

ScrapBook 1.2. (& Add-ons)

A Firefox extension for
gathering information from the
web

Andrew Giles-Peters

June 2007



This work is licensed under the **Creative Commons Attribution-NonCommercial-ShareAlike License**. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/2.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

Acknowledgement

This document builds on the work by Rui Nibau, ***ScrapBook Une extension de Firefox pour archiver et éditer des documents web***, published in the series **Les tutoriels libres** by **FRAMASOFT** (<http://www.framasoft.net>).

Previous editions

ScrapBook 0.16.0 August 2005.

ScrapBook [0.18.0] December 2005

ScrapBook 1.05 May 2006

ScrapBook 1.06 June 2006

ScrapBook 1.1 (& Add-ons) August 2006

Contents

Introduction.....	5
Installation.....	6
The ScrapBook menu.....	6
Accessing the ScrapBook sidebar.....	7
Capturing web pages.....	8
Simple captures.....	8
More complex captures: controlling capture detail.....	10
Title.....	10
Folder.....	11
Download linked files.....	11
In-depth capture.....	11
Comment.....	12
Edit before Capture	13
Bookmarking with ScrapBook.....	13
The ScrapBook sidebar.....	14
Working with folders.....	14
Working with ScrapBook items.....	15
List View.....	16
Searching in the sidebar.....	17
Multi-criteria searches in the metadata	19
Multi-criteria full-text searches	20
Editing captured web pages.....	21
Launching the Edit and Info Toolbars.....	21
The Edit Toolbar.....	21
Highlight tool.....	21
Pencil tool.....	22
Eraser tool.....	22
DOM Eraser.....	22
Undo button.....	23
Save button.....	23
<< Button.....	23
Exiting Edit mode.....	23
The Info Toolbar.....	23
Creating and editing notes.....	25
The ScrapBook Tools menu.....	27
Tools Menu Commands.....	27
Settings.....	27
Manage	27
Combine wizard	28
Import/Export.....	28
Capture Multiple URLs.....	29
Add-on Functions	30
Calculation of Size.....	31
Sort Entire Tree.....	31
Output Tree as HTML.....	31
Emergency Repair.....	31
Online Tips.....	32
Home Page.....	32
About.....	32
The Settings tabs.....	32

General.....	33
Tabs.....	33
Capture.....	33
Edit.....	33
Advanced.....	34
Using ScrapBook in web-based research.....	36
Retrieval by content.....	36
Indexing and the Comment field.....	36
Folders and classification.....	36
Folders and work organisation.....	37
Documenting searches.....	37
Organising notes.....	37
Annotating captures.....	38
Using ScrapBook mark-up for active reading.....	38
Highlighting.....	38
In-line comments and links.....	38
“Sticky annotations”.....	38
Metadata Comments	38
Multiple scrapbooks.....	39
Using Multi-ScrapBook.....	39
Sharing scrapbooks.....	41
Sharing data with ScrapBox.net	41
Exporting text and references from ScrapBook.....	42
Exporting references.....	42
Installing and configuring Copy Page Info.....	42
Using Copy Page Info.....	44
Exporting text.....	44
Using the Combine Utility.....	44
Appendices.....	46
1. ScrapBook file structures.....	46
Export file structures and index.dat.....	47
2. Regular expressions and searching.....	48
3. ScrapBook Keyboard and Mouse Assignments.....	49
General.....	49
Browser Context Menu	49
Capture Detail dialog	49
Drag-and-drop on Sidebar	49
Editing Saved Page	49
Firefox ScrapBook menu.....	49
Info Toolbar	49
Manage window.....	49
Searchbar	49
Sidebar	50
4. 'Help, I have lost my data!'.....	50
5. Getting more information.....	51

Introduction

The World Wide Web is a vast repository of information containing billions of documents. This presents two problems: how do you find what you need to know and what do you do with it when you have found it? Google and its competitors have a workable but by no means perfect solution to the first problem. ScrapBook is the solution to the second problem for those who want to record what they have found.

ScrapBook is aptly named: it is your personal repository for the text snippets, PDFs, web pages, movies and whatever you collect in your travels around the web.

These three points are important. ScrapBook is *personal* in the sense that your scrapbook is part of your Firefox profile by default¹. It is also personal in that you decide how you want to organise your collection. It is a *repository* since it makes a local snapshot of the web content you specify². Finally ScrapBook items can be in virtually *any format*: almost anything Firefox can display, ScrapBook can capture.

Prior to ScrapBook or similar programs, you either book-marked interesting web pages or saved them to your hard disk. In the former case you saved where the page was but not its content; in the latter you saved the content but not where it was from. ScrapBook automatically saves the content, the source URL and the date it was captured. For researchers, who have to document the content they cite, there is no contest. Even the casual user benefits since there are no more saved files lost somewhere on the hard disk: they are all there in the scrapbook!

Add to this that ScrapBook allows selective capture: from a snippet of selected text, to a frame, through a whole page with or without embedded images, to a page together with its linked pages to various depths.

Add, furthermore, that ScrapBook items can be searched and sorted, edited and annotated, classified and filed away in ScrapBook folders. Nor is this all as you will see in the following pages.

As a Firefox extension, ScrapBook only works with the Firefox web browser³. This means however that it works with very slight variations equally well under Linux, Mac OS and Windows. Given that both are free, and Firefox has strengths of its own, ScrapBook and Firefox should be in every modern researcher's toolkit.

¹ I will suggest later why you might change this.

² Although you can keep your scrapbook on a network drive or save to a USB memory key if you like.

³ Firefox is available for download from the Mozilla foundation:
<http://www.mozilla.org/products/firefox/central.html>

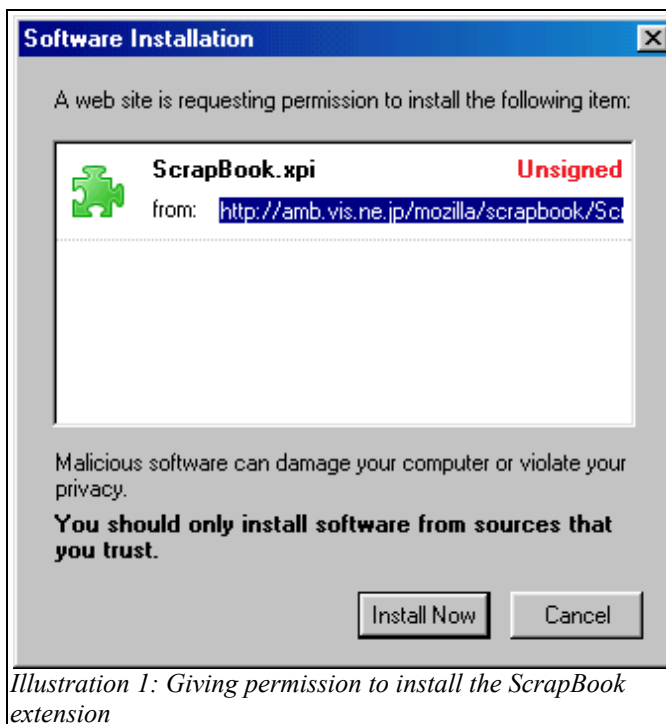
Installation

You can download the ScrapBook extension from two sources

Mozilla Update: <https://addons.mozilla.org/firefox/addon/427/>

The ScrapBook home page: <http://amb.vis.ne.jp/mozilla/scrapbook>

If you download from the ScrapBook site you will have to configure your browser to allow the site to open a pop-up window on your computer. This is discussed later in connection with installing addons from the ScrapBook site. For now use the Mozilla site which is automatically accepted. When the ScrapBook page opens, click on the link "Install now" to start the installation. The Software Installation window will open.



Click on the button **Install now** in the window that has opened. The extensions window will then open.

Once the process is finished, close the extension window and exit Firefox. When you restart Firefox, there will be a new ScrapBook menu in the Firefox menubar between the Bookmarks and the Tools menus and ScrapBook will be listed as installed in the extensions window accessed from the Tools menu.

Illustration 1: Giving permission to install the ScrapBook extension

The ScrapBook menu

Initially the ScrapBook menu just contains four items: three Capture options⁴ and the ScrapBook command to toggle the ScrapBook sidebar on and off.

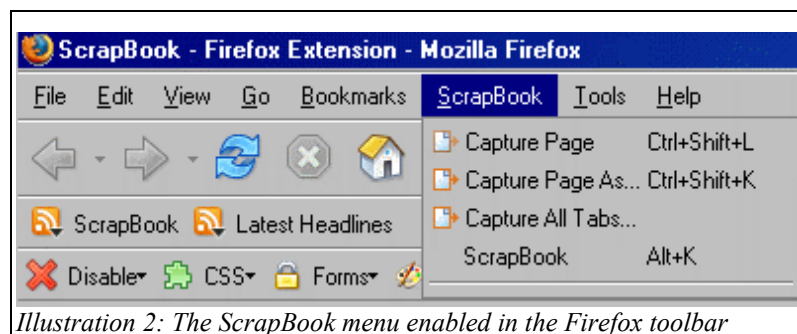


Illustration 2: The ScrapBook menu enabled in the Firefox toolbar

⁴The following discussion assumes that you have not selected text in the web page you are reading. If you have, the **Capture Page (As)** options become **Capture Selection (As)** options.

If you choose '**Capture Page**' (i.e. the current tab) the web page you are viewing is saved immediately to the scrapbook root folder. A few messages will flash through the Firefox statusbar but that is all you will see.

If you choose '**Capture Page As..**' a Capture Detail window will pop-up. This is discussed later. Amongst other things, it lets you save the current page/tab contents in a specified folder.

If you choose '**Capture All Tabs**', i.e. all pages you have open, a Select Folder window will pop-up. Since you haven't yet made any folders the window will be empty. Just click OK to save to the Root Folder. (There is a button 'New Folder' at the bottom left of the window that you can use to make folders. These will appear in the Select Folder window. You can then choose Cancel to exit the Capture dialogue without saving any pages but the folders will be saved.)

After you create folders or save pages these appear below the separator in the ScrapBook menu⁵. If you click on a page in the ScrapBook menu or in a folder under the ScrapBook menu it will open in the current tab – unless you have changed this behaviour in Settings (to be discussed later).

There is also a **Capture Page** command displayed when you select a folder in the ScrapBook menu. Clicking on this saves the current page to that folder without further dialogue. As with the **Capture Page** commands above the separator, this changes to **Capture Selection** if you have selected text on the current page. If you middle click on the command a Capture Detail window will pop up.

Accessing the ScrapBook sidebar

Once Scrapbook is installed you can capture pages by right-clicking in the browser window or using the Capture options from the ScrapBook menu but to access other Scrapbook features you need to activate the ScrapBook entry in the ScrapBook menu. This will show the ScrapBook sidebar. A somewhat longer way is to choose Scrapbook from the Sidebar submenu of the View menu. (This will be relevant if you decide to suppress display of the ScrapBook menu.)

Three other faster ways exist. If you are a keyboard person, use the shortcut Alt-K. Otherwise, middle click on the ScrapBook icon on the extreme right in the Firefox statusbar on the bottom right of the Firefox window. Or, finally, if you prefer using the top of the window, you can put a launch icon for ScrapBook on the Bookmark toolbar and access this by left-clicking.

To do this, chose View – Toolbars – Customise and drag the ScrapBook icon from the display of available icons to the toolbar. (You will probably have to scroll down the display to find it.)

All of these methods toggle the sidebar on and off.

⁵Strictly speaking only those in the current scrapbook appear in the ScrapBook menu. This will be discussed in relation to Multi-ScrapBook later on.

Capturing web pages

ScrapBook has a number of capture options available from the ScrapBook menu, from a context menu in the browser window, or for “drag and drop”.

Simple captures

Usually you do these by right-clicking to bring up a context menu, from the Scrapbook menu or by drag-and-drop into the Scrapbook sidebar (if visible).

Go to a page that you want to save. Place the mouse anywhere in the browser window. (Do not however select text or activate a link!) Right click to see the context menu. Choose **Capture Page**.

You will then have the choice between filing into the Root Folder, into up to five of the most-recently used folders, or of selecting another folder through a dialogue. The Select folder dialogue will also allow you to create an appropriate folder.

You can do the same from the ScrapBook menu in the Firefox menubar.

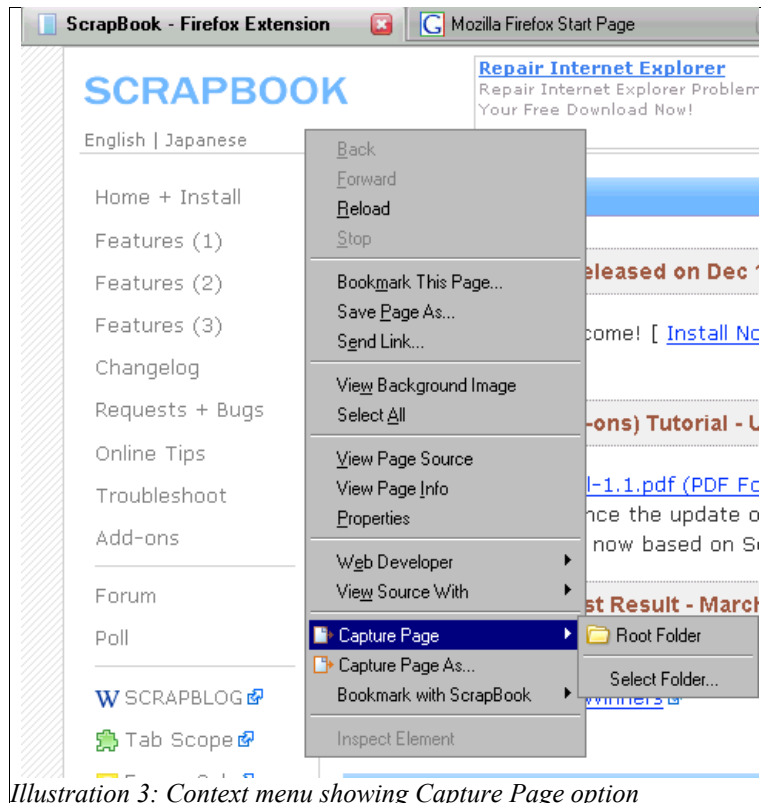


Illustration 3: Context menu showing Capture Page option

You can also save a full Web page by drag-and-drop of its favicon⁶ in the browser location bar into the ScrapBook Sidebar. Note that you cannot drag and drop the URL but only the icon to its left. A tool tip will pop up if you hover the mouse over this and the pointer will change to a hand for grabbing. The names of folders that you drag and

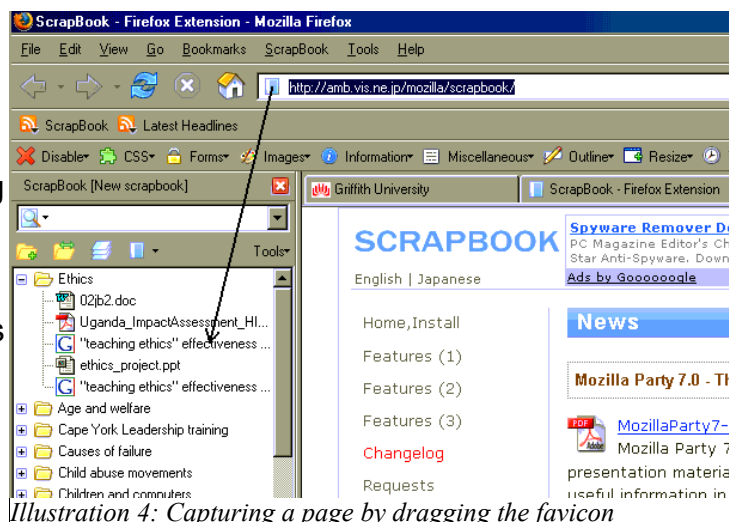


Illustration 4: Capturing a page by dragging the favicon

⁶ Short for 'favorites icon' – an icon created for a web page or site and displayed by most browsers.

drop to are added to the list of most recently used folders in the Capture dialogues.

Any of the above methods will capture the entire page with its images and styles. The quality of the capture is often better than **Save Page As**. The name of the captured page will be the page title or its URL if it has no title.

To capture part of a web page, select the part you want, right click anywhere in the window to bring up the context menu, and choose **Capture Selection**. Alternatively choose **Capture Selection** from the ScrapBook menu or drag the selection onto the ScrapBook sidebar. ScrapBook will create a new HTML file containing the selection using the first few words in the selection as the item name.

