

専門的書籍の  
新しい制作・流通のケーススタディとして  
『PDFインフラストラクチャ解説』  
学びと課題

2016年2月16日  
アンテナハウス株式会社  
小林 徳滋

# はじめに

- 『PDFインフラストラクチャ解説』は:
- プリントオンデマンド(POD)方式とKDPによるセルフ出版の事例であり、
- ワンソースマルチユース方式制作の実践事例でもある。
- この事例から学んだことや、今後の課題を報告する。

# 本書の性格

- 紙に文字を書くから、デジタルのPDFに文字を出すことへの変化について解説
- PDFに関する教養書、専門技術書ではない
- 図版が多い 150点以上
  - インライン配置のイメージ
  - ブロック配置の図
- 表が多い
  - HTMLのtable要素でマークアップ
- 横組
- 参考文献、索引

# 執筆経過

- 2005年10月～2008年7月まで「PDF千夜一夜」ブログ
- 2011年末～2012年1月 アウトラインを決めて記事を整理
- 2012年1月EPUB版を初公開ダウンロード配布
- 2013年2月からEPUB版、PDF版を同時配布
- 2015年12月の0.55版まで随時更新
  - 無償ダウンロードは2016年1月15日に終了しました

# 内容の見直し

- 執筆から出版まで10年を経過している部分を含む
- PDF技術ベースは20世紀に完成している
  - フォント技術
  - オブジェクト表現(シンタックス)
  - ファイル構造
- 基本部分は変わっていないが、周辺環境が変わった
  - 仕様の扱い 2008年にISO 32000-1となった
  - OS、ブラウザ、Officeアプリが標準サポート

# 推敲・仕上げ 検索 & 置換

- 2015年11月～12月
- 用語・書記方法の統一
  - 漢数字とアラビア数字
  - 音引き
  - 「・」の使用。2語は原則として中黒無し
  - 箇条書きの末尾に読点を付けるか
  - カタカナ語と和文語
  - 全角・半角(括弧、数字、ラテンアルファベット)
  - 空きスペース
  - 送り仮名
- プロジェクト管理システムで用語管理

例)  
中黒つかうかどうか

(1)2語は、原則として中黒無し

- アウトライン・フォント⇒アウトラインフォント
- デジタル・フォント⇒デジタルフォント
- フォント・ファミリー⇒フォントファミリー
- フォント・ファイル⇒フォントファイル
- フォント・オブジェクト⇒フォントオブジェクト
- アウトライン・ツリー ⇒アウトラインツリー
- メタデータ・ストリーム ⇒メタデータストリーム
- ストリーム・オブジェクト ⇒ストリームオブジェクト
- バイナリ・ファイル⇒バイナリファイル
- バイナリ・データ
- グリフ・データ⇒グリフデータ
- リソース・ファイル⇒リソースファイル
- フォーム・コントロール ⇒フォームコントロール
- スクリーン・リーダー ⇒ スクリーンリーダー
- スクリーン・リーダ⇒スクリーンリーダー
- セキュリティ・ハンドラー⇒セキュリティハンドラー
- プリンター・ドライバー⇒プリンタードライバー

# 索引

- 索引語を拾う
- 約750箇所・600語
- 索引のマークアップ
- 索引頁は自動生成

EPSのファイルは、PostScriptによるページ記述命令と、そのページの[[[:index プレビューイメージ]]]の両方を1パッケージ化できる。[[[:mindex [[[:prim:nodisp EPS]]] [[[:second プレビューイメージ]]]]]]はPostScript命令を解読してレンダリングする機能をもたないソフトでも、EPSファイルの内容を表示できるようにするためのもの。EPSファイルを作成する時に、オリジナル作成ソフトが用意して埋め込むのが普通である。

# POD版の試作とレイアウト改良

## • 0.24版 (初回POD)

## 最終版

表 8 Adobe-CM10

範囲	CID範囲	Adobe-CM10	説明
範囲 2 (14288-14289) (15711)	Adobe-CM10 CID:10	Adobe-CM10	1992年11月10日現在、Unicode 3.0.0の255個の漢字が収録されている。

表 9 Adobe-CM11

範囲	CID範囲	Adobe-CM11	説明
範囲 2 (14288-14289) (15711)	Adobe-CM11 CID:11	Adobe-CM11	1992年11月10日現在、Unicode 3.0.0の255個の漢字が収録されている。

