Introduction to HTML

José Faria, September 2022

Content

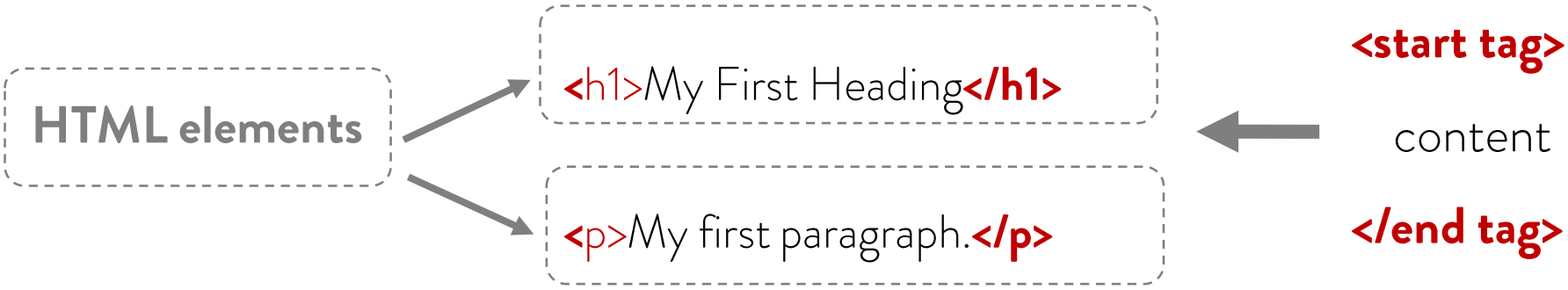
1. HTML basics
2. Publishing HTML pages in a web server
3. Complementary tips

1. HTML basics

Introduction

- **HTML** stands for Hyper Text Markup Language and is the **standard language for creating Web pages**.

- An HTML document consists of a **series of elements** that specify the content and style to be displayed, for example:

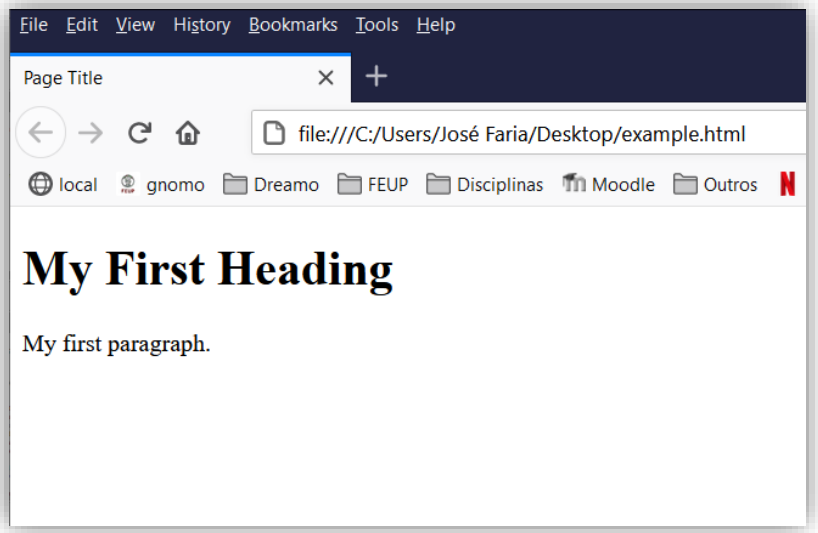


- An HTML document is a plain text file containing HTML code.
- Browsers "read" the HTML code and display it the user interface, for example:

HTML doc [text file containing HTML code]

```
<html>

<head>
<title>Page Title</title>
</head>

<body>

<h1>My First Heading</h1>
<p>My first paragraph.</p>

</body>

</html>
```

Display of the doc in a browser



HTML document (code)

```
<html>
<head>
<title>Page Title</title>
</head>

<body>

<h1>My First Heading</h1>

<p>My first paragraph.</p>

</body>
</html>
```

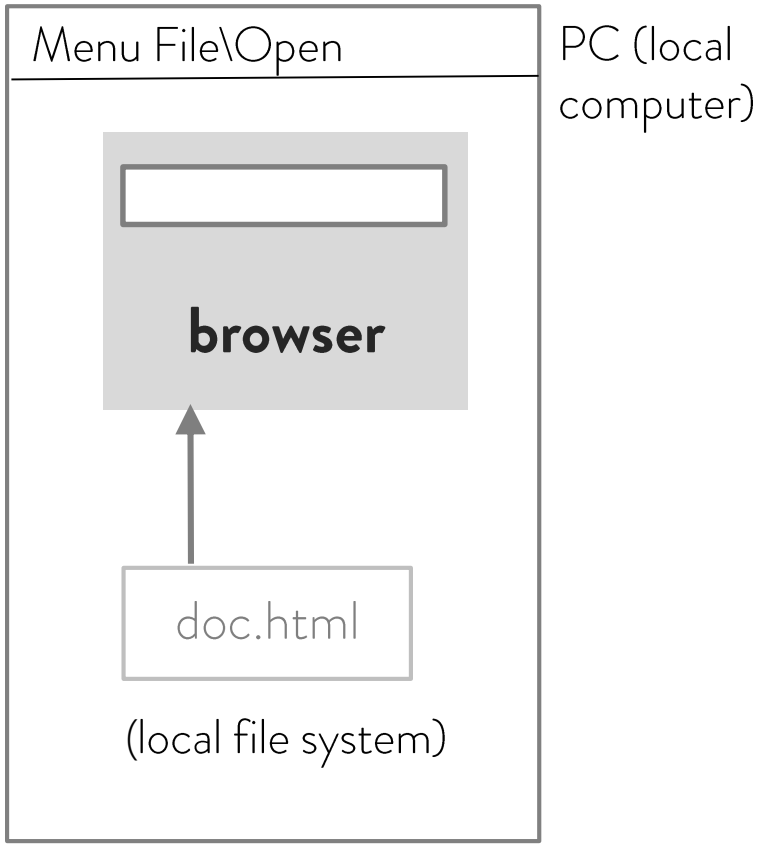
Display of the doc in a browser

HTML document (code)

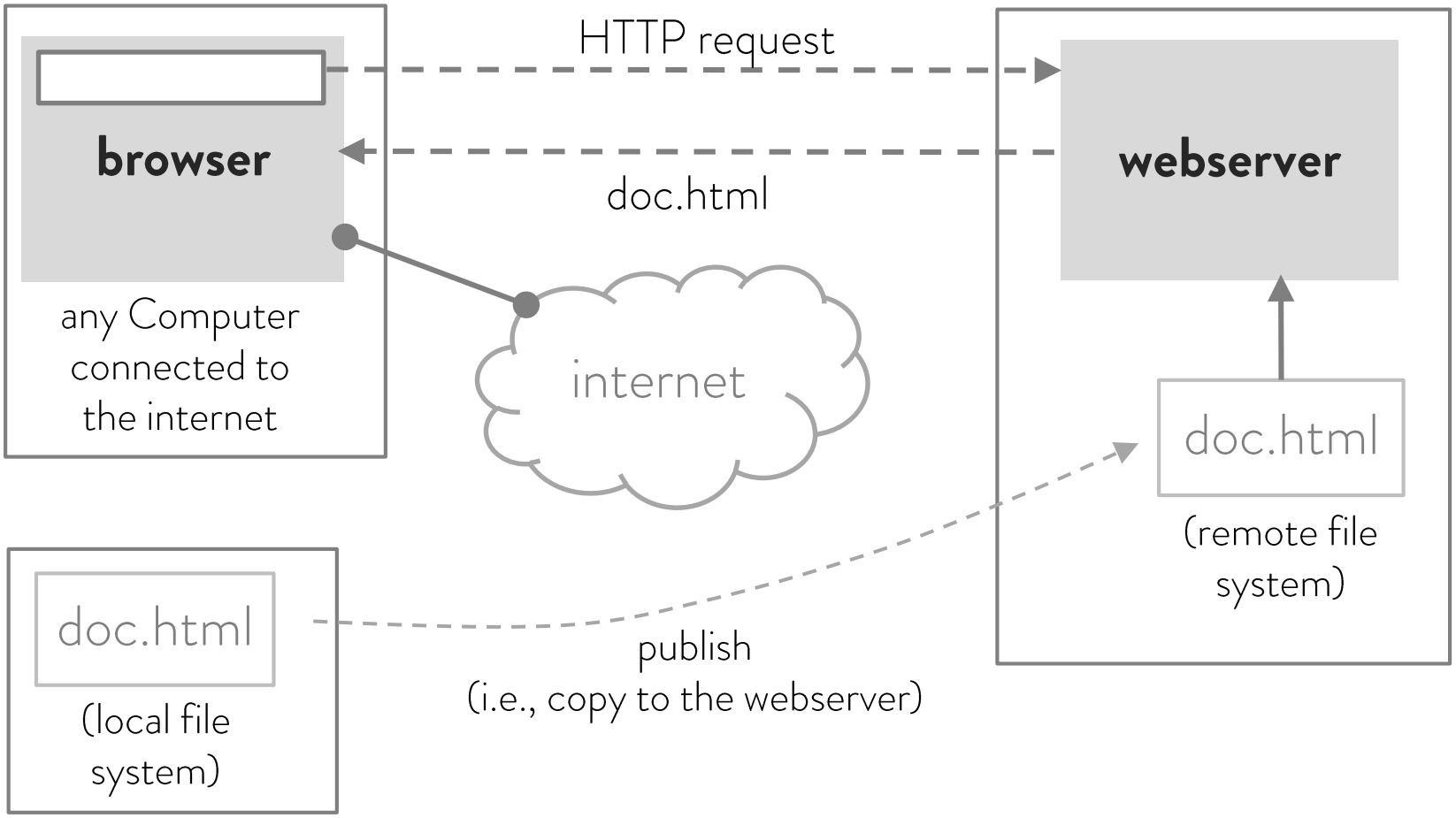
```
<html>
<head>
<title>Page Title</title>
</head>

<body>

<h1>My First Heading</h1>

<p>My first paragraph.</p>

</body>
</html>
```

Display of the doc in a browser

- We can have the html document in our **local computer** and display it in a browser with **File\Open**.



- If we **publish** the document in a **webserver** it will be **available in the internet** through the HTTP protocol.



- Later, we'll see in detail how you'll publish your documents in the web.
- By now, we'll introduce the **basics HTML elements** through the following **example**:

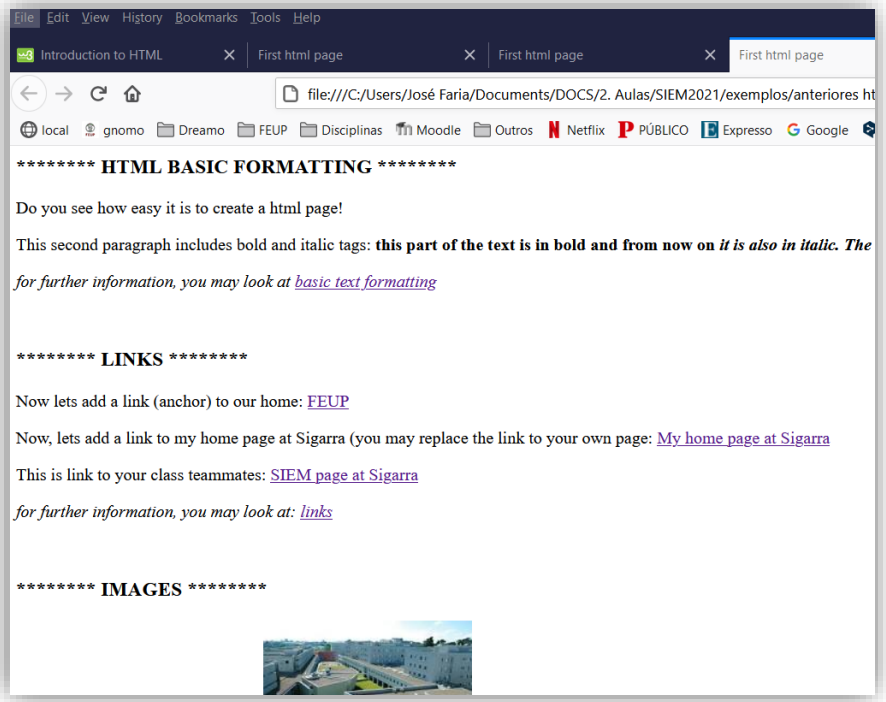
1. **download basicHTMLexample-v1.zip** from Moodle
2. unzip it and **open HTMLbasics.html** in a browser to display its content.

