

Short help for LibreWeb (document version 1.0.0)

This text is valid for version **1.0.x**, up to **1.1.0 (not included)**.

LibreWeb is an extension for LibreOffice, version 5.2 and above.

Tested on Windows, I have no information about the other operational systems.

The idea is catching data from internet and then insert it in a spreadsheet.

For a correct work you have to understand a basic idea of what a web page is.

Each page that your browser opens is a **text file**, ending with **.html**.

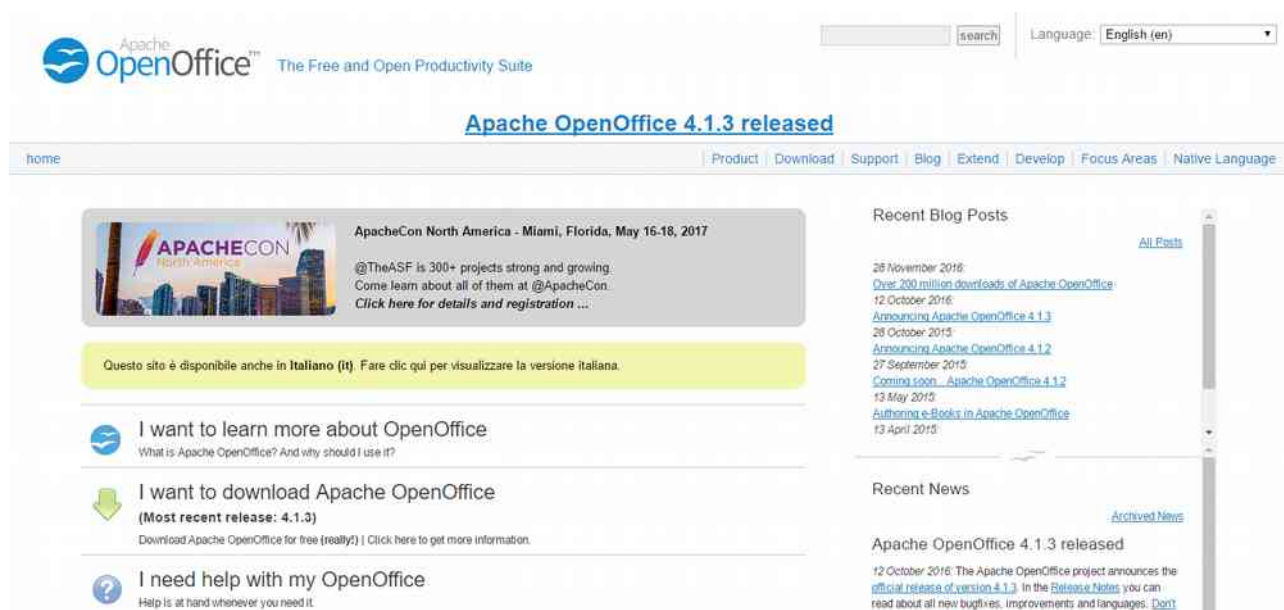
In this file there are a lot of tags: tags for links, tag for text, tag for images and so.

For more information search on wikipedia.