表 10 Adobe-CM12

範囲	CID範囲	Adobe-CM12	説明
範囲 2 (14288-14289) (15711)	Adobe-CM12 CID:12	Adobe-CM12	1992年11月10日現在、Unicode 3.0.0の255個の漢字が収録されている。

表 11 Adobe-CM13

範囲	CID範囲	Adobe-CM13	説明
範囲 2 (14288-14289) (15711)	Adobe-CM13 CID:13	Adobe-CM13	1992年11月10日現在、Unicode 3.0.0の255個の漢字が収録されている。

表 12 Adobe-CM14

範囲	CID範囲	Adobe-CM14	説明
範囲 2 (14288-14289) (15711)	Adobe-CM14 CID:14	Adobe-CM14	1992年11月10日現在、Unicode 3.0.0の255個の漢字が収録されている。

表 13 Adobe-CM15

範囲	CID範囲	Adobe-CM15	説明
範囲 2 (14288-14289) (15711)	Adobe-CM15 CID:15	Adobe-CM15	1992年11月10日現在、Unicode 3.0.0の255個の漢字が収録されている。

表 14 Adobe-CM16

範囲	CID範囲	Adobe-CM16	説明
範囲 2 (14288-14289) (15711)	Adobe-CM16 CID:16	Adobe-CM16	1992年11月10日現在、Unicode 3.0.0の255個の漢字が収録されている。

## • 表のスタイル変更など 左: レポート形式、右: JIS形式



# POD版の販売

- アマゾンe託方式とPOD取次方式の比較

取り扱い	POD 製作	取り扱い費用	在庫負担
e 託	街の業者	年会費 9,000 円 定価の 40%	在庫負担あり。アマゾンの倉庫に納品するコストもかかる
POD 取次	アマゾンが 印刷製本	定価の 38%	PDF で取次に渡すので、在庫負担と納品コスト負担なし

# アマゾンPODのためのPDF

- 頁数： 24 ~ 746
- 判型： 文庫はできない。新書 ~ A4 (高さ269mm)
- 周囲に3mmの裁ち落とし余白 (塗り足し)
- PDFバージョンは1.3 (透明が使えない)
- CAS-UBではPDF/X-1a:2001とする

# オンライン配布PDFとの相違

- 裁ち落とし余白有無
- PDF/X-1aでは、印刷範囲には注釈を置けない
- POD版はリンク(PDFでは注釈のため)を設定できない
  - 目次 本文見出し
  - 索引語 本文の索引語
  - 内部参照にリンクができない
    - 見出し参照 (詳しくは\*\*\*を参照など)
    - 図参照
    - 表参照
    - 引用 参考文献

# PDF:POD版設定メニュー

- パラメータをチェックしてPOD用に変更する

## PDF: POD版設定

設定 1 ▼ 読み込み

設定 1 ▼ に 設定名 設定 1 で保存します

POD(プリント・オン・デマンド)の規格に合わない値は赤字で表示されます

確認・保存

設定項目	値	推奨値	説明
作成するPDFのバージョン	PDF1.7	PDF/X-1a:2001 ▼	PDF1.3 CAS-UBではPDF/X-1a:2001を指定するとPDF1.3となります
トンボ	トンボを出力しない	塗り足しのみ ▼	断ち切りは3mmを周囲につけ、トンボはつけない CAS-UBでは塗り足しのみを指定します
表紙	生成する	生成しない ▼	表紙を出力しない。別途包みの表紙(PDF)を用意してください
用紙	B6判 JIS (128 x 182 mm)	B6判 JIS (128 x 182 mm) ▼	横が102mm~215mm、縦が165mm~296mm
基本版面の指定方法	マージン指定	マージン指定 ▼	
上下マージン	24.5mm	24.5mm	ハンブルや柱、ツメなどを除き6.35mm以上にします
左右マージン	17.96mm	17.96mm	小口余白は6.35mm以上にします
左右位置調整	0mm	0mm	基本版面を小口側に移動して調整します
予想ページ数	24 - 150	24 - 150 ▼	ド余白はページ数により以下の値以上にします 24~150 ページ 9.53mm 151~300 ページ 12.7mm 301~500 ページ 15.88mm 501~700 ページ 19.05mm

