

DETI - ENGENHARIA DE DADOS E CONHECIMENTO

RELATÓRIO FINAL - TRABALHO PRÁTICO 1

1. Introdução ao Tema

Neste primeiro trabalho prático, optamos por desenvolver uma aplicação web onde é possível realizar uma gestão de multimédia, nomeadamente gestão de filmes e séries, onde é possível para além de saber a informação destes, adicionar novos filmes e séries, eliminar, avaliar ou até editar.

2. Dados e suas Fontes

Após pesquisarmos vários ficheiros com as mais variadas informações, acabamos por escolher o ficheiro <https://www.kaggle.com/shivamb/netflix-shows> (CSV) no domínio <https://www.kaggle.com/datasets>. O ficheiro não era XML, o que levou a que fosse necessária uma conversão para XML através de um programa de Python. O novo ficheiro ficou a chamar-se “streamData.xml”.

3. Esquema dos Dados

Através do IDE PyCharm, geramos de forma automática um XML Schema para o ficheiro “streamData.xml”.

Na aplicação web, colocámos um botão “Verify XML” onde é possível validar dois ficheiros XML, o “streamData.xml” e o “streamDataError.xml”, sendo este último um pequeno ficheiro XML com poucos filmes e com um erro propositado na sua estrutura. Ambos são validados através do mesmo ficheiro “streamData.xsd”, mas apenas um deve passar a validação. Esta é feita através de uma *view*, que importa algumas funções da biblioteca *etree*.

4. Transformações sobre os dados

Para transformar documentos XML noutros formatos, utilizámos XSLT mais concretamente na página do perfil do utilizador e no feed RSS.

Toda a informação para as transformações está síncrona com a base de dados e é feito o parse com a a funcao *fromstring()* da biblioteca *etree* para o manuseamento da informação a ser apresentada no website.

Nos ficheiros .xsl tentámos explorar as funcionalidades do xslt como o `<xsl:value-of>`, o `<xsl:for-each>`, o `<xsl:sort>` e o `<xsl:if>` para controlar os dados em questão.

5. Operações sobre os dados

Grande parte das funcionalidades da nossa aplicação envolvem *XQuery* e *XQuery Update*, pois todas as alterações realizadas são guardadas na Base de Dados.

Criamos múltiplos *Library Modules*, vários destes devolvem informação filtrada, facilitando a pesquisa na base de dados, outros servem para fazer um *Update* à Base de Dados.

Os *Library Modules* implementam várias expressões *FLWOR*, localizam vários nós e algumas devolvem o texto de vários elementos. Exploramos também a criação de novos nós, a eliminação e a substituição de alguns nós.

6. Funcionalidades da aplicação

A nossa aplicação tem diversas funcionalidades, entre estas a capacidade de filtrar programas por Filme/Séries e também a hipótese de filtrar por país onde foi realizado.

Em cada programa podemos ainda marcá-lo como visto, permitindo assim ao utilizador gerir os filmes que já viu de forma bastante simples, e avaliá-los numa escala de 0 a 5. Os filmes e séries serão enumerados por ordem decrescente de rating, ou seja os filmes avaliados com maior score estarão nos primeiros lugares (topo da página).

O administrador terá ainda a hipótese de remover um filme/série da base de dados ou de adicionar um filme/série.

O utilizador tem também a hipótese de consultar as últimas novidades do mundo do cinema no tab *News*, esta página é atualizada diariamente.

Outro separador bastante útil é o do perfil, neste o utilizador pode ver os seus dados (foto e nome) e duas listas, uma que contém os filmes vistos e outro as séries vistas.

7. Conclusão

Pensamos que o nosso site possui todas as funcionalidades necessárias para ser utilizado de forma útil, talvez faltando apenas uma barra de pesquisa para encontrar mais facilmente um filme em específico, e claro um sistema de login que não chegou a ser implementado.

8. Configuração para executar a aplicação

- Abrir BaseX e criar nova base de dados com o nome “streamData” com o ficheiro ‘app/data/streamData.xml’
- Instalar todos os packages da pasta app/queries’
- Abrir a pasta ‘baseX/bin’ e correr o ficheiro ‘baseXserver’
- Instalar o package requests e restantes que não estejam já instalados através do PyCharm
- Correr o projeto através do PyCharm
- Conectar-se ao browser através do endereço ‘<http://127.0.0.1:8000/>’

PEDRO GONÇALVES (88859)
RUI OLIVEIRA (89216)
PEDRO SILVA (89228)
NOVEMBRO 2020