

Katalog archivischer Dateiformate KaD E-Book Formate: Situationsanalyse und Perspektiven

1 Bekannte E-Book-Formate

E-Book-Formate zeichnen sich durch dynamischen Zeilen- und Seitenumbruch aus, sie sind nicht seiten- und layoutorientiert wie PDF oder XPS. Sie verhalten sich also grundsätzlich wie HTML-Seiten aus der Anfangszeit des Webs.

Zwei Formate können heute als marktführend im Sinne eigentlicher E-Book-Formate bezeichnet werden: EPUB und Kindle (siehe <http://www.E-Booknet.de/know-how/E-Book-formate> und https://en.wikipedia.org/wiki/Comparison_of_e-book_formats).

- **EPUB** ist ein offener Standard des *International Digital Publishing Forum*.
<http://idpf.org/epub>
- **Kindle Format (KF8)** ist ein proprietäres Format, das von Amazon zusammen mit dem Kindle Reader vermarktet wird. Heute kann es auch mit Apps gelesen werden.
<https://www.amazon.com/gp/feature.html?docId=1000729511>

Sowohl EPUB wie auch das Kindle-Format existieren in verschiedenen Versionen. Es gibt Programme zur Konvertierung zwischen beiden Formaten.

Daneben spielen noch folgende Formate eine Rolle:

- **DjVu** ist eher von untergeordneter Bedeutung, obwohl es eigentlich ebenfalls als E-Reader-Format, besonders im Bereich retrodigitalisierte Bücher, geeignet ist.
<http://www.djvu.org/>
- **iBook**, das iBook-Format von Apple, ist ein abgewandeltes EPUB-Format.
- **PDF** mit Adobe-DRM (Digital Rights Management) ist kein eigentliches E-Book-Format, wird aber häufig auf E-Readern gebraucht.
<http://www.lehmanns.ch/page/E-Bookadobedrm>

Open E-Book und andere ältere E-Book-Formate von Microsoft, Sony und haben heute keine Bedeutung mehr.

2 Digital Rights Management

Das Problem mit den E-Book-Formaten sind nicht so sehr die Formate oder die Formate-Vielfalt. EPUB, iBook und Kindle basieren auf HTML/HTML5 und CSS und können im Prinzip von jedem Browser gelesen werden. Es ist auch ohne grössere Schwierigkeiten möglich, diese Formate in *Plain HTML* umzuwandeln. PDF und DjVu fallen hier insofern aus dem Rahmen, als dass sie nicht eigentliche E-Book-Formate sind und nicht auf HTML basieren.

Digital Rights Management aber ist für das Archiv ein grundsätzliches Problem, beschränkt doch DRM die Benutzung in der Regel auf einen Anwender auf einem Gerät (oder maximal zwei bis drei Anwender auf ein paar Geräten). Zudem kann DRM auch die Benutzungszeit einschränken. Betroffen davon sind nicht nur E-Book-, sondern auch Audio- und Videodateien. Darum erlauben Archive in der Regel prinzipiell keine DRM-geschützten Dateien.

EPUB, iBook, PDF und Kindle implementieren das *Digital Rights Management* unterschiedlich, in der Regel als DRM-Container.

Das Entfernen eines DRM-Schutzes, wenn auch technisch möglich, ist rechtlich nicht statthaft. Das Archiv muss sich also bemühen, die Dateien ohne DRM-Schutz zu erhalten – ein Problem, mit dem sich auch die Bibliotheken herumschlagen müssen¹.

3 EPUB, iBook und Kindle

Ist das Archiv auf legalem Wege zu nicht DRM-geschützten EPUB-, iBook- oder Kindle-Dateien gelangt, stellt sich die Frage nach einer Konvertierung in HTML oder EPUB, da Kindle² und iBook proprietäre und nur partiell offene Formate sind. Die Konvertierung ist ohne Verlust möglich und es existieren etliche Tools dafür.

Ob EPUB zu HTML konvertiert werden soll, hängt davon ab, ob im Archiv die Anzahl der Formate reduziert wird und ob grosse EPUB-Kollektionen vorliegen, z.B. in Bibliotheken. Eine weitere Konvertierung von EPUB/HTML zu PDF ist ebenfalls denkbar und möglich. Dabei handelt es sich aber um eine echte Konvertierung und nicht nur um das Entfernen eines Containers. Bei dieser Konvertierung geht auch der Aspekt des dynamischen Seitenumbruchs bei E-Book-Formaten verloren.

4 PDF und DjVu

PDF und DjVu sind beides seiten- und layoutorientierte Formate. DjVu erreicht bei der Retrodigitalisierung von Büchern viel bessere Komprimierungsraten³, ist ein offenes Format, bewährt und einfach sowie PDF technisch überlegen. Es ist aber in der Archivwelt nicht sehr verbreitet. Aus diesem Grund ist eine Konvertierung in PDF allenfalls notwendig.

Stand: 03.04.2017

¹ Bei der wissenschaftlichen Literatur wird das Problem im Allgemeinen so gelöst, dass die Bibliotheken eine DRM-geschützte Version von den Wissenschaftsverlagen erhalten mit der Option auf eine ungeschützte Version zu einem späteren Zeitpunkt. Alternativ schliesst die Bibliothek mit einem eBook-Vertrieb eine entsprechende Vereinbarung ab, welche die eingeschränkte Verfügbarkeit garantiert. Siehe dazu <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-52746.html>.

² Das Kindle-Format ist von Amazon bereits 2010 offengelegt worden.

³ DjVu-Dateien können 3- bis 5-mal kleiner sein als die entsprechende PDF-Datei, siehe <https://en.wikipedia.org/wiki/DjVu>.