

E-Book-Erstellung

aus \LaTeX und HTML

Dominik Wagenführ

Ubucon, Berlin, Oktober 2012

Eigenschaften eines Buches

- feste Seiten mit Seitenzahlen
- Rand
- Textkörper, Kopf- und Fußzeile
- unveränderbare Schriftgröße
- trifft teilweise auch auf digitale Formate wie PDF zu

- Fließtext
- keine festen Seiten
- Schriftgröße und Schriftart lässt sich ändern
- Verlinkungen (Hyperlinks) möglich

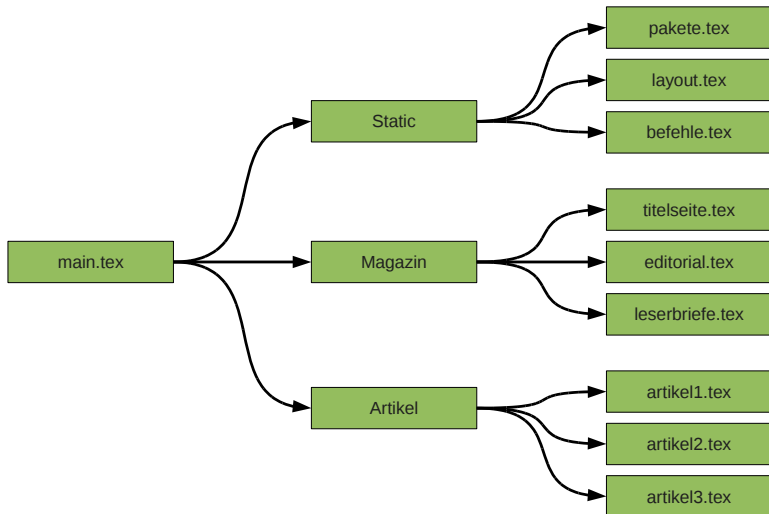
- MOBI-Format¹
 - nur für Amazon Kindle
 - proprietäres Binärformat
- EPUB-Format²
 - offenes, standardisiertes Container-Format
 - packt XHTML-kodierte Dateien und Medieninhalte zusammen
 - unterstützt Digitalen Rechteverwaltung (DRM)
 - ab EPUB 3 auch Audio- und Videoinhalte

¹<https://de.wikipedia.org/wiki/Mobipocket>

²<https://de.wikipedia.org/wiki/EPUB>

- Magazin für Linux und Open Source³
- Bestehen seit März 2006
- erste Ausgabe in ODT, danach L^AT_EX
- seit Januar 2008 als HTML
- seit Januar 2012 als EPUB

³<http://www.freiesmagazin.de/>



openSUSE 12.1 von Mirko Lindner

In der Version 12.1 will openSuse noch besser, leichter, stabiler und überhaupt das beste openSUSE sein. Der Hersteller selbst verspricht, mit dem neuesten Produkt ein weitgehend universelles System für Desktop-, Netbook- und Serveranwender zu liefern. Der Artikel wirft einen genauen Blick auf die Distribution und testet sie in produktiven Umgebungen.

Redaktioneller Hinweis: Der Artikel „openSUSE 12.1“ erschien erstmals bei Pro-Linux [1].


Einführung

Mit der aktuellen Version 12.1 von openSUSE erscheint nun erstmals eine Ausgabe der bekannten Distribution unter dem Dach des neuen SUSE-Eigentümers Attachmate [2]. Über acht Monate lang werkten die Entwickler an der neuen Version. Begleitet von zahlreichen Alpha- und Beta-Versionen gelang es dem Team, den ehrgeizigen Plan einzuhalten und openSUSE 12.1 pünktlich zum versprochenen Termin auszuliefern.

