



**Universidade do Minho**  
Escola de Engenharia

# Processamento de Linguagem Natural em Engenharia Biomédica

**Grupo 1:**

Alícia Soares Oliveira, PG50169

Ana Carolina Veloso, PG50178

Margarida Ferreira Fernandes, PG50588

**Docente:**

Luís Filipe Costa Cunha,

José João Antunes Guimarães Dias Almeida

1 de Junho de 2023

# Índice

<i>1. Introdução .....</i>	<i>3</i>
<i>2. Desenvolvimento .....</i>	<i>3</i>
2.1. Enriquecimento dos ficheiros JSON .....	3
2.2. Melhoria da Aplicação .....	4
<i>3. Conclusão .....</i>	<i>5</i>
<i>4. Bibliografia .....</i>	<i>5</i>

# 1. Introdução

O presente projeto surge como continuação do trabalho desenvolvido na primeira parte da UC de Processamento de Linguagem Natural em Engenharia Biomédica, tendo como objetivo principal aplicar e aprofundar o conhecimento adquirido em sala de aula. Assim sendo, pretende-se aumentar a informação presente no documento gerado no trabalho anterior, focando a informação numa melhoria de um determinado caso de estudo.

Desta forma, durante a realização do projeto foram consultadas diferentes fontes de onde foi retirada informação relevante que, através de uma aplicação *Flask*, permite receber relatórios médicos e ajudar os profissionais de saúde a esclarecer diferentes dúvidas relativas a termos médicos, sistemas do corpo humano e doenças. Assim sendo, no total, existem 3 ficheiros JSON.

## 2. Desenvolvimento

### 2.1. Enriquecimento dos ficheiros JSON

Relativamente aos ficheiros JSON gerados durante o processo de recolha de informação, o primeiro deriva da junção do documento resultante do primeiro trabalho com duas informações distintas: informações provenientes do *website Wikilivros* [1] em relação a termos médicos e traduções de todos os termos em inglês, espanhol, alemão e italiano. O segundo guarda informações referentes a diferentes doenças, tais como uma descrição da doença, a causa e o tratamento. Por fim, o último ficheiro guarda informações respeitantes aos componentes dos diferentes sistemas do corpo humano como, por exemplo, o sistema respiratório e o sistema circulatório.

Com o objetivo de recolher a informação relevante que permite enriquecer o ficheiro JSON e a plataforma desenvolvida, recorreu-se à biblioteca *BeautifulSoup4* do *Python*. Em primeiro lugar, foram extraídos termos médicos provenientes do *website Wikilivros* [1], com vista a aumentar a base de dados construída no primeiro trabalho, relativa a termos. De seguida e, tendo em conta que alguns termos tinham traduções em inglês e espanhol e outros não tinham qualquer tradução, recorreu-se a uma *script* que faz uso do *Google Translate* de modo a preencher as informações relativas a cada um dos termos. Assim sendo, no final gerou-se um ficheiro JSON no qual todos os termos apresentam uma descrição, resultante da concatenação de duas entradas do primeiro trabalho, bem como traduções de inglês, espanhol, alemão e italiano.

Posteriormente, a equipa identificou dois temas ausentes na plataforma do dicionário de termos médicos, nomeadamente, os termos e significados de órgãos e componentes dos diferentes sistemas do corpo humano e informações sobre as principais doenças que atingem a população. Assim, de modo a extrair informações

sobre os sistemas do corpo humano, recorreu-se ao *website* [2] e, no caso das doenças, utilizou-se o *website* [3].

O processo de devolver apenas as informações relevantes de cada um dos três *links* é idêntico. O mesmo é descrito por inicialmente identificar onde se encontra a informação pertinente, nomeadamente, o elemento *html* (*div*, *a*, *p*, *h2*, *table*, ...) e o nome da classe desse elemento, como por exemplo *class\_="entry-content"*. A partir desta etapa foi necessário percorrer cada um dos *html* do *link* em questão até se alcançar a informação anteriormente identificada. Assim, e com vários ciclos de procura, a informação é guardada em variáveis de interesse, para deste modo ser mais fácil o processo de armazenamento numa lista ou dicionário. Por fim, escreveu-se a lista ou dicionário de *output* num ficheiro JSON.

Ainda de salientar que nas *scripts* utilizaram-se funções *Python* como *get()*, *find\_all()* e *find()* para auxiliar no processo de captura da informação de interesse.

