

Help Contents

Introduction

[About WebAnchor](#)

[Registering WebAnchor](#)

Getting Help

[Readme File](#)

[Customer Support](#)

Quick Start Guide

[Trial Run of WebAnchor](#)

What WebAnchor Does

[WebAnchor: An Overview](#)

[Phase 1: Analyzing the Document](#)

[Phase 2: Inserting Anchors into the Document\(s\)](#)

[Phase 3: Creating Index Pages](#)

Using WebAnchor

[Starting the Program](#)

[The Tool Bar](#)

[Choosing Files to be Indexed](#)

[Output Settings](#)

[Generating the Index](#)

[Viewing the Results](#)

[Using Projects to Save Settings](#)

Changing the Appearance of the Index Page

[Index Page Overview](#)

[Page Layout Options](#)

Customizing Output with Lexicons

[Lexicon Overview](#)

[Launching LexEdit](#)

[Loading Lexicons](#)

Converting RTF to HTML

RTF Conversion Overview

Setting Conversion Options

Hints and Troubleshooting

Getting the Best Entries

Known Problems

Appendix A

Guides and Reference Work on HTML

About WebAnchor

Welcome to the world of automatic indexing. WebAnchor analyzes your HTML documents, finds significant words and phrases, and compiles an index that will hyperlink to these words and phrases in your documents. The compiled index is displayed on Index Pages which are hyperlinked to your original documents allowing quick and easy browsing of document sets. You control the volume of the output, how selective WebAnchor is in choosing what to include in your output, and the content of the output through enabling and disabling such features as personal and geographical names and through using personal lexicons created in LexEdit (See LexEdit section).

Registering WebAnchor

Please take a few moments to fill out your registration card and send it in. As a registered user, you will be eligible for free technical support, reduced rates on upgrades, and special offers on new Iconovex software. In addition to upgrading the software itself, we are continually working to upgrade the quality of the primary lexicon that WebAnchor uses to make its semantic decisions, and these lexicon upgrades will also be sent out periodically to registered users.

If you have any questions regarding registration, please call (800) 943-0292. You can also e-mail us at fyi@iconovex.com or post questions to the Iconovex Web server at <http://www.iconovex.com>.

Readme File

Since this documentation must go to the printer several weeks before the software itself is in its final form, there may be changes in the software that are not covered in the *User's Guide*. These late changes, and any other additions to the *User's Guide*, will be covered in the `README.TXT` file which will be placed in the `WBANCHOR` directory by the installation program. You will be given the option of reading this file immediately after installation. If you do not read it then, you can open it later in Notepad or in any text editor, including the DOS Edit program.

Customer Support

At Iconovex, we attempt to provide our customers with as much help as possible. If you have any questions about using WebAnchor, or if you have a technical problem that you can't solve, please call us at 1-800-943-0292. Please have your serial number, which can be found on the **Help–About** screen, available when you call. Our technical support staff will be happy to assist you.

You can also contact us by:

Fax: (612) 896-5101

U.S. Mail: 7900 Xerxes Ave. S., Suite 550, Bloomington, MN 55431

E-mail: 74064.440@compuserve.com

or

techsupp@iconovex.com

WWW: <http://www.iconovex.com>

Trial Run of WebAnchor

The following trial run of WebAnchor will guide you through the steps of creating an index as well as give you a brief view of the options available to you. This trial run uses sample HTML files provided in the `SAMPLES` directory that was created when WebAnchor was installed.

Step 1: Open WebAnchor from the Iconovex program group or folder.

Step 2: At the main screen, use the directory browser on the left side to find the `WBANCHOR\SAMPLES` directory and double-click it. Several `.HTM` files should now appear in the Files box.

Step 3: We will be processing `BBS05.HTM`, `BBS06.HTM`, `BBS07.HTM`, `BBS08.HTM`, and `BBS09.HTM`. Select each file by highlighting it with your mouse and then clicking on the single right arrow button to move it to the Files to be Indexed box on the right side of the screen. You can also select files with your keyboard by using the Tab key to navigate around the screen and the spacebar to select the files.

Step 4: Choose **Options—Settings**. In the Index File Name box, there will be a default index name (`INDEX.HTM`) and directory (`OUTPUT`) for your processed files. You may change these settings if you like. When finished, click **OK**.

Step 5: Now simply click on the **Generate Index** button. WebAnchor will present a message that reads "Index entries will now be generated and placed into your source files. Continue?" WebAnchor needs confirmation from the user to alter the source files by adding index entries to them. Click **Yes** and WebAnchor will proceed to generate Index Pages for your sample files.