Doch was darf der Anwender von dem neuen Produkt erwarten? Der Versionssprung von 11.4 auf 12.1 überrascht zuerst, war es doch Tradition, die Hauptversionen immer mit der Unterversion 0 beginnen zu lassen. Doch dies will das Team mit der aktuellen Version gerade ändern. Der Eindruck einer „Hauptversion“ soll mit der Nomenklatur 12.1 eliminiert werden. So wurde offenbar den

„x.0“-Versionen in der Vergangenheit mehr Aufmerksamkeit geschenkt als anderen. Da jedoch jede neue openSUSE-Version viel Neues enthält, verzichten die Entwickler zukünftig auf „x.0“.



Das Startbild von openSUSE 12.1. 

Lieferumfang

Die Gemeinschaft rund um die freie Distribution liefert auch mit openSUSE 12.1 einen gewohnt großen Umfang an Software. So setzt das Produkt in der aktuellen Version auf die im Oktober veröffentlichte Version 3.1 des Kernels auf, die unter anderem Optimierungen beim Zugriff auf RAM enthält und gegenüber der letzten Version zahlreiche Änderungen erfuh.

Die Basis der Distribution stellen glibc 2.14.1, gcc 4.6 und der X-Server 1.10.4 dar. Ferner flossen GNOME 3.2 und KDE 4.7 in den Lieferumfang von openSUSE ein. Zur Grundausstattung gehören daneben ALSA 1.0.24.1, CUPS 1.5.0, Postfix 2.8.5 und Mesa 7.11. Bei Grub setzt die Distribution weiterhin auf die alte Version des Boot-Loaders auf, was unter anderem dazu führt, dass bei der Auswahl von Btrfs nun eine dedizierte Boot-Partition erstellt werden muss.

Erweitert man die Quellen um die allseits bekannten Alternativquellen, so gibt es kaum eine Applikation, die nicht für openSUSE angeboten wird. Es spielt dabei kaum eine Rolle, ob man ein Programm, eine Bibliothek oder um ein Modul für eine Sprache sucht, denn openSUSE bietet fast alles auch in den neuesten Versionen an. Eine schier unerschöpfliche Quelle von neuen und aktualisierten Paketen ist auch der Build-Service des Projektes.

Anwender, die sich für eine Downloadversion [3] entscheiden, stehen vor der Qual der Wahl, denn zur Auswahl stehen wie gewohnt mehrere Varianten. So bietet das Team DVD-Medien für 32- und 64-Bit-Systeme. Darüber hinaus gibt es auf der Software-Seite noch spezielle Live-CD-Versionen mit GNOME oder wahlweise KDE

- Querformat
- grafische Kopf- und Fußzeile
- dreispaltig
- zahlreiche texumflossene Bilder
- viele Listings