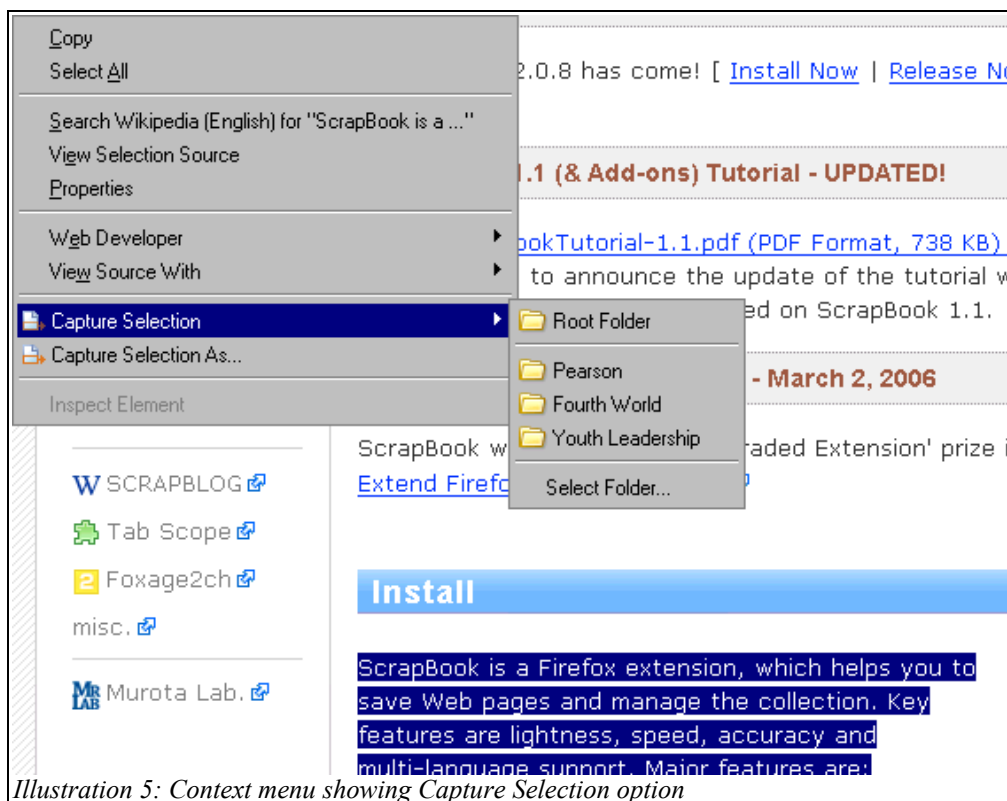


Illustration 5: Context menu showing Capture Selection option

To capture a linked web page or file, right click on the link you are interested in, and select **Capture Link**. Alternatively drag the link into the ScrapBook sidebar. There is no ScrapBook menu option for doing this since selecting a link to identify it would also activate it. If the linked file is HTML then the name of the captured item will be the page title or the URL if the former is absent. If it is not HTML the name will be the filename.

You can capture a frame by right-clicking inside it to bring up the context menu and selecting **Capture Frame**. You are also able to select **Capture Page** in this case. If however you have selected text in that frame you will only be able to **Capture Selection**. There is no ScrapBook menu option for capturing frames since a particular one could only be temporarily identified by making a selection within it and this gives the **Capture Selection** option.

More complex captures: controlling capture detail

As the above screenshots show, for each of the commands mentioned there is an additional command **Capture [whatever] As**.

The **Capture ...As** commands allow you to do three things you can't do with the simple capture commands. You can change the name under which the file is captured, specify what linked or embedded files come with it, and write a comment to attach to it if you desire. The second of these you can **only** do at the moment of capture.

You can of course also specify where to put the capture if you have defined a folder structure in the sidebar. The whole list of folders is shown in the drop-down folder box with the most recently-used five at the top. You can also create new folders for your captures.

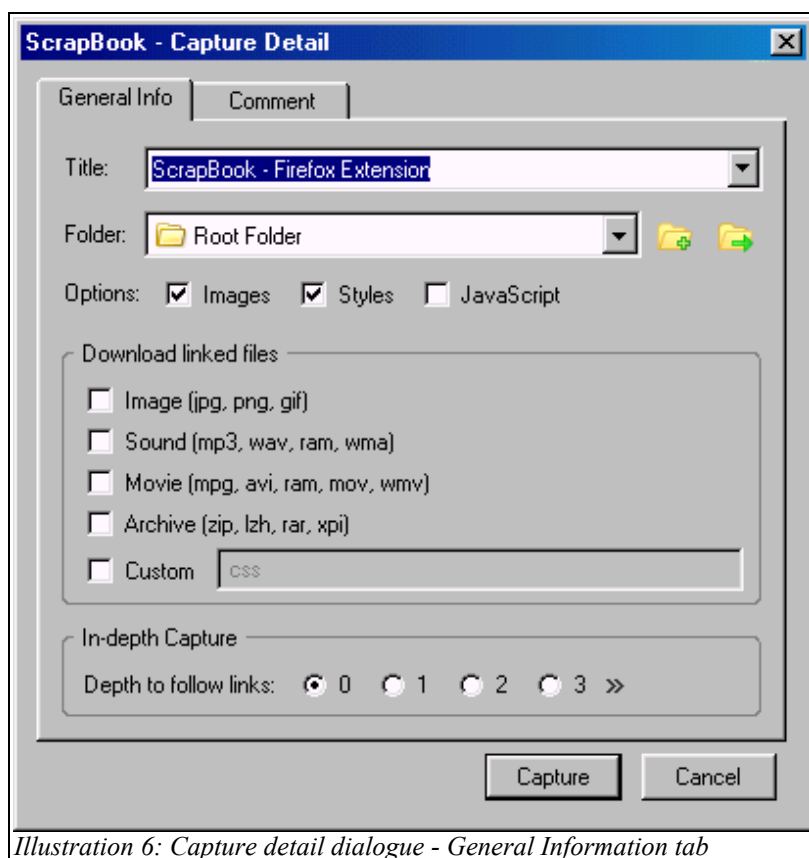


Illustration 6: Capture detail dialogue - General Information tab

The Capture Detail dialogue gives very fine control over the capture process.

Title

The default name of the captured item will usually be the page title in the case of a web page capture, but the URL if there is no title, the file name in the case of a link and the initial text in the case of a selection. In the latter cases it will be sensible to choose a name more meaningful to you. If you want to change the name later you can do so by right-clicking on the item in the sidebar and choosing **Properties**.

Folder

ScrapBook captures to the Root Folder of the sidebar by default. If you have defined a folder structure to classify your captures then you can choose which folder to capture the item to. As with renaming the item this can be done later, but, unless the name makes the classification obvious, it will be faster to file it when your memory of what the item is about is freshest.

To facilitate classification, the **Capture Detail** dialogue includes buttons for making new folders and selecting from the existing ones. The dropdown folder list includes the five most recent choices at the top.

Download linked files

By default no linked files are captured in Capture Detail. You can choose what sort of linked files should be included such as images, sound, movies, archives and document files or PDFs. ScrapBook remembers your settings between captures. You can also capture linked web (HTML) pages to a specified depth from zero to three or set your own depth. ('Choose '>>' in the screenshot above.)

In-depth capture

When you choose a 'Depth to follow links' option other than zero, ScrapBook generates a list of the files that would be captured. You can then **Pause** the process and **Filter** the list or directly modify it by unchecking the boxes to the left of the filenames. The most powerful filter is the string filter which enables you to select only pages matching a URL fragment. Other options are to restrict the files to a single domain or directory, to select all or none, or to invert the selections. You can also **Skip** a file that is loading when you have re-started the process. (The **Pause** button shown below becomes a **Start** button when the process is paused.)

Obviously these inclusion decisions have to be made at the moment of capture.

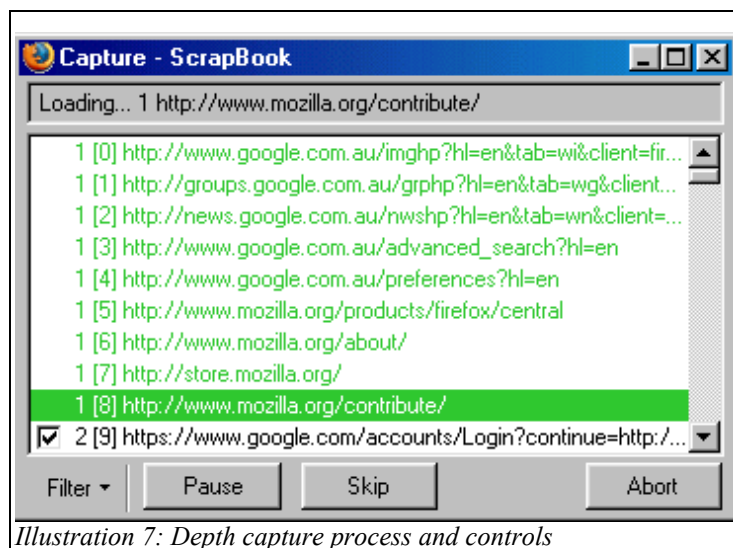


Illustration 7: Depth capture process and controls

Whilst the In-depth Capture option works well for simple links, it does not work so well for JavaScript-based menus of calculated links. You might find that

the files are captured but not linked. In this case try checking the JavaScript option on the Capture Detail dialog. Even then things may not work that well.

A point to note here is that capturing linked pages to depth three might bring in a lot of pages. In the case of the Firefox Start Page it is some 435! To allow you to Pause and Skip – and to trick websites that are armed against automatic download – ScrapBook pauses for three seconds between each page. So this capture process would take over twenty minutes - and much longer over a slow connection.

There has been some discussion of these limitations in the ScrapBook forum so the current version of ScrapBook allows you to capture to a depth greater than 3 and also relaxes the three seconds rule for the file: protocol. In general however capturing to a depth of more than 3 will take a long time.

Another point to note is that capture in depth only saves one entry to the ScrapBook sidebar. In this it is quite different from Capture Multiple URLs to be discussed later. You access the pages from depth capture from the page you have captured exactly as if it were a remote web page. The difference is that the embedded URLs have been changed to point to local files. Hence the copy is frozen at the moment of capture. If instead you capture to depth zero and attempt to access links from the saved page, you might find that they are updated, no longer valid or even inaccessible if the page used a relative URL. Note also that if you skip a file in the depth capture process the link will point to the original location and not to a local copy.

Comment

As mentioned previously, ScrapBook automatically keeps the source URL and date/time of capture. You might however want to attach some additional information using the second, Comment, tab of the Capture Detail dialogue. If you decide to do this later the Comment field is available as one of the item properties accessible through the context menu in the ScrapBook sidebar .

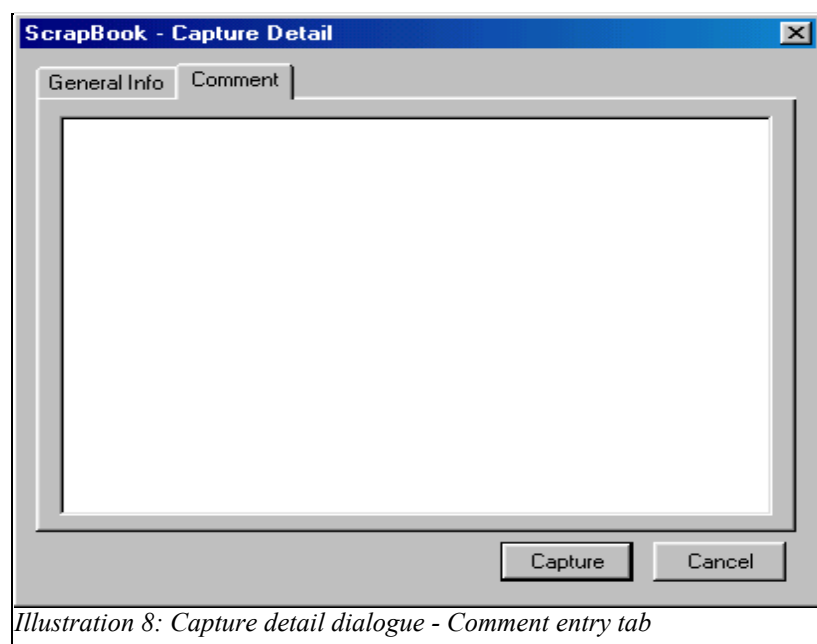


Illustration 8: Capture detail dialogue - Comment entry tab

Edit before Capture

If you right or left click on the ScrapBook icon in the bottom right of the Firefox statusbar when you are browsing a web page, you get a pop-up menu whose only active option is **Edit before Capture**. If you choose this, the Edit Toolbar appears and you can perform most of the ScrapBook Edit operations on the web page prior to capture. The Edit Toolbar will be discussed later; for now it is only important to note that the Save button on the Edit Toolbar gives another way of capturing a page.

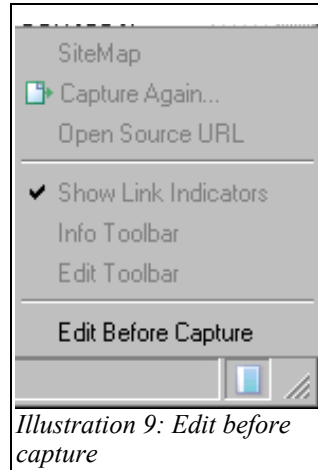


Illustration 9: Edit before capture

The **Edit before Capture** feature is quite general and can be used to clean up or annotate pages prior to printing or saving to disk as well. ScrapBook does not automatically capture the edited page. (Naturally the page has to be HTML rather than a PDF, DOC or TXT format.) The available operations are highlighting or deleting selections, inserting annotations and deleting parts of the document structure. You can also edit the document title and comment but these will be lost if you don't capture the document into ScrapBook.

If you click on the Scrapbook icon when you are viewing a captured (ScrapBook) item then the **Edit before Capture** option is dimmed out and the other menu choices are active. These will be discussed later.

Bookmarking with ScrapBook

Bookmarks are saved using the **Bookmark with ScrapBook** command in the context menu. The process is the same as **Capture Page**. The bookmark is saved to the scrapbook and not to the Firefox bookmarks file. The visible difference between a scrapbook bookmark and other scrapbook items is that the bookmark appears in the sidebar as a green title and, as there is no content, opening the item displays the remote web page rather than a local item.

A ScrapBook bookmark is slightly different from a Firefox bookmark in that it has a date metadata field in its properties but no keyword one. (Expect however ScrapBook to include a keyword metadata field some day.) Apart from that you can do very similar things with them – open them, refile them, etc. Scrapbook treats them as standard items so you can also export them or apply the **Capture Again** command to over-write them with the web page they point to. It does not seem however that you can change a captured item back to a bookmark!

The ScrapBook sidebar

The sidebar initially has four areas: at the top the title 'ScrapBook'⁷ and a Close button; then a search box; next a toolbar/menu area with icons for New Folder/Note, Expand/Collapse All folders, List View, Multi-ScrapBook (if enabled – shown here) and a Tools menu; finally a scrolling region containing the folders and captured items. When a note is created or edited, a further editing area may open below the scrolling region.

I will discuss the ScrapBook Tools menu, Multi-ScrapBook and notes later and concentrate for now on folders and List View.

Working with folders

It has already been mentioned that simple captures can be affected by dragging and dropping icons, selections and links into the ScrapBook sidebar. If you have defined a folder structure in the sidebar then you can drag and drop directly to the desired folder, which makes classification of your captures very easy.

Similarly to move a document from one folder to another within the sidebar, click on the document within the sidebar folder and drag it to its destination. If the destination folder is not open it will file at the bottom of the folder.⁸ If it is open you can place it where you like in the folder.

The New Folder icon is actually a menu (without the down arrowhead) which allows you to create a new folder in the classification tree whose name you can then edit and to which you can add an optional comment. Note that the date/time of creation is also saved.

The Expand/Collapse All Folders icon is useful when you are re-organising as you can drag and drop whole folders as well as documents. You open and close particular folders in the normal way by clicking on the folder name or on the '+' sign beside the folder icon.

