

CA1569

動向レビュー

米国政府刊行物アーカイビングの進展

1. 走り続けるGPO

米国政府印刷局（GPO）の活動が目ざましい。CA 1548で既報のとおり、GPOは2004年12月に将来計画『21世紀への戦略ビジョン』を発表し、保存と利用の拡大のため、媒体種別や時代を問わず、全ての政府刊行物をデジタル形式で提供するというビジョンを描いた。このビジョンを実現するため、矢継ぎ早に実行計画や実施案を策定・公表し、一般からのコメントを公募している。本稿では、CA1548の後の展開を追い、GPOの目下の取り組み、特にFDLP ECと呼ばれる連邦政府刊行物寄託図書館制度（FDLP）による電子コレクション（Electronic Collection）について、概要、課題などを紹介する。

2. ナショナル・コレクションの構築

GPOの究極の使命は「アメリカを知らしめ続けること（Keeping America Informed）」である。この使命のもとで、米国政府刊行物の刊行だけでなく、米国政府刊行物の全国書誌の作成、また連邦政府刊行物寄託図書館（FDL）やGPO Access（CA1149参照）を通じた一般への米国政府刊行物の提供業務を行っている。2005年4月、GPOは『2005-2006会計年度 情報提供・普及（Information Dissemination: ID）実行計画』のドラフトを発表し、これらの業務を支援し、デジタル化に対応させていくための基盤として、政府刊行物のナショナル・コレクション（National Collections）の構築を掲げた。

ナショナル・コレクションとは、すでにFDLに寄託され一般に提供されている資料、GPO Accessで提供している資料に加え、全国書誌の対象となるその他

の資料、すなわち、各政府機関がGPOを通さずに刊行した資料や、FDLへの寄託対象ではない資料—機密文書、業務用文書なども含めた、すべての米国政府刊行物のコレクションを指す。

表1は、コンテンツの媒体種別と目的によるナショナル・コレクションの分類である。有形出版物、デジタル出版物のそれぞれで、一般利用に供されるコレクションと、保存用のコレクションを構築するという。

一般利用コレクションと保存コレクションの関係を、有形出版物を例にとって説明すると、利用者がまず最初にアクセスするのが一般利用コレクション、さしあたって最寄りのFDLのコレクションとなる。ここにはない場合には、地域レベルのFDL、さらにライト・アーカイブ^①という順で探すことになる。このいずれでも見つからない場合に限り、保存コレクションであるダーク・アーカイブからデジタル媒体かFAXコピーが作成され、供される。ダーク・アーカイブは最後の砦（last resort）である。

2005年度には、この有形出版物用のダーク・アーカイブを1か所立ち上げる予定であり、各図書館や関係機関と調整を行っている。有形出版物用のダーク・アーカイブは、安全性を考慮してワシントンD.C.を避けた東西両地域で1か所ずつ構築することとされている。このうちの一方の運営に関して、米国国立公文書館（NARA）との協議が行われている。NARAが保管しているGPO提供資料をダーク・アーカイブに転用し、ダーク・アーカイブ用の場所に移して、適切に保存するというのである^②。

また有形出版物については、向こう3~5年かけて、完全なコレクションとしていったん集めた上で、保存とデジタル化の措置を取りたい、としている。収集方法としては、NARA資料の転用のほか、FDLからの移管や寄贈、各政府機関がGPOに委託して刊行する印刷物の印刷数にダーク・アーカイブ分の2部を追加

するなどが予定されており、このための資料の輸送費や倉庫・書架等の保存設備用の費用が計上されている。

なお現段階では、この有形出版物のナショナル・コレクションの維持費として、向こう5年間は1年あたり150万ドル（約1億5千万円）程度かかると見込まれている。

3. FDLP ECの構築

ナショナル・コレクションのデジタル出版物の部分、保存コレクションと一般利用コレクショ

表1 ナショナル・コレクションの概念的概略

	保存コレクション	一般利用コレクション
有形出版物	保存用コピーからなるダーク・アーカイブ	最小限の利用は許容する ライト・アーカイブ 従来のFDLコレクション
デジタル出版物	デジタル保存マスターからなるダーク・アーカイブ	GPOやパートナー機関がコピーを提供するサービス用 アーカイブ