- direkte Wandlung von PDF nach EPUB nur schwer möglich

Von L^AT_EX nach HTML: Werkzeuge

- TeX4ht⁴
- LaTeX2HTML⁵
- LaTeXML⁶
- Hyperlatex⁷
- Pandoc⁸
- Hevea⁹
- tth¹⁰

⁴<http://www.tug.org/applications/tex4ht/>

⁵<http://www.latex2html.org/>

⁶<http://dlmf.nist.gov/LaTeXML/>

⁷<http://hyperlatex.sourceforge.net/>

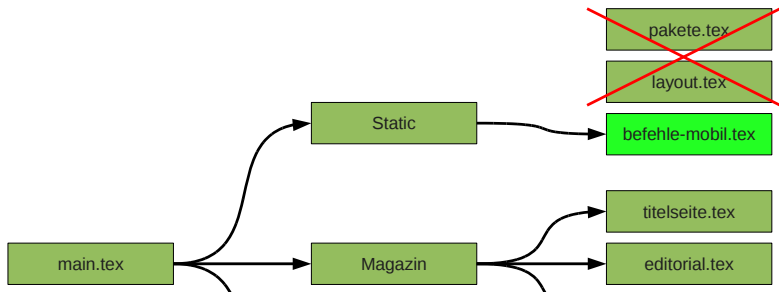
⁸<http://johnmacfarlane.net/pandoc/>

⁹<http://para.inria.fr/~maranget/hevea/>

¹⁰<http://silas.psfc.mit.edu/tth/>

Von \LaTeX nach HTML: tth

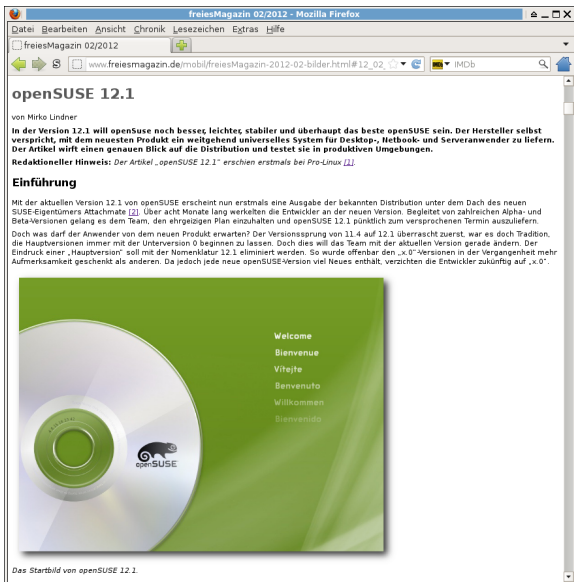
- letzte Version: Juli 2012
- fest eingebauter Sprachschatz
- liest keine Pakete per `\usepackage`



befehle-mobil.tex

- keine Kopf- oder Fußzeile
- einspaltiges Layout
- alle Längenberechnungen entfernen
- keine Spalten- oder Seitenumbrüche
- „normale“ Positionierung von Bildern
- einige Sonderzeichen ($\backslash \sim |$) maskieren
- kein `\ifthenelse`
- keine optionalen Argumente
- Artikelüberschriften als `\chapter`
- keine Listings (Paket `listings`)
- nur Standardtabellenformat
- keine absolute Positionierung (Paket `textpos`)

- unnötige Umbrüche entfernen
- doppelte Absatzabstände entfernen
- maskierte Sonderzeichen zurückwandeln (siehe oben)
- echte Anführungszeichen setzen: „...“ anstatt „...“
- Tabellenlayout korrigieren (feste Breite raus, einfacher Rahmen)
- Bildergröße mittels `style="max-width:100%;"` anpassen
- Meta-Daten (Titel, Herausgeber, etc.) des Dokuments setzen



The screenshot shows a Mozilla Firefox browser window with the address bar containing `www.freiesmagazin.de/mobil/freiesMagazin-2012-02-bilder.html#12_02`. The page title is "freiesMagazin 02/2012". The main content area features the article title "openSUSE 12.1" and the author "von Mirko Lindner". The article text includes a bolded introductory sentence, a redaction note, and an "Einführung" section. Below the text is a large image of the openSUSE 12.1 welcome screen, which shows a CD-ROM on the left and a list of welcome messages in various languages on the right. The CD-ROM has the openSUSE logo and the text "openSUSE" on it. The welcome messages are: "Welcome", "Bienvenue", "Vítejte", "Benvenuto", "Willkommen", and "Bienvenido". Below the image is the caption "Das Startbild von openSUSE 12.1."

freiesMagazin 02/2012 - Mozilla Firefox

freiesMagazin 02/2012

`www.freiesmagazin.de/mobil/freiesMagazin-2012-02-bilder.html#12_02`

openSUSE 12.1

von Mirko Lindner


In der Version 12.1 will openSUSE noch besser, leichter, stabiler und überhaupt das beste openSUSE sein. Der Hersteller selbst verspricht, mit dem neuesten Produkt ein weitgehend universelles System für Desktop-, Netbook- und Serveranwender zu liefern. Der Artikel wirft einen genaueren Blick auf die Distribution und testet sie in produktiven Umgebungen.