## 2.2. Melhoria da Aplicação

No que concerne à aplicação *Flask*, cujo desenvolvimento se iniciou no primeiro trabalho, esta foi aprimorada em diversos níveis, tendo em consideração os novos JSON produzidos. A aplicação consiste num dicionário médico multilínguas, que tem como principal objetivo dar suporte a todos os profissionais de saúde, mais concretamente na análise de relatórios médicos.

Deste modo, a aplicação está subdividida em três páginas principais: português, inglês e espanhol, e cada página contém o mesmo conteúdo, diferindo unicamente na língua padrão utilizada. Além do *Menu*, que especifica a utilidade da aplicação, estão presentes outras 7 páginas: *Termos Médicos*, *Procura de termos*, *Tabela de termos*, *Sistemas do Corpo Humano*, *Doenças Comuns*, *Adicionar Relatórios* e *Analisar Relatórios*.

As páginas *Termos Médicos* e *Procura de termos*, que já estavam presentes no primeiro trabalho realizado nesta Unidade Curricular, foram desenvolvidas. Para tal, à página *Termos Médicos* adicionou-se a possibilidade de, além de consultar os termos, inserir e apagar termos do dicionário. A *Procura de termos* permite pesquisar os termos médicos presentes no dicionário, devolvendo os termos, a respetiva descrição, caso exista, e a tradução em 4 línguas diferentes: inglês, espanhol, alemão e italiano.

Relativamente à página *Tabela de termos*, cuja base surgiu do conteúdo lecionado nas aulas da Unidade Curricular, esta permite consultar de forma mais prática e intuitiva os termos do dicionário, sendo proficiente na comparação de conceitos e até definições. Nos mesmos moldes, a página *Doenças Comuns* visa obter informação sobre as doenças presentes no JSON já anteriormente especificado.

No que tange à página *Sistemas do Corpo Humano*, esta é firmada na organização e subdivisão dos termos *JSON* por sistemas, viabilizando um estudo e uma pesquisa mais condensados, paralelamente a uma conceção mais harmoniosa de todos os conceitos.

Focalizando no suporte que a aplicação promete fornecer, as páginas *Adicionar Relatórios* e *Analisar Relatórios* fornecem a base deste suporte à decisão médica, uma vez que permitem adicionar relatórios médicos, em formato PDF, e analisá-los em formato XML, respetivamente. A análise consiste na leitura de relatórios e, por cada palavra presente em cada um deles, verifica-se se o termo está presente no dicionário médico. Caso seja afirmativo, então o termo fica sublinhado e, clicando no mesmo, redireciona-se para a página com a definição do termo correspondente, bem como as traduções nas diversas línguas.

### 3. Conclusão

Em suma, este projeto resultou numa aplicação *Flask* altamente funcional e de grande utilidade para os profissionais de saúde. Através desta aplicação, é possível aprimorar a compreensão de termos médicos, de sistemas do corpo humano e de doenças, proporcionando uma melhor interpretação de relatórios médicos. O projeto permitiu a aplicação prática e o aprofundamento dos conhecimentos adquiridos ao longo da Unidade Curricular, podendo ser uma ferramenta eficiente no processo de tomada de decisão médica.

Além disso, a aplicação foi desenvolvida de forma a atender às necessidades específicas dos profissionais de saúde, fornecendo recursos e funcionalidades úteis. A possibilidade de adicionar, excluir e pesquisar termos médicos, juntamente com a consulta prática de tabelas e informações sobre as doenças, amplifica o suporte oferecido pela aplicação. Adicionalmente, a análise de relatórios, destacando os termos presentes no dicionário médico e fornecendo as definições e as traduções correspondentes, agrega um valor significativo no processo de interpretação.

Concluindo, este projeto representa um avanço significativo no campo da Engenharia Biomédica, proporcionando uma ferramenta poderosa para auxiliar os profissionais de saúde nas suas atividades diárias.

### 4. Bibliografia

- [1] [https://pt.wikibooks.org/wiki/Dicionário\\_de\\_termos\\_médicos](https://pt.wikibooks.org/wiki/Dicionário_de_termos_médicos)
- [2] <https://beduka.com/blog/materias/biologia/sistemas-do-corpo-humano>
- [3] <https://www.mdsaude.com/doencas-a-z/>