Sistema de Bases de Dados para o Tombo da Mitra

Bruno Arieira, [1,2] João Gomes, [1,3] Tiago Fraga, [1,4]

- ¹ Universidade do Minho, Braga, Portugal
 - a70565@alunos.uminho.pt
 a74033@alunos.uminho.pt

 - 4 a74092@alunos.uminho.pt

Resumo O Projeto apresentado insere-se na unidade curricular de Laboratório de Engenharia Informática e tem como principal objetivo a implementação de um Sistema de Base de dados capaz de suportar a informação contida nos documentos do Tombo da Mitra. Com a criação deste sistema existe ainda o desenvolvimento de uma interface web para importação, análise e gestão destes documentos. O Tombo da Mitra é constituído por 644 fólios escritos em português contendo informação relativa ao tipo de terras, acidentes do terreno, nomes de ruas, proprietários e apontamentos biográficos e genealógicos, etc. O sistema proposto têm de possuir as ferramentas necessárias para o tratamento desta informação de modo a permitir que estudiosos deste tema consigam fazer uma análise de forma mais simples e aprofundada. Nos pontos seguintes irá ser apresentada de forma mais detalhada e consistente, todas as etapas do desenvolvimento deste projeto.

Keywords: Base de dados · Tombo da Mitra · Fólios.

1 Introdução

No primeiro quartel do século XVIII, as propriedades da mesa arcebispal cruzavam todo o Minho e Trás-os-Montes, estendendo-se para fora dos seus domínios religiosos, chegando mesmo a Santarém e passando também as fronteiras do reino. Todas as propriedades eram alistadas num inventário, que ficou conhecido por Tombo da Mitra (também designado de Livro das Propriedades) que regista 1288 páginas manuscritas de tamanho A3, de terras e casas. Nos 644 fólios (folhas), escritos em português, contém informação relativa a tipos de terras, acidentes do terreno, nomes de ruas, proprietários e apontamentos biográficos e genealógicos, etc. O objetivo deste projeto passa pela criação de um sistema de Base de dados capaz de suportar este tipo de documentos, acompanhada de uma interface web que permita a importação destes documentos, a análise dos mesmos e a sua gestão de uma forma simples, facilitando o trabalho do grupo de estudiosos do Tombo da Mitra. Como motivação temos tornar o tombo acessível a um grupo de estudiosos e estudar de forma mais precisa o 'Tombo' recorrendo a mecanismos de análise de dados e de processamento de linguagens. Com este projeto delineamos um conjunto de objetivos a atingir:

- Criar um sistema de bases de dados capaz de acolher o conteúdo do 'Tombo';
- Providenciar um conjunto de mecanismos de indexação e pesquisa que permitam auxiliar o processo de estudo e análise do conteúdo do 'Tombo';
- Desenvolver uma plataforma de trabalho flexível, acessível a todos aqueles que pretendam estudar o 'Tombo';

Nas secções seguintes vai ser abordado o trabalho relacionado em que iremos identificar sistemas semelhantes e apresentar as suas características base; conceção e implementação do sistema em que explicaremos com mais detalhe as características da aplicação bem como as suas funcionalidades e serviços; recursos computacionais onde identificamos as ferramentas escolhidas para o desenvolvimento deste projeto, bem como, explicaremos o porquê da sua escolha; instalação e configuração explicar como se faz a instalação do sistema desenvolvido num computador localmente, bem como a sua configuração base, e os recursos que são necessários; funcionamento do sistema apresentar todo o funcionamento do sistema desenvolvido; e por último teremos as conclusões e trabalho futuro onde realizaremos uma exposição e análise crítica de aspetos positivos e negativos relacionados com o sistema desenvolvido, bem como identificaremos novas características, funcionalidades ou serviços para desenvolvimento futuro.

2 Trabalho Relacionado

Depois de efetuada a devida pesquisa, acreditamos que não existam aplicações com funcionalidades idênticas a nossa, por isso trata-se de um produto único, dando mais valor ao projeto realizado.

