

E-Books selbstgebaut

Jens Pönisch

poenisch@isym.tu-chemnitz.de

<http://www.tu-chemnitz.de/~poenisch>

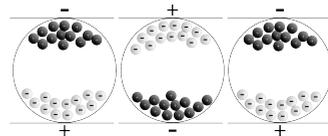
Der Beitrag beschreibt sowohl den Aufbau und die Arbeitsweise handelsüblicher E-Book-Lesegeräte als auch das Erstellen eigener E-Books im EPUB-Format. Neben Software zum Lesen und Verwalten von E-Books am PC werden auch Tools zum Konvertieren und Prüfen vorgestellt.

1 E-Book-Reader

Die Idee, Bücher in elektronischer Form anzubieten, ist nicht neu. So startete das Projekt *Gutenberg* bereits in den 1970-er Jahren mit der Digitalisierung von Büchern, deren Urheberrecht abgelaufen war. Seit dieser Zeit existiert auch der Wunsch, handliche elektronische Lesegeräte bereitzustellen, z.B. ist das die zentrale Idee in Douglas Adams' *The Hitchhiker's Guide to the Galaxy* [1]. Erste Versuche mit LCD-Displays und Disketten- oder Mini-CD-Laufwerken in den 1990er Jahren hatten nur mäßigen Erfolg. Erst mit der Erfindung des elektronischen Papiers gelang Sony 2004 der Durchbruch zu wirklich benutzbaren Geräten. Eine größere Verbreitung der E-Book-Reader wurde 2007 durch *Amazons* reichhaltiges Angebot von elektronischen Büchern gemeinsam mit dem E-Book-Reader *Kindle* erreicht. Auf dem europäischen Markt verbreiten sich diese und eine Vielzahl alternativer Geräte seit 2009 [2].

E-Book-Reader möchten die Vorteile von Taschenbüchern mit den Vorteilen von Computern vereinen. Sie sind klein, leicht, haben eine hohe Laufzeit und hochauflösende Displays mit einem guten Kontrast. Die Bereitstellung von Büchern erfolgt je nach Gerät entweder über einen stationären Computer mittels Micro-SD-Karte, USB bzw. WLAN oder per Mobilfunk.

Für die Displays kommt derzeit meist auf *Elektrophorese* beruhendes elektronisches Papier zum Einsatz. Zwischen zwei leitend beschichteten Glasplatten schwimmen in dunklem Öl elektrisch geladene weiße Kugeln. Durch entsprechende Polung der Platten werden die Kugeln nach oben bzw. unten angezogen und erzeugen so die Anzeige.



E-Paper – Funktionsweise
(Quelle: Wikipedia)

Der große Vorteil dieser Technik ist der extrem geringe Energieverbrauch, lediglich beim Umladen der Platten fließt Strom. Nachteile sind die recht trägen Schaltzeiten von fast einer Sekunde und das komplette Schwärzen der Anzeige während des Schaltens. Ebenso sind derzeit noch keine Farbdisplays erhältlich [3].

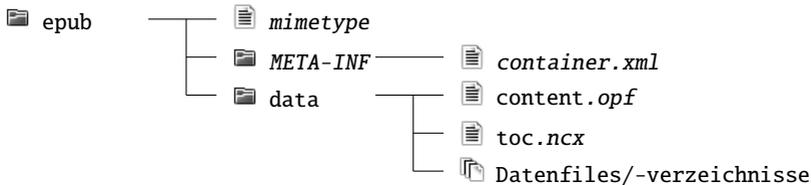
Obwohl größere Displays erhältlich sind, werden derzeit im wesentlichen 5 und 6 Zoll große Displays mit 600×800 Pixeln in E-Book-Readern verbaut, womit sich eine Auflösung von 160 bis 200 dpi ergibt, das entspricht der üblichen Druckqualität von Taschenbüchern. Je nach Anbieter wird entweder das Amazon-Format *Mobipocket* oder das Format *EPUB* unterstützt, wobei letzteres in Europa weiter verbreitet ist. Alternativ können meist noch andere Formate wie HTML, Text und PDF angezeigt werden. Eine Übersicht derzeit erhältlicher Geräte ist in [4] zu finden.