# 販売開始

- POD 1月6日納品 1月21日発売
- KDP 1月14日登録 1月16日発売



# 紙と電子 (EPUB) で本を作る

- 世間一般のワークフロー
  1. DTPで制作完了
  2. DTPデータやPDFを入手
  3. テキスト、画像など素材取り出し
  4. EPUBを制作する
- つまり: 2回の制作作業 & 2回の校正作業
  - 時間の無駄
  - 作業工数の無駄

# ワンソースマルチユースのねらい

- 制作という労働集約的部分を1回で済ませる
- 出版までの時間を短縮
- コストの削減
- 制作者依存からの脱却
  - 外注管理不要
  - 品質の一定化

# ワンソースマルチユース 実現の課題

- ソース文書を制作・編集する仕組み作り
- マルチ出力の仕組み作り
  - 出力レイアウトをコンピュータ支援で最適化する方法
  - PDFの自動生成組版エンジンとスタイルシート開発
    - POD用PDF
    - オンライン配布PDF
    - アクセシブルなPDF/UA(将来)
  - EPUBの自動生成
  - その他





# ソース文書は： XMLが最適

- コンテンツをXML形式で用意すると何が良いか？
  - XMLタグ(要素と属性)は文書をプログラムで処理するためのもの
  - トランスフォーメーション(変換)が自在になる
- 最近、HTML5 + JavaScriptで自動処理という対案も出てきた
  - HTML5はメタ言語的に使える
  - これで、効率的にできるかどうかは、まだわからない

編集

ソース  
文書

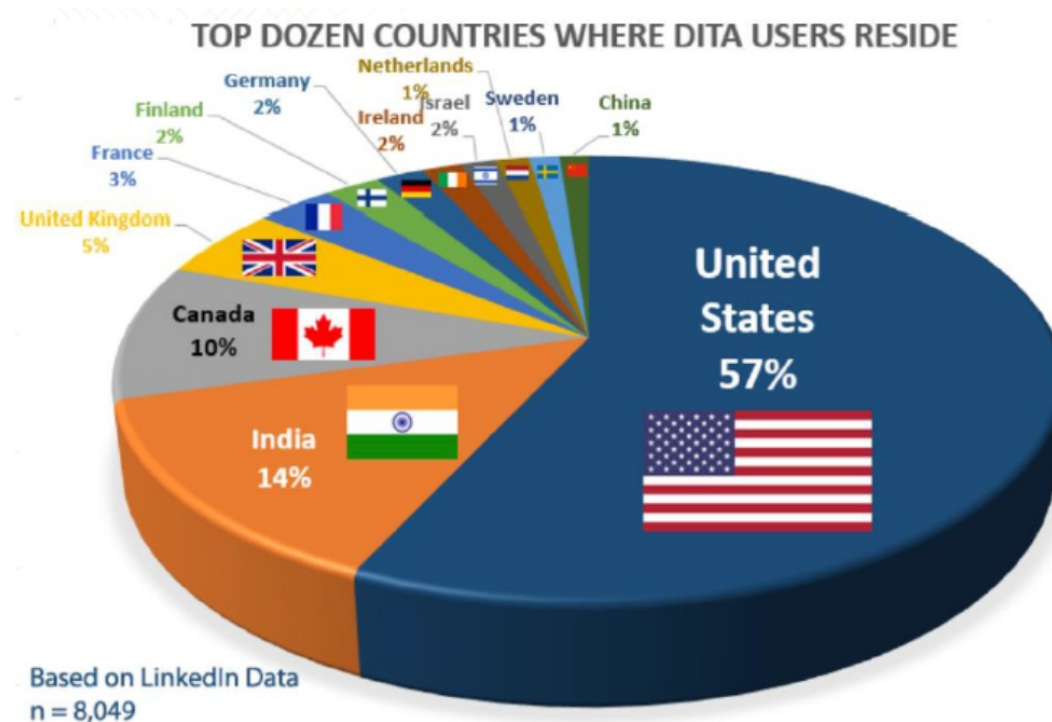
出力

XML文書 = トランスフォーマー

# XML制作をどうするか？ (大きな課題)

- XMLエディタを使うのは本格派
- 利用者は、XMLの基礎知識、ツールの用意、マークアップの訓練が必要である
- DITAの場合：  
    オフショア？

(Robert, Keith.  
@DITA festa 2015)