3 Conceção e Implementação do Sistema

Nesta secção vamos explicar detalhadamente o sistema criado, identificando as características gerais, funcionalidades e serviços. Iremos abordar ainda o modelo computacional do projeto explicando todos os seus componentes.

3.1 Arquitetura do Sistema

O sistema desenvolvido apresenta a seguinte arquitetura:



Figura 1. Arquitetura do Sistema

Tal como se pode observar na imagem a nossa arquitetura apresenta o módulo Extrator, Indexador e Interface.

O extrator tem como objetivo extrair o texto do ficheiro e gravar na base de dados, para isso usa como filtro as *tags* presentes no texto e introduzidas anteriormente na aplicação, assim o extrator obtêm retira o texto por completo do ficheiro e retira o conteúdo presente em cada *Tag* bem como o número de ocorrências de cada *tag*.

O indexador tem como objetivo indexar todas as palavras presentes no texto, retirando linha em que aparecem e ainda número de ocorrências.

A interface é o módulo que vai permitir a visualização e a interação com a aplicação desenvolvida, sendo simples e intuitiva, de modo a que o grupo de estudiosos possa utilizar corretamente o sistema.

A seguir são apresentadas as funcionalidades da aplicação com detalhe, para melhor compreensão

3.2 Sistema

Tal como foi dito nas secções anteriores trata-se de um sistema de base de dados capaz da armazenar os documentos pertencentes ao Tombo da Mitra, com isto temos ainda um sistema com uma interface web, para poder ser utilizado por um grupo de estudiosos. De modo a solucionar-mos o problema proposto desenvolvemos uma interface web com as seguintes funcionalidades:

- Importação de Fólios

Permite ao utilizador importar um documento de texto que será armazenado na base de dados, introduzindo assim dados no sistema.

- Visualização de Fólios

Permite visualizar o conteúdo dos Fólios armazenados.

Análise de Fólios

Permite a análise de fólios, ou seja, analisar texto completo ou frações de texto, conforme o critério do utilizador.

Visualização de Estatísticas

Visualização de estatísticas referentes ao número de documentos introduzidos, número de palavras, número de *tags* presentes no sistema e número de índices.

- Gestão de Utilizadores

Permite adicionar ou remover utilizadores.

4 Recursos Computacionais

Nesta secção vamos explicar quais as tecnologias usadas para o desenvolvimento do projeto, bem como, o porquê de escolhermos tal tecnologia.

4.1 Base de Dados

Para sistema gestor de base de dados usamos *MongoDB*.

Este sistema trata-se de uma base de dados orientada a documentos, sendo não relacional, não possuindo uma estrutura rígida para o armazenamento dos dados, como acontece em bases de dados relacionais.[1]

Como o conteúdo dos textos não era conhecido na integra, optou-se por este sistema, que permite essa flexibilidade na inserção de dados.

Para poder ser utilizado este sistema juntamente com o nossa aplicação basta realizar o download da ferramenta.[2]

4.2 Linguagem de Desenvolvimento

Para linguagem de programação foi utilizado python.

Python é uma linguagem de programação de alto nível, multi-paradigma, isto é, suportada pelo paradigma imperativo, orientado a objetos, funcional e de procedimentos.[3]

A escolha desta linguagem deve-se a facilidade oferecida por ela para processamento de texto, que é grande parte deste nosso projeto.

Para uso é necessário download da ferramenta.[4]

4.3 Framework de Desenvolvimento

Como framework de desenvolvimento foi usado Flask.

