

E-Books im EPUB-Format – Grundlagen und Ansätze der Erstellung

Arbeitstreffen der AG für Universitätsverlage 2014 – Merseburg, 03. April

Dr. Thomas Meinike

Hochschule Merseburg | FB Informatik und Kommunikationssysteme

<http://www.iks.hs-merseburg.de/~meinike/>

thomas.meinike@hs-merseburg.de

Zur Person

- Lehrkraft für besondere Aufgaben seit 1997
- Tätig in den Studiengängen im Fachbereich IKS:
 - Bachelor Technische Redaktion und E-Learning-Systeme
 - Master Technische Redaktion und Wissenskommunikation (Informationsdesign und Medienmanagement)
- Lehr- und Arbeitsgebiete:
 - Auszeichnungs- und Skriptsprachen
 - Online-Dokumentation und Web-Entwicklung
 - XML-Technologien



Technische Redaktion in Merseburg

- Diplomstudiengang 1995 bis 2008
Kommunikation und Technische Dokumentation
- Masterstudiengang seit 2006
Technische Redaktion und Wissenskommunikation
ab 2014: Informationsdesign und Medienmanagement (4 Sem. / M. A.)
- Bachelorstudiengang seit 2010
Technische Redaktion und
E-Learning-Systeme
(6 Sem. / B. Eng.)
- E-Book-Themen sind Teil
des TR-Studiums.



EPUB im Überblick (1/7)

- Kurzform für **E**lectronic **PUB**lication (alternativ ePUB, ePub).
- Herausgeber: IDPF (International Digital Publishing Forum, idpf.org).
- Version 2.0 (2007) / Revision **2.0.1** (2009), **3.0** (2011).
- Verwendet etablierte Technologien wie XML, XHTML, CSS und ist somit in existierende Publikationsprozesse integrierbar.
- Typischer Einsatz auf E-Ink-Lesegeräten sowie zunehmend auf Smartphones und Tablets über Apps.
- Nicht auf der Kindle-Plattform, dort MOBI- bzw. KF8-Format.

EPUB im Überblick (2/7)

→ EPUB-Dateien (.epub) sind ZIP-gepackte Archive („Mini-Websites“).



→ Im Archiv sind die Verzeichnisse **META-INF** und **OPS** enthalten sowie die unkomprimierte (!) Datei **mimetype** (*application/epub+zip*).

→ **META-INF** enthält **container.xml** (verweist auf **content.opf** mit Paketinformationen, meistens im Verzeichnis **OPS** gelegen). Zusätzlich können dort DRM-Informationen liegen.

→ Die eigentlichen Inhalte in Form von XHTML-Dokumenten, CSS-, Bild- und ggf. Font-Dateien werden ebenfalls in **OPS** hinterlegt.

Hinweis: Pflichtangaben sind rot markiert, sonstige Namen können frei gewählt werden!

EPUB im Überblick (3/7)

→ content.opf beschreibt detailliert die Paketstruktur (XML)

- ✓ Metadaten wie mindestens Buchtitel, Sprache, ID (ISBN) usw.
- ✓ Auflistung der Navigations-, Inhalts- und Zusatzdateien des Buches
- ✓ Steuerinformationen und optionale Angaben zur Bedeutung der Einträge

```
<?xml version="1.0" encoding="UTF-8"?>  
<package xmlns="http://www.idpf.org/2007/opf"  
  xmlns:dc="http://purl.org/dc/elements/1.1/"  
  unique-identifier="BookId" version="2.0">
```

```
<metadata> ... </metadata>
```

```
<manifest> ... </manifest>
```

```
<spine toc="toc"> ... </spine>
```

```
<guide> ... </guide>
```

```
</package>
```

EPUB im Überblick (4/7)

→ content.opf ... metadata (dc = Dublin Core)

```
<metadata>
  <dc:title>Titel</dc:title>
  <dc:language>de</dc:language>
  <dc:identifizier id="BookId">http://epub.example.net</dc:identifizier>
  <!-- weitere optionale Angaben ... -->
</metadata>
```

→ content.opf ... manifest (Inhaltsdateien, IDs und Medientypen)

```
<manifest>
  <item id="toc" href="toc.ncx" media-type="application/x-dtbncx+xml"/>
  <item id="CoverDesign" href="CoverDesign.jpg" media-type="image/jpeg"/>
  <item id="CoverPage" href="CoverPage.html" media-type="application/xhtml+xml"/>
  <item id="style" href="style.css" media-type="text/css"/>
  <item id="kap1" href="kapitel1.html" media-type="application/xhtml+xml"/>
  <!-- weitere item-Elemente ... -->
</manifest>
```

EPUB im Überblick (5/7)

