

chmProcessor. A Html / Word converter to Compiled HTML Help / Java Help (v 1.7.2)

chmProcessor is a tool for creating HTML help (CHM files) from MS Word or HTML files. If the file is a Word document, the section titles must use the "Title 1", "Title 2", etc styles that offers by default Word. If the file is a HTML file, the titles must use the <H1>, <H2>, etc. tags.

From this original file, each section of the document is split to different HTML files. Each of these files will be a topic page at the help. From the original document, you can generate:

- A Microsoft Help Workshop project.
- A compiled HTML help (CHM file).
- A web site.
- An Adobe PDF file.
- A Microsoft XPS file.
- A Java Help JAR file.

Requirements

To use this program, you must to have installed the following software:

- **Microsoft Windows.** Tested versions: Windows 2000 (32 bits), Windows XP (32 bits), Windows 7 (32 and 64 bits), Windows 8. Console interface will not work over windows 2000. Versions 1.4 and earlier of chmProcessor do not work on 64 bits operating systems.
- **Microsoft .NET Framework 2.0.** You can download the runtime version here: <http://www.microsoft.com/download/en/details.aspx?displaylang=en&id=19>
- **HTML Help Workshop.** If you don't have it, you cannot compile the help, you can generate a web site only. You can it download from here: <http://msdn2.microsoft.com/en-us/library/ms669985.aspx>
- **Microsoft Word** if you will work with Word files. Ms Word 2003, 2007 and 2013 has been tested. **Starter versions will not work**, they don't support automation.

Optionally, you can generate PDF and XPS files. To generate a PDF, there is two ways: PDFCreator or the Microsoft Office 2007 Add-in to generate PDF and XPS. **The last one is the recommended.** To generate a XPS file, the MS Office 2007 Add-in is mandatory:

- **PDFCreator. Version 1.2 ONLY.** You can download it from here: <http://www.pdfforge.org/products/pdfcreator/download>. **This kind of generation will not be supported on the future.**
- **Microsoft Office 2007 Add-in.** It's an optional tool of the Office suite to generate PDF / XPS files. It's not needed with Office 2013. You can download it from <http://www.microsoft.com/downloads/details.aspx?FamilyID=4D951911-3E7E-4AE6-B059-A2E79ED87041&displaylang=en>.

If you want to generate a Java Help, you will need the following software:

- **Sun JDK.** 1.4 or upper. Only 1.6 has been tested. You can download it from here: <http://www.oracle.com/technetwork/java/javase/downloads/index.html>.
- **Java Help** 2.0 or upper. You can download from here: <http://javahelp.java.net/>.

If you are going to generate CHM files with a character codepage different than your system default, for example if you live on Spain and your help will be written in Cyrillic, you will need this:

- **Microsoft AppLocale.** Can be downloaded from here: <http://www.microsoft.com/download/en/details.aspx?id=13209>