Flask é uma microframework para python baseada em Werkzeug, Jinja 2 .[5] Flask tem a flexibilidade da linguagem de programação Python e oferece um modelo simples para o desenvolvimento web.[6]

Usar esta microframework é como usar qualquer pacote Python, instalandose da seguinte forma :

```
pip install Flask
```

4.4 Interface

Para a interface foi utilizado *Gentelella*. Esta ferramenta é um *template* de **bootstrap3**, possuindo várias bibliotecas para gráficos, calendários, formularios, etc.[7]

Para poder usar esta ferramenta é necessário ter o instalador de pacotes do *NodeJS*, denominado npm, de modo a realizar a sua instalação.[8]

```
npm install gentelella —save
```

4.5 Sistema Operativo

No desenvolvimento deste projecto foram usados dois tipos de sistemas operativos, macOS e a distribuição $Linux\ Ubuntu\ 18.04$. Havendo no entanto capacidade de utilizar Windows, visto que todos os programas e ferramentas utilizadas têm versão para Windows.

5 Instalação e Configuração do Sistema

De modo a compilar o programa numa máquina local, é preciso instalar todos os requisitos, para isso foi criado um ficheiro de texto com todos os requisitos. Pode ser optado por instalar um a um recorrendo ao instalador de pacotes de *Python* ou então executar o seguinte comando:

```
pip install -r requirements.txt
```

Posto isto para executar a aplicação é preciso executar os seguintes comandos

```
export FLASK_APP=tommi.py
flask run
ou
flask run --host=0.0.0.0
```

6 Funcionamento do Sistema

Nesta secção irá ser apresentado todo o funcionamento do sistema, recorrendo a imagens retiradas da aplicação web.

6.1 Login

Login executado pelo email e pela palavra-passe de utilizador, esta é a página principal da aplicação, só utilizadores autenticados é que conseguem utilizar a aplicação, podendo no entano consultar as páginas informativas como os Créditos, Saber Mais, Termos e Privacidade.



Figura 2. Página de Login

6.2 Home

Na página principal é apresentada uma estatística acerca do número de fólios inseridos na base de dados, número de palavras, número de utilizadores registados, número de tags e o número de definições ativas, como se pode observar na imagem.



 ${\bf Figura\ 3.\ Home}$

6.3 Barra de Navegação

Todas as páginas apresentam esta barra de navegação é através dela que se realiza toda a navegação pela aplicação web.



Figura 4. Barra de Navegação Lateral

6.4 Importação

A importação é dividida em cinco passos, ou seja a gravação do conteúdo do ficheiro só é realizada quando estes cinco passos estiverem todos terminados.



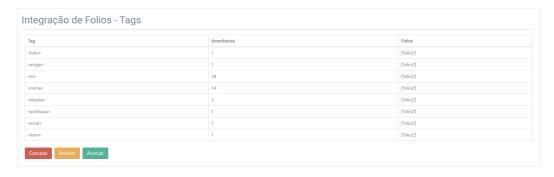
Figura 5. Passo1 - Identificação

Posto este passo temos a visualização dos fólios com tags e sem tags, de modo a permitir que o utilizador reveja o documento que inseriu, verificando se não há enganos.



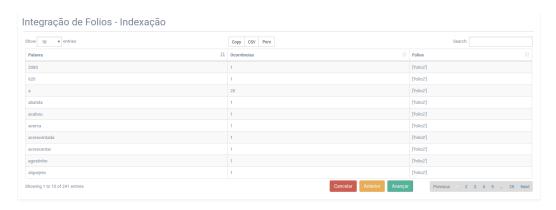
Figura 6. Passo1 - Visualização

Com a verificação do documento verificada é apresentado o calculo das tags encontradas no ficheiro.



 ${\bf Figura~7.~Passo1~-~Tags}$

O passo posterior as Tags é a apresentação do calculo da indexação palavra a palavra do ficheiro.



 ${\bf Figura\,8.}$ Passo1 - Indexação

Para finalizar é apresentado um resumo ao utilizador do processamento do ficheiro inserido carregando no botão gravar, a gravação na base de dados é efetuada, salientando que em cada passo deste processo é oferecido ao utilizador o cancelamento ou mesmo voltar a um passo anterior.



