



Relatório - Trabalho Prático 2

Processamento de Linguagem Natural em Engenharia Biomédica

Mestrado em Engenharia Biomédica - Informática Médica

Grupo:

Afonso Rodrigues (PG55831)
João Esteves Gil (PG56132)
Maria Inês Eusébio (PG57586)

Docentes:

José Almeida
Luís Filipe Cunha



Índice

1	Introdução	1
2	Enriquecimento do Dataset	1
3	Plataforma Web	3
3.1	Gestão e armazenamento dos conceitos	4
3.2	Descrição das Funcionalidades da Aplicação Web	4
3.2.1	Homepage	4
3.2.2	Conceitos	5
3.2.3	Informações do Conceito	6
3.2.4	Eliminar Conceito	8
3.2.5	Editar Conceito	8
3.2.6	Adicionar Novo Conceito	9
3.2.7	Categorias	10
3.2.8	Consulta de Conceitos por Categoria	11
3.2.9	Tabela de Conceitos	12
3.2.10	Relações entre Conceitos	12
3.2.11	Pesquisa Inversa	13
3.2.12	Pesquisa Avançada	14
3.2.13	Lista de Sinónimos	15
3.2.14	Jogo de Sinónimos	16
3.2.15	Barra de Navegação	17
4	Sugestões e Melhorias	17
5	Conclusão	18



1 Introdução

No âmbito da unidade curricular de *Processamento de Linguagem Natural em Engenharia Biomédica*, foi desenvolvido um trabalho prático com o objetivo de explorar técnicas de processamento e enriquecimento de dados textuais aplicados a um glossário biomédico.

Inicialmente, procedeu-se ao enriquecimento do *dataset* original, através da adição de diversas informações complementares a cada conceito. Entre as operações realizadas destacam-se a integração de ligações diretas ao *Google Scholar* para facilitar o acesso a literatura científica relevante, a identificação automática de sinónimos em português e inglês utilizando a biblioteca *NLTK* com o recurso ao *WordNet*, e o cálculo de palavras semanticamente próximas através do treino de um modelo *Word2Vec* sobre os próprios conceitos e descrições existentes.

Para além disso, foi ainda realizado um processo de uniformização e categorização dos conceitos, recorrendo ao apoio de *Large Language Models* (LLMs) para a atribuição automática de categorias a termos que não dispunham dessa informação.

Posteriormente, com os dados enriquecidos, foi desenvolvida uma plataforma *web* interativa, implementada em *Flask*, que permite ao utilizador visualizar, editar, adicionar e explorar os conceitos e as respetivas relações de forma dinâmica.

O trabalho desenvolvido permitiu consolidar os conhecimentos adquiridos ao longo da unidade curricular, aplicando técnicas de processamento de linguagem natural, organização e visualização de dados, bem como desenvolvimento de aplicações web interativas.

2 Enriquecimento do Dataset

De forma a ser possível trabalhar e explorar o dataset de uma forma mais completa, assim como promover todo o seu potencial e valorizar toda a informação que este contém, foi realizado um processo de "enriquecimento do dataset". Este processo permitiu que mais (e nova) informação fosse adicionada aos conceitos já existentes, a fim de completar tudo o que já existia. Este conjunto de processos é relevante para que, numa fase seguinte, seja possível criar uma aplicação web onde todos estes dados estarão presentes e acessíveis, da forma mais completa possível, a todos os utilizadores da mesma.

Inicialmente, e de forma a que um utilizador, caso deseje, possa ter acesso a informações complementares acerca do conceito pesquisado, foi adicionado a cada um destes um link que fará uma pesquisa automática do conceito no Google Scholar. Lá, o utilizador terá um acesso direto e facilitado a um conjunto de artigos científicos e outros artigos que poderão ser relevantes aquando da pesquisa.

```
if nome:  
    conceito["Link Google Scholar"] =  
  
    f"https://scholar.google.com/scholar?q={nome.replace(' ', '+')}"
```

Foi também, para cada conceito, criado um bloco que procura sinónimos em português e inglês usando o WordNet (através do NLTK). Os sinónimos encontrados são adicionados ao campo "Sinónimos", criado para o efeito, caso existam. Caso não seja encontrado nenhum sinónimo para o termo em análise, o campo é retornado com uma lista vazia.



```
for conceito in glossario:  
    nome = conceito.get("Conceito", "")  
    sinonimos_pt = get_synonyms(nome, 'por')  
    sinonimos_en = get_synonyms(nome, 'eng')  
    sinonimos = list(set(sinonimos_pt + sinonimos_en))  
    conceito["Sinónimos"] = sinonimos
```

O script tem a capacidade de treinar um modelo Word2Vec com os nomes e descrições dos conceitos. Para cada conceito, identifica as palavras mais próximas (semanticamente semelhantes) ao nome do conceito, excluíndo o próprio termo, e adiciona-as ao campo "Palavras próximas". De forma a que se perceba a "distância" entre as palavras, também é associado a cada "termo mais próximo" o score calculado.

```
model = Word2Vec(frases, vector_size=100, window=5, min_count=1,  
sg=1, epochs=10, workers=3)  
  
for conceito in tqdm(glossario, desc="Processando conceitos"):  
    nome = conceito.get("Conceito", "")  
    palavras_proximas = []  
    if nome:  
        tokens = [token.lower() for token in word_tokenize(nome,  
language='portuguese') if token.isalpha()]  
        similares = {}  
        for token in tokens:  
            try:  
                similares += [(w, round(score, 2)) for w,  
score in model.wv.most_similar(token, topn=10)]  
            except KeyError:  
                continue  
        # Remove duplicados e o próprio termo, pega os 5 primeiros  
        palavras_proximas = [w for w in dict.fromkeys(similares)  
  
        if w not in tokens][:5]  
    conceito["Palavras próximas"] = palavras_proximas
```

