

Rastreando formulários e logins

Trabalhando com Cookies

Rastreando formulários e logins

O que são Cookies?

O protocolo HTTP não mantém um estado ou conexão (stateless), sendo assim, toda interação que o cliente Web fizer com o servidor Web acarretará em uma nova requisição e resposta.

Cada requisição é independente e possui um tempo de vida. O servidor web não identifica se duas requisições vieram de um mesmo navegador, não fazendo nenhum gerenciamento em memória para que mensagens sejam compartilhadas entre requisições.

Os cookies e as sessões são utilizados para suprir esta necessidade.

Rastreando formulários e logins

O que são Cookies?

Atualmente, em muitos sites existentes na Internet, as configurações principais, como idioma e modo de exibição, são mantidas no navegador, mesmo depois de um tempo sem uso. Por exemplo, caso escolha o idioma Russo na pesquisa do Google, ele será usado até você mudar de ideia.

Para que isso seja possível, o navegador utiliza arquivos de textos chamados Cookies, os quais possuem como principal função armazenar as preferências dos usuários sobre um determinado site na Internet. Cada Cookie em seu PC armazena dados para um endereço web específico.

Basicamente, um Cookie é um arquivo de texto muito simples, cuja composição depende diretamente do conteúdo do endereço Web visitado. Por exemplo, a maioria dos sites armazenam informações básicas, como endereços IP e preferências sobre idiomas, cores, etc. Contudo, em portais como o Gmail e o Hotmail, nomes de usuários e senhas de email também fazem parte dos Cookies.

Rastreando formulários e logins

O que são Cookies?

Quando você visita um site pela primeira vez, este envia um Cookie como resposta para o seu navegador, contendo as suas preferências, em formato de texto. Este pequeno arquivo ficará armazenado em seu computador até que perca sua validade.

Enquanto o cookie estiver salvo em seu PC, toda vez que você digitar o endereço do site, o seu navegador irá enviar este arquivo para o site que você está conectado. Desta maneira, as suas configurações serão aplicadas de maneira automática.

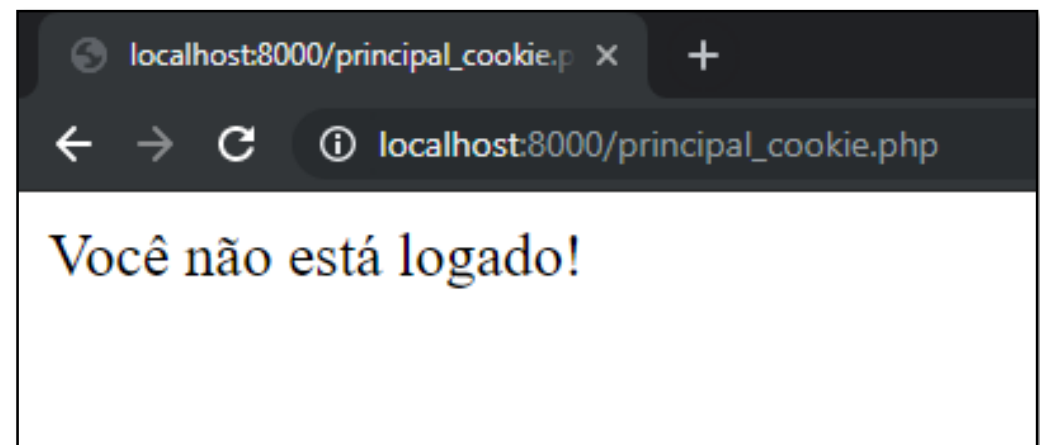
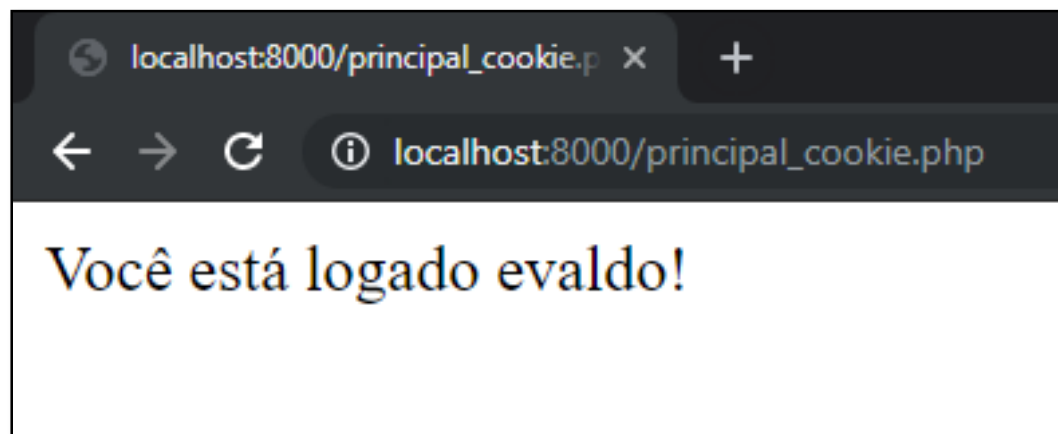
Fonte: <https://www.tecmundo.com.br/web/1069-o-que-sao-cookies-.htm>