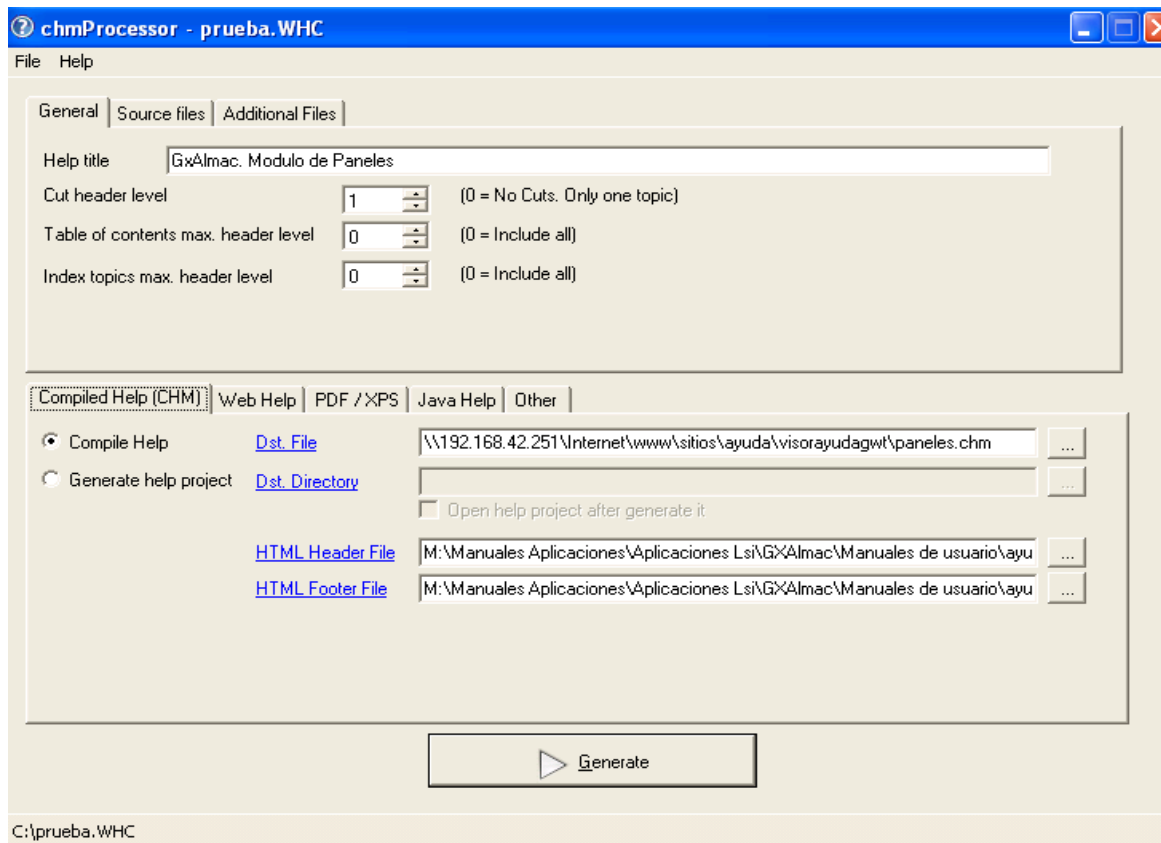
Limitations

- Versions 1.5 and older will not work generating CHM files with source files with encodings different than Windows-1252. Version 1.6 and higher should work.
- **Not all** kind of documents will be well formatted, especially HTML documents. Try to avoid put titles into tables, put images with title styles, and such kind of things, because the application will be confused about how to split and index the document. Avoid Javascript.
- Keep in mind that if you use more than one Word source files, the text styles used should be equal on all documents. So, if some style is different on some document, it can be overwritten by the definition of other document.
- JavaHelp generation will not work if “Use Tidy over the split HTML files” is not checked on [settings](#) window. JavaHelp seems to need 100% standard HTML input files to work.
- To generate CHM files with a character codepage (character set) different than your system default will only work if the source file is a Unicode encoded HTML file. MS Word 2003 / 2007 source files will not work.
- If you have a MS Office starter version, this application cannot make the HTML conversion automatically. You can do it manually using “Save as...” option and using the HTML file as source. **Make sure you save it as filtered HTML** (no HTML), and add the directory with the document images on “Additional files”

Interface

Win Interface

This is the main window:

