

電子公文書等の長期保存に関する調査検討報告書

2024 年 3 月 19 日

公益社団法人日本文書情報マネジメント協会

本報告書は、独立行政法人国立公文書館による委託事業として、公益社団法人日本文書情報マネジメント協会（JIIMA）が実施した、「電子公文書等の長期保存等に関する調査検討」の成果を取りまとめたものである。

改版履歴

版数	発行日	改訂履歴
1.0	2024 年 3 月 19 日	初版発行

目次

1. 調査検討の概要	3
1.1. 調査検討の背景及び目的	3
1.2. 用語	3
1.3. 調査検討の手順	3
2. 技術的な情報の収集及び長期保存リスクの評価	4
2.1. フォーマットのリスク評価の考え方と標準的フォーマット候補	4
2.1.1. 長期保存にあたってリスクが低いフォーマットの考え方	4
2.1.2. 標準的フォーマット候補案	4
2.2. フォーマットの長期保存リスクの評価	5
2.2.1. フォーマットの技術的特性の整理	5
2.2.2. 長期保存リスクの評価方法の検討	7
2.2.3. フォーマットの長期保存リスクの評価結果	9
2.3. 標準的フォーマット案の候補	12
3. 標準的フォーマットによる作成・保存・利用に向けた検討	13
3.1. 標準的フォーマットの判定	13
3.1.1. フォーマット判定アプリケーション	13
3.1.2. フォーマット判定の検証環境	15
3.1.3. フォーマット判定の結果	16
3.2. 標準的フォーマットによる文書作成	19
3.2.1. 標準的フォーマットによる文書作成方法	19
3.2.2. 意図しない情報の混入パターンと点検方法(予備調査)	20
3.2.3. 長期保存のリスク要因となる機能と点検方法(予備調査)	23
3.3. 標準的フォーマットで既に作成されている場合の対処	26
3.3.1. フォーマット変換	26
3.3.1.1. フォーマット変換の検証対象フォーマット	26
3.3.1.2. フォーマット変換方法	26
3.3.1.3. フォーマット変換結果の確認方法	29
3.3.1.4. フォーマット変換の検証環境	29
3.3.1.5. フォーマット変換の結果	30
3.3.1.6. 複数変換先の選択判断	31
4. まとめ	33
4.1. 標準的フォーマット案	33
4.2. 定期的なフォーマットのリスク評価のためのモデル	35
4.3. 標準的フォーマットによる作成・保存・利用のためのモデル	35
4.4. 今後の検討課題	36

1. 調査検討の概要

1.1. 調査検討の背景及び目的

国立公文書館(以下「館」という。)から令和5年7月の公文書管理委員会に「標準的フォーマット候補(案)」が提示された。本調査検討は、「標準的フォーマット候補(案)」を基にした「標準的フォーマット候補」に対する技術的な観点からの長期保存リスクの評価、長期保存の運用に必要な「フォーマット判定」及び「フォーマット変換」の調査並びに実証実験、利用に際して課題となると考えられる「公文書等の作成に利用するアプリケーション」、「意図しない情報混入」及び「長期保存のリスクとなる機能」について調査検討を実施し、長期保存に適した「標準的フォーマット案」を策定することを目的とする。

1.2. 用語

本報告書では、次の用語を用いる。

標準的フォーマット候補(案) 令和5年7月の公文書管理委員会に提示された標準的フォーマット候補の案

標準的フォーマット候補 本調査検討において長期保存リスク評価の対象とした標準的フォーマット候補

標準的フォーマット案候補 本調査検討における長期保存リスク評価の結果選定された標準的フォーマット案の候補

標準的フォーマット案 本調査検討において総合的な判断のもとに策定された標準フォーマット案

OOXML (Office Open XML) Microsoft 社の Office を国際標準化したフォーマット。Office2007 以降に対応する

PRONOM 英国国立公文書館によって開発された、デジタル保存サービスをサポートする Web ベースの技術レジストリ

PAdES 高度電子署名の要件に合うようにした国際標準化された PDF の拡張

1.3. 調査検討の手順

館が提示した「標準的フォーマット候補(案)」を基に選定した「標準的フォーマット候補」(表 2.1.2-1)に対して、長期保存リスクの評価を行うと共に、表 1.3-1 に示す項目に関して技術的な検証を実施し、長期保存に適した「標準的フォーマット案」を提示する。

表 1.3-1 技術的な評価項目

No.	項目	概要
1	作成アプリケーションに関する技術情報の整理	「標準的フォーマット候補(表 2.1.2-1)」による文書作成技術に関する情報の整理
2	フォーマット判定	次のフォーマットで作成された電子ファイルを判定する技術に関する調査、検証 ・「標準的フォーマット候補(表 2.1.2-1)」 ・「その他点検するフォーマット(表 2.2.1-2)」
3	フォーマット変換	「その他点検するフォーマット(表 2.2.1-2)」から、「標準的フォーマット候補(表 2.1.2-1)」へ変換する技術に関する調査、検証
4	意図しない情報混入の対応	「標準的フォーマット候補(表 2.1.2-1)」で文書を作成、保存、移管する上で、意図しない情報混入など留意を要する機能や設定等に関する予備調査
5	長期保存のリスク要因となる機能の対応	「標準的フォーマット候補(表 2.1.2-1)」で作成された文書であっても、長期保存上のリスク要因となる機能が使用されていた場合の対処に関する予備調査

2. 技術的な情報の収集及び長期保存リスクの評価

2.1. フォーマットのリスク評価の考え方と標準的フォーマット候補

2.1.1. 長期保存にあたってリスクが低いフォーマットの考え方

令和5年7月の公文書管理委員会で報告された、表2.1.1-1の「フォーマットのリスク評価の考え方」を、フォーマットのリスク評価の基本方針とした。

表 2.1.1-1 フォーマットのリスク評価の考え方

No.	フォーマットのリスク評価の考え方	リスク評価に用いる理由
1	仕様が公開されていること	米豪では仕様が公開されていることを評価項目としており、仕様が公開されたフォーマットは長期保存に必要な技術情報を確認・取得することができる
2	世界中で広く使用されていること	米豪いずれも「広く採用されている」、「普及性」を評価項目としている
3	可逆圧縮等が可能なこと	圧縮の際に記録の内容が一部欠損するような圧縮技術は長期保存上のリスクとなり得るため
4	メタデータを付与できること	米豪でもメタデータのサポートを評価項目としており、メタデータで管理できる必要がある
5	技術要素が独立していること	特定の環境(ハードウェア, OS, アプリケーション)に依存していること、技術仕様が複雑であることは長期保存上のリスクとなり得るため
6	特許やライセンス等の制約がないこと	特許, ライセンスに過度の条件があると長期保存上のリスクとなり得るため
7	プロテクションの制約がないこと	コンテンツの保護技術(DRM), 暗号化などがあると, 将来利用できなくなる可能性があり, 長期保存上のリスクとなり得るため

2.1.2. 標準的フォーマット候補案

「標準的フォーマット候補(案)」を基に, 利用が想定されるフォーマットのバリエーション(バージョン, コンフォーマンスレベルなど)を考慮し, 表2.1.2-1の「標準的フォーマット候補」を長期保存上のリスクが低いフォーマットとして抽出した。

表 2.1.2-1 標準的フォーマット候補

カテゴリ	通番	フォーマット名	バリエーション	(参考) 拡張子
① 文書	1	Portable Document Format (PDF)	1.7	pdf
	2		2.0	pdf
	3		PDF/A-1	pdf
	4		PDF/A-2	pdf
	5		PDF/A-3	pdf
	6		PDF/A-4	pdf
	7	Office Open XML (OOXML)	Word processing, strict	docx
	8		Word processing, transitional	docx
	9	Open Document Format for Office Applications (ODF)	Open Document Text	odt
	10		Open Document Text,	fodt

			Single XML(FODF)	
② 表計算	11	Office Open XML (OOXML)	Spreadsheet, strict	xlsx
	12		Spreadsheet, transitional	xlsx
	13	Open Document Format for Office Applications (ODF)	Open Document Spreadsheet	ods
	14		Open Document Spreadsheet, Single XML(FODF)	fods
③ プレゼンテーション	15	Office Open XML (OOXML)	Presentation, strict	pptx
	16		Presentation, transitional	pptx
	17	Open Document Format for Office Applications (ODF)	Open Document Presentation	odp
	18		Open Document Presentation, Single XML(FODF)	fodp
③ 画像	19	JPEG 2000	コンフォーマンスレベル:CL-0(低圧縮高品質), プロファイル:Baseline 注1	jp2
	20	Tagged Image File Format (TIFF)	6	tif, tiff
	21	Portable Network Graphics (PNG)	1.2 (ISO/IEC 15948:2004)	png
	22	Graphics Interchange Format (GIF)		gif
	23	Joint Photographic Experts Group (JPEG)		jpg
④ 音声	24	WAVE (Waveform Audio File Format)		wav
	25	MP3 (MPEG Layer III)		mp3
⑤ 動画	26	MPEG-2	MPEG2-PS	mpg, mpeg
	27		MPEG2-TS	ts
	28	MPEG-4 Part14 (MP4)		mp4

2.2. フォーマットの長期保存リスクの評価

2.2.1. フォーマットの技術的特性の整理

フォーマットの長期保存上のリスク評価を実施するにあたって、表 2.2.1-1 に示す項目を基準とし、標準的フォーマット候補である表 2.2.1-1、その他点検するフォーマットとして表 2.2.1-2 に対して予備調査を実施した結果、直近の技術動向及び市場動向に照らして、特段問題となる変化は生じていなかった。よって、調査対象フォーマットとして表 2.2.1-1 及び表 2.2.1-2 とし、整理する。

予備調査として、表 2.2.1-1 及び表 2.2.1-2 の調査対象とするフォーマットについて、情報を収集・整理した。情報は、公開情報を基に収集を行った。収集したフォーマットの情報については、別紙1「長期保存の観点からリスクが低いフォーマットに関する調査結果」、別紙2「その他点検するフォーマットに関する調査結果」を参照。

表 2.2.1-1 フォーマットの技術的な情報の収集・整理項目

No.	項目	説明
1	フォーマット名	フォーマットの名称
2	バージョン	最新バージョン
3	拡張子	拡張子が複数ある場合にはそれぞれの特徴
4	公開日	最初のバージョンのリリース日
5	最終更新日	最新バージョンのリリース日
6	バージョンの更新履歴	バージョンのリリース履歴、特徴

7	標準化対応状況	ISO や JIS 等の認証状況
8	特長	フォーマットの特長
9	課題	フォーマットの課題
10	開発元	製造・開発元
11	ライセンス	フォーマットの特許, ライセンス条件
12	メタデータの埋込可否	メタデータの埋込可否
13	参照情報	関連する URL など
14	PRONOM ID	英国国立公文書館(TNA)が提供する PRONOM で対応するフォーマットの管理番号
15	NARA における ID	米国国立公文書館(NARA)の対応するフォーマットの管理番号
16	作成アプリケーション名	当該フォーマットを作成可能なアプリケーション
17	閲覧アプリケーション名	当該フォーマットを閲覧可能なアプリケーション
18	日本語への対応	一般的な日本語(例:漢字, ひらがな)への対応状況

表 2.2.1-2 その他点検するフォーマット

カテゴリ	通番	フォーマット名	バリエーション	(参考) 拡張子の例
① 文書	29	一太郎	4	jsw
	30		5	jaw, jtw
	31		6	jbw, juw
	32		7	jfw, jvw
	33		8～12	jtd
	34		2004～2008	jtd
	35	Portable Document Format (PDF)	1.4	pdf
	36		1.5	pdf
	37		1.6	pdf
	38	Microsoft Word 95		doc
	39	Microsoft Word 97-2003		doc
② 表計算	40	Microsoft Excel 95		xls
	41	Microsoft Excel 97-2003		xls
	42	Microsoft Excel 2007(マクロ有効ブック)		xlsm
③プレゼンテーション	43	Microsoft PowerPoint 95		ppt
	44	Microsoft PowerPoint 97-2003		ppt

表 2.2.1-3 音声及び動画コンテナに格納可能な推奨コーデック

通番	コンテナ	コーデック	備考
24	WAVE	LPCM (Linear Pulse Code Modulation)	
25	MP3	MP3	
26	MPEG-4	映像:H264, H262 音声:AAC, MP3	
27	MPEG2-PS	映像:H262 音声:AAC, MP3	エラー耐性が必要な環境用
28	MPEG2-TS	映像:H.264, H265 音声:AAC	MPEG-4 Part2 も可 MP3, AC-3 も可

2.2.2. 長期保存リスクの評価方法の検討

フォーマットの仕様に起因するリスクに関しては、開放的(オープン)であることが重要である。オープンであるとは、それを支持するコミュニティがあり、議論され、維持されることを意味し、高品質の仕様が提供される。国際標準であることが望ましいが、仕様が公開されデファクト標準として広く支持されているものがそれに続く。

特許等のライセンスについては、費用対効果の問題であるが、ライセンスが発生するものは、これまでの歴史を見ても使われない傾向にある。

アプリケーションの普及については、仕様開放性、技術独立性、ライセンスと相関関係にあるが、世界中で市場占有率の高いものは、市場を維持するために品質の維持と技術進歩への対応に膨大な投資をした結果であり、今や、国内市場だけを相手にしては維持が難しい。

フォーマットの長期保存リスクは、図 2.2.2-1 に示すように、フォーマットに固有な持続性(Sustainability)リスクと、運用に伴うリスクがある。フォーマットに固有な持続性リスクは更に、技術リスクと実装を含む展開(Deployment)リスクに分けられる。

フォーマットの長期保存リスクにおける持続性リスクの評価の全体像を図 2.2.2-2 に示す。評価の量子化については、表 2.2.2-1 及び表 2.2.2-2 に示す。

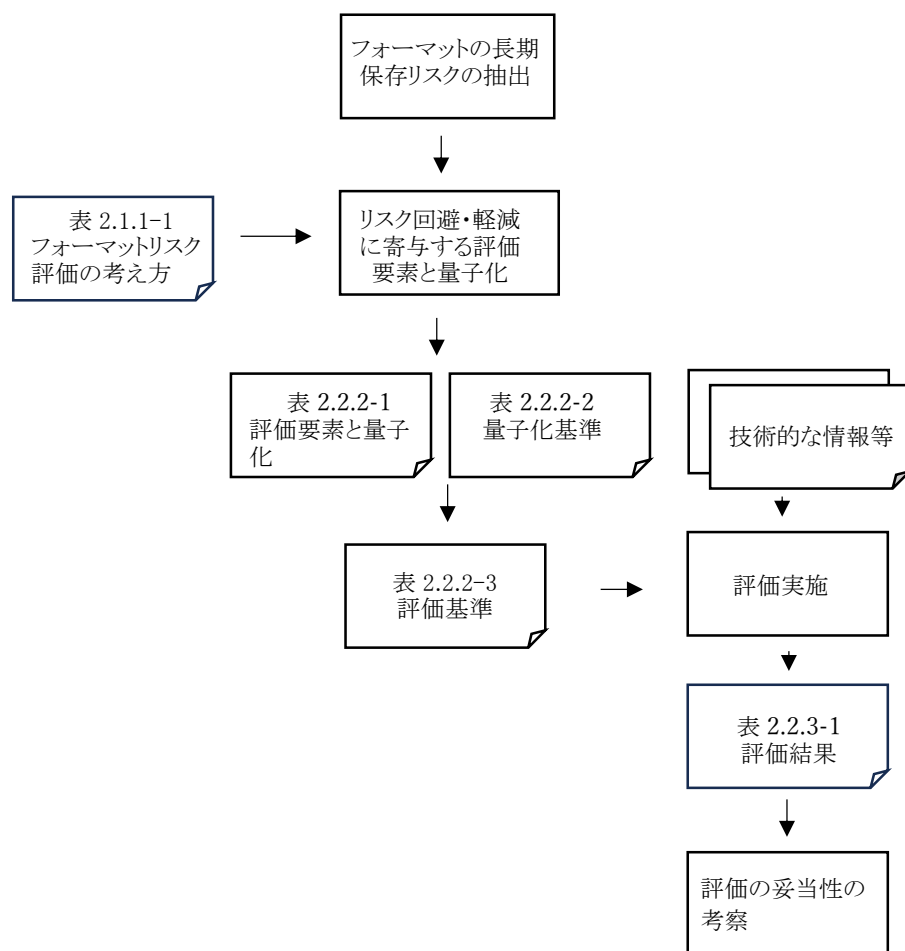


図 2.2.2-2 電子公文書等のフォーマットの持続性リスクの評価の全体像

長期保存リスク	フォーマットに固有な持続性リスク	技術リスク	フォーマット仕様の維持などを含めた技術的なリスク
		展開リスク	フォーマットの展開や実装アプリケーションの普及に関するリスク
	運用に伴うリスク		フォーマットの不適切な使い方に起因するリスク

図 2.2.2-1 長期保存のリスク

表 2.2.2-1 リスク回避・低減策に寄与する評価要素、要素の量子化

カテゴリ	リスク事象	表4 リスク評価の考え方	評価要素
技術リスク	仕様が曖昧で多様な解釈が生まれる。または仕様の改廃訂が特定ベンダの方針に左右される	仕様が公開されていること (No.1)	<u>仕様開放性</u> 2:国際標準 1:仕様公開 0:非公開
	圧縮により画像品質が劣化する	可逆圧縮等が可能なこと (No.3)	<u>圧縮</u> 2:圧縮品質選択可 1:圧縮可 0:圧縮不可
	必要なメタデータが収容できない	メタデータを付与できること (No.4)	<u>メタデータ</u> 2:形式が標準準拠 1:任意要素可 0:固定
	仕様が HW, OS, ライブラリ等の改廃の影響を受ける	技術要素が独立していること (No.5)	<u>技術独立性</u> 2:独立 0:依存
	特許等のライセンスがあり継続的に負担がかかる	特許やライセンス等の制約がないこと (No.6)	<u>ライセンス</u> 2:制約なし 0:制約あり
展開リスク	普及が限定的で相互運用性が損なわれる、またはライフサイクルが終了しアプリケーションの提供が止まる	世界中で広く使用されていること (No.2)	<u>普及度</u> 2:世界中で使用 1:限られた使用 0:使われていない
運用リスク	プロテクションにより読めない、または、マイグレーションができない	プロテクションの制約がないこと (No.7)	個々のフォーマット特性に起因するリスクではないため、評価の対象外

表 2.2.2-2 量子化の基準

No.	評価要素	量子化	基準
1	仕様開放性	2:国際標準	仕様が公開され、仕様を維持するための広く認められた国際的なコミュニティ活動の実態があり、ISO, IEC, ITU, IEEE, IETF, 3GPP, W3C, OASIS などが代表的な国際標準化団体である
		1:仕様公開	仕様が公開されているデファクト標準
		0:非公開	仕様が公開されていないか、公開されていても広く認められていない
2	圧縮	2:圧縮品質選択可	圧縮品質(含む無圧縮)を選択できる
		1:圧縮可	圧縮(可逆圧縮または非可逆圧縮)がサポートされている
		0:圧縮不可	圧縮がサポートされていない、または仕様が不明
3	メタデータ	2:形式が標準準拠	標準準拠(国際またはデファクト標準に準拠した形式の任意のメタデータが付与できる
		1:任意要素可	任意のメタデータが付与できる

		0:固定	任意のメタデータは付与できない, または仕様が不明
4	技術独立性	2:独立	特定の HW, OS 等に依存しない
		0:依存	特定の HW, OS 等に依存するか, 判定不能
5	ライセンス	2:制約なし	特許等が存在しないか, 使用許諾を得ることなく無償で使用するができる
		0:制約あり	特許等の主張があり使用許諾が必要
6	普及度	2:世界中で使用	当該フォーマットをサポートするアプリケーションの実装が相当数 (several)あり世界中で広く使用されている
		1:限られた使用	当該フォーマットをサポートするアプリケーションが国内若しくは一部コミュニティで一定数使用されている又は提供アプリケーションが実質的に1つ
		0:使われていない	当該フォーマットが余り利用されておらずアプリケーションの実装も少ないか, 又はサポート終了が予告, 若しくはサポートを終了している

2.2.3. フォーマットの長期保存リスクの評価結果

表 2.2.3-1に, フォーマット別に収集・整理した情報に基づく, 表 2.2.1-1 及び表 2.2.1-2 のフォーマットに対する評価結果を示す。

フォーマット別の評価理由は次の通りである。

(1) PDF

PDF は PDF 1.4 から 1.6, PDF 1.7 及び PDF 2.0 に分けられる。

PDF 1.4 から PDF 1.6 は, アドビ社の仕様であり, PDF 1.7 (ISO 32000-1) 及び PDF 2.0 (ISO 32000-2) は ISO 規格化されている。

現在の主流は PDF 1.7 であるが, PDF 1.4 から 1.6 もデファクト標準として広く使われている。PDF 1.4 は PDF/A-1 のベース仕様にもなっている。一方, PDF 2.0 は, 2020 年に ISO 規格になったが, アプリケーション提供は進んでいないことから, PDF 2.0 普及度の評価を“1”とした。

PDF 電子署名「PADES」を付与する場合は, PDF 2.0 が前提となる。

(2) PDF/A

PDF/A は, PDF 仕様の使い方 (プロファイル) の規定である。PDF/A には, PDF 1.4 ベースの PDF/A-1, PDF 1.7 ベースの PDF/A-2, A-3, PDF 2.0 ベースの PDF/A-4 があり, いずれも ISO19005-1, ISO19005-2, ISO19005-3, ISO19005-4 として ISO 規格化されている。

現在の主流は PDF/A-2 であるが, PDF/A-2 と PDF/A-3 は埋込ファイルを PDF に限定するか否かだけの違いである。PDF/A-1 には透過性の問題があるが, 紙文書のスキャニングには問題ないので, いまだに広く使われている。一方, PDF/A-4 は, 2021 年に ISO 規格化されたが, PDF 2.0 をベースとしていることもあり普及は進んでいない。このことから PDF/A-4 の普及度の評価を“0”とした。

