

Escola de Engenharia Engenharia Biomédica 1º Semestre 2º Ano

Sistema de Consulta e Análise de Publicações Científicas

Algoritmos e Técnicas de Programação

Docentes orientadores:

Luís Filipe Cunha José Carlos Ramalho

Trabalho elaborado por:

Beatriz Carvalho Figueiras (A102463) Carlos Manuel Filipe Carvalho Araújo(A107187) Paulo Cristiano Fernandes Taveira(A103086)

Índice

Introdução	3
Estrutura do programa	4
Interface de linha de comando (CLI)	4
Importação inicial	4
Criar Publicação	5
Consultar Publicação	5
Consultar Publicações	6
Eliminar Publicação	6
Relatórios	6
Listar autores	7
Atualizar Publicação	7
Sair	7
Interface gráfica	7
wprincipal()	7
wficheirogui()	8
wconsultar()	9
wPublicacao()	10
wrelatorios()	11
wlistar(data)	14
Conclusão	15
Bibliografia	16

Introdução

No âmbito da Unidade Curricular de Algoritmos e Técnicas de Programação, da Licenciatura de Engenharia Biomédica da Escola de Engenharia da Universidade do Minho, 2º ano, 1º semestre, ano letivo 2024/2025, foi proposto aos alunos a elaboração de um trabalho, em sistema Python, que permitisse criar, atualizar e analisar publicações científicas.

Este trabalho tem como objetivo aprofundar o que foi lecionado.

A metodologia deste trabalho consistiu em pesquisas na internet, bem como, materiais fornecidos pelos docentes orientadores e ferramentas do Python.

Neste trabalho, começamos por apresentar uma breve introdução à estrutura geral do programa. De seguida, enumeramos algumas informações sobre o funcionamento tanto da *Interface de linha de comando (CLI)* como da *Interface gráfica*. Terminamos com uma reflexão crítica sobre o que foi mencionado anteriormente.

Estrutura do programa

Utilizou-se a linguagem Python para escrever/desenvolver o nosso programa, em conjunto com módulos como o FreeSimpleGUI (para o desenvolvimento da interface gráfica), datetime (para formatação e gerenciamento de datas), do ficheiro json (como base de dados), matplotlib.pyplot (para a realização dos gráficos), os (para a captação de endereço dos ficheiros para evitar erros no procedimento), PrettyTables (para a realização e apresentação de tabelas no terminal) e Colorama (para atribuir cor ao texto no terminal).

Na pasta "app" temos 4 ficheiros:

- fun.py onde se encontram todas as funções backend para o programa;
- graph.py onde se encontra a interface gráfica e algumas das funções relacionadas à mesma;
- *main.py* quando se corre este ficheiro no terminal, a interação com o programa começa, através da interface gráfica;
- cli.py para correr a linha de comando que permite interagir com o programa através da linha

Interface de linha de comando (CLI)

Executando o ficheiro cli.py no terminal, na respetiva pasta, começa a interação com o programa através da linha de comando:

```
Sistema de Consulta e Análise de Publicações Científicas

1 -> Importar ficheiro
2 -> Criar uma publicação
3 -> Consultar uma publicação
4 -> Consultar publicações
5 -> Eliminar uma publicação
6 -> Relatórios
7 -> Lister Autores
8 -> Atualizar publicação
9 -> Help
9 -> Sair

Escolha uma das opções:
```

Fig.1 - Linha de comando

Inicialmente, pode-se dar input em branco para receber a mensagem de aconselhamento à utilização do comando 9 (Help), para a compreensão do programa e as respetivas instruções.

```
Escolha uma das opções:
Se deseja ajuda para a utilização do programa, use o comando: 9

Escolha uma das opções: 9
Se pretende importar um ficheiro, salvar como novo ou adicionar um ficheiro a outro ficheiro, use o comando: 1
Se pretende criar uma nova publicação, ase o comando: 3
Se pretende consultar uma publicações, use o comando: 4
Se pretende eliminar publicações, use o comando: 4
Se pretende eliminar publicaçõe, ouse o comando: 5
Se pretende visusalizar os relatórios de estatatísticas, use o comando: 6
Se pretende listar todos os autores, use o comando: 7
Se pretende atualizar uma publicação, use o comando: 8
Se pretende sair do programa, use o comando: 8
Se pretende sair do programa, use o comando: 0
```