Step 6: In order to view your results once WebAnchor is finished running, open your browser and point it at the file you specified as the "Index File." If you accepted the default settings, your filename will be `INDEX.HTM` and your output directory will be (`WBANCHOR\OUTPUT`).

Step 7: You will now be looking at the Index Page. There will be a large **INDEX** graphic at the top and a hyperlinked list of the individual letters of the alphabet. Click on the letter "C."

Step 8: WebAnchor will jump you to the "C" Index Page. Once there, scroll down the list until you get to the phrase "Constitution of United States" and click on it.

Step 9: WebAnchor will jump you to the first significant occurrence of that phrase in the processed documents. This should be the very first entry on the page, but that will vary based on the browser you are using. If the entry is not at the top of the page, scroll up or down a few lines and you will find it. You will notice that there is a green ball icon directly in front of the phrase. The green ball icon serves as a circular link that will jump you to every other significant instance of that exact phrase.

Step 10: Click on the link. WebAnchor will jump you to the next occurrence of "Constitution of the United States" in the documents. There will be a green ball icon in front of each occurrence of the phrase that will take you to yet another occurrence. Continue to click on these links until you are returned to the Index Page.

Step 11: WebAnchor should jump you to a total of 5 occurrences of "Constitution of the United States." When you have reached the last occurrence, WebAnchor returns you to the main Index Page where you began. From here, you can continue to search for other terms that you may find interesting in the documents.

This trial run should act as a general basis for processing all other files. As you become more comfortable with WebAnchor, you can customize each setting to guarantee that you are getting the best

personal index.

WebAnchor: An Overview

WebAnchor was created to provide quality, automatic indexing and hyperlinking for HTML documents. WebAnchor also converts any RTF files you may wish to process to HTML automatically. (For more information on the RTF conversion process, see [RTF Conversion Overview](#).) Like back-of-the-book indexes provided in hard copy publications, WebAnchor provides quick reference and information access for online publications. In order to use WebAnchor effectively, it is important to understand exactly what the software does. There are three distinct phases of the process: [analyzing the document\(s\)](#), [inserting hypertext anchors within the document\(s\)](#), and [creating alphabetical Index Pages](#). In each of these phases, you have control over the options that affect the final content and appearance of the index.

Phase 1: Analyzing the Document

WebAnchor analyzes HTML documents and extracts their significant words and phrases. You can control what WebAnchor considers "significant" through the Index Level setting, which allows you to declare how generally selective or inclusive you wish to be when creating Index Pages for your document(s). The higher the Index Level, the more selective WebAnchor will be in its placing of anchors within the document(s). Run-time options in the Settings dialog let you determine whether or not personal names and/or geographical names will become anchors.

With LexEdit, a utility program included with WebAnchor, you can build a personal lexicon of terms you consider important as well as words you don't consider important. This personal lexicon can then be used to instruct WebAnchor about what to index in your documents. For documents with specialized vocabularies, you can instruct WebAnchor to produce a list of words it doesn't recognize. You can then enter these words in your personal lexicon, much like adding words to a spellchecker, and reprocess the document.

Phase 2: Inserting Anchors into the Document(s)

After analyzing the document(s), WebAnchor attaches hypertext anchors to each of the significant words and phrases within the document. These anchors are linked from the Index Pages to the source documents. Within the source documents, each separate occurrence of a significant entry is linked with circular links that allow you to move from one instance of that entry to the next.

Circular links are created for every **exact** match of a key word or phrase. For example, every indexed instance of the word "adoption" will be linked to every other indexed instance of "adoption." However, every indexed instance of "North American rivers: flooded" will be linked to every other indexed instance of "North American rivers: flooded," but not for "North American rivers: navigable."

Phase 3: Creating Index Pages

Once all of the anchors are inserted within the document(s), WebAnchor creates Index Pages that are similar to back-of-the-book style indexes. WebAnchor takes the destination anchors (HREFs) that it finds in the document(s) and builds Index Pages that will link these destination anchors to their source anchors (HREFs). Each entry on the Index Pages is linked to the first significant occurrence of that entry in the document(s). At the top of each Index Page is a Quick Index that lists the individual letters of the alphabet. Clicking on a letter will take you to that particular section of the index.

Note: When WebAnchor processes a single document, it creates Index Pages that are exclusively for that one document. If processing multiple documents, WebAnchor merges the entries together and creates a single set of Index Pages for the entire set of documents.

