

## 中国国家図書館の中国語デジタル資源構築

中国国家図書館  
副館長 陳力

### 1. はじめに

中国国家図書館は、伝統的媒体の中国語文献情報資源を世界で最も多く所蔵している。2006 年末現在、当館の蔵書総数は中国語文献が 1511 万 8936 冊（点）、外国語文献が 1058 万 5424 冊（点）である。このように豊富な文献情報資源に基づいて、国家図書館は政府、研究・教育機関及び一般公衆に対し、良好な文献情報サービスを提供し続けてきた。

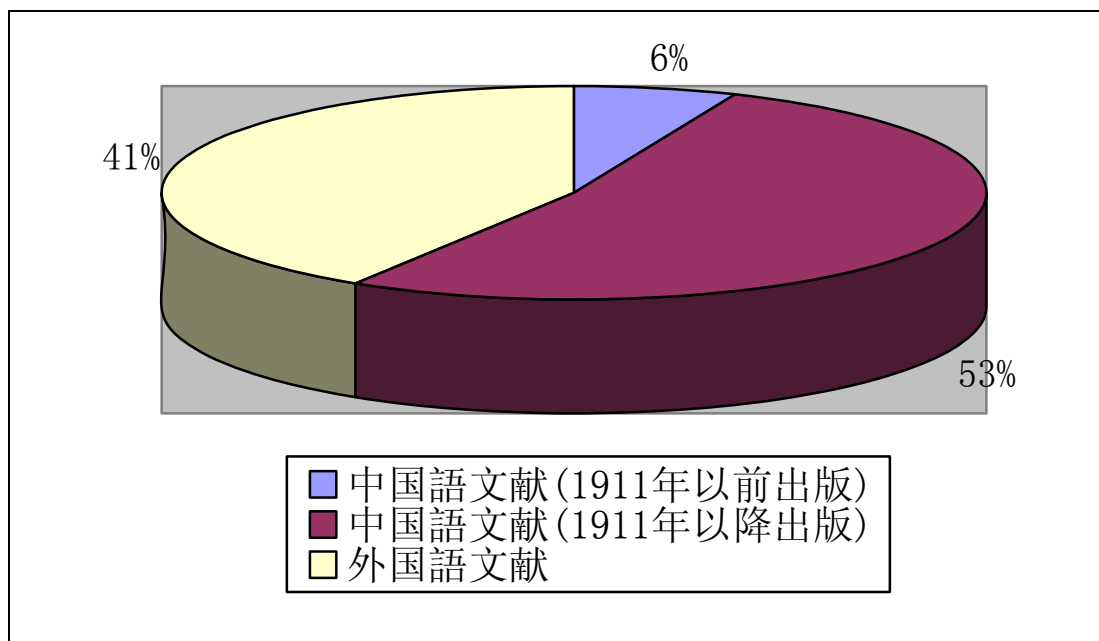


図 1 国家図書館所蔵資料の構成比

しかし、20 世紀の 80-90 年代以降、インターネットの中国における急速な発展に比べ、中国語デジタル資源構築及びサービスの進展の速度は、人々の要求よりはるかに立ち後れた。インターネット上の重要な情報コンテンツの提供者として、国家図書館は中国語デジタル資源構築及びサービスにおける自らの重要な責任を意識するようになった。そこで、1987 年から国家図書館は、デジタル資源の構築とサービス業務に力を入れ始めた。2005

年10月、国家図書館が中心となって進めている国家デジタル図書館プロジェクトの建設開始が、国務院により承認された。本プロジェクトの総目標は、世界で最大の中国語デジタル情報保存基地及び中国語デジタル情報サービス基地を構築することである。プロジェクトの実施により、中国国家図書館のデジタル資源構築及びサービスの基盤となるハード／ソフトのプラットフォームが大きく改善されるであろう。

