Departamento de Engenharia Informática e de Sistemas  
Instituto Superior de Engenharia de Coimbra

Licenciatura em Engenharia Informática

**Integração de Dados**

**2021/2022**

Uma imagem com texto

Descrição gerada automaticamente

Trabalho prático

Autores

Maria José Gonçalves Marcos – 2018018386

Tiago Alexandre Pais Dias - 2019126344

Índice

[Introdução 3](#_Toc104304849)

[Estrutura do projeto 3](#_Toc104304850)

[Análise das fontes de dados 5](#_Toc104304851)

[Esquema Global 6](#_Toc104304852)

[*Wrappers* Implementados 7](#_Toc104304853)

[Ficheiro *XML* 8](#_Toc104304854)

[Validação do modelo 9](#_Toc104304855)

[Pesquisas *XPATH* 10](#_Toc104304856)

[Ficheiros *XSLT/XQuery* 11](#_Toc104304857)

[Interface 12](#_Toc104304858)

[Conclusão 13](#_Toc104304859)

# Introdução

Com a realização deste trabalho pretende-se criar um programa em Java, composto por vários *wrappers* que obtenham dados de fontes heterogéneas, distribuídas e autónomas e possibilitem ao utilizador a visualização dos dados de forma integrada. A seguinte aplicação integradora apresenta uma visão unificada de informações relativas a cidades de diferentes países.

# Estrutura do projeto

# Análise das fontes de dados

Inicialmente foi analisada a distribuição da informação das cidades pelos websites. A decisão foi feita com base na existência, ou inexistência, de determinada informação e na simplicidade das expressões regulares.

A tabela ilustrada abaixo, mostra o website escolhido para a retirada de cada um dos atributos.

|  |  |  |
| --- | --- | --- |
|  | Wikipedia | Db-city |
| Nome da cidade |  |  |
| País a que pertence |  |  |
| Indicação se a cidade é a capital do país |  | x |
| Link para a imagem da bandeira do país |  | x |
| Língua(s) oficial(ais) do país |  | x |
| Link para a imagem da bandeira da cidade | x |  |
| Links para as imagens de monumentos/*landmarks* da cidade | x |  |
| Área da cidade (valor numérico em km2) |  | x |
| Nº de habitantes da cidade |  | x |
| Densidade populacional da cidade (nº de habitantes por km2) |  | x |
| Código Postal da cidade |  | x |
| Presidente da Câmara da cidade |  | x |
| Latitude e Longitude da cidade |  | x |
| Altitude da cidade em metros |  | x |
| Clima da cidade |  | x |
| Fuso Horário da cidade |  | x |
| Website da cidade |  | x |
| Cidades geminadas |  | x |

# Esquema Global

# *Wrappers* Implementados

String obtemLinkdbCity(String pais,String cidade)

A função obtemLinkdbCity recebe o país e o nome da cidade e devolve o link para a página principal da cidade indicada. Para tal, primeiro junta a String do país ao link da página DbCity e obtém a página principal com a informação daquele país, guardando o respetivo source num ficheiro txt (“pais.txt”). As próximas funções fazem algumas das suas pesquisas nesse ficheiro txt. De seguida pesquisa nesse source o link da cidade indicada recorrendo à expressão regular:

String er ="href=\"/([A-Za-z-]+--[A-Za-z-]+[^\"]+)\" title=\""+cidade+"\">"+cidade+"</a>";

String obtemLinkWikipedia(String cidade)

A função obtemLinkWikipedia recebe o nome de uma cidade e junta essa String ao link da página da Wikipédia, obtendo a página principal da cidade indicada r gurdando o seu respectivo source num ficheiro txt (“cidade.txt”). As próximas funções fazem as suas pesquisas no ficheiro txt acima mencionado. Para tal recorre à expressão regular:

String er ="\"wgPageName\":\"([A-Za-z-\_]+)\"";

# Ficheiro *XML*

# Validação do modelo

# Pesquisas *XPATH*

# Ficheiros *XSLT/XQuery*

# Interface

# Conclusão