

5

EPUB 3 : 電子書籍フォーマット



村田 真 (国際大学 GLOCOM)



E バベルの塔の終わり

電子書籍についての狂乱と混迷が国内ではじめて久しい。税金も含めてかなりの投資がなされ、さまざまの動きはあるが、満足のいく結果が出ているとは言いがたい。しかし、電子書籍のフォーマットについては大きな進歩があった。

かつて、電子書籍のフォーマットは世界中で乱立していた。E バベルの塔だと揶揄されたほどである。Microsoft やアップルを含む多くの企業が独自のフォーマットを開発して市場に投入した。Microsoft のフォーマットである LIT については、無償の閲覧システムさえ配布された。しかし、LIT を含めほとんどのフォーマットは歴史の中に消えていった。徐々に注目されようになったフォーマットが EPUB である。2007 年に作られた EPUB 2 は、国際化が不十分である等の問題はあったが、ソニー、アップル、グーグル等に採用され、広く用いられた。

2011 年 10 月に制定された EPUB 3 は、縦書きやルビなどの日本語組版への対応を含めて国際化がなされた。これまでアジアでは EPUB 以外のフォーマット (XPDF, ドットブック, CEBX など) が有力であったが、EPUB 3 によって情勢ははっきり変わった。今後は、日本を含む全世界において、EPUB 3 が中心的な役割を担うと多くの人々が確信している。2012 年 7 月に国内で発売された kobo が EPUB を中心としていること、2012 年 9 月に日本電子書籍出版社協会が EPUB 制作ガイドを公開したことは、その端的な表れといえる。

自らの立場を明らかにするため、私がどのように EPUB にかかわってきたかを述べる。私は、

International Digital Publishing Forum (IDPF) EPUB WG のメンバとして、また国際化担当サブグループのリーダとして EPUB 3 の制定に深く関与した。また、EPUB 3 をデジュール標準として国際標準化すること (後述) にも、SC 34/AHG 4 (EPUB) の共同委員長として関与している。国際投票コメント審議を行う会議 (Ballot Resolution Meeting) の委員長に就任する可能性、技術仕様のプロジェクトエディタを共同で務める可能性もある。

本稿では、まず EPUB を簡単に解説する。次に、情報処理学会の記事であることを踏まえ、文書処理技術の系譜の中での位置づけ、JTC 1 における国際標準化について説明する。

EPUB 3 の概要

EPUB 3 は、HTML5 を筆頭とする Web 技術に基づく電子書籍フォーマットである。EPUB 3 を制定しているのは、IDPF という国際的なフォーラムである。EPUB 3 の仕様書は誰でも無料で入手できるし、EPUB 3 書籍を作成しても誰にもライセンス料を払う必要はない。

EPUB 3 にはいくつもの重要な特徴があるが、それらを単に列挙しても本質はかえって見えにくくなる。EPUB 3 で最も大事なものは、以下の 3 つの理念だと筆者は考える。