HTMLbasics.html example

The HTML code:

```
<html>
<head>
 <title>First html page</title>
</head>
<body>


 <h3>******** HTML BASIC FORMATTING ********</h3>

 <p>Do you see how easy it is to create a html page?</p>

<p>This second paragraph includes bold and italic tags: <b> this part of the text is in b
The italic part ends here </i> and the bold one here</b>. From now on the text continues

 <p><i>for further information, you may look at <a href="https://www.w3schools.com/html/ht
 <br>


 <h3>******** LINKS ********</h3>

 <p>Now lets add a link (anchor) to our home: <a href="http://www.fe.up.pt">FEUP</a></p>

<p>Now, lets add a link to my home page at Sigarra (you may replace the link to your own

 <a href="https://sigarra.up.pt/feup/pt/func_geral.formview?p_codigo=209496">My home page

 <p>This is link to your class teammates: <a href="https://sigarra.up.pt/feup/pt/ucurr_ger

 <p><i>for further information, you may look at: <a href="https://www.w3schools.com/html/b
 <br>


 <h3>******** IMAGES ********</h3>

 <p>Now lets add an image to our page: <img src="pictures\campusfeup.jpg"> </img></p>
 <p>and an image containing a link: <a href="http://www.fe.up.pt"><img src="pictures\campu

 <p><i>for further information, you may look at: <a href="https://www.w3schools.com/html/b
 <br>
```

The display in a browser:



Suggested HTML editors

Notepad++

ATOM

Sublime Text

Examples of HTML elements

HTML element	Exampe of code
paragraph	Do you see how easy it is to create a html page?
link	FEUP
image	<img src="pictures\campusfeup.jpg"></img>
list	 - item 1 - item 2 - item 3

Lists

unordered list

```
<ul type="disk">
    <li>text1</li>
    <li>text2</li>
    <li>text3</li>
</ul>
```

displayed as:

- text1
- text2
- text3

ordered list

```
<ol>
    <li>text1</li>
    <li>text2</li>
    <li>text3</li>
</ol>
```

displayed as:

1. text1
2. text2
3. text3

For the "bullet" there are 3 options:

- disc <ul type="disk">
- circle <ul type="circle">
- square <ul type="square">

Example of table element

HTML element	Code example
table	<table> <tr> <th>City</th><th>Country</th><th>Population</th> </tr> <tr> <td>Lisbon</td><td>Portugal</td><td>1.2 million</td> </tr> <tr> <td>Porto</td><td>Portugal</td><td>0.4 million</td> </tr> ... </table>

City	Country	Population
Lisbon	Portugal	1.2 million
Porto	Portugal	0.4 million
Paris	France	3.2 million
London	UK	6.2 million

Other examples

element	code

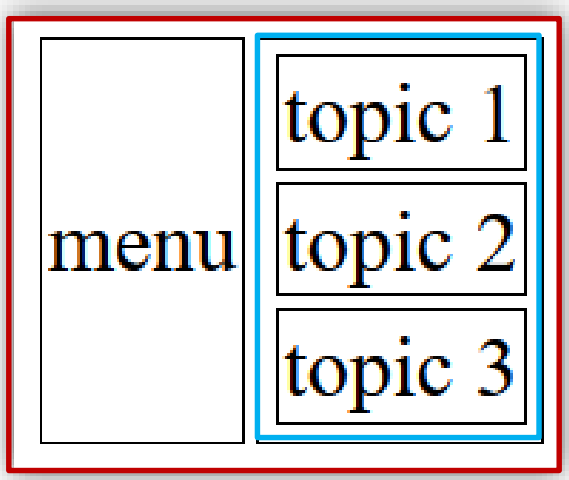
a link in a picture

```
<a href="http://www.fe.up.pt"><img src="pictures\campusfeup.jpg"></img></a>
```

a table inside a table

```
<table>
<tr>
   <td>menu</td>
   <td>
      <table>
         <tr><td>topic 1</td></tr>
         <tr><td>topic 2</td></tr>
         <tr><td>topic 3</td></tr>
      </table>
    </td>
</tr>
</table>
```

in the browser

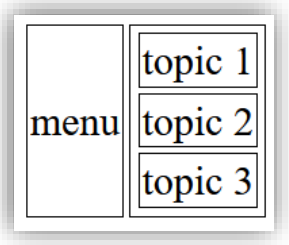


Tables

- Tables are a fundamental element of HTML and can be used for 2 main purposes:
  - present **tabular data**

City	Country	Population
Lisbon	Portugal	1.2 million
Porto	Portugal	0.4 million
Paris	France	3.2 million
London	UK	6.2 million

  - define the main **layout of the page**



- We'll see more on page layouts, later. By now, look HTML tables at https://www.w3schools.com/html/html_tables.asp

Bookmarks

- By default, a link to a page places the cursor at the beginning of the page, but it is **possible to target a specific point** within a page.

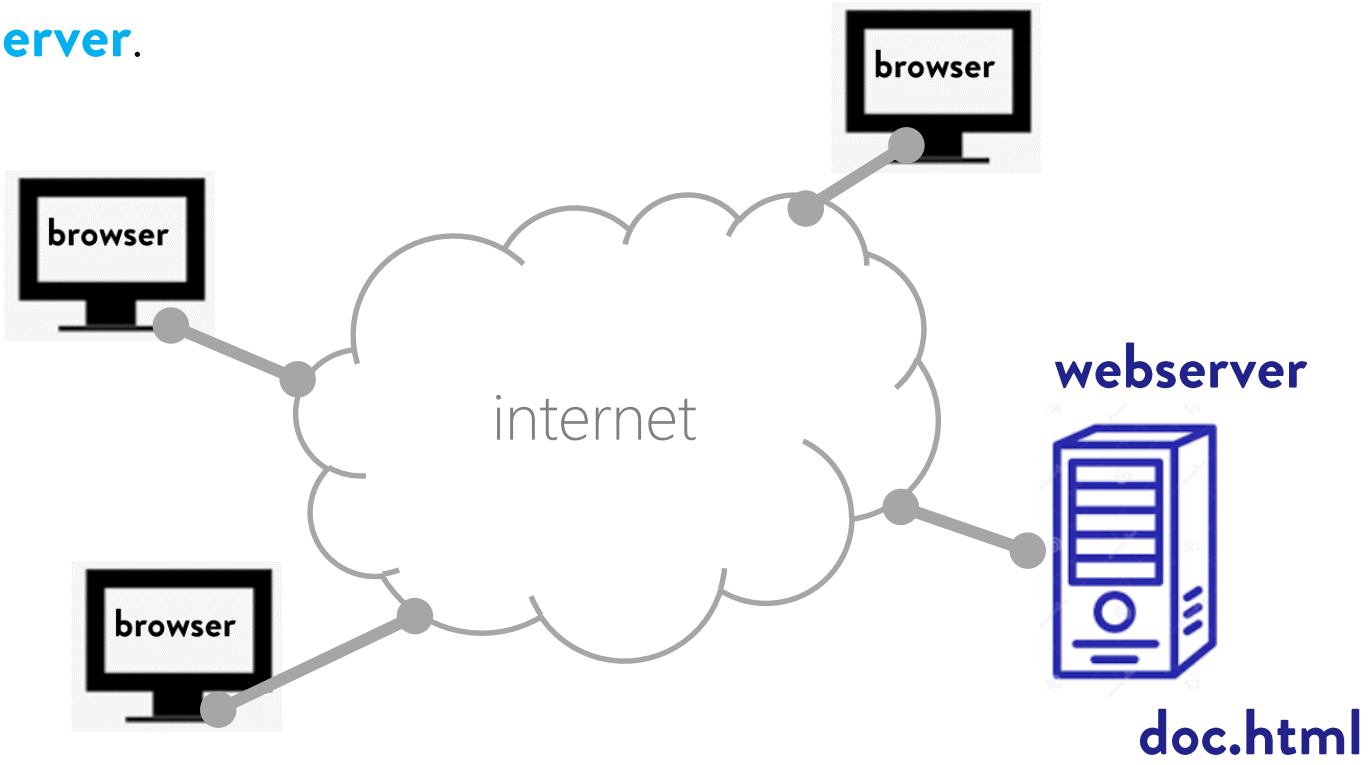
- To do it, a bookmark should be inserted in the target point:

Bookmark example	
Create a bookmark:	<h3 id = “bookmarkDemo”>
url to the beginning of the page:	.../HTMLbasic-v1.html
url to the bookmark:	.../HTMLbasic-v1.html#bookmarkDemo
Link to a bookmark in the same page	<a href="#bookmarkDemo">

2. Publishing HTML pages in a web server

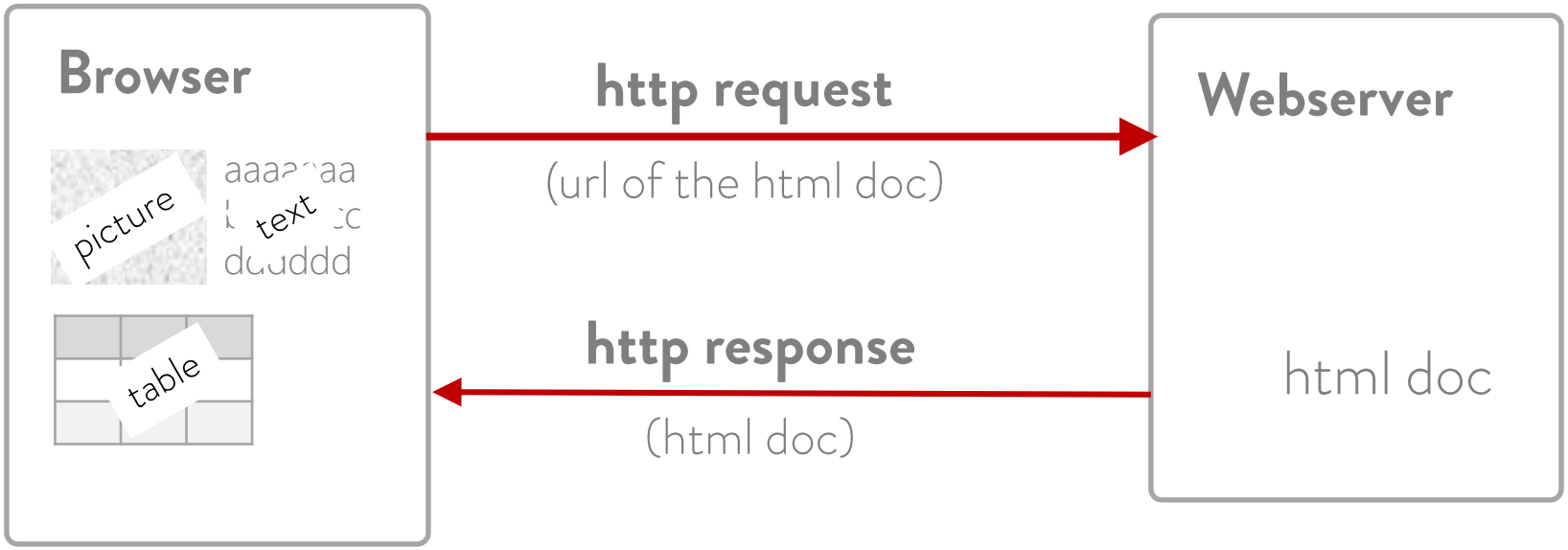
HTTP protocol

- A web **browser** is a software application that displays html documents, normally **stored in a webserver**.