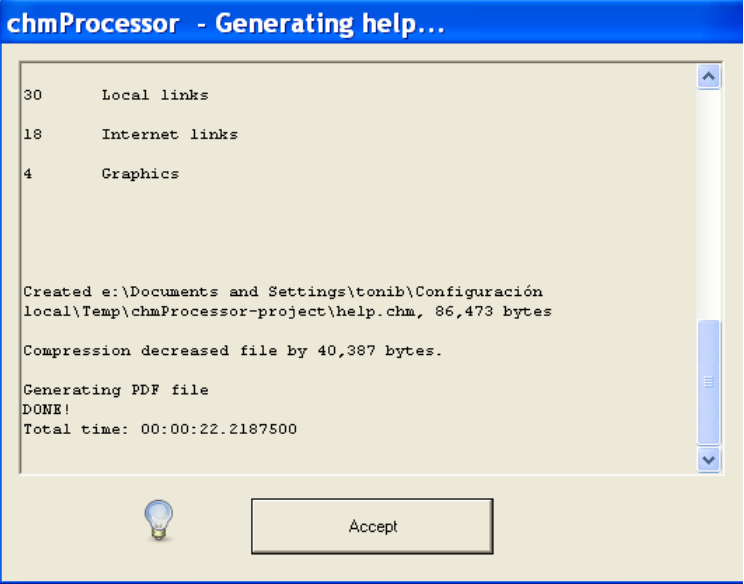


Source files

Here are stored the source document files to generate the help: A single HTML file, or a list of Word files. If there is more than one Word document, they will be joined to create the help.

General

Help Title	Title of the generated help.
Cut level	Level of sections that is used to split the document on topics. As example, if a level 2 is set, the document will be split on many topics as paragraphs with "Title 1" or "Title 2" (or tags <H1> / <H2> if the document is HTML) are found. If is zero, only a topic page will be generated for the entire document.
Table of contents max. level	Maximum level of the sections that will appear on the table of contents of the help. If is zero, all the titles will appear on the table. As example, if a level 3 is set, Titles 1, 2 and 3 will appear on the table of contents.
Index topics max. level	Maximum level of the sections that will appear on the index of the help. If is zero, all the titles will appear on the index. As example, if a level 3 is set, Titles 1, 2 and 3 will appear on the index.
Source files	List of HTML or Word file sources to generate the help. Only one HTML file can be used as source, but multiple Word documents can be used. No HTML and Word documents can be mixed on the list.
Additional files	List of other files / directories that must be included on the compiled help. For example if the source file is an HTML, and it contains images, these images, or the directory with the images must be

	included in this list.
Execute this command line after generate	If you need to make something after the generation, as copy the CHM to other place, put it on a web site by ftp..., you can do it by a command line here.
Generate button	<p>If the source file is a word document, you MUST NOT be editing the file when you press this button. When this button is pressed, the project / help generation will start and a new window will be opened, logging the generation process:</p> 

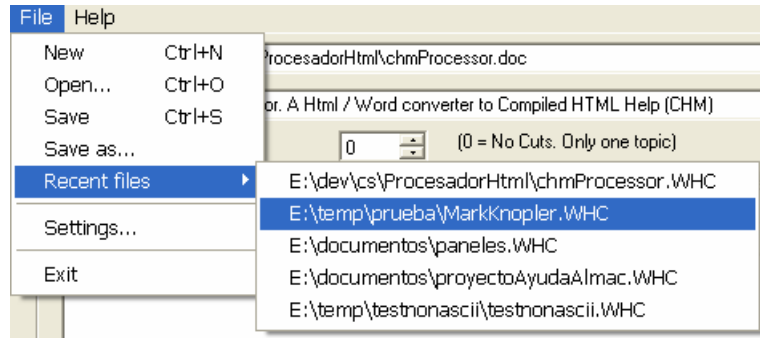
Additional files

Here is stored the list of additional files that are reference by the help document: Images referenced on the header / footers, text files, etc. You can include single files or directories. If you include a directory, all files and subdirectories into the directory will be added. All the additional files and directories will be included at the same directory than the document source files. So, if a hyperlink on the source files needs to reference to some of these files, the URL should be relative to the current directory (= “./theadditionalfiletoreference.txt”)

Important: By default, when MS Word saves a Word document as HTML, all relative links are rewritten. The application does a save as HTML into a temporal directory of the Word file to create the help. So, if you include a relative link “./file.txt”, it will be replaced by a “<temporaldirectory>/file.txt”. To avoid this, there is a MS Word setting called “Update links on save”. You **must** to disable this option to use relative links to additional files. [This article](#) explains where it is located and how to disable it.