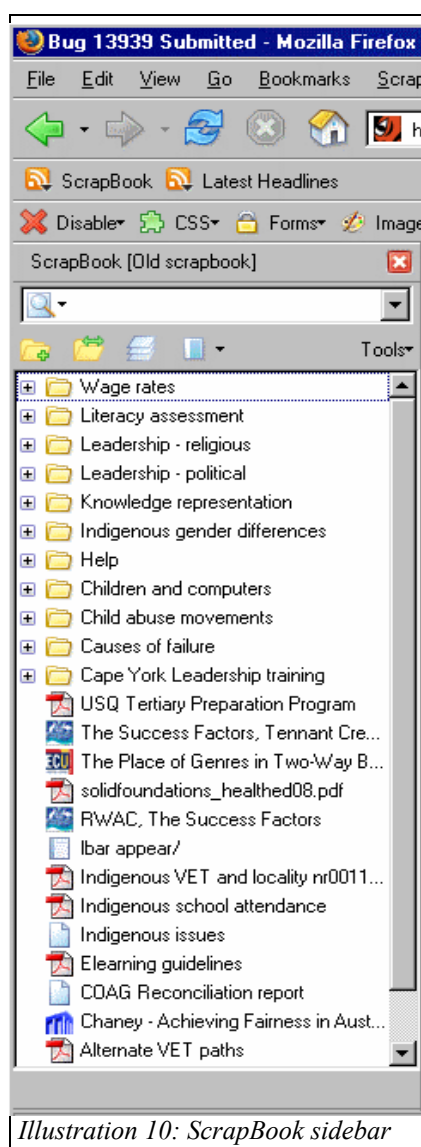


Illustration 10: ScrapBook sidebar

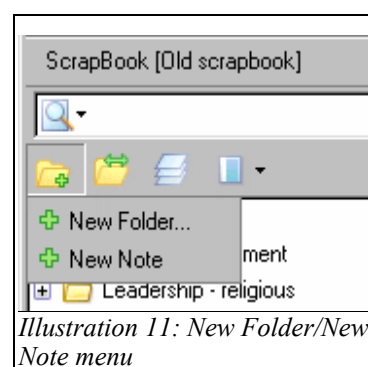


Illustration 11: New Folder/New Note menu

⁷ If you have previously enabled Multi-ScrapBook the title will be 'ScrapBook [ScrapbookName]'. This will be discussed later.

⁸ This default behaviour can be changed by a setting discussed later.

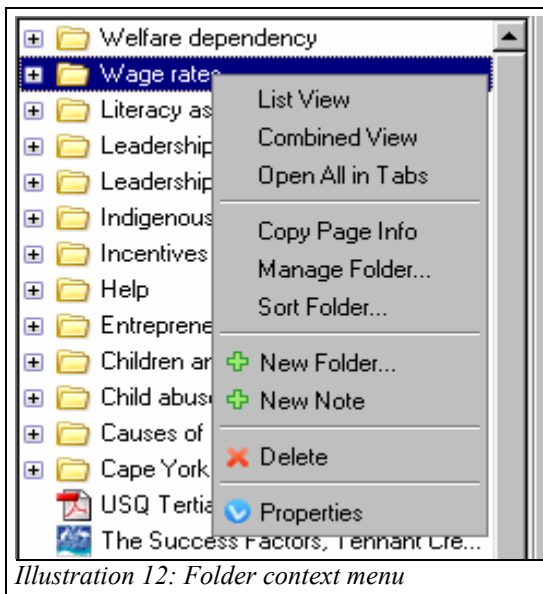


Illustration 12: Folder context menu

Right-clicking on a folder gives you a context menu. This allows you to open the folder in list view (see below) instead of the standard tree view. It also allows you to display all the documents contained in the folder in two ways, horizontally or vertically – or so to speak. You can open its contents in as many Firefox tabs or open its contents in a single window as a combined or concatenated document. The latter option allows you the possibility of printing the combined document but will only work with HTML and text documents like notes. The former can display all types of documents

but will give rise to the usual dialogs if external programs have to be used.

The next group of options are rather varied. **Copy Page Info** is an optional add-on that will copy specified metadata from the properties of each item in the folder to the Clipboard from whence you can paste it into a document. It is discussed in a later section. **Manage Folder** gives access to a view of the folder in which you can select, move or delete groups of items – you can only select one item at a time in the sidebar. The Manage utility is also available for manipulating a whole scrapbook and will be discussed later in the context of the ScrapBook Tools menu. **Sort Folder** allows you to sort the items in the folder by name or capture date, and also optionally to extend the sort to sub-folders.

The **New Note** and **New Folder** options do not work quite as you would expect. The new item is not created in the current folder but rather at the same level in the tree, i.e. as a sibling rather than a child. This is also what happens, rather more reasonably, if you select these commands from the context menu for a scrapbook item. It is only a minor annoyance.

Finally you can also delete a folder and display or change its properties such as its name and any comment attached.

Working with ScrapBook items

The individual items contained in the sidebar folders are bookmarks (these appear as green titles), captures of various file types, combined files, HTML files and notes. All can be opened (displayed), exported, deleted or have their properties displayed and altered (and also copied if Copy Page Information is installed). Clicking on an item will display it – in the relevant sense of display – whilst right clicking will display a context menu which is largely the same for all items. With the exception of notes all support the **Capture Again** command which overwrites the item with a recapture from the source URL. This of course converts a bookmark into an ordinary capture.

If you choose export in the tool sub-menu of the context menu the item is immediately copied to the current export directory without bringing up the Import/Export dialogue. (See the discussion of this utility in the section on the tools menu.)

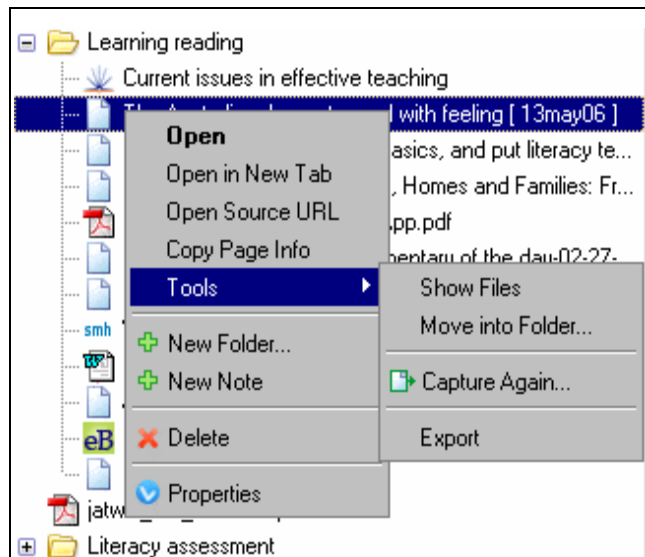


Illustration 13: Item context menu

List View

List View splits the sidebar in two; displaying folders in the top panel and files in the bottom panel. This gives an alternative way of viewing the files in a folder to expanding the folder tree. With List View enabled, clicking on a folder displays its files in the bottom panel. It will also display any sub-folders of the folder in the top panel in the standard tree view.

In the illustration the folder Wage rates has been opened in List View as show by the title of the lower panel and the '-' icon beside the folder name.

With List View enabled, opening folders in the top panel by clicking on the '+' icon displays the sub-folders but does not change the files displayed in the bottom panel. Likewise with closing the folders with the '-' icon. Hence you can navigate the folder tree in the top panel whilst keeping a fixed display in the bottom one.

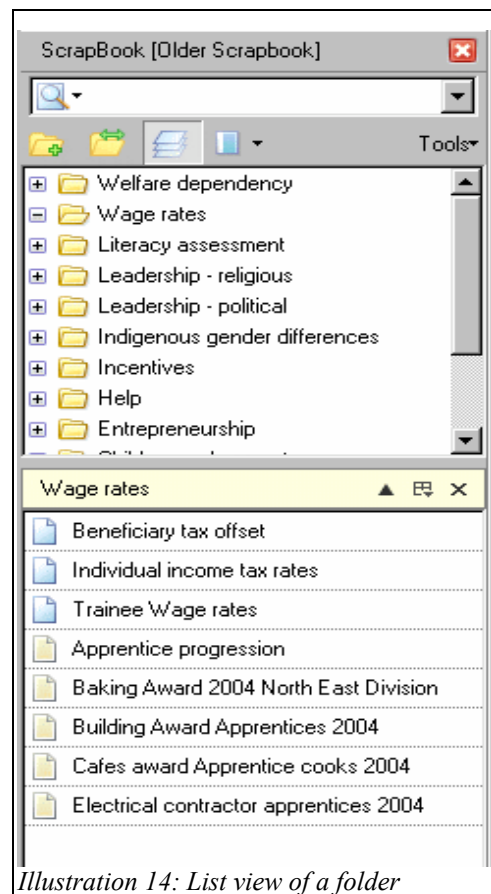


Illustration 14: List view of a folder

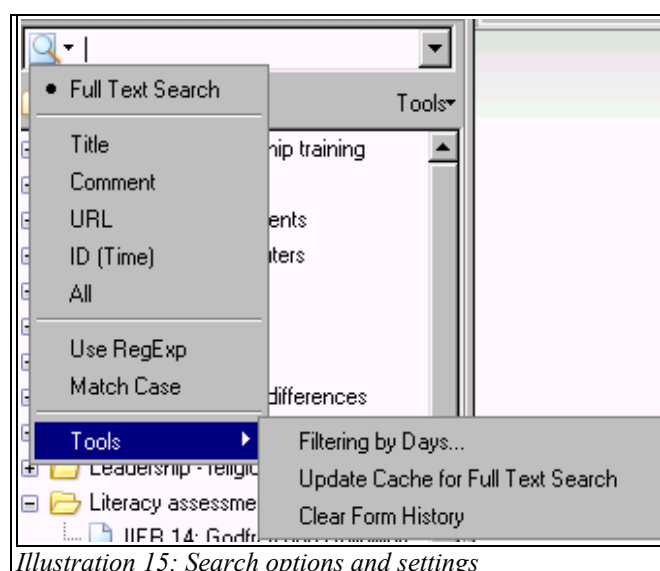
You can also change the bottom display panel independently of the top one by using the up-arrowhead control in the header. This steps up through the folder tree, i.e. it displays the files in the parent folder. (Naturally it is not displayed if you are in the root folder.)

A benefit of List View as compared with the standard tree view is that you can display extra information. The second icon in the lower panel header allows you to choose to display comments as well as titles and also URLs.

You exit List View either by clicking the close control 'x' in the bottom panel header or by clicking the **List View** icon at the top of the sidebar. It will also close if you choose the **New Note** or **New Folder** command.

Searching in the sidebar

The ScrapBook search box is a simple but powerful tool supporting two quite different forms of search through your scrapbook documents. You can search either your data – the HTML pages you have captured, but not text in sticky annotations or inline comments you have added to the data – or your metadata – the information that you and ScrapBook have created to describe it⁹. In the latter case you can search on title, source URL, date/time of capture, comment or all four together. You can also make your search case sensitive and use the powerful pattern-matching language of regular expressions. To access the pull-down menu for these options click on the search icon. (If you instead click on the arrowhead you will get a drop-down list of your past search terms.)



The meaning of the options is as follows:

Full Text Search This searches the captured HTML pages (and any notes). If you have not selected Regular Expressions then a search for 'apples pears' will return exactly the same documents as one for 'pears apples', viz those documents in which both terms appear. If you have selected regular expressions then Scrapbook will look for the exact phrase and the two searches will have potentially different results. You can also do a phrase search without using regular expressions by enclosing the phrase in double quotes. Hence a search for "apples pears" returns those documents containing '...apples pears...' somewhere. You can also search for documents mentioning apples

⁹ You cannot do a Free text search of non-HTML or non-text file types but you can search their metadata.

but not pears by using 'apples –pears'. Note that the minus sign is attached to 'pears' and separated from 'apple'. It is negation not subtraction and the order of the terms is not important. You can have several negated terms. You cannot however only have negated terms.

A point to note is that a full-text search searches in the title and comment as well as the HTML content of a captured page. If you select RegExp then the search is done independently on the three fields and the results combined. If you do not then the search is done on the combined fields.

Title	This will search for the word or phrase or regular expression you nominate anywhere in the title.
Comment	This will search for the word or phrase or regular expression you nominate anywhere in the comment field.
URL	This will search for the word or phrase or regular expression you nominate anywhere in the source web address.
ID (Time)	The ID is the date/time of capture in the form yyyyymmddhhmmss. Hence it only makes sense to search for numbers in this field. You can search for any fragment of a number, or a numerical pattern by using regular expressions, but the normal thing would be to use search expressions like '2005', '200506', '20050601', etc to retrieve all captures in 2005, all in June 2005, all on the 1 st of June 2005 etc.
All	This searches each of the fields above using the search string and amalgamates the results. When used with regular expressions this allows you to do multi-criteria searches.
Use RegExp	Determines whether the search string is interpreted as a regular expression or not. Regular expressions can have a complex syntax. Simple ones are built out of letters and digits together with special characters such as the period ('.') which matches any one character, the question mark ('?') which matches zero or (any) one characters, the asterisk ('*') which matches any number of repetitions of the preceding character in the pattern, the pipe (' ') which separates alternative match expressions, etc. (Note that this quite unlike filename matching in DOS or Unix. See Appendix 2 for details.)
Match Case	If this is selected the search is case sensitive.
Tools	This is a sub-menu that allows you to: Filter by days – this finds the most recent captures (you specify the number of days to go back) Update the cache for full text search – ScrapBook automatically rebuilds its search cache every five days. Clear form history – the search box list retains your search terms unless you clear them

To set the search mode you can either select with the menu or type the corresponding character into the Search box. (See Appendix 3.)

When you carry out a Full Text Search the results will display in the upper pane of a split browser window. If you click on an item it will open in the lower pane. If you right click however, a context menu will appear allowing access to the properties and the files associated with the item and the option to open it in the lower pane or a new tab. The data display in the upper pane is similar to that in the Import/Export and Manage tools: you can choose the items of metadata displayed by clicking on the icon/menu at the far top right of the pane.

When you carry out a search of the metadata, the results appear instead in the sidebar in place of the folder tree structure. Clicking on them opens the item in the normal way and right-clicking brings up the usual sidebar context menu.

The consequence of these differences is that you can do rather more with the results of a metadata search. In neither case can you operate on the search results as a whole but the context menu allows you to refile, export or delete individual items found with a metadata search. If you have the appropriate add-ons installed you can also copy the item metadata or export to Box.net. As we will see later, searching in the Manage interface overcomes the limitation to operating on single items.

Multi-criteria searches in the metadata

In many bibliographic and database systems it is possible to do multi-criteria searches (or filtering) whether by using a query-by-example form or a special search language. This can be quite useful. For example imagine that you know you captured an item last year about some particular topic and want to find it. It appears that you can either search in the Date field for the year and check each one found for the topic or search for words in the Title or Comment fields and check each for the date. However by using the ALL search across all metadata fields and a regular expression you can search for both at once and considerably reduce the amount of manual checking.

If you search All metadata fields using regular expressions you can actually set independent criteria for each metadata field. The trick here is that for an ALL search ScrapBook combines the metadata fields into a single text for searching but separates them with newline codes (“\n”). So you write regular expressions to match what you want in each field and separate them with “\n” to match the newlines. Your final search expression will hence be of the form:

```
title_match\ncomment_match\nsource_match\ndate_match
```

where:

title_match is the expression to match against the Title field,
comment_match the expression to match against the Comment field,
source_match the expression to match against the Source field, and
date_match the expression to match against the date field.

The match expressions can be:

null (an empty expression)	this matches an empty field
.*	this matches any content
literal text	this matches itself as a whole field or at the end of the field for the Title and at the start of the field for the Date
.*literal text.*	this matches itself anywhere in a field
More complex regular expressions	The sky is the limit! (See Appendix 2 on 'Regular expressions and searching'.)

Logically you are ANDing together independent conditions on the fields. To do more complex searches such as OR searches consult **Appendix 2**.

Multi-criteria full-text searches

You cannot do anything quite like this with full text searches. You have the choice of using either regular expressions or included and excluded terms and phrases. If you use the former, the Title, Comment and Content are searched independently and the results combined. If you use the latter, the Title, Comment and Content are combined (as a tab separated text) and then searched. The reason for this is that Title, Comment and Content contain essentially the same type of text – unlike say Date and Source fields. - so it doesn't make much sense to search for different things in different fields.

At present there is no way to combine full text and metadata criteria.

Editing captured web pages

ScrapBook not only allows us to save, classify and document web pages; it also allows us to edit them to a certain extent. More exactly, it supports both highlighting text and adding comments to it and deleting both text and whole web page elements. It does not however support adding text as if it were part of the original web page. It thus maintains the distinction between the material saved and comments upon it. You can of course fully edit a ScrapBook web document by using an extension like ViewSourceWith to call up a web page editor but there is rarely a good reason to do so. (ScrapBook notes, as you will see, are fully editable.)