→ content.opf ... spine (Archivteile und Abfolge beim Blättern)

```
<spine toc="toc">  
  <itemref idref="CoverPage" linear="no"/>  
  <itemref idref="kap1"/>  
  <itemref idref="kap2"/>  
  <itemref idref="kap3"/>  
</spine>
```

→ content.opf ... guide (Bedeutung der Abschnitte wie Cover, Text, Glossar, Index, ...)

```
<guide>  
  <reference type="cover" title="Titelseite" href="CoverPage.html"/>  
  <reference type="text" title="Kapitel 1" href="kapitel1.html"/>  
  <reference type="text" title="Kapitel 2" href="kapitel2.html"/>  
  <reference type="text" title="Kapitel 3" href="kapitel3.html"/>  
</guide>
```

EPUB im Überblick (6/7)

→ toc.ncx (Inhaltsverzeichnis / Navigation)

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ncx ...>
<ncx xmlns="http://www.daisy.org/z3986/2005/ncx/" xml:lang="de" version="2005-1">
  <head> ... </head>
  <docTitle><text>Titel</text></docTitle>

  <navMap>
    <navPoint id="navPoint-1" playOrder="1">
      <navLabel><text>Inhaltsverzeichnis</text></navLabel>
      <content src="inhalt.html"/>
    </navPoint>
    <navPoint id="navPoint-2" playOrder="2">
      <navLabel><text>Kapitel 1</text></navLabel>
      <content src="kapitel1.html"/>
    </navPoint>
    <!-- weitere navPoint-Elemente ... -->
  </navMap>
</ncx>
```

EPUB im Überblick (7/7)

→ Inhalte auf Basis von XHTML 1.1 (EPUB 2)

hx, p, em, strong, a, img, object, dl, ol, ul, table, pre, div, span, style, link, ...

→ Bildformate: GIF, JPEG, PNG, SVG

→ Formatierung: CSS 2.1

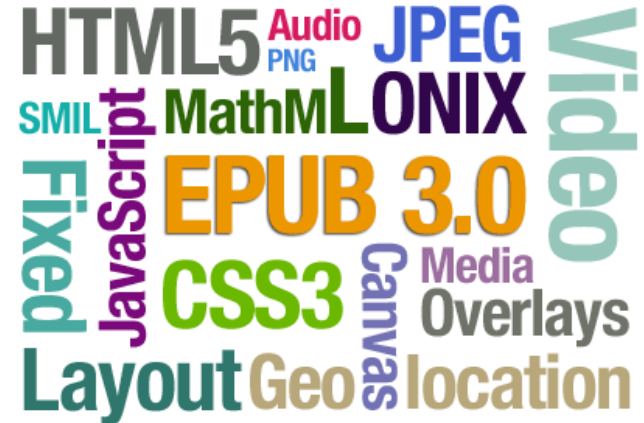
anspruchsvoll in Bezug auf unterschiedliche Reader und Apps