Figura 9. Passo1 - Resumo

6.5 Documentos

Como se pode observar na barra de navegação apresentada anteriormente, este botão é diferente dos demais, visto conter uma lista para três páginas distintas.

Primeiro para a página de Fólios, onde o utilizador poderá visualizar todos os fólios inseridos na base de dados numa primeira instância ou verificar posteriormente o conteúdo de um fólio em concreto.



Figura 10. Listagem de Fólios

Clicando no símbolo do "olho" podemos visualizar o conteúdo do documento inserido.

Fólio: 619

619

Fólio 619v.

2019-06-18 18:33

<folio> [619]</folio> cartigo> Campo da vinha</artigo> O[s] Religiosos da Ordem de São Bento possuem Huas' Casas, que forão do Doutor <nome>Fernão <n|> Mergulhão</nome> são propriedade da Mesa Arcebispal; e dellas lhe pagão de foro en cada <nl> hu' anno trezentos rs' tem dellas prazo en fateosim perpetuo, d' fez o sor Arcebpo' <nome>Don <nl> Baltasar limpo</nome>, o Sitio das quaes Casas he o que se segue. <item> H As Casas, que possuem os Religiosos de são Bento do Mosteiro de <local>Tybães</local> são <nl> sobradadas e Muyto bem acabadas, tem a porta hu' Alpendre co' duas Colunas <nl> de Pedra tem por detras seu quintal cercado de paredes altas, e de comprido de <nl> Norte a sul Casas e quintal setenta e cinco varas, e de largo de Nacente a poente <nl> pella parte do sul dezoito varas e pella do Norte vinte e duas, partem do Nace'te <nl> co' a Rua do Carualhal, e do Norte co' o Caminho, que uay pera a Cerca do <nl> mostejro do Populo, e do poente co' quintal e Casas de <nome>Agostinho Ribejro</nome> que <nl> tambem são da Mesa, e do sul co' o Campo da Vinha. <nome>Agostinho Ribejro</nome> filho, que ficou do Doutor <nome>Pero Ribejro</nome> possue outras <nl> Casas per titulo de Prazo fateosim perpetuo, que apresentou escrito por <nome>Manoel <nl> de Lemos</nome> aos cinco de Mar‡o de Mil quinhentos cincoenta e oyto, o qual fez <nl> o sor Arcebpo <nome>Dom Baltasar limpo</nome> ao Doutor <nome>Manoel Sardinha</nome> paga de <nl> foro en cada hu' anno a Mesa Arcebispal trezentos rs', e as ditas Casas são <nl> as que se seguem.</item> <item> H As Casas que traz o Lecenceado <nome>Agostinho Ribejro</nome> parte' do Nacente co' as <nl> Casas sobreditas, que trazel os Padres de Tybães, e do poente col Casas que ficarão <nl> de <nome>fernão de Sousa</nome> tambem da Mesa, e do Norte col Caminho, que uaj <nl> pera a Cerca do Populo em o qual caminho o quintal dellas uay entestar e <nl> do sul partem com o Campo da Vinha, São Sobradadas e Muyto grandes e <nl> boas' tem de comprido co' o dito quintal de Norte a sul Nouenta e oyto varas; <nl> e de largo na cabe‡a do Norte vinte e duas e na do sul dezenoue. Dona <nome>Helena de Souto Mayor</nome> filha de <nome>Pero da Cunha</nome> viuua, que ficou <ni>por falecimento de <nome>Antonio de Sousa</nome> possue outras Casas per prazo fateosim perpetuo paga dellas trezentos rs' en cada hu' anno as quaes Casas <nl> são as que se seguem.</item> <item> H As Casas de que ora est de posse Dona <nome>Illena Sobradadas</nome> e Muito fermosas <nl> partem do Nacente co' as Casas sobreditas, que traz o Lecenceado <nome>Agostinho <nl> Ribejro</nome>, e do poente com casas tamb,m da Mesa, que traz o Conego <nome>Aluaro Coelho</nome> <nl> e do Norte com o Caminho que uem dos chãos pera a Cerca dos frades, & do sul <nl> com o Campo tem de largo pella parte do sul Na face da Rua dezenoue varas <nl> e na do Norte vinte: e tem de comprido Casas e quintal Nouenta e seis varas. O Conego <nome>Aluaro Coelho</nome> possue outras Casas per titulo de Prazo fateosim perpetuo <nl> feito ao Doutor <nome>Manoel Coelho</nome>, paga dellas de foro en cada hu' anno a <nl> Mesa Arcebispal trezentos rs' as quaes Casas são as que se seguen. </item> <item> H As Casas, que possue o Conego <nome>Aluaro Coelho</nome>, partem do Nacente co' Casas e <nl> quintal de Dona <nome>lilena</nome> acima Nomeadas [sic]#, e do poente co' o Mostejro da frejra <nl> do Saluador, cuio Sitio tambem he da Mesa Arcebispal, e do Norte co' o Caminho <nl> que uay pera a Cerca dos Padres do Populo, e do sul co' o Campo da vinha tem de <nl> Nacente a poente na fa‡e da Rua dezenoue varas e mea, e de Comprido de <nl> Norte a sul com o quintal Nouenta e Noue uaras. O Sitio do Mostejro do Saluador assi Como est co' su