Launching the Edit and Info Toolbars

To edit a ScrapBook document, open the ScrapBook sidebar and select the document or launch it from the ScrapBook menu. Click on its name and it will open in the browser window – or in a new tab of the browser window if you have set that preference. (See later.) It looks like any other web page except that its URL is a local file. You might however also see the Edit Toolbar and/or the Info Toolbar. These can be toggled on and off clicking on the ScrapBook icon in the Firefox statusbar and choosing your preferences.

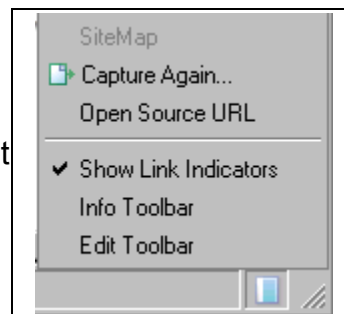


Illustration 16: Accessing the Edit and Info toolbars from the ScrapBook icon

The Edit Toolbar

Reading from the left, the Edit toolbar shown below has two small text boxes, one for the document title and the other for the optional comment as their icons indicate, and a number of tool icons and arrowheads for associated pop-up menus. You can edit the title and comment in the text boxes but it is better to do so via **Properties** in the sidebar context menu since that gives much more room. (Note however that clicking on the comment balloon opens a multi-line edit box.)

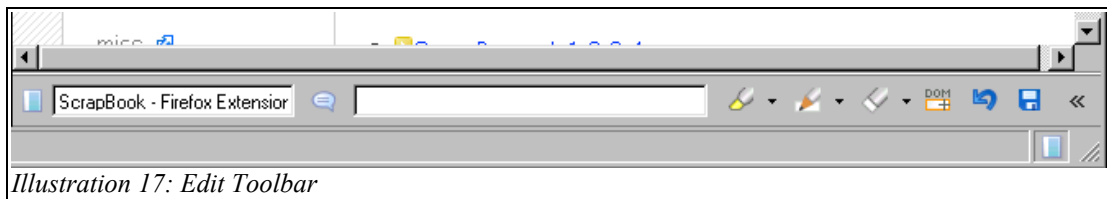


Illustration 17: Edit Toolbar

Highlight tool

Click the highlight icon if you want to highlight a selected passage. There are four pre-set highlight colours. The icon default is the most recently used colour but you can set a different colour by choosing a another option in the pop-up menu attached to the arrowhead to the right of the icon. Note that the colour selected applies to future selections, selections already highlighted retain their original colours. You can change the colours and styles of the highlights using the Customise option on the Settings Edit tab discussed later.

Pencil tool

The Pencil tool icon allows you to add a “sticky annotation” to the captured document. If you click the arrowhead beside it, you will get a more extensive menu allowing you also to add an inline comment to selected text, or to attach a file or external link to a selection to make a hyperlink. (If you haven’t made a selection, these menu items have no effect.)

If you choose to add a sticky annotation, an edit window appears. Type your comment in the window and click on the Save button at the foot of the window. (Scroll down if this is obscured.) The annotation will then appear in a box on a semi-transparent gray background. You can move the annotation block around the file and it will stay where you leave it – hence the “sticky”. You can edit and delete sticky annotations by clicking on them to re-open the edit window.

If you have made a selection and choose to add an inline comment, a single line text box opens in the window. Type your comment in the text box and click on the Enter command to the right of the box. The selection will now have red dotted underlining and the inline comment will appear if you hover the mouse over the selection. The text box will appear again when you click on the selection. To close, choose OK or Cancel. Unlike the Edit sticky annotation window this one is modal. You have to close it to do something else.

Eraser tool

If you make a text selection, and click on the Eraser icon, the selection will be deleted.

The arrowhead menu to the right of the eraser also gives you the options of deleting included scripts, inline frames, all highlights in a selection and all highlights.

DOM Eraser

The DOM (Document Object Model) Eraser icon allows you to delete whole chunks of a web page. If you click on the button, it remains depressed. If you now move the mouse around the web page you will see blocks in a red outline corresponding to HTML divisions, tables, lines, cells, paragraphs etc in the web page. If you click on any of them the outlined design element will be deleted. You can thus very rapidly delete navigation menus, publicity banners, etc and keep only the parts that you’re interested in.

This is however a very powerful tool and you can easily delete part of what you meant to save. If so, use the **Undo** button that appears when you make a deletion. Remember also to click on the depressed DOM Eraser to turn it off when you have deleted what you intended to, for otherwise your next click will take out something you didn’t intend!

If you only want to keep one structural element then you can use the Eraser in Isolate mode. To do this right-click on the element you want to keep and everything else disappears!

If you are a Linux user you might find that the Dom Eraser works a little differently from this description. I find that I have to click once to establish the

context then the second click performs the command. Also the DOM Eraser doesn't appear to switch off when you click it. You can tell of course whether it is off or not since the red boxes appear when it is on.

If you find the mouse-based DOM Eraser makes you nervous there is an useful keyboard command extension for doing the same thing called Aardvark. See its home page at <http://karmatics.com/aardvark/>.

Undo button

This button was previously called the Abandon Changes and Restore/Undo(Z) button. This gave a better idea of its operation. The button has two functions: undoing the last edit operation and reloading the last saved version of the page you are editing. When you highlight, insert annotations or make deletions the Undo button allows you to undo the last change. Once you have undone the last change the button changes its function. You cannot roll back multiple changes one by one.

If you have radically changed the structure of your document and want to restore it, you can undo the operations affected since the last save by clicking the Undo button twice. Unfortunately, however, the dialogue that pops up is a Save dialogue. You are not asked whether you want to reload the document but rather whether you want to save the changes! The normal use of the button hence has you answering negatively which takes some time to get used to. Also you are being asked whether you want to lose the changes or lose the original. You might decide that you want neither but there is no CANCEL button for this. (If you close the dialogue window then the old version is restored, i.e. you lose the changes.)

Clearly this is rather bad design: a button changes its function without changing its appearance; the dialogue seems appropriate to a Save operation rather than a Restore one: you find yourself locked into a destructive operation with no way to backtrack. Hopefully the next version will fix this.

Save button

If you want to commit the changes you have made to date, use the Save button. There is no dialogue. Note that ScrapBook does not currently autosave.

<< Button

This button opens the sidebar if it is closed and selects the document being edited. This useful if you have closed the sidebar to have more editing space.

Exiting Edit mode

To turn Edit mode off click on the ScrapBook icon on the bottom right of the status bar and de-select Edit toolbar. If you have made any changes since the last saved version you will be asked if you want to save your changes. If you cancel this dialogue you exit Edit mode without saving.

The Info Toolbar

The ScrapBook icon gives access to a menu for toggling display of the Info Toolbar, marking local links and accessing the SiteMap of a document captured in depth, recapturing (and hence overwriting) a captured document and displaying the current version of a web page without over-writing the

captured one.

The Info Toolbar is fairly simple: a display of the URL, a Home button, a SiteMap button and the << button to locate the item in the sidebar.

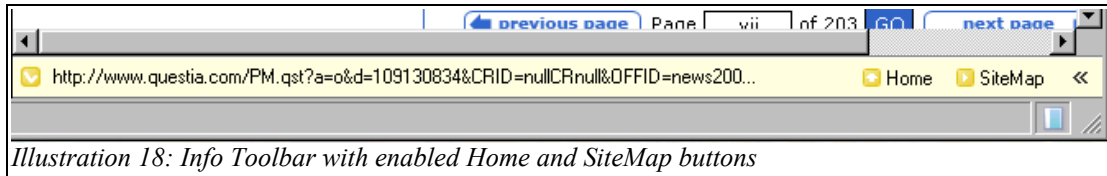


Illustration 18: Info Toolbar with enabled Home and SiteMap buttons

The value of these controls is when you have done an in-depth capture. (If you have not captured the item in depth the Home and SiteMap will be in gray not black.) Only one item is created in the ScrapBook tree in the sidebar for each Capture Detail dialogue so you cannot access the accompanying files for an in-depth capture by navigation in the scrapbook tree. You access them by using the SiteMap button on the Info Toolbar. This brings up a hierarchical (indented) list of the associated pages. You click on the page desired to load your local copy. The Home button takes you back to the main page. The URL on the Info Toolbar remains that of the main captured page; you do not have access to the source URLs of the SiteMap pages.



Illustration 19: Site map (locally saved links) for a page captured with in-depth capture

The Show Link Indicators toggle – shown selected in the illustration above – also only has an effect on pages captured in depth. If it is selected, a small yellow icon appears before links pointing to a local captured file. Unlike the SiteMap command it is not grayed out for pages not captured in depth.

Creating and editing notes

As well as the capture and manipulation of the web documents, ScrapBook also supports note-taking. To create a new note, click on the New Menu button (first icon on the left) at the top of the sidebar and select New Note. A simple note editing window then opens at the bottom of the sidebar. There are only two controls to this window. A double chevron >> opens the note in a new tab of the browser window and the close bottom closes the edit window and saves the note in the scrapbook.

When you type the note, type your intended title as the first line and you will see it as the title of the note in the sidebar when you close the edit window. When you right-click on a note in the sidebar the context menu is similar to that for a captured document except there is no 'Open source URL' option. If you open the note, it opens in the same edit window at the bottom of the sidebar as when you first created the note.

If however you open it in a tab, it opens in the browser window either in two panels with a HTML display window above and an edit window below or just as an edit window – depending on the state when you last closed it. The edit window in

the browser window has some extra controls.

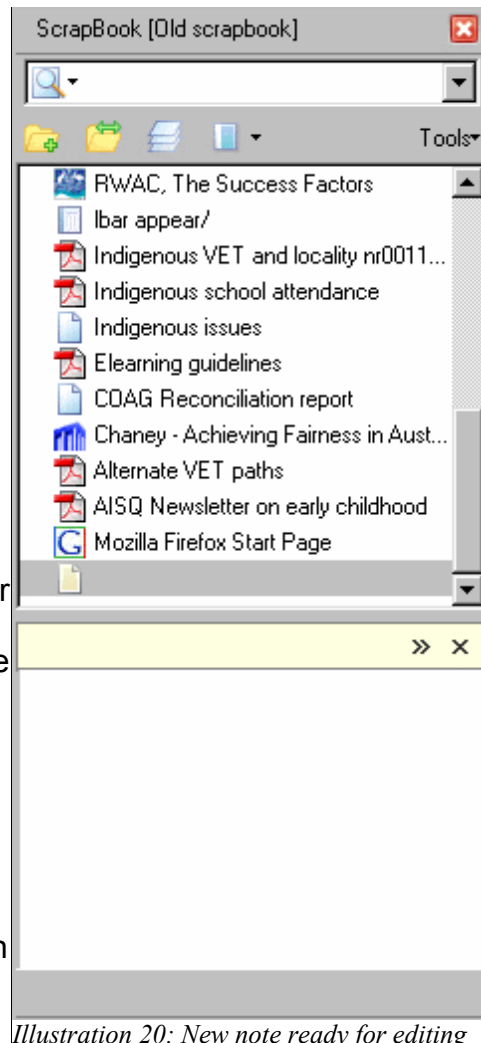


Illustration 20: New note ready for editing

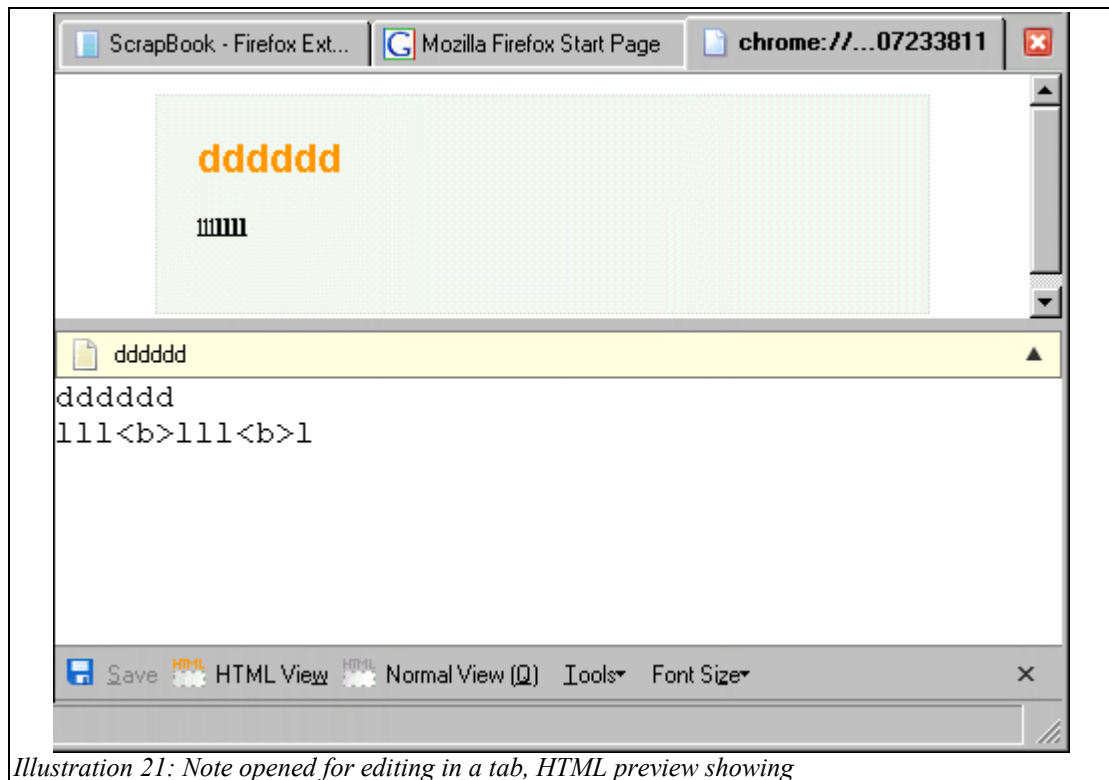


Illustration 21: Note opened for editing in a tab, HTML preview showing

Along the bottom are a Save button; a HTML View button which turns on the two panel display if it is off; the Normal View (Q) button which turns it off, a Tools menu which has the options Help, Auto-linefeed and Edit Template HTML; a menu for choosing Font Size; and a standard Close button. At the top right of the edit window is another button for expanding the Note Edit window thus closing the HTML display window.

The reason for these controls is largely that a note is created and edited as text but stored as HTML. If you chose to display the contents of a folder concatenated into a single page, then any notes in the folder are displayed as HTML. This means that lines in the note separated by carriage returns or line feeds will not display as lines in HTML unless you have selected auto-linefeed when editing. (Selecting it after you have edited will change nothing!) As you can see in the example above you can insert simple HTML markup in the text edit window and preview it in the HTML window. (The latter does not automatically update as you edit.)

As with captured web pages, you can classify your notes into different folders, add comments to them, concatenate them using the Combine utility, etc.

The ScrapBook Tools menu

The Tools button in the top-right of the ScrapBook sidebar gives access to a series of commands that affect the whole scrapbook.

Tools Menu Commands

Settings

This allows configuration of general parameters controlling ScrapBook's behaviour. They will be discussed latter.

Manage

Opens a window displaying the sidebar items in detail in a spreadsheet-like format.

You can set how much metadata is displayed, sort on the metadata by clicking on column heads, delete multiple files or single folders, create new folders, and move multiple selected items simultaneously into and out of folders.

In ScrapBook 1.2 you can do much more than this. To begin with Manage now has a searchbox so that the results of a metadata search can be displayed in the Manage window and selected as a group. This means that you can now move or delete all of the results of a search in a single operation. Even more important, tools such as Combine and Import/Export are now available from the Manage window so that you could combine or export all of the results of a metadata search - or any other selection.