→ Ausblick auf EPUB 3

(X)HTML5, CSS3, JavaScript, MathML, Audio, Video, Animation, Interaktion, Inhalt via nav-Element statt toc.ncx, ...

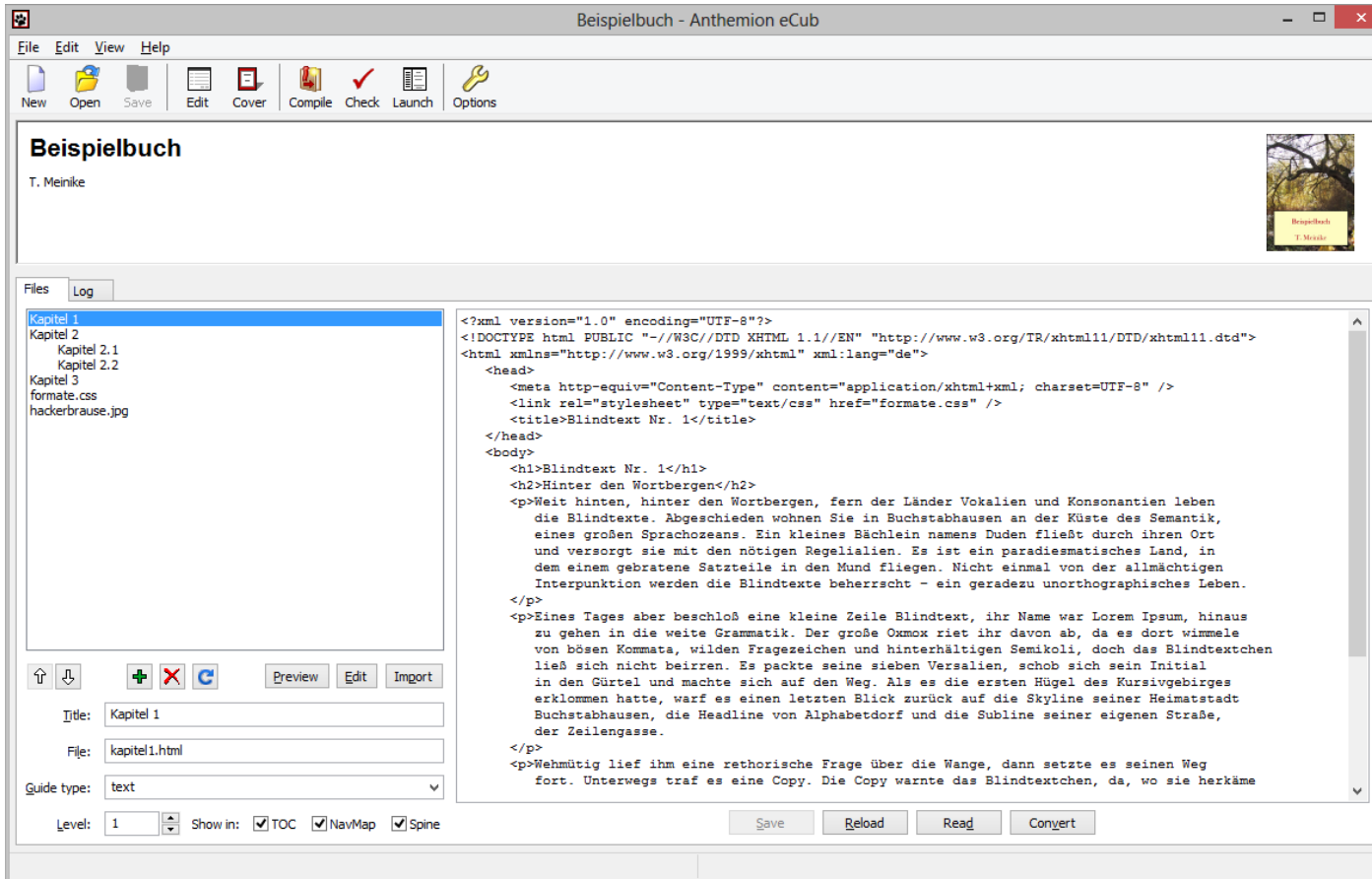
Siehe: <http://epubtest.org/results/>

Bildquelle: <http://www.epub-ebooks.de/>



EPUB-Erstellung (1/7)

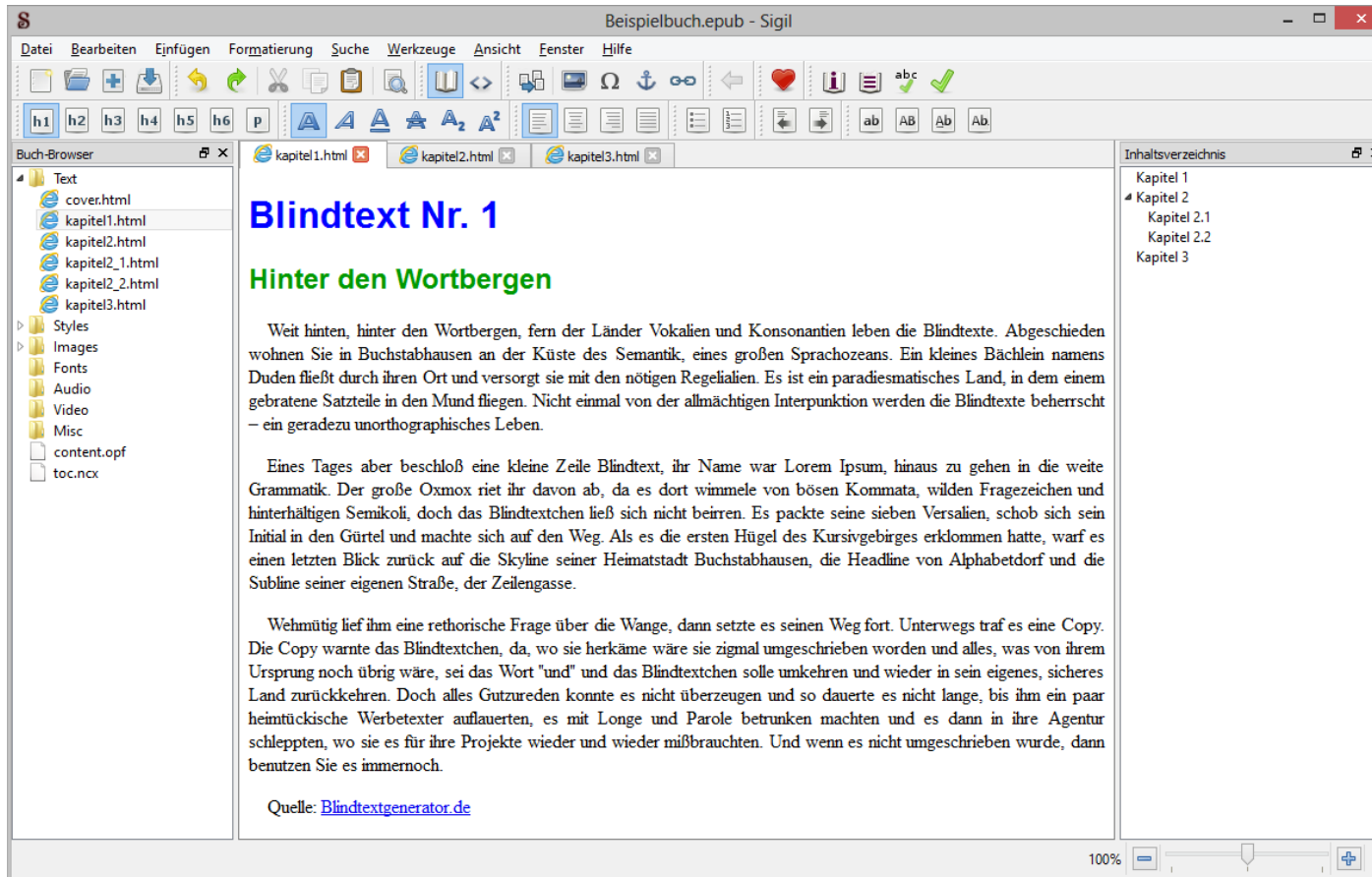
➔ Mit spezieller Software und selbst verfasstem HTML (eCub, frei)



HTML-
Dateien
liegen
separat vor.

EPUB-Erstellung (2/7)

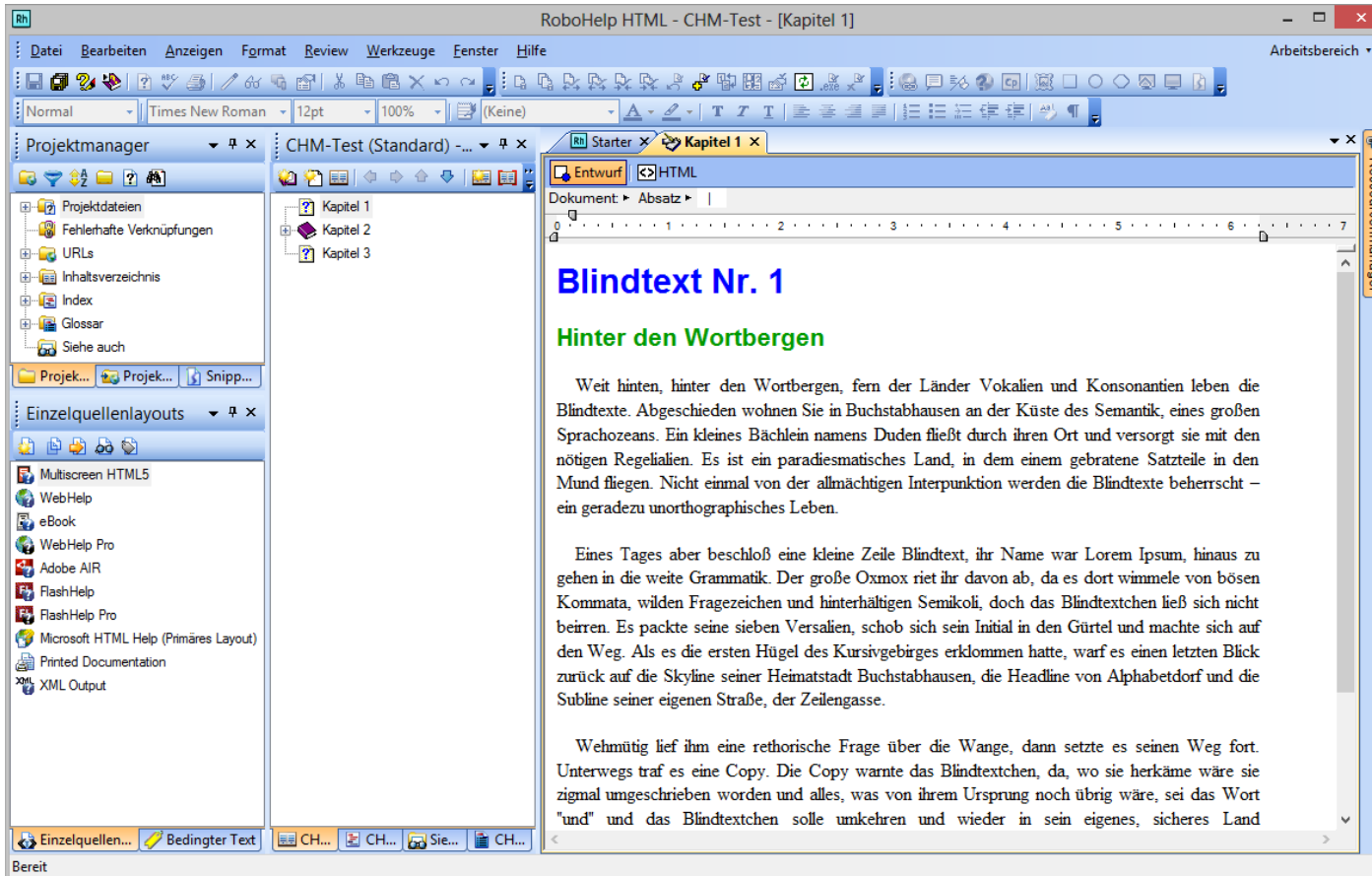
➔ Mit spezieller Software und selbst verfasstem HTML (Sigil, frei)



Direkte
Arbeit am
EPUB-
Archiv.

EPUB-Erstellung (3/7)

➔ Mit spezieller Software und selbst verfasstem HTML (RoboHelp, komm.)

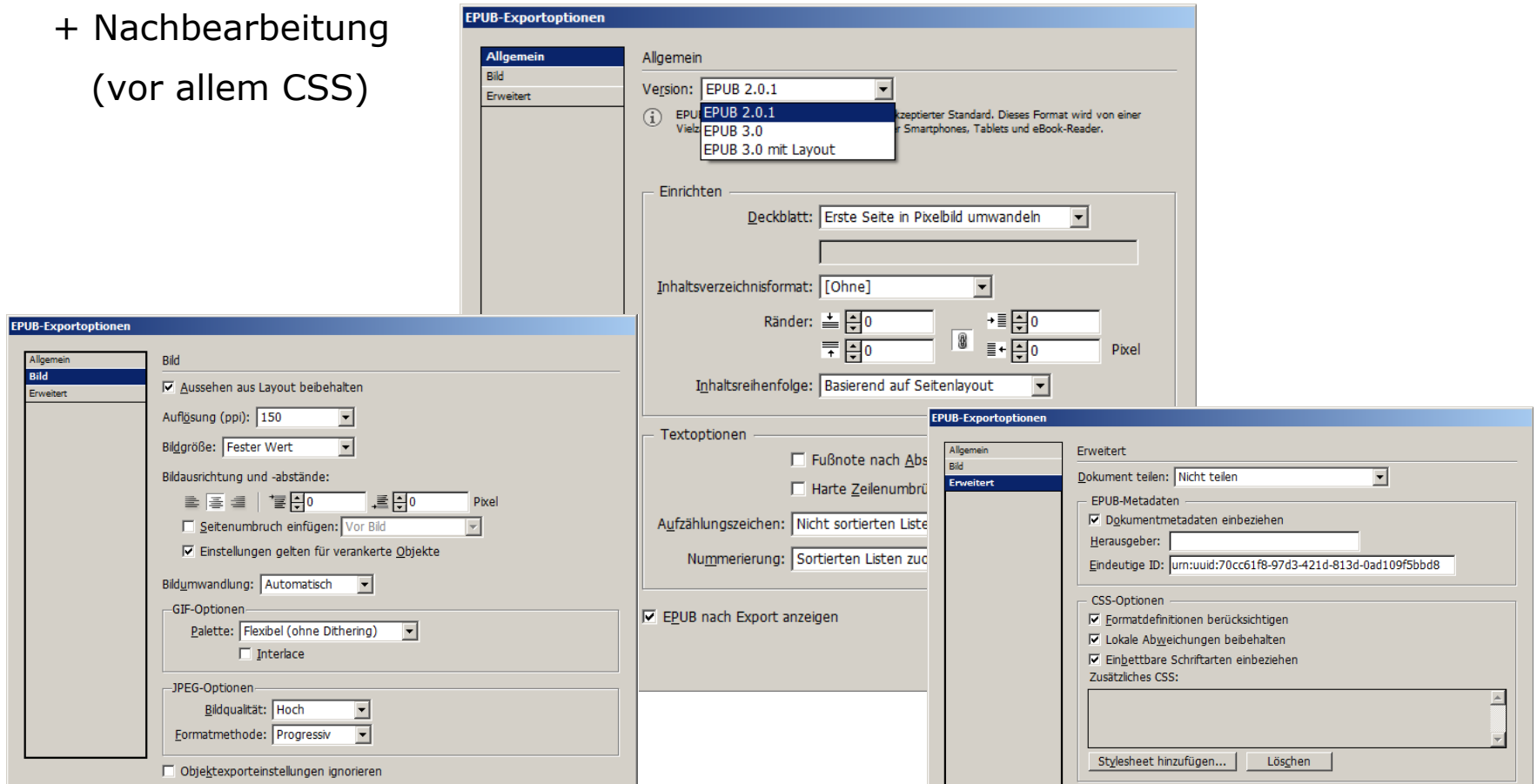


Ausgabe
als CHM,
EPUB,
WebHelp.

EPUB-Erstellung (4/7)

