



webstractor

Table of Contents

Webstractor Overview	2
The Browser	4
Adding Content to the Browser	4
The Browser — List	5
The Browser — Sub-toolbar	6
Deleting Web Pages in Browser	6
Unread Web Pages	6
The Editor	7
Adding Content to the Editor	7
The Editor — List	8
The Editor — Sub-toolbar	9
Text Selection	10
Text Editing pop-up window	10
Radar Overview	11
Document specific settings	12
Web page specific settings	12
Dashboard widget	13
Links Inspector	14
Find & Replace	15
Spotlight	16
Menu Reference	17

Webstractor Overview

A Webstractor document consists of two main sections, the **Browser** and the **Editor**. You can quickly switch between these two sections using the Browse/Edit button on the toolbar, or by selecting the relevant item from the View menu.

The Browser

The main function of the Browser is to navigate through web pages. Web pages you browse within Webstractor are captured automatically into your document. Unlike other web browsers, clicking on an entry in your browser history in Webstractor will display the web page as it was when you visited it, without reloading it again from the web site. This is the case even if you save your document and re-open it later. This powerful feature allows you to research a particular topic on the internet and be sure that all of that information is available to you at a later date even if the original web pages change (or become unavailable) in the meantime. You don't even need to be connected to the internet to browse your saved documents.

Webstractor can even store multiple versions of web pages that change over time.

Webstractor Radar

After capturing web pages into the Browser, Webstractor will periodically check if their original versions on the internet have changed. You can see at a glance which captures have become out-of-date.

Webstractor can be told to automatically add to your document a new instance of a web page when it changes. You can also set a series of keywords to tell Webstractor to only capture a new version when something important to you is mentioned on a changed web page.

The Editor

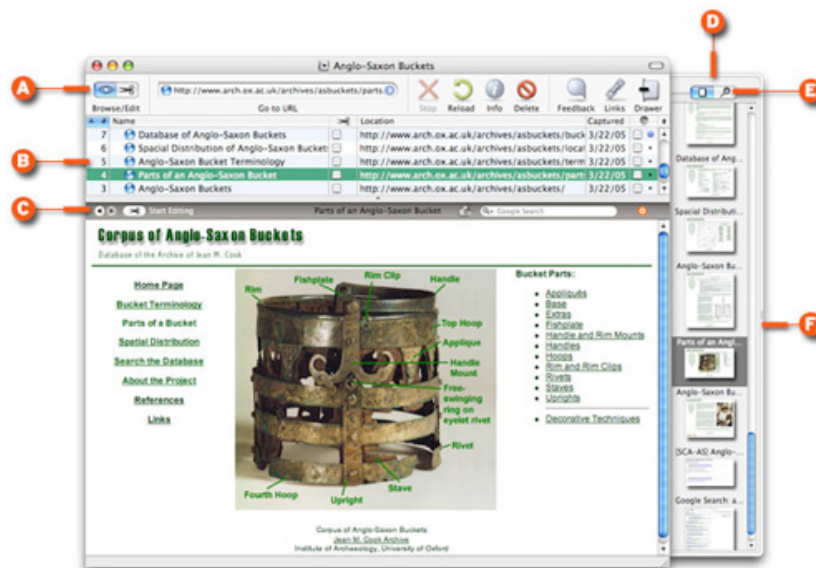
Whilst Webstractor's ability to capture and store web pages in the Browser is great, there are times when you want greater control over the appearance of the captured content. By transferring a web page to the Editor, you are able to quickly do the following:

- Remove unwanted text, advertisements, graphics and backgrounds
- Reformat the text using your own fonts, sizes and colors
- Add your own content to the page
- Prepare multiple web pages for seamless printing
- Export your findings to cross-platform PDF files for sharing with non-Webstractor users

It is easy to transfer a web page to the Editor side. Whilst looking at it in the Browser, click the 'Start Editing' button. After a few moments you will be automatically transferred to the Editor and can start editing the web page using familiar word processing methods. And since you are working on a copy of the captured web page you can still go back to the original in the Browser and interact with it as a web page.

Webstractor Document User Interface

Both the Browser and Editor sections contain the following common user interface elements.



(A) Main Toolbar

Contains commands that are common to both the Browser and Editor sections.

(B) Entry List

Both the Browser and the Editor display their content as a list of entries. As well as allowing access to the entries in each section, this list also

displays additional information specific to the current section. Clicking on an entry will take you to its content in the document window. You can select multiple entries by either shift-clicking or command-clicking them.

(C) Sub-Toolbar

Contains commands that are unique to the Browser and Editor sections.

(D) Drawer

The drawer contains the Thumbnail view of the content and the Find and Replace options. You can hide and show the drawer by selecting Toggle Drawer from the View menu.

(E) Thumbnail/Find and Replace Switcher

Switches the Drawer between the Thumbnail and Find and Replace views.

(F) Drawer Resizer

Drag this to make the Drawer wider or narrower. This will allow you to see larger thumbnail images, or more context for your Find results.

Find and Replace

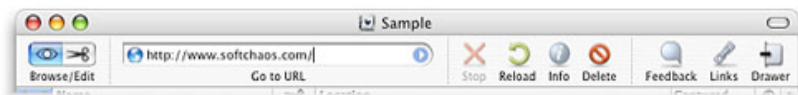
The Find and Replace feature allows for all textual content in either mode (Browser or Editor) to be searched instantaneously. Search results are displayed in the drawer, and clicking on a result will take you to the actual content.

Thumbnails

In either the Browser or Editor, you can see a miniature version of the content in that section. Clicking on a thumbnail will take you directly to that item in the main window. In the Browser, multiple entries can be selected by either shift-clicking or command-clicking the thumbnails.

