas equipes de <b>saúde</b> da família, visando a detecção precoce e tratamento adequado das pessoas[...]" (arquivo 063).</p> <p>“[...] As doenças negligenciadas persistem como um problema de <b>saúde</b> global atingindo, desproporcionalmente, as populações mais vulneráveis e impondo um fardo humano, social e econômico, devastador a mais de 1 bilhão de pessoas em todo o mundo[...]" (arquivo 086).</p> <p>“[...] Desta forma, é um dos grandes desafios da <b>saúde</b> humana atualmente desenvolver, de maneira eficaz, um tratamento e, sobretudo, uma prevenção contra a doença[...]" (arquivo 172).</p>
<b>DEFINIÇÃO</b>	“1 Estado do organismo com funções fisiológicas regulares e com características estruturais normais e estáveis, levando-se em

	<p>consideração a forma de vida e a fase do ciclo vital de cada ser ou indivíduo[...]" (Michaelis, 2024).</p> <p>"Estado do organismo quando funciona otimamente sem evidência de doença. Para a Homeopatia, diz respeito ao estado de bem-estar biopsicossocial do ser vivente, que reflete o equilíbrio da sua força vital, mantendo as características estruturais e funcionais do organismo dentro dos limites normais físicos e emocionais, para sua forma de vida e para a sua fase atual do ciclo vital" (DeCS/MeSH, 2024).</p>
--	--

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	28
<b>TERMO</b>	Sintomas
<b>SINÔNIMOS</b>	Sinal, indício, manifestação
<b>CONTEXTO</b>	<p>"Manchas na pele de cor parda, esbranquiçadas ou avermelhadas estão entre os principais <i>sintomas</i> da hanseníase" (arquivo 034).</p> <p>"Especificamente, os <i>sintomas</i> da oncocerose ocular iniciam com a fotofobia, seguindo para opacidade gradual da córnea e perda da visão[...]" (arquivo 082).</p> <p>"Os resultados mostraram que a população detém uma codificação para identificação de helmintoses, assim como conceituação própria sobre sinais e <i>sintomas</i>, aspectos epidemiológicos e formas de tratamento. Esses achados são de fundamental importância para o planejamento das campanhas e projetos educacionais em verminoses" (arquiv0 160).</p> <p>"Os <i>sintomas</i> do tracoma são sensibilidade e intolerância à luz, coceira nos olhos, sensação de corpo estranho nos olhos, vermelhidão, secreção, dor e lacrimejamento. O período entre o momento da infecção e o surgimento dos <i>sintomas</i> é de 5 a 12 dias. O diagnóstico é clínico-epidemiológico, realizado por exame ocular externo" (arquivo 166).</p>
<b>DEFINIÇÃO</b>	<p>"A forma sintomas é masculino plural de sintoma.</p> <p>1. Sinal que indica uma doença ou mudança no curso de uma doença" (Priberam, 2024).</p> <p>"Sinais e Sintomas</p>

	Manifestações clínicas que podem ser tanto objetivas (quando observadas por médicos) como subjetivas (quando percebidas pelo paciente)” (DeCS/MeSH, 2024).
--	--

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	29
<b>TERMO</b>	Tracoma
<b>SINÔNIMOS</b>	Conjuntivite granulomatosa
<b>CONTEXTO</b>	<p>“[...] O <b>tracoma</b> é a causa mais frequente de cegueira evitável, representando 3% da cegueira do mundo. Globalmente, 21,4 milhões de pessoas têm <b>tracoma</b>, e 1,2 milhões delas são cegas[...].” (arquivo 009).</p> <p>“[...] Companheiras próximas da miséria, essas doenças também mantêm numerosas populações em condições de pobreza. Oncocerose e <b>tracoma</b> causam cegueira[...].” (arquivo 020).</p> <p>“[...] O <b>tracoma</b> é considerado uma doença crônica e atinge em sua maioria crianças em idade escolar que residem em regiões com maior concentração da pobreza[...].” (arquivo 027).</p> <p>“[...] O estado de Pernambuco prevê estratégias de intervenção para redução e eliminação das seguintes: doença de Chagas, hanseníase, esquistossomose, <b>tracoma</b>, filariose linfática, geohelmintoses e tuberculose[...].” (arquivo 158).</p>
<b>DEFINIÇÃO</b>	<p>“MED Infecção contagiosa, caracterizada pela presença de minúsculas granulações sobre a mucosa ocular, provocada pela bactéria Chlamydia trachomatis” (Michaelis, 2024).</p> <p>“Infecção crônica da TÚNICA CONJUNTIVA e da CÓRNEA causada por CHLAMYDIA TRACHOMATIS” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	30
<b>TERMO</b>	Transmissão
<b>SINÔNIMOS</b>	Difusão, irradiação, propagação, contágio, infecção, contaminação