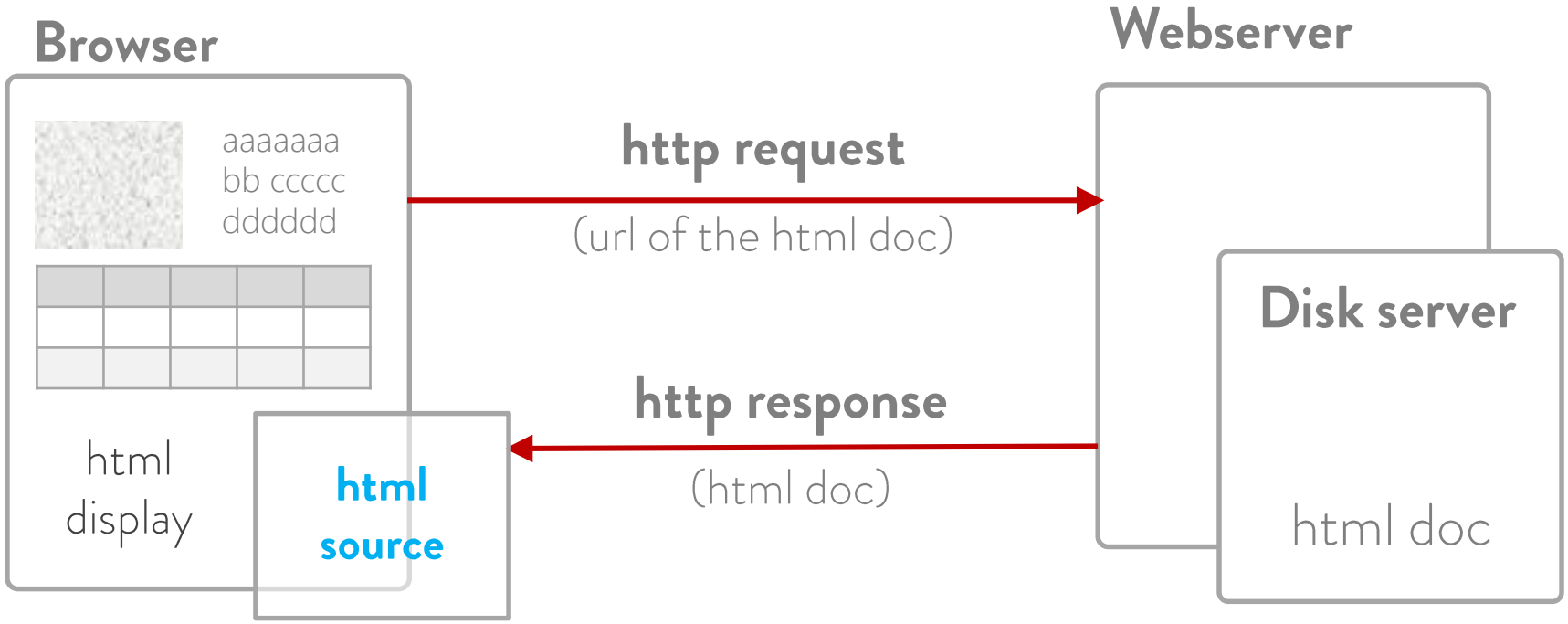


HTTP protocol

- Browsers and webservers interact through the **HTTP protocol**, where HTTP stands for **HyperText Transfer Protocol**.



View page source in the browser



to **view the html doc returned by the webserver** to the browser:

right click within the main window + view page source

HTTP protocol

- Each HTTP interaction proceeds in **4 steps**:
  - the browser **establishes a connection** to the server
  - the browser **sends a request for a document**
  - the server **returns the requested document** to the browser
  - the **connection is closed**.

URL's

- A **Uniform Resource Locator** (URL), currently know as a web address, is a **reference to a web resource** (an HTML document) that specifies its location on a computer network and a mechanism for retrieving it (i.e., a protocol).

- Most commonly, **web pages** referenced in URL's **are accessed through HTTP protocol**, but URL's may also be used with FTP (file transfer) or MAILTO (for accessing messages in an email server), and many other applications.

URL's

- A **typical URL** such as:

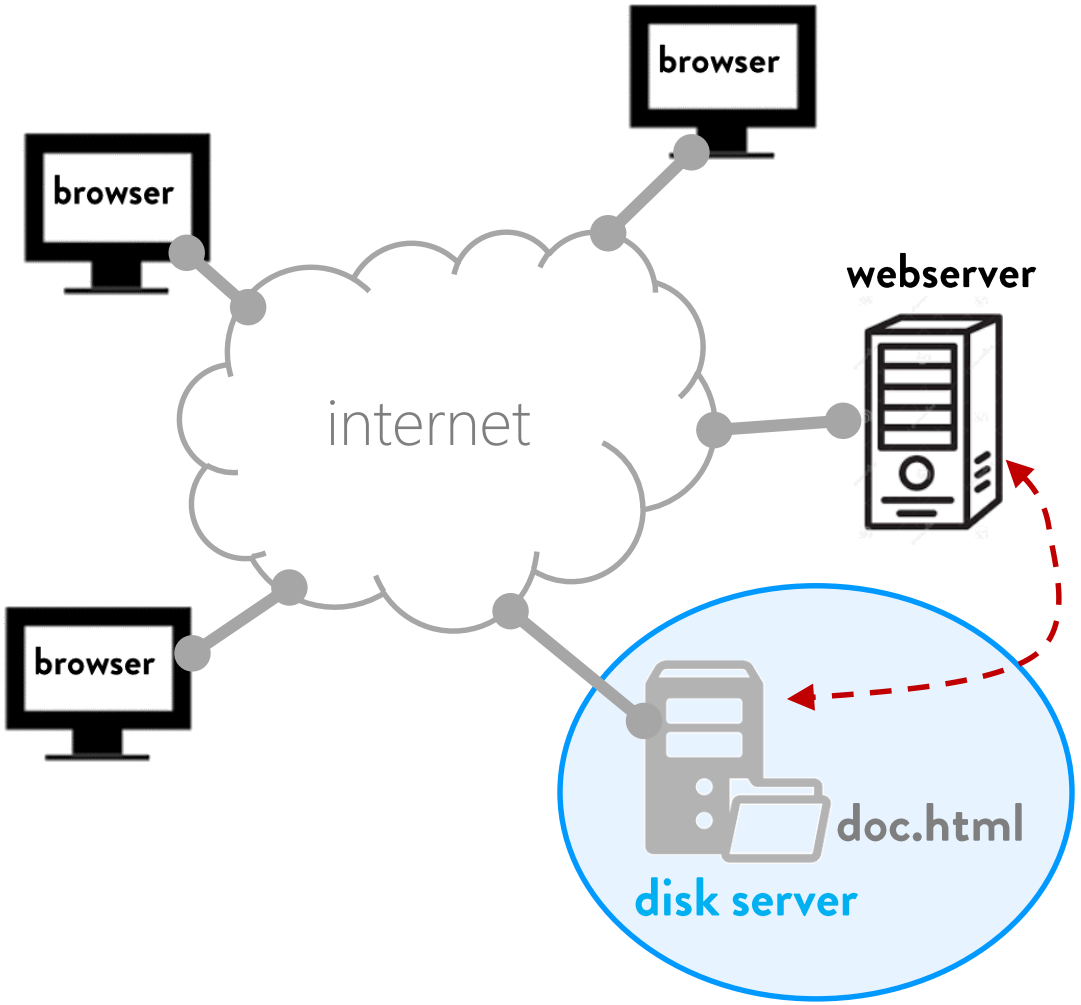
  http://paginas.fe.up.pt/~jfaria/classes/webprog/index.html

- indicates:
  - a **protocol**: http
  - a **webserver** (host name + domain): paginas + fe.up.pt
  - a **path to a folder** in the server: ~jfaria/classes/webprog
  - a **file** located in that folder: example1.html

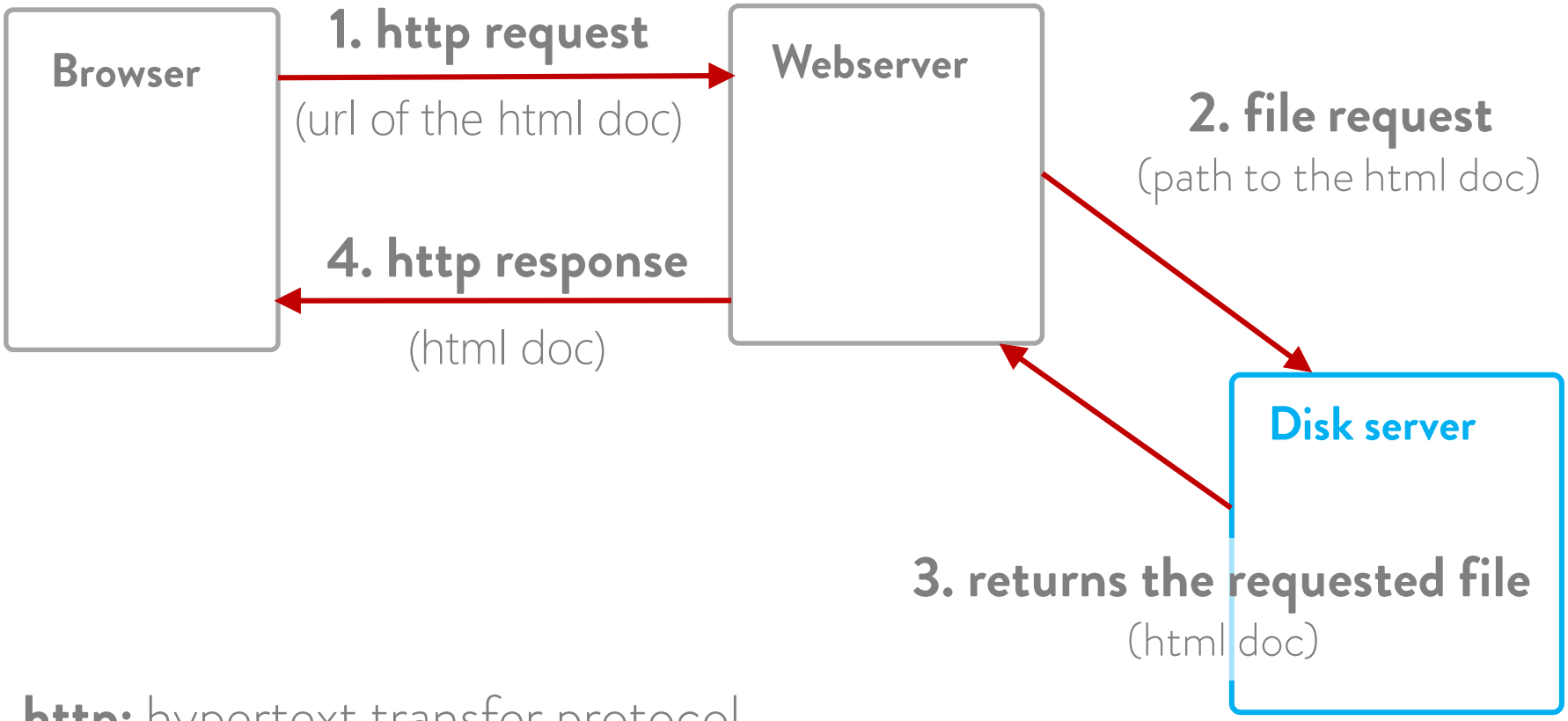
Disk servers

- Normally, the web resources (files) are **stored in a disk server** (or file server, or storage server), other than the webserver.

- This is the case at FEUP as we going to see in a moment.