Starting the Program

To start WebAnchor, open the Iconovex folder or program group and click or double-click on the WebAnchor icon. You will be taken immediately to the main screen.

The Tool Bar

Many of WebAnchor's menus and file choosers may be accessed from the Tool Bar at the top of the screen. When you point your mouse at one of the buttons, the functionality of the button is explained in the status bar at the bottom of the screen.



Clicking on this button will "Create a New Project" and will return all of the previous settings to the defaults.



Clicking on this button will "Open an Existing Project." A file chooser will appear so that you can select the project file you wish to open.



Clicking on this button will "Save the Current Project." A file chooser will appear where you can name the project file and decide where to store it.



Clicking on this button will open the Lexicon Control dialog box where lexicons can be loaded for the current project.



Clicking on this button will open the Page Layout dialog box where you can alter the appearance of the Index Pages.



Clicking on this button will open the RTF Conversion Options dialog box where you can change the way WebAnchor converts from RTF to HTML.



Clicking on this button will open the Settings dialog box where several index content and output options can be determined.



Clicking on this button will open the table of contents listing for the Help File.



Clicking on this button will display the topics list for the Help File so that searches can be made on specific topics.

Choosing Files to be Indexed

The main screen presents you with a file selection window where you can indicate which files you wish to index.

The directory browser on the left allows you to search your system's drive and locate the directory that contains the files you wish to have indexed. If you wish to index all documents contained within a specific directory, simply choose the directory and click on the **Dir** arrow button. All RTF and HTML files contained within that directory will be indexed. If you only wish to index certain files within a directory, double-click the desired directory and as list of files contained within the directory will appear in the Files box at the center of the screen. You may display any file type by selecting the appropriate type in the File Types box. Though you may select all files (*.*), WebAnchor will only properly index HTML and RTF files. Please note that HTML files can have any extension, but RTF files must have an `.RTF` extension to be properly converted.

Individual files can be selected one at a time by first highlighting them and then clicking on the single right arrow button or by double-clicking on the filename. To select all files in a directory for indexing, click on the triple right arrow button. Files can be deselected using the same methods with the single and triple left arrow buttons.

WebAnchor is capable of processing single documents or multiple document sets. Document sets can be composed of files located anywhere on your local system, in the same directory or spread out across many directories.

*** NOTE ***

If you are processing RTF files, WebAnchor will convert those files to HTML and store them in the same directory with the original RTF files. WebAnchor then overwrites that HTML file and replaces it with another HTML file with hypertext links inserted throughout. Your original RTF file is not altered or moved. If you are processing HTML files, WebAnchor overwrites the files while maintaining any previous links that may be in the source documents and adds its own links. The new HTML file will then contain all previous and current links. If you ever wish to remove all index entries from your processed documents, you can do so by using the Remove Index feature on the *Tools* menu. See [Removing Index Entries](#).

Output Settings

The *Settings* dialog box contains several options which may be set by the user that effect the content and the location of the output generated by WebAnchor. To open the *Settings* dialog box, select **Options–Settings** or click the **Settings** button on the tool bar. Each time you run an WebAnchor project, you will want to select a new output directory, and perhaps a new filename. You may also wish to adjust the [Index Level](#), depending on the documents that you are indexing. As you become more comfortable with WebAnchor and find that you would like to further customize the output, you can enable/disable such features as personal and geographical names and unknown words. These settings are explained in detail below. If you decide to use different settings for different projects, you can save any group of settings as a project (see [Using Projects to Save Settings](#)).

More:

[Index Level](#)

[Index Content](#)

[Choosing a File Name and Output Directory for Your Index File](#)

[Creating an Unknown Word List](#)

[Setting the HREFs](#)

Index Level

The Index Level setting allows you to determine the level of detail in the index. The levels of detail span from Level 1 (most comprehensive) to Level 6 (most selective). Level 1 produces a very comprehensive index with almost every significant word indexed. Level 6 produces an index that includes only the most significant terms within the document. Levels 2-5 represent relative increments between Level 1 and Level 6. A typical Level 1 index will include some fairly mundane nouns and phrases such as "toll-free phone number" and "organizational meetings." A Level 6 index will generally include only nouns and noun phrases which indicate a relatively high degree of specialization or complexity such as "limited liability insurance companies." You can choose any of these levels by clicking on the corresponding radio button.

Please note, however, that indexing value is not determined solely on the basis of the words in the indexed term, but also on the basis of the context in which those words were located. Thus, identical index entries can sometimes be assigned different values because of context.