(3) OOXML

OOXML は, Microsoft Office の仕様を XML 化し, ECMA (European Computer Manufacturers' Association) 経由で ISO に提案され ISO/IEC 29500 として ISO 規格となった。

Word processing, Spreadsheet, Presentation の 3 つのパッケージングがあり, それぞれ, Microsoft Office の Word, Excel, PowerPoint に対応する。

OOXML には, strict と transitional の 2 つのコンFORMANCE レベル (互換性を保証する単位) がある。strict は ISO/IEC 29500 完全準拠で, バイナリの Word, Excel, PowerPoint (つまり, doc, xls, ppt) と互換性がなく, 普及が進んで

こなかった。Microsoft Office は、2007、2010 は transitional、2013 以降は strict と transitional の 2 つのコンFORMANCE レベルを実装しているが、デフォルトは、transitional であり、明示的に「完全 Open XML ドキュメント、または strict Open XML スプレッドシート、または完全 Open XML プレゼンテーション」を指定しないと、strict レベルにならない。これらから strict の普及度の評価を“0”とした。

(4) ODF

ODF は OASIS で開発された仕様を ISO/IEC 26300 として ISO 規格化したものである。ODF は、ヨーロッパで広く使われているが、日本では余り使われていない。

ODF は XML 構造に 2 つ表現方法があり、ZIP で包む方法と、SingleXML (FlatODF) がある。FlatODF ファイルは、テキストエディタで編集できることから、一部のコミュニティで日常的に使われているが、通常の行政文書を、XML 言語を駆使して作成ことは考えられず、また、全ての ODF アプリケーションが FlatODF をサポートしている訳ではない。これらの点から、ODF の普及度の評価を“1”とし、FlatODF の普及度の評価を“0”とした。

(5) 画像、音声、動画フォーマット

TIFF6、GIF、WAVE を除き、基本的に、国際標準化された情報交換フォーマットであり、圧縮は数学的な演算であるため、アプリケーションの維持は比較的容易である。TIFF6、GIF、WAVE は国際標準ではないがデファクトとなっているので評価点を“1”とした。なお、TIFF6 対応の国際標準 (TIFF/IT) は、出版向けのサブセットである。その他の要素についてはいずれのフォーマットも、減点要素はない。

MPEG2-TS は、業務用途のフォーマットに分類されるため、標準的フォーマット候補の対象外とする。

(6) 一太郎

仕様が公開されていないこと、仕様が公開されていないことから技術独立性が確認できないこと、国内のみの市場であることから、該当する評価を与えた。

(7) Microsoft Word 95, Excel 95, PowerPoint 95

バイナリフォーマットであり仕様が非公開であることから仕様開放性の評価を“0”に、従って、技術独立性が確認できないことから技術独立性の評価を“0”に、また、既に移行パッケージも提供されていないことから普及度の評価を“0”とした。

(8) Microsoft Word 97-2003, Excel 97-2003, PowerPoint 97-2003

仕様が公開されていることから、仕様開放性の評価を“1”に、技術独立性の評価を“2”に、また、既に仕様が凍結されていることから普及度の評価を“0”とした。

(9) Microsoft Excel 2007 (マクロ有効ブック)

OOXML の一環ではあるが、VBA (Visual Basic for Applications) が最新の環境で正常に動作しない可能性があることが指摘されていることから、技術独立性の評価を“0”とした。

(10) OASYS

「表 2 その他点検するフォーマット」の対象外だが、国立公文書館への移管の実績を踏まえ、別途リスク評価を実施した。仕様が非公開であることから仕様開放性の評価は 0 点とした。圧縮、技術独立性、メタデータ、特許については不明であるため 0 点となった。またアプリケーションについても、個人向けは 2020 年 9 月 30 日、法人向けは 2021 年 5 月 31 日に販売が終了し、サポートについても、個人向けは 2023 年 9 月 30 日、法人向けは 2024 年 5 月 31 日に終了するため、アプリケーションの普及度の評価を 0 点とした。

フォーマットの長期保存リスクを図 2.2.2-2 の全体像に従って評価した結果を表 2.2.3-1 に示す。

表 2.2.3-1 フォーマットの長期保存リスクの評価結果

評価項目	仕様開放性	メタデータ	圧縮	技術独立性	特許等	普及度	評価点
量子化(0,1,2) フォーマット名	2:国際標準	2:形式が標準準拠	2:圧縮品質指定可	2:独立	2:制約なし	2:世界中で使用	
	1:仕様公開	1:任意要素可	1:圧縮可			1:限られた使用	
	0:非公開	0:固定・不明	0:圧縮不可・不明	0:依存・不明	0:制約あり・不明	0:使われていない	
1 PDF1.7	2	2	1	2	2	2	11
2 PDF2.0	2	2	1	2	2	1	10
3 PDF/A-1	2	1	1	2	2	2	10
4 PDF/A-2	2	2	1	2	2	2	11
5 PDF/A-3	2	2	1	2	2	2	11
6 PDF/A-4	2	2	1	2	2	0	9
7 OOXML,Wordprocessing,strict	2	2	1	2	2	0	9
8 OOXML,Wordprocessing,transitional	2	2	1	2	2	2	11
9 OpenDocument Text	2	2	1	2	2	1	10
10 OpenDocument Text,SingleXML	2	2	1	2	2	0	9
11 OOXML,Spredsheel,strict	2	2	1	2	2	0	9
12 OOXML,Spreadsheet,transitional	2	2	1	2	2	2	11
13 OpenDocument Spreadsheet	2	2	1	2	2	1	10
14 OpenDocument Spreadsheet,SingleXML	2	2	1	2	2	0	9
15 OOXML,Presentation,strict	2	2	1	2	2	0	9
16 OOXML,Presentation,transitional	2	2	1	2	2	2	11
17 OpenDocument Presentation	2	2	1	2	2	1	10
18 OpenDocument Presentation,SingleXML	2	2	1	2	2	0	9
19 JPEG2000	2	2	2	2	2	2	12
20 TIFF 6 (Tagged Image File Format)	1	2	2	2	2	2	11
21 PNG 1.2 (Portable Network Graphics)	2	1	1	2	2	2	10
22 GIF (Graphics Interchange Format)	1	1	1	2	2	2	9
23 JPEG (Joint Photographic Experts Group)	2	2	2	2	2	2	12
24 WAVE (Waveform Audio File Format)	1	1	2	2	2	2	10
25 MP3 (MPEG Layer III Audio Encoding)	2	2	2	2	2	2	12
26 AAC (Advanced Audio Coding)	2	1	2	2	2	2	11
27 MPEG-4	2	2	2	2	2	2	12
28 MPEG2-PS	2	2	2	2	2	2	12
29 MPEG2-TS	2	2	2	2	2	2	12
30 一太郎 4	0	0	0	0	2	1	3
31 一太郎 5	0	0	0	0	2	1	3
32 一太郎 6	0	0	0	0	2	1	3
33 一太郎 7	0	0	0	0	2	1	3
34 一太郎 8-12	0	0	0	0	2	1	3
35 一太郎 2004-2008	0	0	0	0	2	1	3
36 PDF1.4	1	1	1	2	2	2	9
37 PDF1.5	1	2	1	2	2	2	10
38 PDF1.6	1	2	1	2	2	2	10
39 Microsoft Word 95	0	0	0	0	2	0	2
40 Microsoft Word 97-2003	1	0	1	2	2	0	6
41 Microsoft Excel 95	0	0	0	0	2	0	2
42 Microsoft Excel 97-2003	1	0	1	2	2	0	6
43 Microsoft Excel 2007(マクロ有効ブック)	2	2	1	0	2	1	8
44 Microsoft PowerPoint 95	0	0	0	0	2	0	2
45 Microsoft PowerPoint 97-2003	1	0	1	2	2	0	6

2.3. 標準的フォーマット案の候補

フォーマットの長期保存リスクを評価し、標準的フォーマット案の候補は表 2.3-1 に示す結果となった。

表 2.3-1 標準的フォーマットリスト案の候補

カテゴリ	フォーマット名	バリエーション	拡張子
① 文書	Portable Document Format (PDF)	1.4～1.6	pdf
		1.7	pdf
		2.0	pdf
		PDF/A-1	pdf
		PDF/A-2	pdf
		PDF/A-3	pdf
	Office Open XML (OOXML)	Word processing, transitional	docx
② 表計算	Open Document Format for Office Applications (ODF)	Open Document Text	odt
	Office Open XML (OOXML)	Spreadsheet, transitional	xlsx
③ プレゼンテーション	Open Document Format for Office Applications (ODF)	Open Document Spreadsheet	ods
	Office Open XML (OOXML)	Presentation, transitional	pptx
④ 画像	Open Document Format for Office Applications (ODF)	Open Document Presentation	odp
	JPEG 2000	コンフォーマンストレベル:CL-0(低圧縮高品質), プロファイル:Baseline	jp2
	Tagged Image File Format (TIFF)	6	tif, tiff
	Portable Network Graphics (PNG)	1.2	png
	Graphics Interchange Format (GIF)		gif
⑤ 音声	Joint Photographic Experts Group (JPEG)		jpg
	WAVE (Waveform Audio File Format)		wav
⑥ 動画	MP3 (MPEG Layer III)		mp3
	MPEG-2	MPEG2-PS	mpg, mpeg
	MPEG-4 Part14 (MP4)		mp4

3. 標準的フォーマットによる作成・保存・利用に向けた検討

3.1. 標準的フォーマットの判定

3.1.1. フォーマット判定アプリケーション

電子公文書等の長期保存の運用においては、長期保存に適していない標準的フォーマット以外のファイルを識別し、標準的フォーマットへの変換を実施する必要がある。本調査検討では、実運用を考慮し、フォーマットの判定と変換を自動的に実施し手動操作を不要となるようアプリケーションを検討した。

フォーマット判定アプリケーションの要件を次に示す。

- 拡張子の有無、誤りに関わらずフォーマットが判定できる
- フォーマット判定でファイルの属性情報が取得できる
 - ファイルパス
 - ファイル名
 - ファイルサイズ
 - チェックサム（ファイルの同一性を確認することができる、例として MD5, SHA256, or SHA512 がある）
 - 最終編集日時
 - PRONOM フォーマット固有識別子（PUID）
 - フォーマット名
 - フォーマットバージョン
 - MIME Type（Web サーバーと Web ブラウザーの間で使用されるデータの形式を文字列で表したもの）
- フォーマット判定またはフォーマット検証で次の情報が取得できる
 - PDF：PDF のバージョン, PDF/A のバージョン
 - 文書, 表計算, プレゼンテーション：OOXML Strict, Transitional
 - 画像：コーデック
 - 音声：コーデック
 - 動画：コーデック

システム化を前提に以下のアプリケーションを採用候補とした。フォーマット判定とフォーマット検証に採用したアプリケーションと採用理由を表 3.1.1-1 に示す。

表 3.1.1-1 フォーマット判定とフォーマット検証に採用したアプリケーションと採用理由

カテゴリ	アプリケーション名	機能	採用	理由
[PDF][文書][表計算][プレゼン][画像][音声][動画]	DROID (Digital Record Object Identification)	メタデータ抽出	○	PRONOM の参考実装であり、様々な文献から採用されていることが確認できているため検証対象として採用する
[PDF][文書][表計算][プレゼン][画像][音声][動画]	fido (Format Identification for Digital Objects)	メタデータ抽出 ファイル形式識別	○	PRONOM の参考実装の1つであり、DROID との比較している文献もあるため検証対象として採用する
[PDF][文書][表計算][プレゼン][画像][音声][動画]	Siegfried	ファイル形式識別	○	PRONOM の参考実装の1つであり、fido と同様 DROID との比較している文献もあるため検証対象として採用する
[PDF][文書][表計算][プレゼン][画像][音声][動画]	JHOVE	メタデータ抽出 ファイル形式識別	○	汎用のファイルフォーマット検証アプリケーションであり、幅広いファイルフォーマットを検証できるため検証対象として採用する
[文書][PDF]	veraPDF	PDF, PDF/A のフォーマット検証	○	PDF, PDF/A の検証をおこなうアプリケーションとして開発されており、Open

				Preservation Foundation でも紹介されているため検証対象として採用する
[文書][PDF]	Adobe Acrobat Pro	PDF のフォーマット検証	○	PDF の検証を行うアプリケーションとしての実績 No.1 であるため検証対象として採用する
[文書][表計算][プレゼン][OOXML]	Office SDK	OOXML のフォーマット検証	○	OOXML の検証を行うアプリケーションとしてオープンソースかつ、約 10 年に渡って提供されている実績があるため検証対象として採用する
[文書][表計算][プレゼン][ODF]	ODF Toolkit	ODF のフォーマット検証	○	ODF の検証を行うアプリケーションとしてオープンソースかつ、約 5 年にわたって提供されている実績があるため検証対象として採用する
[画像]	Adobe Photoshop	画像編集		スクリプトでの実行は可能だが、GUI がバックグラウンドで動作するためエラー時にポップアップウィンドウが表示されることから自動処理には不向き。
[画像]	GIMP	画像編集		GUI による画像情報の表示は可能であるがコマンドラインからの実行は不可。
[画像]	IrfanView	画像閲覧・簡易編集		画像情報出力をコマンドラインから実行可能であるが、エラー時にポップアップウィンドウが表示されることから自動処理には不向き。
[画像]	ImageMagick	画像編集	○	画像ファイルの解析が可能であり、オプション指定により画像コーデック等の詳細情報の抽出が可能。
[音声][動画]	FFmpeg	動画・音声の記録・変換・再生	○	画像、音声、動画ファイルの解析が可能であり、解析時のログレベルを指定することでエラー／警告の表示が可能。

表 3.1.1-2 にフォーマット判定およびフォーマット検証を行うフォーマット毎の対応アプリケーションを示す。

表 3.1.1-2 フォーマット毎の対応アプリケーション候補

カテゴリ	フォーマット名	バリエーション	フォーマット判定	フォーマット検証
① 文書	Portable Document Format (PDF)	PDF/A-1, A-2, A-3, A-4	DROID/fido/Siegfried	JHOVE/veraPDF
		1.7, 2.0	DROID/fido/Siegfried	JHOVE/veraPDF
	Office Open XML (OOXML)	Word processing strict, Word processing transitional	DROID/fido/Siegfried	JHOVE/Office SDK
	Open Document Format for Office Applications (ODF)	Open Document Text	DROID/fido/Siegfried	JHOVE/ODF Toolkit
		Open Document Text, Single XML(FODF)	DROID/fido/Siegfried	JHOVE/ODF Toolkit
	一太郎		DROID/fido/Siegfried	JHOVE
	Portable Document Format (PDF)	1.4, 1.5, 1.6	DROID/fido/Siegfried	JHOVE/veraPDF
	Microsoft Word 95		DROID/fido/Siegfried	JHOVE
	Microsoft Word 97 - 2003		DROID/fido/Siegfried	JHOVE
② 表計算	Office Open XML (OOXML)	Spreadsheet strict, Spreadsheet transitional	DROID/fido/Siegfried	JHOVE/Office SDK
	Open Document Format for Office Applications (ODF)	Open Document Spreadsheet	DROID/fido/Siegfried	JHOVE/ODF Toolkit

		Open Document Spreadsheet, Single XML(FODF)	DROID/fido/Siegfried	JHOVE/ODF Toolkit
	Microsoft Excel 95		DROID/fido/Siegfried	JHOVE
	Microsoft Excel 97 – 2003		DROID/fido/Siegfried	JHOVE
	Microsoft Excel 2007(マクロ有効ブック)		DROID/fido/Siegfried	JHOVE/Office SDK
③ プレゼンテーション	Office Open XML (OOXML)	Presentation strict, Presentation, transitional	DROID/fido/Siegfried	JHOVE/Office SDK
	Open Document Format for Office Applications (ODF)	Open Document Presentation,	DROID/fido/Siegfried	JHOVE/ODF Toolkit
		Open Document Presentation, Single XML(FODF)	DROID/fido/Siegfried	JHOVE/ODF Toolkit
	Microsoft PowerPoint 95		DROID/fido/Siegfried	JHOVE
	Microsoft PowerPoint 97 – 2003		DROID/fido/Siegfried	JHOVE
④ 画像	JPEG 2000		DROID/fido/Siegfried	JHOVE/FFmpeg/ImageMagick
	TIFF 6 (Tagged Image File Format)		DROID/fido/Siegfried	
	PNG 1.2 (Portable Network Graphics)		DROID/fido/Siegfried	
	GIF (Graphics Interchange Format) gif		DROID/fido/Siegfried	
	JPEG (Joint Photographic Experts Group) jpg		DROID/fido/Siegfried	
⑤ 音声	WAVE (Waveform Audio File Format) wav		DROID/fido/Siegfried	JHOVE/FFmpeg
	MP3 (MPEG Layer III Audio Encoding) mp3,mp3 等		DROID/fido/Siegfried	
	AAC (Advanced Audio Coding MPEG-4 LowComplexity Object)		DROID/fido/Siegfried	
⑥ 動画	MPEG-4		DROID/fido/Siegfried	JHOVE/FFmpeg
	MPEG-2 (MPEG2-PS, MPEG2-TS)		DROID/fido/Siegfried	

3.1.2. フォーマット判定の検証環境

フォーマット判定の検証環境は、AWS 上に Windows Server 2022 のインスタンスを以下に示すスペックで構築し、今回検証する判定アプリケーションをインストールしてフォーマット判定を実施した。

OS : Windows Server 2022

CPU : Intel(R) Xeon(R) CPU E5-2676 v3 @ 2.40GHz x 2

メモリー : 8.00 GB

ディスク : 50GB

3.1.3. フォーマット判定の結果

フォーマット判定アプリケーションとしては、英国国立公文書館(TNA : The National Archives)が公開している DROID をはじめいくつかのアプリケーションでフォーマット判定できることを確認したが、表 3.1.3-1 及び表 3.1.3-2 に示すとおり、各アプリケーションの中で DROID が一番幅広いフォーマット判定と関連するファイルの情報を多く取得できたこと、加えて自動化の観点でコマンドラインでの呼び出し、フォルダーを指定しての一括処理、及び結果を CSV ファイルで出力できることから、DROID がフォーマット判定アプリケーションとして最適であった。

また画像、音声、及び動画ではコーデックなどのより詳細な情報を取得する必要があるため、表 3.1.3-3 の結果に示すとおり、フォーマット検証アプリケーションとして画像は ImageMagick、音声及び動画は FFmpeg が最適であった。

表 3.1.3-1 フォーマットの判定結果

カテゴリ	フォーマット名	バリエーション	拡張子	判定結果								
				拡張子あり			拡張子なし			拡張子偽装		
				DROID	fido	Siegfried	DROID	fido	Siegfried	DROID	fido	Siegfried
文書	Portable Document Format (PDF)	PDF/A-1, A-2, A-3, A-4	pdf	○	×	○	○	×	○	○	×	○
		PDF 1.4～1.7, 2.0	pdf	○	○	○	○	○	○	○	○	○
	Office Open XML (OOXML)	Office Open XML(OOXML):Strict	docx	○	○	○	○	○	○	○	○	○
		Office Open XML(OOXML): Transitional	docx	○	○	○	○	○	○	○	○	○
	Microsoft Word 97-2003		doc	○	○	○	○	○	○	○	○	○
	Open Document Format for Office Applications (ODF)	Open Document Text	odt	○	○	○	○	○	○	○	○	○
		Open Document Text, Single XML(FODF)	fodt	×	×	×	×	×	×	×	×	×
	一太郎	4	jsw	×	×	×	×	×	×	×	×	×
		5	jaw	×	×	×	×	×	×	×	×	×
		6	jbw	×	×	×	×	×	×	×	×	×
		7	jfw	○	○	○	○	○	○	○	○	○
		2008～2004／13～8 定型	jtt	○	○	○	○	○	○	○	○	○
		2008～2004／13～8 通常	jtd	○	○	○	○	○	○	○	○	○
		2008～2004／13～11 通常(圧縮)	jtdc	○	○	○	○	○	○	○	○	○

表計算	Office Open XML (OOXML)	Office Open XML(OOXML):Strict	xlsx	○	○	○	○	○	○	○	○	○
		Office Open XML(OOXML): Transitional	xlsx	○	○	○	○	○	○	○	○	○
	Microsoft Excel 97-2003		xls	○	○	○	○	○	○	○	○	○
	Open Document Format for Office Applications (ODF)	Open Document Spreadsheet	ods	○	○	○	○	○	○	○	○	○
		Open Document Spreadsheet, Single XML(FODF)	fods	×	×	×	×	×	×	×	×	×
プレゼンテーション		Office Open XML(OOXML):Strict	pptx	○	○	○	○	○	○	○	○	○
	Office Open XML(OOXML): Transitional		pptx	○	○	○	○	○	○	○	○	○
	Microsoft PowerPoint 97-2003		ppt	○	○	○	○	○	○	○	○	○
	Open Document Presentation		odp	○	○	○	○	○	○	○	○	○
	Open Document Presentation, Single XML(FODF)		fodp	×	×	×	×	×	×	×	×	×
画像	JPEG 2000		jp2	○	○	○	○	○	○	○	○	○
	Tagged Image File Format (TIFF)		tif	○	○	○	○	○	○	○	○	○
	Portable Network Graphics (PNG)		png	○	○	○	○	○	○	○	○	○
	Graphics Interchange Format (GIF)		gif	○	○	○	○	○	○	○	○	○
	Joint Photographic Experts Group (JPEG)		jpg	○	○	○	○	○	○	○	○	○
音声	WAVE (Waveform Audio File Format)		wav	○	○	○	○	○	○	○	○	○
	MP3 (MPEG Layer III)		mp3	○	○	○	○	○	○	○	○	○
	Advanced Audio Coding (AAC)		aac	○	○	○	○	○	○	○	○	○
動画	MPEG2-PS		mpg	○	○	○	○	○	○	○	○	○
	MPEG2-TS		ts	○	○	○	○	○	○	○	○	○
	MPEG-4 Part14(MP4)		mp4	○	○	○	○	○	○	○	○	○

