LINUXUSER



Turn Firefox into an archiving and research tool with ScrapBook WORKSPACE

A handy Firefox extension called ScrapBook lets you save, manage, and annotate web pages. **BY DMITRI POPOV**

often hear readers ask, "What's all this fuss about Firefox? It's just a web browser." Well, not quite. Firefox supports some powerful extensions that make it more than just a browser. One extension that deserves a closer look is ScrapBook [1]. ScrapBook is a convenient tool for saving, manag-

ing, and editing html pages. If you dig a little deeper, you'll discover some practical features that make ScrapBook a useful aide for archiving and research.

Getting Started

To install ScrapBook, point your browser to ScrapBook's website [1] and click the download link. Then restart the browser, and you are ready to go. You can access all ScrapBook's features via the Scrap-Book menu, but you might also want to add a button to Firefox's main toolbar. To do this, right-click somewhere on the toolbar, select *Customize*, then drag the ScrapBook button onto the toolbar and press *Done*. Now you are ready to capture web pages. ScrapBook offers several ways of archiving the page you are currently viewing. Probably the easiest way is to drag the page's URL from the *Address* field onto the ScrapBook's sidebar. You can also choose *Capture Page* from the Scrap-Book menu. This option captures the entire page using the default settings. If you need more control over the capturing process, choose *ScrapBook* | *Capture Page As*. The *Capture Detail* dialog box allows you to specify several capturing options. By default, ScrapBook doesn't



Figure 1: ScrapBook installs as an unobtrusive sidebar.



Figure 2: The Capture Details dialog box gives you more control over the capturing process.

download linked files; you can change that by ticking the appropriate check boxes or specifying custom settings. For example, if the page links to OpenOffice. org documents you want to download, tick the *Custom* check box and enter odt, ods, or any other OpenOffice.org file extensions into the field.

Another important setting defines how many linked pages ScrapBook should capture; you can specify this setting by choosing the desired "link depth" in the Depth to follow links section. If you set the level deep enough, you can capture the entire website. This feature can come in handy if you want to keep a copy of the entire site for off-line viewing. However, you should keep in mind that if the site contains thousands of pages, the capturing process may take a long time, and you can end up with a huge archive. If you configure ScrapBook to follow the links on the page, it will also generate a sitemap that you can use to navigate the captured site.

The third option, *Capture All Tabs*, allows you to capture all pages in the currently opened tabs at once. But what if

Where Does the Data Go?

ScrapBook doesn't lock your data into some obscure archive format: all the captured pages are neatly saved in separate folders, and you can view the pages directly in your browser without ScrapBook. It's reassuring to know that, even if something happens to the Scrap-Book application, you will still be able to access and use your data. you want to capture multiple pages without opening them in the browser? No problem: in the sidebar, select *Tools* | *Capture Multiple URLs*, type the URLs you want to capture, and press *Capture*. The *Capture Multiple URLs* window includes a URL Detector that can capture all links in the open page or links in the current selection.

More Features

ScrapBook can capture not only web pages, but also text snippets, which effectively turns it into a useful notebook tool. To capture a text fragment, select it and drag it onto the ScrapBook's sidebar. But that's not all: ScrapBook can handle links and frames as well as any files supported by Firefox (such as PDF, Flash, and XML). Moreover, instead of capturing the entire page, you can bookmark it, which means that you can use Scrap-Book as a bookmark manager.

Like any archiving tool worth its salt, ScrapBook offers several ways of managing the captured pages and text snippets. The most obvious means for managing entries is with folders: click the *New Folder* button in the sidebar to create a folder, and give the folder a name. You can then move captured items into the folder using drag-and-drop.

ScrapBook also includes a powerful manager *Tools* | *Manage*, which has a couple of clever features of its own. Using the Combine wizard, you can merge multiple pages and notes into one page. Say you have multiple text snippets related to a project you are working on. The easiest way to view them is to

put them on one page. Start the Combine wizard and select the notes you want to merge; the tool takes care of the rest. The manager also allows you to import and export your Scrap-Book data. You can use this feature to backup the contents of ScrapBook, as well as to transfer the contents to another computer. Better yet, you can export and import

data selectively, meaning you can import and export only certain pages or folders.

If one ScrapBook is not enough for your needs, you can create several repositories and easily switch between them. To enable this feature, choose *Tools* | *Settings* and tick the *Enable Multi-Scrap-Book* check box. You can then add other data directories using the *Profile* button on the toolbar.

Search

ScrapBook includes a powerful Search feature that you can use to search inside the captured pages and notes. For example, if you want to find all the pages where the word "Ubuntu" occurs, type it into the search field and hit Enter. Scrap-Book then displays the Results page divided into two panes: the upper pane contains the found pages, while the lower pane displays the contents of the currently selected page with the search term highlighted. Instead of a full text search, you can make ScrapBook look for the specified search term in titles, URLs, comments, etc. ScrapBook even supports regular expressions.

