



Universidade do Minho
Escola de Engenharia

Aplicação *web* para visualização e manipulação de *datasets* de conceitos

Trabalho Prático 2

Ano Letivo 2023/2024

Mestrado/Área de especialização em Informática Médica

Grupo: Diogo Rodrigues PG54450

Inês Faria A95494

Lia Madurago PG53633

Docentes:

José João Antunes Guimarães Dias Almeida

Luís Filipe Costa Cunha

Braga, 11 de junho de 2024

Resumo

O presente relatório insere-se no âmbito da Unidade Curricular de Processamento de Linguagem Natural em Engenharia Biomédica da Universidade do Minho. Destina-se à descrição e explicação detalhada das tarefas desenvolvidas na resolução da segunda parte do trabalho de grupo, que consiste em desenvolver um sistema que permita a utilização dos dados em formato JSON que foram extraídos dos PDFs médicos durante o trabalho prático anterior e criar uma ferramenta que permite manipular e representar os mesmos.

Índice

1	Introdução	4
2	Desenvolvimento	4
2.1	Enriquecimento do <i>dataset</i>	4
2.2	Plataforma <i>Web</i>	6
3	Conclusão	11

Índice de Figuras

Figura 1:	Ficheiro JSON final concatenado.	5
Figura 2:	Página inicial da plataforma <i>web</i>	6
Figura 3:	Página Conceitos.	7
Figura 4:	Consultar os conceitos.	7
Figura 5:	Adicionar Novo Conceito.	8
Figura 6:	Motor de busca.	8
Figura 7:	Página Contactos.	9
Figura 8:	Página Sobre Nós.	10

1 Introdução

O avanço das tecnologias de Processamento de Linguagem Natural (PLN) têm proporcionado uma revolução na forma como os dados textuais são analisados e utilizados, especialmente no domínio da Engenharia Biomédica. A capacidade de extrair, enriquecer e representar dados textuais de maneira eficaz é crucial para o desenvolvimento de sistemas inteligentes que possam auxiliar em diagnósticos médicos, pesquisa clínica, e gestão de informações de saúde.

Neste contexto, o presente trabalho prático tem como objetivo desenvolver uma aplicação *web* para manipulação e visualização de dados médicos extraídos de PDFs, em formato JSON, obtidos durante o trabalho prático anterior. O projeto divide-se em duas etapas: o enriquecimento do dataset e a criação de uma plataforma *web* para visualização e manipulação dos dados.

Este relatório documenta detalhadamente o desenvolvimento do sistema, desde a fase de enriquecimento de dados até a implementação da plataforma *web*, incluindo as metodologias utilizadas, os desafios enfrentados e as soluções adotadas. Assim, o trabalho apresentado contribuirá significativamente para a área de PLN em Engenharia Biomédica, oferecendo uma ferramenta poderosa para gerir e analisar dados médicos textuais.

2 Desenvolvimento

A primeira etapa envolveu a expansão do conjunto de dados existente através da adição de novas informações provenientes de fontes externas, como *websites* e dicionários médicos. Esse processo de enriquecimento visa não apenas introduzir novos conceitos, mas também adicionar atributos aos conceitos já existentes, como sinónimos, definições, termos relacionados e categorias, facilitando análises mais complexas.

A segunda etapa do projeto consistiu no desenvolvimento de uma ferramenta *web* interativa que permite a visualização, manipulação e atualização dos dados enriquecidos. Esta plataforma oferece uma navegação intuitiva, explorando as relações entre os conceitos para permitir uma organização eficiente dos dados por diversos atributos ou categorias.

2.1 Enriquecimento do *dataset*

No processo de enriquecimento do *dataset* original, como falado anteriormente, foram adicionados termos de outras fontes, nomeadamente de *websites* e glossários pertinentes para

o estudo. Para além disso, adicionou-se também um campo que especifica a categoria do termo em estudo aos datasets que não possuíam este campo. Este campo de categoria fornece ao utilizador uma indicação sobre a área em que se insere o termo em estudo, isto é, se estamos perante uma doença, uma administração de um fármaco, um acidente, alimentação, nutrição, entre muitos outros.