Menu File

Here you can load and save the current help generation configuration from / to a “project” file, to access to the last open projects, and edit the settings of the application.



Compiled Help

Compile help / Dst. File	If it's checked, the help will be compiled on a CHM file, on the path selected at the Dst. File field.
Generate help project	If it's checked, a help project for the Microsoft Help Workshop will be generated at the directory selected on Dst. Directory, and it will not be compiled.
Generate web / Dst. Web directory / Description meta / Keywords meta	If "Generate" it's checked, a simple web with the help content will be generated on the Dst. directory. The description and meta fields are the values of this "meta" tags at the index of the web page generated.
HTML Header File	Can be null. If not, the content of this file will be added to the start of each topic page as header. The content must be HTML code. As example, this file can contain an IMG tag with the banner of the application title.
HTML Footer File	Can be null. If not, the content of this file will be added to the end of each topic page as footer. The content must be HTML code. As example it can be a tag to add an end bar to the page or a script to count visits at the generated web.
Language	Language of the CHM contents. If you select a language with a character codepage distinct of the system default, you must to install Microsoft AppLocale and check "Use Microsoft AppLocale to comple CHM files" on the Settings window.

Web Help

Generate Web / Dst. Web directory	If "Generate" it's checked, a simple web with the help content will be generated on the Dst. directory.
Description meta / Keywords meta	Values of the "meta" tags at the web pages generated. They should explain to the web indexers (as Google) the topic of the document.
HTML Header File	Can be null. If not, the content of this file will be added to the start of each topic page as header. The content must be HTML code. As example, this file can contain an IMG tag with the banner of the application title. This can be different to the header of the CHM.
HTML Footer File	Can be null. If not, the content of this file will be added to the end of each topic page as footer. The content must be HTML code. As example it can be a tag to add an end bar to the page or a script to count visits at the generated web, or a javascript code to

	register the visits. This can be different to the footer of the CHM.
Language	<p>Language used on the texts of the web site. If your favourite language is not here, you can create your own translation. To do this check at the “webTranslations” folder of the program. Copy the “English.txt” to your “mylanguage.txt”, open it with the notepad, and change the text of each second line. Example:</p> <p>“English.txt” file: Contents Contents Topics Topics ... “mylanguage.txt”: Contents <translation of Contents> Topics <translation of Topics> ... If you do this, please or put an entry at the patches tracker and it will be added to the application.</p>
Make Full Text Search (Require ASP.NET application)	<p>If it’s checked, a full text index of the document will be created, and all the document text can be searched. To run the searches, the web site must be hosted into an ASP.NET server. Only Microsoft IIS 6 and 7 has been test, but probably will work with the Mono server (only over Windows).</p> <p>If this field it’s not checked, a simple search at the web help will allow you to search at the topic titles of the document only.</p>
<head> include file	<p>Can be null. If not, the content of this file will be added into the <head> tag of each topic page. The content must be valid HTML code for that head.</p> <p>As example, it can be the javascript tag needed by Google Analytics to register the visits.</p>
Make Sitemap / Web Base / Change Freq.	<p>If “Make sitemap” is checked, a file called “sitemap.xml.gz” will be added to the web site. This file is a help for the web indexers as Google to know what pages contains your web site. See http://www.google.com/webmasters/sitemaps/ to know more about this file. If you check this field, you must to put the root of your web site at the field “Web base” (as example, “http://chmprocessor.sf.net”). At the field “Change freq.” you say the frequency of change of the document: daily, monthly, etc.</p>
Web template	<p>Here you select the template to generate the webhelp. chmProcessor comes with some standard templates that can be chosen here. If you have built your own template, you can chose here the value “Custom”, and select the location of your template at the “Template directory” field.</p>
Template directory	<p>If “Web template” is “Custom”, here you define the location of the custom template for this help project.</p>

The standard “template” pages for the web help are stored at the “webFiles” folder into the chmProcessor installation. If you want to change the look of the web help, you can edit them, but don’t change the texts %XXX% because they are replaced by the

application during the generation process. If you write a nice template and you want to share it with the world, please [put an entry at the patches tracker](#) and it will be added to the application.