Also Editing

As you can see, ScrapBook includes almost every imaginable feature for storing and managing web pages. But that's not all. What makes ScrapBook a truly unique extension is its editing tools. You can use ScrapBook's editing tools to modify and annotate captured pages. When you open a captured page in the browser, ScrapBook displays the *Edit* bar at the bottom of the main window. The

| Contract Contract | | | | 104340 | | |
|---|--|---|---|-------------------|--|--|
| -1 | | | | | | |
| G C - GHee. | a Califia | Continue stand | impot/Expert | Bound Transporter | | |
| Title | LPL. | | 7,0+ | [m | | |
| Stating Solution S | http://www.inuer.ne http://www.inuer.ne http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. http://www.inuer. | wiferge zonosnicke pitral wiferge zonosnicke pitral wiferge zonosnicke pitral mail converte by 1008 wifer converte by 1008 wiferge zonosnicke wifer pipel 2004 (2015) strate o wiferge zonosnicke app pipel zonosnicke by 2005 wiferge zonosnicke pitral pipel zonosnicke by 2005 wiferge zonosnicke pitral pipel zonosnicke by 2005 wield a model by 2005 wield a model by 2005 wield a model by 2005 with a model by 2005 with a model by 2005 | Nider | | | |

Figure 3: ScrapBook's Manager offers a few useful tools to help you keep tabs on your stuff.



Figure 4: ScrapBook features powerful search capabilities.

bar contains the Comment area, which allows you to add comments to the page, as well as four tool buttons: Highlight, Pencil, Eraser, and DOM Eraser. There are also self-explanatory Undo and Save buttons. As the name suggests, the High*light* tool highlights selections in the text, and you can choose between four different highlight colors. The Pencil button contains several handy tools. The default tool is Sticky Annotation, which allows you to add sticky notes anywhere on the page. Using the Inline Comments tool, you can add notes to selected text fragments. For example, you can add a definition of a term or a translation of an unknown word. When you add an inline comment, the selected text is marked with a dotted line, and you can see the

comment by hovering the mouse over the marked text.

You can also attach a file to a selected text snippet using the *Attach File to Selection* tool. Last but not least, you can use the *Eraser* and *DOM Eraser* buttons to remove unwanted tags and text fragments from the page. The latest version of ScrapBook even offers the ability to perform all the described operations before capturing the page. To do this, press

the ScrapBook button in the status bar in the right lower corner of the browser window and select *Edit Before Capture*.

Extending ScrapBook

You can extend ScrapBook's already impressive functionality using add-ons. While there are only a handful of add-ons available on Scrap-Book's website,

one of them, ScrapBox.net, is worth mentioning. If you have a Box.net account [2], you can turn it into a Scrap-Book repository using the ScrapBox.net add-on. (You can get a 1GB Box.net account free of charge). This feature allows you to exchange ScrapBook data



Figure 5: Edit captured pages using ScrapBook's editing tools.

with other users and maintain online backup copies of the captured material. Once you've installed ScrapBox.net, you can access the tool via *Tools* | *Add-on Functions* | *Box.net Transporter*. Enter your Box.net account information, and the add-on creates a separate repository for your ScrapBook data. To upload captured pages or notes, simply drag them onto the window to the right. To import pages into ScrapBook, connect to your

| • | دىغا 44 inter دىغا 44 inter | | | | | |
|--|--------------------------------|-------------------------|----------------|---------------|---------|--|
| . | | | | | | |
| IS IT IT ISN'S. ADM | ta . | Cardine Second | hight/Equit | Bas rid Tran | operter | |
| tile in | Uphadrop | 1021 - index cos | | - | - | |
| Environmon Lipha M., | réame | | 04 | C+ | 10 | |
| - Course Lauff tals to | C fieldes ? | setable, gaids protica- | Mecalin II. Mo | NOM THE CO O | 4 | |
| Str Kaustorge Litten soft | ET 8-1942 | D-plained | | 4 42 DHT 2006 | - | |
| W Let of words having it | C HepSkine | | 840 | a 62 Det 2006 | - | |
| Check Constantion Constantial Constantia Constantial Constantia Constantia Constantia Constantia | (• cmat) | • | area da | Cys | - 100 | |

Figure 6: Use your Box.net account as an online SrapBook repository.

Box.net account, select the pages you want, and press *Download*.

Final word

To call ScrapBook an extension is a bit of an understatement: this little software gem turns Firefox into a full-blown archiving tool. If you just want to keep web pages for off-line reading, or even if you are doing online research, Scrap-Book can be an indispensable tool.

INFO

- Scrapbook: http://amb.vis.ne.jp/ mozilla/scrapbook/
- [2] Box.net: http://www.box.net

Dmitri Popov holds a degree in Russian language and computer linguistics. He has been working as a 6 UTHO technical translator and freelancer contributor for several years. He has published over 500 articles covering ٩ productivity software, mobile computing, web applications, and other ш н computer-related topics. His articles have appeared in Danish, British, US, and Russian magazines and websites.