

# 基于 Calibre 应用平台的移动图书馆管理与网络共享搭建

杨孙超 (石狮市图书馆 福建 362700)

**摘要** 随着 5G 通信网络、各类移动阅读设备的蓬勃发展, 移动阅读的读者数量不断增长, 如何高效地管理电子图书, 打造移动图书馆, 作为读者和图书馆管理者都需要认真考虑。Calibre 应用平台基于元数据和 OPDS 标准整合技术对各类电子图书进行有效的全面管理, 可以作为未来数字化图书馆的一个重要组成部分。

**关键词** 移动阅读 移动图书馆 电子书 元数据

## 0 引言

近年来, 数字化阅读呈现逐年递增的发展趋势, 改变着人们的文化生活方式。第十七次全国国民阅读调查报告指出, 我国成年公民 2019 年通过各种媒介 (含书报刊和数字出版物) 的综合阅读率为 81.1%, 比增 0.3%; 线上阅读、电子阅读器阅读、Pad 阅读、手机阅读等数字化阅读方式接触率为 79.3%, 比增 3.1%。随着科技的进步, 数字化阅读的快速的发展, 成年公民对电子媒介的依赖度上升, 综合阅读率和数字化阅读方式接触率同步提升, 使得整体阅读人群持续增加, 从而产生了数字化阅读持续增长、纸质阅读率增长放缓的新趋势。<sup>[1]</sup> 亚马逊中国 2020 全民阅读报告显示, 超过 29% 的读者将电子书作为阅读的主要介质, 比增 6%。<sup>[2]</sup> 由于电子书能够提供个性化和多媒体阅读功能, 在成本、收藏实践性、可访问性、存储空间等综合因素考量下, 越来越多的图书馆选择购买电子书。国内各类图书馆每年购置的电子书数量也在不断攀升中。据不完全统计, 各大电子书提供商的资源量如下: 中文在线电子图书拥有 10 万册电子书和 3 万集听书, 方正阿帕比 (Apabi) 电子书库现有 220 余万册可供阅读的电子图书, 其中可全文下载的有 68.9 万册。目前世界最大的中文在线数字图书馆“超星数字图书馆 (汇雅电子书)”有电子图书 100 万种以上。曾号称是世界最大的电子书图书馆 ZLibrary, 现有电子图书 508 万本。谷歌 (Google) 数字图书馆更是计划建设世界上最大的数字图书馆。

在电子书和纸质图书并行的多元化阅读背景下, 个人和移动图书馆要实现系统地管理海量电子书, 要求管理平台的设计必须统筹考虑移动阅读的硬件、软件、文件格式、扩展应用等方面。笔者经过实际应用, 提出基于 Calibre 应用平台的管理方案。

## 1 移动图书馆客户端的软硬件支持

移动图书馆客户端的阅读硬件发展, 经历了从早期的基于微软 Windows Mobile 操作系统的掌上电脑, 到以安卓和 IOS 平台的智能手机、平板电脑为主流代表的阅读设备, 以及以亚马逊的 Kindle 和国内的掌阅 iReader、京东 JDRead、文石科技 BOOX 为代表的沉浸式墨水屏电子书阅读器这样一个过程。

从软件上看, 除了各电子书阅读器自带的应用, 限于版权和格式支持有限, 第三方阅读 APP 也是不错的选择。目前国内主流的移动阅读软件包括微信读书、QQ 阅读、掌阅、百度阅读、多看阅读、网易蜗牛阅读等。国外支持格式较多的阅读软件包括: FBReader、Mantano Reader、静读天下、EbookDroid、Icecream Ebook Reader 等。衡量各个阅读软件的质量, 一般通过自带书城、轻量化界面、笔记功能、自定义排版等功能来评判。

## 2 常用移动图书馆电子书格式

由于移动图书馆电子书的格式林林总总有数十种, 本文就移动设备使用率最高的格式 TXT、EPUB、MOBI、AZW、PDF 共 5 种进行探讨, 一些随着移动设备更替而较少使用的格式如 PRC、