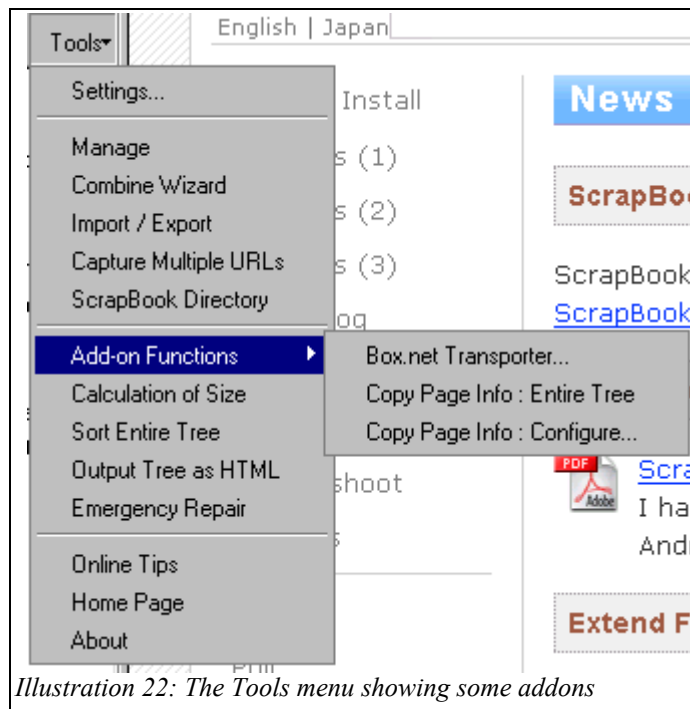


Illustration 22: The Tools menu showing some addons

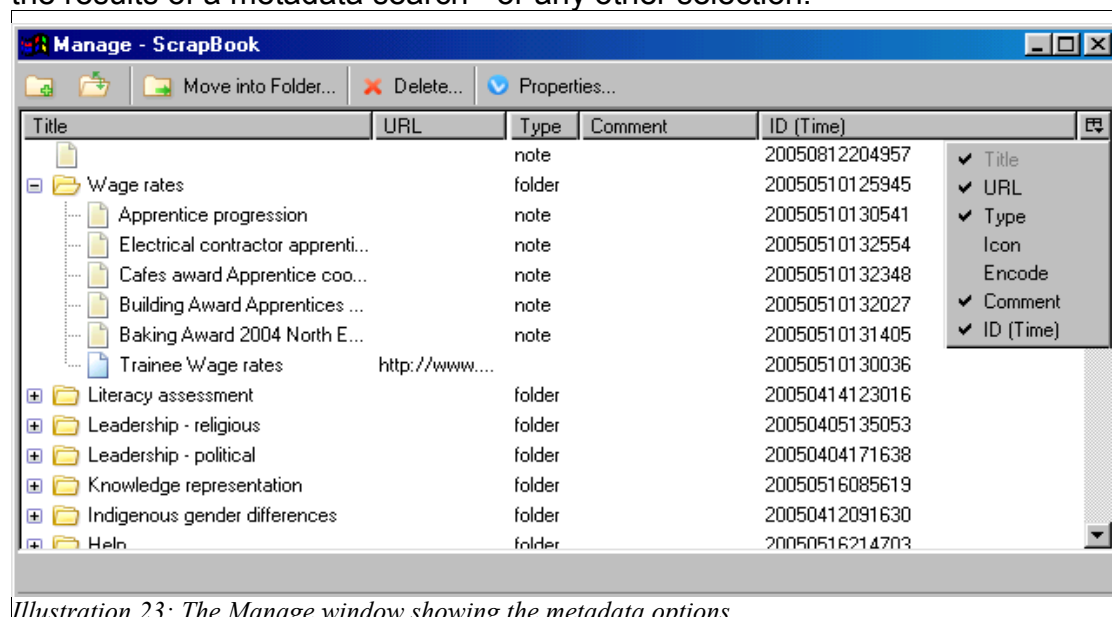


Illustration 23: The Manage window showing the metadata options

Combine wizard

The Combine Wizard opens in the Manage window and can be started either from the Tools menu or from the Manage window. It allows you to combine the scrapbook items you select to display them and save them as a single web page. You can directly combine HTML items and notes. Other file types appear as hyperlinks to the local copy except for bookmarks which are hyperlinks to their source. You can review your combination before saving it to the sidebar as a new item. One use of this is that once you have the pages combined into a single page you can save that externally to ScrapBook using the Firefox File menu. In particular you can usefully save the page as a text file if it is mainly notes. You can choose to remove the source items once the combined file is saved. The titles of combined files are shown in blue. Their URL property is that of the first item chosen. You can combine combined files.

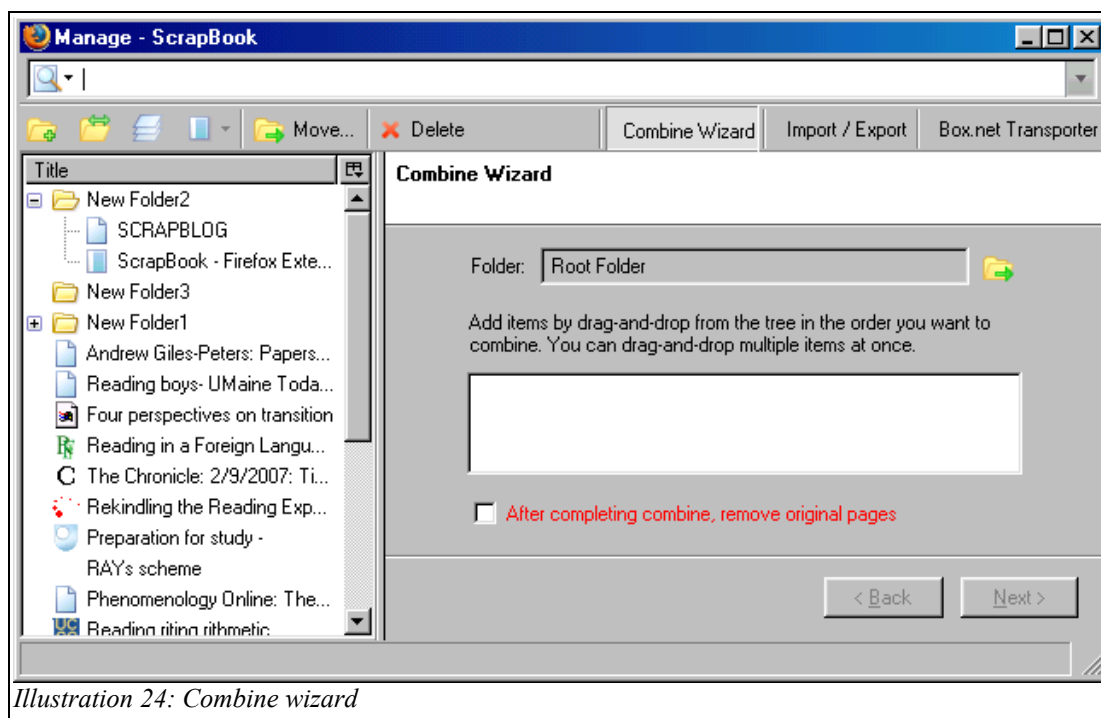


Illustration 24: Combine wizard

Import/Export

This utility also opens in the Manage window . It allows you to export (copy) scrapbook pages to a directory – say on removable media – and import from the directory into another scrapbook. You can export a folder of items and re-import the items as a folder: full “pathnames” are preserved if you ask for this. The folder names of the exported items can be displayed in the right panel of the Import/Export window by using the pulldown menu on the right for metadata columns displayed.

To use the utility set the destination directory (the last one used is saved but may be invalid if it was on removable media) and drag the items you want into it from the ScrapBook item list on the left or just select folders or items and click the Export (>>) button.. To import drag the items from the right panel into the ScrapBook list or select the items and click the Import (<<) button. When you import you can choose to remove the items from the export

directory. If you try to import a page that already exists (same ID) you will get an error message.

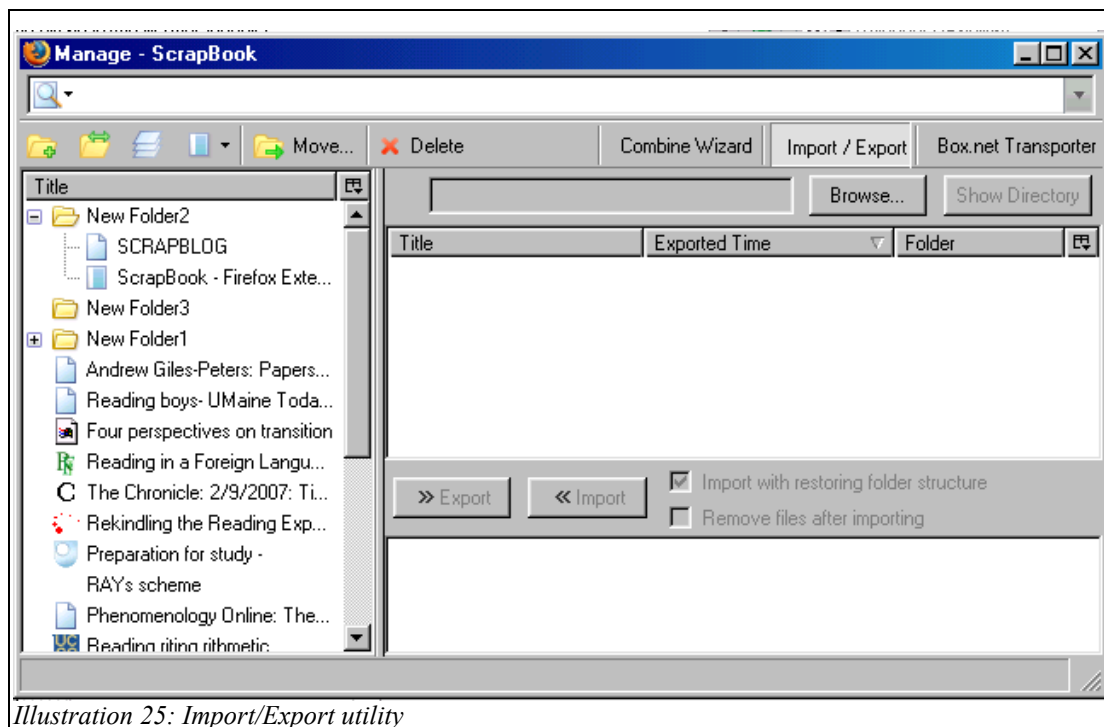


Illustration 25: Import/Export utility

To import items as a folder, select the files in the right panel, check 'Import with restoring folder structure', and click Import (<<). This will re-create the original folders as sub-folders of the Root folder in the left panel if they don't already exist and import the items to them. You can import in this way to several folders at once. If you want the imported folders more deeply nested, you can then drag them to other folders in the left panel. If you import using drag and drop, you can insert the imported items directly wherever you wish.

If you have installed the Multi-ScrapBook extension, its icon appears at the top of the left panel so that you can quickly switch scrapbooks between export and import.

Some users try to import directly from one scrapbook to another. This is not a good idea! It will work (if you open the data directory) because the files index.dat discussed in Appendix 1 are present in the individual item directories BUT they will not have the folder information you need to reconstruct the organisation of the first scrapbook. (This is set by the export process.) Furthermore if you check the setting to delete on import (to simulate a move between scrapbooks) you will certainly delete the data items but leave the references to them in the sidebar of the first scrapbook. This will be rather annoying later on when you get 'File not found' errors. Import/Export is designed to work with an intermediate temporary directory and not with scrapbook directories.

Capture Multiple URLs

This tool opens a window for entering multiple URLs and saving them in the sidebar. You can select the folder you are saving to and make a new folder if required.

You can also do bulk entry by using the URL Detector button.

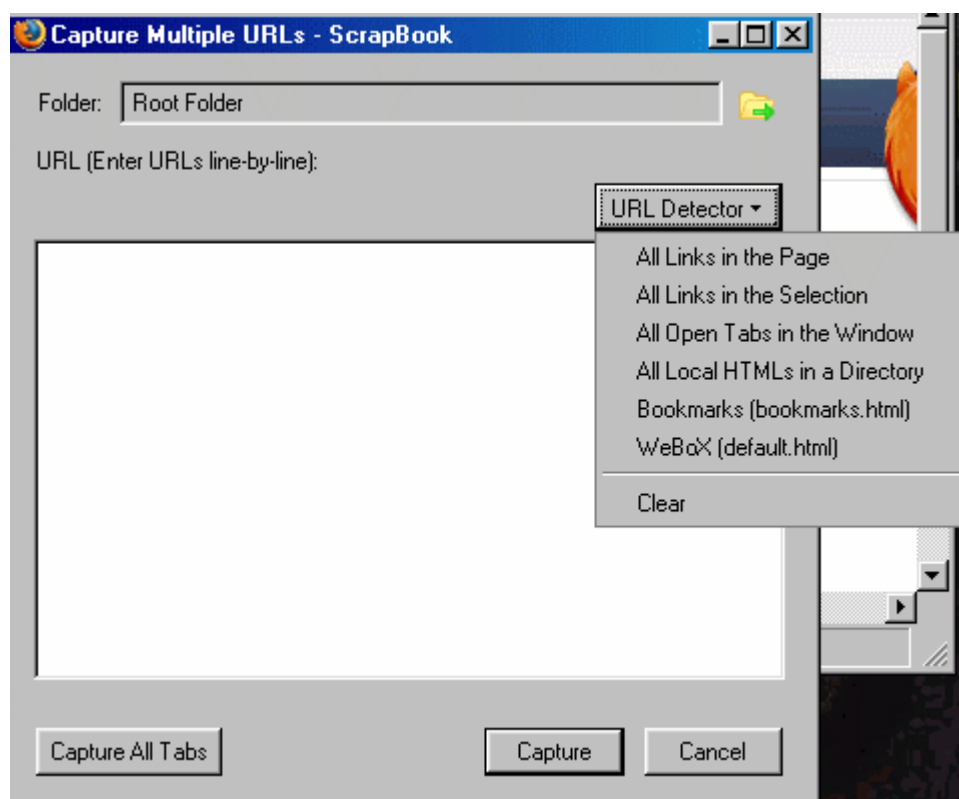


Illustration 26: Capture Multiple URLs window showing automatic entry options

The fourth and sixth choices give rise to a dialogue for selecting the location. WeBoX appears to be a Japanese product in a similar web-capture vein to ScrapBook so this seems to be a way of importing WeBoX's data.

Once you have selected one of these options you can edit the URLs collected in the edit window or clear the window and start again. When you select Capture the same multiple capture window pops up as with Capture in depth.

Scrapbook Directory

Opens the directory of the current scrapbook in the file browser defined in Advanced Settings or in the program default. The directory contains the controlling scrapbook.rdf file and further subdirectories for data, backup etc. If you appear to be missing your data look in here to find out why. (See appendices 1 & 4.) This command is also useful if you don't know where your scrapbook is stored.

Add-on Functions

This submenu will appear if you have installed any add-ons or ScrapBook extensions. Those currently available are AutoSave 1.1, BackupHelper 1.0, CopyPageInfo 1.4.3.2, and ScrapBox.net 1.1. When installed these give rise to commands in the add-ons menu.

Auto Save Mode

Turning AutoSave on allows you to automatically save web-pages visited. Saved pages are stored in a folder named by the date.

Copy Page Info: Entire Tree

Copy Page Info: Configure

Copy Page Info copies the metadata of non-folder items. Copy Page Info is discussed in detail in a later chapter.

Box.net Transporter

ScrapBox.net helps you to transport ScrapBook data to and from the free online storage Box.net. When it is installed there is an additional option to upload the page in the Tools sub-menu of the item context menu. Choosing the Transporter command opens a window into which you can drag items for upload or select items for download to the current scrapbook. Effectively Box.net is used as a remote import/export directory which can be shared with others.

Backup with Compression

Lets you produce a compressed archive of the pages you have captured. You have to specify the program to use in the archiving, the destination of the archive and the command line arguments for the compression. Obviously you have to know a bit about your compression software to do this.

Calculation of Size

Counts the number of scrapbook documents, calculates the space occupied by the totality of files in a scrapbook and may indicate whether the scrapbook is getting too large. If there are INVALID scrapbook items, i.e. documents no longer accessible through the sidebar, it will tell you this too. This can happen if ScrapBook cannot fully delete an item. Normally this is because you have imported an item from some media such as CD or from a directory you don't have write access to. All you have to do is globally set your scrapbook directory to writable and then delete the lost items.

If your back-up directory is getting out of hand, delete the earlier entries. Scrapbook backs up your RDF file (See Appendix 1) daily and should delete entries more than 30 days old but file permissions can interfere with this.

Sort Entire Tree

Sorts all documents in alphabetical order or by date of the capture or saving in either forward or reverse order. You can choose to sort only the top level or all sub-folders recursively.

Output Tree as HTML

Exports the captured pages in the form of an HTML tree structured according to your folders. This can then be viewed by people without ScrapBook or Firefox. You can choose whether to export the whole scrapbook or just some folders – select by clicking, shift-clicking or Control-clicking in the usual manner. You can also choose whether to use frames or not. The frames alternative gives a two panel display reminiscent of the sidebar/browser window combination.