If you want to add navigation links on the header / footer of the help topics, the application can replace some texts with the URL for the previous/next topic page. Here are the texts to put:

%NEXTLINK%	Link to the previous topic page. If there is no previous page, it will be replaced by a “#”
%PREVLINK%	Link to the next topic page. If there is no next page, it will be replaced by a “#”
%HOMELINK%	Link to the first topic page.

So, a tag to the previous page would be like this:

```
<a href="%PREVLINK%">Previous</a>
```

PDF / XPS

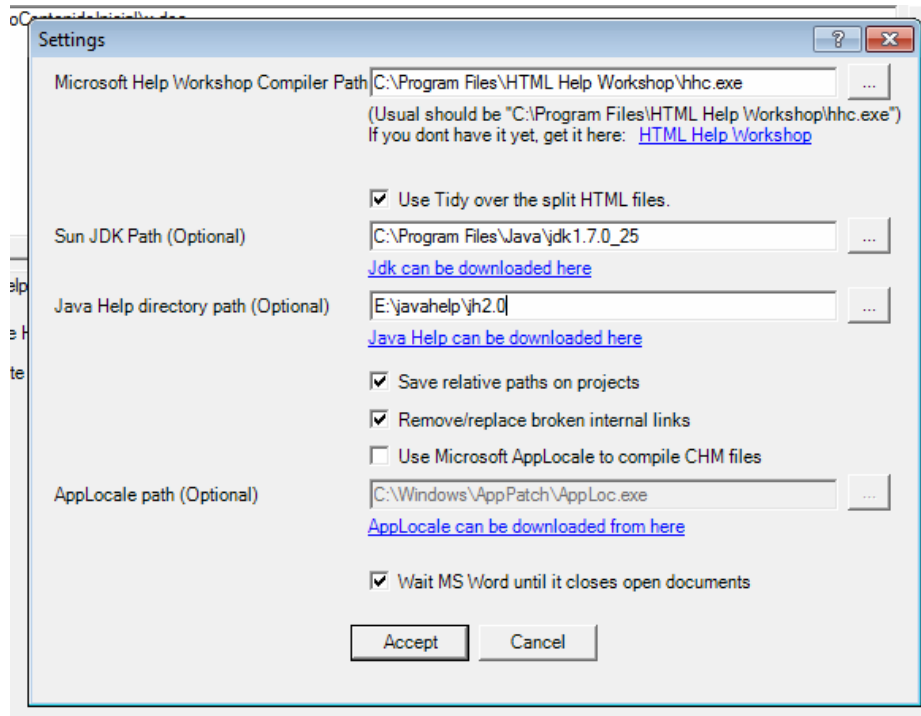
Generate PDF / Dst. File	If it's checked, a PDF file will be generated with the printing of the source Word / Html document at the file on Dst. File. You will need to have the PDFCreator or the “MS Office 2007 Add-in: Save as PDF / XPS” installed to use this.
Generate with PdfCreator / Generate with Office 2007 add-in (recommended)	Choose of how to generate the PDF file: With PDFCreator o with the MS Office 2007 Add-in. The last one is the recommended. Only version 1.2 of PdfCreator will work, and sometimes hangs up the PDF print.
Generate XPS / Dst. File	If it's checked, a XPS file will be generated with the printing of the source Word / Html document at the file on Dst. File. You will need to have the the “MS Office 2007 Add-in: Save as PDF / XPS” installed to use this.