Main Toolbar

The Main Toolbar is common to both the Browser and the Editor, and is used to switch between the two sections as well as controlling and loading entries, setting general properties of the selected entries in the list, and other functions.



Browse/Edit Toggles between the Browser and Editor sections of your document.

URL field Type a web page address here and press Enter to load that page in the current section.

Stop Stop loading web page.

Reload Reload the selected entry in the list.

Info Display the properties of the selected item(s) in the list.

Delete Delete the selected entry or entries from the list.

Feedback Allows you to submit a bug report or a feature request to Softchaos.

Links Inspector Display the Links Inspector window.

Drawer Shows and hides the drawer containing Thumbnails and Find & Replace.

The Browser — Introduction

The Browser allows you to navigate through web pages in a way that will be familiar to users of other web browsers. Web pages you browse within Webstractor are captured automatically into your document. Unlike other web browsers, clicking on an entry in your browser history in Webstractor will display the web page as it was when you visited it, without reloading it again from the web site. This is the case even if you save your document and re-open it later. This powerful feature allows you to research a particular topic on the internet and be sure that all of that information is available to you at a later date even if the original web pages change (or become unavailable) in the meantime. You don't even need to be connected to the internet to view and browse the captured web pages in your saved documents.

If you need to ensure that you have the latest version of a web page that is in your document, click the Reload button on the main Toolbar. If the page has changed since the last time you visited it, Webstractor will create a new entry for it in its list, whilst keeping the version you currently have. In this way Webstractor can maintain different versions of the same page.

Adding Content to the Browser

You can use Webstractor as a web browser, and it will capture every web page you visit into the current document. Start browsing by either typing in a URL (web address) into the "Go to URL" field in the main Toolbar, or by selecting a bookmark from the Bookmarks menu.

There are also other ways you can add content to your Webstractor documents:

From the internet:

While looking at a web page in Webstractor, hold down the ⌘ (command) key while clicking a link to capture it without leaving the current page.

Drag a link from a web page into either the list or the main section of the Browser.

Use the Links Inspector window to automatically browse and capture one or more web pages linked from the currently viewed page.

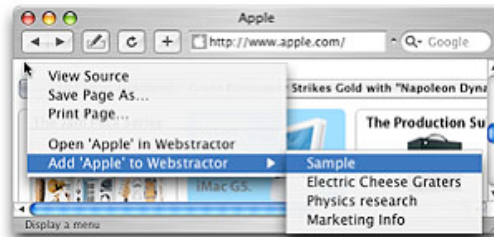
From Disks:

Drag an HTML file into either the list or the main section of the Browser. You can also drag an HTML file directly to the Webstractor icon, which will create a new Webstractor document containing it.

From another web browser:

Drag a URL from the other browser's URL entry field, an underlined link, an Internet bookmark/favorite or history list into either the list or the main section of the Browser.

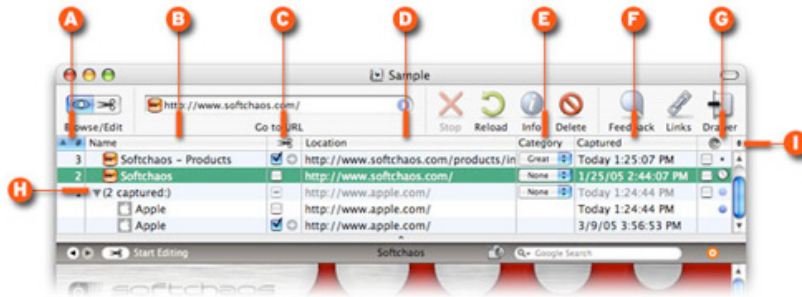
From a Contextual menu in Safari or OmniWeb:



Content can be also added from Webstractor's additions to the Contextual Menus provided by Safari and OmniWeb. From these menu items, the current web page in Safari or OmniWeb can be opened in a new Webstractor document, or added to one of your recently opened Webstractor documents.

In either case, Webstractor will be launched if it is not already running.

The Browser — List



The list shows all of the web pages that have been captured in the document.

(A) Number Column

When the list is sorted by this column, you are able to rearrange the list items manually via drag-and-drop. When sorted by any other column, the order is dependent upon the information in the column. The order of the items in the list is used when you want to add multiple selected items to the Editor in one go. They will be added in the same order that they appear in the list (from the top).

(B) Name Column

This column displays the icons and titles of the captured web pages. You can double-click a name to edit it.

(C) Editable Column

This column contains a check box that allows you to choose whether or not an editable version of the web page should exist in the Editor. When checked, an encircled arrow button will appear next to this check box to allow you to quickly switch to that item in the Editor.

(D) Location Column

This column displays the internet address (URL) of the item.

(E) Category Column

This column allows you to assign your own categories to items in the list. Click on the pop-up menu and either choose an existing category or define your own ones. Any categories you define are stored on a per-document basis. E.g. If you were researching where to go for your next vacation, you could give each web page relating to the same country the same category. Later you could sort the list by category and have all of the web pages with the same country next to each other for easier perusal.

(F) Captured Column

This column shows the date and time the web page was captured.

(G) Source Status Column

This column shows whether the web page has changed since it was captured into the document. Possible states are as follows:

- The captured version is up-to-date.
- 🕒 The captured version is out-of-date.
- ⋮ Webstractor is checking for changes.
- ⬇️ Webstractor is capturing the web page.
- 🔵 The web page has been captured but has not been read yet.
- ⚠️ Webstractor could not determine the state.

