

PDF/Aの概要

2007年9月21日



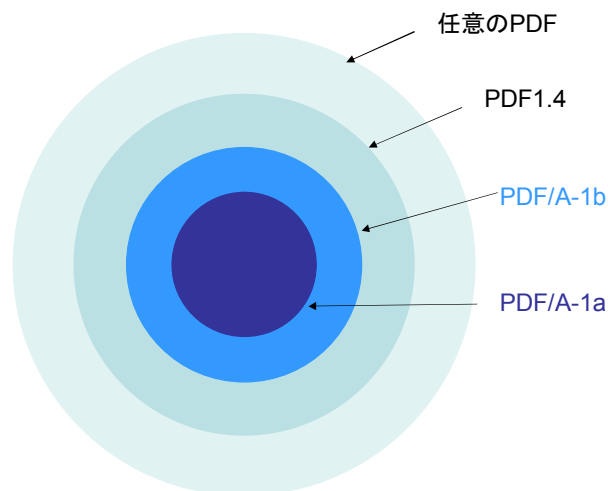
アウトライン

- PDF/Aとは
- PDF/Aの目標
- 仕様の概要
- 要求項目、禁止項目、制限事項
- タグ付きPDFとPDF/A-1a
- 製品動向、Acrobatのサポートなど
- PDF/Aの生成
- PDF/A-2
- デモ

PDF/A

- PDF/A : ISO 19005の同義語
 - PDF/A-1 : ISO 19005-1の同義語
 - PDF/A-1a : ISO 19005-1 レベルA適合
 - PDF/A-1b : ISO 19005-1 レベルB適合
-
- ISO 19005-1 Document Management -
Electronic document file format for long
term preservation - Part 1: Use of PDF 1.4
(2005年9月出版)

PDF、PDF/Aの関係



PDF/A-1a ⊂ PDF/A1-b ⊂ PDF/1.4 ⊂ PDF

PDF/Aの目標

- 電子文書の見かけを、作成・蓄積・可視化ツールから独立にし、時間が経過しても維持できる。
- 電子文書の文脈と履歴をメタデータとして記録するフレームワークを提供する。
 - XMPメタデータの埋め込み
- 電子文書の論理的な構造と意味に関する情報を記録するフレームワークを提供する。
 - タグ付きPDFを利用する

仕様の概要

- PDFの仕様(PDF Reference 1.4)で規定しているPDFのオペレータとオブジェクトに対し次の5つの制約を課す。
 - 要求(Required): 必須である
 - 禁止: 準拠PDFに出現してはならない
 - 制限: 使用方法、内容、値に制限がある
 - 勧告: 準拠ファイルに現れるべき
 - 無視: 準拠ファイルに現れても無視される
- ※非常に細かい規定がなされている

主な要求項目

- カラーの再現性を保証
 - デバイス独立カラーまたはPDF/A-1 OutputIntent指定でカラー特性を指定する
- フォント埋め込み
 - 基本14フォントを含む全てのフォントの埋め込み
 - 使用している全グリフがあれば、部分埋め込みでも良い。
 - 埋め込むフォントは汎用可視化用途での埋め込み許可が必要
 - リーダは埋め込みフォントで表示する
- XMPメタデータの埋め込み

主な禁止項目

- 暗号化
 - パスワードによるアクセス許可はできない
 - LZW圧縮
 - 文書の代替可視化
 - 埋め込みファイル(ファイルを添付操作)
 - PostScriptコード
 - 外部コンテンツへの参照など外部依存性を排除
 - 透明
- ※外部リソースへの依存性を排除する

主な制限事項

- 注釈の制限
 - 隠し注釈、印刷不可の注釈
 - FileAttachment(添付ファイル注釈)
 - Sound、Movie(マルチメディア)
- アクションの制限
 - Launch, Sound, Movie, ResetForm, ImportData, JavaScriptアクションは禁止。
 - NextPage, PrevPage, FirstPage, LastPage以外の名前付きアクションは禁止。
 - 対話フォームからアクションを実行禁止。

※外観を変更したり、外部からファイルをインポートするのは禁止ということ。

主な無視事項

- リニアライズ(Web表示用に最適化)
 - リニアライズされていても良いが、リニアライズ用のオブジェクトは無視する。

PDF/A-1bの要求(まとめ)

- カラースペースのデバイス独立
- フォントは合法的に埋め込む
- 標準的メタデータ
- 完全性(外部ファイル等非依存)
- PostScript禁止、添付ファイル禁止などPDF以外を含まない
- オーディオ、ビデオ禁止
- JavaScript、実行ファイルの起動禁止
- 暗号を許可しない