Vendo que os conceitos foram retirados de várias fontes diferentes, cada um contém um conjunto de informações que se pretendia que fossem o mais uniforme possível. Foi considerado que a categoria de cada termo seria um campo de relevante interesse para análise. Sabendo que nem todos os termos o possuíam, foram criados, com auxílio de um Large Language Model (LLM) um conjunto de categorias para os termos que ainda não as possuíam, contando também com aqueles que já eram disponibilizados.

```
tem = []  
naotem = []
```

```
for item in dados:
    conceito_nome = item.get('Conceito')
    fontes = item.get('Fontes', {})
    categoria_encontrada = None
    for fonte_info in fontes.values():
        if 'Categoria' in fonte_info:
            categoria_encontrada = fonte_info['Categoria']
            break
    if categoria_encontrada:
        tem.append((conceito_nome, categoria_encontrada))
    else:
        naotem.append(conceito_nome)

print("Com categoria:", tem)
print("Sem categoria:", naotem)
```

Também com recurso ao LLM, estes conceitos foram associados à categoria que oferecia melhor correspondência. Assim, todos os termos passam a ter uma categoria associada e todos são passíveis de serem analisados através deste campo.

```
fontes = conceito.get('Fontes', {})
categoria = None
for fonte in fontes.values():
    if 'Categoria' in fonte:
        categoria = fonte['Categoria']
        break
if not categoria:
    categoria = 'sem categoria'
if categoria not in categorias:
    categorias[categoria] = []
categorias[categoria].append(conceito)
time.sleep(0.1)
```

3 Plataforma Web

Com o objetivo de permitir a interação com os conceitos presentes no glossário, foi desenvolvida uma aplicação web. A implementação desta aplicação recorreu ao *framework Flask*, devido à sua simplicidade de aprendizagem e à sua ampla versatilidade para desenvolvimento rápido de aplicações web.

A plataforma disponibiliza um conjunto de funcionalidades organizadas em diferentes páginas (rotas), permitindo ao utilizador aceder, pesquisar, visualizar, editar e eliminar conceitos, bem como explorar relações entre palavras próximas, descobrir sinónimos e realizar pesquisas inversas de forma intuitiva e interativa.



3.1 Gestão e armazenamento dos conceitos

Os conceitos utilizados na aplicação web são armazenados num ficheiro em formato *JSON*, que serve como base de dados para o carregamento e persistência dos dados do glossário. Este formato permite uma estrutura simples e flexível, facilitando a organização dos conceitos por categorias e respetivos atributos.

Para manipular estes dados, foram implementadas funções de leitura e escrita do ficheiro *JSON*, garantindo a correta codificação dos caracteres e a preservação da informação de forma estruturada.

A leitura dos conceitos é efetuada através da função `carregar_conceitos(file)`, que abre o ficheiro especificado e carrega o seu conteúdo para memória utilizando a função `json.load()`. Esta operação permite disponibilizar os dados para posterior processamento e apresentação na aplicação.

```
def carregar_conceitos(file):
    with open(file, 'r', encoding='utf-8') as file:
        conceitos = json.load(file)
    return conceitos
```

Adicionalmente, foi implementada a função `guardar_conceitos(conceitos, filename)` que permite gravar alterações ou atualizações efetuadas ao glossário, escrevendo novamente os dados num novo ficheiro *JSON*.

```
def guardar_conceitos(conceitos, filename):
    with open(filename, 'w', encoding='utf-8') as file:
        json.dump(conceitos, file, ensure_ascii=False, indent=4)
```

3.2 Descrição das Funcionalidades da Aplicação Web

Com o objetivo de melhor detalhar o funcionamento da aplicação desenvolvida, de seguida são apresentadas e descritas as diferentes páginas e funcionalidades disponibilizadas na plataforma, bem como o seu papel na interação com o glossário de conceitos.

3.2.1 Homepage

A página *Homepage* constitui o ponto de entrada principal da aplicação web, oferecendo uma visão geral sobre o propósito e funcionalidades do *Atlas de Conceitos Médicos e Clínicos*. Com um design apelativo e intuitivo, apresenta um cabeçalho com o título da aplicação e uma breve descrição do seu objetivo, destacando a disponibilização de conhecimento médico estruturado, acessível e multilingue.

Para facilitar a navegação, a *Homepage* disponibiliza um painel de botões centrais que direcionam o utilizador para as principais funcionalidades da plataforma, nomeadamente: consulta de conceitos, categorias, palavras próximas, pesquisa inversa, visualização tabular dos conceitos, lista de sinónimos, adição de novos conceitos e o jogo de sinónimos. Desta forma, o utilizador pode rapidamente aceder à funcionalidade pretendida de forma eficiente e organizada.



Figura 1: Homepage da plataforma.

3.2.2 Conceitos

A página *Conceitos* apresenta uma listagem alfabética dos termos existentes no glossário, permitindo ao utilizador consultar e explorar os diferentes conceitos de forma organizada e intuitiva.