Assim, nesse sentido, foram usados os *datasets Glossário- Dicionário Médico, Cruz Verde- Glossário de Saúde* para o aumento do leque de termos e descrições dos mesmos que seriam englobados na totalidade do dataset explorado. Para além disso foi englobado o *Glossário de Siglas do Infarmed* para adicionar, como o próprio nome indica, siglas úteis ao nosso leque de termos. Para isso foi feita uma extração inicial para ficheiros *json* distintos da correlação entre termo e respetiva descrição para cada um dos *websites* explorados.

Inicialmente, após a extração dos termos, foi adicionado o campo "categoria" a cada um dos glossários mencionados anteriormente, visando uma melhor correlação entre eles. Em seguida, os arquivos no formato *JSON* foram concatenados em um único arquivo. Este arquivo consolidado contém a totalidade dos termos, categorias e descrições que serão alvo do estudo no restante do trabalho.

```
{
  "ADJAdjuvante farmacêuticoAids pediátrica": {
    "categoria": "Alimentação e Nutrição",
    "definicao": "É o mesmo que dieta equilibrada ou balanceada e pode ser resumida por três princípios: variedade, moderação e equilíbrio. Variedade significa comer",
  },
  "AIDS II": {
    "categoria": "Drogas de Uso Terapêutico e Social",
    "definicao": "Este termo refere-se tanto aos bebedores-problema quanto aos dependentes do álcool.",
  },
  "ALCALcoólatraAlimentação exclusivaAmebiase": {
    "categoria": "Doenças",
    "definicao": "Denomina-se genericamente anemia como a situação clínica onde ocorre a diminuição do número de eritrócitos circulantes e/ou da quantidade de hemoglobina",
  },
  "AMEAmebiaseAntibioticoterapia": {
    "categoria": "Doenças",
    "definicao": "Doença inflamatória poliarticular crônica, incidindo principalmente nas articulações dos dedos e dos punhos, com caráter extensivo, deformante e angustioso",
  },
  "ANTAnticoncepçãoAssistência à saúde": {
    "categoria": "Administração e Planejamento em Saúde",
    "definicao": "O sistema de atendimento pré-hospitalar tem a finalidade de atender vítimas em situação de urgências e emergências, antes de sua chegada ao hospital",
  },
  "ARC": {
    "categoria": "Atenção à Saúde",
    "definicao": "Medicamentos, Vacinas e Insumos Compreende um conjunto de atividades relacionadas ao acesso e ao uso racional de medicamentos, destinado a complementar",
  },
  "ASSAssistência ambulatorialAtendimento pré-hospitalarAtendimento pré-hospitalar a x0": {
    "categoria": "Administração e Planejamento em Saúde",
    "definicao": "Este é o principal formulário usado pelo SIH (Sistema de Informação Hospitalar) que permite o diagnóstico da produção de serviços. Este instrumento",
  },
  "ATEAtendimento pré-hospitalar a x0Autorização de exportação": {
    "categoria": "Medicamentos, Vacinas e Insumos",
    "definicao": "O banco de preços visa garantir uma política global de acesso a medicamentos para a aids ao maior número de pessoas, com preços justos e adequados à realidade",
  },
  "AUTAutorização de importaçãoAZTBaixo peso ao nascer": {
    "categoria": "Atenção à Saúde",
    "definicao": "Comunicação em Saúde Humanização da assistência à criança hospitalizada através da melhoria da qualidade da atenção nos períodos em que as crianças",
  },
  "AVC": {
    "categoria": "Doenças"
  }
}
```

Figura 1: Ficheiro JSON final concatenado.

2.2 Plataforma Web

A plataforma *web* desenvolvida possui 4 páginas - uma página inicial, com opções que direcionam para as restantes páginas, uma página Conceitos, uma página Contactos e uma página Sobre Nós.

Na página inicial, é possível verificar uma pequena descrição da aplicação *web*, a imagem da Escola de Engenharia, um *header*, um *footer* e 3 botões referentes às páginas Conceitos, Contactos e Sobre Nós, que redirecionam para as mesmas respetivamente. Estes botões são dinâmicos, recorrendo a funcionalidades de *HTML* e mais propriamente *CSS*, tornando os mesmos mais interativos para o utilizador. Para além disso, no *header*, as várias opções de páginas possuem efeitos, também criados através de *HTML* e *CSS*, para tornar a experiência mais interativa para o utilizador.



Figura 2: Página inicial da plataforma *web*.