<b>CONTEXTO</b>	<p>“ [...] O Brasil tem avançado no processo de redução da carga de DTNs, como a Doença de Chagas, que foi incluída na estratégia de eliminação da <b>transmissão</b> vertical, e do tracoma” (arquivo 006).</p> <p>“A esquistossomose, também conhecida como barriga d'água, é caracterizada como uma das doenças negligenciadas. Ela é causada por um parasita que se aloja nas veias do sistema porta, que conecta o intestino ao fígado. Sua <b>transmissão</b> ocorre por meio de caramujos de água doce infectados, que liberam larvas na água” (arquivo 039).</p> <p>“O tratamento da hanseníase é feito a partir da associação de três antibióticos (rifampicina, dapsona e clofazimina). A indicação do tratamento depende da condição clínica de cada paciente, podendo ser de seis meses a um ano. O tratamento melhora a qualidade de vida, reduz os riscos de complicações e interrompe a <b>transmissão</b> da doença” (arquivo 095).</p> <p>“Oncocercose é uma infecção causada pelo nematódeo filarioide por Onchocerca volvulus. Borrachudos (espécies Simulium) são responsáveis pela <b>transmissão</b>[... ]” (arquivo 121).</p>
<b>DEFINIÇÃO</b>	<p>“1 Ação ou efeito de transmitir(-se). [...] 7 MED Propagação de um vírus ou doença por contato com pessoa infectada” (Michaelis, 2024).</p> <p>“Transmissão de Doença Infecciosa – conceito preferido Transmissão de doença ou patógenos infecciosos. Quando a transmissão for dentro da mesma espécie, o modo pode ser [chamado] horizontal ou vertical (TRANSMISSÃO VERTICAL DE DOENÇA INFECCIOSA)” (DeCS/MeSH, 2024).</p>

FICHA TERMINOLÓGICA	
<b>NÚMERO</b>	31
<b>TERMO</b>	Tratamento
<b>SINÔNIMOS</b>	remédio, procedimento, terapia
<b>CONTEXTO</b>	<p>“ [...] O <b>tratamento</b> é disponibilizado gratuitamente. O não tratamento da doença pode levar à morte[...]” (arquivo 015).</p> <p>“ [...] O <b>tratamento</b> preconizado pelo Ministério da Saúde é o antibiótico Azitromicina na dose de 20mg/Kg de peso em dose única, via oral, dose máxima 1g. O SUS disponibiliza tratamento gratuito de azitromicina nas apresentações em comprimidos de 500mg e suspensão de 600mg[...]” (arquivo 027).</p>

	<p>“[...]O <b>tratamento</b> melhora a qualidade de vida, reduz os riscos de complicações e interrompe a transmissão da doença[...].” (arquivo 095).</p> <p>“[...] Na última década, os doadores e os programas nacionais tornaram-se cada vez mais envolvidos e o acesso ao <b>tratamento</b> foi ampliado[...].” (arquivo 122).</p>
<b>DEFINIÇÃO</b>	<p>“1 Ato ou efeito de tratar alguém, alguma coisa ou a si mesmo; trato [...]3 MED Conjunto de meios de que lança mão o médico para cura ou alívio do doente [...]” (Michaelis, 2024).</p> <p>“Terapêutica – Conceito preferido Procedimentos com interesse no tratamento curativo ou preventivo de doenças” (DeCS/MeSH, 2024).</p>