Index Content

Personal Names

By default, WebAnchor will recognize most personal names when they stand alone. With the default setting, you might see the following entries in an index:

- Bailey, F. Lee
- Parkinson's disease
- Schuster, Captain William
- Gascomb, French artist Henri
- Danny Glover

The Personal Names checkbox works much the same way as the Geographical Names checkbox, except that it pertains exclusively to recognized personal names. Where a personal name is recognized as such, it will be rotated around the last name, as in "Bailey, F. Lee." Titles and descriptors, when recognized, will be included and rotated appropriately with personal names. Single names standing alone in isolation (Bob or Wilson) will generally not be extracted with or without this checkbox marked.

Problems can occur, however, when any element of the name is also a common noun. The system will treat the common noun first, unless instructed (through a user lexicon) to do otherwise. Thus, you may get some entries like "Danny Glover" in the example above. "Glover" is here being treated as a common noun and therefore will not be rotated. The solution to this is to enter "Glover" in a user lexicon as a last name (if you feel confident that it is not used in the same document(s) as a common noun).

Geographical Names

The Geographical Names checkbox affects only those geographical names that refer to geopolitical entities such as cities, states, and nations. When the Geographical Names checkbox is unmarked, these names will not be indexed where they appear in isolation in a sentence ("Henry IV went to France.") They will, however, be indexed when associated with another indexable term ("Mrs. O' Leary's cow is said to have caused Chicago's Great Fire.")

Geographical names which refer to topographical features - lakes, mountains, etc. - are not generally effected by this checkbox and are usually indexed. Thus, with Geographical Names unchecked, you might get the following entries in the index:

- Colorado Supreme Court
- Greater Baltimore Chamber of Commerce
- Chicago's great fire
- Common Wealth Nations of Europe
- Southeast Conference State
- Mount Rushmore
- Lake Erie

When the Geographical Names option is checked, your index might include entries such as "Utah" and "Singapore" as well as the entries in the first set of examples.

Where geographical names appear in larger noun phrases or in proper nouns, as in the first set of

examples above, they will still be extracted in that context, just as they would be with the default settings. It should also be noted that this option is entirely dependent on the system recognizing the word as a geographical name. Where the system does not recognize a geographical name as such, that name will be treated more generally as a proper noun, and may therefore show up in your output even if you have not checked the Geographical Names checkbox.

Compound Word Length

For WebAnchor's purposes, a "compound word" is several words that together take on a single meaning and therefore act as one word. An example of a compound word is "control tower." These words do not have much value individually, that is, you would probably not want "control" and "tower" to appear separately as part of a hyperlinked entry in an WebAnchor document. But when placed together, these words take on a more important meaning that you may wish to include in your index. The accuracy of your output increases when you enable Compound Words because WebAnchor will look for compounds, defined in the primary lexicon and in any user lexicons you have created, before choosing what to extract for output.

By default, WebAnchor will search for any compounds, containing up to three words, that are defined in the primary lexicon or loaded user lexicons. A Compound Word Length of "1 Word" will disable the Compound Words function. A value ranging from "2 Words" to "8 Words" will instruct WebAnchor to search for compounds containing as few as 2 words or as many as 8 words; as this value is increased, however, the indexing process becomes slower. The value you choose must match the length of the longest compound word in your lexicon(s) that you wish WebAnchor to consider during processing; for example, if you want WebAnchor to search for "limited liability insurance companies," the Compound Word Length must be set, at least, at "4 Words."

Note: Setting the Compound Word Length at more than 3 words will not necessarily provide a much more accurate index, depending on the documents you are processing, unless you have loaded a lexicon in which you have entered longer compounds.

Choosing a File Name and Output Directory for Your Index File

Before your files can be indexed, you must choose a name for the index file and a directory in which to store it.

In the Index File Name text box, click your cursor in the box and type in the desired output directory and file name. Clicking on the file chooser button will allow you to browse your system's directory structure in order to find the desired output directory. If no directory or filename are chosen, WebAnchor will name your file `INDEX.HTM` and store it in the `WBANCHOR\OUTPUT` directory.

Note: When creating new index files, remember to change the output directory and, if you wish, the index name. If you have any index files that are already present in the selected output directory, WebAnchor will warn you and ask if they are to be overwritten.

Creating an Unknown Word List

The Unknown Word feature is an essential element for refining your indexes. Any unknown words in your document(s) will make it more difficult for WebAnchor to process them and will force WebAnchor to make guesses about how to treat those words it doesn't recognize. Defining unknown words in a lexicon will insure quicker WebAnchor processing and more accurate indexes.