Clicking the check box in this column turns on Radar automatic capture for the web page. This will automatically capture a new version of the web page when Webstractor notices that it has changed.

(H) Individual Web Page History

The availability of multiple versions of a page is indicated by the presence of a disclosure triangle. For an entry with multiple versions, if the disclosure triangle is closed, the most recent version of the entry is displayed in both the list and in the Thumbnail drawer. Clicking the triangle to open it will list all the versions of the web page.

(I) Column selection menu

Clicking this menu allows you to choose which columns are shown in the Browser list.

The Browser — Sub-toolbar



(A) Back

Steps backward through the web pages you have visited.

(B) Forward

Steps forward through the web pages you have visited.

(C) Start Editing

Adds the selected web page or pages to the Editor (if not already added) and automatically switches to the first selected page within the Editor.

(D) Web Page Title

Show the title of the current web page.

(E) Document Properties

Displays properties which affect the entire document. See Radar for more information about settings here which affect the Browser.

(F) Google Search

Type a search term or phrase here to perform a web search using Google's search engine.

(G) Mark Page for SnapBack

Marks the current web page as the one to SnapBack to (see below).

(H) Page SnapBack

After you have marked a page for SnapBack and visited other web sites, click this button to return to the marked page. This allows you to follow a chain of links from one site and easily return there later. The SnapBack information is saved with the document.

(I) Loading Indicator

This animating indicator will appear while the current page is loading.

Deleting Web Pages in the Browser

Web pages can be deleted from the Browser in the following ways:

To delete an entry with a single version:

Select the entry in the list or the Thumbnail drawer and press the delete key.

To delete an entry with a multiple versions:

Collapse the title entry (the entry with the disclosure triangle), select it, and press the delete key.

To delete specific versions of an entry with multiple versions:

Expand the title entry (the entry with the disclosure triangle) to show all the versions of the entry, select the version(s) to be deleted either in the list or in the Thumbnail drawer, and press the delete key.

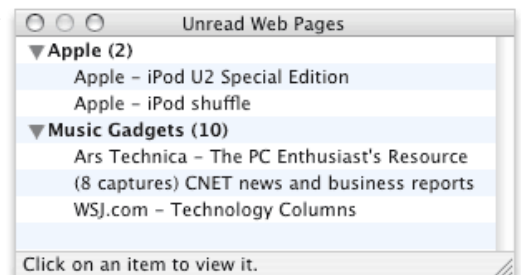
In any case, you can either click the Delete button in the Main Toolbar or use the backspace key.

Unread Web Pages

In Webstractor, it is possible to capture web pages into your document without having ever seen them. These pages may have been captured by using the Links Inspector to browse multiple pages, automatically captured using Radar, or simply as a result of clicking web page's link while holding down the ⌘ (command) key.

This window shows a list of every Webstractor document you have open, with a number indicating how many unread captured web pages each one has. In the example here, you can see that while the Music Gadgets document has 10 unread web pages, 8 of those are multiple captures of the same location at different times.

You can use this window to quickly navigate to those web pages you have not yet read by simply clicking a specific web page in the list. It will become "read" immediately, and the Unread Web Pages window will reflect this.



Dashboard widget

In OS X 10.4 or later, Webstractor also includes a widget with similar functionality. See the Radar section entitled "Dashboard widget" for more information.



The Editor — Introduction

Whilst Webstractor's ability to capture and store web pages in the Browser is great, there are times when you want greater control over the appearance of the captured content. By transferring a web page to the Editor, you are able to quickly do the following:

- Remove unwanted text, advertisements, graphics and backgrounds
- Reformat the text using your own fonts, sizes and colors
- Add your own content to the page
- Prepare multiple web pages for seamless printing
- Export your findings to cross-platform PDF files for sharing with non-Webstractor users

It is easy to transfer a web page to the Editor portion of your Webstractor document. Whilst looking at it in the Browser just click the 'Start Editing' button. After a few moments you will be automatically transferred to the Editor and can start working with the page using familiar word processing techniques. And since you are working on a copy of the captured web page you can still go back to the original in the Browser and interact with it as a web page.

As well as web pages from the Browser, you can also drag-and-drop Microsoft Word, RTFD, RTF and text files to the Editor's list of entries. The files added to the Editor are imported into the Webstractor document, and subsequent editing of their content within Webstractor does not effect the original files, nor is the continued presence of the originals necessary.

Adding Content to the Editor

Content can be added to the Editor in the following ways:

When in the Browser:

Click the Start Editing button whilst browsing the web page you want to edit.
Click a checkbox in the Editable (scissors) column.

When in the Editor:

Enter the URL (Web address) of the web page in the Add URL field in the Main Toolbar and press Enter.

From Disk:

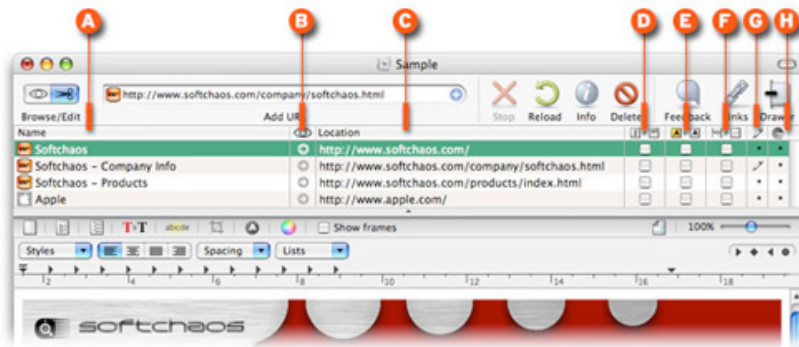
Drag a file (or files) from Finder into the list of entries in the Editor.