## 2. 国家図書館のデジタル資源構築及びサービスの現状

デジタル資源の量及び質に対するユーザーの要求をより一層満足させるために、国家図書館は『2003－2005年国家図書館デジタル資源構築計画』『2006－2010年デジタル資源構築計画』を相次いで策定し、デジタル資源構築の目標を打ち立てた。その目標とはすなわち、①国立図書館としての機能を発揮し、中国語デジタル資源の収集と長期保存を重点的に行い、中国語デジタル資源の保存基地を構築すること、②中国語デジタル資源の検索基地を構築すること、③デジタル資源の収集・整理・保存・管理・サービスの技術面をサポートするプラットフォームを構築すること、④ネットワークを通じて国内外の一般公衆に中国語デジタル資源サービスを提供すること、である。

国家図書館は1987年から中国語書誌データベースの構築を開始し、2000年から計画的なデジタルデータベース構築を開始した。その基本的な考え方は次のとおりである。すなわち、①成熟した市場メカニズムを有する（＝成熟した市場で流通している）資源は、購入により国家図書館が組織化した上で一般公衆に直接サービスを提供する、②特色のある資源は、当館での所蔵状況とサービス対象からの要求に基づき、時期、資料群を分けて選択的な構築を行う、③インターネット上での送信が可能な著作権のない資源、ニーズの高いデジタル資源を優先的に構築する。

国家図書館のデジタル資源の主な収集源は次の4つのルートである。すなわち、①納本により受け入れるデジタル資源で、電子出版物・視聴覚資料の納本、冊子体資料の附録として納本される電子出版物など、②購入による視聴覚資料、中国語CD-ROM・オンラインデータベース、③当館が構築した中国語書誌・雑誌記事索引・全文画像・マルチメディアの各データベース、及び購入または納本された音声・映像資料のデジタル化、④ネットワーク資源の収集。

### (1) 当館が構築したデジタル資源の概況

2006年末現在、国家図書館が構築したデジタル資源のコンテンツは次のとおりである。

- 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余万ページ、約62万余冊（点）

- 中国の古典籍 173 タイトル、4,700 余万字
- 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 万点 (約 5 万時間)

当館が構築したこれらのデジタル資源のうち、中華文化の古典的作品と特色ある所蔵資料が主要な部分を占めているが、中でも特色のあるものとして、以下のデータベースが挙げられる。

- デジタル地方志データベース

当館所蔵地方志の全文画像 4000 余タイトル 7 万余巻、248 万コマを搭載。そのうち 2,000 タイトル 3 万 4,000 巻、120 万コマは全文テキスト変換済みである。現在既に完成している画像データは、インターネット上で全世界の利用者に最初の 24 ページの閲覧サービスが提供されている。<http://res1.nlc.gov.cn:9080/chronic/index.htm> 国家図書館デジタル図書館プロジェクトの建設に伴い、これらの加工済み資源は次々と全てインターネット上でサービス提供されることになる。

- 当館所蔵甲骨実物画像及び拓本実物画像データベース

国家図書館は各種の石刻拓本を 3 万余種、13 万余点所蔵し、現在既に約 2 万種がメタデータ付与済み、2 万 5,000 コマの画像が作成済みである。当館所蔵の甲骨約 6,000 種は、現在既に全てメタデータ付与が完了し、画像 1 万 2,000 コマの作成が完了している。以上のデータは既に全てインターネット上で全世界の利用者にサービス提供されている。

<http://res2.nlc.gov.cn:9080/jiagu/index.htm>

- 国際敦煌プロジェクト

2001 年 3 月 7 日、中国国家図書館と英国図書館は期間 5 年の国際敦煌プロジェクト (IDP) の協力事業を開始した。2002 年 11 月 11 日、IDP プロジェクトの中国語ウェブサイトが国家図書館で正式に公開された。現在、本プロジェクトでは敦煌写本 652 巻のデジタル化が既に完成し、インターネット上で全世界の利用者に検索・閲覧サービスが提供されている。