Na página Conceitos, observa-se um texto inicial com uma pequena apresentação dos tipos de conceitos que temos apresentar, bem como as funcionalidades possíveis de efetuar.

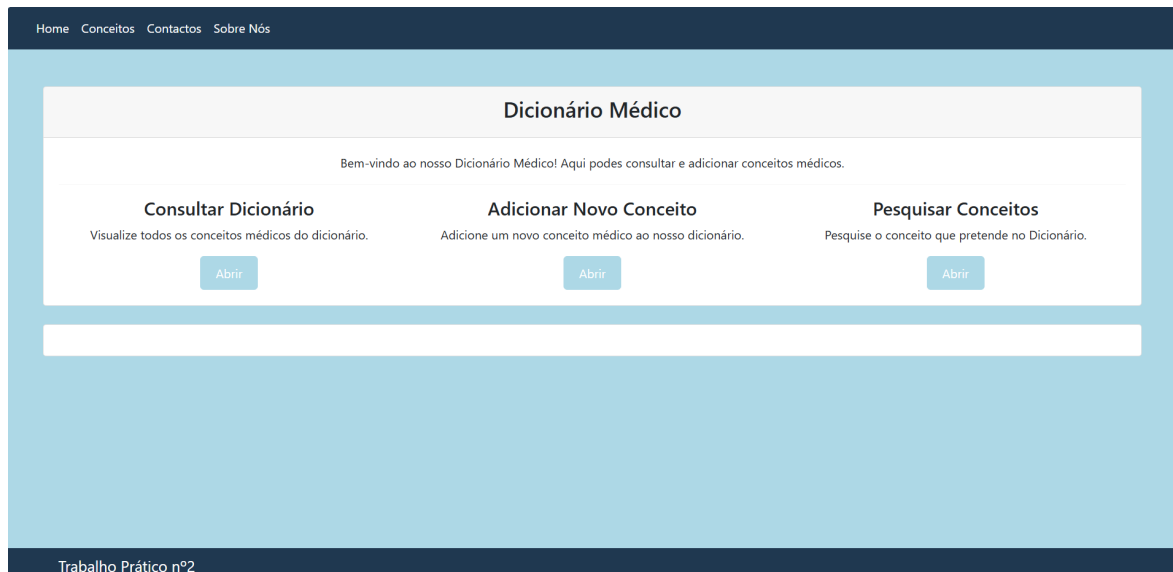


Figura 3: Página Conceitos.

A primeira funcionalidade trata-se da Consultar Dicionário, em que, ao clicar no botão Abrir, é possível listar todos os conceitos, assim como a sua categoria e a sua designação.



Figura 4: Consultar os conceitos.

Já a segunda funcionalidade, a Adicionar Conceito, permite ao utilizar inserir um novo conceito ao dicionário (ficheiro JSON), sendo que é obrigatório inserir o termo, a sua categoria e a designação do mesmo.

The screenshot shows the 'Dicionário Médico' web application. At the top, there's a header with the title 'Dicionário Médico'. Below it, a welcome message says 'Bem-vindo ao nosso Dicionário Médico! Aqui podes consultar e adicionar conceitos médicos.' There are three main buttons: 'Consultar Dicionário', 'Adicionar Novo Conceito', and 'Pesquisar Conceitos'. The 'Adicionar Novo Conceito' button is highlighted. Below these buttons, there's a form with three input fields: 'Título', 'Categoria', and 'Descrição'. At the bottom of the form, there are two buttons: 'Adicionar Conceito' and 'Cancelar'.

Dicionário Médico

Bem-vindo ao nosso Dicionário Médico! Aqui podes consultar e adicionar conceitos médicos.

Consultar Dicionário
Visualize todos os conceitos médicos do dicionário.
[Abrir](#)

Adicionar Novo Conceito
Adicione um novo conceito médico ao nosso dicionário.
[Abrir](#)

Pesquisar Conceitos
Pesquise o conceito que pretende no Dicionário.
[Abrir](#)

Título

Categoria

Descrição

[Adicionar Conceito](#) [Cancelar](#)

Trabalho Prático nº2

Figura 5: Adicionar Novo Conceito.

Por fim, a terceira funcionalidade consiste num motor de busca, onde o utilizador pode pesquisar qualquer palavra, sendo que, caso essa exista, irá remeter todos os termos que possuam na sua designação, na sua descrição ou na sua categoria essa palavra. Desta forma, retorna uma tabela com todos esses conceitos.