(凡例) ○:判定可能, ×:判定不可能

表 3.1.3-2 フォーマット毎の取得情報詳細

属性情報	DROID	fido	Siegfried
ファイルパス	○	×	×
ファイル名	○	○	○
ファイルサイズ	○	○	○
チェックサム (e.g. MD5, SHA256, or SHA512)	○	×	×
最終編集日時	○	×	○
PRONOM フォーマット固有識別子 (PUID)	○	○	○
フォーマット名	○	×	○
フォーマットバージョン	○	×	○
MIME Type	○	○	○

表 3.1.3-3 フォーマット毎の取得可能な属性情報

カテゴリ	フォーマット名	バリエーション	取得可能情報					
			JHOVE	veraPDF	Office SDK	ODF Toolkit	Image Magick	FFmpeg
取得する属性情報			ファイル構造エラーチェック (Status)	PDF, PDF/Aバージョン	フォーマット検証結果 (エラー, 警告)	フォーマット検証結果(エラー, 警告)	コーデック	コーデック
文書	Portable Document Format (PDF)	PDF/A-1, A-2, A-3, A-4	○	○	-	-	-	-
	Office Open XML (OOXML)	Microsoft Word 2007以降 (OOXML)	○	-	○	-	-	-
	Open Document Format for Office Applications (ODF)	Open Document Text	○	-	-	○	-	-
表計算	Office Open XML (OOXML)	Microsoft Excel 2007以降 (OOXML)	○	-	○	-	-	-
	Open Document Format for Office Applications (ODF)	Open Document Spreadsheet	○	-	-	○	-	-
プレゼンテーション	Office Open XML (OOXML)	Microsoft PowerPoint 2007 以降 (OOXML)	○	-	○	-	-	-
	Open Document Format for Office Applications (ODF)	Open Document Presentation	○	-	-	○	-	-
画像	JPEG 2000		○	-	-	-	○	△(※1)
	TIFF 6 (Tagged Image File Format)		○	-	-	-	○	△(※1)
	PNG 1.2 (Portable Network Graphics)		○	-	-	-	○	△(※1)

	GIF (Graphics Interchange Format) gif		○	-	-	-	○	△(※1)
	JPEG (Joint Photographic Experts Group) jpg		○	-	-	-	○	△(※1)
音声	WAVE (Waveform Audio File Format) wav		○	-	-	-	-	○
	MP3 (MPEG Layer III Audio Encoding) mp3,mp3 等		○	-	-	-	-	○
	AAC (Advanced Audio Coding MPEG-4 LowComplexity Object)		○	-	-	-	-	○
動画	MPEG-4		○	-	-	-	-	○
	MPEG-2 (MPEG2-PS, MPEG2-TS)		○	-	-	-	-	○

※1 ImageMagick がより詳細情報を取得可能のため、FFmpeg の評価は「△」とした。

3.2. 標準的フォーマットによる文書作成

3.2.1. 標準的フォーマットによる文書作成方法

標準的フォーマットで文書を作成する方法を以下に示す。また、標準的フォーマットのファイルを作成するアプリケーションの例を表 3.2.1-1 に示す。

表 3.2.1-1 標準的フォーマットのファイルを作成するためのアプリケーション例

カテゴリ	フォーマット名	アプリケーション名	開発元	ライセンス	最終更新日
①文書	Portable Document Format (PDF)	Adobe Acrobat	Adobe Systems	有償	2023 年:DC
		JUST PDF	ジャストシステム	有償	2022 年:V5
		SkyPDF	スカイコム	有償	2023 年:7
	Office Open XML (OOXML)	Microsoft Word	マイクロソフト	有償	2021 年:2021
	Open Document Format for Office Applications (ODF)	LibreOffice Writer	The Document Foundation	LGPL	2022 年:7.4.0
		OpenOffice Writer	Apache ソフトウェア財団	Apache License	2023 年:4.1.14
②表計算	Office Open XML (OOXML)	Microsoft Excel	マイクロソフト	有償	2021 年:2021
	Open Document Format for Office Applications (ODF)	LibreOffice Calc	The Document Foundation	LGPL	2022 年:7.4.0
		OpenOffice Calc	Apache ソフトウェア財団	Apache License	2023 年:4.1.14
③プレゼンテーション	Office Open XML (OOXML)	Microsoft PowerPoint	マイクロソフト	有償	2021 年:2021
	Open Document Format for Office Applications (ODF)	LibreOffice Impress	The Document Foundation	LGPL	2022 年:7.4.0
		OpenOffice Impress	Apache ソフトウェア財団	Apache License	2023 年:4.1.14
④画像	JPEG 2000, Tagged Image File Format (TIFF), Portable Network Graphics (PNG)	Adobe Photoshop	Adobe Systems	有償	2023 年:24.7
		GNU Image Manipulation Program	The GIMP Development Team	GPL	2023 年:2.10.34
		IrfanView	Irfan Skiljan	個人:フリー 商用:有償	2022 年:4.62

	, Graphics Interchange Format (GIF) , Joint Photographic Experts Group (JPEG)				
⑤音声	WAVE (Waveform Audio File Format) MP3 (MPEG Layer III) AAC (Advanced Audio Coding)	VLC	VideoLAN project	GPL	2022 年: 3.0.17
		Audacity	The Audacity Team	GPL	2023 年: 3.3.3
		Adobe Premiere Elements	Adobe Systems	有償	2023 年: 24.0
⑥動画	MPEG-2 MPEG-4 Part14 (MP4)	FFmpeg	FFmpeg チーム	LGPL	2023 年: 6.0
		Windows Media Player	マイクロソフト	Windows 標準	2009 年: V12
		Adobe Premiere Elements	Adobe Systems	有償	2023 年: 24.0

3.2.2. 意図しない情報の混入パターンと点検方法(予備調査)

意図しない情報混入パターンについて、ファイル分類ごとに混入パターンおよび混入情報を表 3.2.2-1 に示す。

意図しない情報混入の点検方法は各フォーマット及びアプリケーションによって異なる。以下に各フォーマットにおける点検方法を示す。

Word, Excel, PowerPoint では、以下手順で点検できる。

1. 「ファイル」の「情報」を選択する
2. 「文書の検査」の「問題のチェック」より「ドキュメント検査」を選択する
3. 「ドキュメントの検査」ウインドウで「検査」を選択する
4. 検査結果を確認し、自動的に削除できない情報については「詳細情報」を確認する

一太郎では、以下手順で点検できる。

1. ドキュメント検査をしたい文書を開き、「ファイル」→「文書補助」→「ドキュメント検査」を選択する
2. 最新の状態で保存されていない場合、メッセージが表示される。「はい」または「いいえ」を選択する
3. 検査できる4つの項目が表示されるので、検査した項目の「検査」または、「すべて検査」を選択する
4. 検査が実行され、結果が表示される

PDF では、以下手順で点検できる。

1. Adobe Acrobat で検査したい PDF ファイルを開き、「すべてのツール」の「PDF を保護」を選択する
2. 「PDF を保護」から「非表示情報を検索して削除」を選択する
3. 非表示情報の検索が実行され検査結果が表示される。ツリーを開くことで詳細情報の確認を行う

画像(JPEG, TIFF)では、以下手順で点検できる。

1. エクスプローラーで点検したい画像ファイルを右クリックし「プロパティ」を選択する
2. プロパティウインドウで詳細タブを選択し詳細情報を確認する

表 3.2.2-1 意図しない情報混入パターンと混入情報

カテゴリ	フォーマット名	情報混入パターン	混入情報
①文書	Portable Document Format (PDF)	非表示情報としてコメントや注釈、変更履歴や修正履歴などが含まれていることがある。	メタデータ、コメント、履歴、非表示テキストなど

②表計算	Microsoft Word 2007 以降 (OOXML)	複数人が共同で作成した文書には、変更履歴、コメント、インク履歴、バージョンなどが含まれている場合があり、文書に手を加えた人の名前、校閲者のコメント、文書に加えられた変更が含まれていることがある。	コメント、変更履歴、バージョン、インク注釈
		文書のプロパティには、作成者、件名、タイトルなど、文書の詳細情報が含まれる。また、文書を最後に保存した人の名前、文書が作成された日付などが自動的に保存される。作成者が使用する機能によっては電子メールヘッダー、校閲者への送信に関する情報、回覧用紙、テンプレート名などの個人を特定できる情報が文書に含まれていることがある。	文書のプロパティと個人情報
		文書のヘッダーやフッターに情報が含まれていることがある。また、Word 文書には透かしを追加する場合がある。	ヘッダー、フッター、透かし
		Word 文書には隠し文字に設定されたテキストが含まれている場合がある。	隠し文字
		Windows SharePoint Services をベースにしたドキュメントワークスペースサイトやライブラリなど、ドキュメント管理サーバー上の場所に保存された文書には、サーバーの場所に関する文書のプロパティや情報が含まれていることがある。	ドキュメントサーバーのプロパティ
		文書自体には表示されないカスタム XML データ(*1)が含まれていることがある。	カスタム XML データ
	一太郎	複数人が共同で作成した文書には、添削、指摘、注釈／注釈行が含まれている場合があり、文書に手を加えた人の名前、校閲者のコメント、文書に加えられた変更が含まれていることがある。	添削、指摘、注釈／注釈行
		文書のプロパティには、作成者、件名、タイトルなど、文書の詳細情報が含まれる。また、文書を最後に保存した人の名前、文書が作成された日付などが自動的に保存される。	文書情報と個人情報
		文書のヘッダーやフッターに情報が含まれていることがある。また、Word 文書には透かしを追加する場合がある。	ヘッダー、フッター、透かし
		絞り込みなどにより非表示となっている情報が含まれる場合がある。	非表示の情報
	Microsoft Excel 2007 以降 (OOXML)	複数人が共同で作成したブックには、コメントやインク注釈などが含まれている可能性があり、ブックで作業した人の名前、校閲者からのコメント、ブックに加えられた変更が混入する。	コメント、インク注釈
		Excel のメタデータまたはドキュメントプロパティには、作成者、件名、タイトルなどの情報が含まれる。また、最後に保存したユーザーの名前、ドキュメントが作成された日付、ドキュメントの場所が自動的に保存され、メールのヘッダー、校閲者への送信に関する情報、回覧用紙、プリンターのプロパティ(プリンターのパス、セキュリティで保護された印刷パスコードなど)、Web ページを発行するためのファイルパス情報など、ほかにも個人を特定できる情報が含まれていることがある。	ドキュメントのプロパティと個人情報
		ブックのヘッダーやフッターに情報が含まれていることがある。	ヘッダー、フッター
		ブックでは、行、列、ワークシート全体を非表示にできるため、これらの行、列、またはワークシートを再表示することで入力されたデータを見ることができる。	非表示の行、列、ワークシート
		ドキュメントワークスペースサイトや、Windows SharePoint Services に基づくライブラリなど、ドキュメント管理サーバー上の場所に保存されたブックには、サーバーの場所に関連するドキュメントプロパティや情報が含まれている場合がある。	ドキュメントサーバーのプロパティ
		文書自体では表示されないカスタム XML データが含まれていることがある。	カスタム XML データ

		非表示として書式設定されているため、表示されないオブジェクトが含まれていることがある。	非表示のコンテンツ
		他のブック内のデータへのリンク(外部リンク)が含まれている場合、そのデータを含むワークシートの名前はブックと共に保存されるが、表示されない場合がある。	外部リンク
		埋め込みファイルや、表示されないデータを含む埋め込みオブジェクト(グラフや数式など)が含まれていることがある。	埋め込まれたファイルまたはオブジェクト
		マクロ、VBA モジュール(*2)、COM(*3)または ActiveX コントロール(*4)、ユーザーフォーム、または非表示データを含む可能性がある User-Defined 関数が含まれていることがある。	VBA コードのマクロ
		表示されない可能性のあるピボットテーブル、ピボットグラフ、スライサー、タイムライン、キューブ式のキャッシュされたデータ(*5)が含まれていることがある。	データがキャッシュされている可能性がある項目
		Web 用 Excel で入力され、ブックと共に保存され、ブックには表示されない非表示の Excel アンケートの質問が含まれていることがある。	Excel アンケート
		シナリオマネージャーを使用して定義されたシナリオにはキャッシュされたデータや非表示のデータが含まれていることがある。	Scenario Manager のシナリオ
		アクティブなオートフィルターまたはテーブルフィルターにより、キャッシュされたデータや非表示のデータがブックと共に保存される可能性がある。	フィルター
		非表示のデータのソースとなる可能性がある非表示の名前が含まれていることがある。	非表示の名前
③プレゼンテーション	Microsoft PowerPoint 2007 以降 (OOXML)	複数人が共同で作成したプレゼンテーションには、コメントやインク注釈などが含まれている可能性があり、他のユーザーが行った変更や、プレゼンテーションに取り組んだユーザーの名前、レビュー担当者からのコメントが含まれていることがある。	コメントおよびインク注釈
		クラウドに保存されている共有ドキュメントで共同作業を行った場合、ドキュメントに変更を加えたユーザーと変更が加えられた時期にかんするデータが含まれていることがある。	リビジョン追跡データ
		メタデータまたはドキュメントプロパティには、作成者、件名、タイトルなどの情報が含まれる。ドキュメントのプロパティには、プレゼンテーションを最後に保存したユーザーの名前、ドキュメントが作成された日付、ドキュメントの場所が自動的に保存される。特定の機能を使用した場合、電子メールヘッダー、校閲者への送信に関する情報、回覧用紙、Web ページを発行するためのファイルパス情報など個人を特定できる情報が含まれていることがある。	ドキュメントのプロパティと個人情報
		非表示に設定されているため表示されないオブジェクトが含まれていることがある。	スライド上の非表示の内容
		PowerPoint プレゼンテーションには、スライドからスライド外領域にドラッグされたため、すぐには表示されないオブジェクトが含まれていることがある。	スライドオフコンテンツ
		PowerPoint プレゼンテーションのノートセクションに情報が含まれていることがある。	プレゼンテーションノート
		ドキュメントワークスペースサイトや、Windows SharePoint Services に基づくライブラリなど、ドキュメント管理サーバー上の場所に保存されたプレゼンテーションには、サーバーの場所に関連するドキュメントプロパティや情報が含まれている場合がある。	ドキュメントサーバーのプロパティ
		文書自体では表示されないカスタム XML データが含まれていることがある。	カスタム XML データ

④画像	JPEG 2000, Tagged Image File Format (TIFF), Portable Network Graphics (PNG), Graphics Interchange Format (GIF), Joint Photographic Experts Group (JPEG)	JPEG, TIFF 形式の画像ファイルに Exif 情報(*6)が含まれていることがある。	撮影日時, 緯度/経度など
-----	---	--	---------------

(*1) 独自に作成された、XML (eXtensible Markup Language) 形式で記載されたプログラムデータ

(*2) Microsoft のプログラミング言語である VBA (Visual Basic for Applications) を用いた簡易プログラム

(*3) Office アプリケーションを拡張するための機能

(*4) コマンドボタンやコンボボックスなどの機能追加するためのソフトウェア部品

(*5) あらかじめ集計されたものを効率よく検索するためにキャッシュされたデータ

(*6) 画像データに撮影時の状況や設定などを表す様々な情報

3.2.3. 長期保存のリスク要因となる機能と点検方法(予備調査)

長期保存上のリスク要因となる機能の検出方法を表 3.2.3-1 に示す。フォーマット及び作成アプリケーションでサポートされない機能については表には記載せず削除している。また、画像ファイル、音声ファイル、動画ファイルについては長期保存上にリスク要因となるような機能を認めなかった。一方、文書、表計算およびプレゼンテーションにおいては、ファイルフォーマット毎にサポートされている機能が異なっているため、運用時にルール化することが望ましい。

表 3.2.3-1 長期保存上のリスク要因となる機能の検出方法

カテゴリ	フォーマット名	課題	機能	検出方法
①文書	Portable Document Format (PDF)	コンテンツの埋め込み	JavaScript, 動画, 音声	JavaScript は, [ツール]→[JavaScript]→[すべての JavaScript]で埋め込まれている JavaScript を確認できる。 動画および音声は, [コンテンツ]欄に埋め込まれているすべてのオブジェクトを確認できる。
		ファイル間のリンク	ファイルのリンク	[添付ファイル]欄に添付されているファイルを確認できる。
		フォントが見つからない, または間違っ て置換されている	フォントの埋め込み	[ファイル]メニューを開き, [プロパティ]を開くことで [文書のプロパティ]ダイアログが表示される。[フォント]タブを選択し表示されるフォント名の横に「(埋め込みサブセット)」と表示されている場合はフォントが埋め込まれている。
		デジタル著作権管理	権限パスワード (PDF)	[ファイル]メニューを開き, [プロパティ]を開くことで [文書のプロパティ]ダイアログが表示される。[セキュリティ]タブを選択し [セキュリティ方法]が「セキュリティなし」以外が選択されており, かつ「文書の印刷および編集を制限。これら の権限設定を編集するにはパスワードが必要」にチェックがされている場合に設定されていることが確認できる。
		暗号化	文書を開くパスワード	ファイルを開いた際にパスワードを要求される。
			電子署名, タイムスタンプ	ファイルを開くと電子署名およびタイムスタンプが付与されている場合には, 署名済みであることが文書上部に表示される。

	Microsoft Word 2007 以降 (OOXML)	コンテンツの埋め込み	マクロ, 動画, 音声	マクロが含まれる場合は, マクロ有効文書であり拡張子 docm で判定可能, かつマクロが含まれる場合にはファイルを開く際に警告が表示される。音声, 動画については, [レイアウト]タブの[オブジェクトの選択と表示]を選択すると表示されるリストで確認することができる。
		ファイル間のリンク	ファイルのリンク	[ファイル]メニューを開き, [情報]タブを選択する。[文書の検査]の[ドキュメント検査]を選択し, [埋め込みドキュメント]が検出されないことを確認する。
		フォントが見つからない, または間違っ て置換されている	フォントの埋め込み	[ファイル]メニューを開き, [情報]タブを選択する。[オプション]を開き, [保存]を選択し, [ファイルにフォントを埋め込む]がチェックされていることを確認する。
		デジタル著作権管理	Information Rights Management (IRM)	[校閲]メニューを開き, [編集の制限]を選択し[書式の制限]および[編集の制限]がチェックされていないことを確認する。または, [ファイル]メニューを開き, [情報]タブを選択する。[文書の保護]の中のメニューが選択されていることを確認する。
		暗号化	パスワードを使用して暗号化	ファイルを開いた際にパスワードを要求される。
			電子署名(デジタル署名)	文書が最終版であることが文書上部に表示される。
	Open Document Format for Office Applications (ODF) : Open Document Text	コンテンツの埋め込み	マクロ, 動画, 音声	マクロは LibreOffice, OpenOffice 共通で, ファイルを開いた際に警告が表示される。動画および音声は LibreOffice では, [ナビゲーター]欄および[ナビゲーター]ウィンドウで確認できる。OpenOffice では[ナビゲーター]欄および[ナビゲーター]ウィンドウに表示されないため, 個別に確認する必要がある。
		フォントが見つからない, または間違っ て置換されている	フォントの埋め込み(LibreOffice)	[ファイル]メニューを開き, [プロパティ]を選択すると表示される文書のプロパティダイアログで[フォント]タブを選択した際に[フォントの埋め込み]の配下のチェックボックスがチェックされている場合はフォントが埋め込まれている。
		暗号化	パスワード付きで保存する (OpenOffice, LibreOffice), GPG キーで暗号化する (LibreOffice)	ファイルを開いた際にパスワードを要求される。
			電子署名(デジタル署名)	ファイルを開いた際に画面上部の文書名の後に「(署名済み)」の表記, および画面下部に署名済みのアイコンが表示される。
		一太郎	コンテンツの埋め込み	ハイパーリンクを右クリックし, [ハイパーリンク変更]で表示される[ハイパーリンク]ダイアログで確認できる。
			暗号化	パスワード付きで保存する
			電子署名	文書を開くと, 文書上部に<署名済み・暗号化>と表示される。
②表計算	Microsoft Excel 2007 以降 (OOXML)	コンテンツの埋め込み	マクロ, 動画, 音声	マクロが含まれる場合は, マクロ有効ブックであり拡張子 xlsm で判定可能, かつマクロが含まれる場合にはファイルを開く際に警告が表示される。音声, 動画については, [ページレイアウト]タブの[オブジェクトの選択と表示]を選択すると表示されるリストで確認することができる。
		ファイル間のリンク	ファイルのリンク	[ファイル]メニューを開き, [情報]タブを選択する。[文書の検査]の[ドキュメント検査]を選択し, [埋め込みドキュメント]が検出されないことを確認する。

