

Converting Word Documents to HTML/Web Page Files for WebCT, Online Courses, and Other Web Sites

Introduction

Microsoft Word is by far the most common word processing software in use at NWACC and in America. Thus, there's a temptation to create documents in Word and then upload them to the Internet as Word documents. The problem with this is that, despite the huge popularity of Word, not everyone uses Word. Additionally, even those who use Word aren't always on the same computer platform. A person using Word in Macintosh is probably not going to be able to read a Word document created in Windows.

The best solution is to create documents in a format that is universally accessible and designed for WWW usage. While there are a few file formats that could do this, the best format for this purpose is HTML. Unfortunately, most people don't know how to create web pages using HTML. Fortunately, Microsoft has provided a command in Word that will do this for you.

Preparing a Word Document for Conversion to HTML

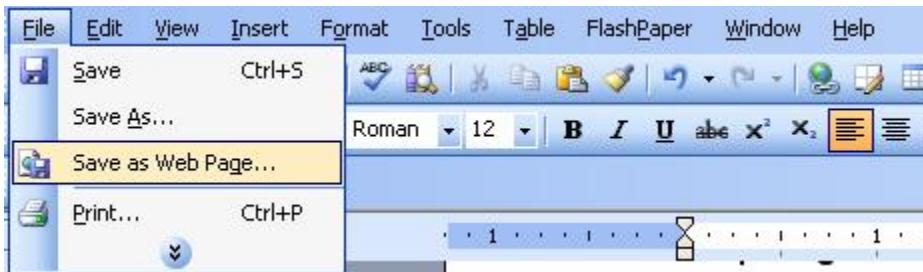
There's very little that's needed to be done to ready a Word document for conversion into HTML. Word simply looks at the document and writes the HTML needed to create the document as a web page on the WWW. However, sometimes people format Word documents using commands that don't neatly translate to HTML. When creating a Word document that will be converted to HTML, keep these things in mind:

- ***Don't use tabs to create and line up tables.*** They turn out uneven in HTML. Use the table tool in Word instead.
- ***Don't overuse line breaks.*** Allow the word wrap to carry text over to the next line. This will prevent unusual line break on the webpage. Only use line break when you really need the text to automatically go to the next line, or to create a paragraph break in the text.
- ***Don't set up bulleted lists using keyboard symbols.*** Instead, used the bulleted list selections in Word.
- ***Don't use exotic fonts.*** Web pages rely on the fonts that the viewer's computer has. So if they don't have the font you choose, the web browser will replace it with something else, usually something plainer. Use common fonts.

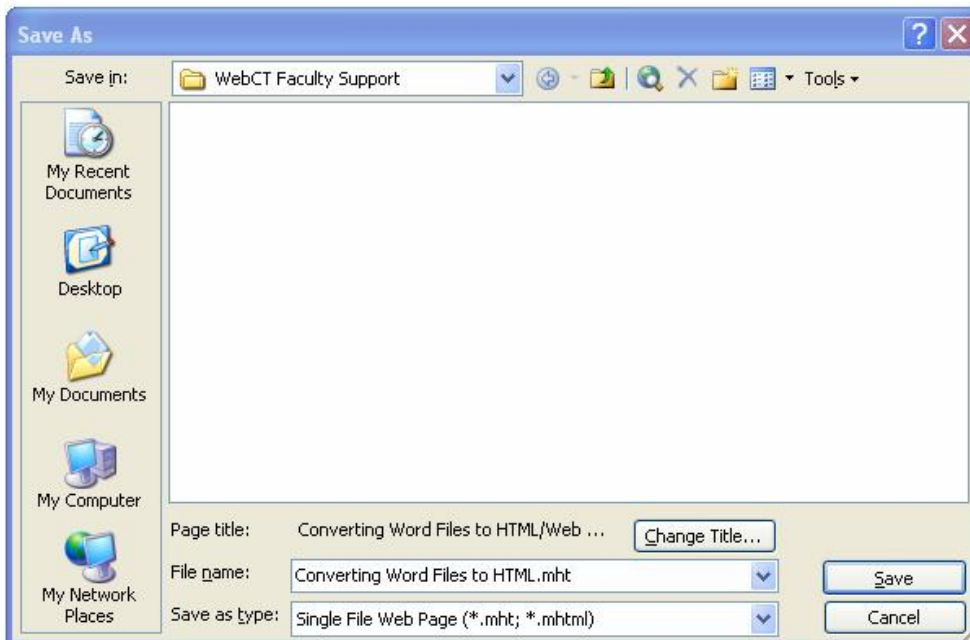
Saving as a Webpage

Once your Word Document is ready, the next step is to actually save the document as an HTML file, in Word.

First, go to the File menu and choose “Save as Web Page.” (In earlier versions of Word, this might read “Save as HTML.”)