Fig.2 - Comando vazio e comando 9

Importação inicial

O comando 1 - importar - designa a variável "filepath" como sendo o resultado da função *file()* que, através do "import tkinter as tk / from tkinter import filedialog", abre uma janela que para selecionar um arquivo json de entre os vários documentos do user.

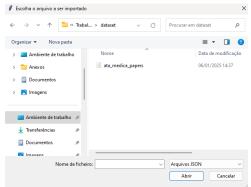


Fig.3 - Janela para selecionar os arquivos

Posteriormente, à variável data é atribuída a função *importar(filepath)*, que acede ao conteúdo do arquivo e verifica a sua estrutura, chamando as funções *verifcestrutura(item)* e *verifcautores(autores)*. Se o conteúdo respeitar o formato, o user pode escolher 3 opções:

- Trabalhar com o arquivo que importou;
- Salvar o conteúdo deste para um novo arquivo e trabalhar com este último;
- Adicionar o conteúdo do primeiro arquivo ao fim de outro arquivo já existente.

```
Escolha uma das opções: 1
Importação concluída. Escolha a próxima ação:

1. Salvar ficheiro.

1. Salvar como novo ficheiro.

2. Adicionar a um ficheiro existente.

Digite 1, 2 ou 3: 1
Ficheiro salvo: C:/Users/bcfig/OneDrive/Ambiente de Trabalho/algoritmo/Trabalho gr.13/novo.json
Importação concluída.
```

Fig.4 - Importação concluída com verificação do formato e escolha de trabalhar com o documento importado

Caso o user comece por selecionar outra opção sem ser 1, 9, 0 ou " ", é feito um print a indicar para iniciar com o comando 1.

```
Escolha uma das opções: 3
Primeiro, importe um arquivo com dados para consultar uma publicação.
```

Fig.5 - Mensagem que aparece caso o user ainda não tenha importado um ficheiro

Criar Publicação

O comando 2 chama a função *createPub()*, que permite adicionar uma nova publicação ao ficheiro. O utilizador fornece os dados necessários, e o código valida os campos obrigatórios (título e pelo menos 1 autor), verificando também se a data está no formato válido. Caso contrário, o sistema solicita a inserção correta. No final, a nova publicação é imediatamente guardada no ficheiro.

```
Escolha uma das opções: 2
Escreva o título da publicação:
O título é obrigatório. Por favor, insira um título válido.
Escreva o título da publicação: Demência nos adultos
Escreva a data de publicação (YYYY-MM-DD) ou deixe em branco: 2025-17-19
A data deve estar no formato YYYY-MM-DD.
Escreva a data de publicação (YYYY-MM-DD) ou deixe em branco: 2025-01-01
Escreva as palavras-chave (separadas por vírgula): demencia
Ouantos autores deseia adicionar?
Entrada inválida. Por favor, insira um número inteiro válido.
Quantos autores deseja adicionar? 1
Escreva o nome do autor: Beatriz Figueiras
Escreva a afiliação do autor:
Escreva o orcid do autor:
Escreva o DOI (ou deixe em branco):
Escreva o link do PDF (ou deixe em branco):
Escreva o link da URL (ou deixe em branco):
Escreva o resumo da publicação (ou deixe em branco):
Publicação criada com sucesso!
```

Fig.6 - Passo a passo para criar uma nova publicação

Consultar Publicação

O comando 3 atribui à variável *filtered_data* a função *tabfilterpt(data)*, que através de uma filtragem de 5 categorias relevantes, é apresentada uma tabela com toda a informação que se enquadre nos filtros.