→ Über Exportoptionen von Layoutsoftware wie InDesign (komm.)

+ Nachbearbeitung
(vor allem CSS)



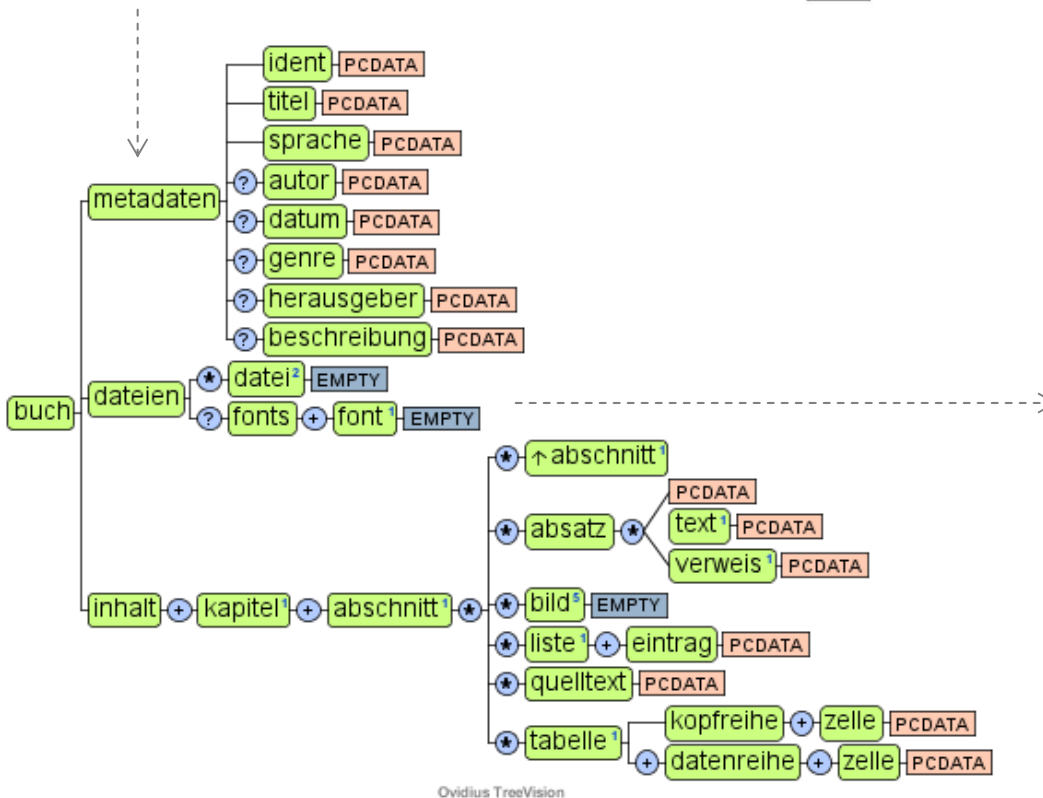
EPUB-Erstellung (5/7)

- Konvertieren von Formaten nach EPUB mit Calibre (frei).
- Speichern und aufwändiges Nachbearbeiten von (X)HTML aus Office-Anwendungen sowie manuelles Zusammenbauen zum E-Book.
- Komplexere Verarbeitung von Office-Archiven wie DOCX (enthaltene XML-Datenstrukturen wie document.xml verarbeiten).
- Nutzung von Konvertierungs-Workflows um aus eigenen XML-Daten die EPUB-relevante Struktur aufzubauen (erfordert Kenntnisse in der XML-Verarbeitung mit XSLT).
- Übergabe von Autorendokumenten an einen spezialisierten Dienstleister für Verlagsproduktionen.

EPUB-Erstellung (6/7)

→ datenverdrahten.de/epubMinFlow

XML-Dokument + XSLT 2.0 →