Java Help

Generate Java Help / Dst. File	If it's checked, the help will be compile a Java Help JAR file, on the path selected at the Dst. File field. You will need to have the JDK and the Java Help installed to use this.
--------------------------------	---

Settings

Going to the menu File > Settings..., you can access to the settings of the program:



MS Help Workshop Compiler Path	Path to the compiler exe contained into the HTML Help Workshop. It must to point to the installation directory of this package.
Use Tidy over the split HTML files	Tidy is software than cleans and repair HTML files, according to the W3C standards. If it's checked, when a HTML page of the help is generated, tidy is executed over this file. Usually works VERY well, but I have found some problem with some XHTML files. If you have troubles with the generated HTML (strange characters, javascript errors, etc) try to uncheck this field.
Sun JDK Path	Installation path of the Sun JDK. It's needed if you want to generate Java Help.
Java Help directory path	Installation path of the Sun Java Help. It's needed if you want to generate Java Help.
Save relative paths on projects	If is checked, when you save a project, the paths to directories and files of the project will be stored relative to the location of the project file. So, if you store your source files and the project file on the same folder, you will can move that folder safely to other locations, and reopen the project and the reference to the source files will still work. If is not checked, paths will be saved as absolute, so, if you move your source files, they will not be found.
Remove/replace broken internal links	If it's checked, when an internal link (a link pointing to a place into the same document) is broken, the application will try to search a section into the document with the same title as the broken link text. If it's found, the link will be changed to point to that section. If no section is found, the link will be removed, but its content (text, image, etc) will be kept. If it's unchecked, the broken link will be kept as is.
Use Microsoft AppLocale to compile CHM files / path	If it's checked, the execution of the CHM compiler will be executed by the AppLocate application. This is only needed if you will generate a CHM file for a character codepage distinct than the system default.

	<p>This is the short story: AppLocate is an application that changes the codepage of the execution of other application. If you don't use it generating a CHM file for a character codepage distinct than the system default, the generated table of contents and index of the CHM file will show wrong characters.</p> <p>The long story: The MS Help workshops, and so the CHM compiler, are very outdated software, and they don't know anything about Unicode. More sad, if you set a Language on the help workshop project, you encode the source files for the TOC and index with the right encoding, all of this is ignored, and the local system locale is used to generate the CHM file. The patch for this is to use AppLocate, to lie the CHM compiler about the local codepage. More info about this topic can be found here: http://blogs.msdn.com/b/sandcastle/archive/2007/09/29/chm-localization-and-unicode-issues-dbcifix-exe.aspx</p>
Wait MS Word until it closes open documents	<p>When the help is generated from a Word file, MS Word is called to save it as HTML, then the Word file is closed and then MS Word is closed. It has been reported that sometimes the application hangs waiting Word do the closes.</p> <p>If this field is checked, the application will wait (up to 15 seconds) to Word closes files and itself. If it's not checked, the application will not wait. Uncheck this field can generate troubles with large Word files, and it's not recommended.</p>

Console Interface

The application can be called from the command line to generate or open project files. There are two exes for that: ChmProcessor.exe and ChmProcessorCmd.exe. The first one is the Windows application, and the second is a console application. ChmProcessor.exe can be used to open or generate help files, ChmProcessorCmd.exe only to generate help files. The command line parameters for both are the same:

```
Use XXXXX.exe [<projectfile.WHC> | <wordfile.DOC>
|<wordfile.DOCX> | <wordfile.HTM> | <wordfile.HTML>] [/g]
[/e] [/y] [/?][[/q] [/l1] [/l2] [/l3] [/l4]
Options:
/g          Generate help sets (chm, javahelp, pdfs,...)
specified by the project
/e          Exit after generate
/y          Dont ask for confirmations
/?         Print help and exit
/q         Prevents a window being shown when run with the
/g command line and logs messages to stdout/stderr
/l1 /l2 /l3 /l4 Lets you choose how much information is
output, where /l1 are errors, /l2 warnings, /l3 application
status information and /l4 are debug messages
```

If you are going to generate a help file from a batch file, ChmProcessorCmd.exe is preferred. Windows applications have limitations to interact with the console.

As you can see, you can generate a CHM directly from a Word file with a command line call, but this has a lot of limitations: The CHM generated will not have cut levels (a single help page will be generated), no title, etc. The preferred way to use the command line is create a WHC file, set there the title, cut levels, etc. and to generate the help with that file.