		フォントが見 つからない、 または間違っ て置換されて いる	フォントの埋め込 み	フォントの埋め込み機能は無い
		デジタル著作 権管理	Information Rights Management (IRM)	[校閲]メニューを開き、[編集の制限]を選択し[書式 の制限]および[編集の制限]がチェックされてい ないことを確認する。または、[ファイル]メニュ ーを開き、[情報]タブを選択する。[文書の保護]の中 のメニューが選択されていることを確認する。
		暗号化	パスワードを使 用して暗号化	ファイルを開いた際にパスワードを要求される。
			電子署名(デジ タル署名)	文書が最終版であることが文書上部に表示される。
	Open Document Format for Office Applications (ODF) : Open Document Spreadsheet	コンテンツの 埋め込み	マクロ、動画、音 声	[マクロ]は LibreOffice, OpenOffice 共通で、フ ァイルを開いた際に警告が表示される。 動画および音声は LibreOffice では、[ナビゲ ーター]欄および[ナビゲーター]ウインドウで確 認できる。 OpenOffice では[ナビゲーター]欄および[ナビ ゲーター]ウインドウに表示されないため、個別 に確認する必要がある。
		ファイル間の リンク	ファイルのリンク	外部リンクが設定されている場合に、シート上 部の枠外に「セキュリティ警告 外部リンクの自 動更新は無効になっています」が表示される。
		フォントが見 つからない、 または間違っ て置換されて いる	フォントの埋め込 み(LibreOffice)	[ファイル]メニューを開き、[プロパティ]を選 択すると表示される文書のプロパティダイア ログで[フォント]タブを選択した際に[フォ ントの埋め込み]の配下のチェックボックスが チェックされている場合はフォントが埋め込 まれている。
		暗号化	パスワードを使 用して暗号化	ファイルを開いた際にパスワードを要求される。
			電子署名(デジ タル署名)	ファイルを開いた際に画面上部の文書名の 後に「(署名済み)」の表記、および画面下部 に署名済みのアイコンが表示される。
③プレゼ ンテーショ ン	Microsoft PowerPoint 2007 以降 (OOXML)	コンテンツの 埋め込み	マクロ、アニメ ーション、音声	マクロが含まれる場合は、マクロ有効プレ ゼンテーションであり拡張子 pptm で判定可 能、かつマクロが含まれる場合にはファイル を開く際に警告が表示される。 アニメーションは、[アニメーション]タブの [アニメーション ウインドウ]を選択するとリス トが表示される。 音声、動画については、[ホーム]タブの[配 置]から[オブジェクトの選択と表示]を選択す ると表示されるリストで確認することができる。
		ファイル間の リンク	ファイルのリンク	[ファイル]メニューを開き、[情報]タブを選 択する。 [文書の検査]の[ドキュメント検査]を選 択し、[埋め込みドキュメント]が検出されな いことを確認する。
		フォントが見 つからない、 または間違っ て置換されて いる	フォントの埋め込 み	[ファイル]メニューを開き、[情報]タブを選 択する。 [オプション]を開き、[保存]を選択し、[フ ァイルにフォントを埋め込む]がチェックされ ていることを確認する。
		デジタル著作 権管理	Information Rights Management (IRM)	[校閲]メニューを開き、[編集の制限]を選 択し[書式の制限]および[編集の制限]がチェ ックされていないことを確認する。または、 [ファイル]メニューを開き、[情報]タブを選 択する。[文書の保護]の中のメニューが選 択されていることを確認する。
		暗号化	パスワードを使 用して暗号化	ファイルを開いた際にパスワードを要求さ れる。
			(PowerPoint), 文 書を開くパスワ ード(PDF), パス	

			ワード付きで保存する (OpenOffice)	
			電子署名(デジタル署名)	文書が最終版であることが文書上部に表示される。
	Open Document Format for Office Applications (ODF) : Open Document Presentation	コンテンツの埋め込み	マクロ, アニメーション, 音声	[マクロ]は LibreOffice, OpenOffice 共通で, ファイルを開いた際に警告が表示される。 動画および音声は LibreOffice では, [ナビゲーター]欄および[ナビゲーター]ウインドウで確認できる。 OpenOffice では[ナビゲーター]欄および[ナビゲーター]ウインドウに表示されないため, 個別に確認する必要がある。 アニメーションについては, LibreOffice では, [表示]メニューの[アニメーション]を選択すると表示されるリストで確認できる。OpenOffice では, [スライドショー]メニューの[アニメーション設定]を選択すると表示されるリストで確認することができる。
		フォントが見つからない, または間違っ て置換されている	フォントの埋め込み	[ファイル]メニューを開き, [プロパティ]を選択すると表示される文書のプロパティダイアログで[フォント]タブを選択した際に[フォントの埋め込み]の配下のチェックボックスがチェックされている場合はフォントが埋め込まれている。
		暗号化	パスワード付きで保存する	ファイルを開いた際にパスワードを要求される。
			電子署名(デジタル署名)	ファイルを開いた際に画面上部の文書名の後に「(署名済み)」の表記, および画面下部に署名済みのアイコンが表示される。

3.3. 標準的フォーマットで既に作成されている場合の対処

3.3.1 フォーマット変換

長期保存するにあたって, 標準的フォーマット以外のフォーマットについては標準的フォーマットに変換することが望ましいため, フォーマット変換の方法, 変換結果について検証した。

3.3.1.1. フォーマット変換の検証対象フォーマット

変換の検証対象フォーマットを表 3.3.1.1-1 に示す。

表 3.3.1.1-1 フォーマット変換対象

○：変換先候補、－：変換先として適切でない

No.	変換元フォーマット		変換先フォーマット (標準的フォーマット)									
	フォーマット	バリエーション	PDF 1.4	PDF 1.5	PDF 1.6	PDF 1.7	PDF 2.0	PDF/A-1	PDF/A-2	PDF/A-3 ※1	PDF/A-4	OOXML Transitional
1	一太郎	JTD/JFW/JBW/JAW/JSW	○	○	○	○	○	○	○	－	○	○
2	Microsoft Word	97-2003	○	○	○	○	○	○	○	－	○	○
3	Microsoft Excel	97-2003	○	○	○	○	○	○	○	－	○	○
4	Microsoft Excel	2007 (マクロ有効ブック)	○	○	○	○	○	○	○	－	○	○
5	Microsoft PowerPoint	97-2003	○	○	○	○	○	○	○	－	○	○

※1：PDF/A-2に添付ファイルを付けた場合の規格の為、PDF/A-3が直接変換対象になることは無い

3.3.1.2. フォーマット変換方法

フォーマット変換は自動化することを前提とし, 変換ツールを作成して実証実験を行った。
フォーマット変換ツールの構成は以下の通りである。マイクロソフトが提供するツール類を使用した「フォーマット変換ツール」を作成して変換を行った。

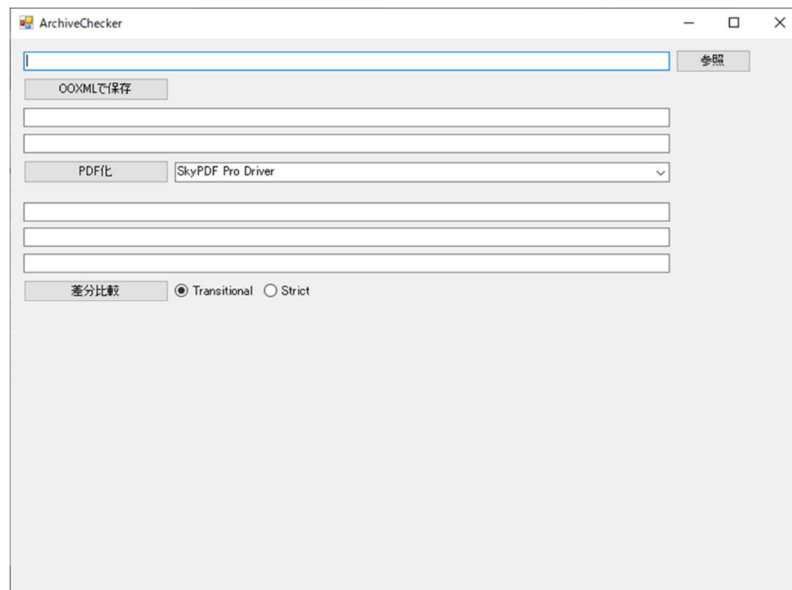


図 3.3.1.2-1 フォーマット変換ツール画面イメージ

フォーマット変換は以下の流れで作成される。

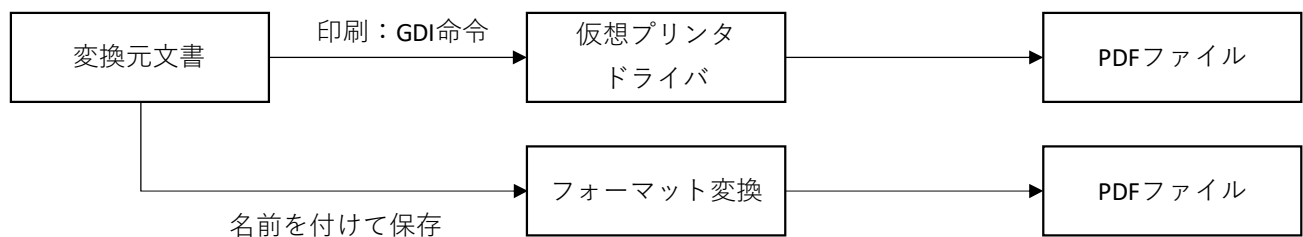


図 3.3.1.2-2 フォーマット変換フロー

「OOXML で保存」ボタンで指定されたファイルを以下のフローに基づき変換元文書から変換先文書へ変換保存を行う。以下の処理で変換可能な文書は Office 文書のみで一太郎は Office 文書への自動変換はできず、手動のみとなる。Office 文書の変換には「Microsoft Office Professional 2021 Plus」を使用している。

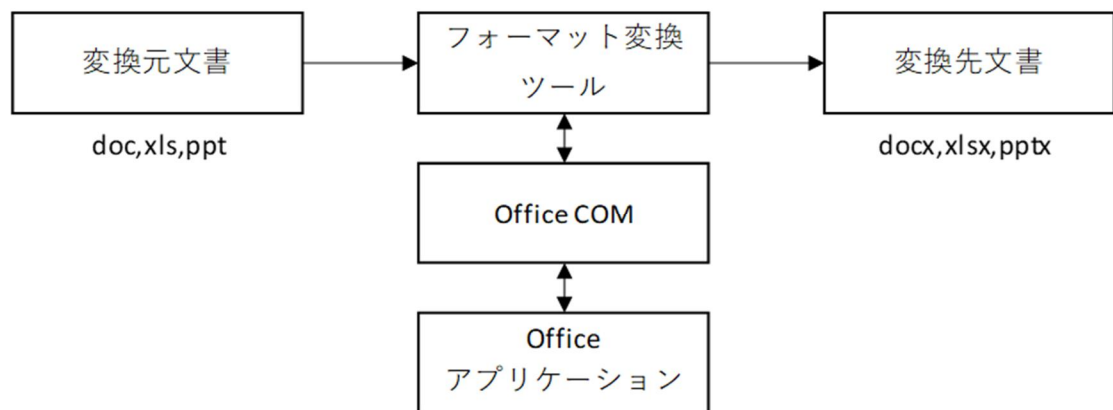


図 3.3.1.2-3 Office文書のフォーマット変換を自動化

「PDF 化」ボタンで各文書を以下の処理フローに基づき PDF に変換する。変換元文書を指定された仮想プリンタドライバへ印刷することで PDF ファイルが作成される。

仮想プリンタドライバは「Adobe Acrobat」「JUSTPDF」「SkyPDF」を使用した。

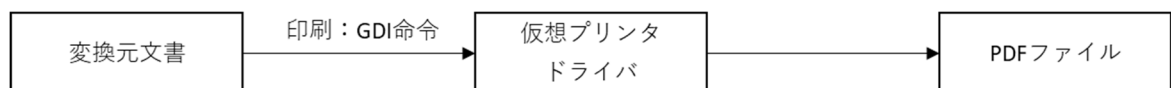


図 3.3.1.2-3 Office文書の PDF 変換フロー

参考としてフォーマット変換が可能なサーバ製品の一覧を表 3.3.1.2-1 に示す。

※注記 サーバ製品を利用した実証試験は行っていない。(利用可能な製品候補として一覧作成となる)

表 3.3.1.2-1 PDF へのフォーマット変換が可能なサーバ製品の一覧

製品名	SkyPDF Web Server	Antenna House PDF Server	Office Server Document Converter
バージョン	9	3.5	7.1
開発元	株式会社スカイコム	アンテナハウス	アンテナハウス
生成方式	Office利用	Office利用	ファイルコンバート
説明	クライアントからWebブラウザーを操作することでOfficeファイルやその他ドキュメントを高速・高圧縮・高品質のPDFファイルへ変換する機能を持つ、オールインワンパッケージのサーバソフトウェアです。	『PDF Server』はサーバ-OSで動作するファイル変換・編集サーバソリューションソフトウェアです。オフィス文書や画像ファイルなどを、あらかじめ設定した変換設定を利用して自動でPDF等に変換/出力します。また、OCR処理による透明テキスト付きPDF(サーチャブルPDF)の生成も可能。	Office Server Document Converter (OSDC) は、Microsoft Word/Excel/PowerPoint、及びRTF文書をサーバサイドでPDF/SVG/TIFF/JPEG/PNG/TIFF/INX/XPSファイルへ変換するライブラリです。グループウェアやWebメールシステム等でOSDCを利用することにより、Microsoft Office がインストールされていない、および動作しない環境でもMicrosoft Office文書、及びRTF文書を簡単に変換することが可能です。
インターフェース	SOAP I/F フォルダ監視	コマンドライン フォルダ監視	コマンドライン 各種言語用API

※Office アプリケーションに関する注意事項

- ・サーバ側で Office アプリケーションを実行するタイプの製品に関しては Office ライセンス条項に留意する
- ・Office アプリケーションを利用しないモデルの場合は, PDF 変換品質は Office での表示結果と同一にならない場合がある

3.3.1.3. フォーマット変換結果の確認方法

変換結果の確認, 変換前と変換後のファイルの表示内容および各情報の保持を目視で確認する。今後変換前後のチェックを自動化することも考慮して以下の方法を用いる。

1. 変換前ドキュメントを PDF ファイルに変換する
2. 変換後ドキュメントを PDF ファイルに変換する
3. 上記1. 2. の PDF ファイルを画像ファイル (PNG 形式) に出力する
4. 上記1. 2. の PDF ファイルからテキストデータを出力する
5. 3 で出力された画像データをピクセル単位で比較し完全一致するか判定する。その際に差分画像を出力する
6. 4 で出力されたテキストが同一になるかを判定した

データ抽出 (画像抽出、TEXT ファイル) には「SkyPDF Tools SDKⅢ」を使用し、比較ツールとして「WinMerge」を使用して確認を行った。

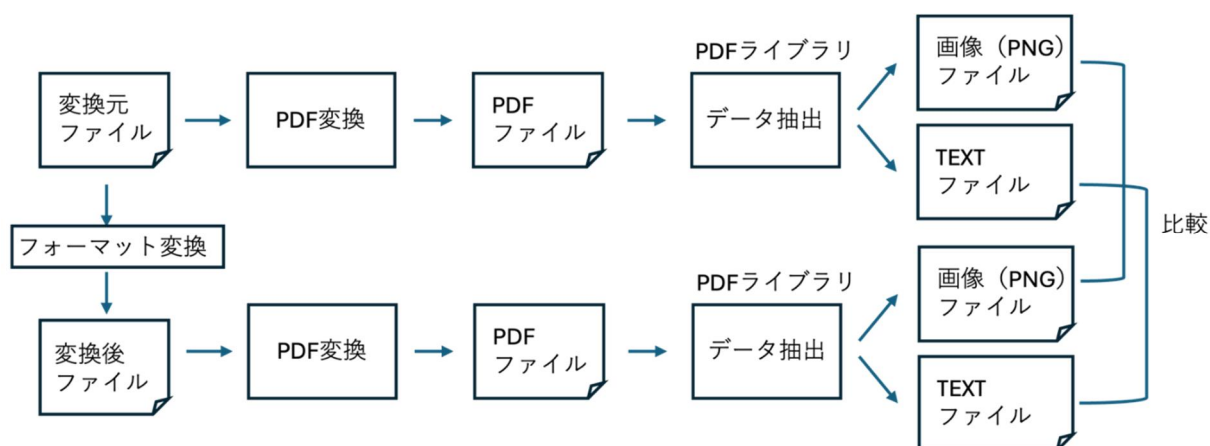


図 3.3.1.3-1 変換結果の確認方法

3.3.1.4. フォーマット変換の検証環境

フォーマット変換の検証環境を図 3.3.1.4-1 に示す。検証用ファイルと試験ツール以外は市販のものを利用した。

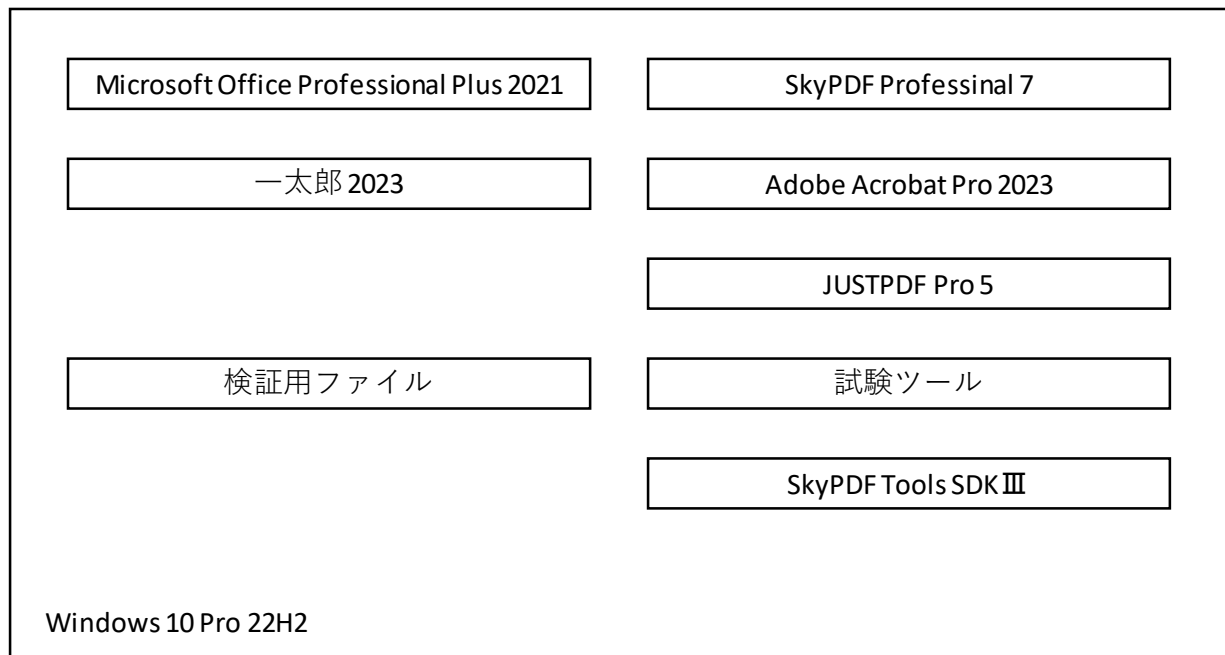


図 3.3.1.4-1 フォーマット変換の検証環境

フォーマット変換を行った PC のスペックを以下に示す。

OS : Windows 10 Pro 22H2

CPU : 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz

メモリー : 16.00 GB

ディスク : 50GB

3.3.1.5 フォーマット変換の結果

各フォーマットの実証実験を行った結果を表 3.3.1.5-1 に示す。詳細は別紙 3(フォーマット変換結果)を参照。

表 3.3.1.5-1 フォーマット変換及び検証の結果

変換前	変換後	自動 変換可否	自動検証	変換後の検証結果
一太郎 (4～12, 2004～2008) 拡張子: jtd, jfw, jbw, jaw, jsw	PDF/A-1～4 PDF1.4～2.0	△注 3	一部可	注釈, 変更履歴, 隠し文字は設定によって出力されない 外字・異体字及び機種依存文字が PDF/A に変換すると維持された
	Word 2007 以降 (OOXML)	×		
Word 97～2003 拡張子: doc	PDF/A-1～4 PDF1.4～2.0	○	一部可	注釈, 変更履歴, 隠し文字は設定によって出力されない 外字・異体字及び機種依存文字が PDF/A に変換すると維持された

	Word 2007 以降 (OOXML)	○		微細だがレイアウトがずれる 手動でフォントの埋め込みで OOXML でも外字・異 体字及び機種依存文字が維持された
Excel 97- 2003 拡張子:xls	PDF/A-1~4 PDF1.4~2.0	○		注釈, 変更履歴, 隠し文字は設定によって出力され ない 外字・異体字及び機種依存文字が PDF/A に変換 すると維持された
	Excel 2007 以降 (OOXML)	○		差異なし
Excel 2007 (マクロ有効ブ ック) 拡張子:xlsm	PDF/A-1~4 PDF1.4~2.0 注1	△		・注釈, 変更履歴, 隠し文字は設定によって出力され ない ・マクロは移行されない 外字・異体字及び機種依存文字が PDF/A に変換 すると維持された
	Excel 2007 以降 (OOXML) 注 2	△		マクロ機能は移行されるが動作保証されない
PowerPoint 97-2003 拡張子:ppt	PDF/A-1~4 PDF1.4~2.0	○		・注釈は設定によって出力されない ・アニメーションは最終状態となる 外字・異体字及び機種依存文字が PDF/A に変換 すると維持された
	PowerPoint 2007 以降 (OOXML)	○		差異なし 手動でフォントの埋め込みで OOXML でも外字・異 体字及び機種依存文字が維持された

注1 マクロは移行されない

注2 マクロの動作は保証されない

注3 複数シートの場合, 自動変換ができない

3.3.1.6. 複数変換先の選択判断

その他確認すべきフォーマットから標準的フォーマットへ直接変換する場合に, 変換先が複数ある場合の選択判断基準を以下に示す。

- 一太郎ファイルは Word(OOXML)に変換できないため, PDF/A ファイルに変換する。
- すべての文書に於いて外字・異体字及び機種(環境)依存文字を使用している場合は PDF/A 形式に変換する。ただし, 手動でフォントの埋め込みができる Word 及び PowerPoint については, OOXML 形式に変換する選択肢もある。
- Office 文書は再利用を行う場合は OOXML形式で保存することが推奨されるが, Word に関しては若干の変換誤差が生じるため, 許容できない場合は国際標準化されている PDF 1.7 をベースとしている PDF/A-2 への変換が推奨される。