2 EPUB-Format

Außer dem *Kindle-Reader* von *Amazon* unterstützen derzeit praktisch alle erhältlichen Geräte die EPUB-Spezifikation mit den Teilspezifikationen *Open Publication Structure* (OPS), *Open Packaging Format* (OPF) und *Open Container Format* (OCF) [5].

Das Erstellen neuer oder die Anpassung vorhandener E-Books in diesem Format ist nicht schwierig. Benötigt werden lediglich ein UTF-8-fähiger Editor, der Packer *Zip* und zur Kontrolle des Ergebnisses der Validator *epubcheck*¹ sowie ein E-Book-Viewer. Für die Bearbeitung des eigentlichen Inhalts sind HTML- und XML-Transformationstools hilfreich.

Ein E-Book im EPUB-Format ist ein gepacktes Archiv mit folgender Verzeichnisstruktur (kursiv gedruckte Namen sind fest definiert):



Nach dem Erstellen dieser Verzeichnisstruktur müssen alle Inhalte in das XHTML-1.1-Format konvertiert werden. Für die Codierung ist UTF-8 vorgegeben, CSS wird in der Version 1.0 mit einigen Erweiterungen unterstützt. Alle Links müssen relativ bezüglich des Datenverzeichnisses sein. Bilder können in den üblichen Formaten und im SVG-Format bereitgestellt werden.

Die Datei `mimetype` ist immer gleich und enthält lediglich die Zeile:

```
application/epub+zip
```

Im Verzeichnis `META-INF` befindet sich mindestens die Datei `container.xml`, die den Pfad der Beschreibungsdatei festlegt.

¹<http://code.google.com/p/epubcheck/>

```

1 <container version="1.0"
2     xmlns="urn:oasis:names:tc:opendocument:xmlns:container">
3   <rootfiles>
4     <rootfile media-type="application/oebps-package+xml"
5               full-path="data/content.opf"/>
6   </rootfiles>
7 </container>

```

Alle weiteren Verzeichnisse werden relativ zum Verzeichnis dieser Datei angegeben.

Bei kopiergeschützten Inhalten enthält dieses Verzeichnis weiterhin Informationen zur Verschlüsselung.

Die wichtigste Steuerdatei ist die Paketbeschreibungsdatei. Neben Metadaten wie Titel und Autor listet sie sämtliche, im Paket enthaltene Ressourcen auf und definiert die Anzeigereihenfolge des E-Books. Die Datei hat folgenden Aufbau:

```

1 <package xmlns="http://www.idpf.org/2007/opf"
2         xmlns:dc="http://purl.org/dc/elements/1.1/"
3         version="2.0"
4         unique-identifier="bookId">
5   <metadata>
6     <dc:identifier id="bookId">Book-URI</dc:identifier>
7     <dc:title>Buchtitel</dc:title>
8     <dc:language>sprachcode</dc:language>
9     ...
10  </metadata>
11
12  <manifest>
13    <!-- Alle Ressourcen -->
14    <item href="datei" id="dateiId" media-type="mime-type"/>
15    ...
16  </manifest>
17
18  <spine>
19    <!-- Lesereihenfolge -->
20    <itemref idref="dateiId"/>
21    ...
22  </spine>
23  ...
24 </package>

```

Zusätzlich kann hier auf eine Inhaltsverzeichnis-Datei verwiesen werden, die eine Navigation innerhalb des E-Books ermöglicht. Zum Aufbau dieser beiden Dateien sei auf [6] und [7] verwiesen.