# 日本語本の制作を： 途上国へのオフショアは？

- 中国はコスト高
- 海外では日本語が通じない
- 管理のためのコミュニケーションのオーバーヘッドが大きすぎる



ソース  
文書

出力

# 誰でも簡単にXML文書 を書ける方法はないだろうか？

- 簡易マークアップ: CAS記法を開発 (2010 ~ )

<p><u>ラベル無し簡条書き</u></p>	<p>番号なし簡条の最初の項目を*::nolabelとすると簡条書きにラベルを付けない。</p>	<p>*::nolabel 一、天下の政権を朝廷に奉還せしめ、政令宜しく朝廷より出づべき事。 *一、上下議政局を設け、議員を置きて万機を参賛せしめ、万機宜しく公議に決すべき事。</p>	<pre>&lt;ul class="nolabel"&gt; &lt;li&gt;一、天下の政権を朝廷に奉還せしめ、政令宜しく朝廷より出づべき事。&lt;/li&gt; &lt;li&gt;一、上下議政局を設け、議員を置きて万機を参賛せしめ、万機宜しく公議に決すべき事。&lt;/li&gt; &lt;/ul&gt;</pre>
<p><u>用語定義リスト【B】</u></p>	<p>用語とその定義を対で記述する簡条書き。 アスタリスク '*' の後のハイフン2つ '--' に挟まれた語句が用語、その後に用語の説明を書く。</p>	<p>*-- 用語 1 --用語 1 についての説明 *-- 用語 2 --用語 2 についての説明</p>	<pre>&lt;dl&gt; &lt;dt&gt;用語1&lt;/dt&gt; &lt;dd&gt;用語1についての説明&lt;/dd&gt; &lt;dt&gt;用語2&lt;/dt&gt; &lt;dd&gt;用語2についての説明&lt;/dd&gt; &lt;/dl&gt;</pre>
<p><u>ブロック範囲指定【B】</u></p>	<p>3つの開き大括弧('[[[')と3つの閉じ大括弧(']]]')で囲む。 開始マークアップが行頭から始まる。 終了マークアップは行頭から始まる。 開始マークアップに続き、属性と、'='を置くことができる。</p>	<pre>[[[:value =caption ブロック範囲指定 ]]]</pre>	<pre>&lt;div class="value" &gt; &lt;div class="caption"&gt;caption&lt;/div&gt; &lt;p&gt;ブロック範囲指定&lt;/p&gt; &lt;/div&gt;</pre>

<http://www.cas-ub.com/howto/reference/index.html>

# CAS記法とマークダウン

- 最近流行りのマークダウン、しかし機能が弱すぎ
  - 簡単なものしかできない
  - 属性の記述ができない
  - タグのネストが弱い
- CAS記法はひと味違います
  - タグのネストある程度OK
  - 属性を自由に書ける
  - 埋め込みも簡単
    - XHTML表
    - MathML数式

# 本には構造がある (Webとは違う)

- 前付
  - 半扉・扉
  - 前書
  - 目次
  - 謝辞
- 本文
  - 章・節・項(階層構造)
- 後付
  - 参考文献
  - 索引
  - 奥付

# 構成を編集する

- 通常制作では
  - 目次と台割り？
- CAS - UBでは構成編集機能

## ヒント

### 構成編集

#### 目録 出版物 PDFインフラストラクチャ解説

- 1 はじめに (201112290806) [前書]
- 2 PDFってどんなもの？ (201112250825) [章]
  - 2.1 PDFとは (201112290805) [節]
  - 2.2 PDFの歴史 (201112250952) [節]
- 3 PDFの国際標準 (201112250924) [章]
  - 3.1 PDFの国際標準仕様 (201302111036) [節]
  - 3.2 ISO 32000の制定と動向 (201112250952a) [節]
  - 3.3 PDFをベースとするプロファイル仕様 (201201020954) [節]
- 4 PDFを作成する方法 (201112250835) [章]
  - 4.1 ボーンデジタルPDFとターンドデジタルPDF (201112292325) [節]
  - 4.2 PostScriptからPDFを作る (201112292326) [節]
  - 4.3 PDF出カライブラリー (201112292330) [節]
  - 4.4 Windowsの印刷機能でPDF作成 (201112292331) [節]
  - 4.5 仮想PDFドライバーで対話機能を設定する (201112292331a) [節]
  - 4.6 直接PDF作成 (201112292332) [節]
  - 4.7 いくつかの技術的課題 (201112292332b) [節]
  - 4.8 スキャナー・複合機によるPDF作成 (201112292332a) [節]
- 5 PDFのナビゲーション機能 (201112310814) [章]
  - 5.1 ナビゲーション機能とは (201112310834) [節]

