

国家デジタル図書館の技術実現

中国国家図書館自動化部主任
国家デジタル図書館
プロジェクト技術指導グループ長 魏大威

一、実現の目標

国家デジタル図書館プロジェクト¹は、我が国の文化界における非常に重要な事業である²。国家図書館は、国家デジタル図書館プロジェクト建設を通じて、我が国の文化建設における地位と役割とを十分に発揮し、インターネット環境における国家図書館の機能をより良く果たし、国内外への影響力を増強し、利用サービスのレベルとサービス能力を極大にまで上げ、中国語のデジタル資源に対する社会の要求を最大限に満たすことになる。

国家デジタル図書館プロジェクトの目標は、「国家図書館の機能を全面的に発揮し、中国語デジタル情報を重点的に収集し、構築し、長期保存を図ること」、「インターネット上に大規模中国語デジタル資源のアーカイブを構築すること」、「中国語デジタル情報の保存拠点とサービス拠点となること」、「デジタル資源の収集、加工、保存技術をサポートするプラットフォームを構築すること」、「国家基幹ネットワークを通じて、国内外の公衆に中国語デジタル情報を提供すること」、「国家デジタル図書館の科学的管理システム、デジタル資源の構築システム、デジタル資源の発信とサービスシステム、デジタル資源の組織的管理及び保存システムを確立し、国際的先進的レベルのデジタル図書館の初期構築を行うこと」である。

二、実現の方向

国家デジタル図書館プロジェクトの建設過程では、主要四方面において、重点的にデジタル資源とサービスシステムの構築事業を進行する。

インターネット上のデジタル資源の組織化に基づいた公共文化情報サービスシステムを

¹ http://ndl.go.jp/jp/aboutus/data/pdf/nlc26_6_shen.pdf 第 26 回日中業務交流テーマ報告 デジタル図書館管理所副所長 申曉娟「中国国家図書館デジタル図書館建設と伝統的図書館業務の調整」

² <http://www.dap.ndl.go.jp/ca/modules/cae/item.php?itemid=357> カレントアウェアネス-E No.62 2005.07.20 「中国国家図書館がデジタル資料の寄贈と著作権譲渡を依頼」に「国家数字図書館」のプロジェクトの一部が紹介されている。

構築する。「情報ネットワーク伝播権保護条例」³の枠組みの下、国家デジタル図書館を、大規模かつ高度な中国語デジタルアーカイブとし、インターネットを通じて、民族文化及び先進的文化の発揚と普及を実現し、インターネット利用の文献情報資料取得への公衆の要求を満たすこととする。

国家図書館をサービスセンターとし、国内大規模公共図書館をサービスポイントとしたデジタル資源の伝達、サービスシステムを形成し、国家図書館と国内大規模図書館のデジタル資源のシームレスな伝達とサービスシステムを構築、各職場の分館と地域分館のサービス体制を形成する。

国家デジタル図書館資源のサービス範囲を拡張し、全国文化情報資源共有プロジェクト⁴の基礎ハードウェア設備を利用し、同プロジェクトの基層センターの現況に基づき、デジタル資源の基本的サービスのプラットフォームを組み立て、基層センターの文献資料不足の現状を有効的に改善する。

国家図書館そのもののデジタル資源の構築と管理レベルを更に高め、来館利用者に先進的サービス手法と閲覧環境を提供、重点的にデジタル資源サービスセンターやデジタル加工センターなどの八大センターを建設し⁵、国家図書館を国内最大の中国語文献情報の生産とサービスのセンターとする。

三、実現への枠組み

国家デジタル図書館プロジェクト建設の目標を順調に実現するためには、情報技術の発展と国家図書館自身の情報資源管理やサービスの豊富な経験により、ハードウェア環境、ソフトウェアのシステム、所蔵コレクション、対外サービスの四方面を完全に結合し、デジタル図書館の運行体系を完備する。

1. 基礎プラットフォーム

業務上の要求に基づきホストを構築する。これにはハイパフォーマンスのコンピューター及びクラスタを含む。インターネット環境を構築する。これにはネットワーク間を接続するルーター、外部接続の光ファイバー、無線LAN、衛星及びデジタル放送を含む。館内LANにはギガビットイーサネット（GbE）⁶を標準設計として採用し、10ギガビット・イーサネット技術でネットワーク各層の高速接続の需要を満たすことを考慮に入れている。第二期館（訳注：2008年に開館する新館のこと）の閲覧室と室外休憩スペースでは、無線LAN技

