

The HTMLToWord Object



HTMLToWord is an open source project from SoftArtisans. It can be downloaded from [SourceForge.net](https://sourceforge.net/projects/htmltoword/).

Table of Contents

- [Examples:](#)
- [HTMLToWord Constants](#)
- [HTMLToWord Enumerations](#)
- [HTMLToWord Properties](#)
- [HTMLToWord Methods](#)

Use the HTMLToWord object to insert well-formed HTML (XHTML) text into a WordWriter Documentobject. You can specify a style to use when inserting text or use the document's default style. You can also specify whether to insert text after a specific element or at the end of a document.

The contents of HTML tags not listed on the [supported tags page](#) may be inserted into the document as plain text. The behavior is dependent on the settings of the AcceptUnknownTags and IncludeContentsOfUnknownTags properties found in the HTMLInsertProperties property container.

For an overview of how to use the HTMLToWord object, see [Using HTMLToWord](#).

Examples:

C#

```
//--- Well-formed HTML fragment to be inserted into the document.
string htmlFragment = @"<p><b>Hello,</b><i> World</i>!</p>";

//--- Open an existing Word document with the SoftArtisans WordApplication object.
WordApplication wApp = new WordApplication();
Document doc = wApp.Open("myDokument.doc");

//--- Create the HTMLToWord object.
HTMLToWord h2w = new HTMLToWord();

//--- Insert the HTML fragment at the end of the document.
//--- Use the default style.
h2w.InsertHTML(htmlFragment, doc, null, doc);
```

VB.NET

```
'--- Well-formed HTML fragment to be inserted into the document.
Dim htmlFragment As String = "<p><b>Hello,</b><i> World</i>!</p>"

'--- Open an existing Word document with the SoftArtisans WordApplication object.
Dim wApp As WordApplication = new WordApplication
Dim doc As Document = wApp.Open("myDokument.doc")

'--- Create the HTMLToWord object.
Dim h2w As HTMLToWord = new HTMLToWord

'--- Insert the HTML fragment at the end of the document.
'--- Use the default style.
h2w.InsertHTML(htmlFragment, doc, Nothing, doc)
```

HTMLToWord Constants

Name	Description
HTMLToWord.Delegates.InsertElementDelegate	Type for user-defined delegate method called by HTMLToWord before inserting an element. HTMLTagAction InsertElementDelegate(Document doc, Element insertAfterElement, XmlNode node)
HTMLToWord.Delegates.FormatElementDelegate	Type for user-defined delegate method called by HTMLToWord after inserting an element. void FormatElementDelegate(Element insertedElement, XmlNode node)
HTMLToWord.Delegates.GetImageDelegate	Type for user-defined delegate method called by HTMLToWord when an HTML img tag is encountered. System.IO.Stream GetImageDelegate(string URI)

HTMLToWord Enumerations

Name	Description
HTMLToWord.HTMLTagAction	Return value from the user-supplied InsertElementDelegate delegate method
HTMLToWord.LoggingLevel	A LoggingLevel value specified the amount of information written to the log file

HTMLToWord Properties

Property	Type	Access	Description
DebugLogLevel	HTMLToWord.LoggingLevel	read write	Sets the amount of information written to the debug log file
FormatDelegate	FormatElementDelegate	read write	User-assigned delegate method that is called after HTMLToWord inserts an Element into the Word document
ImageDelegate	GetImageDelegate	read write	User-assigned delegate method that allows users to provide an implementation for retrieving the contents of an image and passing it back to HTMLToWord for insertion into a document
InsertDelegate	InsertElementDelegate	read write	A user-assigned delegate method that HTMLToWord calls before it processes an XHTML tag
InsertProperties	HTMLInsertProperties	read write	Contains settings that control how unsupported HTML tags are treated and provides an override to the default font used when inserting text into the document.
LogFileName	String	read write	The full path to the HTMLToWord debug log file

HTMLToWord Methods

Returns	Signature and Description
Element	HTMLToWord.InsertHTML(String, Document) Processes the XHTML string and inserts it at the end of the specified Word document as formatted text using the document's Normal style as the default font. Returns the last Element inserted into the document
Element	HTMLToWord.InsertHTML(String, Document, Element) Processes the XHTML string and inserts it into the specified Word document as formatted text after the given WordWriter Element using the document's Normal style as the default font. Returns the last Element inserted into the document
Element	HTMLToWord.InsertHTML(String, Document, NamedStyle, Element) Processes the XHTML string and inserts it into the specified Word document as formatted text after the given WordWriter Element using the given NamedStyle. Returns the last Element inserted into the document