Ao aceder à página, os conceitos são agrupados automaticamente pela sua letra inicial, permitindo a sua visualização por ordem alfabética. No topo da página, é disponibilizado um índice de letras, que permite ao utilizador navegar rapidamente para a secção correspondente a cada letra.

Para facilitar a pesquisa de um termo específico, a página disponibiliza também uma caixa de pesquisa interativa, onde o utilizador pode introduzir o termo desejado, filtrando de forma dinâmica a lista de conceitos apresentados.

Cada conceito surge como uma ligação clicável. Ao selecionar um conceito, o utilizador é redirecionado para uma página individual, dedicada à apresentação detalhada da informação associada a esse termo. Esta página de detalhe será abordada na secção seguinte do presente relatório.

Adicionalmente, a interface apresenta um botão flutuante de *scroll*, que surge automaticamente quando o utilizador percorre a página, permitindo regressar ao topo de forma rápida e intuitiva.

A implementação desta funcionalidade recorreu à organização dos conceitos num dicionário, agrupados pela sua letra inicial. Este agrupamento permite gerar automaticamente os títulos de cada secção e o índice de letras apresentado no topo da página. A página é gerada de forma dinâmica através de um *template* parametrizado, onde são percorridos todos os grupos de letras e respetivos conceitos. Cada conceito é associado a uma hiperligação dinâmica, cuja rota corresponde ao seu nome, permitindo a navegação para a respetiva página de detalhe através de URLs.



Índice de Letras

A B C D E F G H I J K L M N O P Q R S T U V W X Z

A

Ad20
Aids II
Amy-101
Arc
Asr09f
Atpr1923
Art, Artodidínia
Abordagem Médica Tradicional Do Adulto Hospitalizado
Abuso Incentivo
Abuso Financeiro Dos Idosos
Abuso Sexual Na Infância, Abuso Sexual Na Adolescência
Agão Racional
Acidentes Ampliados
Acidente De Trabalho
Acidente De Trabalho Fatais
Acidente De Trabalho Graves
Acidente De Trânsito
Acidentes De Transporte
Acidos Graxos
Agões Estratégicas
Avançamento Do Crescimento E Desenvolvimento Infantil

Figura 2: Página de listagem dos conceitos organizada alfabeticamente.

3.2.3 Informações do Conceito

A página de *Informações do Conceito* apresenta, de forma individualizada, as informações específicas associadas a cada conceito existente no glossário. Esta página pode ser acedida a partir de diferentes listagens de conceitos disponíveis na aplicação, nomeadamente a listagem alfabetica, a tabela de conceitos, as categorias, ou outras visualizações, dependendo do contexto de navegação do utilizador.

Nesta página, são apresentados vários elementos descritivos do conceito, nomeadamente:

- **Categoria:** categoria associada ao conceito.
- **Descrição:** definição e explicação do termo.
- **Citação:** referência bibliográfica associada (caso exista).
- **Sigla:** sigla ou acrônimo correspondente (caso aplicável).
- **Traduções:** traduções do conceito para diferentes idiomas.
- **Palavras próximas:** termos semanticamente relacionados com o conceito e respetivo valor de proximidade.
- **Sinónimos:** termos alternativos ou equivalentes ao conceito.
- **Google Scholar:** ligação direta para consulta adicional na plataforma Google Scholar (quando disponível).

Adicionalmente, a página disponibiliza um conjunto de botões de navegação que permitem ao utilizador retroceder à página anterior, editar as informações do conceito ou proceder à sua eliminação, sendo neste último caso solicitada uma confirmação antes da remoção definitiva do conceito.

A implementação da funcionalidade de consulta de conceitos consiste na leitura dos dados dos conceitos, seguindo-se a pesquisa pelo conceito pretendido. O sistema percorre as



categorias e os respetivos conceitos, comparando o nome do conceito com o termo especificado na rota, de forma a localizar o item correspondente. Após a identificação do conceito, são extraídas e tratadas as respetivas informações, incluindo os eventuais sinónimos existentes, sendo estas posteriormente enviadas para o *template* responsável por gerar a página de detalhe.

The screenshot shows a web application interface for managing medical concepts. At the top, there's a navigation bar with links: Home, Conceitos, Categorias, Palavras Próximas, Pesquisa Inversa, and Tabela. Below the navigation, the main content area has a dark header bar with the text "Conceito: Mal de cap". Underneath this, there's a section titled "Categoria: Sintomas e Sinais". A large blue capsule-shaped icon is positioned on the left side of the content area. The main content includes a "Descrição" field containing the text "Dor localizada al cap, que pode ter diversas causas, das tensões musculares migrâncias finas a problemas neurológicos e infecções.", a "Síntese" section with the text "Sem Síntese Associda", and a "Google Scholar" link. Below these, there's a "Traduções por idioma:" section listing translations in various languages. Further down is a "Palavras Próximas:" section with a list of terms and their scores. At the bottom of the page are three buttons: "Retroceder", "Editar", and "Eliminar". To the right of the content area, there's a large blue silhouette of a microscope.

Figura 3: Exemplo 1 de uma página de descrição de um conceito.