出典：『2005-2006会計年度 ID実行計画』p.13より。
説明の都合上、筆者が一部補足・省略した。

表2 FDLP ECの機能的概略

	保存コレクション	一般利用コレクション
寄託またはハーベスティングで収集したボーン・デジタル資料 (各政府機関が作成)	<ul style="list-style-type: none"> デジタル保存用マスター CD/DVD等の媒体 バックアップ用にコピーした印刷物等の媒体 	公式／真正なコピーをサービスアーカイブから提供 <ul style="list-style-type: none"> GPO Access パートナーのサイト
デジタル化資料 (GPOが作成)	<ul style="list-style-type: none"> スキャンして作ったデジタル保存用マスター 	公式／真正なコピーをGPOや各機関が管理するサービスアーカイブや、他のリポジトリから提供
デジタル化資料 (FDL等が作成)	<ul style="list-style-type: none"> スキャンして作ったデジタル保存用マスター 	真正なコピーをサービスアーカイブや、他のリポジトリ、FDL等から提供

出典：『2005-2006会計年度 ID実行計画』pp.19-20より。
説明の都合上、筆者が一部補足・省略した。

ンの総体は、特にFDLP ECと呼ばれている。2005年度は、このFDLP ECの調査費も計上されている。

FDLP ECの構想は、1998年にさかのぼる⁹⁾。この構想を、85%以上の政府刊行物がオンラインで提供されているという実態、5年以内にFDL寄託資料の全てをデジタル化できるだろうという予測にあわせて改訂した。

表2は、コンテンツの媒体種別と目的によるFDLP ECの分類である。原則的には有形出版物と同じく、永久保存コレクション（ダーク・アーカイブ）と、一般利用コレクションに分かれている。CDやDVDといった媒体で刊行された出版物も、このFDLP ECの一部に含まれる。

有形出版物と異なる点として、コレクションの対象が挙げられる。有形出版物の場合、FDLへの寄託対象ではない資料も、米国政府刊行物の全国書誌の対象であるとしてダーク・アーカイブに収集される。ところが、FDLP ECについては、収集対象は文字どおり、FDLへの寄託対象資料に限られている。収集方法としては、すでにGPO Accessで提供している資料の利用、各機関のウェブサイトからのハーベスティングや寄贈などのほか、FDL所蔵資料のデジタル化が予定されている。

デジタル出版物に特有の問題としては、当該資料の来歴の確認が挙げられている。GPOは、FDLP ECの全ての資料について真正性の保証を行い、また可能な限り公式な資料を提供する、としている。これらの来歴はメタデータに記録される。またファイルフォーマット

トについては、さしあたり文字テキストと画像を表現するものとし、TIFF、PDF、TMLやASCIIフォーマットなどが予定されている。将来的には動画や音楽なども含むという。

デジタル出版物のダーク・アーカイブは複数のオフライン環境で構築され、サーバ、ストレージとも複数用意される。セキュリティに配慮するとともに、災害等が発生しても継続して保存・提供できる仕組みが必要である。また詳細は検討中とされているが、ダーク・アーカイブ用の資料については、バックアップの

ために印刷物やマイクロフィッシュなどの有形出版物を用意するとしている。

4. FDLP ECの具体化(1)－有形出版物のデジタル化

実は、4月段階の『2005-2006会計年度 ID実行計画』案では、全8章のうち、ナショナル・コレクションに関する1つの章と、米国政府刊行物の全国書誌の課題を記載した1つの節のみが完成しており、資料の入手、メタデータの基準、版の管理、真正性の保証、提供サービスなど、多くの部分がペンディングとして先送りされている。

この一方で、有形出版物のデジタル化事業については、先行して仕様の具体化が行われている。2007年末までに全体の70%をデジタル化するというビジョンを実現するために、2004年、有識者を招き、デジタル保存用マスターファイルのフォーマットと、メタデータについての検討会が行われた。

