

平成 20 年度「電子公文書等の管理・移管・保存・利用システムに関する調査」概要版

1. はじめに

本報告書は、平成 20 年度に内閣府が実施した「電子公文書等の管理・移管・保存・利用システムに関する調査」の成果をとりまとめたものである。

内閣府においては、電子公文書等の国立公文書館への移管及び保存を平成 23 年度から開始するという計画に基づいて、平成 22 年度までの間段階的に、国立公文書館における電子公文書の長期保存に適したフォーマット等に関する実証的実験やフォーマット等の策定、電子公文書の移管・移送等に係るルールを検討、国立公文書館における長期保存システムの構築等を実施することとしている。このうち、平成 20 年度においては、プロトタイプシステム等により、移管された電子公文書の受入れから一般利用までのプロセス全体を見渡した調査・研究を行うこととした。

2. 平成 19 年度までの内閣府及び国立公文書館による取組み

電子公文書等の移管の開始へ向けた内閣府の取組みは、平成 17 年度から着手された。

「電子媒体による公文書等の管理・移管・保存のあり方に関する研究会」が、平成 17 年 6 月から 18 年 3 月まで計 7 回、内閣府において開催された。同研究会による専門的・技術的見地に基づく検討の成果は、平成 18 年 4 月、「公文書等の適切な管理、保存及び利用に関する懇談会」に報告された。同懇談会は、この報告に総合的検討を加えて、平成 18 年 6 月、「中間段階における集中管理及び電子媒体による管理・移管・保存に関する報告書」（以下「懇談会報告書」という。）をとりまとめて、内閣官房長官に提出した。

懇談会報告書の提言内容は多岐にわたるが、電子公文書の管理、移管及び保存のあり方について提示した 9 つの「基本的視点」のうち、次の 3 点が鍵になっている。

第 1 に、電子公文書を、紙、マイクロフィルム等の非電子媒体に変換せず、電子媒体のまま保存する原則を明確に打ち出している。また、電子公文書そのもの、すなわち、公文書の内容・コンテキスト（背景・状況・環境）・構造・機能と、記録媒体とを明確に区別し、電子公文書そのものを保存対象とするとしている。

第 2 に、電子公文書の記録としての価値を維持するのに必要不可欠な「エッセンス」を保存するという考え方を提唱し、電子公文書内容及び作成のコンテキストは「エッセンス」として保存する必要があるが、電子公文書の構造・機能等については、記録様式等に応じて、その「エッセンス」を特定した上で、これを保存することが適切であるとの考え方を示している。

第 3 に、電子公文書の長期保存の安定性・効率性等の観点から、作成時から歴史公文書としての保存・利用段階までのライフサイクル全体の管理を行う必要があり、そのメタデータ、ファイル・フォーマット（以下「フォーマット」という。）等については、公文書の

作成前から適切な対応をとる必要があると提言している。

以上のような視点に基づいた検討を踏まえて、懇談会報告書は、

- ・ 保存すべき「エッセンス」の範囲の明確化
- ・ 長期保存に適したフォーマットの策定
- ・ 電子公文書の「原本性」確保のための方法の検討

などについて、より一層専門的かつ実証的な研究を直ちに開始すべきであると提言した。

そこで、内閣府では、平成 19 年度に、電子公文書のメタデータ及びフォーマットのあり方等について、実証実験を含む調査研究を実施した。

その結果、長期保存に必要なメタデータのうち、記録の作成段階で付与し得るメタデータのかなりの部分が、比較的簡便なツールによって電子ファイルのプロパティ情報から抽出できることが分かった。

また、特定の OS に依存する文字コード及び文字フォントを用いて作成した電子ファイルを別の OS で閲覧しようとする、コード及びフォント体系の違いから、漢字の字体が変化してしまう場合が発生したほか、文書レイアウトや図表の乱れ、プロパティ情報の喪失などの事例も確認された。