³ <http://journal.mycom.co.jp/news/2006/05/31/381.html> を参照した。

⁴ <http://www.dap.ndl.go.jp/ca/images2/ca/ca289.pdf>
<http://www.dap.ndl.go.jp/ca/modules/ca/item.php?itemid=1034> カレントアウェアネス（季刊）No.289
（2006.9.20）湯野基生「中国「全国文化情報資源共有プロジェクト」の現状」に詳しい。

⁵ 注1に建設計画の説明あり。

⁶ <http://e-words.jp/w/E382AEE382ACE38393E38383E38388E382A4E383BCE382B5E3838DE38383E38388.html> から。

術を採用し、ノートパソコンを持ち込んだ利用者は、セキュリティ認証を介して、直接デジタル図書館の資源やインターネット上の資源を使用することができる。国際衛星放送のプログラムを受信し、デジタルテレビ放送技術を介することで新館館内の視聴室と第一期館（現在の館舎）のリサーチルームにおいて、セットトップボックス⁷（STB）から指定した衛星放送のプログラムを視聴し、より多くの調査情報を収集することができる。ファイバーチャネルを通じて、中国教育科学研究コンピューターネットワーク（CERNET）⁸、中国科学技術ネットワーク（CSTNET）⁹、中国共産党中央党校デジタル図書館¹⁰などのシステムに、1G（ギガ）bpsの通信を行い、現在の大規模システム間の通信障害を軽減することで、国家図書館は全国の図書館ネットワークの中心となった。大規模ストレージ¹¹を構築する。これにはファイバーチャネル技術¹²を用いたストレージエリアネットワーク（SAN, Storage Area Network）¹³、SATA¹⁴（Serial ATA）、CD-ROM庫、磁気テープ庫を含む。最も利用率の高い発信及びサービスデータはオンラインストレージ（online storage）¹⁵に保管し、その他のデータは、ニアラインストレージ(nearline storage)¹⁶に置き、ニアラインストレージでの保存期間満了後、永久保存のデータを順次オフラインストレージ¹⁷に格納することができる。このような循環で、デジタル資源の利用と保存を動態平衡状態におくことを保証できる。各システムの安定運行のため、ネットワーク管理センターに統一したネットワーク管理・システム保全・全体の連結運行のプラットフォームを構築する。

2. ソフトウェアシステム

⁷<http://e-words.jp/w/E382BBE38383E38388E38388E38383E38397E3839CE38383E382AFE382B9.html> 「テレビに接続して様々なサービスを受けられるようにする機器の総称。」

⁸ <http://www.net.edu.cn/>

⁹ <http://www.cstnet.net.cn/shouye.jsp>

¹⁰ <http://tsg.ccps.gov.cn/home.jsp?id=szzy>

¹¹ <http://e-words.jp/w/E382B9E38388E383ACE383BCE382B8.html> 「存儲」を「ストレージ」とした。

¹² <http://e-words.jp/w/Fibre20Channel.html> ファイバーチャネルはSANを構築する技術。「光ファイバー」は光ファイバーなので、「光ファイバー陣列」はファイバーチャネルと類推。

¹³ <http://www.keyman.or.jp/3w/prd/33/30001433/?vos=nkeyadww10000852> 「存域网」をSANとした。「通常のLANとは別に、ストレージ用のネットワークを構築してバックアップの効率化やストレージ統合などを実現することができるSAN(Storage Area Network)。」

¹⁴ <http://e-words.jp/w/SATA.html> 「SATA」は、コンピューターと補助記憶用ドライブとの接続方式。

¹⁵ 「在线保存」はオンラインストレージとした。

<http://e-words.jp/w/E382AAE383B3E383A9E382A4E383B3E382B9E38388E383ACE383BCE382B8.html> 国家図書館が商業ストレージを持つとも思にくい。オンライン接続用のストレージを保持するくらいの意味か。