By placing a check in the Unknown Words box, you will prompt WebAnchor to produce a list of words that it doesn't recognize when a document or document set is run. If you leave this option selected, unknown words will be added to this lexicon every time you run a new document set and these new words will be merged alphabetically with the existing words. This list will take the form of an ASCII text list which can be opened and edited in LexEdit. The default name is simply `UNKNOWN`, and will be placed in the `WBANCHOR` directory. You can indicate another name if you wish. If you choose a different filename, you can supply only an 8-letter (or less) filename with no extension. This list can be viewed in any text editor, but it will not be sorted and may contain duplicates. Importing it in LexEdit will alphabetize it and eliminate any duplicates.

Once you have created an unknown word list for a given project and imported and edited it in LexEdit, you need to load the resulting lexicon (see [Loading Lexicons Loading Lexicons](#)) and rerun your project.

Setting the HREFs

HREFs are HTML tags that instruct your system to search for a chosen Internet address. By default, WebAnchor writes these HREFs in lowercase. However, some operating systems, such as UNIX, are case sensitive. If you plan to place the finished project on a computer which has a case-sensitive operating system, you must make sure that the HREFs in your project are written in the same case as the filenames and directory names on the host computer.

Generating the Index

Once you have selected the appropriate settings in the Settings window, you are ready to generate an index for your document(s). If you wish to save your settings as a project, choose **File—Save Project As**.

To create an index, simply click on the **Generate Index** button and let WebAnchor do the rest. WebAnchor will analyze your document(s) and build Index Pages for them.

More:

[Removing Index Entries](#)

Removing Index Entries

You can remove index entries from processed documents as easily as you inserted them. If you have saved a project that contains a list of files from which you wish to remove index entries, you can simply open the project and make sure that all desired files are in the Files to be Indexed box. If your files are not saved in a project, you can find the directory containing the desired files and click those files into the Files to be Indexed box. Now choose **Tools—Remove Index** and WebAnchor will automatically remove all entries from the files listed.

Viewing the Results

In order to view the output generated by WebAnchor, start your browser and point it at the Index file in the directory in which you saved the file. (The default name for the Index file is `INDEX.HTM`.) You can now use your browser to explore the hyperlinked index entries and the Index Pages.

Using Projects to Save Settings

A project is any set of options that you choose to give a name and save as a configuration file (.IIX). Saving a configuration file is useful if you have found a particular configuration that you like and wish to use repeatedly for certain documents or document sets, without trying to memorize that configuration or resetting it each time. Projects save all user settings such as page layout options, RTF conversion options, output settings, loaded lexicons, and, if you wish, a file list.

More:

[Creating a New Project](#)

[Saving a Project](#)

[Opening a Project](#)

Creating a New Project

To create a new project, choose **File—New Project** or click on the **New Project** button on the tool bar. This will allow you to select a new group of run-time options and other settings to save as a project for your processed file(s). You may create a new project for each file or group of files that you index.

Saving a Project

Once you have created a project, you may save those settings for use in future projects. To save a file list along with your project, use the directory browser on the main screen and click all of the desired files into the Files to be Indexed box. These files will be the ones saved with the project. The project will be saved as a configuration file (.IIX). Select **File—Save As** or click on the **Save Project As** button on the tool bar and, in the dialog box, type in a filename and use the directory browser to select a directory in which to save the file. If you make changes to an existing project and re-run WebAnchor, you can save those changes by choosing **File—Save Project**. WebAnchor will automatically save the project while maintaining the same filename.

Opening a Project

Once you have created a project, this project can be reopened and used again. If, for instance, you want to add new files to the project, select **File—Open Project** or click on the **Open Project** button on the tool bar and select the `.IIX` file that you saved. When you bring up a project with all of the saved settings, select the new files that you wish to process using those settings and click **Generate Index**. You may also apply projects to other document sets.

Index Page Overview

After all of your documents have been processed, WebAnchor creates an Index Page . This page will be in a "Home Page" format where there is a graphic at the top of the page, a title or heading for the page, and various buttons including a **Next**, **Previous**, **A-Z**, and **Home** button.

The Index Page acts as a "Home Page" for the indexed document(s). Each letter of the alphabet has its own Index Page. For example, all of the "A's" have their own Index Page, all of the "B's" have their own Index Page, etc. From each Index Page, you can jump to the next and previous Index Pages. That is, from the "A" page, clicking on the **Next** button will take you to the "B" page and clicking on the **Previous** button will take you to the "Z" page. You can jump to any Index Page by clicking on the desired letter in the Quick Index. From the bottom of any Index Page, you can click on the **A-Z** button and be returned to the top of the "A" page. If you click on the **Home** button, WebAnchor will transport you to the home page location that you specified (see [Changing Home Page Location](#)). All of the features of the Index Page are described below.