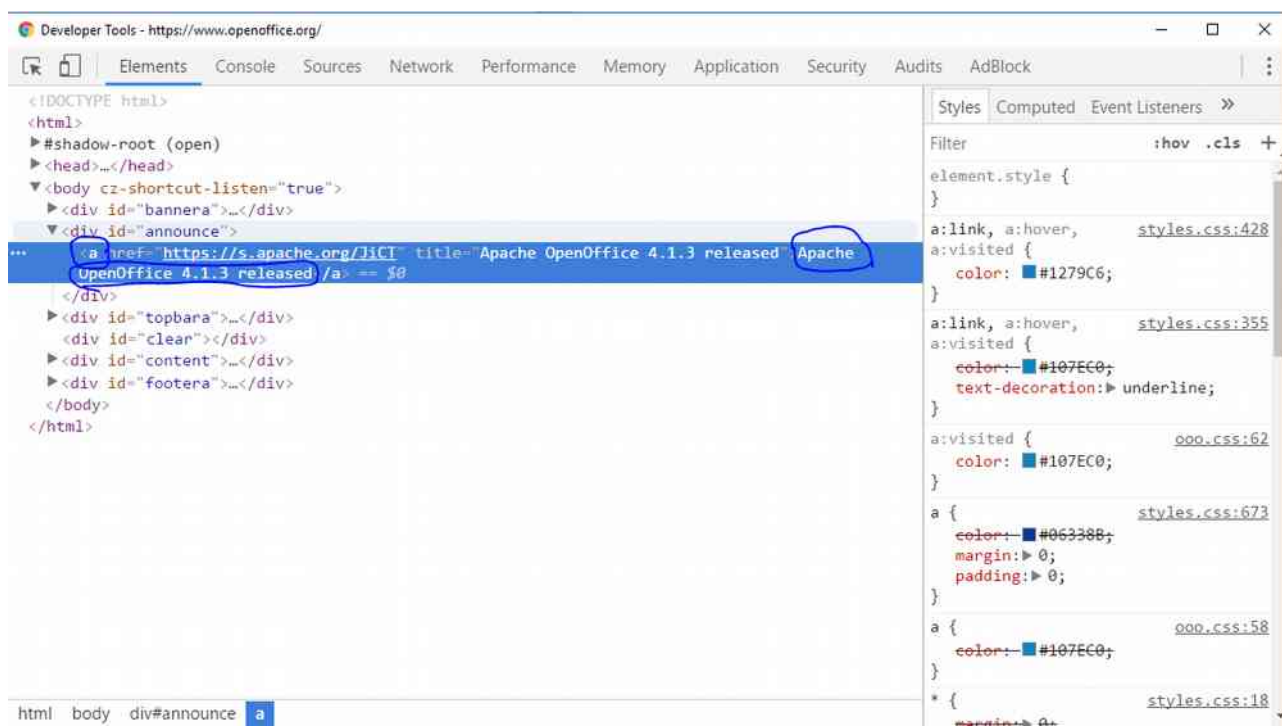
Many of this tags are nesting other tags, so you have to go ahead and find the **final tag**, that shows us a needing information.

Fortunately, each modern browser can show the source code of visited pages.

For example, open a web page with Chrome, maybe <https://www.openoffice.org/>



and put the mouse on the expression Apache OpenOffice 4.1.3 released. After make a **right-click** and in the context menu chose Inspect (or something similar). It is important to do this with mouse, because in this case you will get a new window, with emphasized string, that is our needing tag.



Character “**a**” means a link and inside this tag is our target text.
You have to remember the name of the tag, in this case it is **a**.

Next step, install **LibreWeb**.

There are two versions, one for user with root rights and other normal.
Feel free to try both, because surely one will give you some error at start.
Just uninstall it, after go ahead and install the second one.

Then open a Calc document, in **Menu** you'll see a new Item, **LibreWeb**,
with three commands, **Start service**, **Settings** and **Support Forum**.

For a correct work, **LibreWeb** create in current user folder a directory,
LibreWebSettings. There will be a small database for web-settings.

So after opening a document, push on menu **LibreWeb** → **Settings**.

You'll get the settings dialog:

LibreWeb Settings

libreweb.ods

https://www.openoffice.org/

Sheet Control

Foglio1

Tag Control

Tag Result

No data.

Control Panel

Check URL Save as: Save

1. In the first text field you must insert a valid URL.
2. In section **Tag Control**, there is also a text field. Insert the name of target tag, in our case **a**.
3. Then **Check URL**. You will get an error message if tag is not supported or URL is wrong, else **ListBox**, in **Tag Control**, turns on **Enable** (**Disable** by default), that means **LibreWeb** has found some quantity of your tags.
4. Just search on Items of **ListBox**, and field **Tag Result** will display the content of that tag. In my case this is the first **a** tag (number 0).

LibreWeb Settings

libreweb.ods

https://www.openoffice.org/

Sheet Control

Foglio1

Tag Control

a

0

Tag Result

Apache OpenOffice 4.1.3 released

Control Panel

Check URL

Save as:

Save

6. At this point, just a few things to do:

- In section **Sheet Control** set the cell where will be displayed next time the result, (**setting the Column and the Row**).
- Select a type of inserted data, for the moment you can choose between **String** and **Value**.
- Of course save the settings, **Save** button.

LibreWeb remembers the document name, so don't worry about namespace: you will see only the information, saved for a concrete document.

I hope to add some features, in future, like read and modify/delete settings directly from database. Feel free to tell me what would you like to see in **LibreWeb**.

Warning:

Some versions of LibreOffice come with a broken ssl module, if you continuously get error "*Unknown url: http*" or "*Unknown url: https*" let me know, there is a solution.

LibreWeb at this moment works only with LibreOffice, because python version of OpenOffice is 2, instead of needing 3.