This screenshot shows another example of a concept description page. The layout is similar to Figure 3, with a dark header bar and a main content area. The header bar contains links: Home, Conceitos, Categorias, Palavras Próximas, Pesquisa Inversa, and Tabela. The main content area has a dark header bar with the text "Conceito: Adolescente". A large blue capsule-shaped icon is on the left. The content includes a "Categoria: Ciências Sociais em Saúde", a "Descrição" field with the text "Considera-se adolescente a pessoa entre 12 e 18 anos de idade.", a "Google Scholar" link, a "Palavras Próximas:" section with a list of terms and scores, and a "Sinónimos:" section with a list of terms. At the bottom are "Retroceder", "Editar", and "Eliminar" buttons. To the right of the content area is a large blue silhouette of a microscope.

Figura 4: Exemplo 2 de uma página de descrição de um conceito.

Foi também necessário desenvolver a função `parse_traducoes` para uniformizar as traduções associadas aos conceitos, independentemente do formato em que são recebidas (texto, lista ou dicionário). Em todos os casos, a função converte as traduções para um dicionário padronizado, com os idiomas como chaves e os termos capitalizados. Esta normalização facilita a apresentação consistente das traduções na aplicação.



3.2.4 Eliminar Conceito

A aplicação permite a remoção de conceitos individuais através da funcionalidade de eliminação. Ao aceder à página de detalhe de um conceito, o utilizador pode optar por eliminar o conceito selecionado. Quando esta ação é confirmada, o sistema percorre os dados armazenados, localiza o conceito correspondente ao nome indicado na rota e remove-o da respetiva lista. Antes de efetuar a eliminação, é automaticamente criada uma cópia de segurança do ficheiro original. Caso o conceito seja encontrado e eliminado com sucesso, o ficheiro de dados é atualizado e o utilizador é redirecionado para a listagem principal dos conceitos. Se o conceito não for encontrado, é devolvida uma mensagem de erro apropriada.

3.2.5 Editar Conceito

A aplicação disponibiliza uma página dedicada à edição dos dados de cada conceito, permitindo ao utilizador atualizar de forma simples as informações existentes. Ao aceder à página de detalhe de um conceito, o utilizador pode selecionar a opção de edição, sendo então apresentado um formulário pré-preenchido com os dados atuais do conceito.

O formulário permite a modificação de diversos campos, nomeadamente: o nome do conceito, a categoria, a descrição, a sigla, a citação, as traduções, os sinónimos e o link associado ao Google Scholar. Para além de possibilitar a edição dos campos já existentes, o formulário permite igualmente adicionar novas informações a campos que anteriormente não possuíam dados, como por exemplo introduzir uma sigla para conceitos que não a tinham definida inicialmente.

No caso das traduções, é necessário que o utilizador siga um formato específico de introdução, onde cada tradução deve ser inserida na forma Termo[idioma], sendo possível introduzir múltiplas traduções separadas por ponto e vírgula. Por exemplo: Dog[en]; Perro[es]; Hund[de]. Este formato facilita posteriormente o processamento automático das traduções, permitindo a sua correta estruturação e apresentação.

No caso dos sinónimos, a introdução é igualmente efetuada de forma simples e estruturada, recorrendo a uma lista de termos separados por ponto e vírgula. Cada sinónimo é inserido apenas pelo seu nome, não sendo necessário indicar idioma ou qualquer outro identificador adicional. Por exemplo: *termo1; termo2; termo3*. Esta abordagem permite a posterior conversão automática para o formato interno de armazenamento e facilita a sua correta apresentação e atualização no glossário.

Após a submissão do formulário, as novas informações são validadas e aplicadas sobre o conceito previamente existente. Antes de guardar as alterações, é criada automaticamente uma cópia de segurança do ficheiro de dados original, de forma a garantir a integridade da informação em caso de erro. Os dados atualizados são então persistidos, e o utilizador é redirecionado para a página de detalhe já com a nova informação aplicada.

Esta funcionalidade permite garantir a manutenção contínua e atualizada do glossário, assegurando a flexibilidade na gestão dos conceitos existentes.



Home | Conceitos | Categorias | Palavras Próximas | Pesquisa Inversa | Tabela

Editar Conceito: Adolescente

Conceito:
Adolescente

Categoria:
Ciências Sociais em Saúde

Descrição:
Considera-se adolescente a pessoa entre 12 e 19 anos de idade.

Sigla:

Citação:

Traduções:
Ex: Terceiro(terceira); Terceiro(terceira)

Sinônimos:
Ex: terceiro; terceira; terceira

Link Google Scholar:
<https://scholar.google.com/scholar?q=Adolescente>

Cancelar **Guardar Alterações**

Figura 5: Página de edição de um conceito.

3.2.6 Adicionar Novo Conceito

A aplicação disponibiliza uma página dedicada à adição de novos conceitos ao glossário, permitindo ao utilizador introduzir diretamente novos termos e respetivas informações associadas. Esta funcionalidade garante a expansão contínua e dinâmica da base de dados de conceitos.

Na página de adição, é apresentado um formulário onde o utilizador pode introduzir os diferentes campos do novo conceito, nomeadamente: o nome do conceito, a categoria, a descrição, a sigla, a citação, as traduções, os sinónimos e o link para o Google Scholar. Todos os campos são opcionais, com exceção do nome do conceito, sendo possível introduzir apenas a informação disponível no momento.