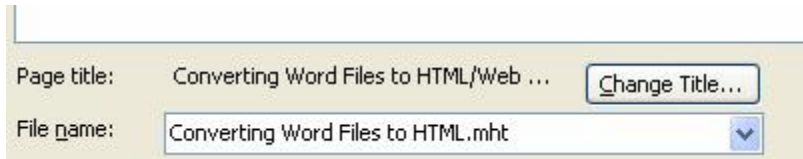


This will open the “Save as Web Page” window. When this window opens, you need to do three (4) things to set up the file:

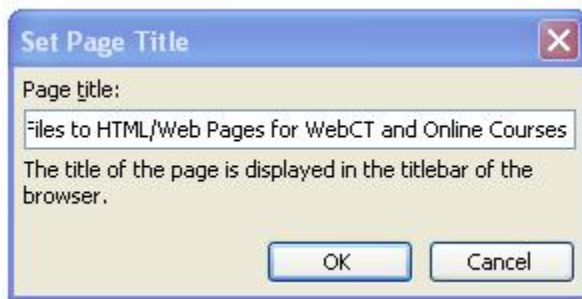


1.) Set the Page title.

The **Page title** is the text that will show in the blue Titlebar at the top of the web browser window. The text should reflect the title of this particular page. To change the Page title, select the button “Change Title” directly underneath the main part of the “Save As” window.

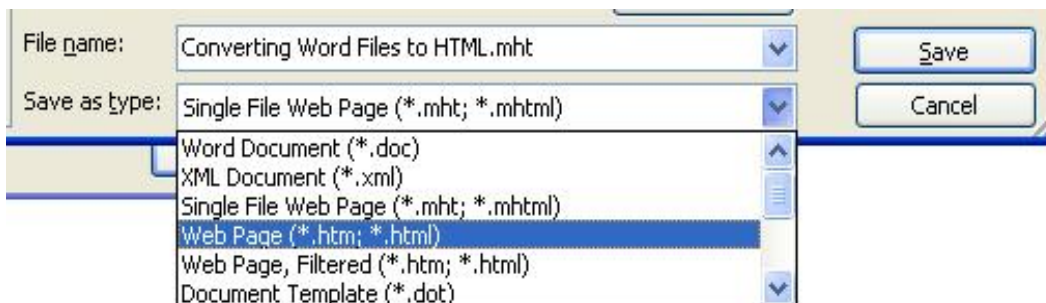


Then, simply enter the appropriate text to title the page, and select the “OK” button.



2.) Choose the appropriate File type.

Next, you need to make sure that the appropriate file type for a Web page is chosen. Word offers a few different file types that relate to HTML. However, the only one I feel works consistently enough to be considered is the file type “.htm” or “.html”. To choose this, select the drop down menu “Save as type” and choose “Web page (*.htm; *.html)”

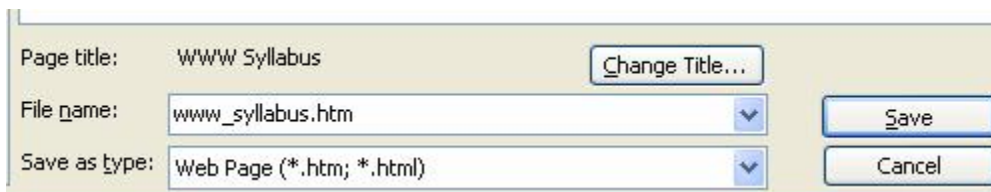


3.) Name the file.

Finally, you should give the file an appropriate name. First of all, make sure the file extension is either “.htm” or “.html”.

Also, one important thing to remember about HTML/Web files is that spaces in filenames do not work well on the WWW. On the web, a space in a filename is replaced with a “%20” which creates some awkward looking web addresses (URLs), as the filename is included in the web address. Therefore, if you feel you must separate words in the filename, use an underscore instead of a space.

Example: The filename “WWW Syllabus.doc” could be renamed “www_syllabus.htm”



The image shows a 'Save As' dialog box from a web browser. It has three rows of input fields and two buttons on the right. The first row is 'Page title:' with the text 'WWW Syllabus' and a 'Change Title...' button. The second row is 'File name:' with the text 'www_syllabus.htm' and a dropdown arrow. The third row is 'Save as type:' with the text 'Web Page (*.htm; *.html)' and a dropdown arrow. On the right side, there are two buttons: 'Save' and 'Cancel'.

All lowercases are not necessary, but they do make for a cleaner URL for that page.

4.) Choose a destination folder for the file.

Just as with any other file, you will need to determine a folder to save the new HTML file in. It could be the same folder as the one the original source Word .doc came from or a separate one for your web documents. Whichever folder you choose, be sure and remember the name of the folder so you can find the HTML file easily.