The screenshot shows the 'Dicionário Médico' web application with search results. The header and navigation buttons are the same as in Figure 5. Below the buttons, there's a table with three columns: 'Termo', 'Categoria', and 'Definição'. The table contains three rows of medical concepts.

Dicionário Médico

Bem-vindo ao nosso Dicionário Médico! Aqui podes consultar e adicionar conceitos médicos.

Consultar Dicionário
Visualize todos os conceitos médicos do dicionário.
[Abrir](#)

Adicionar Novo Conceito
Adicione um novo conceito médico ao nosso dicionário.
[Abrir](#)

Pesquisar Conceitos
Pesquise o conceito que pretende no Dicionário.
[Abrir](#)

Termo	Categoria	Definição
Centro de Atenção Psicossocial	Doenças	Também chamada de clamídiase, é uma doença sexualmente transmissível causada pela bactéria <i>clamydia trachomatis</i> , que afeta homens e mulheres. Se não for adequadamente tratada poderá provocar infertilidade permanente e, em casos de gravidez, haverá riscos de infecção nos olhos ou nos pulmões para o feto ou para o recém-nascido.
Intestino grosso	Doenças	Última parte do intestino, se estende do ceco ao reto. O cólon é um órgão longo em forma de tubo, com cerca de 1,5 metro de comprimento, cujo papel é remover a água dos alimentos digeridos e formar o resíduo sólido denominado fezes, que se move do cólon para o reto e deixa o corpo através do ânus.
Osso púbico	Doenças	Um dos ossos da pelve, localizado na região central e anterior da bacia.

Trabalho Prático nº2

Figura 6: Motor de busca.

Na página de Contactos, além do *header* e *footer* mencionados anteriormente, apresentamos diversas formas de contato para cada elemento do grupo. São fornecidas informações detalhadas, como o nome, número mecanográfico e endereço de correio eletrónico institucional.

Adicionalmente, foi implementada uma funcionalidade que permite enviar um *email* diretamente a qualquer membro do grupo. Ao clicar no endereço de *e-mail*, uma nova mensagem será automaticamente aberta na conta do Outlook, com o destinatário já definido.

Além disso, inclui-se um mapa interativo que contem a localização da Universidade do Minho, Campus de Gualtar, onde grande parte do projeto foi desenvolvido.

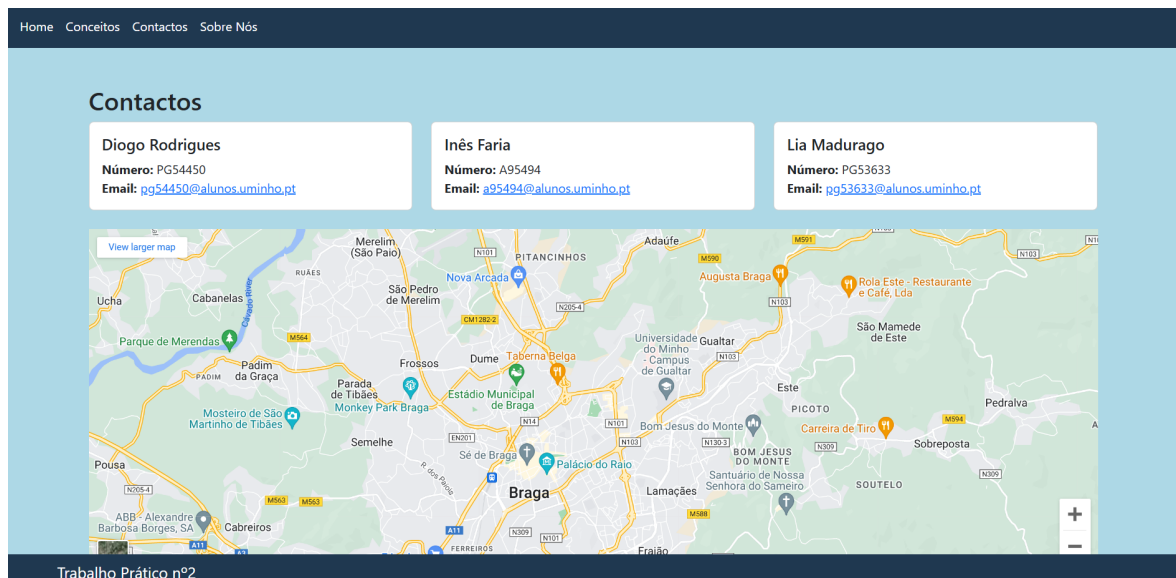


Figura 7: Página Contactos.