From other applications:

Drag a URL from any other browser's URL entry field, underlined link, Internet bookmarks, or history entry into the list of entries in the Editor.

The order of the items in the Editor list determines the order in which they appear in the continuous editable document. Drag and drop one or more entries in the list to a new location in the list to change this order.

The Editor — List



The columns in the list are used as follows:

(A) Name

This column displays the name of the entry and any associated icon. The name can be changed by double-clicking on it. Changing the name of a file-based entry will only affect the entry in the Webstractor document, and does not rename the original file. The name of the entry is used when generating a Table of Contents.

(B) Browse

If an entry in the Editor has a captured version in the Browser, an arrow button will be shown here. Clicking the arrow will take you to the captured version in the Browser. Note that the Browser will always show the original version and does not reflect any changes you have made to the web page in the Editor.

(C) Location

This column displays the address of the original source, either on the Internet or on disk.

(D) Fit To Page

Many web pages use specific widths for items that are within multiple columns. By default, Webstractor will respect these widths when displaying such web pages in the Editor. Turning on the checkbox in this column lets Webstractor override this setting to try and fit the columns to the full width of the document's paper size.

Note: This checkbox is inactive for non-web page based entries.

(E) Show/Hide Background

Many web pages use background images and background colors to make their pages more appealing in a web browser. By default, Webstractor will respect these backgrounds in the Editor. Turning on the checkbox in this column makes Webstractor hide such backgrounds and use white instead. Doing this helps provide a more consistent look to a Webstractor document if it contains information from many web pages with different backgrounds. It can also improve the look of the printed version (and save ink). For web pages that use a dark background, you may also want to use the Strip Text Color button so that light text becomes visible against a white background.

Note: This checkbox is inactive for non-web page based entries.

(F) Continue From Previous Entry

By default, every web page that you add to the Editor will start on a new page within the document. Click the checkbox in this column if you want an entry to continue on the same page as entry preceding it.

(G) Edit State

A pencil appears here if you have made any changes to the entry since it was first made editable.

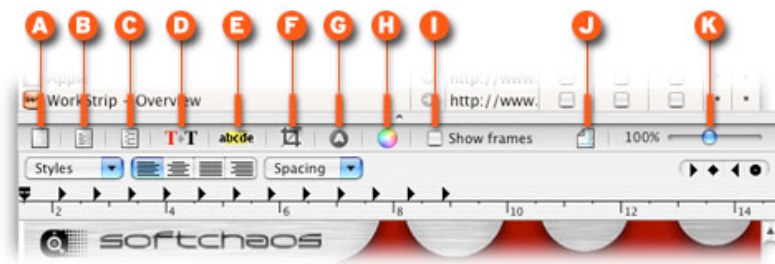
(H) Source Status

This column shows whether the original file or web page has changed since it was captured into the document. Possible states are as follows:

- The captured version is up-to-date.
- 🕒 The captured version is out-of-date.
- ⚠️ Webstractor could not determine the state.

The Editor — Sub-toolbar

The Sub-toolbar contains a set of buttons that assist the editing process.



(A) Insert Empty Entry

This button inserts a blank entry before the selected entry. On creating the entry, the text editor window will be invoked automatically, allowing you to start typing right away.

(B) Table of Contents

This button inserts a simple Table of Contents for the selected entry and those that follow it. Webstructor uses the contents of the text in the Name column to name each entry in the Table of Contents. Once it has been inserted, you can edit the contents of this table like any other item in Webstructor.

Note: Because it is editable, the Table of Contents does not automatically update itself to reflect later changes to the document. If you later wish to replace it with up-to-date information, you will need to Reload it.

(C) Bibliography

This button appends a Bibliography at the end of your edited pages. Webstructor uses the Name, Location, and date of capture to create the Bibliography. Once it has been inserted, you can edit the Bibliography like any other item in Webstructor.

Note: Like the Table of Contents, a Bibliography does not automatically update itself to reflect later changes to the document. If you later wish to replace it with up-to-date information, you will need to Reload it.

(D) Strip Text Color

This changes all of the text in the selection to black. If you have selected some text in the Editor, only that text is changed. If you have selected one or more entries in the Editor list, then all of the text in the selected entries will be changed. This feature is especially useful when used in conjunction with the Show/Hide Background option. If you would only like to strip text colors whilst printing, you may want to turn off the 'Print text in color' option in Webstructor's section of the Print dialog instead of using this.

(E) Highlight Selection

This adds a yellow background to the selected text. If any of the selected text is already highlighted, this will remove the highlighting from it. Use highlighting to draw attention to particular parts of your document.

(F) Crop

This removes all of the non-selected content within a web page. This is very useful when you only want to keep a small portion of a web page. Instead of having to manually delete all of the unwanted items yourself, just select the portion you are interested in and use this. Webstructor will keep the original formatting of the selection intact.

(G) Font Palette

This displays the standard Font palette which allows you to change the font, style, size etc. of selected text.

(H) Color Palette

This displays the standard Color Palette which allows you change the color of selected text.

(I) Show Frames

Turning this checkbox on will show layout boxes around all editable content. This can often make selecting items easier. These boxes will not be printed.

(J) Document Properties

This displays a dialog that allows you to set some options for the entire document. Items that affect the editable display of your document include:

- Page margins for your printed output.
- The scale of the document content within the page. To make everything smaller and fit more on a page, decrease the scale. To make everything larger and fit less on a page, increase the scale.
- Whether to include page numbers on each page.

(K) Zoom

This slider allows you to view your document at smaller and larger scales. Clicking the text next to the slider will toggle between the last non-100% zoom and 100%. When the zoom is not 100%, the text ruler becomes inoperable and is automatically hidden.