<http://idp.nlc.gov.cn/>

- 当館所蔵西夏文献データベース

国家図書館の所蔵する西夏文献は大部分が西夏・元代の孤本であり、700 年から 800 年の歴史を有する。それは宋・元代の貴重書と同じ時代に属するものである。当館所蔵西夏文献は、古代西夏・元代の紙や絹の材料と製法を研究するための貴重な実物資料であり、西夏仏教史の研究にとっても重要な価値を有している。国家図書館は 2002 年にこの貴重な所蔵資料と関連研究著作のデジタル化を開始した。現在、当館所蔵西夏文献の書誌データ 125 件、当館所蔵西夏文献の原資料の画像約 5,000 コマ、西夏研究論文の論題索引データ 1,200 余件、西夏研究論文の全文スキャン画像 9,700 余コマが搭載され、全てインターネット上で全世界のユーザーにサービス提供されている。

<http://res2.nlc.gov.cn:9080/wenxian/index.htm>

- 「年画」データベース

「年画」とは中国の伝統的な絵画様式で、中国の人々から深く愛されている。中国国家図書館では長年にわたって収集してきた 4,000 枚余の年画作品について、2006 年からデジタル画像化作業を進めている。それぞれの年画には、物語のあらすじ、歴史背景、作品の特徴、関連知識の紹介を中心とした説明文を付している。現在、全部で 163 件 190 枚の年画作品がインターネット上で公開されている。

(<http://res2.nlc.gov.cn:9080/nianhua/index.htm>)

●当館所蔵博士論文データベース

中国国家図書館は、我が国の学位論文を網羅的に収集整理する国家指定の専門機関であり、かつポストドクターの研究報告の網羅的な収集に責任を負っている専門機関である。現在、中国国家図書館では博士論文を 12 万件近く所蔵しており、中国国内における博士論文の最大の所蔵機関となっている。このほか、一部の大学院の修士学位論文も所蔵している。

学位論文の保存と利用の利便性を高めるため、中国国家図書館では 2004 年に学位論文全文画像データベースの構築を開始した。現在では 99,000 冊（点）の博士論文の電子化作業を完了し、全文画像データは 1,200 万ページに達している。これらの電子化作業が完了した博士論文について、インターネット上で各論文の先頭 24 ページを閲覧できるサービスを提供している。

(<http://202.96.31.40:9080/doctor/index.htm>)

●当館所蔵民国期雑誌データベース

民国期中国語雑誌は中国国家図書館が保存している中華民国時代の文献の重要な構成部分である。文献の紙質の原因により、これらの文献の保存状況は楽観視できない状態にある。これらの文献の効果的な保存と利用のため、中国国家図書館では早くも 1991 年には当館所蔵民国期雑誌のマイクロフィッシュの製作を開始した。1999 年からはマイクロフィッシュの電子化・スキャニング作業を始めており、3 年以内に 600 万コマ近くのマイクロフィッシュの電子変換を完成させる見込みである。現在までに 2,800 余タイトル 8 万号余りの中華民国雑誌のデジタル化作業を完了し、全文画像データは 480 万余ページに達している。現在、インターネット上で全世界のユーザーに向けて先頭 24 ページを閲覧できるサービスを提供している。

(<http://202.96.31.40:9080/magazine/index.htm>)

●オンライン展示とオンライン講座データベース

中国国家図書館には文化の伝承と社会教育という重要な機能がある。社会や大衆に向けた各種の所蔵貴重コレクションの展示や、学術文化シリーズ講座の開催は中国国家図書館の優れた伝統となっている。中国国家図書館ではこれらの展示を電子化し、講座のビデオ映像とともにインターネット上で公開して、それらの優れた文化資源をより多くの人々が享受できるようにしている。現在インターネット上で 308 回の講座、17 本の展示を公開している。

(オンライン講座 <http://vod.nlc.gov.cn/>)

(オンライン展示会

<http://www.nlc.gov.cn/service/jiangzuozhanlan/zaixianzhanlan.htm>)

#### ●インターネット情報資源収集と保存

日々増加するインターネット情報資源の収集と保存は国家図書館の重要な責務である。そこで中国国家図書館では2003年にインターネット情報資源の収集と保存に関する実験的事業を開始した。現在までに収集・保存したインターネット情報は中国各レベルの行政府サイト 2 万サイト、電子新聞 105 タイトル及びオリンピック、中国学等 7 つの主題に関するインターネット情報である。