The Heading Image is the graphic that appears at the very top of the Index Page. This graphic can be any size that you choose. The image will not be hotlinked or clickable; it is only for decorative purposes. This image can be eliminated all together if desired.

The Heading Text is what you choose as a heading for your index. This heading is centered directly below the Heading Image.

The Quick Index is a listing of each individual letter of the alphabet. The Quick Index appears at the top of each Index Page just before the alphabetical listing of key words and phrases for a given letter. If the index is segmented, you can simply click on the beginning letter of a word that you wish to search for and you will be jumped to that particular section of the index.

The alphabetical listing of key words and phrases begins just below the Quick Index on each Index Page. Each key word or phrase will be directly linked to its first significant occurrence in the document(s). Once you are at the first occurrence of your chosen word, you will be able to jump to every other occurrence of that word via circular links. Circular links are marked with icons and appear within the document(s).

The Previous Button Image is the hot linked image that acts as the Previous button. The Previous button takes you to the previous Index Page. For example, clicking on the Previous button from the "B" page will jump you back to the "A" page.

The Next Button Image is the hot linked image that acts as the Next button. This button performs the same function as the Previous button, only in the opposite direction. Clicking on this button will take you to the next Index Page.

The Index Button Image is the hot linked image that acts as the A-Z button. Clicking on the A-Z button will take you back to the beginning of the Index Page (the "A" page). For example, if you are on the "J" page, clicking on the A-Z button will take you to the Quick Index at the top of the "A" page.

The Home Button Image is the hot linked image that acts as the Home button. Clicking on this button will take you to the address that you have specified as your home page.

Page Layout Options

The *Page Layout* dialog box contains layout options you can use to change the appearance of the Index Page created by WebAnchor. These options must be set before WebAnchor is run. If you wish to change the appearance of the Index Pages after it has been generated, you must make the appropriate changes in the *Page Layout* dialog and then rerun WebAnchor again or you must edit the HTML page directly. If you are not familiar with HTML, please refer to [Guides and Reference Work on HTML](#).

More:

[Changing Images](#)

[Changing Heading Text](#)

[Changing Home Page Location](#)

[Changing Index Styles](#)

[Eliminating All Graphics](#)

[Segmenting the Index](#)

Changing Images

You can change any of the images that will be placed on the Index Pages by clicking your mouse in the text box and manually typing in the directory and filename of the image you wish to substitute or by clicking on the file chooser button to the right of the text boxes and locating the image file in your directory structure. You can use the images (Index Styles) that we have provided or choose images of your own. When you choose an index style, all images will be automatically filled in for you. If you do not want any graphics at all on your Index Pages, you can check the Text Only box .

Changing Heading Text

You can change the heading of your index by clicking your cursor in the text box and entering a new heading. By default, the heading text will be in Heading 2 style, boldface, and centered. You will need to edit the HTML file directly to change the size or style of the heading; only the words themselves can be changed at the *Page Layout* dialog.

Changing Home Page Location

To change the Home Page Location, which is linked to the **Home** button, simply click your cursor in the text box and enter any valid Web site address that you wish. You must enter a valid Internet address here for the **Home** button to work properly. By default, there is a placeholder address in this box and the **Home** button will not work unless you replace it with a valid one.

Changing Index Styles

WebAnchor offers several different styles for the Index Page. The Index Style box allows you to choose from the various icon sets we have provided. Each set of icons includes six images: (1) Heading Image, (2) Previous Button Image, (3) Next Button Image, (4) Index Button Image, (5) Home Button Image, and (6) Circular Link Image. Once a style is chosen, all of the image boxes will automatically be filled in for you. For example, if you choose the Golf style, the appearance of the Index Page will appear with a "golf" motif, the Heading Image being a golf green and all coordinating buttons having a "golf-related" theme. You can, however, choose an index style and then substitute individual images with your own. To view the icon sets in your browser, you can open the `VIEW.HTM` file in the `WBANCHOR/TEMPLATE` directory. A Template Master will come up where you can view each set of icons individually or see how they would appear on an Index Page.

