

Inserting HTML with WordApplication

Introduction

SoftArtisans' open source project HTMLToWord allows users to insert well-formed HTML(XHTML) snippets into Word documents as formatted text.

The project can be downloaded from [SourceForge.net](#).

What is HTMLToWord?

HTMLToWord is an Open Source Microsoft Windows .NET C# project by SoftArtisans, Inc. that allows users to insert fragments of well-formed HTML (XHTML) into a Word document as formatted text. HTMLToWord uses SoftArtisans' OfficeWriter for Word (WordWriter) WordApplication object.

HTMLToWord parses XHTML fragments and inserts the contents into a Word document at the location you specify. In general, you can use HTMLToWord to insert headings, paragraphs, lists and tables into a new or existing Word document that you open or create with the SoftArtisans WordApplication object. The current version of HTMLToWord is not designed to convert entire HTML pages to Word documents.

HTMLToWord understands a subset of the tags defined in the HTML specification. See the documentation page on [supported tags](#) for a complete list of supported tags and their support level. For any HTML tag not listed in the table, or if your XHTML documents have custom tags, HTMLToWord will attempt to insert the text contents of the tag into the Word document. You can override this default behavior by setting the [AcceptUnknownTags](#) and [IncludeContentsOfUnknownTags](#) properties found in the [HTMLInsertProperties](#) property container and by implementing the [InsertElementDelegate](#) and [FormatElementDelegate](#) delegate methods.

To get started with HTMLToWord see [Using HTMLToWord](#).



For information on inserting HTML with WordTemplate see [Inserting an Embedded Document](#)