表 3.3.1.6-1 複数変換先の選択判断

No.	変換元		変換先									
	フォーマット	バリエーション	PDF 1.4	PDF 1.5	PDF 1.6	PDF 1.7	PDF 2.0	PDF/A-1	PDF/A-2	PDF/A-3	PDF/A-4	OOXML Transitional
1	一太郎	JTD/JFW/JBW/JAW/JSW				○	○	○	◎			
2	PDF	1.4~2.0 PDF/A-1~4	○	○	○	○	○	○	○	○	○	
3	PDF	外字、異体字が含まれる						○	◎	○	○	
4	Microsoft Word	97-2003				○	○	○	○			◎
5	Microsoft Excel	97-2003				○	○	○	○			◎
6	Microsoft Excel	2007 (マクロ有効ブック)				○	○	○	◎			
7	Microsoft PowerPoint	97-2003				○	○	○	○			◎
8	Officeファイル全般	外字、異体字が含まれる						○	◎			

○: 変換先として妥当, ◎: 変換先として推奨

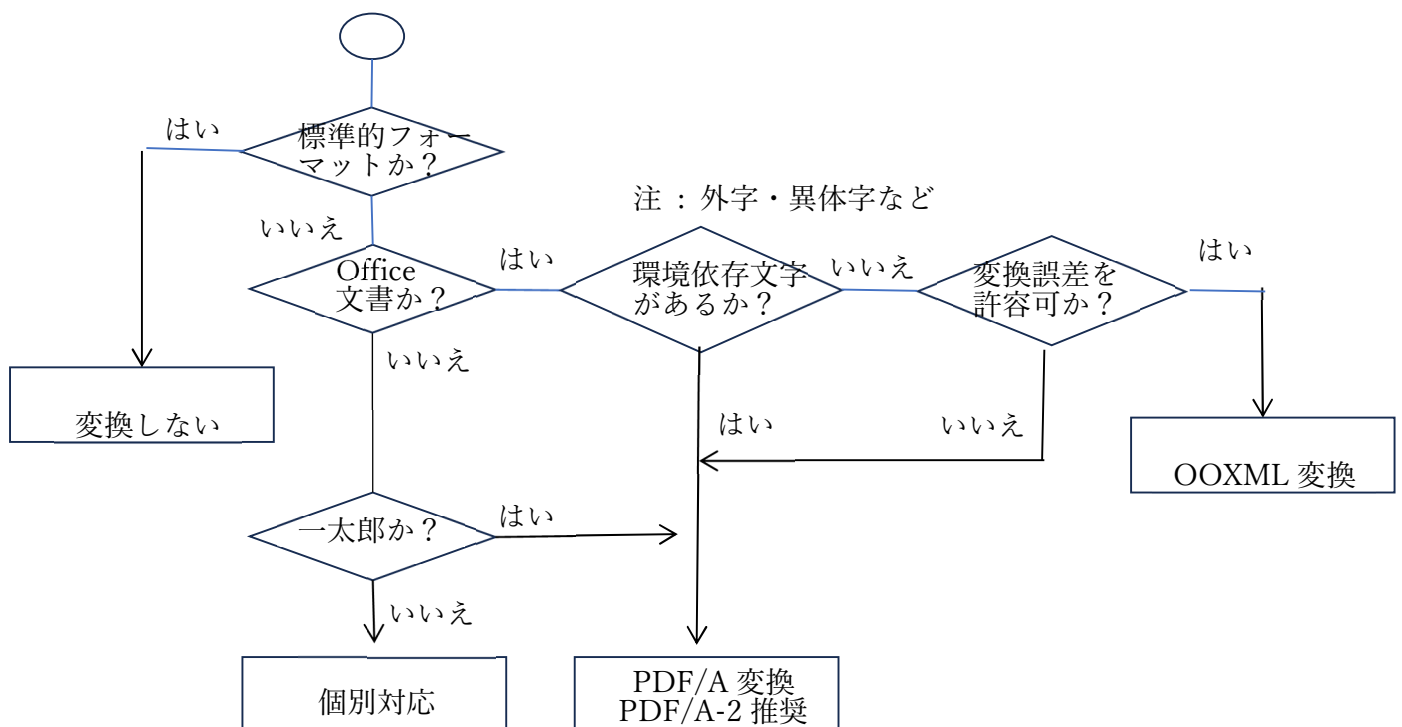


図 2.3-1 複数変換先の選択判断

4. まとめ

4.1. 標準的フォーマット案

本調査検討及び総合的な判断の結果、標準的フォーマット案候補のうち、表 4.1-1 のフォーマットを「標準的フォーマット案」として策定した。

表 4.1-1 標準的フォーマット案

リスクが低い フォーマット名	バリエーション	拡張子	カテゴリ	評価 点 合計	評価コメント	採否
Portable Document Format (PDF)	PDF/A-1	pdf	文書・ プレゼン テーショ ン	10	ISO19005-1。PDF 1.4 をベースに見読性の維持に適さない機能を制限。フォント埋込必須	○
	PDF/A-2			11	ISO19005-2。PDF 1.7 をベースに見読性の維持に適さない機能を制限。フォント埋込必須。長期保存にあたり安定しているフォーマット	○
	PDF/A-3			11	ISO19005-3。PDF/A-2 に加えて、多様なファイル (PDF/A への変換前のファイルなど) の埋込みが可能となったが、埋込みファイルの見読性やセキュリティなどに考慮が必要	
	PDF/A-4			9	ISO19005-4。PDF 2.0 をベースに見読性の維持に適さない機能を制限している。普及は限定的	
	1.7			11	ISO32000-1。広く普及している。一般的なソフトで作成可能。フォント埋込が必須ではない	○
	2.0			10	ISO32000-2。最新の PDF フォーマットであるが、普及は限定的	
Office Open XML (OOXML)	Word processing, transitional	docx	文書	11	ISO-IEC29500。国際標準化されており、一般的なオフィスソフトで作成可能 (※1)	○
Open Document Format for Office Applications (ODF)	OpenDocument Text	odt		10	ISO-IEC26300。無償で提供されているソフトで作成可能だが普及していない (※2)	
Office Open XML (OOXML)	Spreadsheet, transitional	xlsx	表計算	11	(※1)と同様	○
Open Document Format for Office Applications (ODF)	OpenDocument Spreadsheet	ods		10	(※2)と同様	
Office Open XML (OOXML)	Presentation, transitional	pptx	プレゼン テーショ ン	11	(※1)と同様	○
Open Document Format for Office Applications (ODF)	OpenDocument Presentation	odp		10	(※2)と同様	
JPEG2000		jp2	画像	12	ISO-IEC15444-1。可逆圧縮で圧縮品質が選択可能。長期保存の目的で広く利用	○
Portable Network Graphics (PNG)	1.2	png		10	ISO-IEC15948。可逆圧縮に対応。広く普及	○
JPEG		jpg		12	ISO-IEC10918。非可逆圧縮だが圧縮品質が選択可能で、広く普及	○
Tagged Image File Format (TIFF)	6	tif, tiff		11	仕様は公開されているが、国際標準ではない。可逆圧縮で圧縮品質が選択可能。	

Graphics Interchange Format (GIF)		gif		9	仕様は公開されているが、国際標準ではない。可逆圧縮に対応。解像度は低い、データサイズが小さくなるため普及している	
WAVE (Waveform Audio File Format)		wav	音声 動画	10	非圧縮方式であり、幅広く利用されているが、国際標準ではない。	
MP3 (MPEG Layer III)		mp3		12	ISO-IEC11172-3。非可逆圧縮だが圧縮品質が選択可能で、高音質。様々なデバイスで再生可能のため広く普及	○
MPEG-4 Part14 (MP4)		mp4		12	ISO-IEC14496-14。非可逆圧縮だが、圧縮品質が選択可能で、高音質。広く普及	○
MPEG-2		mpg, mpeg		12	ISO-IEC13818-1。非可逆圧縮だが、圧縮品質が選択可能で、高音質。広く普及	○

4.2 定期的なフォーマットのリスク評価のためのモデル

技術の進歩や環境の変化, 特に, b 及び c の理由から, 定期的な評価の見直しが必要, また, 見直し時は, a を考慮する。見直し時のフローを図 4.2-1 (図 2.2.2-2 を再掲) に示す。

- 取り扱うフォーマット数の桁が増えると級数的に運用負荷が増加し, 全く別の運用方法を考えなければならなくなるため, 現状通り, カテゴリごとに必要最小限 (理想的には1つ) のフォーマットに限定することが望ましい
- 標準化から普及までは 10 年以上かかる。とはいえ, 音声, 動画のカテゴリのフォーマット技術は日々進歩していることからシステム更改を考慮し, 計画的な見直しを検討する必要がある。
- 設計に CAD が広く使われており, 建築等では長期保存が必要なため, PDF への変換も含め, 今後, フォーマットリストに CAD を追加するか否かを検討する必要がある。

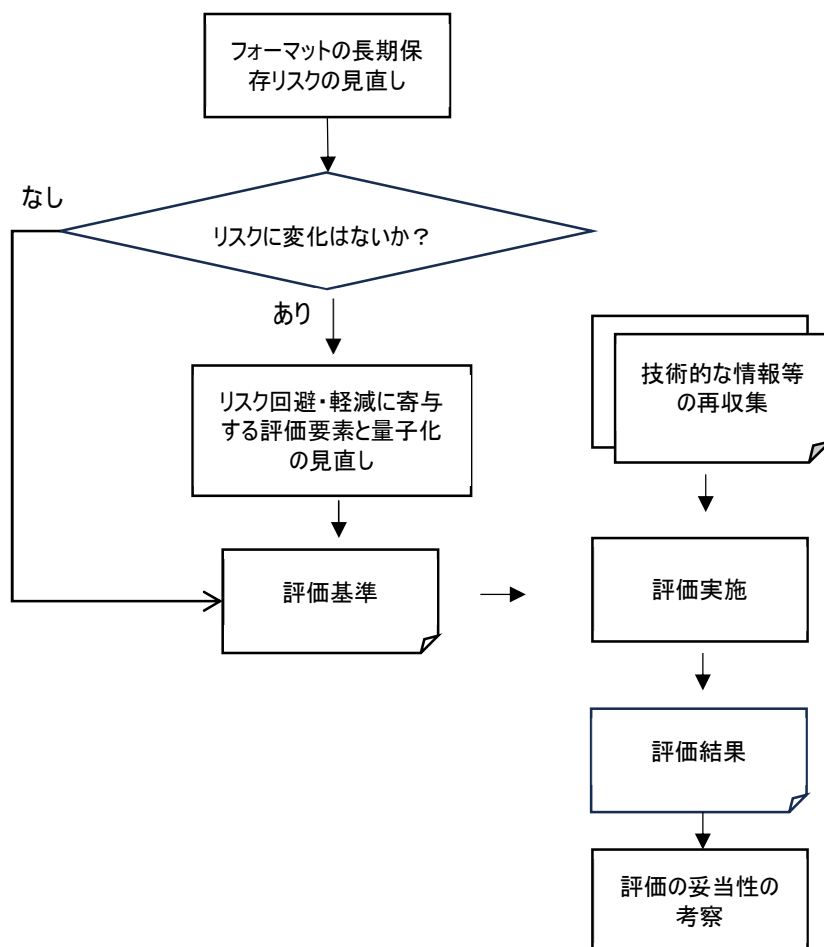


図 4.2-1 電子公文書等のフォーマットの持続性リスクの評価の全体像

4.3 標準的フォーマットによる作成・保存・利用のためのモデル

標準的フォーマットによる作成・保存・利用にあたって, 文書管理システムを利用した場合の以下に示す文書ライフサイクルを考慮した運用フローを図 4.3-1 及び図 4.3-2 に示す。

- 文書発生時(作成・収集・受領)タイミングでのフォーマット判定
- すでに文書管理システム内にて保存されている文書

- アーカイブ対象として保存されている文書

上記を発生したタイミング及び、クローリングしながらの定期的なチェックを自動的に行えることが望ましい。

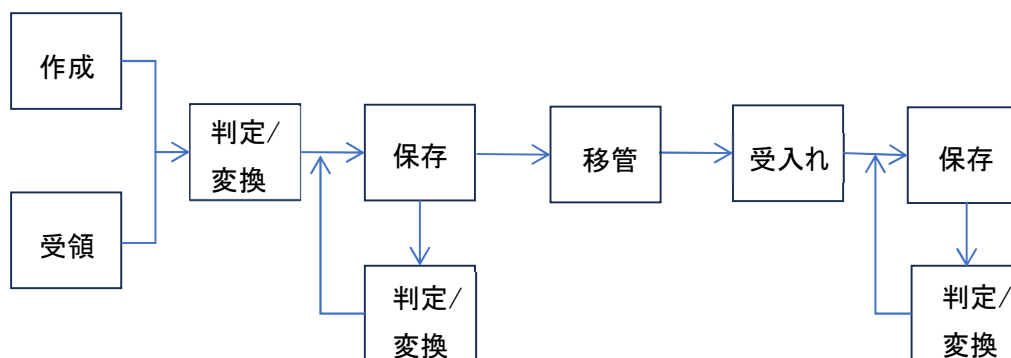


図 4.3-1 文書ライフサイクルと運用フロー

上記図内のフォーマット判定／変換部分の詳細を図式化したものを以下に示す。

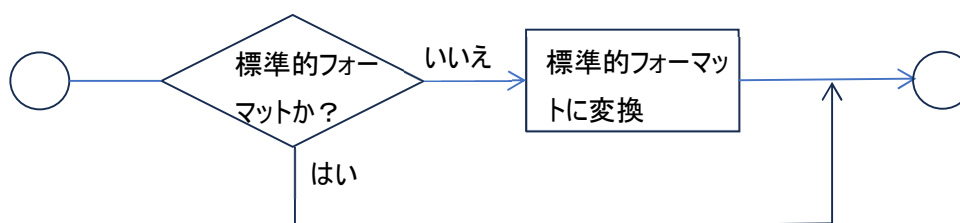


図 4.3-2 フォーマット判定と変換処理

4.4 今後の検討課題

3.2.2 で「意図しない情報の混入パターンと点検方法(予備調査)」, 3.2.3 で「長期保存のリスク要因となる機能と点検方法(予備調査)」を実施した結果を踏まえた今後の検討課題を表 4.4-1 に示す。

表 4.4-1 今後の検討課題検討課題

今後の検討課題検討課題	内容
「標準的フォーマット」で移管, 保存したものを利用に供する方法の検討	- 利用制限情報のマスキング方法 - 電子的に紛れ込んでしまう情報の検出方法の確認は, 手動操作によるものだったため, それぞれのフォーマットにおいて, 紛れ込んでしまう情報を自動検出する方法の確認方法
「標準的フォーマット」による保存等の留意事項に関する技術検証	特定の機能(マクロ, 関数, 外字, 異体字等)を使用していた場合の「標準的フォーマット」による長期保存の安定性を損なうマクロの削除や外字・異体字を使用している場合のフォントの埋め込みなどを自動でおこなう方法

別紙1 長期保存の観点からリスクが低いフォーマットに関する調査結果

通番	1
フォーマット名	Portable Document Format (PDF)
バージョン	1.7
拡張子	pdf
公開日	2008-07
最終更新日	2008-07
バージョンの更新履歴	2008
標準化対応状況	国際標準として制定 (ISO 32000-1)
特長	PDF1.7 は, Adobe Systems Incorporated によって開発された Portable Document Format 仕様のバージョン 1.7 を意味する。PDF1.7 は, ISO 32000-1 として国際標準化された最初の PDF 仕様であり, テキスト, 画像, ビデオ, オーディオなどのさまざまなコンテンツを一つのファイルに統合できる柔軟性がある。PDF1.7 は, 文書の属性やメタデータをタグとして埋め込むことができ, これにより, 文書のカタログ化や検索などが容易になる。PDF1.7 は, Windows, Linux, macOS などの主要な OS で使用でき, また, 多くの文書閲覧ソフトや編集ソフトでも対応している。
課題	特になし
開発元	Adobe Systems Inc. (現 Adobe Inc.)
ライセンス	特許の有無: 無
メタデータの埋込可否	PDF のメタデータは, Adobe Systems によって開発された, XML ベースの XMP (Extensible Metadata Platform) 形式で保存される。メタデータには, PDF 文書のタイトル, 作成者, 作成日時, 変更日時などの情報を含む
参照情報	https://www.iso.org/standard/51502.html
PRONOM ID	fnt/276
NARA における ID	NF00369
作成ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	2
フォーマット名	Portable Document Format (PDF)
バージョン	2.0
拡張子	pdf
公開日	2017-07
最終更新日	2020-12
バージョンの更新履歴	2017, 2020
標準化対応状況	国際標準として開発 (ISO 32000-2)
特長	PDF2.0 は, 国際標準化機構 (ISO) が定める PDF の最新バージョンである。PDF1.7 から 10 年ぶりにリリースされた。PDF2.0 は, PDF 言語の進化と洗練を目指し, すでに使用されなくなった機能を廃止し, adobe の独自仕様を国際標準に整合化させた。PDF2.0 では, 3D アートワークやマルチメディアファイルの表示や再生が強化され, また, 暗号化やデジタル署名などのセキュリティ機能も改善された。 また, XML や CSV などの任意のファイル形式を PDF ファイル内に埋め込むことができ, これにより, 文書の構造化やデータ交換などが容易になる。

課題	2020 年に標準化され、現時点で実装製品は限定されている。 また、PDF2.0 は、PDF1.7 とは非互換であるといっても差し支えないほど変更されているため、PDF2.0 をサポートしない PDF 処理ソフトでは開くことができない場合がある。
開発元	ISO Tc171/SC2
ライセンス	特許の有無:無
メタデータの埋込可否	PDF のメタデータは、Adobe Systems によって開発された、XML ベースの XMP (Extensible Metadata Platform) 形式で保存される。メタデータには、PDF 文書のタイトル、作成者、作成日時、変更日時などの情報を含む
参照情報	https://www.iso.org/standard/75839.html
PRONOM ID	fnt/1129
NARA における ID	NF00662
作成ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	3
フォーマット名	PDF
バージョン	A-1
拡張子	pdf
公開日	2005-10
最終更新日	2005-11
バージョンの更新履歴	2005, 2007/Cor1, 2011/Cor2
標準化対応状況	国際標準として開発 (ISO 19005-1)
特長	PDF/A シリーズは、電子文書の長期保存を目的とした PDF 規格である。PDF/A では、フォントや画像などの外部リソースをすべて PDF ファイル内に埋め込むことで、文書の表示や印刷に必要な情報を保持し。これにより、文書が環境や時間によって変化することを防ぐ。 PDF/A-1 は、PDF1.4 をベースに、ISO 19005-1 として国際標準化されたプロファイル仕様 (使い方を規定した仕様) である。PDF/A-1 には、PDF/A-1b と PDF/A-1a の二つのコンFORMANCE レベルがある。PDF/A-1b (b は basic を意味する) では、暗号化や JavaScript などの機能を禁止することで、文書の互換性を高め、これにより、文書が様々な PDF ビューアーやアプリケーションで正しく開けることを保証する。PDF/A-1a (a は accessibility を意味する) では、タグやメタデータなどの構造情報を付与することで、文書のアクセシビリティを向上させ、これにより、文書がスクリーンリーダーや音声合成ソフトなどで読み上げられることを可能にしている。
課題	初期の実装において、透明効果に問題が発生する場合があることが報告されている。 https://blog.adobe.com/jp/publish/2022/02/09/dc-20-years-of-transparency-in-pdf
開発元	Adobe Inc.
ライセンス	特許の有無:無
メタデータの埋込可否	PDF のメタデータは、Adobe Systems によって開発された、XML ベースの XMP (Extensible Metadata Platform) 形式で保存される。メタデータには、PDF 文書のタイトル、作成者、作成日時、変更日時などの情報を含む
参照情報	https://www.iso.org/standard/38920.html
PRONOM ID	fnt/354
NARA における ID	NF00372

作成ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	4
フォーマット名	PDF
バージョン	A-2
拡張子	pdf
公開日	2011-07
最終更新日	2011-07
バージョンの更新履歴	2011
標準化対応状況	国際標準として開発 (ISO 19005-2)
特長	PDF/A-2 は、PDF1.7 をベースに ISO 19005-2 として国際標準化された。JPEG 2000 や PDF ポートフォリオ機能をサポートしている。PDF ポートフォリオとは、複数の PDF 文書やその他のファイル形式を一つのファイルにまとめることができる機能をいう。PDF/A-2 では、透明効果やレイヤーなどの PDF1.7 で追加された機能が利用できる。これにより、よりリアルな画像表現やデザインが可能になった。PDF/A-2 には PDF/A-2b, PDF/A-2a, PDF/A-2u の 3 つのコンFORMANCE レベルがある。PDF/A-2 の、文書の完全性・互換性・アクセシビリティに関する要件は PDF/A-1 と同じである。PDF/A-2u (u は Unicode を意味する) は、テキストが Unicode でエンコードされていることを要求する。
課題	すべての PDF ビューアーや編集ソフトウェアが PDF/A-2u で使用される特定の Unicode エンコーディングに完全に対応しているわけではない。特に古いまたは限定的な機能を持つソフトウェアでは、PDF/A-2u 文書の正確なレンダリングやテキストの抽出に問題が生じることがある。
開発元	ISO TC171/SC2
ライセンス	特許の有無: 無
メタデータの埋込可否	PDF のメタデータは、Adobe Systems によって開発された、XML ベースの XMP (Extensible Metadata Platform) 形式で保存される。メタデータには、PDF 文書のタイトル、作成者、作成日時、変更日時などの情報を含む
参照情報	https://www.iso.org/standard/50655.html
PRONOM ID	fnt/477
NARA における ID	NF00634
作成ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	5
フォーマット名	PDF
バージョン	A-3
拡張子	pdf
公開日	2012-10
最終更新日	2012-10
バージョンの更新履歴	2012
標準化対応状況	国際標準として開発 (ISO 19005-3)
特長	PDF/A-3 は、PDF1.7 をベースに ISO 19005-3 として国際標準化された。PDF/A-3 には PDF/A-3b, PDF/A-3a, PDF/A-3u の 3 つのコンFORMANCE レベルがある。