<b>FICHA TERMINOLÓGICA</b>	
<b>NÚMERO</b>	32
<b>TERMO</b>	Tuberculose
<b>SINÔNIMOS</b>	tísica, febre hética
<b>CONTEXTO</b>	<p>“[...]Embora as doenças tropicais e a <b>tuberculose</b> sejam responsáveis por 11,4% da carga global de doença, apenas 21 (1,3%) dos 1.556 novos medicamentos registrados entre 1975 e 2004, foram desenvolvidos especificamente para essas doenças[...].” (arquivo 004).</p> <p>“[...] Não há uma vacina específica para prevenir a hanseníase. No entanto, a vacina BCG, normalmente aplicada no nascimento para prevenir <b>tuberculose</b>, também reduz o risco de hanseníase, pois os agentes causadores de ambas as doenças são semelhantes[...].” (arquivo 034).</p> <p>“[...] Desenvolvimento Sustentável (ODS) da Organização das Nações Unidas (ONU). A iniciativa tem como meta o combate às epidemias da doença e de outros agravos transmissíveis, como <b>tuberculose</b> e malária até 2030[...].” (arquivo 095).</p> <p>“[...] Oitenta por cento dos casos de <b>tuberculose</b> no mundo estão concentrados em 22 países, entre os quais o Brasil é o único americano presente, segundo a OMS[...].” (arquivo 140).</p>
<b>DEFINIÇÃO</b>	“1 Doença infectocontagiosa causada pelo bacilo de Koch ( <i>Mycobacterium tuberculosis</i> ), que pode afetar qualquer órgão ou

	<p>tecido, especialmente os pulmões, caracterizada pela formação de tubérculos caseosos” (Michaelis, 2024).</p> <p>“Qualquer uma das doenças infecciosas do ser humano e de outros animais causadas por espécies de MYCOBACTERIUM TUBERCULOSIS” (DeCS/MeSH, 2024).</p>
--	--

## **APÊNDICE B – Links/sites do *corpus* levantado pelo BootCat utilizados nas fichas terminológicas**

arquivo 000

[http://cienciaecultura.bvs.br/scielo.php?script=sci\\_arttext&pid=S0009-67252010000100003](http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252010000100003)

arquivo 002

[http://www.diaadiaeducacao.pr.gov.br/portals/cadernospde/pdebusca/producoes\\_pde/2014/2014\\_unicentro\\_port\\_pdp\\_mirian\\_izabel\\_tullio.pdf](http://www.diaadiaeducacao.pr.gov.br/portals/cadernospde/pdebusca/producoes_pde/2014/2014_unicentro_port_pdp_mirian_izabel_tullio.pdf)

arquivo 004

<https://agencia.fiocruz.br/doen%C3%A7as-negligenciadas>

arquivo 005

<https://www.bbc.com/portuguese/articles/c25qy2yn9evo>

arquivo 006

<https://agenciaaids.com.br/noticia/oms-17-bilhao-de-pessoas-no-mundo-podem-estar-sobrisco-das-doencas-tropicais-negligenciadas/>

arquivo 007

<https://agenciabrasil.ebc.com.br/saude/noticia/2022-01/doencas-negligenciadas-tem-aumento-em-taxa-de-mortalidade#:~:text=ouvir%3A,29%2C3%25%20nas%20interna%C3%A7%C3%B5es.>

arquivo 008

<https://m.folha.uol.com.br/cotidiano/2013/01/1210971-areas-pobres-do-pais-ainda-sofrem-com-doencas-esquecidas.shtml>

arquivo 009

<https://consulta.camara-arq.sp.gov.br/arquivo?Id=276127>

arquivo 012

[https://brasil.elpais.com/brasil/2015/11/04/ciencia/1446663259\\_269082.html](https://brasil.elpais.com/brasil/2015/11/04/ciencia/1446663259_269082.html)

arquivo 013

[https://brasil.elpais.com/brasil/2015/11/24/politica/1448394485\\_338466.html](https://brasil.elpais.com/brasil/2015/11/24/politica/1448394485_338466.html)