Fig.7 - Filtragem inicial e tabela correspondente

Posteriormente, o user introduz o número do ID da publicação que pretende ver com maior detalhe, chamando a função *conspub(filterID, filtered_data)* e é apresentada uma nova tabela só com os dados dessa publicação.



Fig.8 – Input do ID e tabela com informação detalhada da publicação

Consultar Publicações

O comando 4 é semelhante ao 3, distinguindo-os o facto de poder consultar mais do que um ID.

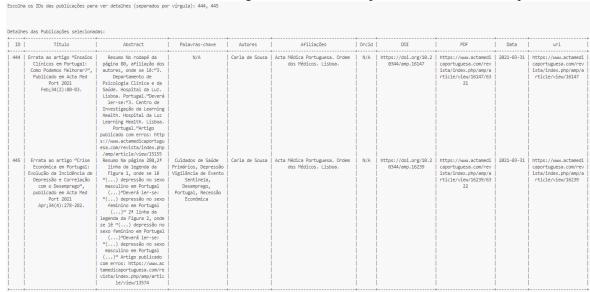


Fig.9 - Input dos ID e tabela com informação detalhada das publicações

Eliminar Publicação

O comando 5 é de igual forma semelhante ao comando 3. Contudo, substitui a função *conspub()* pela *DelPub()*, que permite eliminar a uma publicação pelo input do seu ID. Ainda nesta função, é garantido que os ID são reordenados, não havendo saltos entre eles.



Fig.10 - Filtragem das publicações e Input do ID da publicação a eliminar

Relatórios

O comando 6 começa por pedir ao user para escolher qual relatório pretende visualizar. Dentro dos disponíveis estão: 1- Distribuição de publicações por ano; 2-Número de publicações por autor e 3-Distribuição de palavras-chave pela sua frequência.

```
Escolha uma das opções: 6

1 -> Distribuição de publicações por ano.
2 -> Número de publicações por autor.
3 -> Distribuição de palavras-chave pela sua frequência (top 20 palavras-chave).
- -> Voltar para o menu anterior

Escolha o número do relatório: N
```

Fig.11 – Lista de comandos para escolher o relatório

Os resultados são exibidos no terminal com o auxílio das seguintes funções: *relPubAnoster()*, *relPubAutor()* e *relKeywords()*. Se o user introduzir o comando "-", retoma ao menu anterior (lista()).

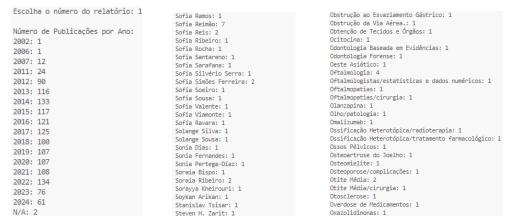


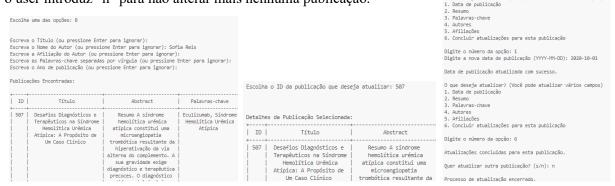
Fig.12 - Exemplos dos relatórios 1, 2 e 3, respetivamente

Listar autores

Ao utilizar o comando 7, a função *autores()* lista no sistema todos os autores por ordem alfabética.

Atualizar Publicação

No comando 8, a estrutura do código é igual à do 3, sendo apenas chamada a função *updatePubs(pub)* imediatamente a seguir à função *conspub*. Na função *updatePubs* é possível alterar os dados de 5 categorias, nomeadamente: data de publicação, resumo, palavras-chave, autores e afiliações. Esta função termina quando o user introduz "n" para não alterar mais nenhuma publicação.



 $Fig. 13, 14\ e\ 15 - Filtragem\ das\ publicações, seleção\ do\ ID\ da\ publicação\ a\ alterar\ e\ alterações\ da\ publicação, respetivamente$

Sair

Com o comando 0, é encerrada consulta com a mensagem "Saiu do programa".