Beim Zusammenpacken ist zu beachten, dass die Datei `mimetype` unkomprimiert am Anfang des Zip-Archivs steht. Das wird mit folgender Befehlsfolge erreicht:

```
zip -X0 buch.epub mimetype
zip -Xr9 buch.epub META-INF datenverzeichnis
```

Vor dem Anschauen oder Übertragen auf den E-Book-Reader sollte mit dem Java-Programm `epubcheck` die Korrektheit des E-Books überprüft werden.

3 Anzeigeprogramme und Werkzeuge

Zur Verwaltung und Anzeige von E-Books auf dem PC existieren verschiedene Anzeigeprogramme mit unterschiedlichem Leistungsumfang.

Käuflich erworbene Bücher enthalten häufig einen personenbezogenen Kopierschutz, wobei in Europa meist das *Digital Rights Management* (DRM) von *Adobe Inc.* zum Einsatz kommt. Zum Kauf und zur Verwaltung solcher Bücher sowie zur Aktivierung des E-Book-Readers wird zwingend das Programm *Adobe Digital Editions*² benötigt, das allerdings nur für Windows und MacOS X erhältlich ist. Neben diesen Aufgaben erlaubt es auch die Anzeige und den Ausdruck von E-Books am Computer, eine komfortable Verwaltung sowie die Übertragung von E-Books an den E-Book-Reader.

Eine qualitativ hochwertige Anzeige ist auch mit der Firefox-Erweiterung *EPUBReader*³ möglich. Sie ermöglicht das Verwalten einer Bibliothek, jedoch keine Kommunikation mit E-Book-Readern.

Ein vielseitiges Werkzeug für Anzeige, Verwaltung und Bau von E-Books ist *Calibre*⁴. Es unterstützt die Übertragung auf E-Book-Reader, die Umwandlung verschiedener Formate, Korrekturen der Metadaten sowie mustergestütztes Bauen von E-Books aus einer Sammlung von HTML-Dateien. Ebenso können Meldungen verschiedener Newsticker automatisch zu einem E-Book zusammengefasst werden.

Weiterhin soll noch das Programm *FBReader*⁵ genannt werden, das gern auf PDAs, Mobiltelefonen und ähnlichen Geräten eingesetzt wird. Die Anzeigequalität ist jedoch nicht mit den beiden vorher genannten Programmen vergleichbar, DRM wird nicht unterstützt.

²<http://www.adobe.com/de/products/digitaleditions/>

³<http://www.epubread.com/en/>

⁴<http://calibre-ebook.com/>

⁵<http://www.fbreader.org/>

4 Quellen für freie E-Books

DRM-geschützte Bücher haben gegenüber gedruckten Exemplaren eine Vielzahl von Nachteilen: Sie können nicht verliehen oder weiterverkauft werden, bei Defekt des Rechners sind der kryptographische Schlüssel und damit auch alle digitalen Bücher verloren. Zu diesen Nachteilen kommt ein meist nur sehr bescheidener Preisnachlass von 1 bis 2 € (s. auch [8]). Damit stellt sich die Frage nach DRM- bzw. kostenfreien Alternativen.

Da das Urheberrecht 70 Jahre nach dem Tod des Autors erlischt, werden Bücher nach dieser Zeit gemeinfrei. Sie können also problemlos digitalisiert und als E-Books bereitgestellt werden. Das bereits erwähnte *Project Gutenberg*⁶ stellt inzwischen über 30 000 freie Bücher digital in verschiedenen Formaten bereit.

Ein ähnliches Projekt startete 1994 *Spiegel-Online* mit dem Projekt *Gutenberg-DE*⁷. Inzwischen sind ca. 2500 deutschsprachige Bücher im HTML-Format verfügbar. Ebenso bietet die Firma *Contumax* im Projekt *Zeno.org*⁸ eine Vielzahl von gemeinfreien digitalen Büchern an. Eine Konvertierung in das EPUB-Format ist z.B. mit *Calibre* oder einem kleinen selbstgeschriebenen Tool⁹ möglich. Wer den Konvertierungsaufwand scheut, kann E-Book-Bibliotheken gemeinfreier Bücher preiswert auf Datenträgern erwerben.