Após a submissão do formulário, os dados introduzidos são processados e integrados na estrutura de dados já existente. Sempre que necessário, é criada a categoria correspondente caso esta ainda não exista. Antes de guardar os novos dados, é igualmente criada uma cópia de segurança do ficheiro de dados, assegurando a integridade da informação. Após a adição, o utilizador é redirecionado para a listagem completa de conceitos, onde o novo termo fica imediatamente disponível para consulta e edição.

The screenshot shows a web-based application interface for adding a new concept. At the top, there's a navigation bar with links: Home, Conceitos, Categorias, Palavras Próximas, Pesquisa Inversa, and Tabela. Below the navigation is a dark header bar with the text "Adicionar Novo Conceito". The main area contains several input fields: "Concepto", "Categoria", "Descrição", "Síntese", "Citações", "Tradução" (with an example entry: "Ex: Termo1[termo1]; Termo2[termo2]"), "Sinônimos" (with an example entry: "Ex: termo1; termo2; termo3"), and "Link Google Scholar". At the bottom left are two buttons: "Cancelar" (Cancel) and "Adicionar Conceito" (Add Concept).

Figura 6: Página de adição de um novo conceito.

3.2.7 Categorias

A aplicação disponibiliza uma página de consulta de categorias, permitindo ao utilizador explorar os conceitos agrupados por categoria temática. Ao aceder à página de categorias, é apresentada uma listagem de todas as categorias existentes, sendo cada uma representada por um botão individual de navegação.

O utilizador pode selecionar qualquer uma das categorias disponíveis, sendo automaticamente redirecionado para a página correspondente onde são listados os conceitos pertencentes à categoria selecionada. Esta página permite uma visualização organizada e filtrada dos conceitos, facilitando a consulta de termos relacionados por tema.

A apresentação das categorias baseia-se na estrutura de armazenamento dos dados, onde cada categoria corresponde a uma chave do ficheiro *JSON*. Para a obtenção da lista de categorias, foi desenvolvida a função `obter_categorias()`, que percorre o conjunto de dados, extrai as categorias existentes, normaliza a capitalização de cada categoria para garantir uma apresentação consistente e ordena alfabeticamente a lista resultante.

Adicionalmente, foi implementada uma lógica de validação para garantir que, ao aceder diretamente a uma categoria, esta existe efetivamente no conjunto de dados.



[Home](#) | [Conceitos](#) | [Categorias](#) | [Palavras Próximas](#) | [Pesquisa Inversa](#) | [Tabela](#)

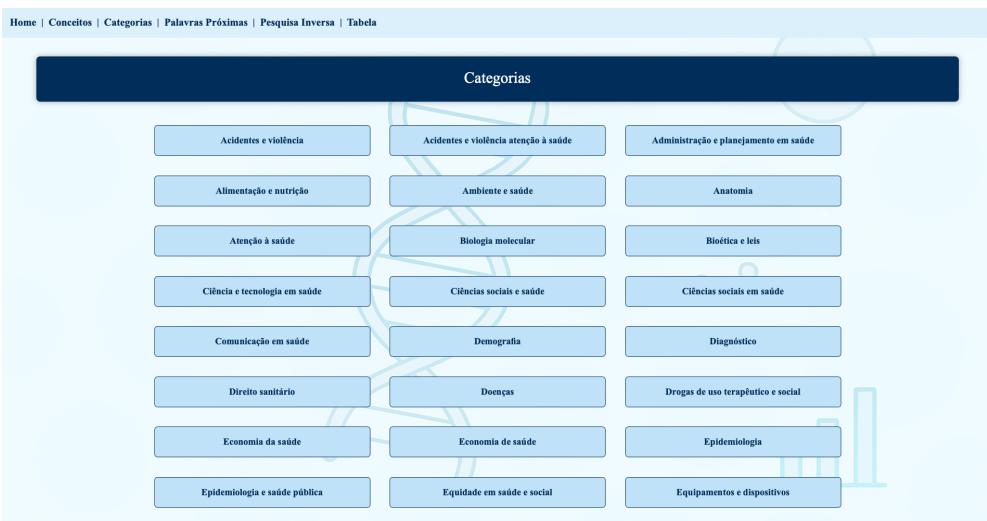


Figura 7: Página de listagem das categorias.

3.2.8 Consulta de Conceitos por Categoria

Nesta página, o nome da categoria é destacado no topo, sendo apresentada a lista de conceitos existentes nessa categoria, organizada alfabeticamente. Cada conceito surge como uma ligação clicável, permitindo o acesso à respetiva página de detalhe do conceito.

Para facilitar a pesquisa de um termo dentro da própria categoria, é disponibilizada uma caixa de pesquisa interativa, que permite ao utilizador filtrar de forma dinâmica a lista de conceitos apresentada com base no texto introduzido.

Foi ainda implementado um botão de navegação que permite ao utilizador regressar à página de categorias, bem como um botão flutuante para retorno rápido ao topo da página após a realização de *scroll*.



Figura 8: Página de listagem dos conceitos pertencentes à categoria selecionada.

3.2.9 Tabela de Conceitos

A aplicação disponibiliza uma página adicional de consulta dos conceitos sob a forma de tabela, permitindo uma visualização compacta e global de todos os conceitos existentes no glossário, independentemente da categoria a que pertencem.

Nesta página, todos os conceitos são apresentados em formato tabular, distribuídos por três colunas principais: *Conceito*, *Categoria* e *Descrição*. Esta organização permite ao utilizador obter uma visão geral rápida e organizada dos conceitos disponíveis e respetivas informações associadas.