Interface gráfica

wprincipal()

Esta é, como o próprio nome indica, a janela principal da interface, constituída por 6 botões:

- Help abre uma janela reduzida, a whelp(), que contêm listado o propósito dos 4 botões centrais;
- Sair fecha a janela;
- Importar abre a janela wficheirogui();
- Consultar abre a janela wconsultar(data, filepath);
- Listar Autores abre a janela wlistar(data);
- Relatórios abre a janela wrelatorios(data).



Fig.17 - Janela Principal - wprincipal()



Fig.18 – Janela whelp()

Se o user começar por selecionar os botões **Consultar**, **Listar Autores** e **Relatórios**, é adicionada uma linha de texto vermelha à janela principal a indicar: "Ainda não selecionou um ficheiro". Por outro lado, se começar por ****Importar**** e a importação for bem sucedida é adicionada uma linha verde: "Ficheiro importado com sucesso!", se for mal sucedida é uma linha vermelha: "Erro ao importar ficheiro!".



Fig. 19 e 20 - Janela Principal quando ainda não se importou um ficheiro e quando o ficheiro é importado, respetivamente

wficheirogui()

Nesta janela, existe apenas um botão - **Browse** - que, semelhante ao cli, atribui à variável filepath função *file()*, abrindo a pasta de arquivos do user, mas à variável data atribui a função *importargui(filepath)*, que apenas salva o ficheiro e utiliza o seu conteúdo.

É retomado à wprincipal o filepath e a data, que nada mais é do que uma cópia do conteúdo do ficheiro.

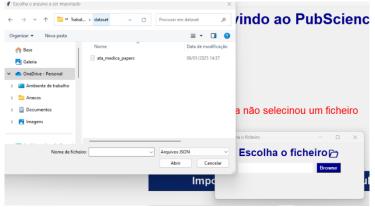


Fig.21 - Janela wficheirogui()

wconsultar()

Esta é a janela mais complexa da interface. Antes mesmo do layout, o formato do data é simplificado e transformado no formato de publicações, através da função *publicacoesc(data)*.

O layout da janela é formado por um título, 4 botões: 2 superiores (**Filtrar** e **Limpar filtros**) e dois inferiores (**Sair** e **Atualizar**), uma tabela e uma linha de texto informativa.

A tabela possui 5 colunas: "Título", "Autores", "Afiliações", "Data de publicação" e "Palavras-chave". Se o user clicar sobre as headings: "Título", "Autores", e "Data de publicação" o conteúdo da tabela será ordenado consoante a heading selecionada. Isto é possível através do evento == ('-TABELA-', '+CLICKED+', (-1, 0)), mais especificamente, '-TABELA-' identifica o elemento do layout, '+CLICKED+' indica o tipo de evento (clicar) e o (-1, 0) representa as coordenadas (linha, coluna), aqui usamos linha igual a -1 pois os headins não fazem parte das linhas do conteúdo da tabela. Para a ordenação são chamadas as funções: ordpubtitulo(publicacoes,ori), ordpubnome(publicacoes,ori) e ordpubdata(publicacoes,ori).



Fig.22 - Janela wconsultar()

O botão **Filtrar** abre a janela *wFilterPub(publicacoes)*, que através do preenchimento de pelo menos um dos inputs, filtra pelas publicações que se enquadram, retomando à janela anterior com apenas as publicações filtradas na tabela.

💸 Filtro de Publicações		×
Título		
Autor		
Afiliação		
Palavras-chave (separadas por vírgulas)		
Ano de Publicação		
Filtrar		

Fig.23 – Janela wFilterPub()

O botão **Limpar filtros** volta a formatar o data em publicações e a tabela volta a ter as publicações iniciais.

O botão **Atualizar** usa exatamente o mesmo código que o ****Limpar filtros****, contudo, é utilizando quando se retoma da janela *wPublicacao()*, pois o data já é diferente em algum aspeto e a tabela não atualiza de imediato.