Text Selection

Due to the complexity of the underlying structure of many web pages, Webstractor has a number of different selection modes to simplify selecting parts of your document. These different selection modes can be set in the Selection Mode menu in the Edit menu. All the Marquee selection modes draw a marquee showing the area affected by the selection as it is being dragged out. When the selection mode is set to one of the Marquee selections, it is also possible to toggle between the different Marquee modes by pressing the appropriate keys outlined below.

All marquee modes are more readily understood with the Show Frames checkbox enabled in your document.

Each selection mode has different behavior and a specific cursor:

Layout Flow

This mode is the most similar to selecting text in a standard word processor. If the content is laid out in a column structure the content on the left hand column gets selected before selecting anything in the next column to the right.

+ ■ Marquee Touch

This mode selects items that intersect the rectangle that you drag out. While using this mode, clicking in any text box will select the entire content of that text box. To temporarily change to this mode, hold down the Option key while dragging.

+ □ Marquee Enclosed

This mode selects any items that are fully enclosed within the rectangle that you drag out. To temporarily change to this mode, hold down the Command key while dragging.

+ Marquee Hybrid

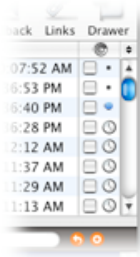
Whilst dragging out a selection, this mode behaves like **Layout Flow** for the text boxes that contain the start and end points of the selection. Any other boxes that are within the selection area will be selected as if the selection mode is Marquee Touch. **Marquee Hybrid** is the default selection mode.

Text Editing pop-up window

In the Editor, upon either typing at an insertion point or double-clicking on some text, the Edit Text pop-up window will be invoked. This window allows for text to be added, deleted, and otherwise formatted. When you have finished typing, click OK to apply your changes. To abandon any editing performed in this window, click Cancel.

Radar Overview

Webstractor periodically compares the web pages that you have added to your documents with their original versions on the internet. If a page has changed it will show that there is a newer version available by changing the icon in the Radar status column for that page.



In this example we can see that there are two web pages that are up-to-date and have already been read (•), one that has not yet been read (●) and five that are "out-of-date" or different from what is currently on the internet (🕒).

The default interval for this check can be changed in Webstractor's Preferences. You can also set a per-document interval in the Radar Automation panel of the Document Properties.

As well as being told when a web page changes, you can also get Webstractor to automatically capture new versions of web pages that have changed. You can enable this for pages in the Browser by turning on the checkbox in the Radar column. Once Webstractor has noticed that the web page has changed it will automatically capture a new version and mark it as being unread. By default it will keep all older captures, though you can change this behavior (see below). As such, you could set up a document with web sites that you visit often, set them to being automatically captured and let Webstractor tell you when any of them contain new information.

There are two sets of settings for Radar; those that are specific to the whole document and those that are specific to individual web pages. The document specific settings can be accessed from the File > Document Properties. and the web page specific ones by selecting a web page in the Browser and choosing Entries > Get Info.

Document-Specific Radar settings

To configure the settings that apply to the whole document choose File > Document Properties... and select the Radar Automation panel. The settings are as follows:

Automatic web page update checking

These settings control how often Webstractor will check for changes to the web pages in the document.

Disabled

Webstractor will not automatically check for any updates to the web pages. You can still manually do an update check using Entries > Check for Updates Now.

Use default interval

Webstractor will use the Radar update checking setting from Webstractor's Preferences.

Use custom interval

Enter your own interval specifically for the web pages in this document.

After Radar automatically captures a web page

These settings control what Webstractor will do once it has automatically captured a newer version of a web page.

Only keep the new capture if it contains a keyword

The newly captured web page will be discarded if it does not contain any of the user-specifiable keywords or phrases added to the associated list. This can help you focus your web browsing on a news site only when topics of interest are mentioned, for example.

Discard all older captures of the page

Older versions of the newly captured web page will be removed from the document's Browser side. This helps to keep your document from getting cluttered with out-of-date captures. Webstractor will never automatically remove a capture for which you have an editable counterpart.

...except unread captures

When this is on in conjunction with the previous option, only older versions that have been read will be removed. Once you have read a captured version it will be discarded the next time a newer version of the web page is captured.

Autocapture on by default

Turning this on will cause all browsed web pages to be set to being autocaptured when they change. By default, only pages that are manually marked for autocapture will do so.

Web page-Specific Radar Settings

To configure the settings that apply to an individual web page, select the page in the Browser's entry list and choose Entries > Get Info. You may also do this on multiple selected web pages to change all their settings at once. The settings are as follows:

Capture a new instance of this page when it changes

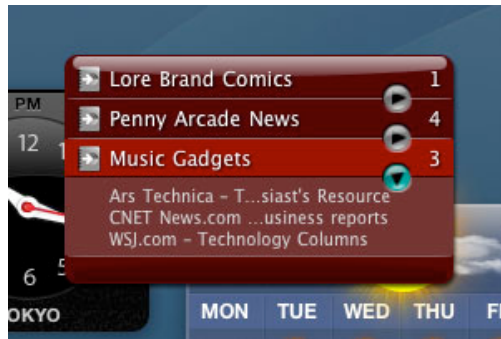
Turn this option on to get Webstractor to automatically capture the latest version of the web page when Webstractor notices that the original internet version has changed. This setting is equivalent to the checkbox in the Browser's Status column.

Custom change detection method

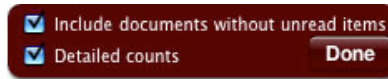
By default, Webstractor uses various information about the web page to determine if it has changed since the captured version. For some web pages this method might miss very subtle changes or think that a page has changed when it hasn't. By turning on this setting you can fine-tune how specific web pages are considered to have changed.