¹⁶ 「近線保存」はニアラインストレージとした。コストパフォーマンスのためのストレージらしい。安価だが処理速度が遅いらしい。

http://en.wikipedia.org/wiki/Nearline_storage

<http://www.asahipress.com/eclub/dic/dic46.txt>

¹⁷ http://en.wikipedia.org/wiki/Offline_storage#Off-line_storage 「离线保存」は、オフラインの記憶装置と思われる。媒体は様々らしい。

デジタル資源のライフサイクルの管理は、国家デジタル図書館の収集デジタル資源管理における重要な戦略であり、この戦略の下、国家図書館はコアとなるソフトウェアシステムを構築した。主要なソフトウェアシステムは、次のとおりである。

デジタル資源収集システム。デジタル資源は、様々なフォーマット、キャリアー、媒体のデジタル資源に対して収集、選定及び登録することで館のデジタルコレクションとなる。デジタル資源収集業務の必要に基づき、電子資源の納本とネットワーク資源の収集機能を増設した。デジタル資源の収集及び業務管理においては、印刷文献と非印刷文献収集の重複防止、相互補完を図っている。

デジタル資源の加工と統合システム。デジタル資源の加工、記述、インデキシングと、メタデータ及びメタデータ作成対象データの統合により、専門情報データベースとナレッジ組織化システムを構築。

デジタル資源の組織化と管理システム。デジタル化した所蔵コレクションについては順序立った管理を行い、他機関構築のデジタル資源については効率的に調整を行い、デジタル資源サービスを確実に保障する。

デジタル資源の発信とサービスシステム。デジタル資源発信プラットフォームと統一的なデジタル資源サービスポータル構築を通じて、検索サービス、ウェブサイトサービス、電子閲覧室サービス、電子ビジネスサービス、バーチャルレファレンスサービス、図書館間貸出サービス、ドキュメントデリバリーと情報発送サービスを実現する。

3. 所蔵資源

館所蔵デジタル資源を構築する上では、資源構築の二本の柱を明確にした。一つ目は館所蔵の伝統的文献のデジタル化による構築及びこれに基づくデジタル資源の統合と開発であり、二つ目はネットワーク資源及びその他のボーンデジタル資源の収集と再加工である。2006年末までに、国家図書館が自ら作成したデジタル資源の内容は以下のとおりである。

蔵書の書誌データ 380万件
典拠データ 70万件
雑誌記事索引データ 275万件
図書全文画像データ 1億700万ページ余、約62万冊余
中国の古典籍 173部、4,700万字余
デジタル化した音声資料 約50万曲
デジタル化した映像資料 約2万件点（約5万時間）
専門主題データベース 18件

4. 対外サービス

国家デジタル図書館完成後には、主に以下のようなサービスを外部に提供する。
メタデータ集中検索サービスの実現。中国語の印刷体及びデジタル文献の書誌データ、

雑誌記事索引データ、目次データ、抄録などを国家図書館が網羅的に収集・整理して、基層図書館にメタデータのサービスを提供する。電子情報を有する文献は、すべてネットワーク上で最初の20ページ(ママ)を表示して、利用者の選択に供する。

デジタル放送及びデジタル放送リクエストサービス。デジタル図書館プロジェクト用の衛星・デジタル放送システム及び全国文化情報資源共有プロジェクトが現在保有するチャンネルを利用して、全国の図書館の貴重資料を組織化し、デジタル放送とデジタル放送リクエストサービスを開設する。例えば、国家図書館文津講壇(訳注:国家図書館古籍館が一般向けに開講している文化講座)や国立図書館館長フォーラムなどの活動は衛星とネットワークを通じて生放送ができる。

文献デジタル化サービス。国家図書館全国総合目録システムにより、各基層図書館の印刷体文献の所蔵状況に応じて、各図書館に対し文献デジタル化サービスを提供する。国家図書館は、デジタル化済みの電子資料の目録をネット上で公開し、基層図書館の選択とレファレンスに供する。基層図書館が収蔵する印刷体文献や、国家図書館が既にデジタル化した印刷体文献についてはすべて、その図書館が要求するだけで国家図書館は電子版の全文提供が可能であり、基層図書館の電子資源の保存と・提供サービスを満たすことができる。