The application exit codes are:

- 0: OK
- 1: There were warnings.
- 2: There were errors.

Web Site Generation

The application can generate a simple web site with a similar format to the compiled help. Any topic indexed at the help can be opened directly using the URL. The way to open it depends of the template used.

For the "Frames" template, a "topic" parameter can be used on the URL. As example, if you have a topic called "Download", you can open it like this:

<http://chmprocessor.sf.net/index.html?topic=Downloads>.

For the "jQuery" template, you can use the "topic" parameter and the URL hash. The last one is the preferred, because the hash will be updated each time a new help topic is loaded. An example is <http://chmprocessor.sf.net/index.html#Downloads>.

Here are some examples of web sites generated with chmProcessor:

- [AdminUCV NGN manual](#)
- [SiteFX User's Guide](#)
- [Microrenewables Toolkit](#)
- [Bijlagen: de Vergelijkingeditor in Word 2007](#)
- [SprinxCRM online help](#)
- [Agile Salesforce to Exchange Synchronizer](#)

Note about Chrome

The standard webhelp templates of chmProcessor use frames (or iframes) and javascript to show the help content. Chrome has a security control that disable some javascript functions of frames/iframes when the page is on the local file system (=when the protocol is file:///). The standard webhelp templates of chmProcessor use those functions, so you will get an error:

```
Blocked a frame with origin "null" from accessing a frame
with origin "null". Protocols, domains, and ports must
match
```

This will not happen when the page is located on a web server (=when the protocol is http://). To fix this, you must run Chrome with the option "--allow-file-access-from-files":

```
"C:\Program Files\Google\Chrome\Application\chrome.exe" --
allow-file-access-from-files
"thewebhelpdirectory\index.html"
```

License

The license of chmProcessor is [GPL](#). It uses the following software:

- [Tango icons](#)
- [Tidy](#)
- [SQLite](#)

Downloads

You can download the chmProcessor installation binaries from [here](#). Source code is available from subversion repository [here](#). You can get this document as a PDF [here](#). See the [changelog here](#).

Bug reporting

The bug reporting can be done at the SourceForge page:

<https://sourceforge.net/p/chmprocessor/bugs/>. **Please, post always the following elements with the bug report:**

- An example file (the Word/Html document, additional files, the WHC help project file, etc.) to reproduce the bug.
- The chmProcessor version.

It's the faster way to fix bugs.

Info for developers

Here come some tips if you want to open a help file generated with chmProcessor from your application. Usually there is three ways to show a help topic on your application:

- Open directly the help file.
- Open a help topic by its title.
- Open a help topic by its HTML file name.

The last one is discarded, because you cannot control how chmProcessor will generate the file names, and more, those file names can change if you add more titles to your source help document. So, if you want to open a topic help directly you should use the topic text titles to open the help. Here are the functions you can use to open the help by one topic name on C# and Java:

JavaHelp

Your code should look something like this, without any warranty:

```
// The generated JAR must to be on your class path.
ClassLoader cl = this.getClass().getClassLoader();
URL hsURL = HelpSet.findHelpSet(cl, "help.hs"); //
help.hs if the fixed name for the generated help.
HelpSet hs = new HelpSet(null, hsURL);
HelpBroker hb = hs.createHelpBroker();
hb.setCurrentID("The topic title to show");
```

```
hb.setDisplayed(true);
```

Much, much more info on the JavaHelp documentation.

CHM file with C#

```
System.Windows.Forms.Help.ShowHelp(null,  
"C:\path\to\my\helpfile.CHM" ,  
System.Windows.Forms.HelpNavigator.KeywordIndex , "The  
topic title to show");
```

Web help with C#

Keep in mind that the topic title must be encoded for the URL:

```
System.Diagnostics.Process.Start("http://chmprocessor.sourc  
eforge.net/index.html?topic=Web%20Help");
```