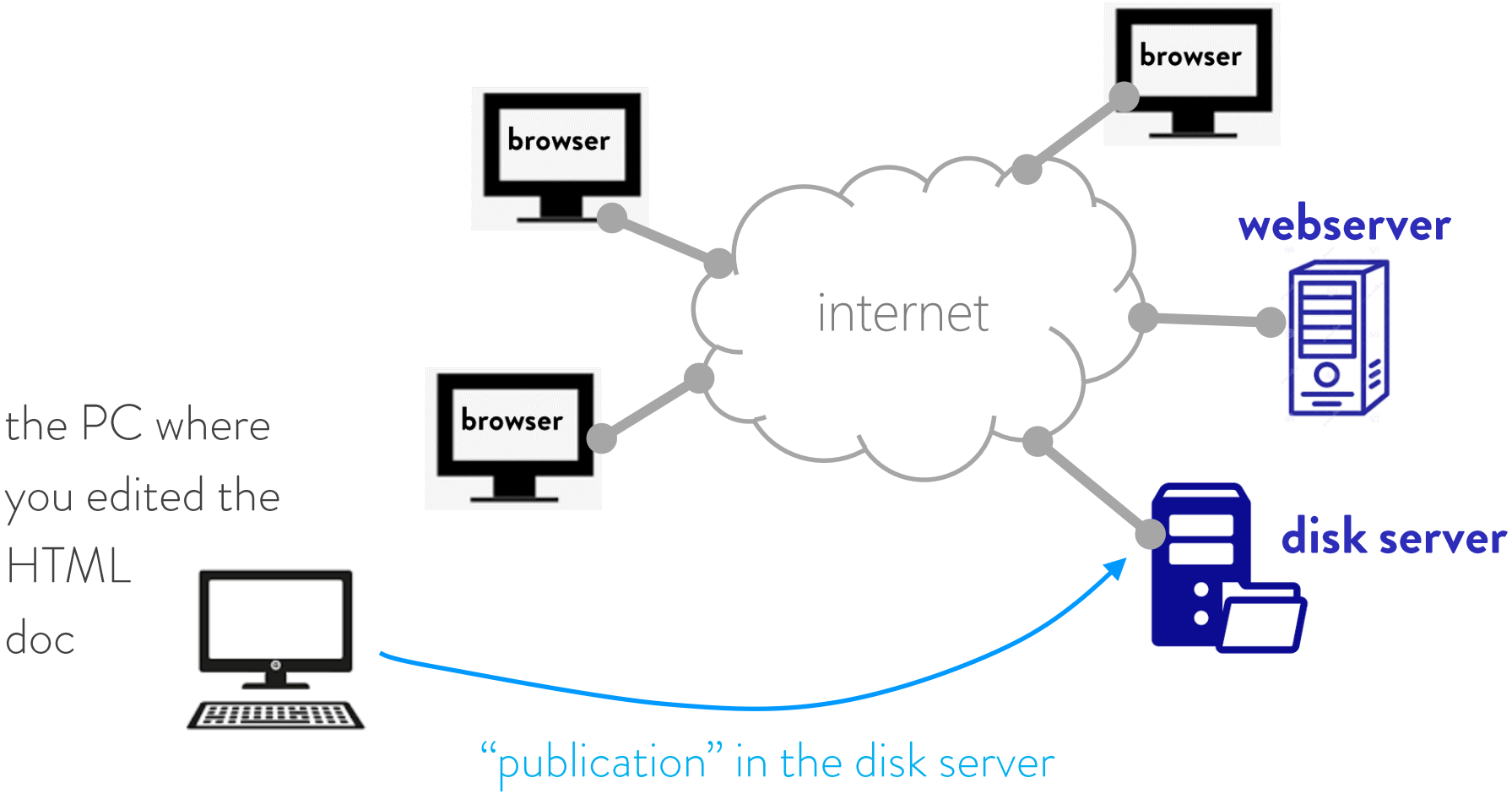
Disk server



http: hypertext transfer protocol

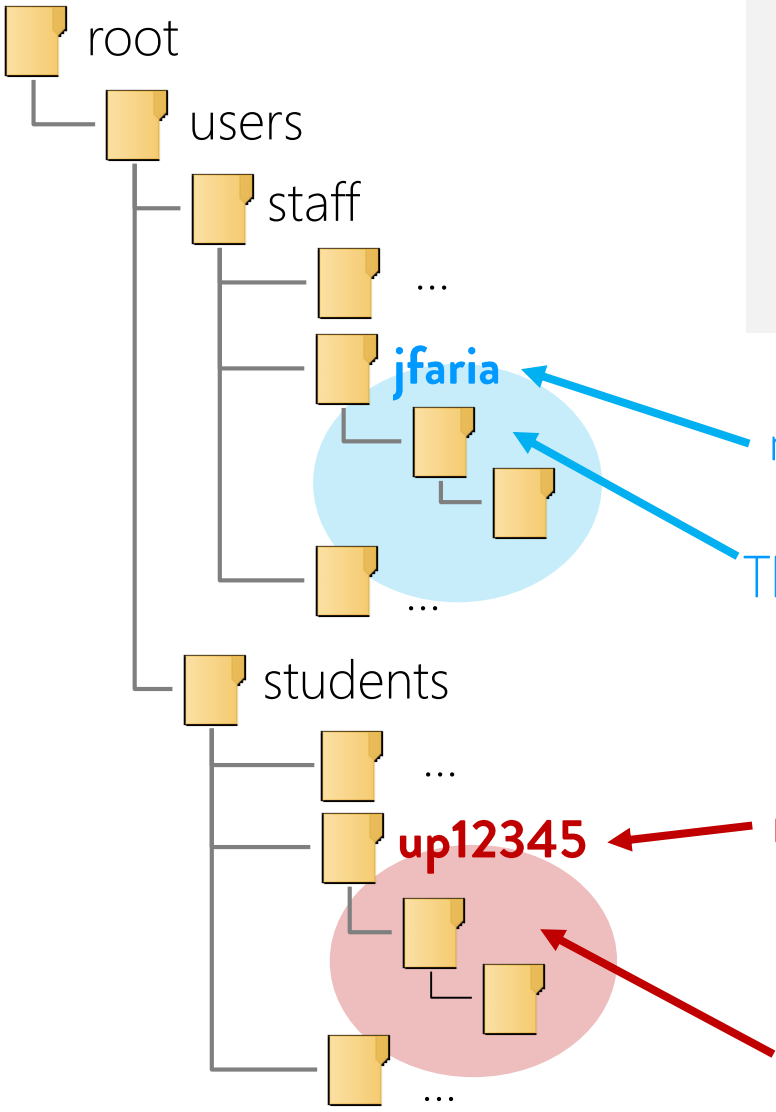
html: hypertext markup language

- So, in order to make **your html pages available in the internet** you have to **store them in the disk server**.



FEUP's disk server

- contains a storage area per Sigarra user (student or staff/teacher)
- the access to this area requires authentication using Sigarra's credentials



root folder of user jfaria (~jfaria)

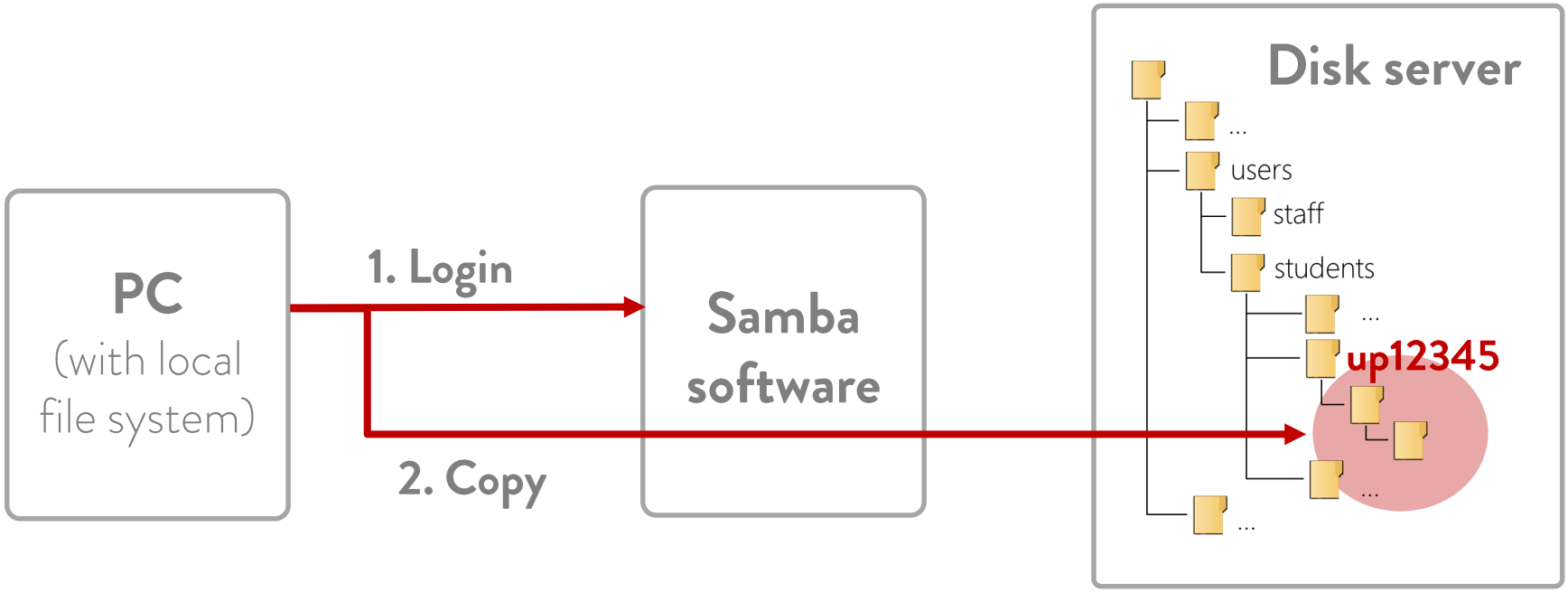
This content can only be managed by jfaria

root folder of student up12345 (~up12345)

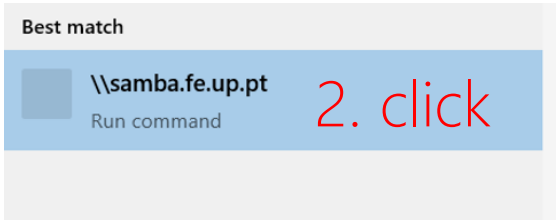
This content can only be managed by up12345

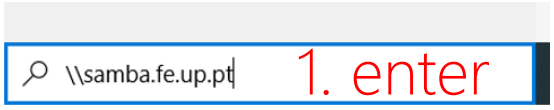
Samba software

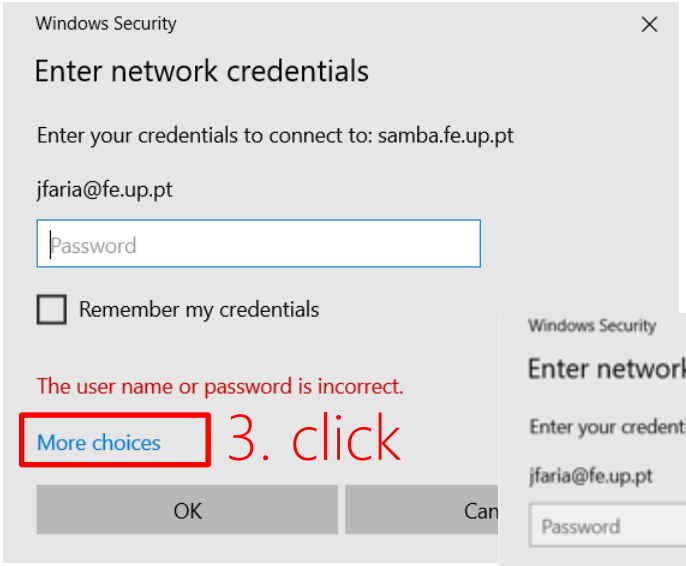
- You can copy your html files to the disk server very easily with **samba software**:

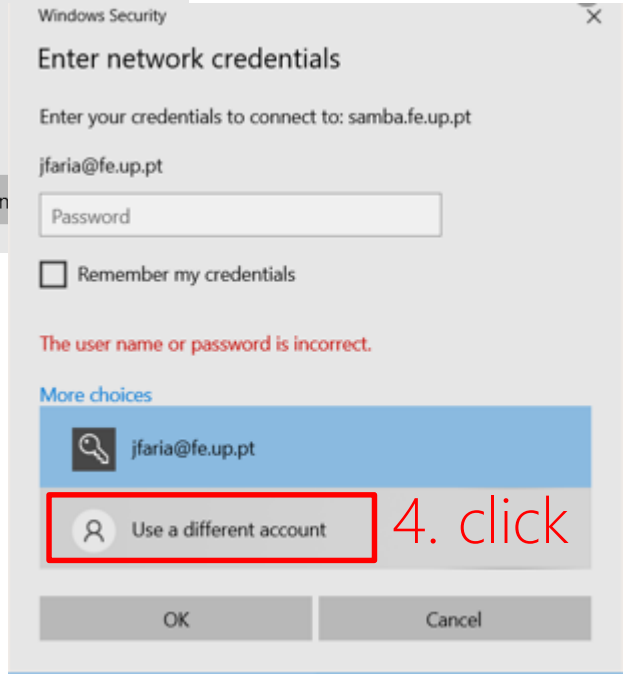


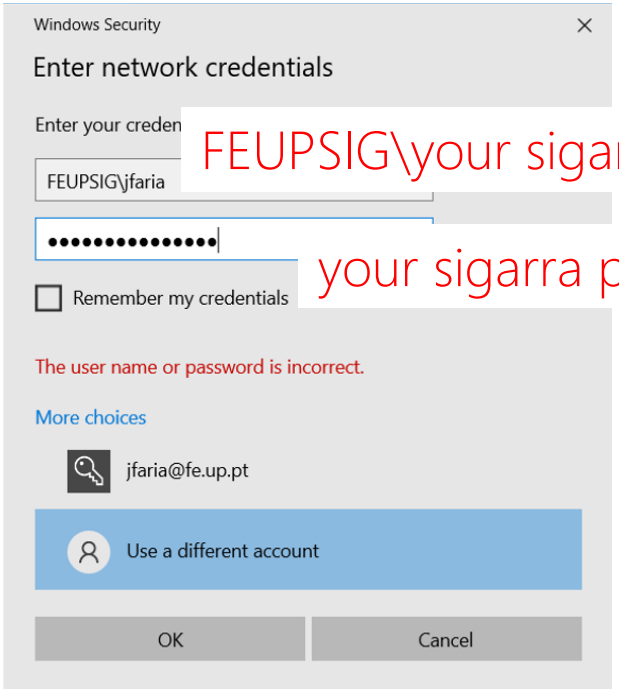
Samba login


Best match
\\samba.fe.up.pt
Run command
2. click


\\samba.fe.up.pt
1. enter


Windows Security
Enter network credentials
Enter your credentials to connect to: samba.fe.up.pt
jfaria@fe.up.pt
Password
Remember my credentials
The user name or password is incorrect.
More choices
3. click
OK
Can

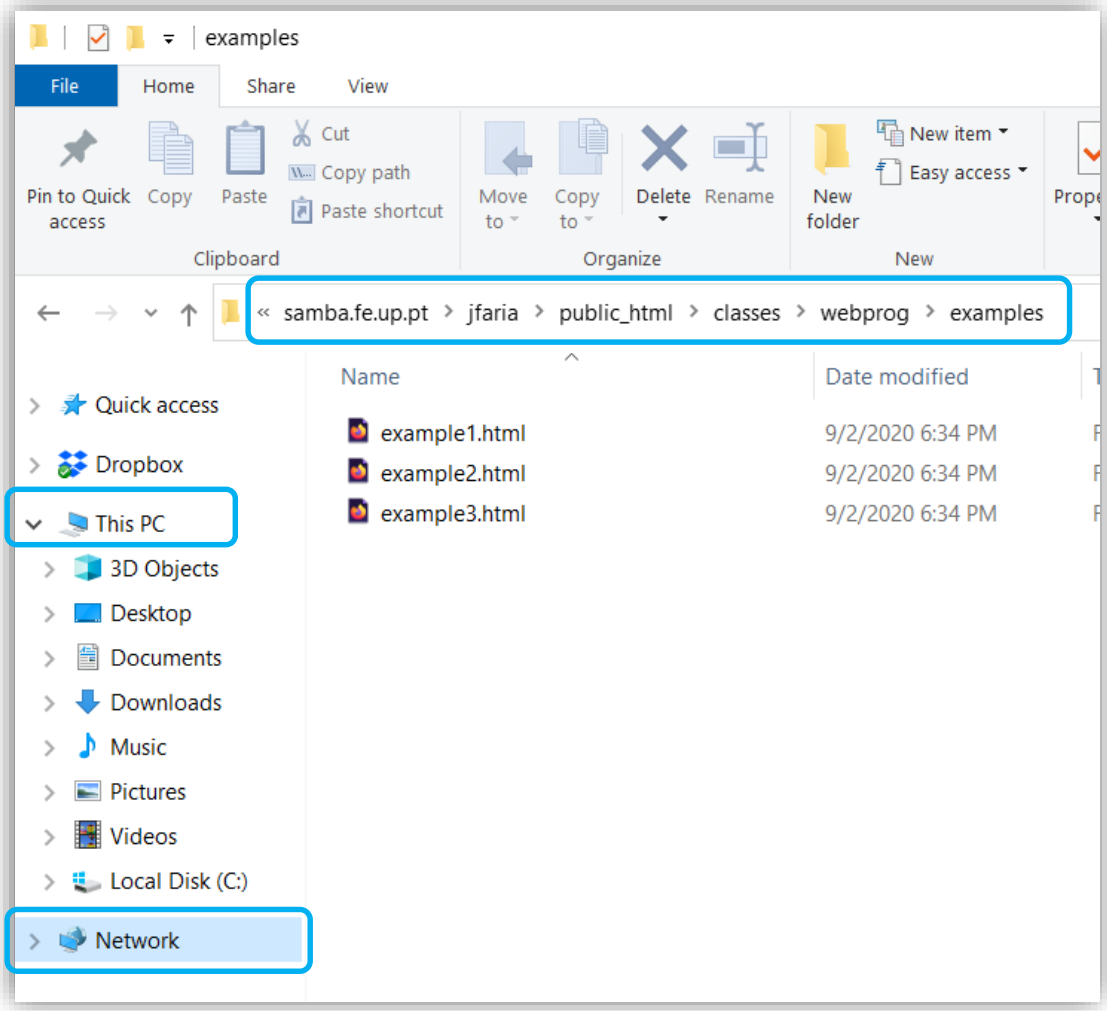

Windows Security
Enter network credentials
Enter your credentials to connect to: samba.fe.up.pt
jfaria@fe.up.pt
Password
Remember my credentials
The user name or password is incorrect.
More choices
jfaria@fe.up.pt
Use a different account
4. click
OK
Cancel


Windows Security
Enter network credentials
Enter your creden
FEUPSIG\jfaria
FEUPSIG\your sigarra username *
your sigarra password
Remember my credentials
The user name or password is incorrect.
More choices
jfaria@fe.up.pt
Use a different account
OK
Cancel

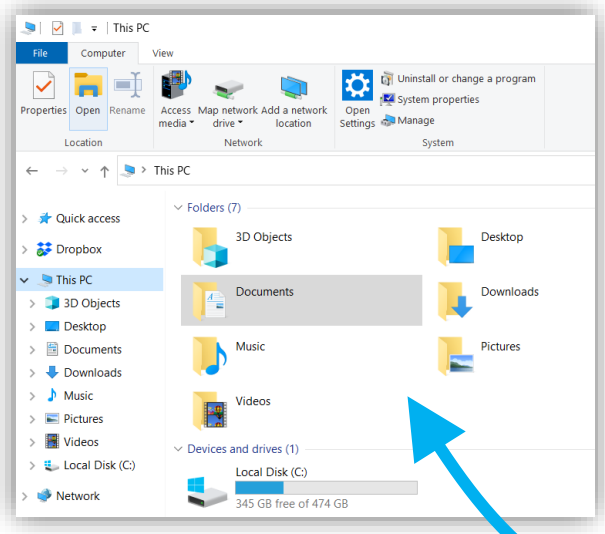
Samba software

- Once logged in samba, **your storage area** in the remote server will be **mounted in the local file system** of your PC.

- This allows you to manage the **remote files as local files**.

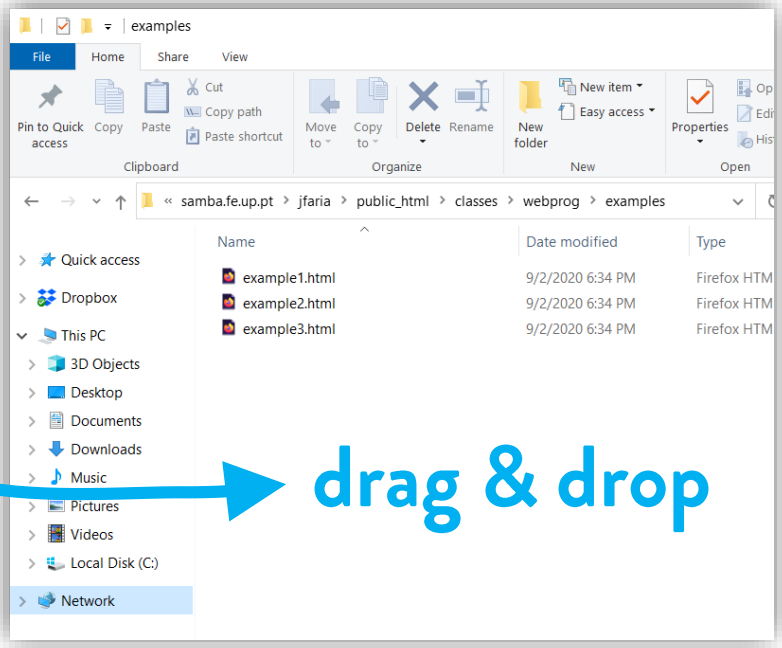


- Therefore, to **publish** a html files you just need to **drag and drop it** (or copy/paste) **from your PC to your storage area** in the disk server.



Local (your PC)

Remote (your storage area)



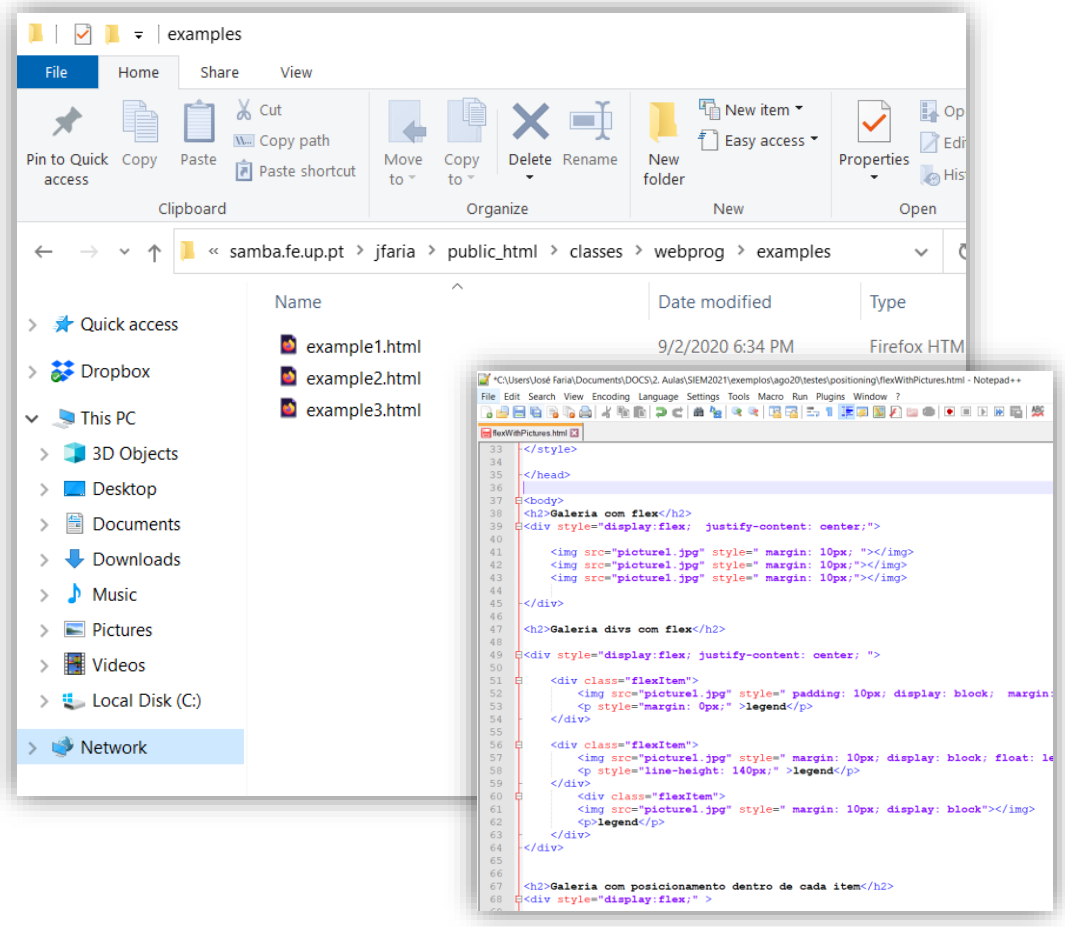
drag & drop

- Samba also allows you to **create and edit the html files directly in the remote server**,