Eliminating All Graphics

If you do not want any graphics on the Index Page, place a check in the Text Only box and the page will appear in an all-text form. There will be a hyperlinked @ for the circular links within the document and all graphic buttons on the Index Page will be replaced with hyperlinked text. This option is particularly useful for end-users who may have slower modems or simply want a straight index with no graphics. By default, this box is unchecked.

Segmenting the Index

WebAnchor automatically segments the entries from each letter of the alphabet into its own file. Clicking on the **Next** button from the "J" page will take you to the next file which would be the "K" page. If you uncheck the Segment Index box, your Index Page will be one continuous list of entries for all of the letters.

Lexicon Overview

You can shape the specific content of WebAnchor's output by raising or lowering the indexing values of selected words in a user lexicon. User lexicons can be created with LexEdit, a lexicon maintenance tool supplied with WebAnchor. Lexicons are used to assure that terms you consider important are indexed and that terms you consider unimportant are suppressed. If you are using many technical terms, you may wish to enable Unknown Words (see [Output Settings](#)) and then import the resulting list of unknown words into a new user lexicon to edit.

Launching LexEdit

There will be a separate LexEdit icon included in the Iconovex program group where you can start the program; or, you can launch LexEdit from within WebAnchor by choosing it from the **Tools** menu on the main screen.

Loading Lexicons

Once you have created a user lexicon, you will need to activate it for the WebAnchor project that you wish to use it with prior to running that project. Choose **Options—Lexicon Control** or click on the **Load Lexicons** button on the tool bar to open the Lexicon Control dialog. A file selection window will appear that closely resembles WebAnchor's main screen. The Lexicon Control file selector works the same way as the one on the main screen. First access the directory in which you stored your lexicon(s) on the left, then choose the filename(s) of the lexicons that you want loaded. Finally, highlight the desired files and move them into the Loaded Lexicons box. Use the single right arrow or **Tab—Spacebar** to move individual lexicons into the list or the right triple arrow button to move all lexicons into the list. Lexicons will be searched in the order they appear in this box, so the first values located for any given word will be applied to that word. This means that values assigned in the first lexicon in the list will override values assigned to the same word in any subsequent lexicons, including the primary lexicon.

When a lexicon is created within LexEdit, it is given a (.DLX) extension. If you are concerned with processing speed, you can compile your user lexicon so it will be processed faster. LexEdit will maintain your (.DLX) lexicon and create the same lexicon, but with an (.LDB) extension. If you compile a lexicon to (.LDB) and wish to make changes to it, you must make the changes to the (.DLX) lexicon in LexEdit and then re-compile the lexicon again. Since (.LDB) lexicons are "read-only" files, you may not make direct changes to them. After compiling, you will have two lexicons with the same name, but with different extensions. Make sure to choose the (.LDB) lexicon when you are loading lexicons.

You may load lexicons created in earlier versions of LexEdit, but you will first need to convert them using the convert utility supplied with LexEdit 2.0.

Lexicons can be unloaded by using the left single arrow and triple arrow buttons. Unloading a lexicon does not delete it. It simply deactivates it so that it will not be used as a basis for indexing decisions.

Note: Before creating an index using a user lexicon, you must close LexEdit.

RTF Conversion Overview

WebAnchor looks to see if any of your files have an RTF extension and, if they do, converts them to HTML during the indexing process. When WebAnchor converts your RTF files to HTML, it tries to match the corresponding HTML styles to the RTF styles in your document(s). WebAnchor allows you to change the way that RTF styles, tables, and lists are converted to HTML through the RTF Conversion Options dialog.

HTML uses "tags" to mark certain styles. For example, the HTML tag for a Heading 1 style would be <H1> at the beginning of the heading and </H1> at the end of the heading. WebAnchor recognizes the following RTF style indicators and translates them to the appropriate HTML tags:

- Headings 1 through 6

- Body paragraphs

- Special characters (&, <, >, and non-breaking space)

- ISO Latin -1 alphabetic characters (ñ, è, ü, etc.)

- Character emphasis styles (bold and italic)

- Bullet lists

- Number lists

- Tables

Setting Conversion Options

You can change the way that WebAnchor converts RTF styles to HTML by choosing **Options—RTF Conversion**.

More:

[Changing Heading Conversions](#)

[Converting Tables](#)

[Setting Tab Size](#)

Changing Heading Conversions

In the RTF Conversion Options dialog box, there is a section entitled Convert RTF Style Names to HTML Headings where six different text boxes are shown. From this menu, you can change the way that WebAnchor will convert the headings contained within RTF document(s).