AZK、UMD、JAR 和一些更新的格式如 KFX 暂不展开说明。这 5 种电子书格式 (TXT 格式除外) 本质上都是从 HTML 文档转换而来, 大多数 HTML 标签和 CSS 样式表特性都有支持, 它们之间的主要区别是对排版和新功能的支持上。<sup>[3]</sup>

### 2.1 TXT 格式

TXT 格式是基本的文本。体积小, 是各种电子书的原始载质, 易于传输。缺点是只支持纯文字, 不支持图像。

### 2.2 EPUB 格式

EPUB 格式基于国际数字出版论坛 (IDPF) 提出的 EPUB 电子图书标准, 得益于 EPUB 标准的自由开放和完善的数位版权管理 (DRM) 功能, EPUB 已经成为非常普遍的电子书格式, 国内外大多数电子书制造商和数字图书馆都使用特定方式加密, 而 2011 年亚马逊公司推出 MOBI KF8 格式后, 大部分的 MOBI 和 AZW3 格式电子书同样由 EPUB 格式转码而来。<sup>[3]</sup> EPUB 格式对于复杂的图表、公式、排版等元素的兼容性比 MOBI 格式好很多, 在矢量图形、各类公式、脚本的支持方面也优于 MOBI 格式, 最大优势体现在图片嵌入字体、图文混排等, 未来可预测的优势是 EPUB 格式对于音像、视频等多媒体内容互动的支持上。

### 2.3 MOBI 格式

MOBI 格式分两种标准: KF7 和 KF8, 也称

MOBI 7 和 MOBI 8, 主要区别是, KF8 相较 KF7 增加了丰富的排版格式, 如更换字体、加粗、行距等。

### 2.4 AZW 格式

AZW 格式目前主要以 AZW3 格式为主, 是亚马逊公司在 2011 年推出 Kindle Fire 平板设备时建立的格式。主要解决 MOBI 格式对于复杂排版支持的问题, 尤其支持了很多 HTML5 (目前尚不支持 HTML5 的视频和音频标签) 和 CSS3 的语法, 这就大大改善了原来 MOBI 或 AZW 内容排版上的一些缺陷, 单纯从阅读角度来讲, 不输 EPUB 格式。目前亚马逊提供 Kindle 阅读器的电子书大多基于 AZW3 格式, 而其另一专有 MOBI 格式的电子书则越来越少。

### 2.5 PDF 格式

PDF 格式是由 Adobe Systems 公司为交换丰富排版的文件于 1993 年开发的便携式文档格式, 无关平台、操作系统和硬件, 具有通用性的特点。<sup>[4]</sup> PDF 文件使用了工业标准的压缩算法, 通常比 PostScript (适用于列印图像和文字的语言) 文件小, 易于传输与储存, 是公开的 ISO 国际标准。它支持各种复杂的排版, 如图表、插图、嵌入式字体等, 具有纸书一般的质感和阅读效果, 能逼真地展现原书风貌。所以也是图书印刷前通用的电子格式。

表 1: 常用移动图书馆电子书格式比较

格式	开放性	通用性	排版	图表及多媒体扩展	DRM 保护	文件大小排名*
TXT *	公开	所有移动设备	差	不支持	不支持	1
EPUB	公开	所有移动设备	优	支持	Apple Fairplay, KOBO, Adobe Adept, B&N Nook 等 DRM	2
AZW *	亚马逊专有	亚马逊的 kindle	优	支持	Amazon DRM	3
MOBI *	亚马逊专有	亚马逊的 kindle, 部分第三方阅读软件	中	较差	Amazon DRM	4
PDF	公开	所有移动设备	优	支持	Adobe acrobat DRM	5