バーチャルレファレンスサービス。国家図書館及び全国の図書館のバーチャルレファレンスサービスプラットフォームを利用して、サービス体制を共同で組織することで、各レベルの公共図書館が所在地の行政府や立法機関の政策決定に資するためのバーチャルレファレンスサービスを実現する。また、大型図書館や専門図書館が基層図書館のレファレンス担当職員のために提供するレファレンス支援や研修用プラットフォームを実現する。さらに、システムやレベルが異なるバーチャルレファレンスサービスセンターを設立することで、一定の範囲内で一般公衆のためにバーチャルレファレンスサービスを提供する。

全国レベルでの図書館デジタル資源へのオープンアクセスと統合検索を実現させ、国家デジタル図書館のデジタル資源サービスプラットフォームを利用してネットワーク上で全国の図書館のデジタル資源を統合する。全国の各種図書館のデジタル資源(公開可能な資源)のオープンアクセスと統合検索を実現させ、全国図書館検索エンジンを立ち上げ、利用者のデジタル資源検索や情報取得を満足させる。

図書館デジタルチャンネルの構築。様々な方式でデジタル資源を組織化し、異なる館種の図書館が既に知的財産権を有するデジタル資源を共有し、図書館デジタルチャンネル上で専門デジタル情報サービスを提供し、テレビとデジタルチャンネルを通じて広く公衆のために図書、新聞・雑誌、視聴覚資料のデジタル資源に関するサービスを提供する。

デジタル図書館の標準規格の共同作成・共有。「中国デジタル図書館標準規格」の基礎の上に、国家デジタル図書館プロジェクトの標準規格制定事業を結合し、全国の図書館の専門職員が共同してデジタル図書館標準規格を構築する。オープンに構築して成果を共有する。

電子メールとショートメッセージによる情報発信サービスについては、目下中国のネット人口と移动通信機器所有者人口が共に世界第一位である。このため国家デジタル図書館プロジェクトでは電子メールやショートメッセージなどの方式を利用してサービスのルートを広げ、一般市民が十分快適かつ迅速にパーソナライズされたサービスを受けられるようにする。

電子ビジネスサービス。ユーザーのためにネット決済やモバイルバンキング決済などの方式を提供し、図書館へ行かなくても国家デジタル図書館が提供する相互貸出、ドキュメントデリバリー、電子資源取得などの各種付加価値サービスを利用できる。

四．実現へのプロセス

国家デジタル図書館プロジェクトはおおまかに二段階に分けて実現する。

第一段階は、新館開館後のデジタル化管理とサービスの基本的な実現である。来館利用者にセルフ貸出・返却サービスや自館作成資源の応用サービス、OPAC検索サービスを提供する。デジタル資源加工・統合プラットフォームを開発し、ナレッジ組織化システム管理を実現し、中国語データベースを組織し、セキュリティシステムとシステム維持管理プラットフォームなどの構築を完成させ、国家図書館の館内LANにおいてすべての購入デジタル資源と著作権による制限があるデジタル資源を公開する。

まず、一般ネットユーザー向けの全方位サービスを形成する。一般公衆のためにデジタル資源ポータルサービス、ネット決済サービス、ドキュメントデリバリーサービス、バーチャルレファレンスサービスを提供し、ネット上に自館作成のデジタル資源をすべて公開する。

大型公共図書館を重点的に選んで国家デジタル図書館のサービスポイントを設置する。分散型のデジタル化文献情報資源サービス体系を構築する。

第二段階では、デジタル資源構築の速度を上げ、インターネット上のデジタル資源サービスの容量をさらに向上させる。デジタル資源収集、加工・統合、公開・サービス、保存・管理のシステム設計と応用を完成し、関連ハードウェアの調達と入札作業を完了し、より完備したデジタル図書館システムを構築することで、デジタル資源の収集、加工、統合、公開、サービス、管理、保存のニーズを満たす。

基本設備が比較的好条件の基層図書館を選び、全国文化情報資源共有プロジェクトの情報チャンネルを利用して、それらの館にデジタル資源基層サービスプラットフォームを提供する。県、郷クラスの基層図書館には専門情報デジタル資源サービスを提供し、併せて資源の構築と更新作業を行う責務を負う。

各種ソフトウェアとハードウェアの完備と検収を進め、国家デジタル図書館プロジェクトの最終段階を実施する。