- **Simple** – Webstractor will query the web server to see if the page has changed. This option is the fastest but will be unavailable if the web server does not support this method.
- **Thorough** – Webstractor will check all of the data associated with the web page to see if anything has changed. This option is the slowest, but is the only way to detect changes to images on web pages when the text itself does not change.
- **Textual** – Webstractor will compare the text on the web page with the captured version whilst ignoring any non-textual content (e.g. images and animations). The slider below it controls how much the text must have changed before it is considered to be different. If you find that Webstractor is capturing too many versions of the page when very little seems to have changed, move the slider to the right. For example, many web pages will display the current date or different advertisements even if no other information has changed. The further the slider is to the right, the more textual changes are needed before the page is considered to have changed.

Dashboard widget



Webstractor includes a Dashboard widget if you are using Mac OS X 10.4 or later. This widget provides a bird's-eye view of your Webstractor documents, with particular attention to those containing unread web pages. By default, Webstractor's widget simply shows only those documents which contain one or more unread web pages, and shows a total count of how many different web page locations have been captured in each document. You can use it to quickly have a look at what web pages have been automatically captured using Radar, and go straight to viewing a capture with a single mouse click. Clicking on a specific web page title will dismiss Dashboard and make that web page visible in Webstractor, marking the web page as "read". Clicking on a document title will merely bring that document to the front. For general information about activating and managing Dashboard widgets, please refer to your system's "Mac Help".



The "back" of the widget.

Switching on the **Include documents without unread items** feature on the back of the widget does what you would expect. Here we have a document entitled "Anglo-Saxon Buckets" which does not have any unread web pages; with this feature turned off, this document is not shown on the widget.



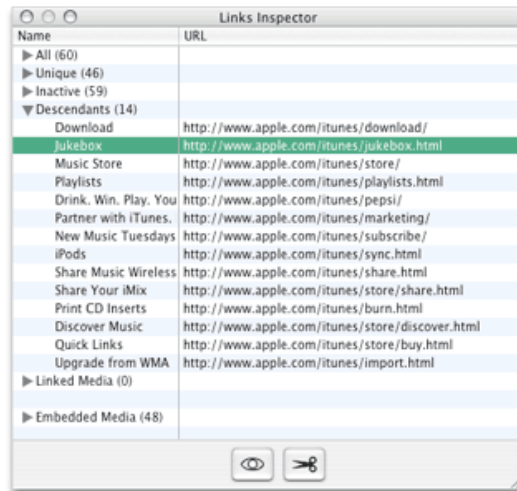
Documents containing unread web pages are shown with a triangular "disclosure" button which will allow you to inspect the specific locations that have been captured. Here you can see the titles of three different locations that have been captured in the "Music Gadgets" document.



Switching on the **Detailed Counts** feature allows you to see specific counts of how many captures each document contains, instead of merely how many unique web page locations have unread captures. In this case, by turning this feature on you can see that while there are only three locations in "Music Gadgets" which have unread captures, there are multiple captures of two of them.



Links Inspector



When you have one web page selected in either the Browser or the Editor, the Links Inspector will list all of the web pages that the page has links to. It will also list all of the images and other media that were used to display the web page. You can show the Links Inspector by choosing it from the Window menu, or by clicking the Links icon in a document's Toolbar.

Any of the items in the Links Inspector can be added to either the Browser or to the Editor:
Select and drag the links into an appropriate location in the Browser or the Editor.
Select the links and click either the Browse button (Eye) or the Edit button (Scissors) at the bottom of the Links Inspector window.

There are several sections within the Links Inspector:

All: The accumulation of all of the links on the web page (excluding Embedded Media).

Unique: This section shows all of the links within the selected web page that do not exist in any of the other web pages in the same side of the document (Browser or Editor). This is most useful after having visited several web pages from the same web site as it quickly allows you to see links that are only available from the selected web page.

Inactive: This section shows all of the links within the selected web page that have not already been captured in the same side of the document (Browser or Editor). This is useful when you want to ensure you have captured all of the related information in the selected web page.

Descendants: This section shows all of the links within the selected web page that are within the same web server directory (or deeper) as that web page.

For example, if the selected page is at

<http://www.utsr.com/info/about.html>

then the Descendants section would show these links:

<http://www.utsr.com/info/location.html>

<http://www.utsr.com/info/products/gizmo.html>

but it would not show these links:

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

<http://www.utsr.com/gadgets.html>

This is useful as often other web pages that are directly related to a web page are kept in the same directory structure on the web server; unrelated pages would be kept in a different directory structure. As such you can more easily locate related information and add it to your document.