```
C:\Windows\system32\cmd.exe - run
D:\epubMinFlow>run
EPUB-Build-Prozess gestartet ...

* Vorbereitung ...

style\standard.css
  1 Datei(en) kopiert.
bilder\cover.jpg
bilder\ereader.png
bilder\grafik.svg
bilder\pizza.jpg
  4 Datei(en) kopiert.
fonts\Delicious-Roman.otf
fonts\Tallys_15.otf
  2 Datei(en) kopiert.

* XSLT-Prozess ...

* Kompression ...

adding: mimetype (stored 0%)
adding: META-INF/container.xml (deflated 32%)
adding: OPS/bilder/cover.jpg (deflated 3%)
adding: OPS/bilder/ereader.png (deflated 0%)
adding: OPS/bilder/grafik.svg (deflated 63%)
adding: OPS/bilder/pizza.jpg (deflated 0%)
adding: OPS/content.opf (deflated 73%)
adding: OPS/fonts/Delicious-Roman.otf (deflated 43%)
adding: OPS/fonts/Tallys_15.otf (deflated 18%)
adding: OPS/html/buchcover.html (deflated 39%)
adding: OPS/html/inhalt.html (deflated 52%)
adding: OPS/html/kapitel1.html (deflated 49%)
adding: OPS/html/kapitel2.html (deflated 50%)
adding: OPS/html/kapitel3.html (deflated 44%)
adding: OPS/html/kapitel4.html (deflated 60%)
adding: OPS/style/standard.css (deflated 57%)
adding: OPS/toc.ncx (deflated 79%)

... EPUB (buch.epub) wurde erzeugt.

* EPUB-Tests ...

