

**M2Env**

**COLLABORATORS**

	<i>TITLE :</i> M2Env		
<i>ACTION</i>	<i>NAME</i>	<i>DATE</i>	<i>SIGNATURE</i>
WRITTEN BY		January 2, 2023	

**REVISION HISTORY**

NUMBER	DATE	DESCRIPTION	NAME

---

# Contents

<b>1</b>	<b>M2Env</b>	<b>1</b>
1.1	HTMLtoTeX - Main . . . . .	1
1.2	HTMLtoTeX - Usage . . . . .	1
1.3	HTMLtoTeX - command line options . . . . .	2
1.4	HTMLtoTeX - The built-in header . . . . .	4
1.5	HTMLtoTeX - Miscellaneous, Contact . . . . .	4

---

# Chapter 1

## M2Env

### 1.1 HTMLtoTeX - Main

This is

```
HTMLtoTeX Version 1.00
```

```
~~~~~
```

```
a CLI based tool for converting HTML documents  
into LaTeX2e textfiles.
```

```
written by Axel Sodtalbers, Goettingen, Germany.
```

```
~Documentation~
```

```
COPYRIGHT 1997 BY AXEL SODTALBERS, GOETTINGEN.
```

```
written 1997 by Axel Sodtalbers  
Philipp-Reis-Strasse 2/49  
37075 Goettingen  
Germany
```

```
email: asodtal@gwdg.de
```

HTMLtoTeX is freely distributable.

It may not be modified in any way. This especially includes changing the copyright notice.

No guarantee of any kind is given that the programs described in this document are 100% reliable. You are using this material at your own risk. The author can not be made responsible for any damage which is caused by using this program.

### 1.2 HTMLToTeX - Usage

---

HTMLtoTeX has been designed to convert HTML text (!) documents taken from the WorldWideWeb to LaTeX2e documents.

The program can not convert any type of HTML files. Trying to convert a web page full of graphics, applets and so on into LaTeX2e files is obviously useless!

In return, HTMLtoTeX recognises the following HTML elements and tries to convert them to suitable LaTeX2e tags:

- nearly all numeric and character entities
- most common font styles (such as <B>)
- information style elements (<CODE> for example),
- headings (<H1>)
- paragraphs (<P>) and linebreaks (<BR>)
- ordered, unordered and definition lists
- preformatted text (<PRE>), block quotes (<BQ>)

As you see, tabulars are (currently) not supported. Unknown HTML tags will just be filtered out.

HTMLtoTeX needs a minimum stack size of 20000 bytes. To use the program, open a shell and type the following:

```
1> stack 20000
1> HTMLtoTeX mywebtext.htm
1> latex mywebtext.tex
1> showdvi mywebtext.dvi
```

More easier is the following way:

```
1> stack 20000
1> HTMLtoTeX mywebtext
1> latex mywebtext
1> showdvi mywebtext
```

You can change the way the conversion is done by using some

optional~arguments

.

### 1.3 HTMLToTeX - command line options

The template:

FROM/A, TO, HEADER, Q=QUIET/S, GERMAN/S, LINKS/S, RULES/S, PLAIN/S, BUFFER/L

---

The options:

- FROM/A      Your source file in HTML format.  
If the file cannot be found, HTMLtoTeX appends the ".htm" appendix to the filename automatically and tries to read it again.
- TO            The name of the target file produced by HTMLtoTeX after conversion.  
If there is no filename supplied, HTMLtoTeX takes the one of the sourcefile and changes the ".htm" appendix to ".tex".
- HEADER       HTMLtoTeX automatically produces a complete latex file with a  
              ~suitable~header~  
              .  
If you want to use your own header, use this option. Then, HTMLtoTeX reads the header file, appends the converted text and saves the whole file to the target file.
- Example: HTMLToTeX Test.htm header=TeX:Text/myheader.tex
- Q=QUIET/S    When using this option, HTMLtoTeX does not print any output into the CLI output stream. Useful within DOS scripts.
- GERMAN/S     The option "german" will be used in the LaTeX2e option.
- LINKS/S      The HTML hyperlinks will be marked in the latex2e textfile by using the `{\sffamily [link]}` tag.
- RULES/S      Some HTML text documents are containing horicontal rules (`<HR>` tag). This looks quiet strange in latex2e documents, so the horicontal rules will be filtered out normally. If you want to keep them, you have to use this option.
- PLAIN/S      By using this option, you can force HTMLtoTeX to convert only the pure text information. There won't be any style conversion.
- BUFFER/N     The size (in KBytes) of the input filebuffer used while converting. The default is 10, which means a filebuffer of 10 KBytes.  
The larger the filebuffer, the faster the conversion will be. While loading the source file, HTMLtoTeX will print out a dot (.) every time the file buffer is full.
-

## 1.4 HTMLtoTeX - The built-in header

Here you can see the built-in standard header.  
It uses the "fancyheadings" package for easy page layout control.

You can use your own header by using the "header"

~option~

.

```
~~~~~  
  
% Converted with HTMLtoTeX 1.00 © by Axel Sodtalbers.  
  
\documentclass [a4paper] {article}  
  
\usepackage[latin1]{inputenc}  
\usepackage{fancyheadings}  
\usepackage[T1]{fontenc}  
  
\pagestyle{fancy}  
\lhead{[...title...]}  
\chead{}  
\rhead{}  
\lfoot{}  
\cfoot{}  
\rfoot{}  
  
\parindent0pt  
  
\begin{document}  
\sloppy  
  
[...converted body...]  
  
\end{document}
```

## 1.5 HTMLtoTeX - Miscellaneous, Contact

While programming HTMLtoTeX, the aim was to write not a fast but a compatible code. For example all attributes of the HTML tags are read (but currently not processed).

If you have any ideas, suggestions for enhancement or error reports, please send an email to me:

asodtal@gwdg.de