# ~~DTP (WYSIWYG) 批判~~

- WYSIWYGは、対話的に、画面上で印刷 = 紙への出力を作る
  - 紙の世界の延長にある
- InDesign
  - 機能が高く、なんでもできる
  - ソフトの利用価格も安い
- 現代は草の根DTPの時代
  - 本の制作者は玉石混交
  - おかしなレイアウトの本が店頭一杯ある
    - 極端に余白が狭い
    - ノンブルや柱の位置がおかしい
    - 余計な飾り(読むのに集中できない)

編集

ソース  
文書





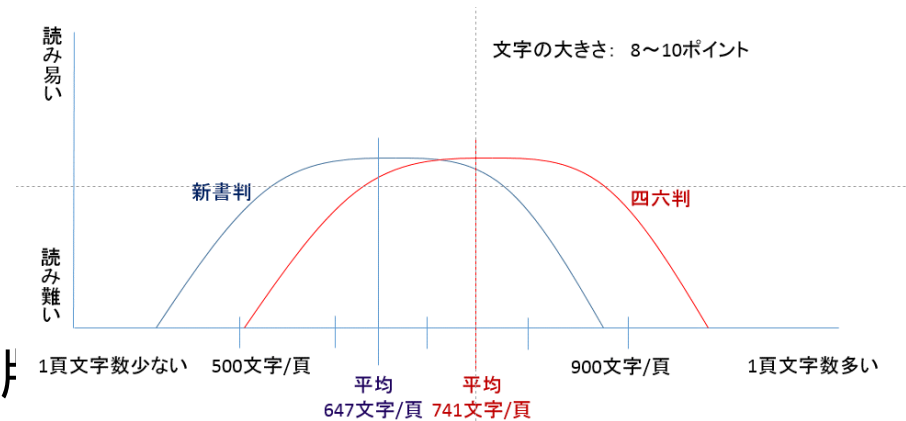
# 紙の本を自動的に作る時： (配慮すべきパラメータ)

- 基本版面(日本語)のパラメータ

- 判型
- 余白
- 文字サイズ
- 一行の文字数
- 行数
- 行間

- 本作りの教科書に載っている基本版面  
(ほとんど使われていない)

- 基本版面のアウトラインフォント向けの見直し



<http://blog.cas-ub.com/?p=9140>

# 記事別スタイル

- 例) 参考文献の書き方とレイアウトのスタイル
- 欧米には参考文献の書き方のスタイル教科書が膨大にある
  - MLA Handbook (7版) は、本の半分が引用と参考文献の書き方
  - CSL (Citation Style Language) では1,000種以上のスタイルがある
- 日本では比較的少ない
  - SIST (科学技術情報流通技術基準) にある



# 扉と改丁

- 紙の本には扉が多い。扉で改丁、本の大きな区切りとなる。
- 扉の種類
  - 半扉(英語の本)
  - 化粧扉(本文と別紙)または本扉(本文と同じ紙)
  - 目次扉(縦組のみ)
  - 書名扉(半扉)
  - 章扉
- 日本語の本は扉を置くか、その位置が勝手・きまま
- 英語の本は歴史的に扉の位置が明確

# ノンブル

- ノンブルのカウント開始位置
- ノンブルの区切り(カウント・リセット)方法
  - 前付・本文・後付全部通し
  - 前付はノンブルなし、本文と後付で通し
  - 縦組中の横組頁のノンブル
    - 順方法(右から左へ)
    - 逆方法(左から右へ:横組と同じ)
    - 両方(二つのノンブル)
- 印字位置(上下、左右)
- 基本版面からの距離

