

## OTS 2 EPUB – Anatomie eines elektronischen Buchformates

Fachvortrag

Dr. Thomas Meinike, Hochschule Merseburg (FH), Merseburg

### Einstieg

EPUB ist zurzeit das populärste Format zur Publikation von E-Books. Diese lassen sich sowohl mit eigenständigen Lesegeräten auf E-Ink-Basis als auch über spezielle Anzeigeprogramme am PC oder auf Tablets / Smartphones wiedergeben. Die Inhalte passen sich im Gegensatz zu seitenorientierten Formaten wie PDF an die jeweilige Darstellungsumgebung an („reflowable content“). Auch und gerade das in diesem Jahr erschienene Apple iPad hat in Verbindung mit der Anwendung iBooks zur Akzeptanz von elektronischen Büchern im EPUB-Format beigetragen. Für die Technische Redaktion ergibt sich mit der Umsetzung von entsprechenden Produkten zugleich ein erweitertes Aufgabenfeld. Obwohl es bereits Software mit direkter Ausgabe auf Knopfdruck gibt, sind grundlegende Kenntnisse des Konzeptes von Vorteil, insbesondere für eine automatisierte Produktion nach dem Single-Source-Prinzip oder für die Nachbearbeitung.

### Allgemeines

Es handelt sich grundsätzlich um ein ZIP-komprimiertes Archivformat mit der Dateierweiterung .epub. Somit lässt sich ein erster Einblick durch Umbenennen der Endung einer vorhandenen Buchdatei in .zip und Öffnen mit einem geeigneten Entpacker gewinnen. Abbildung 1 veranschaulicht den typischen Dateiaufbau.

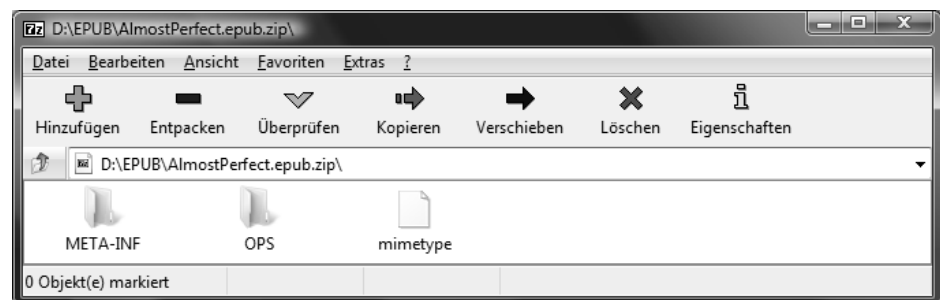


Abb. 1: Einblick in ein EPUB-Archiv (entpackt mit 7-Zip).

Die konkreten strukturellen und inhaltlichen Anforderungen werden durch die vom International Digital Publishing Forum (IDPF) herausgegebenen Spezifikationen definiert [1]. Aktuell wird EPUB 2.0 verwendet, eine fehlerbereinigte Version 2.0.1 ist bereits als Arbeitsfassung verfügbar. An Erweiterungen hinsichtlich neuer Marktentwicklungen wird ebenfalls gearbeitet.

Der EPUB-Standard subsummiert drei miteinander verwobene Einzelspezifikationen:

- Open Publication Structure (OPS): Beschreibt verwendbare Datenformate, u. a. für Texte, Bilder, Formatierungen und Schriftarten.
- Open Packaging Format (OPF): Enthält Angaben zu Metadaten sowie zum Aufbau von Inhaltsverzeichnis und Navigation.

- Open Container Format (OCF): Legt die formale Datei- und Verzeichnisstruktur sowie die Art der Kompression fest.

### Inhalte: XHTML, XML und CSS

Das EPUB-Format verwendet im Kern XHTML in Version 1.1 (alternativ DTBook) zur Auszeichnung der Inhalte und zusätzliche XML-Dokumente für die Steuerinformationen. Formatierungen lassen sich mit Cascading Stylesheets (CSS2) vornehmen. Übliche Raster- bzw. Vektorbildformate wie GIF, JPEG, PNG und SVG sowie über @font-face-Regeln eingebettete Schriftarten werden unterstützt. Die Einstiegshürde zur Erstellung von Inhalten liegt also für Autoren und Entwickler von Websites und Onlinehilfen relativ niedrig. Wesentlich ist jedoch standardkonformes Arbeiten auf allen Ebenen. Dabei helfen Validierungswerkzeuge wie EpubCheck [2].