Finally, to save the file as an HTML page, simply select the “**Save**” button.

Checking Your HTML File

Once you have saved your Word document to HTML, you will want to make sure that it is formatted correctly. Check for the following to make sure that the formatting will present correctly in a web browser:

- Make sure there are no unusual or awkward line breaks. If this happens, it's usually due to a manual line break in the original Word document. Remember to use Word Wrap to allow Word to set up your line breaks within paragraphs. Only put in manual line breaks for paragraphs or when a separate line is absolutely necessary. If there are unwanted line breaks, simply remove them using the Backspace key.
- Make sure indented items line up properly. While you can tab indentations, they will not always line up correctly. You can use tables to line up items more accurately.
- Make sure bulleted lists are organized properly.

Once you have done this, you can **Save** your changes immediately. See **Updating Your HTML/Web Page File** below.

Updating Your HTML/Web Page File

Now that you've created an HTML file/Web Page, you will almost certainly want to update it and save it again at some point. Fortunately, once you've converted the file to an HTML file, you won't need to do that again. Microsoft Word let's you edit the file as an HTML file and save it again, just as you would with any other document in Word.

So, if you wanted to make changes to the HTML file, simply do the following:

- 1.) **Open** the HTML file in Word
- 2.) Make the changes needed to the file.
- 3.) **Save** the file.

You do not need to choose Save as Web Page/Save as HTML again for that file. In fact, if you do this, you will alter the file in a way that will prevent it from showing up properly in a web browser. So remember that once you've saved the file as a Web Page, you only need to do a **regular "Save"** after that.

Web Pages with Graphics

One other thing you will want to know. If your Word Document has graphics in it (pictures, clip art, etc.), **those graphics will not be saved to the HTML file** you create from your Word document. The reason for this is that HTML files are really just text files (.txt) and cannot support graphics. Web pages with graphics actually pull the graphic files into the web page when the web browser downloads them. So, **the graphics are actually saved separately from the HTML file.**

So, let's say that we have a Word document called **Syllabus.doc** with two pictures in it, a graphic at the top, and a clip art image in the body of the document. When we Save as Web Page, and create **syllabus.htm**, the graphics will be saved as separate files. It's important for use to know how this works and especially where those graphics files are located.

Word takes the name of the file and creates a folder to put the graphics in. This folder is named after the HTML filename with the appendage “_files”. So, our HTML file “**syllabus.htm**” will have a folder with the two graphics in it called “**syllabus_files**”. Word will create the folder “**syllabus_files**” in the same exact folder that you put “**syllabus.htm**” into. The HTML file then has commands in it that will tell the web browser how to find the graphics and where to insert them in the finished Web Page. The location of the graphics listed in the HTML file is called the “**path**.” For every graphic in your HTML file, there will be a path listed, so the web browser reading the HTML file can find the graphic. The nice thing is that Word puts all the graphics into the same folder, so the path is the same for all the graphics.

Here's the really important part that you need to remember: Whenever you move the HTML file to the web, either in WebCT, some other online course page (Course Compass or Course Studio), or your faculty webpage, the folder with the graphics must be moved into the same location as the HTML file. **This means that wherever “syllabus.htm” winds up, there must be a folder there as well with the two graphics in it in order for the images on that web page to load properly.**

If you are able to drag and drop or upload folders unto your web server, then it's simply a matter of moving the HTML file and the graphics folder together at the same time. If you are not able to drag and drop or upload a folder on to your web server, then you will need to create a folder with the exact same name as the one the graphics were created into (in this example “**syllabus_files**”) and copy or upload all the graphics files into that folder.

Again, the folder with the graphics in it must be copied or created in the same folder in which the HTML file resides. If not, the path in the HTML to the graphics files will be incorrect and the graphics will not properly load on the web page.

Uploading Your File to the WWW

Finally, you will eventually want to upload your HTML page (and any graphics on the page) to the WWW. Depending on where you are uploading the web page to, you may need to take various things into account. For NWACC Faculty, here are some notes regarding the various options:

- **WebCT** – When uploading HTML files into WebCT, you will need to either set up a “**link**” on an “**Organizer Page**”, or add the HTML file to a “**Content Module**.” See WebCT Help or contact the Distance Learning department for assistance as needed.

- **Course Studio** – Course Studio allows the upload of course documents. These documents are linked within Course Studio.
- **Course Compass** – Contact Course Compass documentation or help for more information on how to upload and display web pages.
- **Faculty Homepage** – You will need to have the IT Department, usually the NWACC Webmaster, set up a **faculty.nwacc.edu** account for you. Once this is done, you can upload/copy the files and folders to that account via Windows Explorer (Right click on Start Menu and select “Explore.”)

Further Assistance

For further assistance, contact the **Distance Learning Department** at (479)619-4382 or e-mail dl@nwacc.edu.