Redaktioneller Hinweis: Der Artikel „openSUSE 12.1“ erschien erstmals bei Pro-Linux [\[1\]](#).

Einführung

Mit der aktuellen Version 12.1 von openSUSE erscheint nun erstmals eine Ausgabe der bekannten Distribution unter dem Dach des neuen SUSE-Eigentümers Attachmate [\[2\]](#). Über acht Monate lang werkten die Entwickler an der neuen Version. Begleitet von zahlreichen Alpha- und Beta-Versionen gelang es dem Team, den ehrgeizigen Plan einzuhalten und openSUSE 12.1 pünktlich zum versprochenen Termin auszuliefern.

Doch was darf der Anwender von dem neuen Produkt erwarten? Der Versionssprung von 11.4 auf 12.1 überrascht zuerst, war es doch Tradition, die Hauptversionen immer mit der Unterversion 0 beginnen zu lassen. Doch dies will das Team mit der aktuellen Version gerade ändern. Der Eindruck einer „Hauptversion“ soll mit der Nomenklatur 12.1 eliminiert werden. So wurde offenbar den „x.0“-Versionen in der Vergangenheit mehr Aufmerksamkeit geschenkt als anderen. Da jedoch jede neue openSUSE-Version viel Neues enthält, verzichten die Entwickler zukünftig auf „x.0“.



Das Startbild von openSUSE 12.1.

Programme:

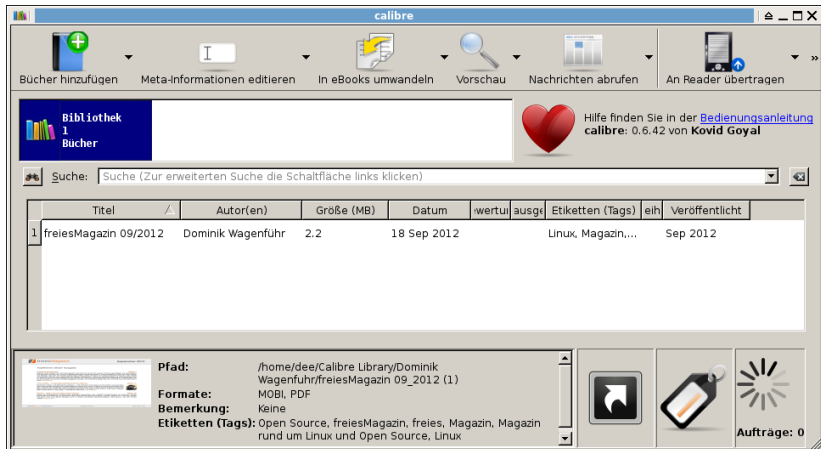
- Calibre¹¹
 - Format und Aufteilung nicht gut
 - fehlerhafte Umbrüche
- Pandoc¹²
 - ähnliche Probleme wie Calibre
 - CSS wird teilweise ignoriert
- Sigil¹³
 - inoffizielle Konsolenversion
 - benötigt ein paar Patches im Quellcode
 - kann leicht Inhaltsverzeichnis erstellen
 - Artikel an SGF-Chapter-Marks trennbar

¹¹<http://calibre-ebook.com/>

¹²<http://johnmacfarlane.net/pandoc/>

¹³<https://code.google.com/p/sigil/>

EPUB-Konvertierung: Calibre



The screenshot shows the Calibre application window. The title bar reads "calibre". The menu bar includes: Bücher hinzufügen, Meta-Informationen editieren, In eBooks umwandeln, Vorschau, Nachrichten abrufen, and An Reader übertragen. Below the menu bar, there is a "Bibliothek" section with "1 Bücher". A search bar contains the text "Suche: Suche (Zur erweiterten Suche die Schaltfläche links klicken)". A table displays a single book entry:

	Titel	Autor(en)	Größe (MB)	Datum	wertul	ausge	Etiketten (Tags)	eih	Veröffentlicht
1	freiesMagazin 09/2012	Dominik Wagenfuhr	2.2	18 Sep 2012			Linux, Magazin,...		Sep 2012

Below the table, the book's metadata is shown:

- Pfad:** /home/dee/Calibre Library/Dominik Wagenfuhr/freiesMagazin 09_2012 (1)
- Formate:** MOBI, PDF
- Bemerkung:** Keine
- Etiketten (Tags):** Open Source, freiesMagazin, freies, Magazin, Magazin rund um Linux und Open Source, Linux

On the right side of the metadata panel, there are three icons: a square with a white arrow pointing up and left, a tag icon, and a loading spinner icon. Below the spinner icon, it says "Aufträge: 0".

EPUB-Konvertierung: Sigil

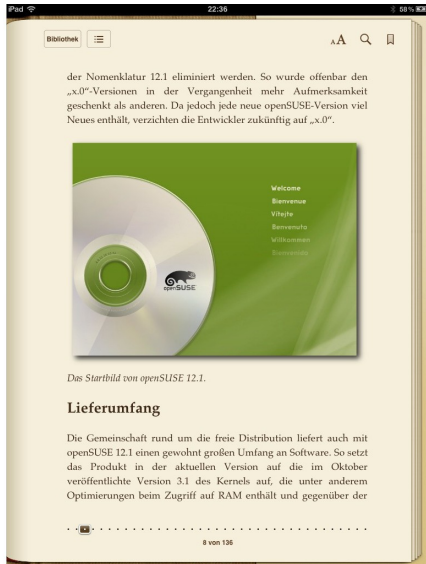
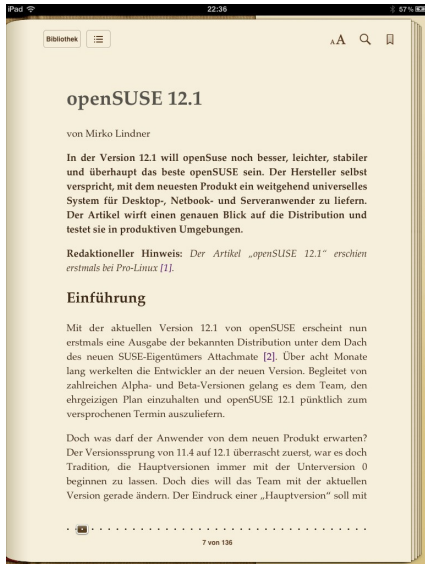
The screenshot displays the Sigil EPUB editor window titled "freiesMagazin-2012-09-bilder.epub - Sigil". The interface includes a menu bar (File, Edit, View, Format, Quick Settings, Window, Help), a toolbar with various editing tools, and a status bar at the bottom showing "Line: 0, Col: 0" and "100%".

The main editing area is divided into three panes:

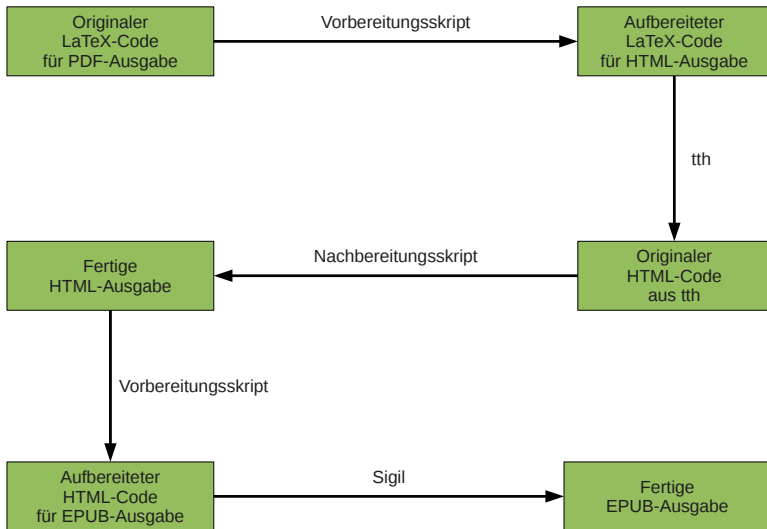
- Book Browser:** Shows a tree view of the EPUB's internal structure, including a "Text" folder with the current file "freiesMagazin-2012-09-bilder.epub.html" selected, and folders for "Styles", "Images", "Fonts", "Misc", "toc.ncx", and "content.opf".
- Main Editor:** Displays the content of the selected file, which is the magazine cover page. The text reads:
freiesMagazin
September 2012
(ISSN 1867-7991)
- Table of Contents:** Lists the contents of the EPUB, including:
 - Topthemen dieser Ausgabe
 - Editorial
 - Passwortsicherheit
 - Der August im Kernelrückblick
 - Tiny & Big - Unterwäschefetischismus de...
 - Redis - Mehr als ein Key-Value-Store
 - Taskwarrior - What's next? (Teil 2)
 - Rezension: Schrödinger programmiert C++
 - Rezension: Seven Databases in Seven W...
 - Rezension: Android - Apps entwickeln
 - Leserbriefe
 - Veranstaltungskalender
 - Vorschau
 - Konventionen
 - Impressum ISSN 1867-7991

At the bottom of the Table of Contents pane, there is a button labeled "Generate TOC from headings".

- Inhaltsverzeichnis entfernen
- „Titelseite“ erstellen
- jeder Artikel auf neue Seite



Ablauf: Von \LaTeX nach EPUB



Gründe:

- E-Book-Reader nicht so großzügig wie Webbrowser
- korrekte EPUB-Formatierung sinnvoll

Werkzeuge:

- Sigil¹⁴
- EpubCheck¹⁵

¹⁴<https://code.google.com/p/sigil/>

¹⁵<http://code.google.com/p/epubcheck/>

EPUB-Validierung: Sigil

The screenshot shows the Sigil EPUB editor interface. The main window displays the HTML source code for a file named 'freiesMagazin-2012-09-bilder-epub_0004.xhtml'. The code includes a `<dl compact='compact'>` tag with a `<dt class='sgc-10'>` element. The text content describes a brute-force attack on a password. The right sidebar shows a 'Table of Contents' with various sections like 'Editorial', 'Passwortsicherheit', and 'Rezeption: Schrödinger programmiert C++'. Below the main editor, a 'Validation Results' window is open, displaying a table of errors.

File	Line	Message
OEBPS/content.opf	5	The <keywords> element is not an allowed child of the <metadata> element.
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0003.xhtml	74	attribute 'name' is not declared for element 'a'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0004.xhtml	70	attribute 'name' is not declared for element 'a'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0004.xhtml	122	attribute 'border' is not declared for element 'img'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0004.xhtml	165	attribute 'compact' is not declared for element 'dl'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0005.xhtml	70	attribute 'name' is not declared for element 'a'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0006.xhtml	70	attribute 'name' is not declared for element 'a'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0006.xhtml	92	attribute 'border' is not declared for element 'img'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0006.xhtml	107	attribute 'border' is not declared for element 'img'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0006.xhtml	132	attribute 'border' is not declared for element 'img'
OEBPS/Text/freiesMagazin-2012-09-bilder-epub_0006.xhtml	167	attribute 'border' is not declared for element 'img'

Line: 165, Col: 3 100%

Vielen Dank für die Aufmerksamkeit!

Die Folien und Inhalte unterliegen (wenn nicht anders angegeben) der CreativeCommons „Namensnennung-Weitergabe unter gleichen Bedingungen 3.0 Unported“.



Copyright 2012 Dominik Wagenführ