5 Hinweise zum Layout

Für ein angenehmes Leseerlebnis sollten Stylesheets für E-Books versuchen, das Layout von gedruckten Büchern so weit wie möglich nachzuempfinden. Selbst gekaufte Bücher vertragen oft eine Nachbesserung, die allerdings nur bei fehlendem DRM-Schutz möglich ist.

Der Seitenspiegel sollte so gewählt werden, dass um den Text etwas Rand bleibt. Eine Seitennummer entfällt, da die Seitenaufteilung von der meist frei wählbaren Schriftgröße abhängt. Aus dem gleichen Grund können keine Fußnoten verwendet werden, diese können aber sinnvoll durch Endnoten ersetzt und im Text verlinkt werden. Aufgrund der beschränkten Anzeigebreite sind Randnotizen und mehrspaltiger Text ungeeignet. Auch beim Tabellensatz ist auf die eingeschränkte Breite zu achten.

Absätze sollten außer nach einer Überschrift einen Einzug haben und im Blocksatz gesetzt werden.

Eingebettete Bilder müssen einen ausreichenden Kontrast aufweisen, es ist sinnvoll, die Größe auf die Breite von 600 Pixeln bzw. 800 Pixeln im Querformat zu optimieren.

⁶<http://www.gutenberg.org/>

⁷<http://gutenberg.spiegel.de/>

⁸<http://www.zeno.org/>

⁹<http://xml2lms.in-chemnitz.de/epub/>

Enthält der Text Zeichen, die nicht im Latin-1-Zeichensatz vorhanden sind, kann es je nach verwendetem Zeichensatz des E-Book-Readers zu Anzeigefehlern kommen. Hier bietet es sich an, passende Fonts in das E-Book einzubetten und diese zu verwenden.

6 Zusammenfassung

E-Book-Reader sind aufgrund des großen Speicherplatzes und der langen Betriebszeit ideale Reisebegleiter und ein guter Ersatz für Taschenbücher. Durch die spezielle Technologie neigen sie weniger als Computerdisplays zu Reflexionen, haben einen größeren Ablesewinkel und können auch bei Sonnenschein problemlos gelesen werden. Die kleinen Displays eignen sich nur bedingt für das Lesen von Fachbüchern. Ebenso lassen die Typographie und die Bildqualität zu wünschen übrig.

Ein weiteres Hindernis für die Verbreitung von E-Books und entsprechenden Lesegeräten sind der Kopierschutz und der meist unverhältnismäßig hohe Preis der elektronischen Publikationen gegenüber der gedruckten Ausgabe. Glücklicherweise existiert jedoch inzwischen eine große Zahl gemeinfreier Bücher, die entweder bereits im passenden Datenformat vorliegen oder mit wenig Aufwand konvertiert werden können.

Literatur

- [1] Adams: *The Hitchhiker's Guide to the Galaxy*. Radio series. BBC 1978.
- [2] Pawlak: *Die Wissenschaft bei Douglas Adams*. Wiley 2010.
- [3] *Elektronisches Papier*. Wikipedia
http://de.wikipedia.org/wiki/Elektronisches_Papier
- [4] *E-Book Reader Matrix*.
http://wiki.mobileread.com/wiki/E-book_Reader_Matrix
- [5] *Epub Specifications*. International Digital Publishing Forum (IDPF)
<http://www.openebook.org/specs.htm>
- [6] *Epub Format Construction Guide*.
http://www.hxa.name/articles/content/epub-guide_hxa7241_2007.html
- [7] Dietrich: *Der kleine E-Autor*. c't 25/2009.
- [8] *Studie zu E-Books und -Readern*. BCG 2010.
<http://www.bcg.com/media/PressReleaseDetails.aspx?id=tcn:12-46750>