Por fim, o evento '-TABELA-' and values['-TABELA-']!=[], seleciona uma linha da tabela que não seja vazia e atribui os seus valores/conteúdo à variável *linha_selec*, que se existir fornece a informação da *linha_selec* em publicacoes à variável pub. Esta última variável entra na chamada da janela *wPublicacao(pub, data, filepath, publicacoes)*.

wPublicacao()

Nesta janela, o layout inicia com 3 botões: **Apagar Publicação**, **Criar Publicação** e **Editar Publicação**, por baixo destes estão vários títulos, textos dentro de retângulos simples e textos dentro de blocos com barras que deslizam, que permitem mostrar toda a informação detalhada da publicação selecionada. Ainda no layout, há o botão **Fechar** para fechar a janela.



Fig.24 - Janela wPublicacao()

Ao selecionar o botão **Apagar Publicação** é aberta a janela *wConfirmacao()*, estilo *popup*. Se a opção for sim, é chamada a função *deletepubgui(pub, data, filepath)*, que compara o ID da publicação selecionada com os ID do data e elimina a que for correspondente. De seguida, o filepath é atualizado com o novo data e o data volta a sofrer a transformação para publicacoes, retomando este ultimo para a janela *wconsultar*, com o fechar automático da janela *wPublicacao()*.



Fig.25 - Janela wConfirmacao()

Por outro lado, se o botão selecionado for **Criar Publicação** é aberta a janela *criar Pub (data, filepath)*. Nesta janela, o user preenche os espaços consoante o seu interesse para criar uma publicação. Esta é, posteriormente, transformada na estrutura do data, o filepath é atualizado e data volta a transformar-se em publicações.



Fig.26, 27, 28 e 29 – Janela criarPub() e respetivos popups para preencher a nova publicação

Por fim, se o botão selecionado for o Editar Publicação é aberta a janela editarpublicacao(pub, data, filepath, original_id=pub['ID']). O layout é semelhante ao wPublicacao() mas mais simplificado, sendo possível alterar o conteúdo dentro de cada quadrado. Nesta janela são chamadas duas funções process_authors(pub) e format_authors(authors). Apesar da informação ser alterada no data, no filepath e em publicacoes, a parte dos nomes, afiliações e orcid é apagada e não conseguimos detetar o problema.

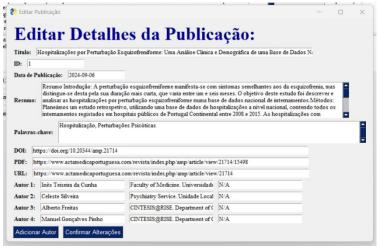


Fig.30 - Janela editarpublicacao ()

wrelatorios()

Nesta janela existem 6 botões centrais mais o botão sair. Nos botões centrais temos: 'Distribuição de publicações por ano', 'Distribuição de publicações por mês', 'Número de publicações por autor', 'Distribuição de publicações de um autor por anos', 'Distribuição de palavras-chave pela sua frequência' e 'Distribuição de palavras-chave mais frequente por ano.' Consoante o botão central selecionado é chamada a função wpubano(data), wpubmes(data), wpubsautor(data), wpubsautor(data), e wkeyfreano(data), respetivamente.



Fig.31 - Janela wrelatorios()

wpubano(data)

Esta janela chama, por sua vez, a função *relPubAnosgui(data)*, que permite criar um gráfico de pontos ligados relativamente ao número de publicações por cada ano, usando *o Counter()* para a contagem das publicações e *pyplot* para gerar o gráfico.

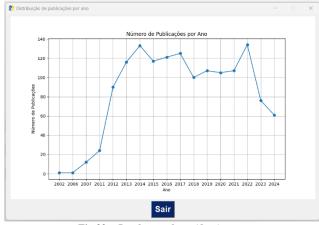


Fig.32 – Janela wpubano(data)

wpubmes(data)

Nesta janela o user introduz o ano que pretende visualizar a distribuição de publicações por mês e chama a função *distribPubPorMesgui(data, ano)*, semelhante à *relPubAnosgui(data)*.