arquivo 014

<https://brasil.un.org/pt-br/170204-doen%C3%A7as-tropicais-negligenciadas-opas-pede-fim-dos-atrasos-no-tratamento-nas-am%C3%A9ricas>

arquivo 015

<https://www.saude.ao/esquistossomose/>

arquivo 016

<https://bvsms.saude.gov.br/agir-agora-agir-juntos-investir-em-dtns-30-01-dia-mundial-das-doencas-tropicais-negligenciadas/#:~:text=IN%C3%8DCIO,%E2%80%9CAgir%20agora.,pobres%20e%20marginalizadas%20do%20mundo.>

arquivo 017

[https://bvsms.saude.gov.br/bvs/publicacoes/caderno\\_tematico\\_pse\\_doencas\\_negligenciadas.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/caderno_tematico_pse_doencas_negligenciadas.pdf)

arquivo 018

[https://bvsms.saude.gov.br/bvs/publicacoes/doencas\\_infecciosas\\_parasitaria\\_guia\\_bolso.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/doencas_infecciosas_parasitaria_guia_bolso.pdf)

arquivo 019

[https://bvsms.saude.gov.br/bvs/publicacoes/guia\\_de\\_hanseniese.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/guia_de_hanseniese.pdf)

arquivo 020

[https://bvsms.saude.gov.br/bvs/publicacoes/primeiro\\_relatorio\\_oms\\_doencas\\_tropicais.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/primeiro_relatorio_oms_doencas_tropicais.pdf)

arquivo 021

[https://bvsms.saude.gov.br/bvs/publicacoes/vigilancia\\_esquistosomae\\_mansoni\\_diretrizes\\_tecnicas.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/vigilancia_esquistosomae_mansoni_diretrizes_tecnicas.pdf)

arquivo 022

<https://bvsms.saude.gov.br/hanseniese-9/>

arquivo 023

<https://bvsms.saude.gov.br/unir-agir-eliminar-30-01-dia-mundial-das-doencas-tropicais-negligenciadas/>

arquivo 027

<https://cosemsto.org.br/Noticias/V/saude-chama-a-atencao-da-populacao-para-a-prevencao-e-diagnostico-do-tracoma>

arquivo 034

<https://drauziovarella.uol.com.br/doencas-e-sintomas/hansenise-lepra/>

arquivo 035

<https://drauziovarella.uol.com.br/doencas-e-sintomas/tracoma/>

arquivo 038

<https://editoraomnisscientia.com.br/catalogos/ciencias-da-saude/doencas-negligenciadas-hansenise/>

arquivo 039

[https://educacaomedica.afya.com.br/blog/doencas-tropicais-negligenciadas-ainda-encontradas-no-brasil?utm\\_source=google&utm\\_medium=organic](https://educacaomedica.afya.com.br/blog/doencas-tropicais-negligenciadas-ainda-encontradas-no-brasil?utm_source=google&utm_medium=organic)

arquivo 040

<https://educapes.capes.gov.br/bitstream/capes/206104/2/parasitos.pdf>

arquivo 041

<https://fapergs.rs.gov.br/upload/arquivos/carga20171220/14092030-doencas-tropicais.pdf>

arquivo 043

<https://gizmodo.uol.com.br/lista-da-onu-com-25-doencas-negligencias-e-puxao-de-orelha-no-brasil/>

arquivo 045

<https://iris.paho.org/handle/10665.2/34290>

arquivo 052

<https://www.bbc.com/portuguese/articles/c25qy2yn9evo>

arquivo 054

<https://agenciabrasil.ebc.com.br/saude/noticia/2022-01/doencas-negligenciadas-tem-aumento-em-taxa-de-mortalidade>

arquivo 061

<https://agencia.fiocruz.br/pesquisadora-da-fiocruz-fala-sobre-oncocerose-nas-am%C3%A9ricas>

arquivo 063

<https://portal.saude.pe.gov.br/programa/secretaria-executiva-de-vigilancia-em-saude/programa-sanar-doencas-negligenciadas>