	スレベルがあるが、PDF/A-2 との違いは、PDF/A-2 が、PDF ファイルのみ埋込める (embedded) のに対して、任意のファイルを埋込める点だけである。
課題	PDF/A-3u については、すべての PDF ビューアーや編集ソフトウェアが使用される特定の Unicode エンコーディングに完全に対応しているわけではない。特に古いまたは限定的な機能を持つソフトウェアでは、正確なレンダリングやテキストの抽出に問題が生じることがある。 また、埋め込むファイルは PDF 仕様のスコープ外であり、長期保存の保証対象外である。
開発元	ISO TC171/SC2
ライセンス	特許の有無: 無
メタデータの埋込可否	PDF のメタデータは、Adobe Systems によって開発された、XML ベースの XMP (Extensible Metadata Platform) 形式で保存される。メタデータには、PDF 文書のタイトル、作成者、作成日時、変更日時などの情報を含む
参照情報	https://www.iso.org/standard/57229.html
PRONOM ID	fnt/480
NARA における ID	NF00644
作成ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	6
フォーマット名	PDF
バージョン	A-4
拡張子	pdf
公開日	2020-11
最終更新日	2020-11
バージョンの更新履歴	2020
標準化対応状況	国際標準として開発(ISO 19005-4)
特長	PDF/A-4 は、PDF2.0 をベースに、ISO 19005-4 として国際標準化された。PDF/A-4 では、PDF 2.0 で追加された機能が利用できる。例えば、暗号化やデジタル署名の強化、画像やテキストの透明効果やレイヤーのサポート、OpenType フォントや JPEG 2000 形式の画像圧縮のサポート、XML や CSV などの任意のファイル形式を PDF ファイル内に埋め込むことができる。 PDF/A-4 には PDF/A-4(b という表示はなくなった)、PDF/A-4a、PDF/A-4u の 3 つのコンFORMANCEレベルがある。PDF/A-4 の、文書の完全性・互換性、アクセシビリティ、ユニコードに関する要件は PDF/A-3 と同じである。
課題	PDF/A-4u については、すべての PDF ビューアーや編集ソフトウェアが使用される特定の Unicode エンコーディングに完全に対応しているわけではない。特に古いまたは限定的な機能を持つソフトウェアでは、正確なレンダリングやテキストの抽出に問題が生じることがある。 また、埋め込むファイルは PDF 仕様のスコープ外であり、長期保存の保証対象外である。
開発元	ISO TC171/SC2
ライセンス	特許の有無: 無
メタデータの埋込可否	PDF のメタデータは、Adobe Systems によって開発された、XML ベースの XMP (Extensible Metadata Platform) 形式で保存される。メタデータには、PDF 文書のタイトル、作成者、作成日時、変更日時などの情報を含む

参照情報	https://www.iso.org/standard/71832.html
PRONOM ID	fnt/481
NARA における ID	NF00646
作成ソフトウェア名	SkyPDF
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	7, 8
フォーマット名	Office Open XML(OOXML) Word processing
バージョン	Word2007 以降
拡張子	docx
公開日	2008-12
最終更新日	2021-08
バージョンの更新履歴	ISO/IEC 29500-1, -2, -3, -4:2008 ISO/IEC 29500-1, -2, -3,-4:2011 ISO/IEC 29500-1, -2, -3, -4:2012 ISO/IEC 29500-3:2015 ISO/IEC 29500-1, -4:2016 ISO/IEC 29500-2:2021
標準化対応状況	国際標準として制定
特長	Office Open XML は, Microsoft Word の文書を表現したものであり, ISO/IEC 29500 シリーズとして標準化されている。ISO/IEC 29500 の作成時点における目標は Microsoft Office アプリケーション(Microsoft Office 97 から Microsoft Office 2008)によって作成された Word 文書, Excel シート, および PowerPoint プレゼンテーションにおける既存の表現を忠実に再現できることであった。また Office Open XML の提供者と利用者の要件も規定しており, 複数のベンダーや複数のプラットフォームにおいての実装を可能にすることで, 拡張性と相互運用性を促進することも目標としている。 Office Open XML(OOXML) Word-processing は, Office Open XML を使って Microsoft Word を表現したものである。 なお, ISO/IEC 29500 では, strict(厳密)と transitional(過渡的)の2つのコンFORMANCE レベルを規定している。
課題	strict(厳密)で作成された文書と transitional(過渡的)で作成された文書に互換性はない。
開発元	Microsoft, ECMA(欧州電子計算機工業会)を介して ISO 標準化
ライセンス	特許の有無:無
メタデータの埋込可否	Office Open XML ファイルには, メタデータを埋め込むことができる。メタデータには, ファイルのタイトル, 作成者, 作成日時, 変更日時などの情報を含む。メタデータは, 文書を補足する情報であり, 文書の表示には影響を与えない。メタデータを埋め込む方法については, Microsoft Office のドキュメントに関する技術情報が提供されている。
参照情報	https://www.iso.org/standard/77818.html
PRONOM ID	fnt/1827
NARA における ID	NF00311
作成ソフトウェア名	Microsoft Word/Microsoft 365/ Google ドキュメント等多数あり
閲覧ソフトウェア名	Microsoft Word/Microsoft 365/ Google ドキュメント等多数あり

日本語への対応	対応済み
---------	------

通番	9, 10
フォーマット名	OpenDocument Text
バージョン	Version 1.2
拡張子	odt, fodt
公開日	2006-12-01
最終更新日	2015-07
バージョンの更新履歴	ISO/IEC 26300:2006 ISO/IEC 26300:2006/Amd 1:2012 ISO/IEC 26300:2006/Cor 1:2010 ISO/IEC 26300:2006/Cor 2:2011 ISO/IEC 26300-1:2015
標準化対応状況	国際標準として制定
特長	Open Document Format for Office Applications を使ってワープロ文書を表示したものであり、ISO/IEC 26300「Open Document Format for Office Applications」(OpenDocument) シリーズとして標準化されている。オフィス文書用の XML スキーマを定義している。オフィス文書には、ワープロ文書、スプレッドシート、図表、図面やプレゼンテーションのようなグラフィカル文書が含まれるが、これらの種類の文書に限定されるものではない。 OpenDocument の XML スキーマは、XSLT や XML ベースのツールで処理できるように設計されている。
課題	ODF の最新 Version は 1.3 である。Version1.3 は OOXNL の strict と親和性が高く、Version1.2 は Transitional と親和性が高いが、完全互換ではない。ODF の XML 構造には 2 つの表現方法があり、ZIP で包む方法と、SingleXML (FlatODF) があるが、すべての ODF ツールが FlatODF をサポートしている訳ではない。
開発元	OASIS, ECMA(欧州電子計算機工業会)を介して ISO 標準化
ライセンス	特許の有無:無
メタデータの埋込可否	OpenDocument ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。メタデータは、文書を補足する情報であり、文書の表示には影響を与えない。一般的に、メタデータは PDF 文書のカタログ化や検索などに利用される。
参照情報	https://www.iso.org/standard/66375.html
PRONOM ID	1.3 NF00801, 1.2 NF00508, 1.1 NF00507, 1.0 NF00345
NARA における ID	1.0 NF00345, 1.1 NF00507, 1.2 NF00508, 1.2 NF00801
作成ソフトウェア名	Google Docs, 一太郎, Microsoft Word, LibreOffice Writer, Apache OpenOffice Writer 等多数あり
閲覧ソフトウェア名	Google Docs, 一太郎, Microsoft Word, LibreOffice Writer, Apache OpenOffice Writer 等多数あり
日本語への対応	対応済み

通番	11, 12
フォーマット名	Office Open XML(OOXML) Spreadsheet
バージョン	Excel2007 以降
拡張子	xlsx
公開日	2008-12

最終更新日	2021-08
バージョンの更新履歴	ISO/IEC 29500-1, -2, -3, -4:2008 ISO/IEC 29500-1, -2, -3,-4:2011 ISO/IEC 29500-1, -2, -3, -4:2012 ISO/IEC 29500-3:2015 ISO/IEC 29500-1, -4:2016 ISO/IEC 29500-2:2021
標準化対応状況	国際標準として制定
特長	Office Open XML を使って Microsoft Excel のブックを表現したものである。ISO/IEC 29500 では、strict(厳密)と transitional(過渡的)の2つのコンFORMANCE レベルを規定している。
課題	strict(厳密)で作成された文書と transitional(過渡的)で作成された文書に互換性はない。
開発元	Microsoft, ECMA(欧州電子計算機工業会)を介して ISO 標準化
ライセンス	特許の有無:無
メタデータの埋込可否	Office Open XML ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。メタデータは、文書を補足する情報であり、文書の表示には影響を与えない。メタデータを埋め込む方法については、Microsoft Office のドキュメントに関する技術情報が提供されている。
参照情報	https://www.iso.org/standard/77818.html
PRONOM ID	fnt/1828
NARA における ID	NF00272
作成ソフトウェア名	Microsoft Excel/Microsoft 365/ Google スプレッドシート等多数あり
閲覧ソフトウェア名	Microsoft Excel/Microsoft 365/ Google スプレッドシート等多数あり
日本語への対応	対応済み

通番	13, 14
フォーマット名	OpenDocument Spreadsheet
バージョン	Version 1.2
拡張子	ods, fods
公開日	2006-12-01
最終更新日	2015-07
バージョンの更新履歴	ISO/IEC 26300:2006 ISO/IEC 26300:2006/Amd 1:2012 ISO/IEC 26300:2006/Cor 1:2010 ISO/IEC 26300:2006/Cor 2:2011 ISO/IEC 26300-1:2015
標準化対応状況	国際標準として制定
特長	Open Document Format for Office Applications を使ってスプレッドシートを表現したものであり、ISO/IEC 26300「Open Document Format for Office Applications」(OpenDocument)シリーズとして標準化されている。オフィス文書用の XML スキーマを定義している。オフィス文書には、ワープロ文書、スプレッドシート、図表、図面やプレゼンテーションのようなグラフィカル文書が含まれるが、これらの種類の文書に限定されるものではない。 OpenDocument の XML スキーマは、XSLT や XML ベースのツールで処理で

	きるように設計されている。
課題	ODF の最新 Version は 1.3 である。Version1.3 は OOXNL の strict と親和性が高く、Version1.2 は Transitional と親和性が高いが、完全互換ではない。ODF の XML 構造には 2 つの表現方法があり、ZIP で包む方法と、SingleXML (FlatODF) があるが、すべての ODF ツールが FlatODF をサポートしている訳ではない。
開発元	OASIS, ECMA (欧州電子計算機工業会) を介して ISO 標準化
ライセンス	特許の有無: 無
メタデータの埋込可否	OpenDocument ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。メタデータは、文書を補足する情報であり、文書の表示には影響を与えない。一般的に、メタデータは PDF 文書のカタログ化や検索などに利用される。
参照情報	https://www.iso.org/standard/66375.html
PRONOM ID	1.3 fmt/1755, 1.2 fmt/295, 1.1 fmt/294, 1.0 fmt/137
NARA における ID	1.3 NF00800, 1.2 NF00514, 1.1 NF00513, 1.0 NF00349
作成ソフトウェア名	Google Spreadsheet, 三四郎, Microsoft Excel, LibreOffice Calc, Apache OpenOffice Calc 等多数あり
閲覧ソフトウェア名	Google Spreadsheet, 三四郎, Microsoft Excel, LibreOffice Calc, Apache OpenOffice Calc 等多数あり
日本語への対応	対応済み

通番	15, 16
フォーマット名	Office Open XML(OOXML) Presentation
バージョン	Powerpoint2007 以降
拡張子	pptx
公開日	2008-12
最終更新日	2021-08
バージョンの更新履歴	ISO/IEC 29500-1, -2, -3, -4::2008 ISO/IEC 29500-1, -2, -3。-4::2011 ISO/IEC 29500-1, -2, -3, -4:2012 ISO/IEC 29500-3:2015 ISO/IEC 29500-1, -4:2016 ISO/IEC 29500-2:2021
標準化対応状況	国際標準として制定
特長	Office Open XML を使って Microsoft PowerPoint のプレゼンテーションを表現したものである。 なお、ISO/IEC 29500 では、strict(厳密)と transitional(過渡的)の2つのコンFORMANCE レベルを規定している。
課題	strict(厳密)で作成された文書と transitional(過渡的)で作成された文書に互換性はない。
開発元	Microsoft, ECMA (欧州電子計算機工業会) を介して ISO 標準化
ライセンス	特許の有無: 無
メタデータの埋込可否	Office Open XML ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。メタデータは、文書を補足する情報であり、文書の表示には影響を与えない。メタデータを埋め込む方法については、Microsoft Office のドキュメントに関する技術情報が提供されている。
参照情報	https://www.iso.org/standard/77818.html

PRONOM ID	fnt/1829
NARA における ID	NF00289
作成ソフトウェア名	Microsoft PowerPoint/Microsoft 365/ Google スライド等多数あり
閲覧ソフトウェア名	Microsoft PowerPoint/Microsoft 365/ Google スライド等多数あり
日本語への対応	対応済み

通番	17, 18
フォーマット名	OpenDocument Presentation
バージョン	Versu1.2
拡張子	odp, fodp
公開日	2006-12-01
最終更新日	2015-07
バージョンの更新履歴	ISO/IEC 26300:2006 ISO/IEC 26300:2006/Amd 1:2012 ISO/IEC 26300:2006/Cor 1:2010 ISO/IEC 26300:2006/Cor 2:2011 ISO/IEC 26300-1:2015
標準化対応状況	国際標準として制定
特長	Open Document Format for Office Applications を使ってプレゼンテーションを表現したものであり、ISO/IEC 26300「Open Document Format for Office Applications」(OpenDocument) シリーズとして標準化されている。オフィス文書用の XML スキーマを定義している。オフィス文書には、ワープロ文書、スプレッドシート、図表、図面やプレゼンテーションのようなグラフィカル文書が含まれるが、これらの種類の文書に限定されるものではない。 OpenDocument の XML スキーマは、XSLT や XML ベースのツールで処理できるように設計されている。
課題	ODF の最新 Version は 1.3 である。Version1.3 は OOXNL の strict と親和性が高く、Version1.2 は Transitional と親和性が高いが、完全互換ではない。 ODF の XML 構造には 2 つの表現方法があり、ZIP で包む方法と、SingleXML (FlatODF) があるが、すべての ODF ツールが FlatODF をサポートしている訳ではない。
開発元	OASIS, ECMA(欧州電子計算機工業会)を介して ISO 標準化
ライセンス	特許の有無:無
メタデータの埋込可否	OpenDocument ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。メタデータは、文書を補足する情報であり、文書の表示には影響を与えない。一般的に、メタデータは PDF 文書のカタログ化や検索などに利用される。
参照情報	https://www.iso.org/standard/66375.html
PRONOM ID	1.3 fnt/1754, 1.2 fnt/293, 1.1 fnt/292, 1.0 fnt/138
NARA における ID	1.3 NF00799, 1.2 NF00512, 1.1 NF00511, 1.0 NF00348
作成ソフトウェア名	Google Slides, Just Slide, Microsoft PowerPoint, LibreOffice Impress, Apache OpenOffice Impress 等多数あり
閲覧ソフトウェア名	Google Slides, Just Slide, Microsoft PowerPoint, LibreOffice Impress, Apache OpenOffice Impress 等多数あり
日本語への対応	対応済み

通番	19															
フォーマット名	JPEG2000															
バージョン	制定年で識別されバージョンの概念はない															
拡張子	Jp2															
公開日	2000-12															
最終更新日	2019-10															
バージョンの更新履歴	2000, 2004, 2016, 2019															
標準化対応状況	国際標準として開発 (ISO/IEC 15444)															
特長	JPEG2000 は、JPEG よりも高い圧縮率で、より高品質の画像を実現するために作成された。JPEG2000 は、非可逆圧縮の JPEG ファイルの問題に対応するために、可逆圧縮を用いて、画像データの損失を最小限に抑えている。JPEG2000 は、周波数変換に離散コサイン変換ではなく、離散ウェーブレット変換を用いる。これにより、画像の局所的な特徴をより効率的に表現できる。JPEG2000 は、一つの圧縮画像を様々な解像度やビットレートで利用できるというスケーラビリティ機能を有している。これにより、画像の一部だけを取り出したり、画質やサイズを調整したりすることが容易になる。JPEG2000 は、送信中または保存中に発生する可能性のあるエラーに対して回復力があるように設計されている。オーディオに可聴アーティファクトやその他の歪みを導入することなく、エラーから回復できる。 JPEG2000 のコンフォーマンスレベルは、CL0 (低圧縮高品質) から CL255 (高圧縮低品質) まで定義できる仕様になっている。また、JPEG2000 の主要なプロファイルにはつぎのようなものがある。 Baseline (BP):最も基本的なプロファイル。コンパクトなファイルサイズと高い圧縮効率を提供。ウェブ画像やデスクトップアプリケーションの画像保存など、一般的な用途に適している。 Extended (XP):Baseline プロファイルの機能を拡張し、より高い圧縮効率と高品質の画像圧縮を提供。一般的な用途に加えて、より専門的な画像処理タスクに使用される。 Main (MP):さまざまな画像品質要件を満たす。一般的に、メディカルイメージング、印刷業界、アーカイブ、プロフェッショナルな画像処理に使用される。 High (HP):最高品質の圧縮を提供し、ロスレス圧縮に焦点を当てている。一般的に、高品質な画像アーカイブや専門的な医療画像などで使用される。 Cinema2K (C2K):デジタルシネマの要件に合わせて設計され 2K 解像度の映画コンテンツをサポート。映画制作およびデジタルシネマで使用される。 Broadcast (BR):放送業界向け。映像制作や放送用の画像コンテンツに特化。高品質の画像と音声の統合をサポート。 Imaging (IM):医療画像、衛星画像、産業用画像処理などの高度な画像処理タスクをサポート。カラー画像、多重解像度、透過情報などを処理できる。 コンフォーマンスレベルとプロファイルの組合せの例を次に示す。 <table><tr><td>用途</td><td>コンフォーマンスレベル (CL)</td><td>プロファイル</td></tr><tr><td>一般用途</td><td>0 低圧縮高品質</td><td>0 Baseline</td></tr><tr><td>医療画像等</td><td>0 低圧縮高品質</td><td>High</td></tr><tr><td>映画製作向け</td><td>0 低圧縮高品質</td><td>Cinema2K</td></tr><tr><td>放送向け</td><td>0 低圧縮高品質 または 1 中程度の圧縮ほぼ高品質</td><td>Broadcast</td></tr></table>	用途	コンフォーマンスレベル (CL)	プロファイル	一般用途	0 低圧縮高品質	0 Baseline	医療画像等	0 低圧縮高品質	High	映画製作向け	0 低圧縮高品質	Cinema2K	放送向け	0 低圧縮高品質 または 1 中程度の圧縮ほぼ高品質	Broadcast
用途	コンフォーマンスレベル (CL)	プロファイル														
一般用途	0 低圧縮高品質	0 Baseline														
医療画像等	0 低圧縮高品質	High														
映画製作向け	0 低圧縮高品質	Cinema2K														
放送向け	0 低圧縮高品質 または 1 中程度の圧縮ほぼ高品質	Broadcast														
課題	特になし															
開発元	ISO/IEC JTC1/SC29, Joint Photographic Experts Group															

ライセンス	特許の有無:有。特許の制約はない。JPEG2000 は「ロイヤリティフリー(無償利用可能)」または「リーズナブルで非差別的な条件(RAND)」で利用できるようにすることを目指したが、すべての関連特許がこの方針に従っているわけではない。
メタデータの埋込可否	JPEG 2000 ファイルには、メタデータを埋め込むことができる。JPEG 2000 ファイルは、JPEG 2000 コアビットストリームを拡張するオブジェクトベースの設計に基づいており、不透明度/透明度チャンネル、カラープロファイル情報、その他のメタデータなどのオブジェクトタイプをラップすることができる。 XMP(Extensible Metadata Platform)は、JPEG 2000 ファイルに埋め込むための XML ベースのメタデータスキーマである。XMP メタデータを JPEG 2000 ファイルに埋め込む方法については、Adobe Photoshop や Adobe Bridge などのアプリケーションで提供される XMP パネルを使用することができる。
参照情報	https://jpeg.org/jpeg2000/ https://www.iso.org/standard/78321.html
PRONOM ID	x-fmt/392
NARA における ID	NF00219
作成ソフトウェア名	Adobe Photoshop, GIMP はじめ多数あり
閲覧ソフトウェア名	Google Chrome, Firefox, Microsoft Edge をはじめ多数あり
日本語への対応	対応済み