\*注: 选用相同内容、图文排版的电子书, 各类格式采用最新版本 (如 AZW 文件选取最新的 AZW3 格式, MOBI 文件选取 MOBI8 格式) 为样本进行比较, 容量由小到大排列, TXT 文件由于不支持图片, 所以容量最小。

### 3 通用移动图书馆的管理

Calibre 应用平台作为一款功能强大的移动图书馆电子书管理软件,支持亚马逊 (Amazon)、苹果 (Apple)、Ectaco、无尽创意 (Endless Ideas)、Bookeen、Hanlin Song、谷歌 (Google) / HTC 等移动阅读设备。<sup>[5]</sup>支持多达 18 种电子书格式,包括 AZW、MOBI、EPUB、PDF 等主流电子书格式,提供电子书管理、联网获取元数据及编目、格式转换、阅读设备图书同步、RSS 新闻下载将其转换成电子书方便阅读等功能。它使用 Python 程序语言开发,可实现网络共享,是一个支持 Windows、MacOS、Linux 各系统跨平台的自由软件。

#### 3.1 图书导入与格式转换

Calibre 应用平台可以建构多个书库,比如中文和外文、儿童和成人、专题等书库,更好地管理专属和专业书库,各书库切换非常方便。

图书导入非常简单,直接把电子书文件拖到 Calibre 窗口进行添加。也可点击工具栏“添加书籍”,批量添加一个目录下的所有文件(包括子目录下的文件)。电子书导入后,会自动读取该书的元数据,并根据书名、作者、丛书、格式、出版社、评分、标签、书号、语言等属性自动建立索引并排序。

Calibre 通用性强大,支持各类电子书文件格式,常见的 TXT、PDF、EPUB 自不用说,亚马逊公司独家的 MOBI、AZW 和最新的 KFX 格式也支持完美导入,并可以转换成不同的格式,方便各类阅读设备读取。支持的导入格式有:TXT、CBZ、CBR、CBC、CHM、EPUB、FB2、HTML、LIT、LRF、MOBI、ODT、PDF、PRC、PDB、PML、RB、RTF、TCR 等,支持的转换格式有:TXT、EPUB、FB2、OEB、LIT、LRF、MOBI、PDB、PDF 等。解决了前述表 1 中的开放性问题。通过内建的阅读器及外挂程序,Calibre 可以即时打开各类电子书甚至进行二次编辑。

Calibre 支持对实体书的管理,点击工具栏“添加书籍→按 ISBN 添加”,输入实体书的 ISBN 号,在添加的时候可以添加标签说明是实体书,

Calibre 会自动查找下载元数据和封面。Calibre 还支持批量导入,通过 Excel 单元格的链接,将作者和标题的内容链接到相应的命令格式,并为每本书生成一个 Calibre 以添加命令行。点击工具栏“添加书籍→添加文件到选中的书籍记录中”,可以手动添加相关文件,这项功能尤其适合管理该书的有声书音频或视频文件。Calibre 目前只支持一种文件格式的文件关联,多个文件需要合并或打包(压缩)为一个文件,或者手动拷贝到相关图书目录下。

#### 3.2 元数据 (Metadata) 管理

出版业方面,全国新闻出版标准化技术委员会于 2010 年组织成立了电子书内容标准的工作组,从电子书内容格式、版权保护、元数据、服务平台、质量检测等方面开展电子书标准研究。在文化行业中,以国家图书馆为代表的图书馆界研制了电子书对象数据和元数据规范。其中 16 个标准规范涉及各种类型电子书的对象数据和元数据,分别在 2012 年和 2014 年被批准为文化部行业标准或行业标准化指导性技术文件。<sup>[6]</sup>

