
Calibre for ebook management

Jon Jerney*

Calibre is a free multi-platform, open-source application designed for the management of ebook collections. It includes searching capabilities, a generic ebook reader, an EPUB editor and internet links to find and retrieve metadata. Calibre will recognise external reading devices and allow their books to be transferred to and from a PC. It can be used as a server to provide access to an ebook collection on a local network or over the internet. Calibre will store and backup ebooks in DRM-protected formats, but is unable to display them in its reader.

THE ORIGINS OF CALIBRE

The open source application that became Calibre was created by Kovid Goyal in 2006 under the name “libprs500”. As it became popular a name change was suggested, and “Calibre” was chosen by Goyal’s wife Krittika. Since its creation, Calibre has increased rapidly in power and scope and, at the time of writing, has reached Release 1.42.0. Beginning as a relatively simple cataloguing application, it has grown into a must-have tool for anyone working with a sizeable ebook collection.

The Calibre installation file is only about 50Mb in size, but it carries an enormous amount of functionality. Users who want even more can download and install various plug-in extensions from the Calibre team and others. These include protocols to deal with reading and converting various ebook formats, connections to various websites as sources of further information about books, controls for the configuration of Calibre itself, and many others. Goyal and his team are prolific with updates, and it’s not unusual to come back to Calibre after a week or two away and find that you are already a couple of versions behind. Users are notified of new releases by the program itself, and the latest version can be downloaded from the Calibre website.¹

BASIC EBOOK MANAGEMENT

Managing an ebook collection is Calibre’s central function. The default screen layout shows a toolbar across the top with a search box underneath: typing text here allows the user to filter the display of books on title, author’s name, other details such as publisher, or the contents of the “Tags” field, which can include a publisher’s blurb and/or subject information. Fields available for display and sorting include author, title, publisher, date published, date modified (when the book was added to this collection), tags and format. There are about 20 predefined fields for each book, but the user can create and display other fields of their own, for example, “Pseudonym” or “ISBN”.

A navigation panel at the left allows for more precise searching. Users can expand an alphabetical list of authors, for instance, and click any one of them to see books by that author only. Other criteria can be added with a “Ctrl-click”. A recent addition is the “Virtual library”, where a specific filter can be set up and stored for re-use via a single click: for example, all mystery books or all non-fiction.

Adding an ebook is done through a toolbar button. Books can be added singly, collectively, or from a whole directory at once. Imported books are copied into Calibre’s own filing system, which consists of top-level folders for each author’s name in direct order, plus one subfolder for each book. Subfolder names are made up of the book’s title plus a number given in sequential order to each new book. The book file itself is renamed if necessary so that all files follow a standard format. This can be customised by the user, but the default is to show the title, followed by a hyphen, followed by the author’s name – again in direct order.

* Jon Jerney is an indexer, computer trainer and writer.

All websites viewed February 2014.

¹ Calibre – E-book management, <http://calibre-ebook.com>.

Each individual subfolder contains the book file itself, plus files for the cover image and any associated metadata. Books by multiple authors are stored in the first author's folder, though Calibre can access them via the second author's name. For example, *The Silver Fox* by Somerville and Ross would be found as "The Silver Fox – Edith O. Somerville.cpub" in the directory *Calibre Library/Edith O. Somerville/The Silver Fox (850)*, assuming it's the 850th book added to the collection. In practice, it's seldom necessary for Calibre users to worry about file storage details, as all access to the books can be done through Calibre itself.

For EPUB and other formats which contain metadata, Calibre will use this information in naming the book file and adding information. For other formats like TXT, PDF and RTF, Calibre will make a guess at the title and author from the file name. Where this is incorrect, it can be revised within Calibre itself, and the program will rename and move the file accordingly. The program also supports batch modifications (for example, changing the author of all "Edgar Box" books to "Gore Vidal"), whereby dozens or hundreds of records can be altered at once.

Once the book details are imported into Calibre, they can be searched, sorted and revised, if necessary. For books from non-commercial sources, which often lack metadata, Calibre can search the web to determine whether a cover image or other details are available. Its success rate on Project Gutenberg books, for instance, is fairly high – around 80%.

Calibre also has features relating to long-term maintenance of a book collection. It can record what you've read and when, find books in a series and put them in the right order, identify and remove duplicate books, and "tidy up" some formats such as EPUB so that inconsistencies in punctuation and spacing are minimised. It can count the words or pages in a book, and convert books between different formats to suit different reading devices. Users can "heart" books to indicate their favourites and delete those that no longer hold their interest.

Batch operations (for example, mass conversions), which take a long time, can be run as background jobs, allowing the user to continue working with Calibre in the meantime. Jobs can be displayed in their own dialog box and paused or cancelled if necessary.

With regard to exporting, ebook files can be selected and copied to a single directory, or to a directory structure organised by title, author, publisher or any other text field. There is no built-in backup facility (yet) but copying the top-level Calibre Library folder with all its subfolders effectively backs up the whole collection. Should the Calibre application itself be lost or corrupted, the collection details can be restored from a new installation with a single command.

Calibre can handle large numbers of books with no difficulty, given the space on disk. My own collection includes more than 10,000 titles, most of which are public domain books supplied on DVDs in the pre-broadband days. Calibre was able to load them in and convert them to EPUB format rapidly and without any fuss. Just as track management programs make it easy to work with music collections that would have been unfeasibly large in the past, so Calibre makes it possible to accumulate and use a vast personal ebook library.