arquivo 064

<https://portugues.medscape.com/verartigo/6510087?form=fpf>

arquivo 065

<https://pressreleases.scielo.org/blog/2022/06/30/complicacoes-cardiacas-da-esquistossomose-podem-causar-incapacidade-a-longo-prazo-e-morte/#:~:text=A%20conclus%C3%A3o%20foi%20que%20as,para%20detectar%20esse%20envolvimento%20cardiovascular.>

arquivo 067

[https://pt.wikipedia.org/wiki/Chlamydia\\_trachomatis](https://pt.wikipedia.org/wiki/Chlamydia_trachomatis)

arquivo 068

<https://pt.wikipedia.org/wiki/Oncocerose>

arquivo 071

[https://repositorio.ipea.gov.br/bitstream/11058/1577/1/td\\_1607.pdf](https://repositorio.ipea.gov.br/bitstream/11058/1577/1/td_1607.pdf)

arquivo 078

<https://revistapesquisa.fapesp.br/uma-agenda-para-as-doencas-esquecidas/>

arquivo 080

<https://revistapesquisa.fapesp.br/resgate-das-doencas-esquecidas/>

arquivo 081

<https://revistapesquisa.fapesp.br/uma-agenda-para-as-doencas-esquecidas/>

arquivo 083

<https://bvsms.saude.gov.br/agir-agora-agir-juntos-investir-em-dtns-30-01-dia-mundial-das-doencas-tropicais-negligenciadas/>

arquivo 085

<https://www.arca.fiocruz.br/bitstream/handle/icict/18687/;jsessionid=ECC631D0C23AEC368BFE6685E24DA0FE?sequence=2>

arquivo 086

[https://www.arca.fiocruz.br/bitstream/handle/icict/51650/paula\\_teixeira\\_pinto\\_et\\_all.pdf;jsessionid=87F24C5957BE506BCF6120ACAC74B3F0?sequence=2](https://www.arca.fiocruz.br/bitstream/handle/icict/51650/paula_teixeira_pinto_et_all.pdf;jsessionid=87F24C5957BE506BCF6120ACAC74B3F0?sequence=2)

arquivo 087

<https://www.atenaeditora.com.br/catalogo/dowload-post/79649>

arquivo 092

<https://bio.fiocruz.br/index.php/en/noticias/2245-doencas-tropicais-negligenciadas>

arquivo 095

<https://www.cnnbrasil.com.br/saude/por-que-a-hansenise-ainda-persiste-como-problema-de-saude-publica-no-brasil/>

arquivo 096

<https://agenciabrasil.ebc.com.br/saude/noticia/2018-10/oms-1-bilhao-de-pessoas-sao-tratadas-contra-doencas-negligenciadas>

arquivo 097

<https://www.arca.fiocruz.br/handle/icict/3940>

arquivo 101

<https://www.dw.com/pt-br/doen%C3%A7as-tropicais-negligenciadas-um-problema-global/a-64558878>

arquivo 104

<https://www.fiocruzbrasilia.fiocruz.br/fiocruz-tem-pesquisadores-citados-em-ranking-mundial-de-cientistas-que-mais-contribuiram-no-combate-a-hanseniese-e-outras-doencas-negligenciadas/>

arquivo 111

<http://antigo.aids.gov.br/pt-br/o-que-e-hanseniese>

arquivo 113

<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/o/oncocercose/situacao-epidemiologica>

arquivo 114

<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/t/tracoma>

arquivo 118

<https://www.ioc.fiocruz.br/noticias/servico-de-referencia-em-oncocercose-recebe-diretora-do-ministerio-da-saude>

arquivo 119

[https://www.mds.gov.br/webarquivos/publicacao/brasil\\_sem\\_miseria/livro\\_o\\_brasil\\_sem\\_miseria/artigo\\_28.pdf](https://www.mds.gov.br/webarquivos/publicacao/brasil_sem_miseria/livro_o_brasil_sem_miseria/artigo_28.pdf)