Figura 11. Fólios - Visualização

Clicando no menu no botão de Índices temos a uma página com o resultado da indexação de todos os Fólios na base de dados, a semelhança do anterior possuímos a capacidade de ver um índice ao pormenor.

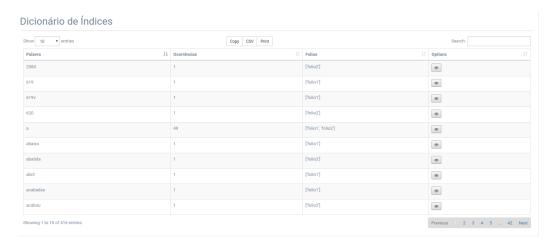


Figura 12. Listagem de Índices

Visualização depois de clicado o botão com o símbolo de um "olho".



Figura 13. Índices - Visualização

Por último como é apresentado na barra de navegação temos o botão de Tags, dando acesso a uma página com todas as tags encontradas nos ficheiros, podendo visualizar o conteúdo de cada tag ao pormenor.

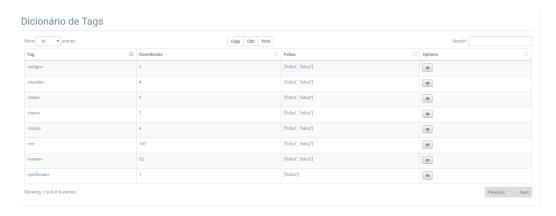


Figura 14. Listagem de Tags

Depois de clicado o símbolo do "olho" temos acesso ao conteúdo de cada
 $tag.\,$



 ${\bf Figura\,15.}\ {\rm Tags}$ - Visualização

6.6 Análise

A análise é uma parte importante da aplicação, é onde o utilizador pode efetuar uma pesquisa com base em parâmetros, tais como: por texto, por palavra, por tag, pode filtrar a um fólio em específico e ainda pode selecionar período, linha ou número de palavras. Como se pode visualizar na imagem a seguir exposta.

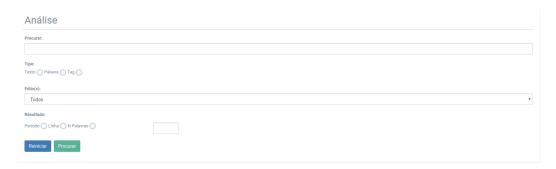


Figura 16. Análise

6.7 Utilizadores

Na página relativa aos utilizadores temos a visualização de todos os utilizadores como página principal.



Figura 17. Lista de Utilizadores

A partir desta página pode-se clicar no símbolo do "olho"e temos acesso a visualização do perfil de utilizador.