READING WITH CALIBRE

Calibre supports a wide range of ebook file formats, including TXT, RTF, EPUB and PDF as well as those used by the Amazon Kindle, Kobo and Nook readers. Generally PDF files will call up a specific application like Adobe Reader for reading, but most others can be opened and read in the built-in Calibre Ebook Reader. This is a simple display window that supports searching, paging or scrolling, adjustable font sizes, and a full-screen view. Users can insert bookmarks and return to them, and a "Reference Mode" allows the user to identify paragraphs by number. Text can be copied or printed, and a CSS-based theme option allows visually impaired or other users to create and use their own display specifications. This should be adequate for most users, but as with PDF files, file formats can be associated with other reading applications if necessary.

Mobile devices

Calibre can communicate with mobile reading devices in several ways: directly via a cable; through a local wired or wireless network system; and even via the internet to remote devices.

Wired connections

Attaching a reading device to your PC with a USB cable while Calibre is running will prompt the program to find the external device, record its ebook contents, and display them on the Calibre screen along with the books in the main library. Thanks to the use of plug-ins, the list of supported devices is enormous, and includes everything from top-of-the-line tablets down to basic MP3 players with ebook capability. The user can copy books back and forth, delete books on the device or edit their details.

Calibre as server

The user with a network at home can go one better, and set Calibre up as a web server. This means that any device on the network with web-browsing capabilities can access Calibre via a local URL as if it were a website, and download books from there directly. This includes wireless devices like Android tablets and iPhones. There is direct support in the program, however, for specialised mobile apps like Calibre Sync, Calibre Library and Calibre Companion, which can connect directly to a Calibre server via Wi-Fi, and provide your mobile device with an intuitive interface which gives control over your ebook collection.

As if that wasn't enough, web-savvy users with a fixed IP address can set Calibre up to serve books from a PC across the web, providing password access to your library from anywhere that has an internet connection. The computer has to be left on, the Calibre program needs to be running, and a specified port must be opened in your firewall. If that's all feasible for you, then Calibre can make your personal library accessible from just about anywhere.

Calibre as newsreader

In rather an odd sideline, Calibre can be set up to download up-to-date information from news feed sites. Several hundred of these are already listed in the Calibre menus, but others can be added by the user. The type and content of the feed will vary depending on the source, but typically headlines and "teasers" for full articles that can then be accessed via a web browser. The most recently updated version of the feed appears as a "book" in the Calibre list, and can also be accessed via servers and the web as described above.

EPUB EDITOR

Calibre generally is format-agnostic, but the popularity of the EPUB format has prompted the Calibre team and other contributors to give this format some special attention. In addition to the standard conversion options, Calibre includes an EPUB modifier dialog which controls some of the most frequently altered settings in EPUB files. This includes an option to "smarten" bad punctuation – a sore point with many readers of carelessly-created books.

More recently Calibre has added a fully-fledged EPUB editor along the same lines as Sigil or Oxygen. This uses a split screen to show file information in a panel at the left, the source code in a central window, and a formatted display at the right. EPUBs can be opened for editing by right-clicking in Calibre or via the editor's "File" menu.

The EPUB is a zipped format, and the editor unpacks this to display the component files. These can be edited, renamed, reordered or deleted. Component files usually include CSS stylesheets which control the formatting; these are plain text and can be edited in the same way as source code. Multiple component files can be opened at once in a tabbed display. Users can create a "checkpoint" at any time and return to it if mistakes have been made in the interim.

The "Tools" menu includes options to fix erroneous source code, identify duplicate entries or misnamed files, improve punctuation, and "beautify" an EPUB by applying Calibre's own CSS file. The user can also construct or modify a table of contents. A search and replace facility allows the user

to find and change errors throughout the book. It includes regular expression searching for more complex expressions. Saving the book and closing the editor automatically updates the copy in the Calibre Library.

In addition to working with existing books, the EPUB editor can also be used to create new ebooks from scratch, although this will require some knowledge of HTML and a good deal of hand coding. (A better approach would be to prepare the text chapters in an HTML editing program, and copy the source code into the Calibre Editor for compilation into an EPUB.) The editor also works with KZW (Kindle) format files, but not if these are DRM-protected.

PROGRAM RELIABILITY AND SUPPORT

Installation and setup is relatively easy, and Calibre works in a virtually identical way on Windows, Macintosh and Linux systems, making it easy to transfer expertise across to different platforms. However, Calibre is not bulletproof. The program hangs occasionally and sometimes stalls in the middle of long and complex tasks, though a simple restart usually solves the problem. Data damage is very rare, but like any other files, your ebook collection should be backed up regularly and often. There are sometimes minor problems with new releases, though these are quickly fixed, so it may be best to hold off updating for a week or two. Changes for each release are documented on the Calibre website.

Although Calibre is aimed at private users, it could easily be used in a library or archive as a means of providing users with simple, intuitive access to large collections of electronic documents. As a free and open-source program, there are no licensing or financial issues involved, although the Calibre team accept and appreciate donations.

There is plenty of Calibre support on the internet, including a User Manual on the website and a MobileRead discussion forum,² but like most open-source software, the program is provided on an “all-care-but-no-responsibility” basis.

Anyone about to buy or upgrade a commercial document management system for ebooks should certainly consider this free and highly effective substitute.

² Calibre – MobileRead Forums, <http://www.mobilerread.com/forums/forumdisplay.php?f=166>.