Each text box contains an RTF style name and the corresponding HTML headings that will be applied by default. "Heading 1" style in RTF is converted to "<H1>...</H1>" (Heading 1 style in HTML) and so on. If your RTF document(s) includes style names other than the default Heading 1 through Heading 6 - if, for instance, you have defined custom styles - you can enter those style names in the appropriate text box to indicate how you wish them to be converted.

Note: If the headings in the original document(s) are set off with manual styles rather than through the use of style tags, the RTF form of the document(s) will not carry the information necessary to make the conversion to heading tags in HTML.

Converting Tables

WebAnchor can convert RTF tables over to HTML in two ways: as a table or a bullet list. Many browsers, however, will not be able to read table tags; we provide the option of converting RTF tables to HTML as nested bullet lists so that all of your tables will be viewable by all browsers.

If you choose to have your RTF tables converted to bullet lists, WebAnchor will number the rows and then place the cell content of each row in a bullet list under the appropriate row number.

Setting Tab Size

HTML does not allow tabs. Therefore, all tabs marked in RTF documents must be replaced by non-breaking spaces.

In the Tab Size box, enter the number of spaces you would like to use for each tab replacement. A typical tab would consist of 5 non-breaking spaces. Due to different font sizes, tab spacing is not always accurate; therefore, we have the default Tab Size set at 0 where tab conversion will simply be ignored.

Getting the Best Entries

- Remember to spellcheck your documents before running WebAnchor on them, especially if you are not the author of the documents and cannot vouch for their condition. The spellchecker will not only catch misspellings, which, if uncorrected, will come up as "unknown" words and produce odd entries in your output, but will also help catch oddities introduced by some word processors when saving documents to RTF or HTML (spaces added in hyphenated words, etc.). In general, the better the condition of the original documents, the better the WebAnchor output will be.
- For the very best results on any document suite, we recommend that you do a preliminary run, make necessary adjustments, and then do your final run. In the preliminary run, you should create a list of unknown words and you should store all your configuration settings in a configuration file. Between runs you can then add the unknown words to a user lexicon using LexEdit and make any adjustments you feel are necessary for obtaining the best results. You may discover, for instance, that you need to raise the Index Level if your output is too large or if it includes entries you do not consider important for your purposes.

Known Problems

- Jumps to the wrong place in the document: We have discovered that some browsers - most notably some versions of Mosaic - have difficulty resolving links that include a reference to a name attribute. If your browser has this difficulty, you will notice it whenever you click on links that jump to a specific point within a segment rather than simply to the top of the segment. The symptom will be that you will jump to the correct segment, but not necessarily to the specific reference point. If you do not jump to the correct spot in the document, try hitting the Back or Previous button on your browser so you are taken back to the Index Page and click on the desired entry again. We have found that it will work, for the most part, on the second try.

Guides and Reference Work on HTML

Books:

HTML Manual of Style by Larry Aronson (Ziff-Davis Press)

Teach Yourself Web Publishing with HTML in a Week by Laura Lemay (SAMS Publishing)

The World Wide Web Complete Reference by Rick Stout (Osborne: 1996)

Sites:

<http://aeiveos.aeveos.wa.com/html/>

Some good references and tutorials for HTML, most supplied by NCSA, where Mosaic was developed.

http://info.med.yale.edu/caim/StyleManual_Top.HTML

Yale University's Style Manual for HTML; not an introductory manual, but contains good information for reasonable experienced users of HTML.

http://www.yahoo.com/Computers/World_Wide_Web/HTML

A good collection of discussions, utilities, and converters (also a fine starting point for Web exploration).

<http://WWW.Stars.com/>

The Web Developer's Virtual Library. Good mix of hard-core and beginner material.

<http://www.ski.mskcc.org:80/browserwatch/>

BrowserWatch, a good site to keep up on the doings and developments of your favorite Web browser software.

<http://info.cern.ch/hypertext/www/MarkUp/MarkUp.html>

More good general information and starting points.

<http://www.ietf.cnri.reston.va.us/home.html>

This is a bit more hard-core; it's the homepage for the Internet Engineering Task Force, some of the people who establish the actual protocols for the Internet.

http://www.yahoo.com/Computers/World_Wide_Web/Special_Interest_Groups_SIGWEB/

Special interest groups and discussion groups about HTML and other Web-related topics.

http://www.w3.org/hypertext/WWW/Daemon/User_3.0/

Information about HTTP and HTML in general, including the CERN hypertext guide.

<http://www.tucows.com/>

Ultimate Collection of Winsock Software: a good collection of software, including browser and utilities, for Windows-based surfing and development.