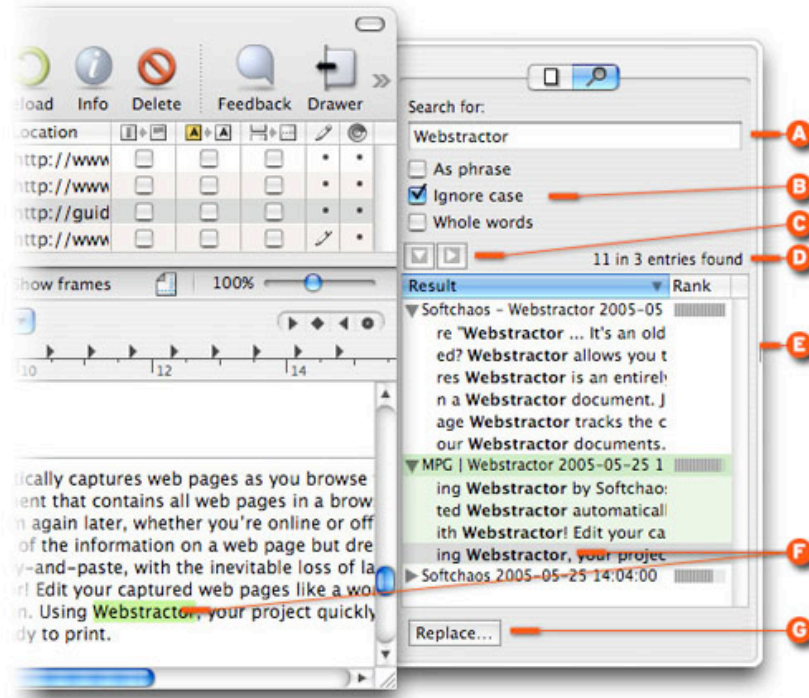
Linked Media: This section shows the links that point directly to recognized images, movies, Flash animations and PDF documents. This section will often be empty as many web designers will embed such items within separate HTML web pages (which will show up in one of the other sections).

Embedded Media: This section shows the images, movies, Flash animations and PDF documents that are embedded directly within the web page. These are the items that are used to display the web page and can include background images as well as content.

Note: On many web pages, what appears to be a single image may in fact be made up of several images positioned next to each other. Each of these sub images will have its own entry in this section.

Find and Replace

Webstractor provides a simple yet powerful search system that allows you to quickly find all occurrences of a word or phrase within all of the captured web pages in a document. For pages in the Editor, you can also replace the found results with different text. To open the Find panel, select Find... from the Edit menu. Search results can be sorted by "Result", which is the order they appear in your document, or by "Rank".



(A) Search term field

Type the word(s) or phrase you wish to search for here and press Enter. Webstractor will display a list of all of the occurrences and take you to the first one found.

(B) As phrase

Check this to treat your search as a single phrase. Without this item checked, Webstractor will find instances of any of multiple words entered.

(B) Ignore case

Check this to treat UPPERCASE and lowercase characters as the same when performing the search. E.g. without this checked, a search for 'imac' would only find 'imac' and not 'iMac', whereas with this checked both versions would be found.

(B) Whole words

Check this to limit your results to whole words and not partial words. E.g. with this checked, a search for "lemon" would not find a result in the sentence "I like lemonade".

(C) Expand All / Collapse All

Click to expand or collapse all of the sections within the search results.

(D) Result count

A display of how many instances were found, and in how many entries.

(E) Resizable Drawer Edge

Drag the edge of the drawer to resize it. A wider drawer will allow you to see more of the surrounding text for each search result.

(F) Search Results

The search results for each web page in the document are listed within their own sub-section. By default these sections are collapsed; clicking on a section header will expand it and take you to the first item. You can click on any of the items in the results list to quickly jump to that occurrence, or you can use the Find Next or Find Previous items in the Edit menu to step forwards or backwards through the results.

(G) Replace

When in the Editor, use this button to replace all of the selected search results with different text. This option is not available in the Browser.

Webstractor & Spotlight

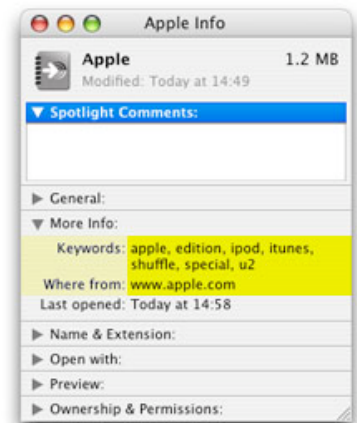
If your Mac is running OS X 10.4 or later, your Webstractor documents are automatically indexed with Spotlight technology. Webstractor provides Spotlight with three important types of information:

- **Keywords:** Automatically chosen from the titles of captured web pages
- **Where from:** Internet domains from which web pages were captured.
- **Text content:** Every word contained on every web page captured in your document.

All information is automatically searched during Spotlight searches, with the first two types visible in Finder's "Get Info" window as shown here.

Simply put, when you do a Spotlight search, if the word you're looking for appears on **any** web page you've ever captured in **any** Webstractor document on your system, it will appear in the results.

Not only that, but when you open a Webstractor document from a Spotlight results window, Webstractor will automatically search that document for the word(s) you used in your Spotlight search, taking you directly to the first instance found in your document.



Webstractor Menu Reference

This section lists all of Webstractor's menu commands.

Note: Standard menu commands that are common to most applications are not listed. E.g. New, Save, Quit etc.

Webstractor Menu

About Webstractor

Displays the standard about window. The version number of Webstractor is displayed in the top right hand corner of the window. This number will be useful if you need to contact Softchaos Support.

Provide Webstractor Feedback...

Use this if you have any problems using Webstractor or suggestions for improvement. If you need a response from Softchaos, please email Softchaos Support instead.

Preferences...

Allows you to change some of the default behavior of Webstractor.

Block Pop-Up Windows

Some web pages display pop-up windows when you open or close the page. For example, a web page might use a pop-up window to request login information from you or display advertising. You can block these pop-up windows.

Blocking pop-up windows stops only pop-up windows that open when you open or close a page. It does not block pop-up windows that open when you click a link.

If you block pop-up windows, you might miss important information for a web page.

Allow Web Page Update Checking

This item controls whether Webstractor will automatically check for updates to any web pages in the open Webstractor documents. With this turned off, you must manually check web pages for updates using Check For Updates Now. If any documents are using the Webstractor Radar to capture changes to web pages, then such changes will not be automatically captured whilst this is turned off.