タグ付きPDF

- PDFは紙のようなページをもつメディアに文字列などを表示開始座標を使って、可視化することを目的とするため、文字列が見かけ通りに繋がっていることも保証されない。
 - テキストが千切れてしまい、音声読み上げなどで困る。
- タグ付きPDF
 - PDFに論理構造を付加するためにPDF 1.4で導入された仕様。
 - 米国では、アクセシブルPDFのためなどに用いられる。
 - PDF/A-1aではタグ付きPDFであることを要求する。

PDF/A-1a

- PDF/A-1bに追加要求(2項目)
 - タグ付きPDFで論理構造を埋め込む
 - ToUnicodeCMapでUnicodeへの対応
(一部フォントを除く)
- 自動作成は困難？
 - タグ付きPDFは論理構造なので、PDFにする前の文書に構造が定義されてないと、自動的に作成できない。
 - 6.8項では、もともとないものを自動的に作ってはならないとされている。もし後で付加するなら対話方式にする。

PDF/Aの製品動向

- 2005年遅くにPDF/Aの仕様が決まってから、対応製品が登場してきた。
- PDF/A対応製品のメーカーの多くは欧州の会社であり、欧州が熱心なように見える。USよりも欧州が正式なISO仕様準拠への要求が強いようだ(下記、米国議会図書館資料)。
 - <http://www.digitalpreservation.gov/formats/fdd/fdd000125.shtml>
- PDF/A Competence Center(ドイツ)
 - <http://www.pdfa.org/>

AcrobatのPDF/Aサポート

- Acrobat 7 ProfessionalでDraft をサポート
 - 7.07 PDF/A-1b(正式仕様)の作成が可能
- Acrobat 8 Professionalで標準をサポート
 - プリフライト機能(PDF/A-1a、PDF/A-1b準拠度)
 - PDF/Aへの変換
 - PDFを全てPDF/Aに変換できるわけではない。
 - PDF DriverでPDF/A-1bを生成可能
 - PDF/A固有情報の削除

その他のツール

2006年～2007年に発売されたPDF/A対応製品

- Microsoft Office 2007
 - Office 2007のアドインPDF変換(SaveAsPDF)は、PDF/A保存をサポートしている。
- PDF Tools AG(スイス)
 - PDF/A対応製品に力を入れており、様々なPDF/Aの生成、検証、表示、変換ツールを揃えている。
- PDFLib(ドイツ)
 - サーバサイドのPDF生成ツールPDFLib7.0でPDF/A-1a、PDF/A-1bの作成
- Nuance(米国)
 - PDF Edit Professional: PDF/Aの生成が可能 (PDF Driver) ㊦
- XSL-FOからPDF
 - アンテナハウス(日本)のXSL Formatter 4.2: PDF/A-1a、PDF/A-1bの作成が可能。

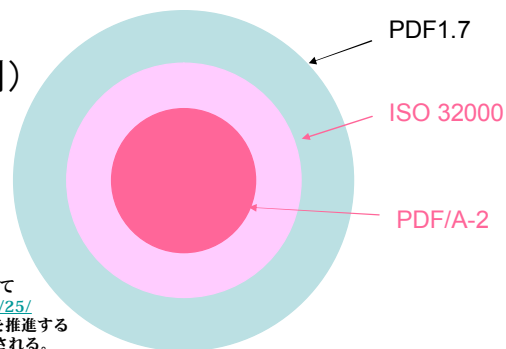
上のリストはすべてをカバーしているわけではありません。

PDFからPDF/Aへの変換

- PDFをAcrobat 8.1のプリフライト機能でPDF/Aに変換するとエラーの山になる
 - PDF/Aを自動変換で作成するのは難しい
 - ソース文書からPDF/A化を配慮して作成しないと無理ではないか？

PDF/A-2

- ISOで策定中
 - 当初はPDF 1.6ベース(注1)
 - その後PDF 1.7ベースに変更(注2)
 - 恐らくISO 32000ベースになる(推測)



出展:

注1) PDF/A - Archive

<http://pdf.editme.com/PDFA>

注2) 2007年08月25日PDF/Aの次期バージョンについて

<http://blog.antenna.co.jp/PDFTool/archives/2007/08/25/>

情報源のWIKIとブログは、アメリカのISO/PDF標準化を推進する団体が開設するWIKI/ブログであり、準公式情報とみなされる。