Emergency Repair

The two repairs concern the major problem of a missing or corrupt RDF file and the minor problem of missing favicons in the sidebar. In the former case

a valid copy of scrapbook.rdf is restored from the backup directory. However if you have multiple scrapbooks the problem might be that you have switched to an empty one.

Online Tips

Selecting this opens the “Online tips” page of the ScrapBook web site in a new tab. This mainly concerns mouse and keyboard equivalents of commands.

Home Page

Selecting this opens the ScrapBook web site in a new tab.

About

This tells you exactly which version of ScrapBook you are using and lists the contributors.

The Settings tabs

Choosing the Settings option in the Tools menu gives access to a tabbed form with five groups of settings.

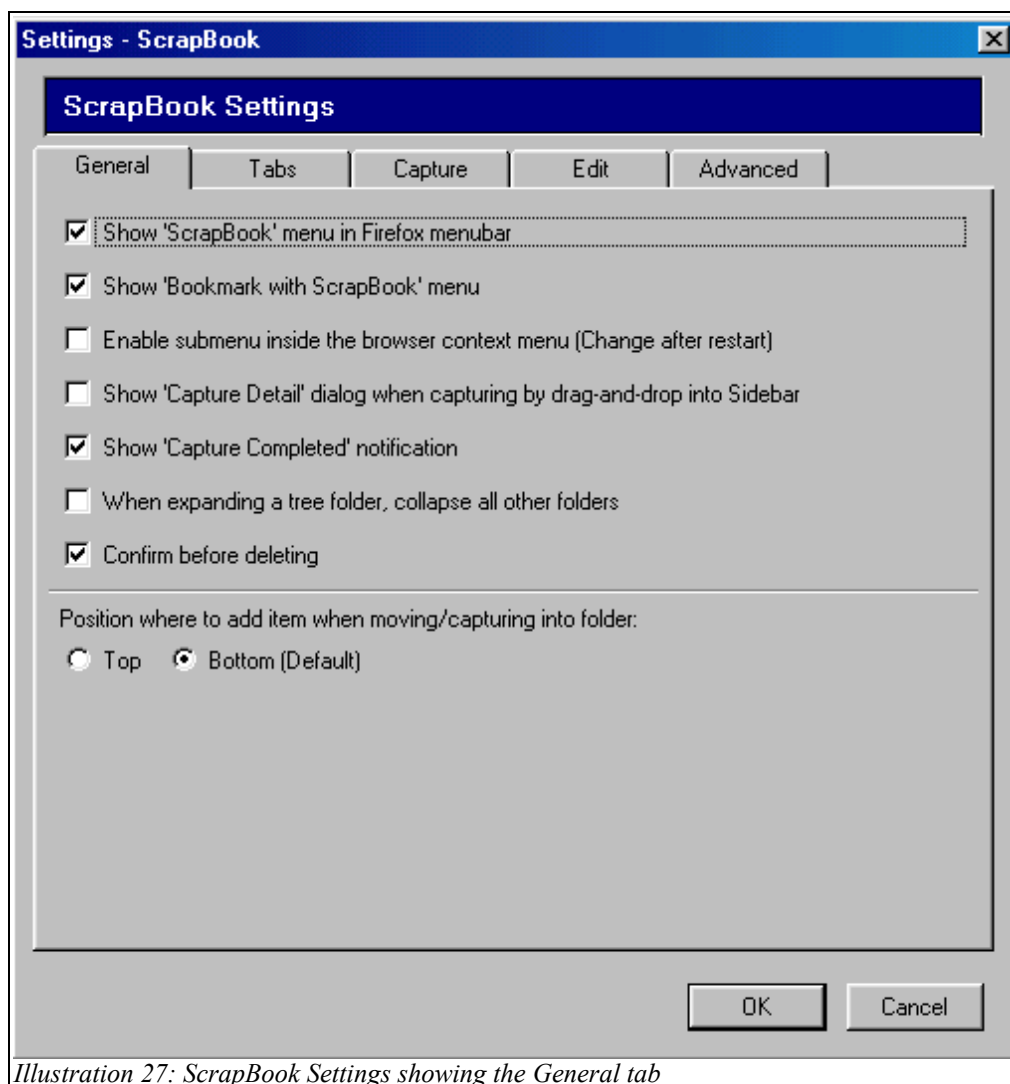


Illustration 27: ScrapBook Settings showing the General tab

General

Most of the settings are fairly straightforward. If you usually capture by drag-and-drop, the third and fourth are probably worth selecting. (The effect of the third option is to slightly remove clutter from the Firefox context menu by replacing the Capture commands by a ScrapBook sub-menu. Obviously you won't do this if you usually use the context menu for capture.)

Tabs

One of the useful features of Firefox is the ability to open several web pages in the browser window with different tabs. This means that you can switch back and forth between different pages without closing any document which is quite useful if you are completing a form whilst consulting other pages. ScrapBook allows you to specify what actions will open a document in a new tab instead of closing and replacing the current document.

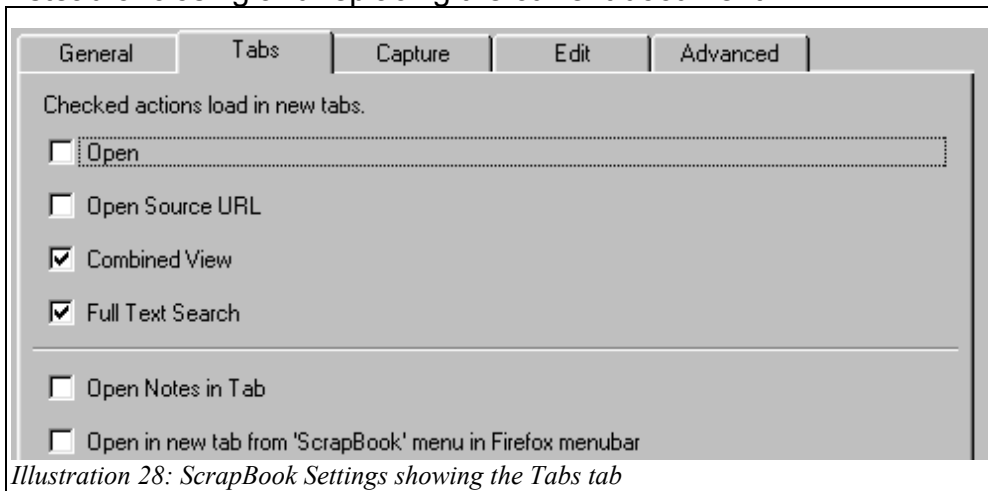


Illustration 28: ScrapBook Settings showing the Tabs tab

Capture

The only option on this panel concerns the conversion of the character encoding on the page. It is recommended to keep the default setting unless you have some good reason not to do so.

Edit

This tab allows you to customise the colours and styles of the four highlighters (so that you can use different ones for different purposes). The Customise buttons give access to eight pre-set styles as well as letting you define your own.

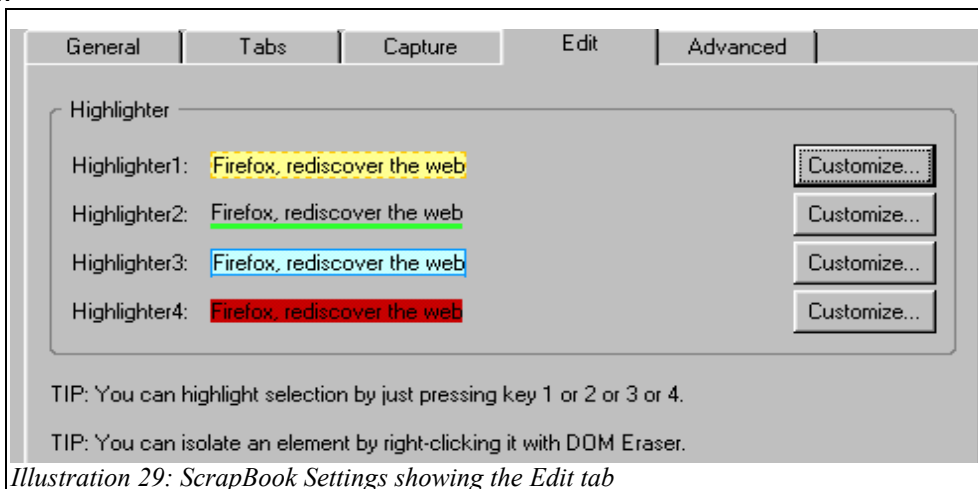


Illustration 29: ScrapBook Settings showing the Edit tab

Advanced

The Advanced tab contains two options: to set the Destination directory where the scrapbook is saved and to set the file browser used by ScrapBook to display directories. You may find however that it only shows the File Browser option if you have opened Settings before closing an Import/Export or Capture Multiple URLs window. ScrapBook sensibly discourages you from switching scrapbooks whilst the current one is in use!

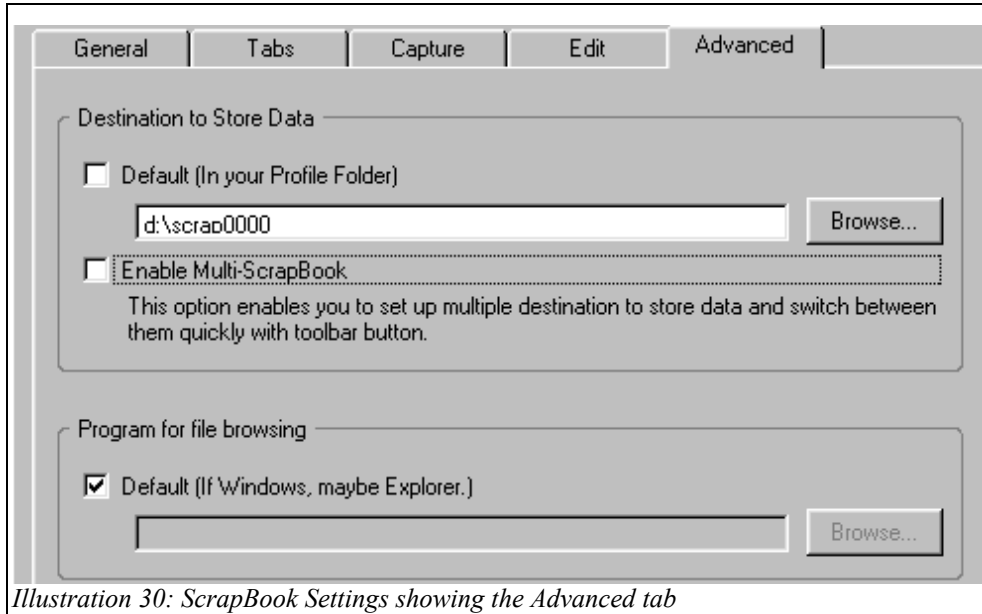


Illustration 30: ScrapBook Settings showing the Advanced tab

If you leave the default checked, ScrapBook saves the pages captured in a ScrapBook directory located under your Firefox profile directory¹⁰. If you uncheck it, you can specify another directory for saving on your hard disk, a network drive or a removable device. If you decide to do this you can copy over the current directory (this is done outside ScrapBook and Firefox) or leave the new one empty and start with a clean slate. You can have as many scrapbooks as you like and switch between them by changing the Destination setting as the directory does not have to be called ScrapBook.. Note however that Multi-ScrapBook is the preferred way of doing this and you can only give your scrapbooks names with Multi-ScrapBook. The Advanced tab method is however useful when you want to use a scrapbook on removable media but won't want to keep it as a permanent option.

If you check Enable Multi-ScrapBook both the Default and alternative Destination will be greyed out. In this case the directory (scrapbook) you are storing data in is given by the option selected in Multi-ScrapBook . The use of Multi-ScrapBook is discussed in the next chapter.

¹⁰ A number of users seem to find this confusing. They think that they are changing the default directory rather than choosing not to use it. When they then use Multi-ScrapBook they expect the Default option to point to their newly chosen directory. It does not and it cannot. What both the Advanced tab and Multi-ScrapBook do is change the current or active scrapbook. The location of this is saved on exit and restored as the active scrapbook on re-starting. There is no default scrapbook to which ScrapBook reverts. All there is is a default initial location. This exists for a good reason. Users should not be forced to choose a scrapbook location. The one they get by default if they do not choose must be in a standard location independent of the users' filing systems. The only thing the installer can be sure of is that there will be directories for Firefox and ScrapBook. Hence these determine the default location.

One point to keep in mind is that the scrapbook name at the top of the sidebar is set by Multi-ScrapBook. Hence if you disable Multi-ScrapBook it will not update if you change scrapbooks by using the Advanced tab.

Another point that may cause you some angst is that if you upgrade from an earlier version of Multi-Scrapbook your scrapbooks may not initially appear in the drop-down menu on the ScrapBook toolbar. In this case choose the configure option (see the detailed discussion of MSB in the next chapter) and hopefully you will see the full list. Now choose one, select Edit and then exit. The scrapbooks should now appear in the drop-down list. If you didn't find them look at Appendix 4 on data recovery.

If you are a Windows user the default file browser will most likely be Windows Explorer which is perfectly adequate. The default on my Linux system has much less functionality so Linux users will probably want to specify their preferred file browser of the many choices available to them.

Using ScrapBook in web-based research

As already explained in the Introduction, ScrapBook does three fundamental jobs automatically:

- It saves the web address
- It saves the web content
- It saves the time of capture

Furthermore it saves these accessibly and all three are searchable.

For some users this will be enough. Others will need more organisation and documentation.

Folders are the most visible form of organisation although document titling comes a close second. The ScrapBook search box gives a third approach to organising documents. A good design principle is to use the three modes to give you independent views of your captures. Since the search box allows you to instantly display all captures made in 2005, May 2005, or 22 May 2005, for instance, there is no point in filing documents by date of capture. Similarly you can instantly display all captures from a particular web address or domain so there is no point in filing by source address. Nor should you file by document type as searching the URL field allows you to pick up file extensions. The exception to this is Notes. Since you cannot automatically find and display them there may be a point in having separate notes folders for documents representing your own thoughts as opposed to snippets taken from elsewhere.

Retrieval by content

Indexing and the Comment field

The intellectual content of documents can be recorded and made retrievable through titling, filing (folder allocation) and indexing. The comment field can be used for quite sophisticated indexing. Since the search box does a phrase search, constructs such as tagged fields, eg 'au:Marx, K', or hierarchical index terms, eg 'politics-socialism-parties', can be used together with descriptive text such as an abstract.

Whether you even use the comment field for indexing or other purposes will depend on the scale and complexity of the research you are doing, the time period over which you will be doing it, and whether you will be sharing your captures with others. If you are part of a collaborative project you should at least store your name or an equivalent code in the comment field. (It would be nice if export did this automatically.)

Folders and classification

If you don't use index terms – and most casual users will not – then the folder system can organise your documents. It has however some important limitations. The first is that you can only have one folder system at a time and a document can only be in one folder at a time. (There is nothing comparable to Windows shortcuts or Linux links in the ScrapBook sidebar.) The second is that you cannot search for folder names and deeply nested folders may be hard to find. The third limitation is that folders are a visual tool for manual

manipulation. The depth of the sidebar limits the number of top-level folders you should have and you should not nest them too deeply either. Hence classification in folders is not a real substitute for indexing if you are interested in exact description and easy or fast retrieval in large document collections.

If you do use folders as your fundamental approach to classification and retrieval then you should name them with some care. If you are sharing your folders you should also document them to explain the rules for filing documents. This information would naturally go in the Comment field. If you rename folders or split them, this historical information should go in the Comment field as well. (The creation date of a folder is always available.)

One good use for folders is as a temporary classification system. Set up the folders and put the corresponding index term into each folder's Comment field. Drag and drop captured documents into the folders. When there are a fair number in each folder go through the folders one by one to put the corresponding index term in the Comment fields of each of the contained documents. To do this open the folder Comment field, select the index term and copy with Control-C. Then go through the documents opening the Comment field, clicking in it to establish focus, and using Control-V to paste. If you then re-classify the documents for another purpose, you can still retrieve them in terms of the original classification.¹¹

Folders and work organisation

Documenting searches

Typically you find the documents you capture by a search but the document you save does not record how it was found. If later you want to know how you found it, and whether your search strategy was complete, you will need to have saved the searches. This can most usefully be done in a folder called Searches. If you capture the first results page of an advanced Google search, say, you will find that the source URL contains the search string, that the Scrapbook title is the text from the search box, and, of course, you have the date of your search. You also have the number of hits and the first group of documents it retrieved on that occasion. Sorted by date your Search folder gives the history of your investigation. You can also annotate your searches by using the comment field.