arquivo 122

<https://www.msf.org.br/noticias/doencas-tropicais-negligenciadas/>

arquivo 130

<https://www.nhrbrasil.org.br/nhr-brasil/nossas-causas/doencas-tropicais-negligenciadas.html>

arquivo 131

<https://www.novartis.com/br-pt/esg/doencas-negligenciadas>

arquivo 132

<https://www.ourofinosaudedanimal.com/ourofinoemcampo/categoria/artigos/parasitas-internos-que-acometem-os-caes-e-o-que-el/>

arquivo 134

<https://www.paho.org/pt/noticias/24-7-2018-oncocerose-ultimo-foco-da-doenca-ocorre-em-comunidade-indigena-entre-brasil-e>

arquivo 135

<https://brasil.un.org/pt-br/170204-doen%C3%A7as-tropicais-negligenciadas-opas-pede-fim-dos-atrasos-no-tratamento-nas-am%C3%A9ricas#:~:text=O%20M%C3%A9xico%20p%C3%B4s%20fim%20%C3%A0,vermes%20parasitas%2C%20nos%20pr%C3%B3ximos%20anos.>

arquivo 139

<https://www.gov.br/saude/pt-br/assuntos/noticias/2023/janeiro/dtns-brasil-tem-mais-de-90-dos-novos-casos-de-hansenise-registrados-nas-americas>

arquivo 140

<https://www.redebrasilitual.com.br/revistas/a-febre-da-exclusao/>

arquivo 144

<https://www.saude.ba.gov.br/suvisa/vigilancia-epidemiologica/hansenise/>

arquivo 155

<https://www.scielo.br/j/csp/a/5BfN4vHxrsS9jj8BSf66ncK/abstract/?lang=pt>

arquivo 156

<https://www.scielo.br/j/csp/a/Wqv3BjYjSC37rcVL68PWVLM/?lang=pt>

arquivo 158

<https://www.scielo.br/j/jped/a/C5LryrRmdnDPb5BmbGGyR9q/>

arquivo 160

<https://www.scielo.br/j/rsp/a/5dXN5vPnchskvG7JvnsMxmG/>

arquivo 161

<https://rsp.fsp.usp.br/artigo/analise-cienciometrica-da-pesquisa-sobre-tracoma-no-brasil-2000-2020/>

arquivo 162

<https://www.scielo.br/j/sausoc/a/HcVVcJn5MrvZqRFVPHxRCG/abstract/?lang=pt>

arquivo 163

<https://www.scielo.br/j/sausoc/a/nXWpzPJ5pfHMDmKZBqkSZMx/abstract/?lang=pt>

arquivo 166

<https://www.to.gov.br/secom/noticias/saude-reforca-cuidados-de-prevencao-e-conscientizacao-sobre-o-tracoma/cits6t4jst5>

arquivo 167

<https://www.tuasaude.com/onchocerca-volvulus/>

arquivo 168

<https://www.uol.com.br/vivabem/noticias/deutsche-welle/2023/01/30/doencas-tropicais-negligenciadas-um-problema-global.htm>

arquivo 169

<https://www.uol.com.br/vivabem/noticias/deutsche-welle/2022/01/30/doencas-negligenciadas-matam-500-mil-por-ano-no-mundo.htm>

arquivo 170

<https://www.bbc.com/portuguese/articles/c25qy2yn9evo#:~:text=O%20aumento%20m%C3%A9dico%20da%20temperatura,e%20chikungunya%20%E2%80%94%20o%20Aedes%20aegypti.>

arquivo 172

<https://www2.ifsc.usp.br/portal-ifsc/a-fisica-e-as-doencas-negligenciadas-o-exemplo-das-esquistossomose/>

arquivo 174

<https://www.estadao.com.br/opiniao/guerra-as-doencas-tropicais-imp-/>

arquivo 175

<https://www3.paho.org/hq/dmdocuments/2013/CD52-INF4-G-p.pdf>

arquivo 177

<https://www5.iqsc.usp.br/2021/doencas-tropicais-negligenciadas-quando-elas-terao-a-devida-atencao/>