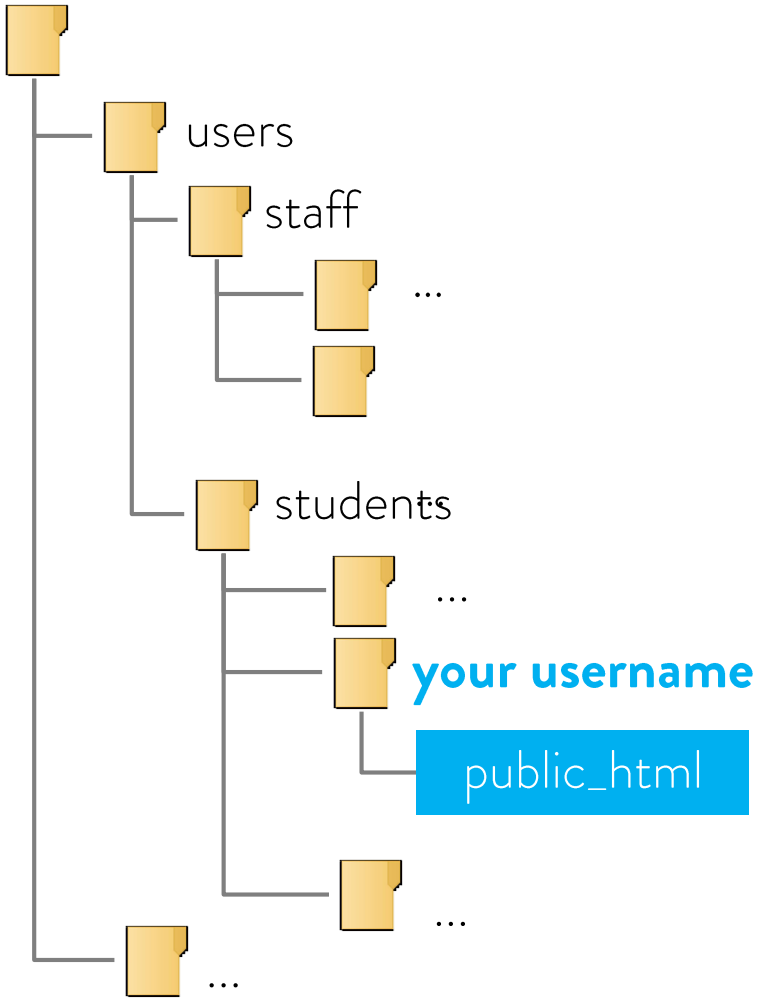
  much like you'd do it

  in your local computer

  …

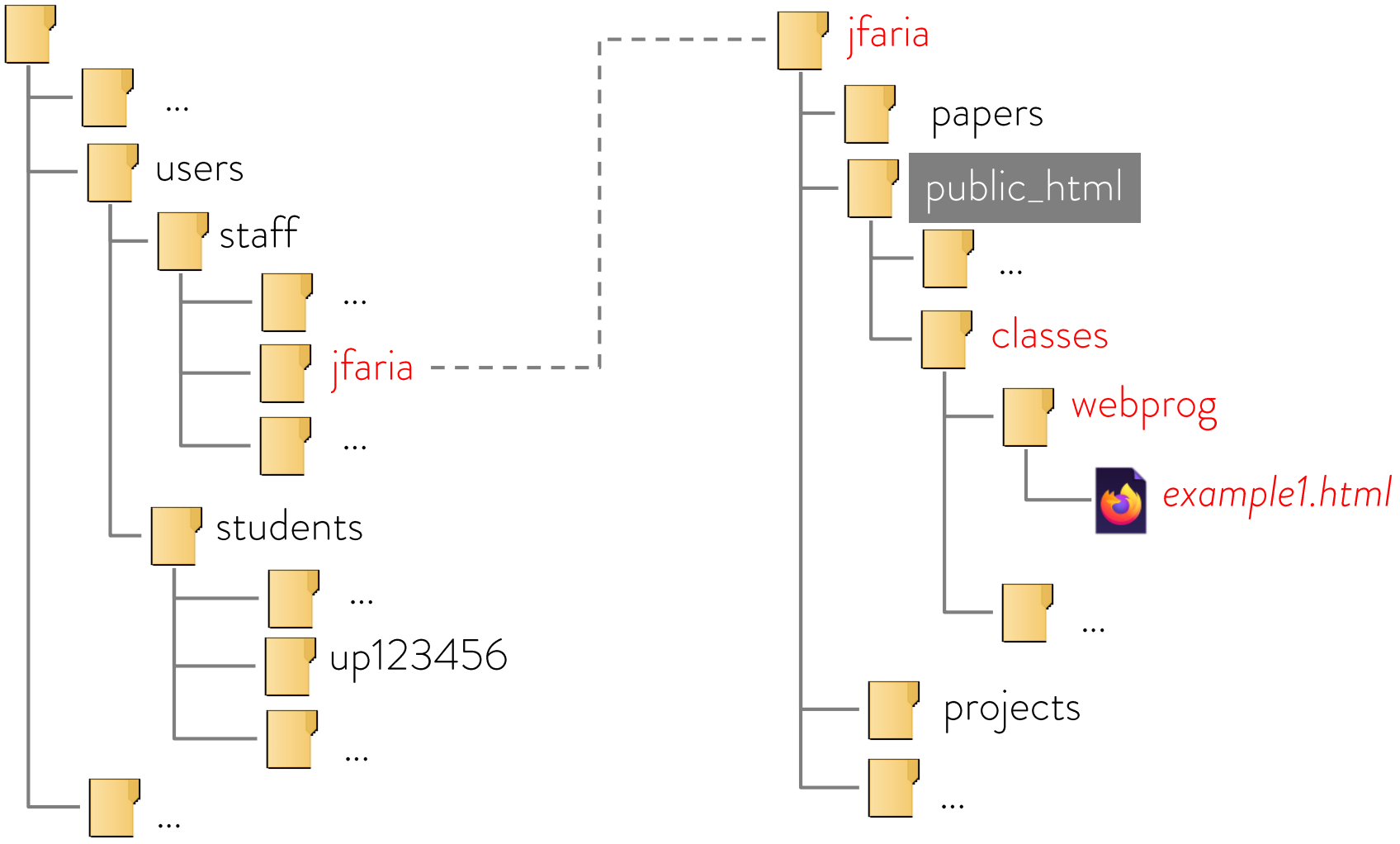


public_html folder



- The **pages** that we want to be made **available in the web** must reside below a folder named **public_html**.
- So, start by **creating the public_html folder** in your storage area.
- Then **copy the html files** to that folder (or a subfolder), for example ...