Cada conceito apresentado na tabela surge como uma ligação clicável, permitindo o acesso direto à página de detalhe do respetivo conceito, onde poderão ser consultadas as informações completas associadas.

Esta funcionalidade complementa as restantes formas de navegação existentes na aplicação, permitindo ao utilizador consultar, de forma centralizada e estruturada, a totalidade dos conceitos registados no glossário.

Tabela de Conceitos:		
10 - 20 items per page		<input type="text"/> Search
Conceito	Categoria	Descrição
abeta	Genética e Biologia Molecular	Proteína que pode ser encontrada em todos os tipos de células do organismo humano. Ao acumular-se excessivamente no córtex cerebral do ser humano pode contribuir para o aceleramento do mal de alzheimer.
Abordagem médica tradicional do adulto hospitalizado	Atenção à Saúde	Focada em uma queixa principal e o hábito médico de tentar explicar todas as queixas e os sinais por um único diagnóstico, que é adequada no adulto jovem mas se aplica em relação ao idoso.
Abuso incestuoso	Acidentes e Violência	Consiste no abuso sexual envolvendo pais ou outro parente próximo, os quais se encontram em uma posição de maior poder em relação à vítima.
Abuso sexual na infância, Abuso sexual na adolescência	Acidentes e Violência	É todo ato ou jogo sexual, relação heterossexual ou homossexual, cujo agressor está em estágio de desenvolvimento póssexual mais adaptado que a criança ou adolescente. Tem por intenção estimular sexualmente ou utilizá-la para obter satisfação sexual. Essas práticas eróticas e sexuais são impostas à criança ou adolescente pela violência física, por ameaças ou pela indução de sua vontade.
Abuso financeiro dos idosos	Acidentes e Violência	Exploração imprópria ou ilegal e/ou uso não consentido de recursos financeiros dos idosos.
acalabrutinib	Fármacos e Terapias	
acidente vascular cerebral isquêmico	Doenças	Lesão que causa a morte de parte do cérebro humano, em virtude da falta de oxigénio e nutrientes que são inseridos no cérebro por meio da circulação sanguínea. Inf. encicl.: Uma vez privados do sangue, os neurônios morrem e liberam glutamato, uma substância química que realiza a comunicação entre as células nervosas. Em concentrações elevadas, porém, o glutamato torna-se tóxico e mata as células vizinhas, ampliando o estrago. O problema pode levar à imobilidade de braços e pernas bem como causar a perda da fala.

Figura 9: Página de visualização em tabela dos conceitos do glossário.

3.2.10 Relações entre Conceitos

A aplicação disponibiliza uma página dedicada à visualização das relações existentes entre os diferentes conceitos, com base nas palavras próximas associadas a cada termo. Esta funcionalidade permite identificar de forma intuitiva os conceitos que partilham ligações semânticas indiretas.

Para cada conceito, são apresentadas as palavras próximas previamente definidas no ficheiro de dados. Em seguida, para cada uma dessas palavras próximas, é identificado e apresentado o conjunto de outros conceitos que também incluem essa mesma palavra na sua lista de palavras próximas. Desta forma, é possível visualizar as ligações cruzadas e relações partilhadas existentes entre diferentes conceitos do glossário.

A página inclui ainda uma caixa de pesquisa no canto superior direito, permitindo ao utilizador filtrar dinamicamente os conceitos a visualizar com base no termo introduzido, facilitando assim a exploração de conceitos específicos.

Inicialmente, é construído um dicionário que associa cada conceito às respetivas palavras próximas. Posteriormente, são percorridos todos os conceitos e, para cada palavra próxima, são identificados os conceitos adicionais onde essa palavra também ocorre. O resultado desta análise é organizado e apresentado de forma agrupada e interativa na interface da aplicação.

Figura 10: Página de visualização das relações entre conceitos.

3.2.11 Pesquisa Inversa

A aplicação disponibiliza uma funcionalidade adicional de pesquisa inversa com base nas palavras próximas associadas aos conceitos. Esta funcionalidade permite explorar as palavras próximas existentes e identificar, para cada uma delas, os conceitos que as incluem.

Numa primeira página, é apresentada uma listagem alfabética de todas as palavras próximas existentes no glossário, agrupadas pela respetiva letra inicial. O utilizador pode navegar por este índice e selecionar uma determinada palavra próxima para consultar em detalhe.

Após selecionar uma palavra, é apresentada uma nova página onde são listados todos os conceitos que possuem essa palavra na sua lista de palavras próximas. Cada conceito surge como uma ligação clicável, permitindo o acesso direto à sua página de detalhe.

Esta funcionalidade foi implementada através da análise prévia de todos os conceitos carregados a partir do ficheiro *JSON*. Inicialmente, é gerada uma coleção única de todas as palavras próximas existentes. Posteriormente, para cada palavra, é identificado o conjunto de conceitos que a contém, permitindo construir esta relação inversa entre palavras próximas e conceitos.

Adicionalmente, a página de listagem alfabética, tal como na página dos Conceitos, inclui uma caixa de pesquisa dinâmica, permitindo ao utilizador filtrar as palavras próximas com base no texto introduzido, facilitando a navegação e a localização rápida da palavra pretendida.

Esta funcionalidade oferece uma perspetiva adicional na exploração do glossário, permitindo identificar rapidamente os conceitos que partilham elementos lexicais comuns e facilitando a análise semântica do conjunto de dados.