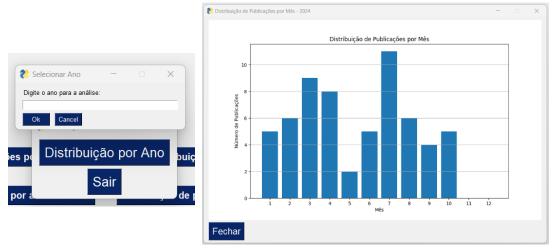


Fig.33 e 34 – Popups para introduzir o ano e janela wpubmes(data)

wpubsautor(data)

Esta janela chama a função *relPubAutorgui(data, top_n=20)*, que cria uma lista de dicionários, sendo os nomes dos autores as chaves e o número de publicações os valores. Esta lista é organizada e seleciona o top 20 autores com mais publicações, para gerar um gráfico de barras.

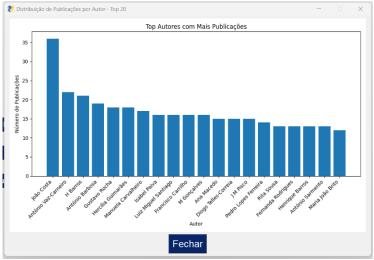


Fig.35 – Janela wpubsautor(data)

wpubautorano(data)

Nesta janela o user introduz o nome do autor, é aberto um popup com a lista de autores que façam correspondência parcial ao input do user, de seguida este seleciona o que pretende e é chamada a função *PubAutoPorAnogui(data, autor)* para gerar o gráfico de distribuição de publicações por anos.



Fig.36 e 37 - Popups

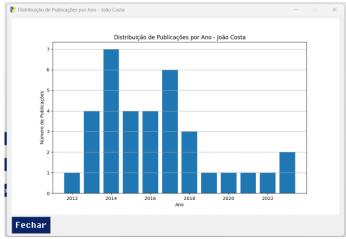


Fig.38 - Janela wpubautorano(data)

wpubskeyword(data)

Esta janela é muito semelhante à *wpubsautor(data)*, chamando a função *relkeywordsgui(data, top_n=20)*. As palavras mais frequentes apresentadas correspondem a caracteres, apesar de que na função *relkeywordsgui(data, top_n=20)* ocorrer uma filtragem para que tal não aconteça.

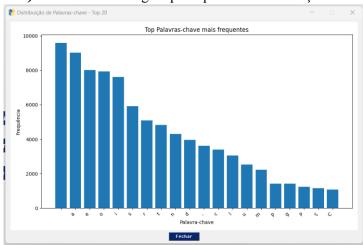


Fig.39 – Janela wpubskeyword(data)

wkeyfreano(data)

Esta janela chama a função keyfreano(data), contudo o gráfico gerado fica em branco.



Fig.40 – Janela wkeyfreano(data)

wlistar(data)

Esta janela não está operacional, mas decidimos desenvolver pelo menos o seu layout. No código são apresentadas algumas funções ao clicar nos botões, no entanto, estas dão erro.

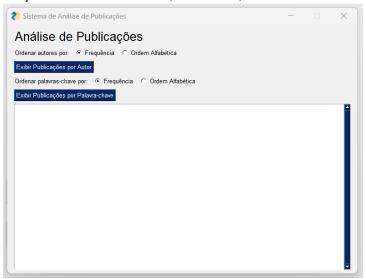


Fig.41 – Janela wlistar(data)

Conclusão

A criação deste sistema de consulta e análise de publicações científicas em Python, irá permitir ao utilizador desta ferramenta, e criar, atualizar e analisar publicações. Ou seja, com base num dataset de publicações, o sistema possibilitar a pesquisa de artigos usando filtros relevantes, tais como a data de publicação, as palavras-chave, autores, etc. E ainda ser possível, gerar relatórios (mostrando gráficos ilustrativos com estatísticas) detalhados para a análise de métricas dos artigos e dos seus autores.

O desenvolvimento deste projeto, permitiu alargar em grande escala os nossos conhecimentos sobre esta disciplina e testar a nossa capacidade lógica e criativa.

Bibliografia

https://www.w3schools.com/