public_html

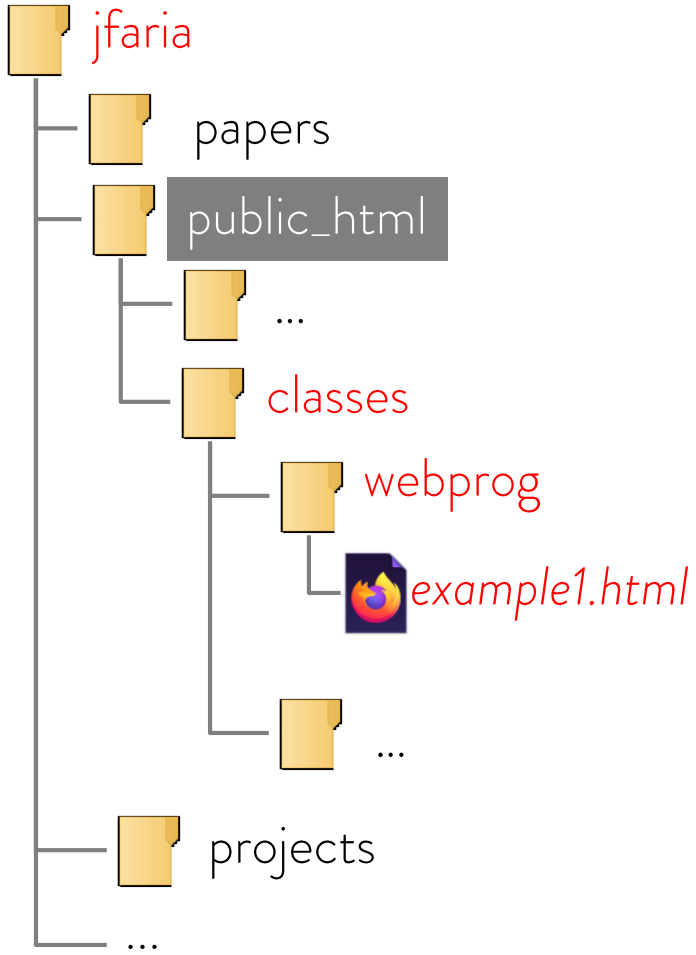


paginas.fe.up.pt /~jfaria /classes/progweb/example1.html

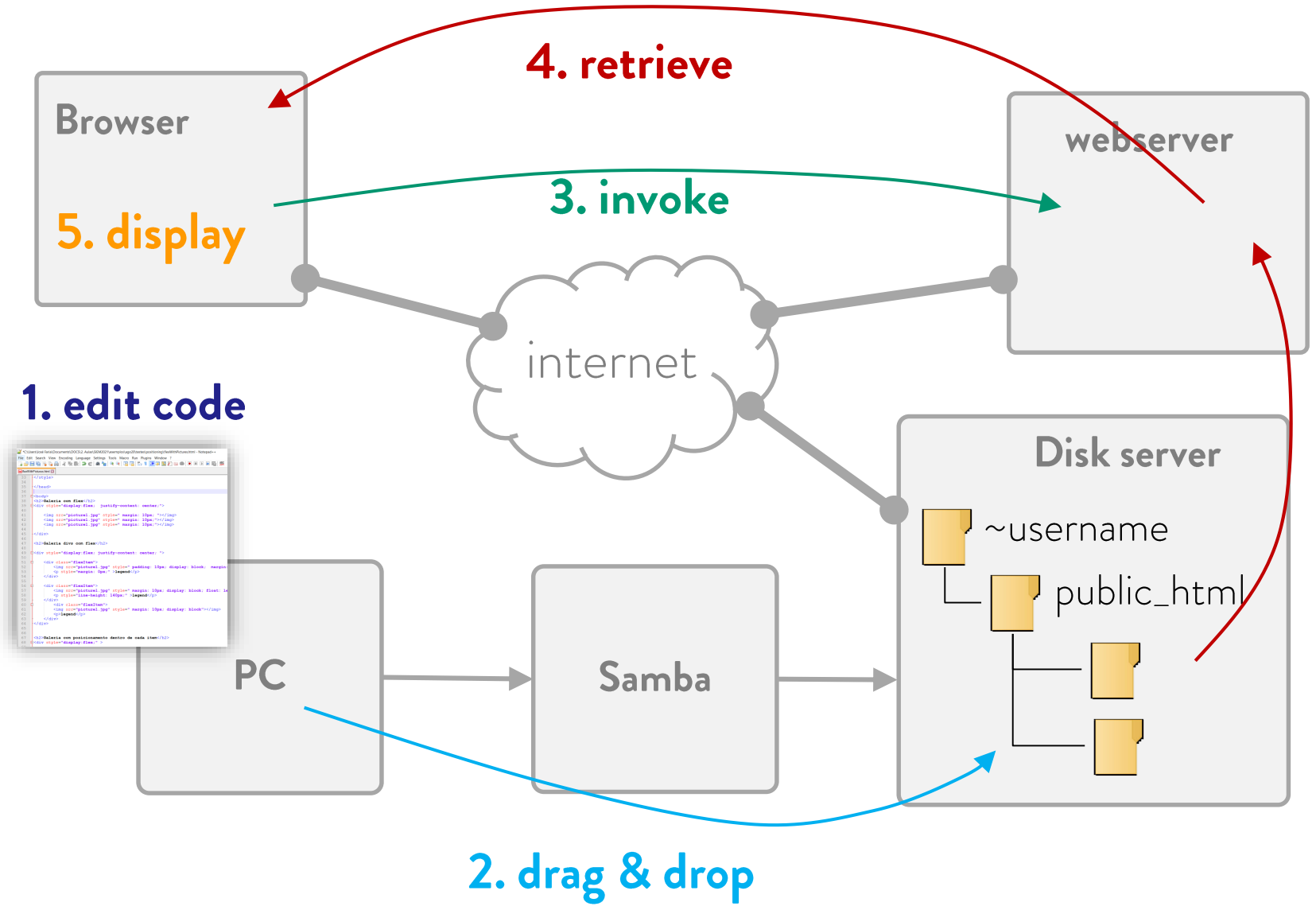
public_html

- Note that, once **public_html** is the **standard root folder** for the web pages,
- it should **not be included in the path** to the HTML file.
- For example, the path to file *example1.html* is:

  ~jfaria/classes/webprog/example1.html



in summary


4. retrieve
Browser
webserver
3. invoke
5. display
internet
1. edit code
Disk server
~username
public_html
PC
Samba
2. drag & drop

3. Complementary tips

- Please, look at the following **tips** as they may help you **avoiding annoying situations** when developing your first web pages

3.1. Case sensitive file names

- Windows is not case sensitive as far as file names are concerned, but Linux is.
- For example, the two file paths:

  class/examples/file.html and **class/Examples/file.HTML**

  are **equivalent in Windows but not in Linux** !

3.1. Case sensitive file names

- So, when you move HTML files from your PC-windows to a Linux-webserver (as *paginas* and *gnomo*) **you'll get broken links** if the links specified in the code and the file names do not have the same capitalization.

- Be attentive to this **annoying and rather common mistake** (every year I hear this from one or more "desperate" students ☺: the page works perfectly in my computer. Why doesn't it work in the webserver as well? For sure there is a problem with the server!)

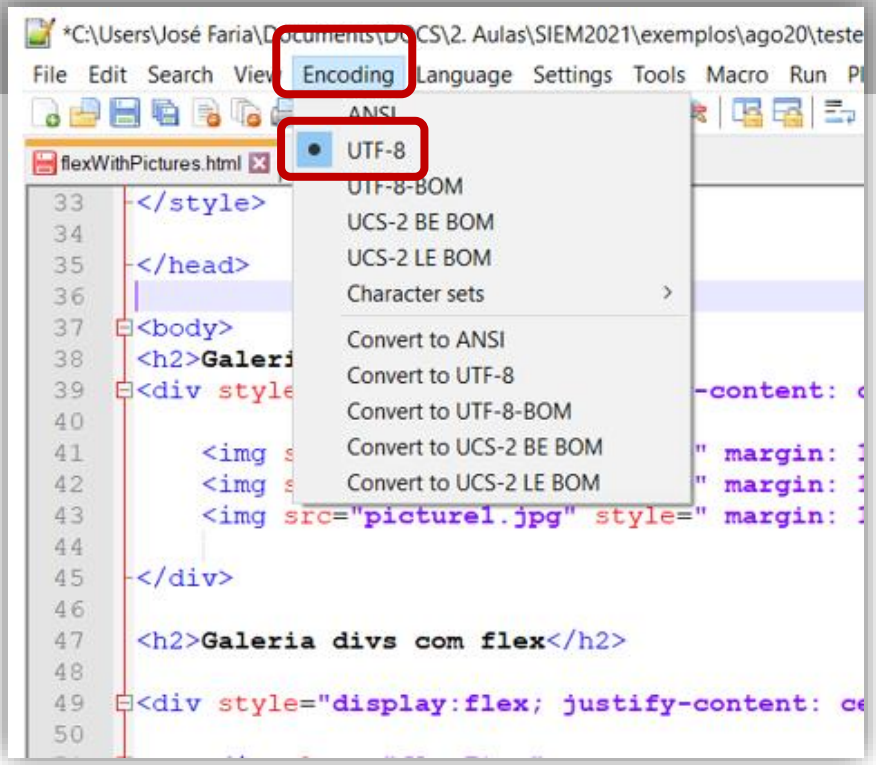
3.2. Charsets

- A charset defines the **internal code associated to each character** to be displayed

)	29	RIGHT PARENTHESIS
*	2a	ASTERISK
+	2b	PLUS SIGN
,	2c	COMMA
-	2d	HYPHEN-MINUS
.	2e	FULL STOP
/	2f	SOLIDUS
0	30	DIGIT ZERO
1	31	DIGIT ONE
2	32	DIGIT TWO
3	33	DIGIT THREE
4	34	DIGIT FOUR
5	35	DIGIT FIVE
6	36	DIGIT SIX
7	37	DIGIT SEVEN
8	38	DIGIT EIGHT
9	39	DIGIT NINE
:	3a	COLON
;	3b	SEMICOLON
<	3c	LESS-THAN SIGN
=	3d	EQUALS SIGN
>	3e	GREATER-THAN SIGN
?	3f	QUESTION MARK
@	40	COMMERCIAL AT
A	41	LATIN CAPITAL LETTER A
B	42	LATIN CAPITAL LETTER B
C	43	LATIN CAPITAL LETTER C

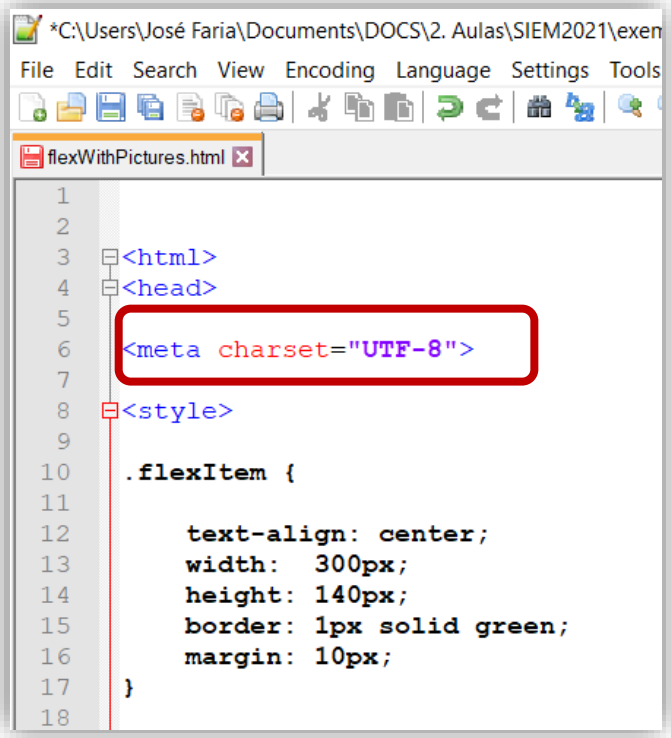
3.2. Charsets

- The **default character set for HTML5 is UTF-8**, which covers almost all of the characters and symbols in the world (UTF-8 employees 8 bytes per character!).
- To display an HTML page correctly, the **web browser must know which character set should use**.
- To provide that information to the browser, **proceed as follows** ...



1. check the character set configured in your HTML editor

2. add a meta command to the HTML file:



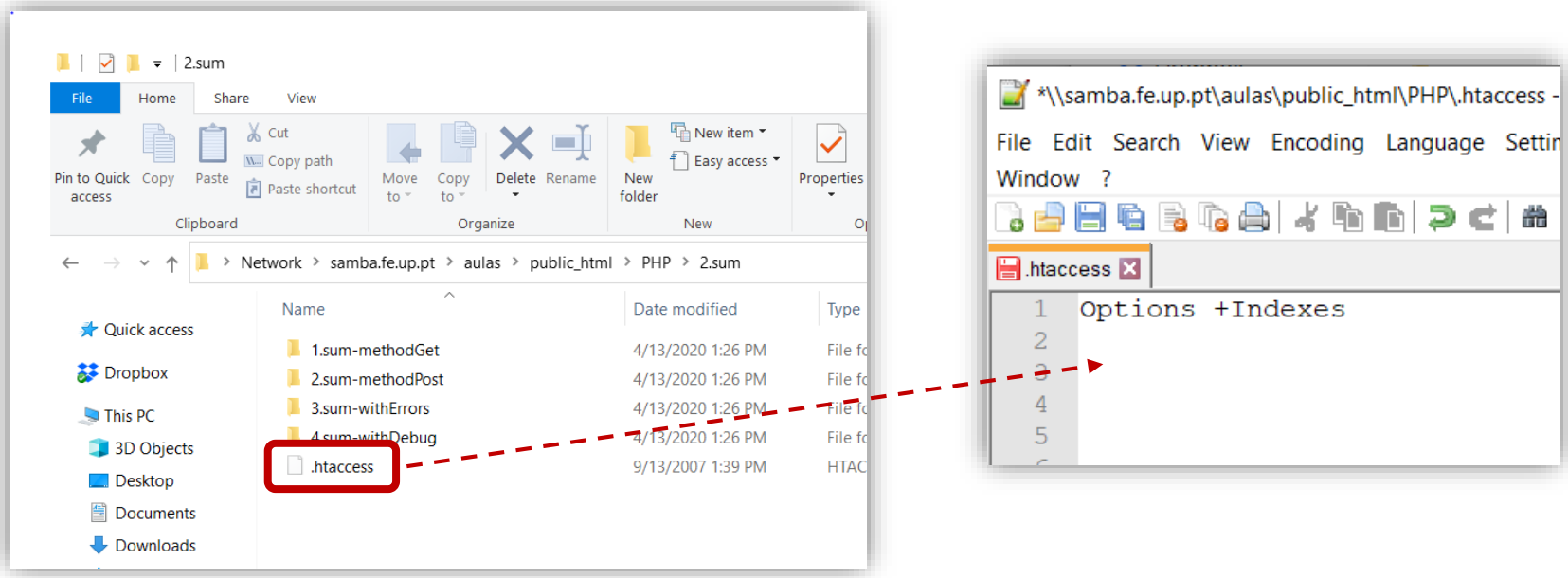
See more on charsets at

https://www.w3schools.com/html/html_charset.asp

3.3. Directory browsing

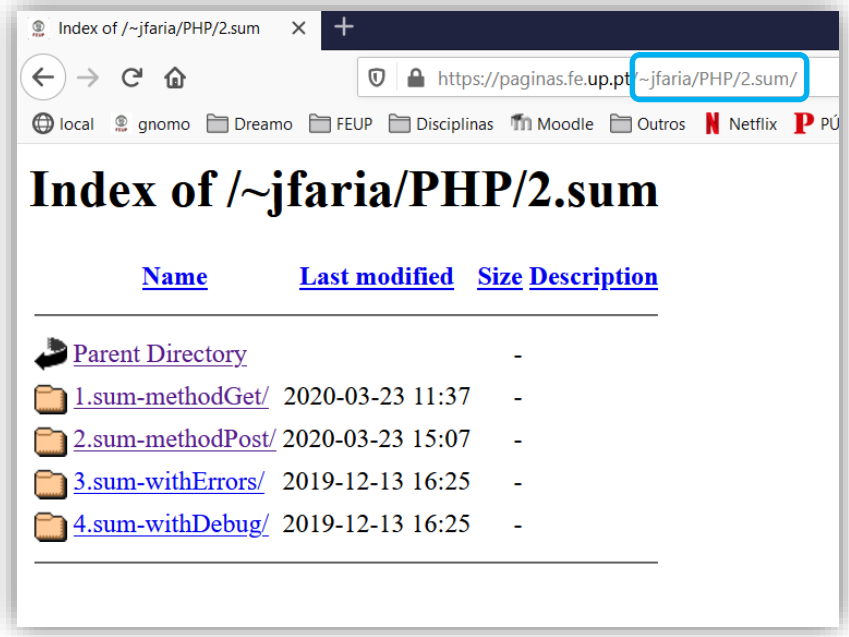
- When developing web applications it is very useful to able to browse the contents of directories.