# 改頁と組版オブジェクト 配置の最適化

- 紙では、基本版面上に、テキスト・図表を配置する。
- 頁内に大きな空きができてしまう：
  - 段落の後に図が入らないので図の前で改頁する
  - 見出しの前で改頁
    - 見出しと本文の間を泣き別れしないため
  - 図表の前で改頁
    - 図表のキャプションと図表本体の間で改頁を許さないため
    - 表の途中で改頁を許さないため
- **ウイドウ** (頁先頭の孤立行)、**オーファン** (頁末尾の孤立行) **を避けたい**

# 図表の大きさと配置戦略

- 図の大きさ決定
- 図の配置戦略
  - なりゆき
  - 基本版面の上部に置くか下部に置く(上下フロート)
  - 基本版面の左右に置く(左右フロート)
  - 基本版面の小口側に置く(フロート)
  - 基本版面の上小口に置く
- 表の方向
  - 縦組なのに表は横組

# 白紙頁を挿入する？

- 扉裏は白紙か、それとも文章を入れる
- 改頁か、改丁か
  - 改丁する位置
  - 改頁する位置

# オブジェクトの最適配置を計算で割り出せないか？

- 配置戦略・アルゴリズムの開発
  - レイアウト最適化戦略
  - 戦略別に配置の良さをパラメータで表して最適化する
- 最適値をどうやって測定するか？
- 頁レイアウトの標準化
  - 良いレイアウトと良くないレイアウトをスコア化する必要がある



Power  
Up!

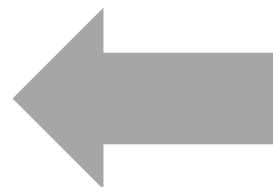


# 今のEPUBではワンソースマルチユース実現が難しい

- 例) 多くのEPUBリーダーのレベルが低いため、改頁禁止の制御が使えない。
- EPUB制作では、一般的に：
  - 図とキャプションをまとめて画像化する
  - 表を画像化する
- 本文と図のキャプションや表の文字サイズが不調和になるため、印刷では使えない
  - ワンソース化できない

# 紙とEPUBでは対応しない 概念が多い

- 扉裏(空白頁)は不要
- 改丁の概念はない
- ノンブルはない
- 柱がない

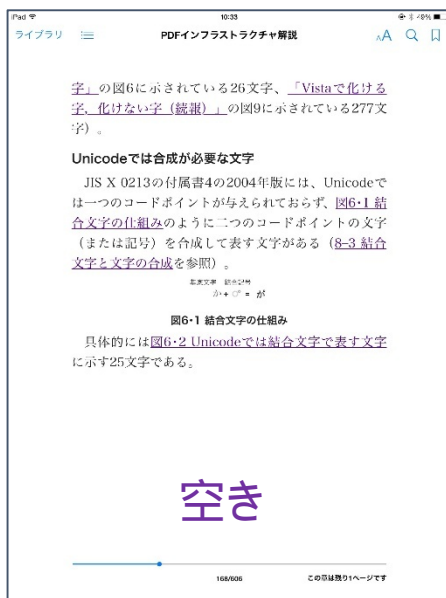


みんな頁  
の関係

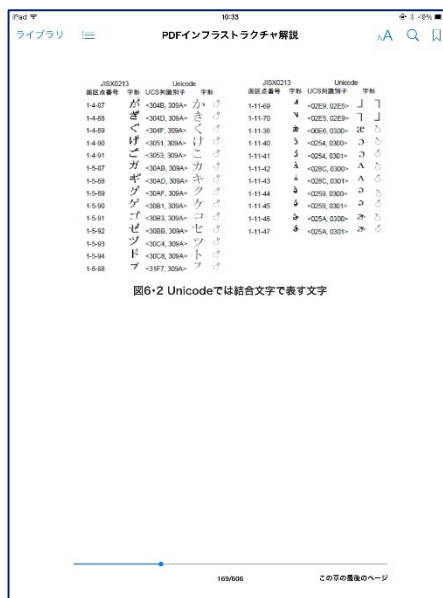
# 電子書籍は頁の概念を見直しすべき

- 画面を頁に見立てる意味はないのでは？
- 巻物方式の方が良いのでは？

## 図への参照



## 図



## 巻物方式なら