理念 1 : Web と電子書籍のシナジーを追求する

理念 2 : 世界中の言語・文化を扱う

理念 3 : アクセシビリティのための特殊なものを作るのではなく、普通のをアクセシブルにする

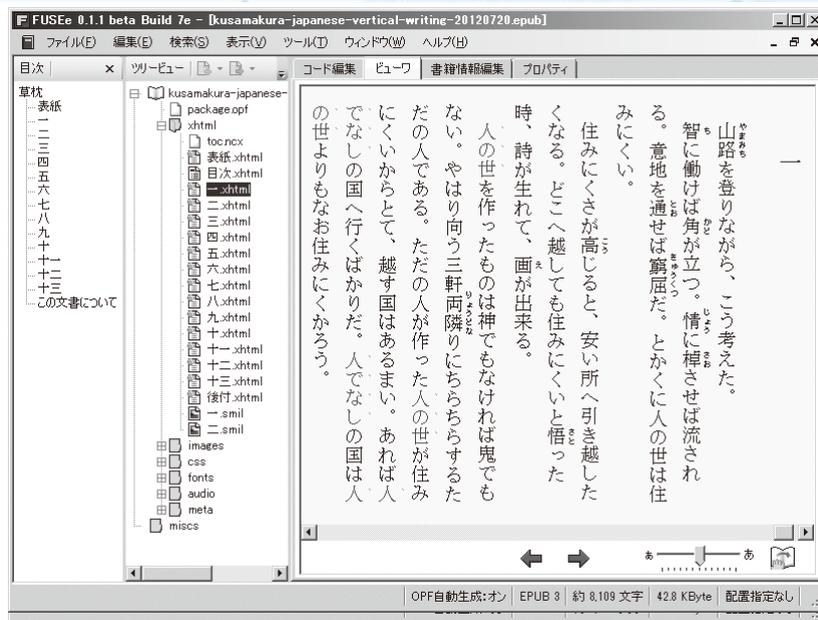


図-1 EPUB エディタ FUSEe β の画面

これらが IDPF の公式文書に理念として記されているわけではないし、これらとは関係ない機能（たとえば数式）も多い。しかし、EPUB 3 はこれらの理念をもとに設計されたと思うし、EPUB 3 がほかのフォーマット（たとえば、ドットブックや XDMF）とは一線を画す理由もここにあると信じる。

EPUB 3 は、リフローと固定レイアウトの両方を扱う。リフローとは、閲覧に用いる機器・ソフトウェアおよびユーザの指定によってページ数が変わることである。逆に言えば、閲覧環境とユーザにふさわしいページレイアウトが自動的に行われる。一方、固定レイアウトとは、どんな閲覧環境であってもページ数が変わらないことである。リフローは文芸書などに、固定レイアウトは絵本やマンガなどに適している。

EPUB 3 書籍は、1つの ZIP ファイルとして表現される。この ZIP ファイル中に HTML 文書などが格納される。図-1 に、EPUB 3 で表現した『草枕』（夏目漱石）を電子書籍オーサリングソフトウェア FUSEe β で表示したところを示す。

- 左側のサブウィンドウに示されているのは目次である。

- 真ん中のサブウィンドウに示されているのは ZIP ファイルの中のフォルダ階層である。
- 右側のウィンドウに示されているのは HTML 文書（一.html）をレイアウトした結果である。HTML5 文書から参照されている CSS スタイルシートで縦書きを指定しているため、縦書きレイアウトになっている。
- styles ディレクトリには CSS スタイルシートが格納されており、images ディレクトリには画像が格納されている。
- 真ん中のサブウィンドウの上部に表示されている package.opf ファイルは、パッケージ文書である。これは、HTML 文書や CSS スタイルシートや画像などをまとめ、読む順番を指定するなどの役割を持つ。

EPUB の詳細については本稿では説明しない。国内で良書（たとえば文献2）と文献3）が発刊されているので、そちらを参照されたい。EPUB 3 の制定、特に国際化の部分の制定については、文献1）を参照されたい。

文書処理技術の中での位置づけ

現在、最も広く使われている文書処理技術は、オフィス製品（Microsoft Office 等）と Web の 2 つであろう。この 2 つと EPUB の関係を考えてみよう。EPUB と Web との親和性が強調されることは多い。しかし、EPUB には、オフィス製品のためのフォーマットである OOXML および ODF とも共通点がある。

■ Web

1980 年代末に始まった Web は、世界中に普及し、世の中を大きく変えた。いまやほとんどの携帯電話が Web に対応しており、テレビなどの家電製品にも入り込もうとしている。

Web はもともと HTML, HTTP, URI から出発したが、画像やスタイルシート（CSS）がまず追加され、いまでは音声、動画、数式、ベクター図形（SVG）も追加されている。とくに重要な HTML は、いまでは HTML5 にまで進化した。

EPUB は、ほとんどの要素技術を Web から借りている（これを体よく言い直したものが先に述べた理念 1 である）。EPUB

では、テキスト部分は HTML 文書によって表現しており、レイアウトを制御するのは CSS スタイルシートである。音声、動画、数式、ベクター図形も含め、ほとんどの要素技術はもともと Web で用いられてきた。EPUB で独自に追加された機構は、パッケージ文書程度に過ぎない。

EPUB が Web 技術に基づいていることは、2 つの大きな利点をもたらす。1 つは、Web コンテンツを電子書籍にすること（もしくはその逆）が容易な点である。Web で公開した内容を電子書籍にすると、通常は HTML 文書を全面的に作り直す必要はない。もう 1 つは、Web のために作成されたツール