通番	20
フォーマット名	TIFF (Tagged Image File Format)
バージョン	Version 6
拡張子	tif tiff
公開日	1998-09
最終更新日	2004-05
バージョンの更新履歴	1998(Withdraw),2004
標準化対応状況	なし。TIFF/ITは国際標準として制定(ISO 12639)
特長	TIFF は非圧縮または可逆圧縮形式をサポートし、画像データの損失を最小限に抑えて写真やスキャン画像などの情報を正確に保存することができる。TIFF は、1 ビットから 32 ビットまでの色深度と、RGB, CMYK, グレースケールなどのカラーモードをサポートし、さまざまな画像タイプに対応できる。TIFF は、Windows, Linux, macOS などの主要な OS で使用でき、また、多くの画像編集ソフトやビューアソフトでも対応している。
課題	Version 6 以前の version は自由度が高い反面、相互運用性に欠ける。
開発元	Aldus Corporation(現在, Adobe Inc.)
ライセンス	特許の有無:無
メタデータの埋込可否	TIFF(Tag Image File Format)ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。メタデータは、文書を補足する情報であり、文書の表示には影響を与えない。一般的に、メタデータは PDF 文書のカatalog化や検索などに利用される。 TIFF ファイルのメタデータには、事前定義されたメタデータ要素が含まれており、ユーザー定義およびカスタムメタデータも許可されている。TIFF ファイルにメタデータを埋め込む方法については、TIFFImageWriter クラスを使用して実行できる。
参照情報	https://developer.adobe.com/content/dam/udp/en/open/standards/tiff/TIFF6.pdf

	https://www.iso.org/standard/34342.html
PRONOM ID	fnt/353
NARA における ID	NF00419
作成ソフトウェア名	Adobe Photoshop, GIMP はじめ多数あり
閲覧ソフトウェア名	Google Chrome, Firefox, Microsoft Edge をはじめ多数あり
日本語への対応	対応済み

通番	21
フォーマット名	PNG (Portable Network Graphics)
バージョン	Version1.2
拡張子	png
公開日	2004-03
最終更新日	2004-03
バージョンの更新履歴	2004
標準化対応状況	国際標準として制定(ISO/IEC 15948)
特長	PNG は非圧縮または可逆圧縮形式で、画像データの損失を最小限に抑えることができるため、写真やスキャン画像などの情報を正確に保存することができる。PNG は、1 ビットから 32 ビットまでの色深度と、RGB, CMYK, グレースケールなどのカラーモードをサポートしている。これにより、さまざまな画像タイプに対応できる。PNG は、Windows, Linux, macOS などの主要な OS で使用でき、多くの画像編集ソフトやビューアソフトでも対応している。PNG は透明な処理だけでなく半透明などの調整も可能であり、ロゴなどの画像を重ねているファイルを扱う際に、縁取りの違和感を消し、なじませることができる。比較的新しい形式のため古いブラウザでは表示されない可能性がある。
課題	特になし
開発元	PNG Development Group (W3C に寄贈)
ライセンス	特許の有無:無
メタデータの埋込可否	PNG ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。メタデータは、文書を補足する情報であり、文書の表示には影響を与えない。一般的に、メタデータは PDF 文書のカタログ化や検索などに利用される。 PNG ファイルのメタデータには、事前定義されたメタデータ要素が含まれており、ユーザー定義およびカスタムメタデータも許可されている ¹²³ 。PNG ファイルにメタデータを埋め込む方法については、多くの画像編集ソフトウェアが提供している。
参照情報	http://libpng.org/pub/png/ https://www.iso.org/standard/29581.html
PRONOM ID	fnt/13
NARA における ID	NF00383
作成ソフトウェア名	Adobe Photoshop, GIMP はじめ多数あり
閲覧ソフトウェア名	Google Chrome, Firefox, Microsoft Edge をはじめ多数あり
日本語への対応	対応済み

通番	22
フォーマット名	GIF (Graphics Interchange Format)
バージョン	89a
拡張子	gif

公開日	1987-6
最終更新日	1990-7
バージョンの更新履歴	87a 89a
標準化対応状況	対応していない
特長	GIF は、ファイルサイズが小さく、軽いのが特徴である。これにより、WEB ページや動画関連ファイルへの埋め込みなど、できるだけデータの読み込み容量を抑えたい時に適している。GIF は、静止画を連続再生させることで短いアニメーションを制作できる。アニメーションは自動再生式で、再生ボタンをクリックする必要がない。GIF は、画像の背景を透過させることができるため、背景の不要な部分をカットできるだけでなく、他の画像に背景透過させた画像を重ねると透明部分に画像が差し替わり、簡易的なグラデーション表現ができる。最大 256 色までしか表現できないため、画像として再現性が低い。
課題	初期のバージョン 87a は相互運用性の問題がある。
開発元	CompuServe
ライセンス	特許の有無:2004 年 6 月に失効
メタデータの埋込可否	GIF (Graphics Interchange Format) ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。GIF ファイルのメタデータには、事前定義されたメタデータ要素が含まれており、ユーザー定義およびカスタムメタデータも許可されている。GIF ファイルにメタデータを埋め込む方法については、多くの画像編集ソフトウェアが提供している。
参照情報	https://en.wikipedia.org/wiki/GIF
PRONOM ID	fnt/4
NARA における ID	NF00203
作成ソフトウェア名	Adobe Photoshop, GIMP はじめ多数あり
閲覧ソフトウェア名	Google Chrome, Firefox, Microsoft Edge をはじめ多数あり
日本語への対応	対応済み

通番	23
フォーマット名	JPEG (Joint Photographic Experts Group)
バージョン	制定年で表示されバージョンの概念はない
拡張子	jpg, jpeg
公開日	1994
最終更新日	2021
バージョンの更新履歴	ISO/IEC 10918-1:1994 ISO/IEC 10918-1:1994/Cor 1:2005 ISO/IEC 10918-2:1995 ISO/IEC 10918-3:1997 ISO/IEC 10918-3:1997/Amd 1:1999 ISO/IEC 10918-4:1999 ISO/IEC 10918-4:1999/Amd 1:2013 ISO/IEC 10918-5:2013 ISO/IEC 10918-6:2013 ISO/IEC 10918-7:2021
標準化対応状況	国際標準として制定
特長	JPEG は非可逆圧縮方式の画像形式で、一度圧縮した画像は元に戻すことができない。また、上書き保存するたびに画像が劣化するという特性がある。フル

	カラー(約 1677 万色)に対応しており、高画質かつ鮮明な画像表現が可能である。圧縮率をユーザーが指定することができ、画像の品質とファイルサイズをバランス良く調整することが可能である。JPEG では、画像の背景を透過させることはできない。そのため、背景を透過させたい場合には他の画像形式を使用する必要がある。JPEG は、PC やスマートフォン、タブレット、デジタルカメラなどデバイスを問わずに扱うことができる。また、ほとんどの画像編集ソフトに対応しているため、画像編集がしやすい点も大きなメリットである。これらの特長から、JPG は写真や高解像度の画像で利用されやすい画像形式である。ただし、何度も編集する場合や背景を透過させたい場合には向いていない。
課題	特になし
開発元	ISO/IEC JTC1/SC29, Joint Photographic Experts Group
ライセンス	特許の有無:無
メタデータの埋込可否	JPEG ファイルには、メタデータを埋め込むことができる。JPEG ファイルにメタデータを埋め込む方法は、多くの画像編集ソフトウェアが提供している。 JPEG ファイルに埋め込まれたメタデータの例としては、Exchangeable image file format (Exif)がある。Exif は、JPEG ファイルに含まれる画像情報を記録するための規格である。Exif には、カメラのモデル名、撮影日時、露出時間などの情報を含む。
参照情報	https://jpeg.org/jpeg/ 10918-1 : https://www.iso.org/standard/18902.html 10918-2 : https://www.iso.org/standard/20689.html 10918-3 : https://www.iso.org/standard/25037.html 10918-4 : https://www.iso.org/standard/25431.html 10918-5 : https://www.iso.org/standard/54989.html 10918-6 : https://www.iso.org/standard/59634.html 10918-7 : https://www.iso.org/standard/80896.html
PRONOM ID	fnt/44
NARA における ID	NF00492
作成ソフトウェア名	Adobe Photoshop, GIMP はじめ多数あり
閲覧ソフトウェア名	Google Chrome, Firefox, Microsoft Edge をはじめ多数あり
日本語への対応	対応済み

通番	24
フォーマット名	WAVE (Waveform Audio File Format)
バージョン	—
拡張子	wav wave
公開日	1991-08
最終更新日	1994-04
バージョンの更新履歴	非公開
標準化対応状況	対応していない
特長	WAVE ファイルは非圧縮形式で、音源そのままのデータを保持する。WAVE ファイルは Microsoft が開発した Windows 用の音声ファイル形式で、Windows が標準で対応している。WAVE ファイルは、複数の音を作る 1 つの波形をそのままの状態データにしているため、CD やレコードなどの高品質な音源をデジタル化する際や、音楽制作・編集によく使用される。データサイズが大きいため、ストレージ容量に余裕がある場合や、高音質が求められる場合に適する。
課題	特になし

開発元	Microsoft
ライセンス	特許の有無:無
メタデータの埋込可否	WAV (Waveform Audio File Format) ファイルには、メタデータを埋め込むことができる。WAV ファイルのメタデータには、事前定義されたメタデータ要素が含まれており、ユーザー定義およびカスタムメタデータも許可されている。WAV ファイルにメタデータを埋め込む方法については、多くの画像編集ソフトウェアが提供している。
参照情報	https://docs.fileformat.com/ja/audio/wav/
PRONOM ID	fnt/141
NARA における ID	NF00438
作成ソフトウェア名	FFmpeg; VLC; Audacity
閲覧ソフトウェア名	FFmpeg; VLC; Audacity
日本語への対応	対応済み

通番	25
フォーマット名	MP3 (MPEG Layer III Audio Encoding)
バージョン	—
拡張子	mp3
公開日	1993-08
最終更新日	1993-08
バージョンの更新履歴	ISO/IEC 11172-3:1993、ISO/IEC 13818-3:1998
標準化対応状況	国際標準として制定
特長	MP3 は、元の音楽データのうち、人間の可聴域から外れた音声をカットすることでデータ容量の圧縮を実現している。その圧縮率は 5 分の 1 から 10 分の 1 となっており、原音の約 10 分の 1 にまで圧縮することができる。MP3 は最もメジャーな音声ファイル形式であり、多くのプレイヤーやデバイスで再生できる。その普及は、iPod やウォークマンに代表されるポータブルプレイヤーの普及に繋がった。一般的な MP3 ファイルは 128kbps～320kbps 前後の容量となっており、320kbps になると CD と遜色ない高音質の音楽ファイルとして利用できる。MP3 は元のデータが完全には保存されない非可逆圧縮（不可逆圧縮）形式であるため、一度圧縮（変換）すると元に戻せない。これらの特長から、MP3 は曲数を多く入れたい場合や、CD とほぼ同じ音質を求める場合に適する。ただし、非可逆圧縮形式であるため、元のデータが完全には保存されない点に注意が必要である。
課題	特になし
開発元	ISO/IEC JTC1/SC29
ライセンス	特許の有無:2017 年 4 月に失効
メタデータの埋込可否	MP3 ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。メタデータは、文書を補足する情報であり、文書の表示には影響を与えない。MP3 ファイルにメタデータを埋め込む方法については、多くの音楽編集ソフトウェアが提供している。 MP3 ファイルに埋め込まれたメタデータの例としては、Exchangeable image file format (Exif)がある 1. Exif は、JPEG ファイルに含まれる画像情報を記録するための規格である。Exif には、カメラのモデル名、撮影日時、露出時間などの情報を含む。
参照情報	https://www.iso.org/standard/22412.html
PRONOM ID	fnt/134

NARA における ID	NF00335
作成ソフトウェア名	VLC Media Player 等多数
閲覧ソフトウェア名	VLC Media Player 等多数
日本語への対応	対応済み

通番	26
フォーマット名	AAC (Advanced Audio Coding MPEG-4 Low Complexity Object)
バージョン	—
拡張子	aac adts
公開日	1997-12
最終更新日	2006-01
バージョンの更新履歴	ISO/IEC 13818-7:1997 ISO/IEC 13818-7:2003 ISO/IEC 13818-7:2004 ISO/IEC 13818-7:2006
標準化対応状況	国際標準として制定
特長	AAC は、広い周波数範囲と高い S/N 比を持つ高品質のオーディオを提供するように設計されている。また、幅広いビットレートに対応しているため、高音質と低帯域の両方のアプリケーションに使用することができる。AAC は高度な音声符号化技術により、高い品質を維持したまま音声データを圧縮する。これにより、他のオーディオフォーマットよりも少ない容量でオーディオデータを伝送・保存することができる。AAC は、スマートフォン、タブレット、メディアプレーヤーなど、さまざまな機器やソフトウェアに対応している。また、Spotify や Apple Music など、多くの人気ストリーミングサービスでもサポートされている。AAC は、送信中または保存中に発生する可能性のあるエラーに対して回復力があるように設計されている。オーディオに可聴アーティファクトやその他の歪みを導入することなく、エラーから回復できる。AAC は、5.1 および 7.1 サラウンドサウンドを含むマルチチャンネルオーディオをサポートし、ホームシアターシステムなどのアプリケーションでの使用に適している。これらの特長から、AAC ファイルは音楽配信やデジタル放送などで広く利用されている。
課題	非可逆圧縮形式であるため、一度圧縮(変換)すると元に戻せない点に注意が必要である。
開発元	ブラウンフォーファ研究所、ドルビー、ソニーにより開発され、ISO/IEC JTC1/SC29 で標準化された。
ライセンス	特許の有無:無
メタデータの埋込可否	AAC (Advanced Audio Coding) ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。AAC ファイルにメタデータを埋め込む方法については、多くの音楽編集ソフトウェアが提供している。また、ffmpeg を使用して AAC ファイルにメタデータを追加することもできる。
参照情報	https://www.iso.org/standard/43345.html
PRONOM ID	fnt/1812
NARA における ID	NF00106
作成ソフトウェア名	FFmpeg; VLC; Windows Media Player
閲覧ソフトウェア名	FFmpeg; VLC; Windows Media Player
日本語への対応	対応済み

通番	27, 28
フォーマット名	MPEG-2 (MPEG2-PS, MPEG2-TS)
バージョン	—
拡張子	mpg mpeg mpeg2 mpv mp2t mp1s mp2p x-mpg x-pn-mpg x-mpeg x-mpeg2a
公開日	1996-04
最終更新日	2022-09
バージョンの更新履歴	ISO/IEC 13818-1:1996 ISO/IEC 13818-1:2000 ISO/IEC 13818-1:2007 ISO/IEC 13818-1:2013 ISO/IEC 13818-1:2015 ISO/IEC 13818-1:2018 ISO/IEC 13818-1:2019 ISO/IEC 13818-1:2022
標準化対応状況	国際標準として制定
特長	MPEG2 は、MPEG-1 を踏襲しつつ、低解像度から高解像度まで 4 種類のサイズに対応する動画圧縮技術であり、デジタルテレビ放送や DVD などのメディアに適したファイル形式である。MPEG2 は、インターレースされた動画を効率的に圧縮するために、フィールド補完という技術を導入している。これにより、動きのあるシーンでも画質の劣化を抑えることができる。MPEG2 は、動画や音声だけでなく、字幕やメタデータなどのさまざまなコンテンツを一つのファイルに統合することができる。また、プログラムストリームやトランスポートストリームという多重化方式を定義し、これにより、蓄積メディアや放送・通信などの用途に応じて柔軟にデータ構造を変更できる。MPEG2 は、国際標準化機構 (ISO) と国際電気通信連合 (ITU) によって定められたオープン標準であり、多くのプラットフォームやデバイスで再生できるようになっている。また、多くの文献やウェブサイトで詳細に説明されており、オープンソースのツールで変換や編集ができるため、保存に適した形式としても認められている。
課題	特になし
開発元	ISO/IEC JTC1/SC29
ライセンス	特許の有無: コンテナについては無し、映像コーデックについては残存
メタデータの埋込可否	MPEG2 ファイルには、メタデータを埋め込むことができる。メタデータには、ファイルのタイトル、作成者、作成日時、変更日時などの情報を含む。MPEG2 ファイルにメタデータを埋め込む方法については、多くのビデオ編集ソフトウェアが提供している。
参照情報	https://www.iso.org/standard/83239.html
PRONOM ID	x-fmt/386
NARA における ID	NF00337
作成ソフトウェア名	FFmpeg; VLC; Windows Media Player
閲覧ソフトウェア名	FFmpeg; VLC; Windows Media Player
日本語への対応	対応済み

通番	29
フォーマット名	MPEG-4
フォーマット名	MPEG-4
バージョン	—
拡張子	mp4 mpa

公開日	2003-11
最終更新日	2020-01
バージョンの更新履歴	ISO/IEC 14496-14:2003 ISO/IEC 14496-14:2018 ISO/IEC 14496-14:2020
標準化対応状況	国際標準として制定
特長	MPEG4 は、MPEG-4 Advanced Video Encoding という技術を用いて、ビデオを高品質で低ビットレートに圧縮することができる。これにより、ストリーミングやダウンロードなどのインターネット上のビデオ配信に適したファイル形式になっている。MPEG4 は、ビデオだけでなく、音声や字幕や画像などのさまざまなメディアコンテンツを一つのファイルに統合することができる。また、インタラクティブな機能や 3D グラフィックスなどの拡張機能もサポートしている。MPEG4 は、国際標準化機構 (ISO) によって定められたオープン標準であるため、多くのプラットフォームやデバイスで再生できるようになっている。また、多くの文献やウェブサイトで詳細に説明されており、オープンソースのツールで変換や編集ができるため、保存に適した形式としても認められている。
課題	特になし
開発元	ISO/IEC JTC1/SC29
ライセンス	特許の有無:コンテナについては無し、映像コーデックについては残存
メタデータの埋込可否	MPEG4 ファイルには、メタデータを埋め込むことができる。MPEG4 ファイルは、ISO Base Media File Format を拡張して、MPEG-4 ビジュアル/オーディオコーデックやオブジェクト記述子、シーン記述子などのさまざまな MPEG-4 システム機能をサポートするように定義されている 1. 他の ISO ベースメディアファイルフォーマットに基づく形式 (例:3GP) でもこれらの拡張機能が使用されている。 MPEG4 ファイルにメタデータを埋め込む方法については、多くのビデオ編集ソフトウェアが提供している。また、ffmpeg を使用して MPEG4 ファイルにメタデータを追加することもできる。
参照情報	https://www.iso.org/standard/79110.html
PRONOM ID	fnt/596
NARA における ID	NF00339
作成ソフトウェア名	FFmpeg; VLC; Windows Media Player
閲覧ソフトウェア名	FFmpeg; VLC; Windows Media Player
日本語への対応	対応済み

別紙2 その他点検するフォーマットに関する調査結果

通番	30～35
フォーマット名	一太郎
バージョン	2023 (33)
拡張子	jdt, jtdc, jfw, jbw, jaw
公開日	1985
最終更新日	2023
バージョンの更新履歴	1985jx-word 一太郎 1985 一太郎 1986 一太郎 Ger. 1987 一太郎 Ver.3 1989 一太郎 Ver.4 1993 一太郎 Ver.5 1995 一太郎 Ver.6 1996 一太郎 7 1997 一太郎 8 1998 一太郎 9 1999 一太郎 10 2001 一太郎 11 2002 一太郎 12 2003 一太郎 13 以降, 2004 から 2023 まで一太郎 20xx
標準化対応状況	対応していない
特長	ジャストシステムが開発した日本語ワープロ。公用文のルールに対応している。
課題	docx 形式では保存できないため、一旦 doc 形式で保存し、doc 形式から docx に変換する必要がある
開発元	ジャストシステム
ライセンス	特許の有無:無
メタデータの埋込可否	一太郎に固定メタデータを埋め込むことができる。ただし、一太郎のバージョンによっては、メタデータを埋め込む機能がない場合がある。
参照情報	https://www.justsystems.com/jp/products/ichitaro/
PRONOMID	fnt/1400
NARA における ID	該当なし
作成ソフトウェア名	一太郎
閲覧ソフトウェア名	一太郎
日本語への対応	対応済み

通番	36
フォーマット名	PDF
バージョン	1.4
拡張子	pdf
公開日	2001
最終更新日	2001
バージョンの更新履歴	2001
標準化対応状況	対応していない 但し、PDF/A-1 のベース仕様として参照されている
特長	PDF1.4 は、Adobe Systems Incorporated によって開発された Portable

	Document Format のバージョン 1.4 であり、2001 年にリリースされた Acrobat5.0 に対応するバージョンである。PDF1.4 では、画像やテキストなどのオブジェクトに透明度やブレンドモードを設定することができ、これにより、よりリアルな画像表現やデザインが可能である。フォントは、TrueType フォントと PostScript フォントの特徴を兼ね備えた OpenType フォントをサポートし、多言語や高度な文字装飾に対応している。また、エクスポート時に Unicode 文字コードを使用することができる。
課題	初期の実装において、透明効果に問題が発生する場合があることが報告されている。 https://blog.adobe.com/jp/publish/2022/02/09/dc-20-years-of-transparency-in-pdf
開発元	Adobe
ライセンス	特許の有無:無
メタデータの埋込可否	固定のみ可
参照情報	公式ページの URL 不明
PRONOMID	fnt/18
NARA における ID	NF00366
作成ソフトウェア名	Adobe Acrobat5
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	37
フォーマット名	PDF
バージョン	1.5
拡張子	pdf
公開日	2003
最終更新日	2003
バージョンの更新履歴	2003
標準化対応状況	対応していない
特長	PDF1.5 は、Adobe Systems Incorporated によって開発された Portable Document Format のバージョン 1.5 であり、2003 年にリリースされた Acrobat6.0 に対応するバージョンである。PDF1.5 では、JPEG2000 をサポートし、また、Flash や QuickTime などのマルチメディアファイルを PDF ファイルに添付し、再生することができ、マルチメディアファイルの再生に必要なプラグインやコーデックを PDF ファイルに埋め込むこともできる。PDF1.5 では、レイヤー機能(文書の一部を表示したり非表示にしたりすることができる機能)をサポートし、文書の内容や見た目を柔軟に変更することができる。PDF1.5 では、暗号化やデジタル署名などのセキュリティ機能や、注釈やブックマークなどの文書管理機能が強化されている。
課題	特になし
開発元	Adobe
ライセンス	特許の有無:無
メタデータの埋込可否	PDF のメタデータは、Adobe Systems によって開発された、XML ベースの XMP(Extensible Metadata Platform)形式で保存される。メタデータには、PDF 文書のタイトル、作成者、作成日時、変更日時などの情報を含む
参照情報	公式ページの URL 不明
PRONOMID	fnt/19
NARA における ID	NF00367