Auf ausführlichere Beschreibungen und Codefragmente wird an dieser Stelle mit Hinweis auf einen kürzlich vom Autor verfassten Artikel zum Thema [3] und eine detailliertere Behandlung im Vortrag verzichtet.

### EPUB-Software

Im Grunde würden ein einfacher Texteditor und ein ZIP-Packer gepaart mit dem Wissen um die Zielstruktur bereits ausreichen. Praktikabler sind erwartungsgemäß spezialisierte Werkzeuge oder Exportfilter. Letztere stehen beispielsweise in Adobe InDesign [4] seit Version 3 und im Programm Help & Manual [5] zur Verfügung. Unter Verwendung bereits vorhandener XHTML-Dokumente und Stylesheets ermöglicht die plattformübergreifende freie Anwendung Anthemion eCub [6] nach kurzer Einarbeitung eine komfortable EPUB-Erstellung (vgl. Abbildungen 2 und 3).

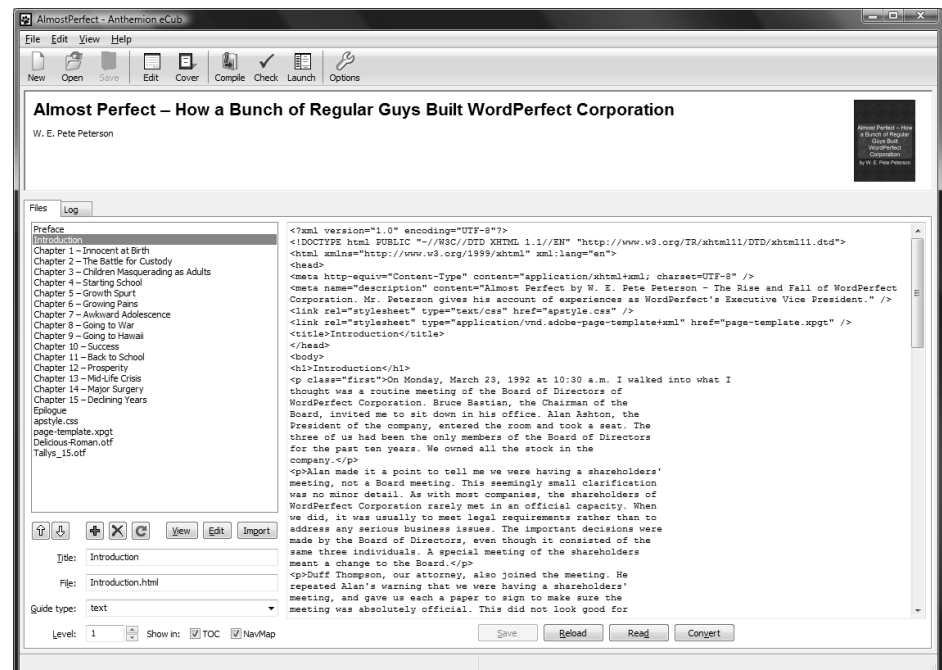


Abb. 2: Programmoberfläche von Anthemion eCub.