を、電子書籍のために利用できることである。Web ページを作るための編集ツールは多々あるが、それらを電子書籍の作製に流用することが可能になる。

しかし、EPUB と Web とでは一点において大きく異なる。それは、パッケージ化である（後述）。

■ オフィス製品

今日のオフィス製品は、OOXML か ODF の両方または少なくとも一方のフォーマットを実装している。OOXML は、Microsoft Office における過去の蓄積のもとに作られたフォーマットであり、ODF は、OpenOffice.org のフォーマットから出発したものである。OOXML と ODF のどちらも、オフィス文書を電子的に表現するためのフォーマットである。

EPUB は、OOXML や ODF と比べられることはほとんどないが、重要な共通点が 1 つある。前述したパッケージ化である。

OOXML も ODF も、ZIP でいくつかのファイルをパッケージ化したものである。ファイルの多くは XML であるが、イメージなども含まれる。XML ファイルに使われるタグについては、OOXML と ODF の仕様書

がそれぞれ規定している。EPUB 書籍も ZIP によって HTML 文書などのファイルをパッケージ化したものである。

パッケージ化によって、商業流通・保存・版数管理を簡単に行うことができる。一方、Web ではこれらは容易ではない。Web にはない EPUB の利点といえる。

デジュール標準としての国際標準化

EPUB を制定している IDPF は、国際的な標準化団体としてはまだまだ実績に乏しい。EPUB 2 は欧



米にすでに普及し、EPUB 3はアジアに急速に普及しつつあるが、さらに国際的な普及を加速するために、ISO、IEC、JTC 1での標準化が検討されている。なお、韓国はすでにEPUB 3を国家規格として制定している。

検討を行ってきたのは、ISO/IEC JTC 1/SC 34/AHG 4 (EPUB)であり、ここにはSC 34とIDPFのほか、ISO TC 46とIEC TC 100が加わっている。筆者は、このAHG 4の委員長を韓国のYong-Sang Cho博士と共同で務めている。

JTC 1/SC 34：文書の記述と処理の言語

ISO TC 46：情報とドキュメンテーション

IEC TC 100：オーディオ、ビデオ、マルチメディアシステムおよび機器

AHG 4の審議の結果、以下のことが決まっている。

- 国際規格ではなく技術仕様 (Technical Specification) としていったん制定し、数年後に国際規格化を検討する。これはEPUB 3が参照しているHTML5などがまだ確定していないためである。
- 韓国から迅速化手続きによってJTC 1に提出し国際投票にかける。コメント審議に基づく改定を行

って、技術仕様として発行する。

- SC 34, TC 46, TC 100によるJoint Working Groupを設立し、そこで技術仕様のメンテナンスを行う。
- IDPFのEPUBとの技術的一致を保証し、無償で入手可能とする。

技術仕様としての成立時期はまだ確定していないが、2013年の前半には成立するものと予測している。これによって、公教育などの公的な用途へのEPUBの採用が加速するものと思われる。ISO TC 46によって図書館でのEPUB利用に関するプロファイルが制定される可能性、JTC 1/SC 36 (学習、教育、訓練)によって教育へのEPUB利用についての国際規格ができる可能性もある。

参考文献

- 1) 村田 真：電子書籍フォーマットEPUBと日本語組版 日本でメインストリームにいる人間は国際標準化の舞台ではまず勝てない、情報管理, Vol.55, No.1 pp.13-20 (2012), <http://dx.doi.org/10.1241/johokanri.55.13>
- 2) 林 拓也：EPUB 3 電子書籍制作の教科書, 技術評論社 (2012).
- 3) 境 祐司ほか：EPUB 3 スタンダード・デザインガイド, マイナビ 2012.

(2012年9月25日受付)

● 村田 真 eb2m-mrt@asahi-net.or.jp

1960年生まれ。1982年京都大学理学部卒業。2006年筑波大学博士(工学)。構造化文書の研究、XMLと電子文書の国際標準化(W3C, IDPF, ISO/IEC JTC 1/SC 34など)に従事する。