Por fim, na página Sobre Nós, foi apresentada uma breve descrição das ideologias do grupo. O nosso compromisso é a inovação constante e a procura pela excelência em todas as nossas atividades. Esta página demonstra os valores pelos quais o grupo se guia e que podem auxiliar o utilizador do *website* a conhecer um pouco melhor sobre a equipa responsável pelo mesmo.

Home Conceitos Contactos Sobre Nós

Sobre Nós

Somos uma equipa composta por profissionais dedicados e apaixonados pela tecnologia e inovação, nomeadamente na área da Engenharia Biomédica.

O nosso principal objetivo é fornecer soluções avançadas que melhorem a qualidade de vida das pessoas, através da utilização das últimas tecnologias em PLN e outras áreas da Engenharia Biomédica.

Com uma equipa composta por três Engenheiros Biomédicos, trabalhamos em conjunto para criar produtos e serviços que atendam às necessidades dos nossos clientes e parceiros.



Missão

Promover o bem-estar e a saúde através de inovações tecnológicas na área, fornecendo soluções eficientes e acessíveis.

Visão

Ser referência mundial em inovação tecnológica na Engenharia Biomédica, reconhecida pela excelência e impacto positivo na sociedade.

Valores

- Inovação
- Qualidade
- Integridade
- Compromisso com o cliente
- Trabalho em equipa

Trabalho Prático nº2

Figura 8: Página Sobre Nós.

3 Conclusão

Em suma, este projeto resultou numa aplicação *web* em Flask altamente funcional e de grande utilidade para os profissionais de saúde. Através desta aplicação, é possível aprimorar a compreensão de termos médicos, assim como as suas categorias, proporcionando uma melhor interpretação dos mesmos. 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.

Desta forma, a aplicação foi desenvolvida de forma a atender às necessidades específicas dos profissionais de saúde, fornecendo recursos e funcionalidades úteis. A possibilidade de adicionar, listar e pesquisar termos médicos, juntamente com a consulta prática de tabelas e informações sobre as doenças, amplifica o suporte oferecido pela aplicação.

A execução deste trabalho permitiu consolidar conhecimentos em *HTML*, *CSS*, *Flask* e *Jinja2*, tecnologias que serão certamente de extrema utilidade no futuro e que não haviam sido muito exploradas até o momento. Além disso, competências relacionadas à extração de informações pertinentes, como a utilização da biblioteca *BeautifulSoup4*, foram de extrema importância. Esse processo permitiu aos membros do grupo desenvolver novas aptidões em toda a etapa de coleta de informações, especialmente em páginas *web*, uma área que ainda não havia sido explorada por nenhum dos integrantes do grupo de trabalho.

Para melhorias futuras, é pretendido a melhoria da estética da aplicação, de modo a que esta fique mais interativa e dinâmica para o utilizador, assim como a melhoria do motor de busca, para que o mesmo consiga fazer pesquisas mais eficazes e direcionadas para objetivos específicos do utilizador. Além disso, visa-se a implementação de mais funcionalidades, como uma opção de tradução dos conceitos em si, bem como outras páginas que permitam tornar a aplicação mais robusta e completa.

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.

Referências

- [1] GLOSSÁRIO – DICIONÁRIO MÉDICO. <https://www.mdsaude.com/glossario/>
- [2] Cruz Verde- Glossário de Saúde. <https://www.cruzverde.pt/apoio-cliente/glossario-saude>
- [3] Glossário Siglas Infarmed. <https://www.infarmed.pt/web/infarmed>
- [4] Glossário Ministério da Saúde. Fornecido pelos professores para o trabalho prático anterior.
- [5] Bootstrap. <https://getbootstrap.com/>
- [6] Learn CSS Animations In 20 Minutes - For Beginners. <https://www.youtube.com/watch?v=SgmNx91WcY>
- [7] How To Make Website Using HTML And CSS | Website Design. <https://www.youtube.com/watch?v=PgAZ8Kzfh08>