(<http://webarchive.nlc.gov.cn/index.html>)

### (2) 購入デジタル資源の概況

デジタル技術とコンピューターネットワークの発展に伴って、図書館が収集する資料の範囲はもはや紙媒体の文献だけに限定できなくなっている。利用者の多様化する情報ニーズを満たすため、国家図書館では90年代半ばからデジタル資源の購入を開始し、現在では相当数にのぼっている。それらは以下のとおりである。

- ・電子図書 うち中国語 14 万タイトル 31 万余冊、欧文 33 万 8,000 タイトル 7,000 万ページ

- ・電子ジャーナル うち中国語 1 万 2,000 余タイトル、欧文 7,000 タイトル

- ・電子新聞 うち中国語 1,000 タイトル、外国語 130 タイトル

- ・中国語・外国語学位論文 113 万 6,000 篇

- ・中国語・外国語会議論文 148 万余篇

- ・データベース 88 種 うち中国語データベース 32 種、欧文データベース 56 種で、ほとんどすべての学問分野をカバーしている。

購入契約に基づき、これらの購入デジタル資源はインターネット上で提供してもよいものもあれば、館内 LAN でしか提供できないものもある。

ユーザーがそれらのデータベースを効率よく利用できるようにするため、2006 年 5 月、中国国家デジタル図書館サービスポータルを公開した。そして国家図書館が購入したデジタル資源及び自館で構築しているデータベースを有機的に組織化し国家図書館デジタルリソースポータル (<http://dportal.nlc.gov.cn:8331/portal/findbase.htm>) を公開し、徐々にリソース間をシームレスに連結できるようにしている。ユーザーは利用者カードでログインすると、多くのデータベースの統合検索ができ、検索結果を統一的に表示することができる。

### (3) データストレージと著作権管理

2006 年末までに、国家図書館が保有するデジタル資源の総容量は 120TB に達した。現在主に光ディスク、磁気テープ、磁気ディスクの 3 種類の媒体によってオフラインで保存している。国家デジタル図書館プロジェクトの完成を待って、国家図書館は資源の利用率

によってオンライン、ニアライン、オフラインの三種類の保存形式を採用し、永久に保存する必要がある資源についてはローカルに保存するだけでなく、リモート・ディザスター・リカバリー・バックアップセンターを設け、リモートにも保存する。

国家図書館ではデジタル資源の知的所有権の保護を非常に重視している。しかし著作権者の利益を保護する一方で、他方では公衆が等しく情報にアクセスする権利も保証しなければならない。国家図書館では関連するデジタル資源構築政策を制定した。そこで明確にしたのは、知的所有権の保護期間内の文献については、主に商用データベースの使用許可を購入するという方法を通じてサービスし、自館で構築するデジタル資源には当館の特別コレクションや知的所有権の問題がないか、もしくは少ない文献を主とする、ということである。このほか、国家図書館ではさらに一年中サイト上で作品の使用権利譲渡の募集をかけるなど、いろいろな方法で著作権者や出版社から直接使用権を得ようとしている。

2008年にオリンピックが北京で開催される時に、中国国家図書館新館も利用者サービスを開始する。国家図書館では現在急ピッチでデジタル資源の組織化作業を進め、特色ある所蔵資料の資源ナビゲーションシステムを構築している。そして自館で構築した資源の統一的なナビゲーションを実現し、ユーザーに多くのルート、多くの形式の検索手段を提供する予定である。

今後3年から5年が中国国家デジタル図書館プロジェクト建設のキーとなる時期である。国家図書館はデジタル資源の収集、整理、管理、公開、サービスのソフト／ハードのプラットフォーム構築を足場として、重要なデジタル資源の長期保存を実現し、知識そのもののデジタル資源の組織とサービス中での応用を模索し、それによって優れた文化遺産の伝承をサポートし、公衆の研究、教育、生活、仕事の情報や知識のニーズを満たすデジタル図書館サービスポータルを構築する予定である。