- To **allow directory browsing** in a folder of a Linux server such as *paginas*, the folder should contain text file named **.htaccess** with ***Options + Indexes*** in its content:

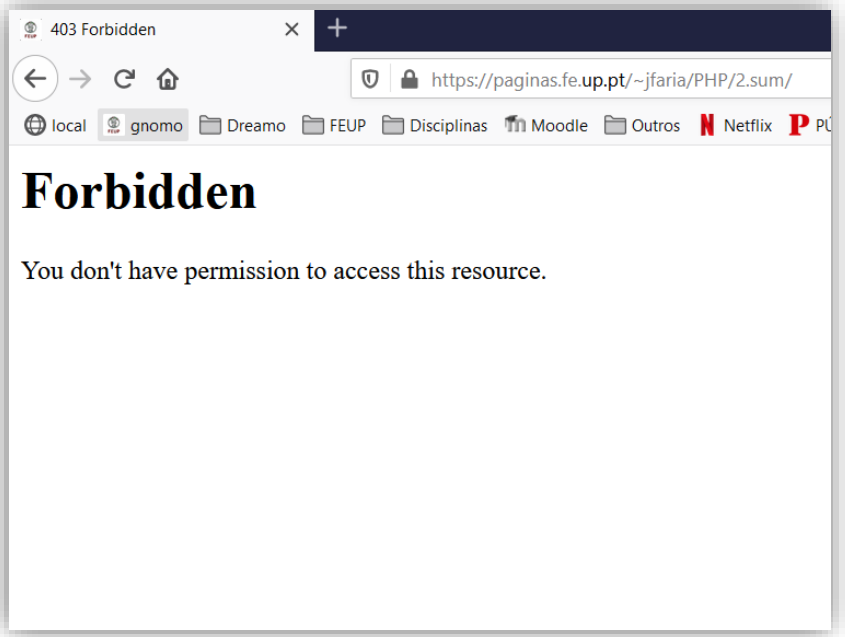


3.3. Directory browsing

with .htaccess in folder ~jfaria/PHP/2.sum you get:



without .htaccess you get:

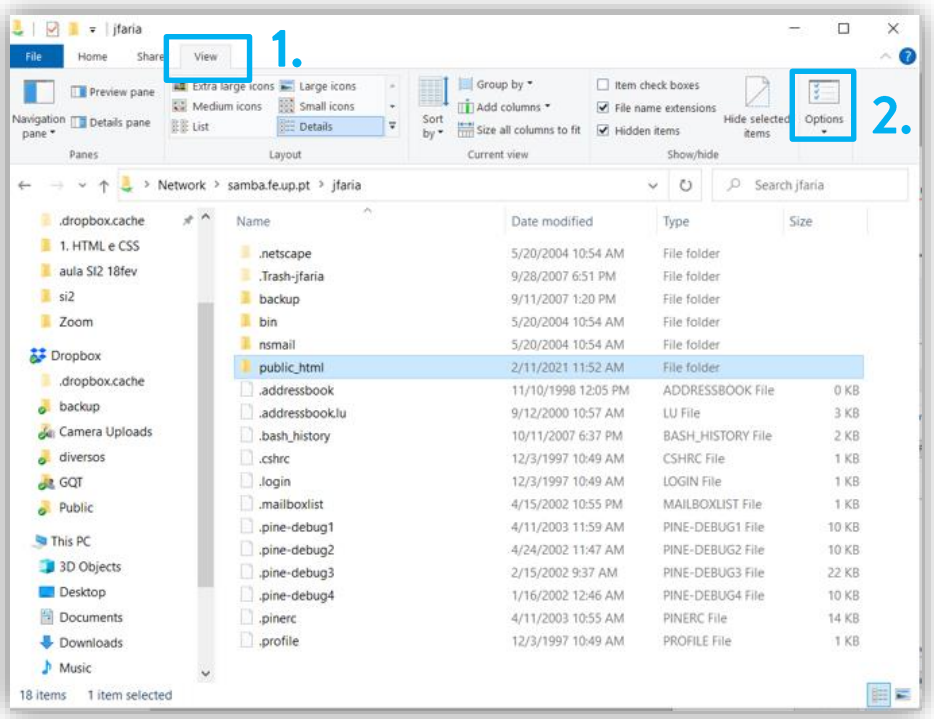


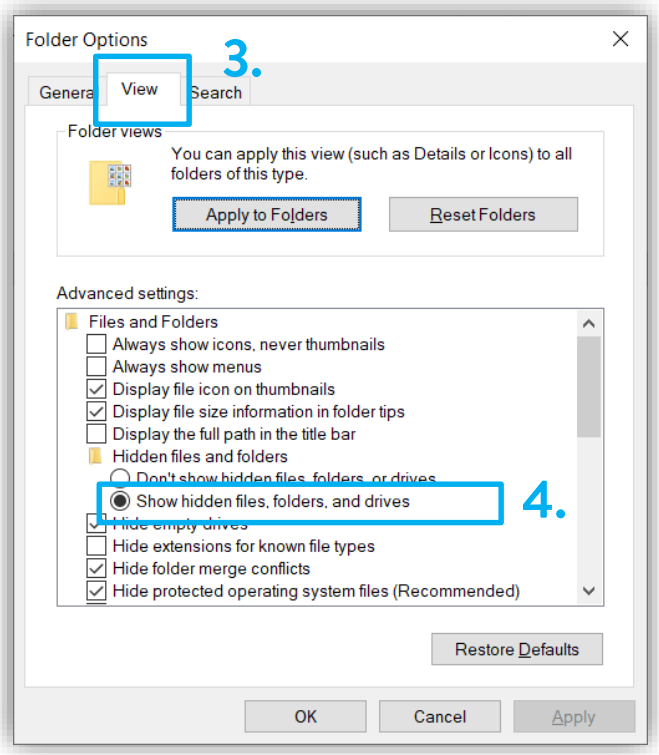
3.3. Directory browsing

- In Windows, you can not create a new file without a filename (such as .htaccess), but you can copy an existing file.
- So, **get .htaccess from Moodle** and **copy it to the remote disk** server using samba.

3.3.View hidden files like .htaccess

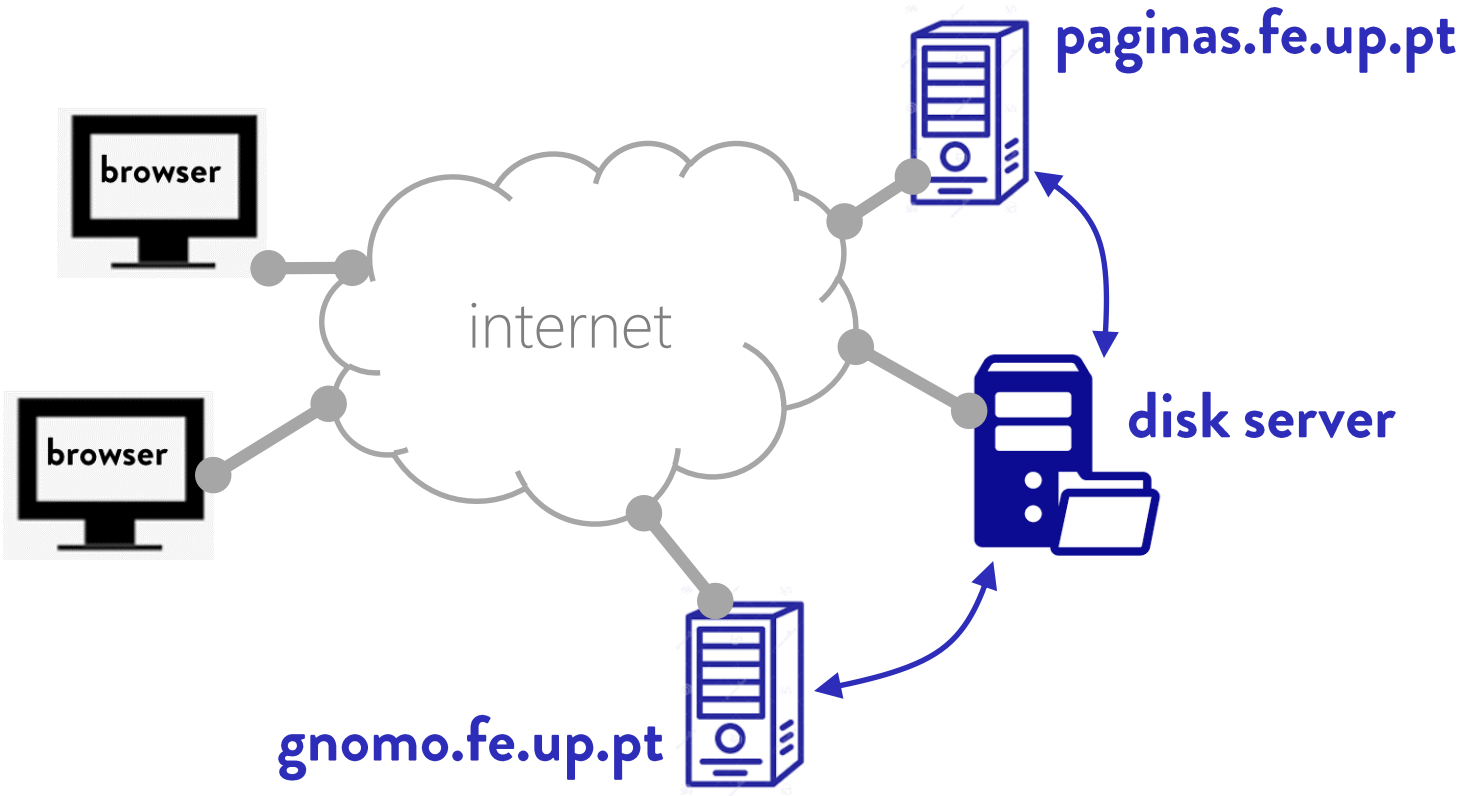
- To view hidden files (i.e., files with name .xxxx such as .htaccess) in Windows, you need to activate the option **Show hidden files, folders and drivers** in Explorer:





4. Folder browsing in gnomo and paginas

- At FEUP, you can run your web pages in **2 alternative webservers** **paginas.fe.up.pt** and **gnomo.fe.up.pt**.



4. Folder browsing in gnomo and paginas

- As you'll soon realize, when you start developing your first web pages and application, it is **very useful to browse the content of the webserver's folders**.

- Be aware that webservers **paginas and gnomo have a different behavior** regarding folder browsing as explained in the following slides.

Directory browsing in gnomo and paginas

- **gnomo** is a **development server** that is **private**, that is, it is only **accessible inside FEUP's** network (or from the outside with a VPN connection),

- **paginas** is a **production server** that is **public**, and so it may be accessed **from any computer connected to the Internet**.

Directory browsing in gnomo and paginas

- As an internal server, **gnomo is configured to offer directory browsing by default**.

  So, to have folder browsing, **you don't need to add the .htaccess** file.

- On the contrary, **paginas does not offer directory browsing by default due to security reasons** (as a public server is more exposed to hackers' attacks).

  In order folder browsing in paginas, you should **add the .htaccess file** to the browsable folders.

thank you !