

## **Ficha de Trabalho nº 4**

### **Extração de dados da Web: Wrappers,**

### **1. Introdução**

Neste ficha de trabalho pretende-se que os **alunos** aprendam a extrair dados da web usando a função `HttpRequest`. No Moodle encontra um ficheiro com a implementação desta função:

- **`HttpRequestFunctions.java`**: para aceder aos dados/source de um determinado site

Existem duas funções, onde a única diferença é o separador usado na codificação do espaço em branco (+ ou \_).

Quando existe uma palavra de pesquisa para a criação do URL a aceder, use a função com duas strings como argumento. Caso não haja palavra de pesquisa, envie a String "" no 2º argumento.

Assim garante que a codificação de todos os caracteres é feita corretamente e é gerado um URL válido que pode ser acedido pela função.

### **2. Exercícios**

#### **2.1. Wrappers: extrair dados da web – Ficha 3**

Copie o ficheiro **`HttpRequestFunctions.java`** para a pasta **`src/ficha3`** do Projeto da Ficha3 desenvolvido nas aulas anteriores. Corrija eventuais erros (nome da *package*).

Elimine o ficheiro **`peessoas.html`** da pasta do projeto. O ficheiro **`peessoas.html`** usado nos wrappers da Ficha 3 passou agora a estar disponibilizado no site: <https://eden.dei.uc.pt/~abs/ID/peessoas.html>

Aos Wrappers implementados na Ficha 3 acrescente a chamada à função **`HttpRequest1`** para que o *source* do ficheiro seja temporariamente armazenado em disco antes da chamada dos *Wrappers*.

Quando gravar o *source* em disco, mantenha o nome do ficheiro antigo (**`peessoas.html`**)

```
String link = "https://eden.dei.uc.pt/~abs/ID/peessoas.html";  
String pesquisa = "";  
HttpRequestFunctions.httpRequest1(link, pesquisa, "peessoas.html");
```

Teste os *wrappers* e verifique se a informação é extraída da Web corretamente.

## 2.2. Wrappers: extrair dados da web, loja Bertrand

No seu Browser, aceda a **bertrand.pt** e faça uma pesquisa usando o **isbn** de um livro, por exemplo, “9789892314044”. Analise o *source* da página web e identifique onde se encontra o título, o autor, o preço do livro devolvido pela pesquisa.

Crie um novo projeto **Ficha4**.

Na **package** *ficha4* crie uma Classe **Livro** com os seguintes campos:

**isbn** - String  
**título** - String  
**autor** - String  
**preço** - double

Acrescente à classe o construtor, getters e setters.

a) Na **package** *ficha4* crie um novo ficheiro de nome **Wrappers.java** e implemente 3 wrappers que, dado um isbn, obtenham do site o título, o autor e o preço do livro. Utilize as ERs corretas.

- **public static String obter\_titulo(String isbn)**
- **public static String obter\_autor(String isbn)**
- **public static double obter\_preco(String isbn)**

Em cada Wrapper deve chamar a função `HttpRequest` e obter o source de cada isbn.:

```
String link = "https://www.bertrand.pt/pesquisa/"  
HttpRequestFunctions.httpRequest1(link, isbn, "bertrand.html");
```

As ERs devem ser aplicadas sobre o ficheiro **bertrand.html**

b) Implemente a função **criaLivro** que chama as funções anteriores e agrega os dados recolhidos num objeto da classe **Livro**

- **public static Livro criaLivro(String isbn)**

c) No *main* teste as funções anteriores.

d) Pesquise os seguintes livros 9789897224607, 9789722129220, 9789722532877, 9789892314044, 9789722533492

Use as quatro funções implementadas anteriormente para obter os dados do site e criar o objeto da Classe **Livro**. Crie um *ArrayList* com os livros pesquisados e imprima-os na consola.

Teste com mais livros, corrija eventuais erros:

9789892348568, 9789899039353, 9789898857071, 9789722055932, 9788416987719