The screenshot shows a table titled 'Índice de Palavras Próximas' with the letter 'A' selected. The table lists various medical terms starting with 'A', such as Abaixo, Abaulamentos, Abdominais, Abdominal, Aberta, Abertura, Abordagem, Abrengue, Absoluta, Absuso, Absusos, Acalabentib, Acerola, Acesso, Acetilulicílico, Acidental, Acrometem, Acromete, Acompanha, and Acromosídeo.

A
Abaixo
Abaulamentos
Abdominais
Abdominal
Aberta
Abertura
Abordagem
Abrengue
Absoluta
Absuso
Absusos
Acalabentib
Acerola
Acesso
Acetilulicílico
Acidental
Acrometem
Acromete
Acompanha
Acromosídeo

Figura 11: Página de listagem alfabética das palavras próximas existentes no glossário.

The screenshot shows a page for the word 'absoluta'. It displays a list of concepts that contain this word: Bócio, sarcoma de ewing, Sarcoma de Kaposi, and Embriaguez. There is also a button labeled 'Voltar à lista de palavras'.

Conceitos que contêm esta palavra:
• Bócio
• sarcoma de ewing
• Sarcoma de Kaposi
• Embriaguez

Figura 12: Página de visualização dos conceitos que incluem a palavra próxima selecionada.

3.2.12 Pesquisa Avançada

Foi desenvolvida uma funcionalidade adicional de pesquisa flexível, permitindo ao utilizador procurar conceitos de forma mais personalizada, com base em diferentes critérios de correspondência.

A página de pesquisa disponibiliza um formulário centralizado e interativo, onde o utilizador pode introduzir o termo desejado e selecionar o modo de pesquisa pretendido:

- **Exata:** apenas são apresentados resultados que correspondam exatamente ao termo introduzido.
- **Contém:** são devolvidos todos os conceitos que contenham, em qualquer posição, o termo introduzido.
- **Começa por:** são apresentados os conceitos cujo nome se inicia pelo termo especificado.

Após a submissão do formulário, os resultados encontrados são apresentados numa lista organizada, com hiperligações que permitem o acesso direto à página de detalhe de cada conceito correspondente. Caso não sejam encontrados resultados para os critérios definidos, é apresentada uma mensagem informativa ao utilizador.



Figura 13: Página de Pesquisa Avançada.

3.2.13 Lista de Sinónimos

A aplicação disponibiliza uma página dedicada à visualização dos sinónimos associados aos diferentes conceitos presentes no glossário. Esta funcionalidade permite ao utilizador consultar, de forma centralizada, todos os conceitos que possuem sinónimos registados.

Na página é apresentada uma listagem organizada onde, para cada conceito com sinónimos, são exibidos o respetivo nome do conceito, a categoria a que pertence e a lista de sinónimos associados. Cada conceito surge como uma ligação interativa, permitindo o acesso direto à página de detalhe do mesmo.

Para facilitar a navegação e pesquisa, a página inclui ainda uma caixa de pesquisa dinâmica. À medida que o utilizador introduz texto na caixa de pesquisa, a listagem de conceitos é automaticamente filtrada com base no conteúdo pesquisado, permitindo localizar rapidamente conceitos ou sinónimos específicos.

A implementação desta funcionalidade recorre ao carregamento inicial de todos os conceitos a partir do ficheiro *JSON*, sendo posteriormente identificados apenas aqueles que possuem sinónimos associados. A informação é organizada e apresentada de forma ordenada, proporcionando uma consulta clara e eficiente dos sinónimos registados no glossário.



The screenshot shows a search interface with a search bar labeled 'Pesquisar' and a placeholder 'Procurar conceito ou sinónimo...'. Below the search bar, there is a list of five concepts:

- Adolescente**
Categoria: Ciências Sociais em Saúde
Sinónimos: moça, moço, rapaz, rapariga
- Aerosol**
Categoria: Equipamentos e Dispositivos
Sinónimos: aerosol can, aerosol container, aerosol bomb, spray can
- Aids**
Categoria: Doenças
Sinónimos: SIDA, Síndrome da Imunodeficiência Adquirida, Síndrome da imuno-deficiência adquirida, tending, help, assistance, assist, economic aid, financial aid, attention, acquired immune deficiency syndrome, aid, care
- Alcoolismo**
Categoria: Drogas de Uso Terapêutico e Social
Sinónimos: Dipomania
- Alcoólatra**
Categoria: Drogas de Uso Terapêutico e Social
Sinónimos: bêbado, bêbedo

Figura 14: Página de listagem dos conceitos com sinónimos registados.

3.2.14 Jogo de Sinónimos

Com o objetivo de introduzir uma componente interativa e dinâmica na exploração do glossário, foi desenvolvida uma página de *Jogo de Sinónimos*. Esta funcionalidade permite ao utilizador testar os seus conhecimentos relativamente aos sinónimos associados aos conceitos presentes, promovendo uma aprendizagem ativa dos conteúdos.

Ao aceder à página, o sistema seleciona automaticamente um conjunto aleatório de conceitos que possuem sinónimos associados. Para cada conceito, é apresentada uma pergunta com várias opções de resposta, onde o utilizador deve selecionar o sinónimo correto correspondente. As opções são geradas de forma dinâmica, incluindo sempre o sinónimo correto e várias opções de distração, também extraídas de outros sinónimos presentes no glossário.