Calibre 支持元数据的下载和编辑功能,完整的元数据格式包括书名、作者、丛书、星级、标签、IDS、出版日期、出版商、语言、封面、摘要等。导入电子书时,文件默认信息可能只包括书名和作者,如果需要更加详尽的元数据内容,可以使用“下载元数据和封面”工具自动下载。目前 Calibre 支持从亚马逊、谷歌、豆瓣、Ozon、Open Library 等平台抓取图书的相关信息,下载电子书的所有元数据。同时可以手动编辑相关元数据,通过选取一定类型的书(如同一作者、同一系列)进行批量修改、搜索替换等。

而高级搜索和排序功能,可以便捷地找到所需电子书。还可以通过“首选项”中的“界面→增加栏目”功能,添加自定义字段。添加后,相应的列将显示在图书列表中,而“自定义元数据”页将显示在“编辑元数据”对话框中。比如下图中笔者对图书设置“是否”标记已经读过、设置“长文本”作为读书笔记、用“标签”功能制作自己喜欢的书单等。



图 1: Calibre 添加自定义元数据

### 3.3 扩展功能

RSS (Really Simple Syndication) 是一种简易描述和同步网站内容实现信息聚合的格式规范，它是目前使用最广泛的 XML 应用程序。<sup>[7]</sup> Calibre 应用平台除了可以导入、编辑、管理电子书之外，还可以从 RSS 源获取新闻内容并抓取，将其转换为支持格式的电子书，推送到阅读器设备上。

各类网站的 RSS 订阅地址，一般会提供 RSS 图标或者 OPML 文件下载，点击 Calibre 的“抓取新闻→添加自定义新闻源”功能，将 RSS 订阅或导入 OPML 文件添加到“订阅列表中的新闻”。在一个订阅列表中可以对多个 RSS 源同时定制抓取，还可通过自定义内置订阅列表，从 Calibre 系统自带的 RSS 库中添加。Calibre 最新的 4.15 版中自带包括《人民日报》在内近 1600 个 RSS 源。设置好订阅源后就可以即时或制定计划抓取新闻，并生成一本 EPUB 格式的电子书。

## 4 移动图书馆网络共享搭建

### 4.1 OPDS 标准

OPDS (Open Publication Distribution System) 是使用 Atom 格式为在线电子书目录创建的开放式出版发布的开放标准。<sup>[8]</sup> OPDS 标准的一大优势是支持各种桌面系统和移动终端，支持各种应用软件 APP，可以在不同的终端和电子书阅读软件上显示和共享电子书的书目信息。读者可以用一款运行于 IOS、安卓等移动平台的阅读器软件如超星阅读器、FBReader、EPUBReader、Aldiko、Infinite Book Reader、PocketBook Reader 等订阅多个来源电子书书目信息，像 RSS 订阅一样做到实时更新。对于一些不涉及版权问题的免费电子书，可以直接下载并阅读。建立 OPDS 格式的目录之后，作者、出版社、图书馆可将其图书做成 EPUB 格式的电子书，链接到书目的元数据。



图 2: Calibre 定期 RSS 订阅下载库

未来,当所有的作者、出版社、图书馆、书店都采用 OPDS 开放式出版物发行系统,读者可以直接从阅读软件搜索每本书目录和下载,直接下载、借阅或购买电子图书,无论他们使用什么样的电子终端设备——电子书阅读器、手机、电脑。而不是从一个软件转到另一个软件,从一个网站跳到另一个网站。对于图书馆来说,EPUB 格式的电子图书馆藏一旦建立,读者可以通过手中电子阅读终端设备搜索图书馆 OPDS 格式的电子图书目录,并通过开放式的电子阅读软件从图书馆的网站下载或借阅电子书。图书馆可以根据需要设置借阅规则,规定可借阅的复本数,实现电子图书的借阅管理。

#### 4.2 根据 OPDS 标准搭建网络图书馆

Calibre 应用平台可以根据 OPDS 标准建立在线 web 书库,聚合本地的电子图书。由于 OPDS 目录可以合并,方便组合更大的 OPDS 目录,从而实现海量电子图书的聚合。用自带的内容服务