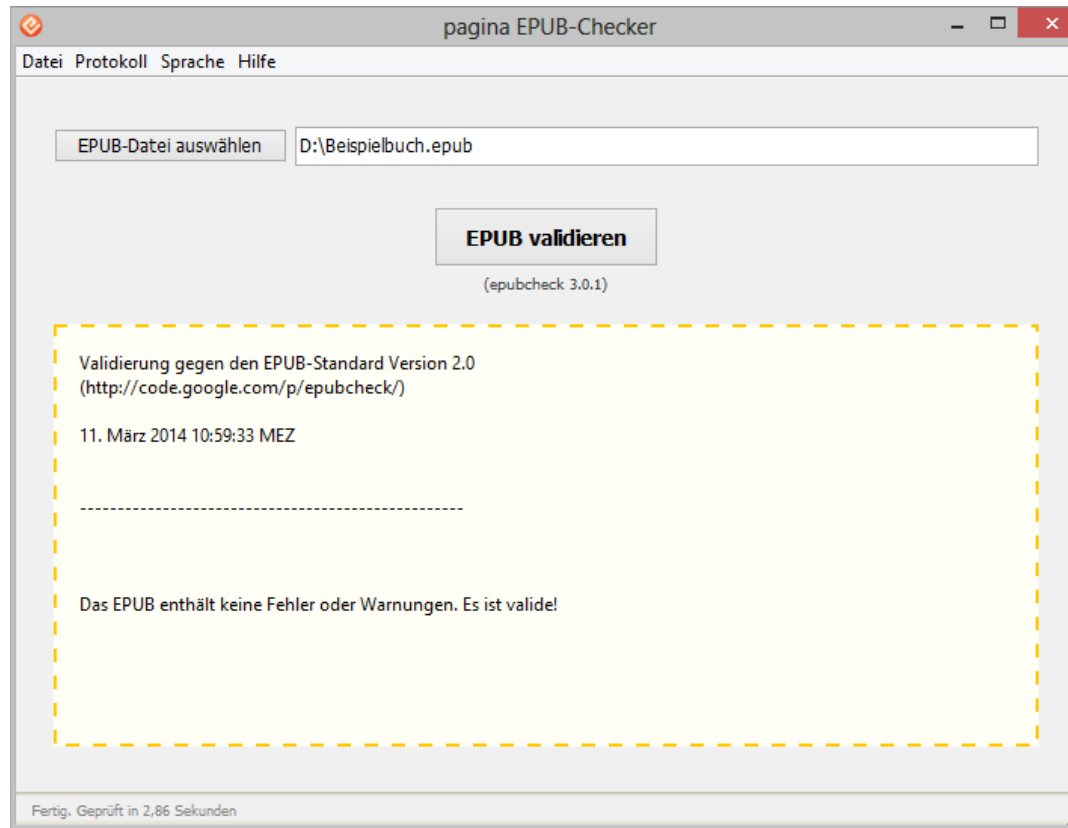
Epubcheck Version 3.0.1

Validating against EPUB version 2.0
No errors or warnings detected.
```

EPUB-Erstellung (7/7)

→ Konformität testen!

EpubCheck als Java-Konsolentool oder mit grafischer Oberfläche von pagina:



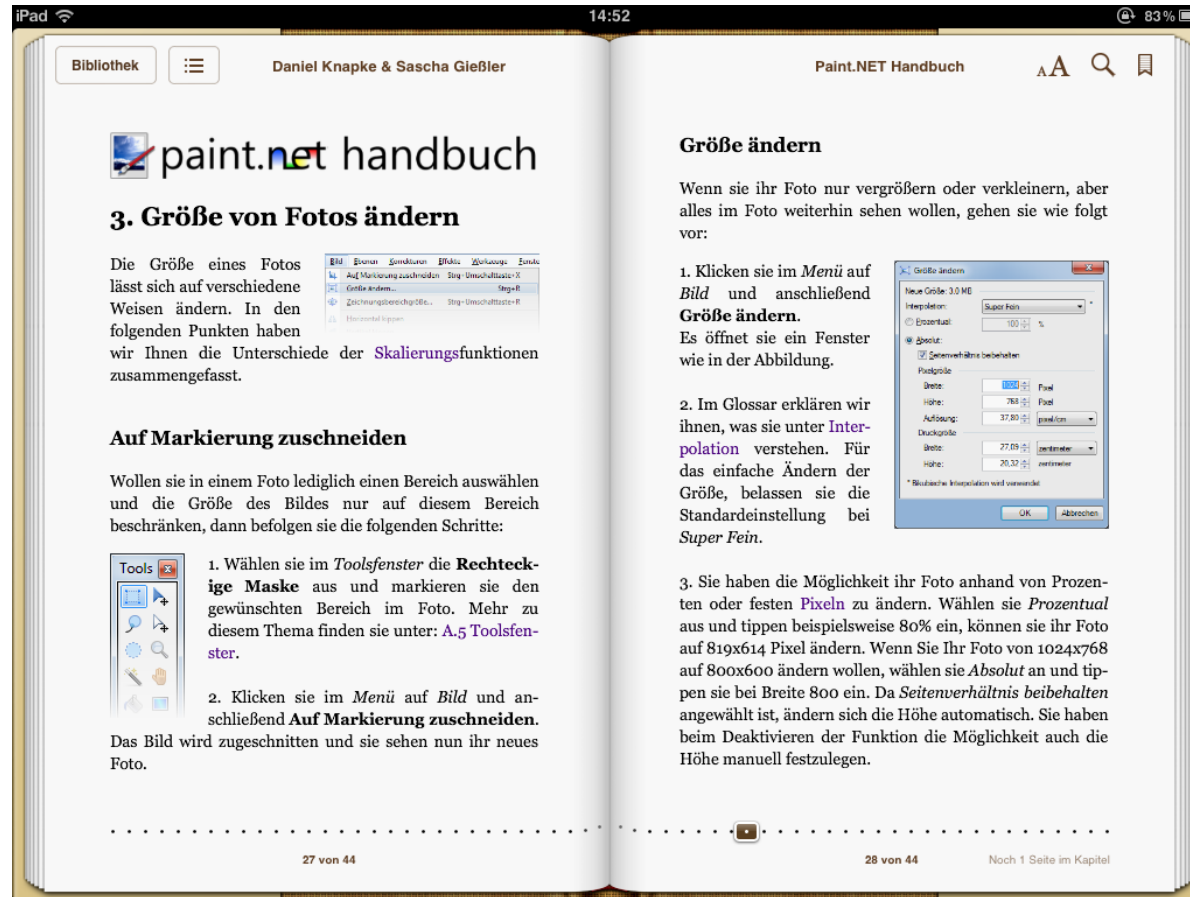
EPUB-Beispiel aus einem Kurs (1/2)

→ Inhaltsverzeichnis (iPad-Ansicht / iBooks-App)



EPUB-Beispiel aus einem Kurs (2/2)

➔ Inhaltsansicht (iPad-Ansicht / iBooks-App)



Zusammenfassung und Ausblick

- EPUB 2 ist ein etabliertes Format für elektronische Bücher.
- Zur Produktion lassen sich je nach Anforderung unterschiedliche Ansätze nutzen (einfache Software bis hin zu komplexen Workflows).
- Die Vielzahl an existierenden Geräten und Programmen zur Darstellung von E-Books erfordert umfangreiche Tests.
- EPUB 3 ermöglicht bei konsequenter Anwendung und Unterstützung auf den Endgeräten künftig neue Publikationsformen mit zusätzlicher Interaktivität, bessere Umsetzung von Vektorgrafiken und Formeln, fixierte Layouts uvm.

Referenzlinks

- ✓ Buchtipp: E-Books mit InDesign CC: <http://www.epub-fuer-indesigner.de/>
- ✓ Software Adobe InDesign: <http://www.adobe.com/de/products/indesign.html>
- ✓ Software Adobe RoboHelp: <http://www.adobe.com/de/products/robohelp.html>
- ✓ Software Calibre: <http://calibre-ebook.com/>
- ✓ Software eCub: <http://www.anthemion.co.uk/ecub/>
- ✓ Software EpubCheck: <https://github.com/IDPF/epubcheck>
- ✓ Software pagina EPUB-Checker: <http://www.pagina-online.de/produkte/epub-checker/>
- ✓ Software Sigil: <https://code.google.com/p/sigil/>
- ✓ Software Adobe Digital Editions: <http://www.adobe.com/products/digital-editions.html>
- ✓ Software Apple iBooks: <http://www.apple.com/de/ibooks/>
- ✓ Software AZARDI: <http://azardi.infogridpacific.com/>
- ✓ Software Adobe Digital Editions: <http://www.adobe.com/products/digital-editions.html>
- ✓ Workflow-Beispiel: epubMinFlow: <http://datenverdrahten.de/epubMinFlow/>