Após a seleção das respostas para todas as perguntas, o utilizador submete o formulário, sendo os seus resultados processados no servidor. O sistema procede então à validação automática das respostas, comparando as escolhas efetuadas com os sinónimos corretos armazenados na base de dados. No final, é apresentada uma página de resultados com a pontuação obtida, indicando o número total de respostas corretas e permitindo ao utilizador reiniciar o jogo para novas tentativas.

Esta funcionalidade recorre a um sistema de validação implementado na própria rota que trata tanto a geração dinâmica das perguntas (método GET) como a validação e cálculo de resultados (método POST).

O *Jogo de Sinónimos* acrescenta uma vertente lúdica à aplicação, permitindo aos utilizadores consolidar o conhecimento dos conceitos de forma prática e divertida.



Figura 15: Página do jogo dos sinónimos.

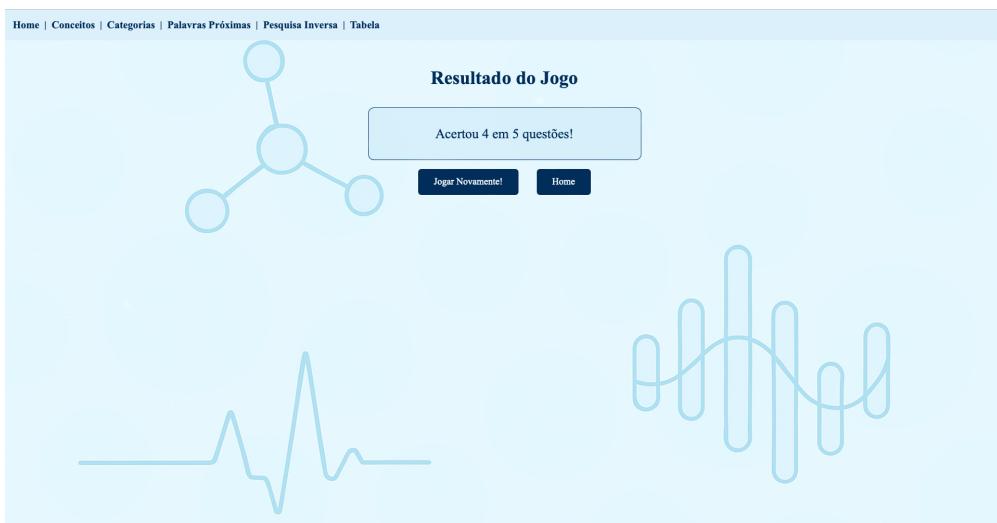


Figura 16: Página de resultados do jogo dos sinónimos.

3.2.15 Barra de Navegação

A aplicação disponibiliza, de forma constante em todas as páginas, uma barra de navegação superior (*navbar*), que permite ao utilizador aceder rapidamente às principais funcionalidades do sistema, promovendo uma navegação intuitiva e eficiente.

A implementação da *navbar* recorreu à utilização de templates dinâmicos em *Flask* com o auxílio do framework *Bootstrap*, proporcionando uma integração visual consistente e uma experiência de navegação fluida ao longo de toda a aplicação.

4 Sugestões e Melhorias

Ao nível das funcionalidades, seria vantajoso o desenvolvimento de funcionalidades de exportação de dados, permitindo ao utilizador descarregar subconjuntos do glossário em



formatos estruturados (*CSV*, *JSON*, *Excel*, etc.), de modo a facilitar a sua utilização em contextos académicos ou de investigação.

Por fim, a integração com recursos biomédicos externos, como a *Unified Medical Language System (UMLS)* ou o *SNOMED CT*, poderia enriquecer ainda mais o glossário, garantindo uma maior cobertura terminológica e interoperabilidade com sistemas clínicos e científicos.

Estas melhorias contribuiriam para tornar a aplicação numa ferramenta ainda mais completa, flexível e aplicável a contextos reais de apoio ao ensino, investigação e prática clínica.

5 Conclusão

O trabalho desenvolvido permitiu a aplicação prática de diversas técnicas de Processamento de Linguagem Natural no contexto biomédico, resultando na construção de um glossário enriquecido e acessível através de uma aplicação *web* interativa.

Ao longo do projeto, foram abordadas várias etapas fundamentais, desde o enriquecimento automático dos conceitos — com a adição de sinónimos, traduções, palavras semanticamente próximas e referências bibliográficas — até ao desenvolvimento de funcionalidades avançadas na aplicação, como a gestão e visualização dos conceitos, a categorização automática, a pesquisa inversa, o jogo de sinónimos e a funcionalidade de pesquisa personalizada.

A implementação da plataforma permitiu consolidar conhecimentos de processamento de dados textuais, integração de modelos de linguagem, organização e estruturação de dados, bem como competências de desenvolvimento *web* com *Flask*.

Apesar da robustez e diversidade de funcionalidades já implementadas, foram também identificadas diversas oportunidades de melhoria e expansão futura, nomeadamente ao nível da integração com terminologias biomédicas externas, da aplicação de modelos semânticos mais avançados e do aumento da flexibilidade do sistema de pesquisa.

Em suma, o projeto revelou-se um exercício enriquecedor, multidisciplinar e alinhado com os desafios atuais da informática médica, demonstrando o potencial da utilização de técnicas de PLN no apoio à gestão e disponibilização de conhecimento biomédico estruturado.