器,点击工具栏的“连接/共享”输入一个 IP 地址和端口,假定 IP 地址是 192.168.0.200,端口是 8888,在浏览器中输入 <http://192.168.0.200:8888>,就可通过浏览器访问书库中所有的资源。再利用 Calibre - web、COPS、BicBucStriim、Calibre2opds 等服务端,就可以打造完美的移动图书馆功能。

Calibre - web 是一款基于 Python 语言设计的服务端,功能全面,界面美观。

COPS 的优势是服务相对比较稳定,相对 Calibre - web 不需要使用 Python 运行环境,虽然功能丰富,但界面相对简陋。

BicBucStriim 界面非常简洁,但对元数据支持不是很友好,如果作为基础的存储服务可以考虑。

作为 Calibre 的插件工具,Calibre2opds 可以读取 SQLite 数据库并转换为符合 OPDS 规范的静态页面,使用也比较方便,但界面也不如 Calibre - web 方便。<sup>[9]</sup>

鉴于 Calibre - web 的通用性和易用性, 选择其服务端开发设置的居多, 只要在 Python 官网下载安装包, 手动加入环境变量, 再下载 Calibre - web 安装, 进行相关配置, 绑定域名实现外网访问。最后通过用户管理, 实现分类读者的访问权限, 就可以开通让读者通过域名访问移动图书馆。

Calibre 应用平台基于元数据和 OPDS 标准整合技术对各类电子书进行全面管理, 集成 RSS 订阅, 并可以通过网络搭建移动图书馆, 在移动终端上实现资源的一站式搜索、导航、书单定制和下载阅读, 其核心具有共享性, 可以作为未来数字化图书馆的一个重要组成部分, 借助 5G 移动无线通讯网络技术的迅猛发展, 其发展前景不言而喻。

#### 参考文献

- [1] 中国全民阅读网. 第十七次全国国民阅读调查成果发布[EB/OL]. (2020 - 04 - 20) [2020 - 05 - 30]. <http://www.nationalreading.gov.cn/ReadBook/contents/6271/414891.shtml>.
- [2] 新浪财经. 亚马逊中国 2020 全民阅读报告: 29% 读者将电子书作为主要阅读介质[EB/OL]. (2020 - 04 - 20) [2020 - 05 - 30]. <http://finance.sina.com.cn/stock/relnews/us/2020-04-20/doc-iirczymi7424793.shtml>.
- [3] 知乎. 一文看懂 epub、mobi、azw(3)、kfx 格式的区别[EB/OL]. (2020 - 04 - 13) [2020 - 05 - 30]. <https://zhuanlan.zhihu.com/p/129807172>.
- [4] 百度百科. pdf 格式[EB/OL]. (2020 - 04 - 02). [2020 - 05 - 30]. <https://baike.baidu.com/item/pdf%E6%A0%BC%E5%BC%8F>.
- [5] OSCHINA 社区. Calibre 4.16.0 发布, 功能强大的电子书管理软件[EB/OL]. (2020 - 05 - 16) [2020 - 05 - 30]. <https://www.oschina.net/news/115686/calibre-4-16-0-released>.
- [6] 谢强, 刘术华. 我国数字阅读最新发展及对图书馆电子书服务的启示[J]. 图书馆, 2017(1): 24 - 27.
- [7] 易晓阳. RSS: 含义、本征与应用[J]. 图书馆学研究, 2006(8): 69 - 72.
- [8] 向明, 刘刚. 用 Calibre 架设符合 OPDS 标准的数字图书馆[J]. 发明与创新: 教育信息化, 2014(4): 53 - 56.
- [9] 刘军. 采用 EPUB 和 OPDS 标准的电子书聚合及在数字图书馆中的应用[J]. 河南图书馆学刊, 2013(7): 89 - 92.

杨孙超 石狮市图书馆馆员。

(收稿日期: 2020 - 06 - 05。龚永年编发。)