Organising notes

Another idea is to use folders to group your notes. You might have a single Notes folder or several such as Ideas, Hunches (Hypotheses), and Results. Another useful pair of folders is ToDo and Done. You could even be more elaborate and base a structure on the Plan-Act-Observe-Reflect cycle and shuttle notes between them adding an annotation at each move. Thus:

- Add a description of what you are looking for and refile under Plan
- Add a specification of the search done and refile under Action
- Add a description of the results obtained and refile under Observation
- Add an evaluation of the search and refile under Reflection

(The notes would then be filed as completed plans, actions, observations and reflections but carry their history with them.)

¹¹This is admittedly a lot of work. I have written an extension Set Page Info to add information to selected metadata fields of all items in a folder with a single click.

Of course just as you would only save interesting searches in the Search folder you would probably write up several phases at once and skip uninteresting dead ends. However the structure gives the possibility of industrial strength documentation and work-flow monitoring.

Annotating captures

Using ScrapBook mark-up for active reading

ScrapBook's editing facilities are especially useful to students and beginning researchers. Captured documents are of little value unless they are read. ScrapBook supports an organised approach to reading and note-taking.

In active reading of hard copy you read through once, highlighting or flagging the points of interest and underlining unfamiliar or key terms. You then read through a second time adding definitions and relevant references and writing in questions and responses to the text. Finally you write your own summary or judgement. All this can be done easily to captured HTML documents using the ScrapBook Edit commands. You start of course with the Highlighters.

Highlighting

As mentioned previously there are four highlighter styles active/available at any one time. And you can customise these to your own taste. You can hence colour code your first reading. Obvious things to flag are main points, unknown or obscure terms, sections you need to re-read, points you want to dispute, and so on.

In-line comments and links

In-line comments are useful for attaching definitions to terms. Cross-references can be implemented by attaching links to local files or web URLs. In line annotations are quite unobtrusive: only the dotted underline to the text is visible until you use the mouse to display the comment or attachment.

“Sticky annotations”

These are ScrapBook's electronic version of post-it notes. They are reasonably obtrusive and cover the text beneath them, albeit semi-transparently. On the other hand they are freely movable and resizable. Because they can be attached to the text they are useful for giving a running commentary.

Metadata Comments

The comment field lends itself to general comments on the capture, not tied to particular sections. Whilst it can be read and edited via the Properties dialogue, you can't scroll through a document whilst the Properties window is open. If you click on the comment icon in the Edit Toolbar however you get an expanded comment panel into which you can easily type large amounts of text since it scrolls when it fills. You can also scroll through the captured document whilst you are writing the comment.

There are three good reasons for using the metadata comment field for your summary comments rather than a sticky annotation at the beginning or end of the document. The metadata comment is searchable, displayable without opening the document, and also exportable if you use Copy Page Info.

Multiple scrapbooks

It makes good sense to have a separate scrapbook for each major research project. It is possible to do this by changing the default scrapbook directory under the Advanced Setting tab but Multi-ScrapBook makes it much easier. You can then switch scrapbooks on the fly with a simple menu choice.

The benefits of using multiple scrapbooks for different broad subject or interest areas are that:

- Each sidebar will have fewer folders
- Nesting in each sidebar may be one level shallower
- Different indexing systems can be used in different scrapbooks
- Searching will be more precise as you will not get hits in irrelevant (out of subject area) documents
- Searching will be faster
- You can exchange whole scrapbooks with co-workers by mounting them as additional scrapbooks
- Notes folders and saved search folders will be clearly associated with one subject/interest area

In short your work will be better organised.

Using Multi-ScrapBook

To make most effective use of multiple scrapbooks, enable Multi-ScrapBook on the Advanced tab of ScrapBook Settings. The Multi-Scrapbook menu will then appear in the ScrapBook sidebar.

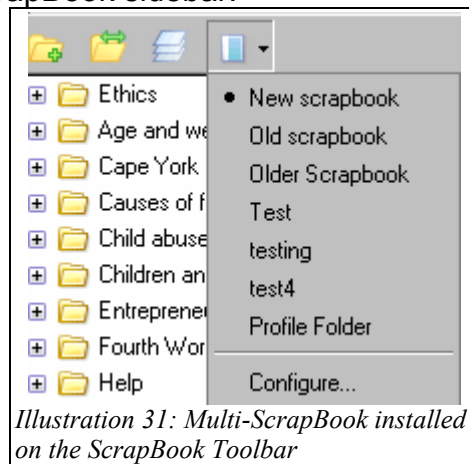


Illustration 31: Multi-ScrapBook installed on the ScrapBook Toolbar

The Multi-ScrapBook menu – accessed by a pull-down menu from the arrow head to the right of the icon - gives a list of the currently installed scrapbooks plus the option 'Configure'. One of the installed scrapbooks will be called 'Profile Folder'. This will be the only one in the list when you first enable Multi-Scrapbook. It is the one under your Firefox profile and cannot be renamed or deleted. Given that it is, or can be, protected under most modern operating systems, you might use it for more private items.

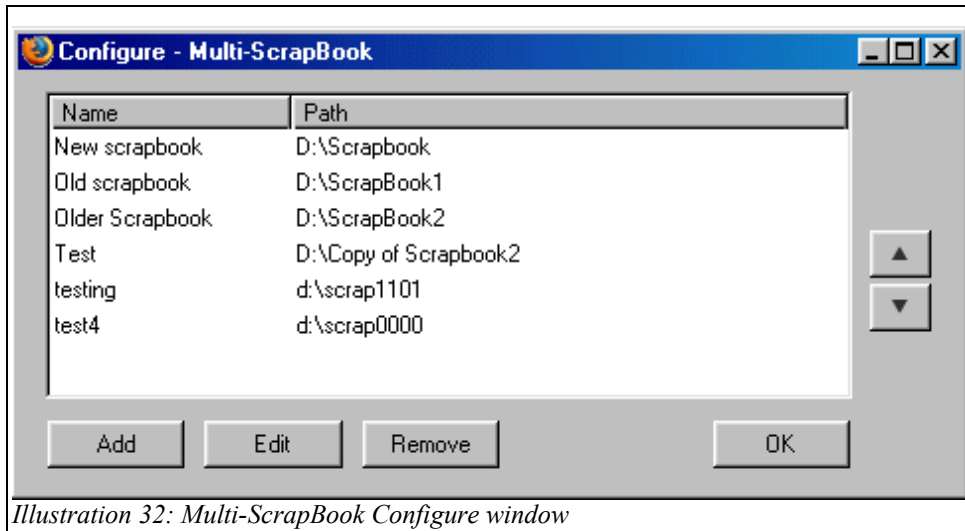


Illustration 32: Multi-ScrapBook Configure window

You add extra scrapbooks by choosing the Configure option. The Multi-ScrapBook Configure window shows the (non-default) scrapbooks you have installed under Multi-Scrapbook. It will not show a scrapbook added under the Advanced tab until you configure it. The reason is that the Multi-ScrapBook menu uses scrapbook names rather than directory names. When you choose 'Add' to add a new scrapbook, you name it and assign its directory. (If it is your current scrapbook, you can discover its directory with the ScrapBook Directory command in the Tools menu.) You can add any number of other scrapbooks with appropriate names – *Project1*, *Project2*, etc – and corresponding directories. (Any non-existent directories will be created and the scrapbook data structures initialised.) Close the Configure window and you will be able to switch between scrapbooks, although the new ones may be empty. Note that because Multi-ScrapBook distinguishes the name of a scrapbook from the name of its containing directory you can name scrapbooks however you like as well as store them wherever you like.

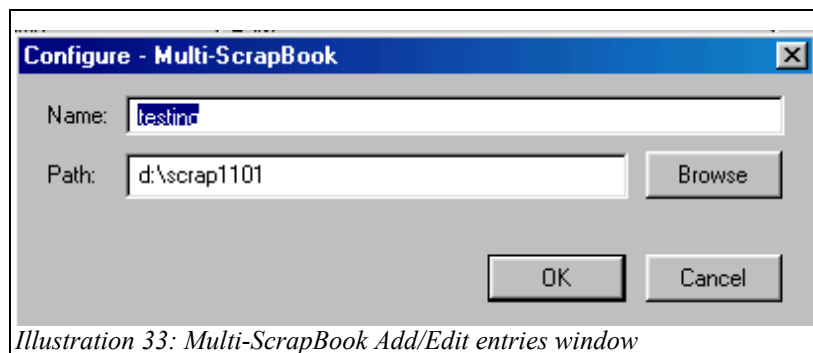


Illustration 33: Multi-ScrapBook Add/Edit entries window

As well as adding new scrapbooks you can edit their names or paths and remove them.

If you want to copy some items to the new scrapbooks, use the import/export utility with a temporary directory. If you want to copy a whole scrapbook to one of the new locations then this is better done at the file system level. Say you have a scrapbook contained in a directory Scrapbook1 that you want to copy to a directory Scrapbook2. Open the Scrapbook1 directory and copy all files and folders. Open the Scrapbook2 folder and paste. Do not do this if

you already have ScrapBook data in ScrapBook2! If you do, the data originally in Scrapbook2 will become inaccessible.

Sharing scrapbooks

Your default scrapbook is stored under your Firefox profile so under any operating system with serious file security it will be inaccessible to other users. If you want your scrapbooks to be accessible you have to set the location using the Advanced settings tab (or Multi-ScrapBook) and put them in shared directories.

There are two scenarios: sharing a scrapbook on the one computer and sharing a scrapbook between two or more computers on a network.

The first scenario gives no technical problems. You only need to agree on the rules about deletion of documents and folders, creation and renaming of folders and re-filing of documents in folders.

There are however technical issues in having a shared scrapbook on a network. What happens, for instance, if two people try to capture at the one time? Each will be running their own browser and their ScrapBook extensions will not be aware of each other. Since ScrapBook is not designed as a multi-user database system it is likely that there could be problems.¹² On the other hand in situations of low usage and infrequent capture there is unlikely to be a problem. One strategy would be to allocate non-overlapping times at which different users could update the shared scrapbook.

Another solution is to have multiple scrapbooks, each having a designated owner. Make all the scrapbooks available to the group using shared directories and Multi-Scrapbook but let only the owner update any particular scrapbook, whether by capture or export and import from the other scrapbooks. This could be enforced through file permissions under Linux and some later Windows systems. A variant of this is a master scrapbook and feeder scrapbooks – a system that might make sense for a researcher with assistants.

Sharing data with ScrapBox.net

With the ScrapBox.net add-on to ScrapBook it is now possible to share Scrapbook data over the web as well as over a local file-sharing network. You install ScrapBox.net, create an account with the free online storage service Box.net, and then upload Scrapbook data using the Box.net Transporter. Anyone supplying the username and password can then download the data to their scrapbooks using the same add-on.

This suggests the following collaboration scenario. The data collector creates a Box.net account and notifies her collaborators by email of the email address and password used to set it up. The collaborators install ScrapBox.net and configure their Transporters to use that account. They then upload data to it. The collector downloads the data and removes it from the ScrapBox. Other scenarios are obviously possible.

¹² If, for instance, ScrapBook put an exclusive file lock on the scrapbook folder when launched this would be a major access problem, an exclusive write lock would be less of a problem. Having no locks would create an inconsistent update problem.

One drawback of the current early version of ScrapBox.net is that it does not save and restore folder information. Hence the classification of items is lost unless it is written into the comment field – a rather tedious process if you aren't already using a keyword indexing system.

One point to remember about the free basic Box.net service is that everyone with the username and password has the same rights, including delete. If you want to differentiate permissions you need a paid plan that allows sub-users. A \$50 pa plan includes three sub-users (ie three distinct groups of permissions) which is surely enough for most purposes.

Exporting text and references from ScrapBook

ScrapBook does a brilliant job of capturing, saving, classifying and documenting web data but researchers also need to produce reports on their data. With the latest version of the ScrapBook Copy Page Information add-on you can list all items in a scrapbook or export all comments as a text file. You can also collect all your notes together using the Combine utility and save them as a text file.

Exporting references

A typical requirement in research work is that you list your sources in an appendix to your report. If your sources are ScrapBook items you will want to export a file of titles, URLs and capture dates for loading into your wordprocessor.

A less usual requirement is for a commentary or annotated bibliography. For this you would want to export a file of titles, URLs and comments.

Both of the above jobs can be done easily with the Copy Page Info extension. All you have to do is configure it to copy the required metadata, apply the Copy Page Info command either recursively from the root or to individual folders, and paste the copied data from the clipboard into a document.

Installing and configuring Copy Page Info

Copy Page Info is a separate program from ScrapBook and should be installed after you have installed ScrapBook. Go to the ScrapBook site and choose Add-ons from the site menu. On the Add-ons page you will find Copy Page Info. Double click on the link and follow the instructions. (Since you install Copy Page Info from the ScrapBook site you will have to add the site to your “white list”. To do this choose Edit Options when the alert message pops up and then click on Allow when the Allowed Sites dialogue appears. Then choose Install Now when the Software Installation dialogue appears.) Once it is installed open the Firefox Extensions window (choose Tools- Extensions), select Copy Page Info and choose options to configure it. Alternatively go to the **Tools** menu in the ScrapBook sidebar and choose **Add-on functions** and then **Copy Page Info: Configure**.

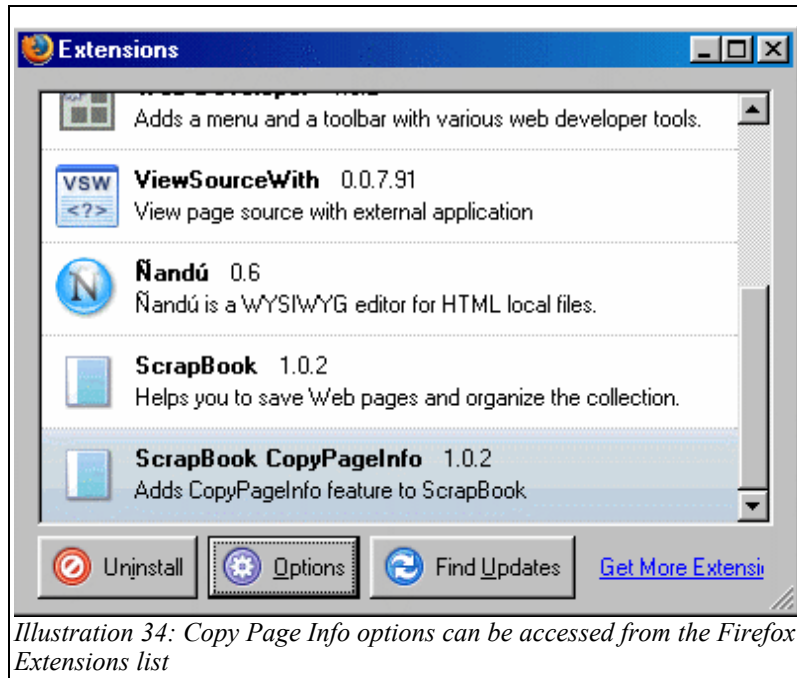


Illustration 34: Copy Page Info options can be accessed from the Firefox Extensions list

Copy Page Info gives access to the metadata items you view through the Properties command in the sidebar item context menu. In particular it will copy title, source URL, date of capture and comment but also some others. The default format is multi-line but you can change this to a single line if desired. When you click on Options above you get an edit window containing the existing settings and a list of available metadata items. What you enter in the edit window is used as a template for each item's data so you can insert any printable characters desired, such as commas instead of carriage returns, between metadata items. This is sufficient if you are just going to import to a Wordprocessor. Copy Page Info includes some prepackaged templates that illustrate the sorts of things you can do.