Rastreando formulários e logins

Trabalhando com Cookies

O nosso ambiente para esta aula será o seguinte:

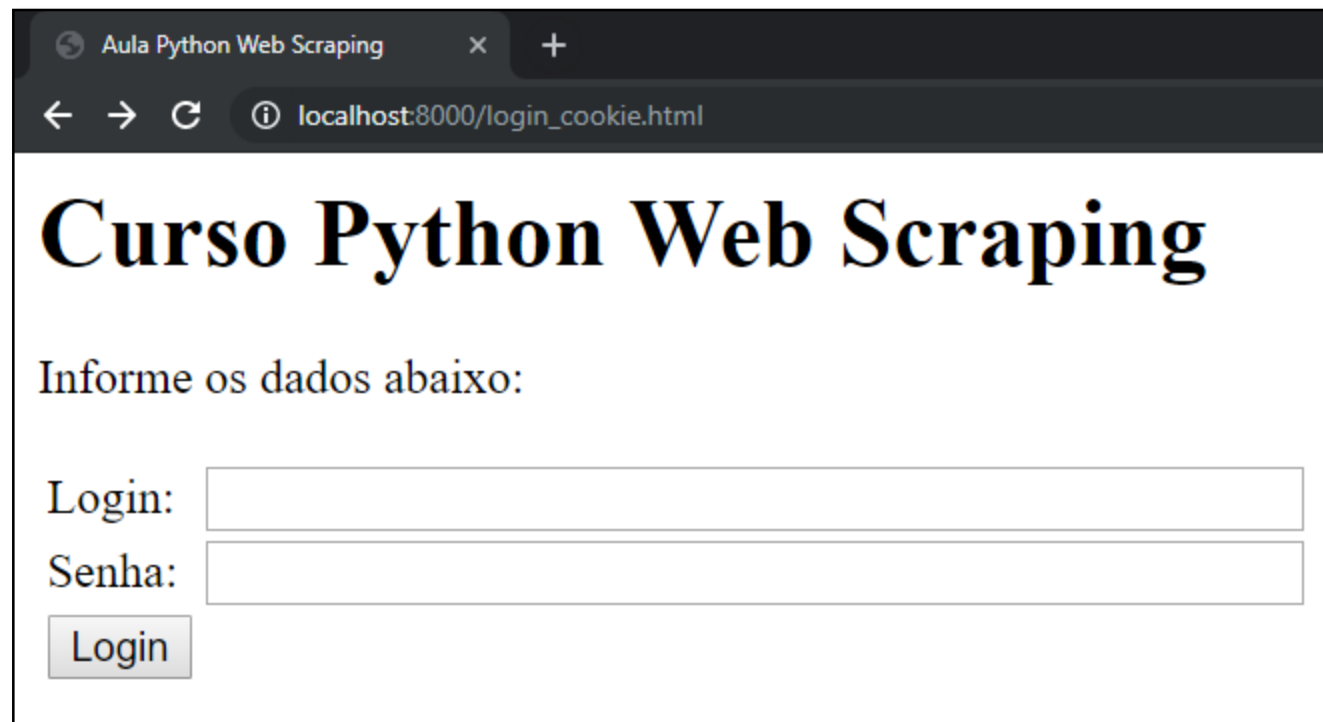
Uma página PHP que valida um cookie para identificar o nome do usuário logado, simulando uma página principal de um sistema que é exibida após o login do usuário. O nome do arquivo desta página é `principal_cookie.php`.