Licenses...

When you buy Webstractor, you will receive a license code that removes the limitations of the trial mode. This menu command will display a window to allow you to enter that license code.

Join Mailing List...

Allows you to join the Softchaos Mailing list. Members of the list will be periodically e-mailed hints and tips for getting the most out of Webstractor as well as any special offers.

File Menu

Open Location...

Takes you to the URL entry field of the current Webstractor document. If no documents are open, a new one will be created.

Export Web Page to PDF... (Browser Only)

Saves the currently viewed web page in the Browser side of the document to the cross-platform PDF format. To export multiple web pages as a single PDF, you will need to make the web pages editable and export while in the Editor.

Export to PDF... (Editor Only)

Saves the contents of the Editor side of the document to the cross-platform PDF format. Use this if you need to share your documents with people who do not have Webstractor. PDF documents are for viewing only and cannot be re-edited within Webstractor.

Document Properties

This displays a dialog that allows you to set some options for the entire document. You can set:

- Page margins for your printed output.
- The scale of the document content within the page. To make everything smaller and fit more on a page, decrease the scale. To make everything larger and fit less on a page, increase the scale. This scale is equivalent to the scale in the Page Setup dialog.
- Whether to include page numbers on each page printed from the Editor side.
- Options for the Webstractor Radar.

Print Web Page... (Browser Only)

Prints the currently viewed web page in the Browser side of the document. To print multiple web pages as one print job, you will need to make the web pages editable and print from the Editor.

Print... (Editor Only)

As well as the standard printing options, Webstractor provides an extra option in the Print dialog. The 'Print text in color' option can be turned off to force any colored or gray text in the document to be printed in black. The original color of the text in the document will not be changed by this option. To permanently change the color of text to black use Strip Text Color.

Edit Menu

Crop

Within the Editor this removes all of the non-selected content within a web page. This is very useful when you only want to keep a small portion of a web page. Instead of having to manually delete all of the unwanted items yourself, just select the portion you are interested in and use this. Webstractor will keep the original formatting of the selection intact.

Start Editing Browsed Page

Within the Browser this adds the selected web page or pages to the Editor (if not already added) and automatically switches to the first selected page within the Editor.

Selection Mode

See the Editor section "Text Selection" for an explanation of these selection modes.

Highlight Selected text

Within the Editor this adds a yellow background to the selected text. If any of the selected text is already highlighted, this will remove the highlighting from it. Use highlighting to draw attention to particular parts of your document.

Insert Page Break

Within the Editor, this inserts a page break at the current text insertion point. Everything following the page break will start on the next page.

<Search Engine> Search...

Takes you to the Internet search field of the current Webstractor document. If no documents are open, a new one will be created. You can choose which search engine you'd like to use from this search field's menu, available by clicking its "magnifying glass" icon.

Find...

Opens up the Find panel for the current Webstractor document. See the section entitled "Find and Replace" for details.

Find Next

Selects the next entry in the search results.

Find Previous

Selects the previous entry in the search results.

Replace...

Within the Editor, use this to replace all of the selected search results with different text.

Font Menu

Contains many standard options to change the font, style and color of selected text. Only available in the Editor.

Bookmarks Menu

Webstractor can use the Bookmarks that you have set up in other Web browsers to allow you to quickly access your favorite web sites from within Webstractor. The supported browsers are Safari, Internet Explorer, OmniWeb and FireFox. Selecting a bookmark from this menu will load that page in the current document's Browser section.

Use Bookmarks From

By default, Webstractor will list the bookmarks from your default Web browser (assuming it is supported). To use the bookmarks from a different supported browser, select the browser from this submenu.

Home

Takes you to your default home page. This page can be configured in Safari's preferences.

Open All

Loads all of the bookmarks in the menu (and any sub-menus) into the current document's Browser.

Entries Menu

Stop Loading

Cancels the loading of the current web page.

Reload

Use this to ensure that you have the latest version of the selected web page (or local file) in either the Browser or Editor. In the Browser, if the web page is different from the existing captured version, Webstractor will create another copy of it, thereby keeping the older and newer version. This behavior can be changed in the Preferences. In the Editor, reloading an entry will always overwrite the existing entry.

Get Info

Displays an information sheet for the selected item or items. The information displayed depends on whether you are in the Browser or the Editor.

Rename as Selected Text

After selecting some text in either the Browser or the Editor, use this to change the name associated with the current web page to this text.

Check for Updates Now

Manually checks whether the selected web pages are up to date with the versions on their original web site. This is useful if you do not want to wait for the next automatic check or if you have Allow Web Page Update Checking turned off.

Mark as Read/Unread

Within the Browser, manually toggles the selected captures between appearing to have been read and not read.

Back

Within the Browser, use this to step backwards through the web pages you have visited in the current document.

Forward

Within the Browser, use this to step forwards through the web pages you have visited in the current document.

Mark Page For SnapBack

Marks the current web page as the one to SnapBack to (see below).

Page SnapBack

After you have marked a page for SnapBack and visited other web sites, click this button to return to the marked page. This allows you to follow a chain of links from one site and easily return there later. The SnapBack information is saved with the document.

View Menu**Browser**

Switches the current document to the Browser view.

Editor

Switches the current document to the Editor view.

Ruler

Shows and hides the text ruler within the Editor.

Show/Hide Toolbar

Shows and hides the main Toolbar in the current document.

Toggle Drawer

Shows and hides the Thumbnail and Find drawer attached to the current document.

Window Menu**Links Inspector**

Shows and hides the Links Inspector window.

Unread Web Pages

Shows and hides the Unread Web Pages window.