フォーマットの在り方については、ワープロソフト、表計算ソフト及びプレゼンテーション・ソフトによって作成されるテキスト文書形式の電子記録を長期保存に適したフォーマットに変換する実験では、国の各機関で電子記録作成に主に用いられているファイル・フォーマットの多くが、国際標準化(ISO19005-1)され、長期保存フォーマットとして各国政府で採用が相次いでいる PDF/A に変換できることが確認された。他方、音声、画像、動画などのフォーマット変換では、品質の劣化、ファイル容量の増大などの問題があることが分かった。

内閣府では、以上の取組みを踏まえて、平成 20 年度には、平成 23 年度に運用を開始するシステムの実際の運用及び業務フローを想定しつつ、(1)受入れ・保存管理機能、(2)行政利用機能及び(3)デジタルアーカイブ連携機能の 3 つのフェーズごとに、実態を把握し、運用上の問題を明らかにすると同時に、システム設計・構築上の要件等を抽出・整理することとした。

3. 本報告書の概要

平成 20 年度の調査においては、上記の 3 機能により構成されるプロトタイプシステムを構築し、検証を実施した。

まずプロトタイプシステムで用いるメタデータ項目及びスキーマの定義を行い、

- ・ 記録管理メタデータ(電子公文書等の作成時の府省名、分類、文書名、作成者等行政文書ファイル管理簿の項目並びに、その他文書内容に係る情報を管理するための情報)
- ・ 技術的メタデータ(電子公文書等の電子ファイルのファイル形式、作成アプリケーション名称とそのバージョン、作成アプリケーションが動作する OS、電子署名やタイム

スタンプ等の作成時の技術的情報)

- ・ アーカイバルメタデータ(電子公文書等の移管後に、非現用文書として管理する受入情報、保存延長情報、公開情報、破棄情報等の情報)

の各メタデータについて、受入・保存管理機能において自動生成が可能であるかの検証を行った。その結果、記録管理メタデータは行政文書ファイル管理簿の情報からある程度の自動生成が可能であり、技術的メタデータはファイルのプロパティ情報から一部は自動的に生成できるものの、移管時にメタデータとして付与されることが期待される項目も存在した。アーカイバルメタデータについては移管後の管理のための情報であるため、自動生成できない項目が多いが、記録管理メタデータから流用することである程度カバーできた。

また、電子公文書等のフォーマットについては、移管時のフォーマット(オリジナルフォーマット)のうち、平成 19 年度までの調査研究において推奨されたものを標準フォーマットとして位置付け、標準フォーマットから長期保存フォーマットへの変換と、変換結果の検証を行った。テキスト文書形式から長期保存フォーマットへの変換では一部機能が利用できなくなるケースや、オリジナルフォーマットのフォントが完全に再現されなかったり、レイアウトが崩れたりするケースがあり、音声・音楽、映像については変換に失敗するケースも見受けられ、システム設計・構築において留意する必要があることが分かった。

次に、行政利用機能では、任意文字列検索、詳細検索、階層検索の 3 方式の検索用インタフェースを用意し、検索機能と閲覧機能の検証を行った。プロトタイプシステムでは性能面での問題は見られなかったが、システム設計・構築では将来的に移管される電子公文書等の件数を考慮する必要がある。また、実際の利用者である各府省担当者の意見も踏まえることが望まれる。

デジタルアーカイブ連携機能では、電子公文書の内容に公開に不適切な情報(個人情報など)が含まれている場合に情報をマスキングするための実現方式を調査し、技術的な課題はクリアされていることを確認できたが、運用方法等の詳細はシステム設計・構築と併せて整備されるべきである。

今後、移管媒体、フォルダ構成についてのルール化、文書作成時のガイドラインなどが進められることにより、システム設計・構築がスムーズに行われることが期待される。

4. おわりに

平成 20 年度までの検討の成果を基にして、内閣府では、今後、電子公文書の移管及び保存の開始へ向けて、移管等に係るルール化等の取組みを行っていくこととしている。また、国立公文書館では、電子公文書等の移管・保存・利用システムの設計及び構築を平成 21 年度から実施していくこととしている。