Rastreando formulários e logins

Trabalhando com Cookies

Uma página HTML contendo um formulário para realização do login. O nome do arquivo é login_cookie.html.



A screenshot of a web browser window. The tab is titled 'Aula Python Web Scraping'. The address bar shows 'localhost:8000/login_cookie.html'. The page content features a large heading 'Curso Python Web Scraping' in a bold, black, serif font. Below the heading, the text 'Informe os dados abaixo:' is displayed. There are two input fields: the first is labeled 'Login:' and the second is labeled 'Senha:'. Below these fields is a button labeled 'Login'.

Rastreando formulários e logins

Trabalhando com Cookies

Uma página PHP contendo a verificação do usuário e senha.

Caso o usuário e/ou a senha estejam errados, será exibida na página a mensagem “Login ou senha inválidos”.

Caso o usuário e a senha estejam corretos, será criado um cookie e redirecionado para a página principal_cookie.php.

O nome desta página é login_cookie.php.

Rastreando formulários e logins

Trabalhando com Cookies

Por fim, um arquivo `login_cookie.py` que será nosso script para acessar a página `principal_cookie.php`, enviando o cookie resgatado da página `login_cookie.php`

Rastreado formulários e logins

Trabalhando com Cookies

principal_cookie.php

```
<?php
    if (isset($_COOKIE['cookie_pythonwebscraping_exemplo'])) {
        $login = $_COOKIE['cookie_pythonwebscraping_exemplo'];
        setcookie("cookie_pythonwebscraping_exemplo","",time()-1);
        echo "Você está logado " . $login . "!";
    }else{
        echo "Você não está logado!";
    }
?>
```

Rastreando formulários e logins

Trabalhando com Cookies

login_cookie.html

```
<html>
  <head>
    <title>Aula Python Web Scraping</title>
  </head>
  <body>
    <h1>Curso Python Web Scraping</h1>

    <form action="login_cookie.php" method="POST">
      Informe os dados abaixo:<br><br>
      <table>
        <tr>
          <td>Login:</td>
          <td><input type="text" name="login" size="50" maxlength="100"></input></td>
        </tr>
        <tr>
          <td>Senha:</td>
          <td><input type="password" name="senha" size="50" maxlength="100"></input></td>
        </tr>
        <tr>
          <td><input type="submit" name="enviar" value="Login"></td>
        </tr>
      </table>
    </form>
  </body>
</html>
```

Rastreado formulários e logins

Trabalhando com Cookies

login_cookie.php

```
<?php
    $login = $_POST['login'];
    $senha = $_POST['senha'];
    if ($login == "evaldo" and $senha == "1234"){
        setcookie("cookie_pythonwebscraping_exemplo", $login);
        header('Location: principal_cookie.php');
        exit();
    }else{
        echo "Login ou senha inválidos!!!";
    }
?>
```

Rastreando formulários e logins

Trabalhando com Cookies

login_cookie.py

```
# Simulando o login_cookie.html
import requests

user_agent = "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_3) " \
             "AppleWebKit/537.36 (KHTML, like Gecko)" \
             "Chrome/35.0.1916.47 Safari/537.36"

params = {'login': 'evaldo', 'senha': '1234'}
#r = requests.post("http://localhost:8000/login_cookie.php", data=params)
r = requests.post("http://www.pythonparatodos.com.br/secao14aula02cookie/"
                 "login_cookie.php", headers={'User-Agent': user_agent}, data=params)
#r = requests.post("http://localhost:8000/principal_cookie.php", cookies=r.cookies)
r = requests.post("http://www.pythonparatodos.com.br/secao14aula02cookie/"
                 "principal_cookie.php", headers={'User-Agent': user_agent}, cookies=r.cookies)
print(r.text)
```

FIM