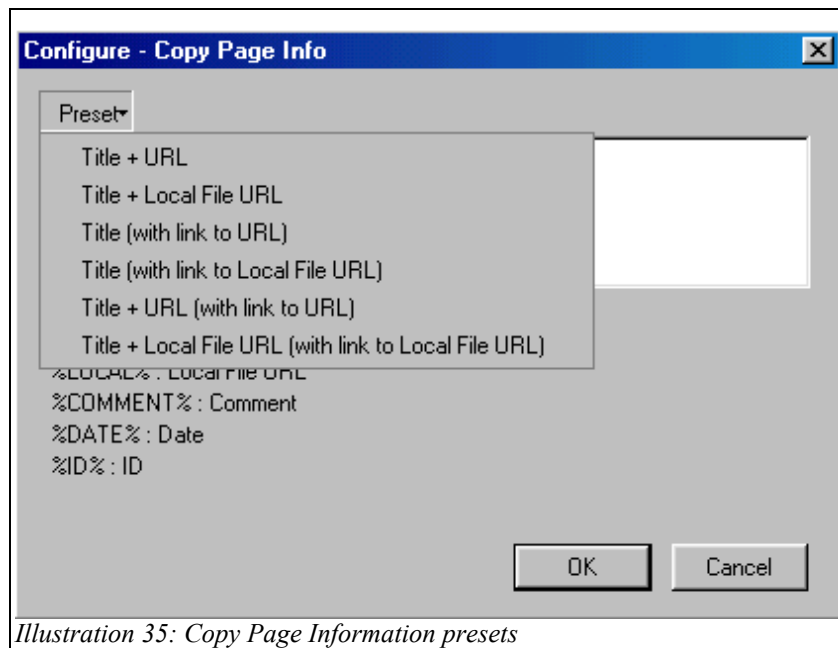


Illustration 35: Copy Page Information presets

If you need to import your reference list to a database or bibliographic program, then Tab Separated Values (TSV) is probably the most useful format. To produce this put the %TAB% metavariabale between each of the desired data values and remember to finish the line with a carriage return.

The %FOLDER% metavariabale will also be useful if you are importing to database but less so if you are importing to a wordprocessor.

Using Copy Page Info

After installation the Copy Page Info command will appear in the Tools menu under the Add-ons section, in the Folder context menu, and in the item context menu. There is also fast access to Copy Page Info configuration via the Add-on Functions section of the Tools menu.

There is no obvious feedback when you choose the command; the information requested is silently copied to the ClipBoard. To use it open any editable document and paste it in. You could for instance paste it into a ScrapBook note, a text editor or a wordprocessor document. Note that each time you use the command you overwrite the previous data.

With the Tools Add-on Functions menu command you can copy metadata from root folder items or recursively through the whole tree by middle-clicking the command instead of left-clicking.

With the folder context menu command you can copy from all items in that folder or recursively down the tree from that folder – again by middle-clicking.

With the item context menu command you can copy the data for that item only.

Exporting text

If you make heavy use of the Notes facility to record ideas or comments there will come a time when you want to collect these notes and drop them into a wordprocessor for further elaboration and formatting. Previously you could only do this by individually copying and pasting the note contents. Now you can use the Combine utility.

Whilst the Combined View utility just concatenates references for display by Firefox, the Combine utility in the Tools menu concatenates the contents of items. It does this both for HTML files and notes. As a result a combine file saved to the scrapbook can be saved to disk as reasonably useful text.

Using the Combine Utility

Unlike Combined View and Copy Page Info, the Combine utility does not operate on whole folders. Rather Combine requires you to select items individually in the order in which you want them to appear. You can choose whether or not to delete the original items after they are combined. When you have finished selecting and choose Next, Combine displays the resulting combined document for Preview. You can then either save it as a new scrapbook item by choosing Combine or Cancel the whole process.

The combined document is saved with the same title as the first item but is in

a blue font to distinguish it. If you click on it, it opens as a standard scrapbook item even if composed of notes. Hence it can be saved as a text file from the Firefox File menu or you can use Edit - Select All and Copy and dump it into a wordprocessor. You can also print it directly from Firefox. Obviously you can rename it.

Another important point is that combined files can themselves be combined so that you can build up a document incrementally within ScrapBook. For instance you may combine all the notes in a folder to a single item and then combine all the combined files in a group of folders.

In deciding whether to use the delete option on the original notes and intermediate combined files there are some conflicting considerations. Using the delete option tidies things up and you can see what notes you haven't used. The combine file keeps the note title as a sub-heading so you can identify the separate notes combined later on. However you lose the other metadata, in particular the date that you wrote the note. If this is important you shouldn't use the delete option or instead use the F5 key to timestamp your notes as you write or edit them. Another point is that combine files are not editable in the same way as notes although if you use an extension like ViewSourceWith you can directly edit ScrapBook items with an external editor. Overall the neatness point probably wins.

Appendices

1. ScrapBook file structures

A scrapbook is a directory containing at a minimum two directories and one file:

- backup
- data
- scrapbook.rdf

There might also be other files such as the `note_template.html` file or the `cache.rdf` used for full text searching.

The **backup** directory contains files with names like `scrapbook_yyyymmdd.rdf`. When you launch ScrapBook the `scrapbook.rdf` file for the current scrapbook is copied to the backup directory under the name `scrapbook_[today's date].rdf`. Backup files more than 30 days old are deleted – or will be if they or the directory are not marked 'Read Only' as might be the case if you have copied a whole scrapbook from a medium like CDROM. (If backup files are not being deleted, check that the Read Only bit for the backup directory properties is cleared.) If your scrapbook becomes corrupted you can try to restore it by copying back an earlier version to `scrapbook.rdf`. (There is an option under the ScrapBook Tools menu to do just this.)

The **data** directory contains directories with names like `yyymmddhhmmss`, ie the directories are named by the ScrapBook document ID. Each of these contains an `index.html` file and various image, CSS and other files depending on what is in the capture. Scrapbook displays documents by loading the html file into the Firefox browser and letting Firefox do all the rest.

The `scrapbook.rdf` file contains the metadata descriptions of captured documents and the current folder structure. This is in a particular XML format known as Resource Description Framework. You edit this file indirectly by changing the folder organisation and properties through the ScrapBook sidebar. Do not try to edit it directly or you will be needing the backup!

If you restore an earlier `scrapbook.rdf` file you might find that some of your captured documents no longer appear. They may still be there in the data directory but not referenced in the RDF file. You can pick these orphans up when you run the scrapbook size calculation as they will be detected as invalid items. To restore them you can use the trick in the next section – providing that you know their titles.

Since many of the scrapbook data files are small, the disk space used by ScrapBook is about twice the amount of data actually involved. A mixed collection of 200 items can easily occupy 90 megabytes in file systems with large allocation blocks. Given today's large disks this hardly matters. You are more likely to encounter problems with keeping track of your captures. A 1500 item scrapbook might back up to a CD, if you were short of space, but would be quite difficult to browse that number of items in the sidebar. By this stage you would need indexing and/or multiple scrapbooks to keep track.

Export file structures and index.dat

When you create a new item ScrapBook creates a file index.dat in the item directory labelled by the ID to record the metadata. Strictly this is redundant as it is what is available in scrapbook.rdf. When you export an item ScrapBook copies the item directory labelled by the ID to an export directory labelled by the item name (with underscores replacing spaces) after updating index.dat to reflect perhaps new folder information. For example, the item Aardvark Keys is exported as a directory Aardvark_Keys and index.dat contains the following:

```
id      20050726235532
type
title   Aardvark Keys
chars   UTF-8
icon    favicon.ico
source   http://karmatics.com/aardvark/
comment
folder  Help
```

Note that since ScrapBook saves and restores complete folder structures, the Folder data item will be a tab separated list of folder names for a nested folder.

When the item is imported elsewhere, index.dat is copied over as well but does not appear to be used after the RDF file has been updated with the new metadata. In particular ScrapBook does not update it when the item is moved to a new folder.

Index.dat can be used to recover invalid items by re-import. To do this just make the Import/Export directory the data directory of the scrapbook you are restoring and import to that scrapbook. (The fact that this works shows that the utility does not look at the directory names but only at index.dat.) Do not check the 'Delete after Import' setting!

The same trick can be used to merge two scrapbooks. Open one of them, open the Import/Export window, and set the directory to the data directory of the other scrapbook. Then select all the items in the import window and import (again without the 'Delete' setting). The items already in the first scrapbook will be skipped and the others imported to the root folder (unless they were previously imported from elsewhere). If you want to import with valid folder information you have to export the second scrapbook to a temporary folder and then re-import it as the current folder information is only written to index.dat at the moment of export.

2. Regular expressions and searching

The table below gives a partial syntax for the regular expressions that work in search expressions.

Character	Description
\	Marks the next character as a special character or a literal. For example, '?' matches the preceding character zero or one times. '\?' matches a literal '?' character. On the other hand 's' matches 's' but '\s' matches any space character.
^	Matches the position at the start of the string.
\$	Matches the position at the end of the string.
*	Matches the preceding character or subexpression zero or more times. For example, zo* matches "z" and "zoo".
+	Matches the preceding character or subexpression one or more times. For example, 'zo+' matches "zo" and "zoo", but not "z". + is equivalent to {1,}.
?	Matches the preceding character or subexpression zero or one time. For example, "do(es)?" matches the "do" in "do" or "does". ? is equivalent to {0,1}
{n}	n is a nonnegative integer. Matches the preceding character exactly n times. For example, 'o{2}' does not match the 'o' in "Bob," but matches the two o's in "food".
{n,m}	m and n are nonnegative integers, where n <= m. Matches at least n and at most m times. For example, "o{1,3}" matches the first three o's in "fooooood". 'o{0,1}' is equivalent to 'o?'. Note that you cannot put a space between the comma and the numbers.
.	Matches any single character except "\n". To match any character including the '\n', use a pattern such as '[\s\S]'.
x y	Matches either x or y. For example, 'z food' matches "z" or "food". '(z f)ood' matches "zood" or "food".
[xyz]	A character set. Matches any one of the enclosed characters. For example, '[abc]' matches the 'a' in "plain".
[^xyz]	A negative character set. Matches any character not enclosed. For example, '[^abc]' matches the 'p' in "plain".
[a-z]	A range of characters. Matches any character in the specified range. For example, '[a-z]' matches any lowercase alphabetic character in the range 'a' through 'z'.
[^a-z]	A negative range characters. Matches any character not in the specified range. For example, '[^a-z]' matches any character not in the range 'a' through 'z'.
\b	Matches a word boundary, that is, the position between a word and a space. For example, 'er\b' matches the 'er' in "never" but not the 'er' in "verb".
\B	Matches a nonword boundary. 'er\B' matches the 'er' in "verb" but not the 'er' in "never".
\d	Matches a digit character. Equivalent to [0-9].
\D	Matches a nondigit character. Equivalent to [^0-9].
\s	Matches any white space character including space, tab, and so on.
\S	Matches any non-white space character.
\w	Matches any word character including underscore. Equivalent to '[A-Za-z0-9_]'.
\W	Matches any nonword character. Equivalent to '[^A-Za-z0-9_]'.

(Adapted from *Introduction to Regular expressions*, Microsoft Developer Network Library, 2005)

Regular expressions obviously allow very precise searching. They do not however easily translate standard bibliographic searching. It is easy to search for 'apples OR pears' – as a first approximation use the regular expression '(apples)|(pears)' – but less easy to do 'apples AND pears'. For this we must use something like '(apples\b.*pears\b)|(pears\b.*apples\b)' A search for 'apples NOT pears', i.e. 'apples -pears', is probably untranslatable into regular expressions.

Full text searching without regular expressions uses lists of included and excluded terms (which may be phrases in double quotation marks.) There must be at least one included term. To exclude a term simply prefix it with '-'. Scrapbook then logically ANDs the search conditions. So for full text searching use regular expressions if your interest is in alternative words or phrases and standard search if you want to examine co-presence or absence of words or phrases.

3. ScrapBook Keyboard and Mouse Assignments

These are listed and illustrated in the Online Tips page of the ScrapBook site. I just list them here for completeness. I have not attempted to treat them in the main text. Most of them are merely alternative to using the various menus but some – eg Isolate – can only be done this way.

General

Alt-K: Toggle ScrapBook sidebar

Browser Context Menu

[Capture Page] > middle-click submenu: Capture Page As...

Same applies to Capture Selection, Capture Frame and Capture Link.

Capture Detail dialog

Middle-click folder list: Select the last used folder

Drag-and-drop on Sidebar

Drop icon in location bar with holding Alt key: Bookmark with ScrapBook

Drop icon in location bar with holding Ctrl or Shift key: Capture Page As...

Drop selection with holding Ctrl or Shift key: Capture Selection As...

Drop link with holding Ctrl or Shift key: Capture Link As...

Editing Saved Page

Alt + A: Create a new sticky annotation

Alt + Z: Undo

Alt + S: Save

Select text and Alt + Q: Highlighter

Select text and press 1, 2, 3 or 4 key: Highlighter

Right-click with DOM Eraser: DOM Isolator (Isolate the clicked element)

Select text and middle-click the textbox for title: Set the selection to textbox

Click the comment icon: Expand/collapse the comment area

Tab/Shift Tab: Steps through the item

Backspace: Opens the source URL

Firefox ScrapBook menu

Middle-click [Capture Page]: Capture Page As...

Middle-click a folder: Open Sidebar and locate the folder

Middle click ScrapBook entry to open the Manage window

Info Toolbar

Click on Source URL: Opens Source URL in same tab

Manage window

Ctrl-A: Selects all visible items

Searchbar

Enter 'f': Change search method to [Full Text Search]

Enter 't': Change search method to [Title]

Enter 'c': Change search method to [Comment]

Enter 'u': Change search method to [URL]

Enter 'i': Change search method to [ID (Time)]

Enter 'a': Change search method to [All]

Sidebar

F2 Key: Properties

Del Key: Delete

Middle-click: Open in New Tab

Click with Shift Key: Open in New Tab

Click with Ctrl Key: Open in New Tab

Middle-click a folder: List View

Middle-click [Capture Again...]: Capture Again without 'Capture Detail' dialog

4. 'Help, I have lost my data!'

If you think you have lost your data DON'T PANIC. It is actually rather unlikely.

To really lose your data you have to delete it at the operating system level and then clear the recycle bin – unless it is a network drive but in that case there is probably a backup anyway. Of course if you delete a folder full of items ScrapBook will do what you ask but even then you should be able to recover the deleted files and restore.

The first thing is to decide what the problem is.

1. The side panel is empty of items.
Click on the Tools menu and chose Scrapbook Directory. If there are not the data structures mentioned in Appendix 1 then you aren't really connected to a scrapbook. If the data directory contains item directories then try Emergency Repair to restore a valid scrapbook.rdf. If there are now items in the sidepanel run Calculation of Size to pick up invalid items. Use the Import trick in Appendix 2 to restore any invalid items detected.
2. The side panel has items but not all those I expect.
Run Calculation of size to see if you can pick up invalid items. Otherwise proceed as for 3-5.
3. The side panel is empty and the data directory is empty.
Are you sure it is the right scrapbook? Are you using Multi-ScrapBook? If so, check the title and check your other scrapbooks. If not, choose Tools-Settings-Advanced tab and see whether you are using the Profile folder default or another location. (If you have used another location this will be shown even if greyed out. Switch between it and your profile folder to see if you can find the data.) Upgrade time is when this problem is most likely to happen.
4. You still can't find the data.
Enable Multi-ScrapBook, search for scrapbook.rdf at the operating system level and add each directory containing a scrapbook.rdf to Multi-Scrapbook if it is not already in the list. Check each scrapbook.
5. Still missing the data.
Restore all directories with ID names (yyyymmddhhmmss) and files called scrapbook.rdf from the recycle bin/trash can. Repeat step 3. If the data is still missing run the Calculation of Size on each scrapbook and import the invalid items. Now STOP – you can do no more!

If you get an error which says "ScrapBook ERROR: Failed to initialize datasource."...

1. Check if your scrapbook.rdf and the destination directory are neither Read-Only nor Hidden Attribute (especially for Windows users). Otherwise, uncheck both attributes and restart Firefox.
2. Check if you have a permission to read/write in the destination directory (especially for Linux users). Otherwise, change the destination directory to a correct one.
3. Still getting the error? It is highly possible that 'scrapbook.rdf' has collapsed. Try to restore from backup by 'Emergency Repair' tool.

5. Getting more information

The ScrapBook site - <http://amb.vis.ne.jp/mozilla/scrapbook> – has some information on modifying ScrapBook's behaviour in the section **FAQ, Tips**. There are also some pages of general information – the **Features** pages – but most of that material has been covered in this Tutorial.

If you are having problems, or have a complex query, then the Forum on the ScrapBook site is the place to ask questions and raise issues. The developer Gomita monitors the list but usually only replies to suggestions and questions about bugs. General help is usually supplied by other users.

Another page of the ScrapBook site you might like to visit is the Questionnaire. Gomita often holds votes on how the user interface should work after issues have been raised in the Forum.

For formal reporting and discussion of bugs there is a link to Bugzilla on the ScrapBook home page.