保存用のフォーマットとしては、他機関の先行事例等をもとに検討した結果、非圧縮または可逆圧縮のTIFF形式が推奨された。1ページごとに1ファイルを作成し、白紙もスキャンされる。さらに白黒テキスト、イラスト付きテキスト、カラーページなどの種別に応じた解像度、色調、色などの推奨値も提示されている。

またメタデータの表現方法としては、現在、図書館の多くがMARCを採用している一方、Dublin Coreなどのような簡易な記述を採用している図書館・機関もある。これらのデータを別々に維持するのではなく、METS (Metadata Encoding Transmission Standard;

CA1552参照)でラッピングする⁽⁴⁾ことにより、記述レベルが異なっても一つのデータベースとして管理・提供できるようにすることとされた。また固有タイトル・MARC番号・固有ID・来歴・派生物の情報・財産権の情報など、メタデータ・パッケージに含めるべき11の要素も提示されている。これらの報告書も公表されており、意見募集を行っている。

2005年に入ってから、6月にデジタル化用機器の仕様書案の第3版、7月には『デジタル化資料優先順位』のドラフトが発表された。後者は、2004年に行った調査をもとに、媒体としてはまず印刷物から始め、資料としてはすでにデジタル化されていて利用が多いものやGPO Accessで提供しているものについて、1990年から現在までのものをオンラインで見られるようにバックファイルを作成することから始める、としている。

このように、着々とデジタル化の準備を進めてはいるものの、印刷物だけでおよそ220万タイトル、6,000万ページもある。目標達成に向けて、仕様策定、計画の具体化などをよりいっそう進めなければなるまい。

5. FDLP ECの具体化(2)ーボーン・デジタル資料のハーベスティング

有形出版物のデジタル化については、上述のとおり以前から準備されてきたところであるが、2005年7月、突如として、IDプログラム73号として政府刊行物のデジタル出版物を一括収集するという方針声明が提示された。これは、FDL制度と全国書誌のために、各政府機関のウェブサイトで一般公開されているデジタル出版物を、合意が得られれば機械的に、得られなくても人手でハーベスティングするというものである。

このハーベスティングの対象は、各機関のサイトのほか、各機関の上位機関や協力関係を結んでいる大学等の機関のサイトも含まれている。さらに、すでに各機関のウェブサイトですり取られてしまっているものについては、インターネットのアーカイビングを行っている非政府機関のサイトなどからも入手するとしている。なお、ハーベスティングの頻度やタイミングについては、各機関の助言に従うことになっている。

収集した刊行物に対しては、正当に取得したアーカイブ用ファイルであることを示すメタデータを付与する。このときに、当該の刊行物が非公開である、制限があるなどと判明した場合には、各機関に問い合わせの上、サーバから永久に削除される。また、各機関は、政府情報資源のGPOプログラムからの削除の手続きを定めたIDプログラム72号に従って申し出を行うことで、当該の刊行物の削除を依頼することができる。

GPOは各機関に対し、事業の目的と機械的なハーベスティングの可否を文書で尋ねており、30日以内に連絡がない場合には、承認されたものとして機械的に

ハーベスティングするという。米国の著作権法は、政府刊行物には適用されないため、このような強権的とも思われる措置が可能なのである。これによって、有形出版物のデジタル化よりも先に、ハーベスティングが始まることとなった。

6. FDLの苦悩

このようなGPOの、大変ではあるが華やかな事業の一方で、FDLの存在意義が問われている。FDLという制度の名のもとで、ナショナル・コレクションの構築事業が推進されているにもかかわらず、資料がデジタル化されオンラインで提供されるようになると、FDLを経由する必要性に乏しくなる。

寄託図書館協会 (Depository Library Council) は7月、10月に行われるFDLの秋季大会に向けて、『21世紀への戦略ビジョンに対して』と題するペーパーを作成した。これは、FDLとしての役割は何か、利用者からはどのように見られているのか、何を収集するのか、どのような付加価値を提供できるのかなど、FDLだけでなくすべての図書館にとっての存在理由にかかわる問いで構成されており、FDLの図書館員に対し、専用のブログや大会で回答するよう希望している。