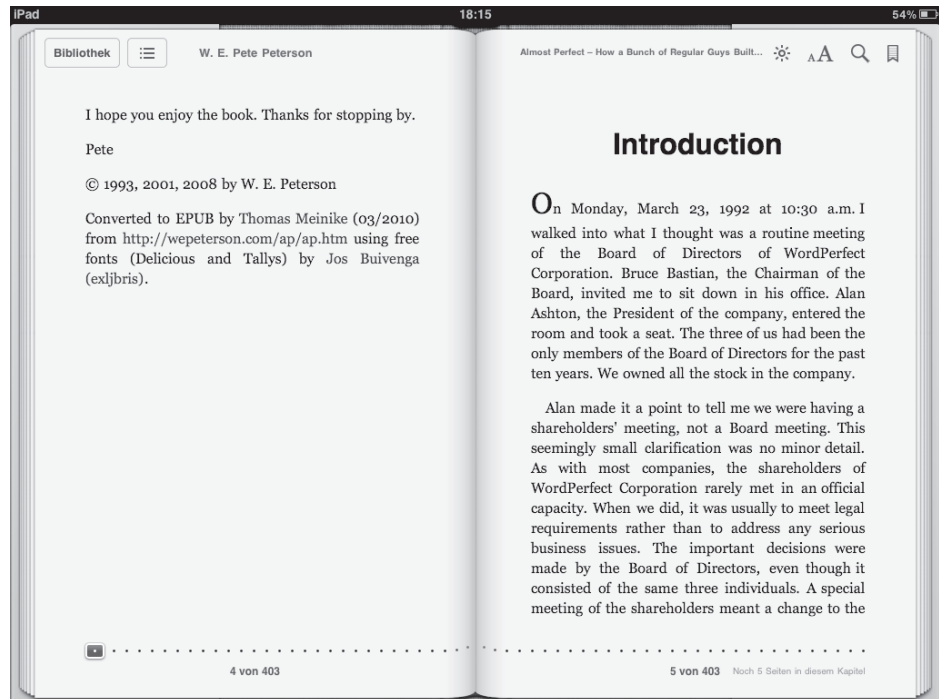


Abb. 3: Anzeige eines mit eCub erstellten E-Books auf dem iPad.

### Ein Ansatz zur Automatisierung

Das eingangs erwähnte Single-Source-Prinzip bietet sich auch zur EPUB-Produktion an. Eine entscheidende Voraussetzung ist die medienneutrale Datenhaltung vorzugsweise in XML-Dokumenten. Es existieren bereits Erweiterungen der bekannten XSLT-Stylesheets für die Arbeit mit DocBook oder DITA. Vom Autor wurde primär für Ausbildungszwecke das Projekt epubMinFlow konzipiert und umgesetzt. Ausgehend von einer vereinfachten XML-Dokumentstruktur ermöglicht eine Batch-Konsolenanwendung die Verarbeitung der XML-Inhalte (buch.xml) mit einem XSLT 2.0-Stylesheet (epub.xsl) zu allen erforderlichen Steuer- und Kapiteldateien bis hin zur finalen Kompression und Ablage des EPUB-Archivs (buch.epub). Weitere Informationen und ein lauffähiges Beispielprojekt stehen unter [7] zur Verfügung.

### Ausblick

EPUB hat sich in der Verlagslandschaft gut etabliert und lässt sich prinzipiell bei moderatem Aufwand kostengünstig produzieren. Durch die Einführung stärker multimedial nutzbarer Abspielgeräte (insbesondere iPad) sind auch neue Wünsche an das IDPF-Gremium herangetragen worden. Eine 2010 einberufene Arbeitsgruppe bemüht sich bereits um einen Nachfolgestandard (2.1 oder 3.0) mit zusätzlichen Möglichkeiten wie Videointegration und Interaktivität [8]. Hier dürfte auch die in Entwicklung befindliche HTML5-Spezifikation Spuren hinterlassen.

Da es aktuell auf dem Buchmarkt zum Thema selbst noch sehr dünn aussieht, sei auf das im Juli erschienene kompakte Buch [9] verwiesen. Es behandelt die EPUB-orientierte Nutzung von Microsoft Word und Adobe InDesign unter Einbeziehung nützlicher Tipps & Tricks für die praktische Arbeit und die damit verbundenen Fallstricke [9].

### Literaturangaben und Links

[1] <http://www.idpf.org/specs.htm>

[2] <http://code.google.com/p/epubcheck>

[3] Meinike, T.: Einfach publizieren und benutzen – EPUB-Format in Theorie und Praxis; Entwickler Magazin 4.2010, S. 99–106, PDF-Version unter: <http://www.iks.hs-merseburg.de/~meinike/artikel.php>

[4] <http://www.adobe.com/products/indesign>

[5] <http://www.helpandmanual.com>

[6] <http://www.juliansmart.com/ecub>

[7] <http://datenverdrahten.de/epubMinFlow>

[8] [http://idpf.org/idpf\\_groups/IDPF-EPUB-WG-Charter-4-6-2010.html](http://idpf.org/idpf_groups/IDPF-EPUB-WG-Charter-4-6-2010.html)

[9] Castro, E.: EPUB Straight to the Point – Creating ebooks for the Apple iPad and other ereaders; Peachpit Press 2010, ISBN 978-0-321-73468-6

*für Rückfragen: [thomas.meinike@hs-merseburg.de](mailto:thomas.meinike@hs-merseburg.de)*