

Googol

Motor de pesquisa de páginas web



Universidade de Coimbra
Faculdade de Ciências e Tecnologia
Departamento de Engenharia Informática

Meta 2

Trabalho realizado por:
Ricardo Guegan nº2020211358
Samuel Machado nº2020219391

Introdução	2
Pesquisa	3
Consultar Urls que apontam para um dado Url	3
Indexar Url	4
Integração com serviço REST - HackerNews	4
Indexar histórias de um utilizador	4
Indexar as top stories	5
Testes realizados	6

Introdução

Na segunda meta do projeto, tivemos como objetivo criar uma nova forma de interagir com o sistema desenvolvido na primeira meta, a aplicação Googol. Trata-se de uma WebApplication criada com SpringBoot, possibilitando que os utilizadores interajam com o sistema sem terem de ter instalado o software do cliente, através da Internet, usufruindo ainda de uma interface gráfica, que torna a interação muito mais intuitiva para o cliente em comparação com a linha de comandos desenvolvida na meta 1.

A interface da página inicial ("/"), possibilita o recurso às diferentes funcionalidades.

- Pesquisa
- Pedir Urls que apontam para um dado Url
- Indexar Url
- Indexar histórias de um utilizador - HackerNews
- Indexar as top stories - HackerNews

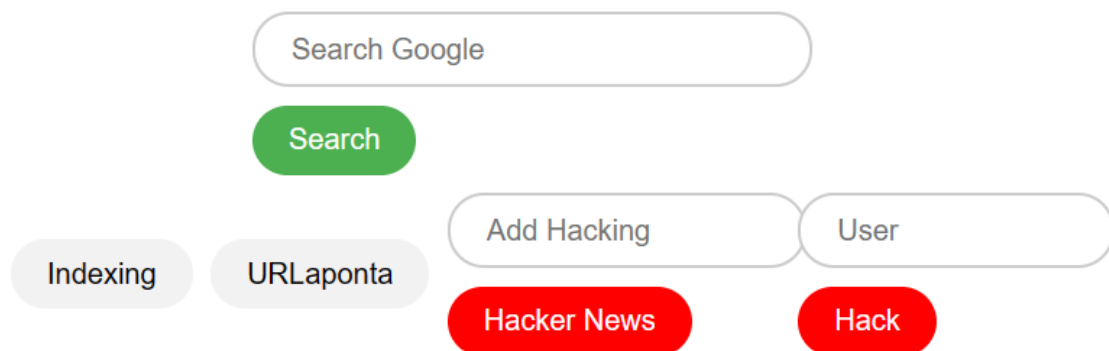
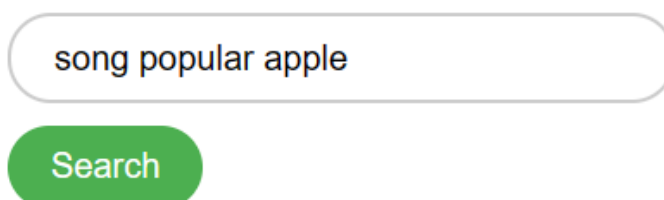


Figura 1 - Interface da página principal

Pesquisa

Ao optar por realizar uma pesquisa, somos redirecionados para uma outra página (“/Pesquisa”) que possibilita a visualização dos resultados, que consistem num conjunto de todos os urls cuja página contém todas as palavras presentes nos termos de pesquisa, sendo que stopwords não são consideradas. Estes resultados são internamente organizados de forma a serem renderizados por ordem de relevância, isto é, urls que tenham mais aparições em outras páginas surgem mais acima na página de resultados.



song popular apple

Search

Figura 2 - Interface com os termos de pesquisa introduzidos



Resultados Pesquisa (10)

<https://workspace.google.com/intl/vi/>

日本語

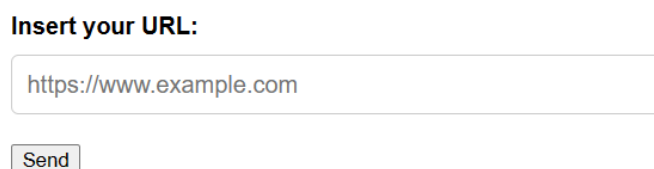
Google Workspace | Các công cụ cộng tác và ứng dụng...

Figura 3 - Resultados da pesquisa apresentados

Consultar Urls que apontam para um dado Url

É possível realizar esta consulta clicando no botão “URLaponta” que nos leva a outra página (“/apontaURL”) onde poderemos fornecer um Url. Assim, após introduzir um Url na caixa de texto e submeter, somos redirecionados para a página (“/ResultadosAponta”), em que podemos encontrar o resultado, todos os Urls cuja página contém o mesmo.

Ver URL que apontam



Insert your URL:

<https://www.example.com>

Send

Figura 4 - Interface para fornecer Url a considerar

URL que apontam

<https://blog.youtube/news-and-events/2022-youtube-creator-economy-impact-report/>

<https://support.google.com/youtube/>

<https://blog.youtube/creator-and-artist-stories/road-to-1-million-marie-stella/>

Figura 5 - Resultados Url que apontam

Indexar Url

O url fornecido pelo utilizador através da interface gráfica é inserido no final da fila de Urls a analisar, de forma a que, brevemente, seja retirado e analisado pelo sistema. Posteriormente, quando este for analisado, todos os Urls presentes na sua página são retirados e inseridos também estes na fila, de forma a alcançar esta recursividade desejada.

Indexar URL

Insert your URL:

Figura 6 - Interface para fornecer Url a indexar

Integração com serviço REST - HackerNews

Para alcançar esta funcionalidade utilizámos uma versão por nós alterada do código disponibilizado da aula PL5.

Indexar histórias de um utilizador

Nesta funcionalidade, o sistema pega em todas as histórias de um dado utilizador e, para todas as que tiverem um Url, este é introduzido na fila de Urls a ser analisados.

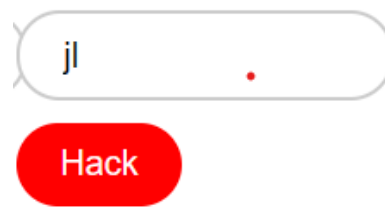


Figura 7 - Interface para fornecer user a considerar

Indexar as top stories

Decidimos ter em conta um top 10 de stories, sendo que o sistema pede ao serviço REST todas as stories disponíveis e vai iterando sobre cada uma até ter recolhido 10 stories que contenham no seu texto os termos fornecidos.

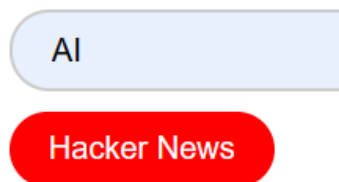


Figura 8 - Interface para fornecer termos a considerar

Testes realizados

Funcionalidade	Input Vazio	Input Normal	Monkey Test
Pesquisa	mensagem: "Preencha este campo"	apresenta resultados da pesquisa	0 resultados
Consultar Urls que apontam para um dado Url	mensagem: "Preencha este campo"	apresenta urls que apontam	0 resultados
Indexar Url	mensagem: "Preencha este campo"	url adicionado com sucesso	Nada é indexado
Indexar stories de um user - HackerNews	mensagem: "Preencha este campo"	indexadas todas as stories do user fornecido	Salvamos tempo de computação no caso de o user não existir, cortando a execução da função
Indexar Top Stories - HackerNews	mensagem: "Preencha este campo"	indexadas no máximo top 10 stories que contém os termos fornecidos	Nada é indexado, porém sem salvar tempo computacional