2005年8月時点でFDLは約1,250館であり、徐々に数が減っている。この秋季大会でどのような展望—あるいは生き残り戦略—が出されるか、注目される場所である。

7. おわりに

ナショナル・コレクションの構築に向けて邁進するGPOであるが、別途電子公文書アーカイブ (Electronic Records Archives: ERA) 計画を進めているNARAとの協力関係はうまく行くのか、各政府機関からのハーベスティングはどのように進んでいくのか、先送りになっている諸仕様は詰まっていくのかなど、不安要素はまだ多いように思われる。このプロジェクトをいかにして成功に導いていくのか。引き続き、GPO、またGPOを取り巻く周囲の今後の動向に注目したい。

(関西館事業部図書館協力課：^{むらかみこうすけ}村上浩介)

- (1) ライト・アーカイブは、ダーク・アーカイブを補完する位置付けの、原則としては保存用だが最小限の利用は許容されるというものである。また将来構想として、寄託図書館コミュニティによる共有リポジトリの創設も視野に入れられている。
- (2) GPOは、政府機関の活動記録として、GPOが目録を作成した政府刊行物一式を、NARAに移管している。
- (3) 古賀崇. アメリカ連邦政府寄託図書館制度の電子化への過程とその背景. 日本図書館情報学会誌. 46(3), 2001, 111-127.
なお、この論文は本稿の「前史」を簡潔に紹介したものと非常に有益である。
- (4) ラッピング (wrapping) とは、原則として既存のデータ構造やアプリケーションを維持しつつ、外部システムから、元のものとは異なる形式 (例えば、外部システムが採用している標準的な形式や新

世代の形式など) のデータやアプリケーションとして取り扱うことができるように、元のものを「包み込む」仕組みを指す情報処理用語である。

- Ref: GPO. Information Dissemination Implementation Plan, FY 2005-2006. Draft. 2005-04-14, 41p. (online), available from <http://www.access.gpo.gov/su_docs/fdlp/pubs/IDPlan.pdf>, (accessed 2005-08-29).
- GPO. A Strategic Vision for the 21st Century. 2004-12-01, 20p. (online), available from <<http://www.gpo.gov/congressional/pdfs/04strategicplan.pdf>>, (accessed 2005-08-29).
- GPO. Report on the Meeting of Experts on Digital Preservation. [2nd Version]. [2004-06-18], 14p. (online), available from <<http://www.gpoaccess.gov/about/reports/preservation2.pdf>>, (accessed 2005-08-29).
- GPO. Report on the Meeting of Experts on Digital Preservation: Metadata Specifications. [2nd Version]. [2005-06-02], [18p]. (online), available from <http://www.gpoaccess.gov/about/reports/metadata_report_final.pdf>, (accessed 2005-08-29).
- GPO. Information Dissemination Implementation Plan: Priorities for Digitization of Legacy Collection. Draft. 2005-07-01, 6p. (online), available from <http://www.gpoaccess.gov/about/reports/digitization_plan_070505.pdf>, (accessed 2005-08-29).
- [GPO]. Harvesting Federal Digital Publications for GPO's Information Dissemination (ID) Programs. [2005-07-05], 3p. (Information Dissemination Policy Statement, ID 73). (online), available from <http://www.access.gpo.gov/su_docs/fdlp/pubs/policies/id73_07-06-05.pdf>, (accessed 2005-08-29).
- Depository Library Council. Toward a Vision of the Government Information Environment of the 21st Century: A Draft Outline. 2005-07-22, [3p]. (online), available from <http://www.access.gpo.gov/su_docs/fdlp/pubs/dlc_vision_outline.pdf>, (accessed 2005-08-29).

視覚障害その他の理由でこの本を活字のままでは読むことのできない人の利用に供するために、この本をもとに録音図書（音声訳）、拡大写本又は電子図書（パソコンなどを利用して読む図書）の作成を希望される方は、国立国会図書館まで御連絡ください。

連絡先 国立国会図書館総務部総務課
住 所 〒100-8924
東京都千代田区永田町1-10-1
電話番号 03-3506-3306