作成ソフトウェア名	Adobe Acrobat6
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	38
フォーマット名	PDF
バージョン	1.6
拡張子	pdf
公開日	2004
最終更新日	2004
バージョンの更新履歴	2004
標準化対応状況	対応していない
特長	PDF1.6 は、Adobe Systems Incorporated によって開発された Portable Document Format のバージョン 1.6 であり、2004 年にリリースされた Acrobat7.0 に対応するバージョンである。PDF1.6 では、3D アートワークを含む PDF ファイルを作成し、表示することができる(3D アートワークは、Acrobat3D や CAD ソフトウェアなどで作成できる)。PDF1.6 では、リッチテキスト(テキストの書式やスタイルを指定できる形式)を含む PDF ファイルを作成し、表示することができる。
課題	特になし
開発元	Adobe
ライセンス	特許の有無:無
メタデータの埋込可否	PDF のメタデータは、Adobe Systems によって開発された、XML ベースの XMP (Extensible Metadata Platform) 形式で保存される。メタデータには、PDF 文書のタイトル、作成者、作成日時、変更日時などの情報を含む
参照情報	公式ページの URL 不明
PRONOMID	fnt/20
NARA における ID	NF00368
作成ソフトウェア名	Adobe Acrobat7
閲覧ソフトウェア名	Adobe Acrobat/JUSTPDF/SkyPDF
日本語への対応	対応済み

通番	39
フォーマット名	Microsoft Word95
バージョン	6.0/95
拡張子	doc
公開日	1995
最終更新日	1995
バージョンの更新履歴	1995
標準化対応状況	対応していない
特長	Wordドキュメントは、複合ドキュメント形式ファイル内の「WordDocument」ストリームとして保存される。この形式は Word95 のリリース以降も変更されていない。
課題	
開発元	Microsoft
ライセンス	特許の有無:無
メタデータの埋込可否	埋込み不可

参照情報	仕様非公開
PRONOMID	fnt/39
NARA における ID	NF00302
作成ソフトウェア名	Microsoft Word for Windows7.0
閲覧ソフトウェア名	Microsoft Word/Microsoft365/ Google ドキュメント等多数あり
日本語への対応	対応済み

通番	40
フォーマット名	Microsoft Word97-2003
バージョン	97-2003
拡張子	doc wbk
公開日	1997
最終更新日	2003
バージョンの更新履歴	1997, 2000, 2001, 2003
標準化対応状況	対応していない
特長	Word97 のリリースに伴い、Microsoft は、汎用の OLE2 複合ドキュメント形式に基づいたネイティブバイナリワードプロセッシング形式を改訂した。この形式は独自のものである。Word ドキュメントは、複合ドキュメント形式ファイル内の「Word Document」ストリームとして保存される。この形式は Word2000, 2002, 2003 のリリースでも変更されていない。代替拡張子.wbk は Word 文書のバックアップファイルを指すが、.wbk ファイルと.doc ファイルの間には物質的または構造的な違いはない。
課題	OOXML との互換性
開発元	Microsoft
ライセンス	特許の有無:無
メタデータの埋込可否	Word ファイルのメタデータには、事前定義されたメタデータ要素が含まれている。Word ファイルにメタデータを埋め込む方法については、Microsoft Office の公式ドキュメントや多くの Web サイトが提供している。
参照情報	https://msopenspecs.azureedge.net/files/OfficeFileFormatsProtocols.zip
PRONOMID	fnt/40
NARA における ID	NF00303
作成ソフトウェア名	Microsoft WordforWindows8.0/8.5/9.0/10.0/11.0
閲覧ソフトウェア名	Microsoft Word/Microsoft365/ Google ドキュメント等多数あり
日本語への対応	対応済み

通番	41
フォーマット名	Microsoft Excel95
バージョン	5.0/95
拡張子	xls xlw
公開日	1995
最終更新日	1995
バージョンの更新履歴	なし

標準化対応状況	対応していない
特長	Binary Interchange File Format(BIFF)は、Microsoft によって開発されたスプレッドシートファイル形式である。BIFF5 は、Windows 用 Microsoft Excel5.0 のネイティブファイル形式である。この形式は独自のものであり、Microsoft はその構造の詳細を公開していない。Excel5.0 以降、BIFF は Microsoft の汎用 OLE2 複合ドキュメント形式に基づいている。Excel ワークブックは、複合ドキュメント形式ファイル内の「ブック」ストリームとして保存される。ブックストリームは、ファイルの始まり(BOF)レコード、ワークブックグローバルレコードを含み、ファイルの終わり(Eof)レコードで終了する。
課題	OOXML との互換性
開発元	Microsoft
ライセンス	特許の有無:無
メタデータの埋込可否	埋込不可
参照情報	仕様非公開
PRONOMID	fnt/59
NARA における ID	該当なし
作成ソフトウェア名	Microsoft Excel for Windows7.0
閲覧ソフトウェア名	Microsoft Excel/Microsoft365/ Google スプレッドシート等多数あり
日本語への対応	対応済み

通番	42
フォーマット名	Microsoft Excel97
バージョン	8
拡張子	xls xlw
公開日	1997
最終更新日	1997
バージョンの更新履歴	1997
標準化対応状況	対応していない
特長	Binary Interchange File Format(BIFF)は、Microsoft によって開発されたスプレッドシートファイル形式である。BIFF8 は Microsoft Excel97 のネイティブファイル形式である。この形式は独自のものであり、Microsoft はその構造の詳細を公開していない。BIFF8 は、Microsoft の汎用 OLE2 複合ドキュメント形式に基づいている。Excel ワークブックは、複合ドキュメント形式ファイル内の「ブック」ストリームとして保存される。ブックストリームは、ファイルの始まり(BOF)レコード、ワークブックグローバルレコードを含み、ファイルの終わり(Eof)レコードで終了するワークブックグローバルサブストリームで構成される。これに、ファイルの先頭(BOF)レコード、ワークシートレコード、ファイルの終わり(Eof)レコードによって終了する。
課題	OOXML との互換性
開発元	Microsoft
ライセンス	特許の有無:無
メタデータの埋込可否	Excel ファイルのメタデータには、事前定義されたメタデータ要素が含まれている。Excel ファイルにメタデータを埋め込む方法については、Microsoft Office の公式ドキュメントや多くの Web サイトが提供している。
参照情報	https://msopenspecs.azureedge.net/files/OfficeFileFormatsProtocols.zip
PRONOMID	fnt/61

NARA における ID	NF00263
作成ソフトウェア名	Microsoft Excel for Windows8.0
閲覧ソフトウェア名	Microsoft Excel/Microsoft365/ Google スプレッドシート等多数あり
日本語への対応	対応済み

通番	43
フォーマット名	Microsoft Excel2007 (マクロ有効ブック)
バージョン	2007
拡張子	xlsm
公開日	2007
最終更新日	—
バージョンの更新履歴	—
標準化対応状況	国際標準として制定(ISO/IEC2950 シリーズ)
特長	「XLSM は、マクロの使用を可能にする XLSX(Office2007 以降で使用されている Microsoft Excel 形式)のバリエーションである。マクロはコンピュータプログラム的一种であるため、これにより、マクロ以外の通常のスプレッドシートファイルよりも大きなセキュリティリスクが生じる。これは、悪意のあるアクションを自動的に実行できるため、マクロを含む可能性のあるドキュメントとマクロを含まないドキュメントを区別するために個別の拡張機能を使用される。マクロは Visual Basic for Applications(VBA)で作成される。Excel 以外のプログラムでは、VBA マクロが意図したとおりに動作しない場合がある。
課題	VBA が最新の環境で動作しない可能性があることが報告されている(技術独立性)
開発元	Microsoft
ライセンス	特許の有無:無
メタデータの埋込可否	Excel ファイルのメタデータには、事前定義されたメタデータ要素が含まれており、ユーザー定義およびカスタムメタデータも許可されている.Excel ファイルにメタデータを埋め込む方法については、Microsoft Office の公式ドキュメントや多くの Web サイトが提供している。
参照情報	OOXML 参照
PRONOMID	fnt/445
NARA における ID	NF00665
作成ソフトウェア名	Microsoft ExcelforWindows12.0
閲覧ソフトウェア名	Microsoft Excel/Microsoft365/ Google スプレッドシート等多数あり
日本語への対応	対応済み

通番	44
フォーマット名	Microsoft PowerPoint95
バージョン	95
拡張子	ppt
公開日	1995
最終更新日	1995
バージョンの更新履歴	PowerPoint95 Version7
標準化対応状況	対応していない
特長	PowerPoint 形式は、Microsoft によって開発されたプレゼンテーション形式で

	ある。 PowerPoint95 (PowerPoint7.0 ともいう。)形式は、PowerPoint95 のネイティブ形式である。この形式は独自のものであり、Microsoft はその構造の詳細を公開していない。PowerPoint95 は、Microsoft の汎用 OLE2 複合ドキュメント形式に基づいている。PowerPoint プレゼンテーションは、複合ドキュメント形式ファイル内の「PowerPoint ドキュメント」ストリームとして保存される。
課題	OOXML との互換性
開発元	Microsoft
ライセンス	特許の有無:無
メタデータの埋込可否	埋込み不可
参照情報	仕様非公開
PRONOMID	fnt/125
NARA における ID	NF00291
作成ソフトウェア名	Microsoft PowerPoint for Windows7.0
閲覧ソフトウェア名	Microsoft PowerPoint/Microsoft365/ Google スライド等多数あり
日本語への対応	対応済み

通番	45
フォーマット名	Microsoft PowerPoint 97-2003
バージョン	8～11
拡張子	ppt
公開日	1997
最終更新日	2003
バージョンの更新履歴	PowerPoint97 Version8 PowerPoint2000 Version9 PowerPoint2001 Version10 PowerPoint2003 Version11
標準化対応状況	対応していない
特長	PowerPoint97-2003 形式は、PowerPoint97 以降のネイティブ形式である。PowerPoint97 (PowerPoint8.0 ともいう)は Office97 スイートの一部であり、他の Office アプリケーションとのより緊密な統合が特徴である。
課題	OOXML との互換性。また、現在、PRONOM は Microsoft PowerPoint Presentation97-2003 と Microsoft PowerPointShow を署名で区別できない。つまり、拡張子が pps の Microsoft PowerPointShow ファイルは、Microsoft PowerPoint Presentation97-2003PUID として識別される可能性が高くなる。さらなる研究により、これに対処できる可能性がある。
開発元	Microsoft
ライセンス	特許の有無:無
メタデータの埋込可否	PowerPoint ファイルのメタデータには、事前定義されたメタデータ要素が含まれている。PowerPoint ファイルにメタデータを埋め込む方法については、Microsoft Office の公式ドキュメントや多くの Web サイトが提供している。
参照情報	https://msopenspecs.azureedge.net/files/OfficeFileFormatsProtocols.zip
PRONOMID	fnt/126
NARA における ID	NF00284
作成ソフトウェア名	Microsoft PowerPoint for Windows8.0/8.5/9.0/10.0/11.0
閲覧ソフトウェア名	Microsoft PowerPoint/Microsoft365/

	Google スライド等多数あり
日本語への対応	対応済み

参考文献

Digital Preservation Strategy 2022-2026

<https://www.archives.gov/preservation/digital-preservation/strategy>

Sustainability of Digital Formats: Planning for Library of Congress Collections

<https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml>

Born-digital file format standards

<https://www.naa.gov.au/information-management/storing-and-preserving-information/preserving-information/born-digital-file-format-standards>

Evaluating File Formats for Long-term Preservation

<https://wiki.dnb.de/download/attachments/83791059/Evaluating+File+Formats+for+Long-term+Preservation+unpubl.pdf?version=1&modificationDate=1382441353000>

TNA Technical registry PRONOM

<https://www.nationalarchives.gov.uk/pronom/BasicSearch/proBasicSearch.aspx?status=new>

別紙3 フォーマット変換結果

注記:

以下の項番(1)～(4)における変換結果の表の見方は以下の通り。

自動:変換確認方法にて自動判定できるもの

目視:変換確認方法にて自動判定できず、ファイルを開いて目視確認が必要なもの

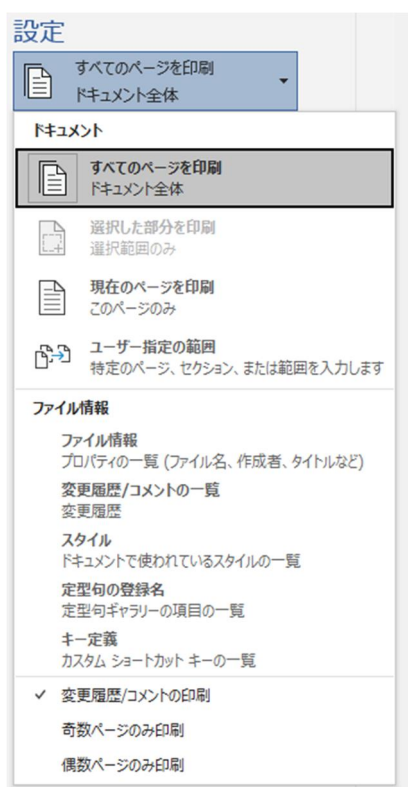
表中のグレー部分は実証試験の対象外を示す。

文字コード:各アプリケーションからテキストファイルを出力する機能であり、「テキスト内容」の確認で可能なため変換結果の確認項目としては除外とした

カスタム XML データ:アプリケーションの標準的な操作で含めることはできないため、対象外とした

なお、(1), (3), (4)における、「変更履歴および注釈」については、以下 a), b)の通り。

- a) 手動操作の場合、デフォルトで変更履歴およびコメントが出力される様になっているが、「変更履歴／コメントの印刷」チェックをオフにすることで PDF 出力時に当該情報は出力されなくなる。



- b) 自動制御する場合、Office の API を利用して、印刷する対象を指定可能であるが、印刷処理を行う際に印刷する対象をパラメータで指定可能であるため印刷時の設計に依存する。

参考リンク)

<https://learn.microsoft.com/ja-jp/office/vba/api/word.wdprintoutitem>

(1) Word97-2003 から標準的フォーマットへの変換

Word97-2003 から標準的フォーマットへの変換結果は次の通りである。詳細を表 2.3-5 に示す。

- Office 文書を旧バージョンのフォーマットから OOXML に変換した場合、微細ではあるが表示位置のずれがみられる。
- 移行されるテキストに空白が追加される場合がある。(ただし、文字が変わることはない。)
- 旧バージョンの表幅にテキストが入りきらず改行が追加されるケースがある。(その結果表が大きく崩れることが予想される)
- OOXML Transitional と OOXML Strict を比較すると「Transitional」の方がより忠実に変換できる。
- 「変更履歴および注釈」は設定または API にて出力制御可能である。(必ず出力されるものではない。)

表 1 文書変換結果 (Office)

No.	項目	確認対象	自動	目視	PDF 全般	OOXML Transitional	OOXML Strict
1	テキスト内容	○	○		○	○	空白が余分に含まれる
2	ノート						
3	ヘッダー	○	○		○	○	○
4	フッター	○	○		○	○	○
5	透かし	○	○		○	○	○
6	レイアウト	○			○	表示位置がずれる	表示位置がよりずれる
7	セルの高さと幅						
8	文字コード	○					
9	フォント	○	○		○	○	○
10	文字色	○	○		○	○	○
11	セル色						
12	文字サイズ	○	○		○	○	○
13	用紙方向	○	○		○	○	○
14	文字の方向	○	○		○	○	○
15	改ページ位置	○	○		○	○	○
16	関数の値						
17	リンク先の値						
18	集計表						
19	集計グラフ						
20	注釈	○	○		出力されない	○	○
21	変更履歴	○	○		出力されない	○	○
22	表	○	○		○	○	表内の改行がずれる
23	画像	○	○		○	○	○
24	アニメーション						
25	埋め込みデータ	○	○		○	○	○
26	埋め込みドキュメント	○	○		○	○	○
27	他ファイルへのリンク						
28	フィルター						
29	隠し文字	○		○	出力されない	○	○
30	非表示項目						
31	非表示シート						
32	カスタムXMLデータ	○					

表内の「○」は正常に変換される、または設定が引き継がれるということを意味している。

(2) 一太郎から標準フォーマット候補への変換

一太郎から標準的フォーマットへの変換結果は次の通りである。詳細を表 2 に示す。

一太郎文書から OOXML 形式に直接変換する方法は存在しない為「一太郎文書」に関しては PDF に変換する方法を採用することになる。

表 2 文書変換結果(一太郎)

No.	項目	確認対象	自動	目視	PDF 全般
1	テキスト内容	○	○		○
2	ノート				
3	ヘッダー	○	○		○
4	フッター	○	○		○
5	透かし	○	○		○
6	レイアウト	○	○		○
7	セルの高さと幅				
8	文字コード	○			
9	フォント	○	○		○
10	文字色	○	○		○
11	セル色				
12	文字サイズ	○	○		○
13	用紙方向	○	○		○
14	文字の方向	○	○		○
15	改ページ位置	○	○		○
16	関数の値				
17	リンク先の値				
18	集計表				
19	集計グラフ				
20	注釈	○	○		出力されない
21	変更履歴	○	○		出力されない
22	表	○	○		○
23	画像	○	○		○
24	アニメーション				
25	埋め込みデータ	○	○		○
26	埋め込みドキュメント	○	○		○
27	他ファイルへのリンク				
28	フィルター				
29	隠し文字	○		○	出力されない
30	非表示項目				
31	非表示シート				
32	カスタムXMLデータ	○			

表内の「○」は正常に変換される、または設定が引き継がれるということを意味している。

(3) Excel97-2003 から標準的フォーマットへの変換結果

Excel97-2003 から標準的フォーマットへの変換結果は次の通りである。詳細を表 3 に示す。

- EXCEL 文書に関しては、特に問題なく移行できた。
- 表示結果に関しても、ピクセル単位で一致した。
- 「変更履歴および注釈」は設定または API にて出力制御可能である。(必ず出力されるものではない。)

留意事項

一般的に EXCEL ファイル自体が出力するプリンタの情報を内包しているため、同一のプリンタを内包している場合の結果が上記結果である。

環境により、出力先のプリンタの設定が変わる(具体的には解像度、印字可能領域など)で表示結果が変わるため、留意が必要である。

表 3 表計算変換結果

No.	項目	確認対象	自動	目視	PDF全般	OOXML Transitional	OOXML Strict
1	テキスト内容	○	○		○	○	○
2	ノート						
3	ヘッダー	○	○		○	○	○
4	フッター	○	○		○	○	○
5	透かし						
6	レイアウト						
7	セルの高さと幅	○	○		○	○	○
8	文字コード	○					
9	フォント	○	○		○	○	○
10	文字色	○	○		○	○	○
11	セル色	○	○		○	○	○
12	文字サイズ	○	○		○	○	○
13	用紙方向	○	○		○	○	○
14	文字の方向	○	○		○	○	○
15	改ページ位置	○	○		○	○	○
16	関数の値	○	○		○	○	○
17	リンク先の値	○	○		○	○	○
18	集計表	○	○		○	○	○
19	集計グラフ	○	○		○	○	○
20	注釈	○	○		出力されない	○	○
21	変更履歴						
22	表						
23	画像	○	○		○	○	○
24	アニメーション						
25	埋め込みデータ	○	○		○	○	○
26	埋め込みドキュメント	○	○		○	○	○
27	他ファイルへのリンク	○		○	○	○	○
28	フィルター	○			最終フィルタ状態	○	○
29	隠し文字						
30	非表示項目	○		○	出力されない	○	○
31	非表示シート	○		○	出力されない	○	○
32	カスタムXMLデータ	○					

表内の「○」は正常に変換される、または設定が引き継がれるということを意味している。

(4) PowerPoint97-2003 から標準的フォーマットへの変換結果

PowerPoint97-2003 から標準的フォーマットへの変換結果は次の通りである。詳細を表 4 に示す。

- PowerPoint 文書に関しては、特に問題なく移行できた。
- 表示結果に関しても、ピクセル単位で一致した。
- PDF にはアニメーションは移行されない。
- 「変更履歴および注釈」は設定または API にて出力制御可能である。(必ず出力されるものではない。)

表 4 プレゼンテーション変換結果

No.	項目	確認対象	自動	目視	PDF 全般	OOXML Transitional	OOXML Strict
1	テキスト内容	○	○		○	○	○
2	ノート	○	○		出力されない	○	○
3	ヘッダー	○	○		○	○	○
4	フッター	○	○		○	○	○
5	透かし						
6	レイアウト	○	○		○	○	○
7	セルの高さと幅						
8	文字コード	○					
9	フォント	○	○		○	○	○
10	文字色	○	○		○	○	○
11	セル色						
12	文字サイズ	○	○		○	○	○
13	用紙方向	○	○		○	○	○
14	文字の方向	○	○		○	○	○
15	改ページ位置						
16	関数の値						
17	リンク先の値						
18	集計表						
19	集計グラフ						
20	注釈	○	○		出力されない	○	○
21	変更履歴						
22	表	○	○		○	○	○
23	画像	○	○		○	○	○
24	アニメーション	○		○	最終状態で出力	○	○
25	埋め込みデータ	○	○		○	○	○
26	埋め込みドキュメント	○	○		○	○	○
27	他ファイルへのリンク						
28	フィルター						
29	隠し文字						
30	非表示項目						
31	非表示シート						
32	カスタムXMLデータ	○					

表内の「○」は正常に変換される、または設定が引き継がれるということを意味している。