Figura 18. Visualização de Utilizador

Voltando a página principal de utilizadores pode-se clicar no símbolo do "lápis"e ter acesso a edição de utilizadores.



 ${\bf Figura~19.}$ Edição de Utilizadores

Ainda na página pode-se clicar no símbolo de "mais" e entramos na adição de utilizadores.



Figura 20. Adição de Utilizadores

Ainda existe a opção de remover utilizador que só remove um utilizador da base de dados sendo ilustrada pelo símbolo "caixote do lixo"

6.8 Definições

A página de definições serve para adicionar tags, visualizar tags e remover tags. Estas tags são tidas em conta na pes quisa nos textos, no momento da importação do fólio.

A página principal de definições é a seguinte:



Figura 21. Definições - Página Principal

Clicando no botão de "adicionar" temos a página com o formulário de adição de tags ao sistema.



 ${\bf Figura\,22.}$ Definições - Adição de ${\it Tags}$

Quando clicado o botão de "ver"
temos uma listagem das $\it tags$ gravadas no sistema.

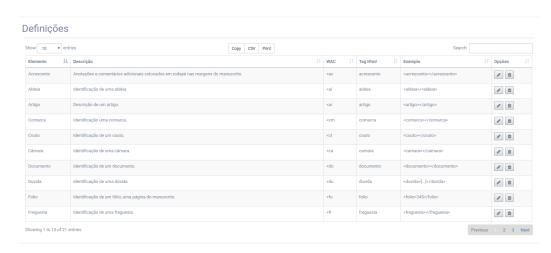


Figura 23. Definições - Ver Tag

Nesta página ainda temos a opção de remover Tag ou editar tag que leva até uma nova página de edição.

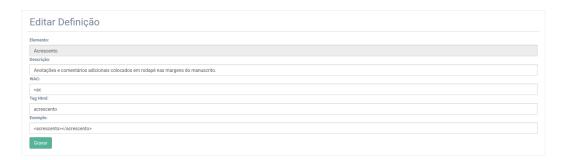


Figura 24. Definições - Ver Tag

7 Conclusões e Trabalho Futuro

A aplicação desenvolvida é capaz de importar fólios, visualizar os mesmos, processar o texto indexando todas as suas palavras, efetuar uma analise aos fólios introduzidos e possuí ainda gestão de utilizadores.

Esta aplicação irá ser muito útil a qualquer grupo de estudiosos que queiram estudar o Tombo da Mitra na sua totalidade, visto se tratar de uma versão digitalizada e modernizada, possuindo assim uma fácil leitura do documento, o que não era possível na sua versão original. Além disto é permitido efetuar uma análise precisa e mais rápida.

Dado por concluído o projeto temos que salientar que visualmente existam algumas melhorias a fazer, tirando melhor proveito do *template* usado. Em contrapartida salientamos que a aplicação está pronta a ser usada, com todas as funcionalidades propostas operacionais.

Como trabalho futuro achamos que a interface podia ser melhorada visualmente, melhorar os mecanismos de indexação e pesquisa recorrendo a técnicas de processamento de linguagens, de modo, tornar a nossa aplicação numa ferramenta com maior capacidade de análise. Tornar a aplicação acessível a toda a gente, visto atualmente só utilizadores registados tenham acesso. Permitir remoção de fólios e edição dos mesmos.

Referências

- 1. MongoDB wikipedia, https://en.wikipedia.org/wiki/MongoDB.
- $2. \ \ MongoDB \ download, \ \texttt{https://www.mongodb.com/download-center/community}$
- Python wikipedia, https://en.wikipedia.org/wiki/Python_(programming_ language).
- 4. Python download, https://www.python.org/downloads/
- 5. Flask homepage, http://flask.pocoo.org/
- 6. Flask wikipedia, https://pt.wikipedia.org/wiki/Flask_(framework_web)
- 7. Gentelella github, https://github.com/ColorlibHQ/gentelella
- 8. NodeJS download, https://nodejs.org/en/download/