

所蔵文書目録データベースの提供について

山口県文書館 伊藤 一晴

はじめに

所蔵文書目録は文書館・公文書館等にとって利用者と所蔵文書を結ぶ架け橋であり、その作成は最も重要な業務の一つである。従って当館においても管理台帳的な性格を持つものから、文書群を分析し階層構造の復元を試みたもの、さらに利用者に対して多様な検索手段を提供するための主題別・形態別によるものまで、昭和34年（1959）の開館以降、多様な目録を作成してきた。

また一方で全国的には1980年代以降、業務にコンピュータが取り入れられ、その役割も冊子体の目録作成ツールから、所蔵文書の管理やインターネットを活用した目録情報の提供へと、その役割を急速に拡大している。さらにこの目録情報の提供にあたっては、グローバル・ボーダレスというインターネットの特性ゆえに、様々な国際標準の動向にも注視せざるをえない状況にある。

このような中、山口県では平成14年度事業として文書館・博物館・美術館の館蔵品をインターネットの活用により広く発信する「館蔵品情報発信事業」を進めることとなった。本稿では約14ヶ月間という限られた時間の中で、当館が如何に考え、積み上げてきた蓄積をどのように生かし、当事業を行ったのかを紹介する。当館の事例が現在問題となっている様々な時代的要請に込えているとは言えないが、他館の参考になれば幸甚である。

1 事業目的とデータベース構築にあたっての4原則

3館同時に進めることとなった当事業の主目的は、インターネットを活用した館蔵品の情報発信であり、館内業務の電子化（自動化）ではない。よって当館では受入・保存・管理等の既存業務の電子化は行わないまま、既に利用に供している目録を電子化した上、ホームページ上で検索可能とし、利用者が所蔵文書にアクセスするまでの時間及び費用の軽減を図ることにした。

次いで、構築にあたり4つの原則を立てた。

①**現状の整理区分を最大限利用する。**：例えば当館において利用頻度の高い『毛利家文庫目録』は、明治期に毛利家編纂所で行われた主題別分類を基にし、文書一件毎に解説を付した全国的に見ても希な目録である。この文書群を分析し階層構造の復元を試みることは有益であるが、そのためには多くの労力と時間が必要となる。今回は再整理の時間も無いため、既存の目録の長所・欠点を認識した上で、現状の整理区分をそのまま生かすことにした。

②**印刷化された目録を対象とする。**：現在、当館では文書の受け入れに際し、出来る限り速やかに仮目録（手書き又はワープロソフトを使用）を作成し、閲覧提供を開始している。その後、印刷化にあたり時間をかけて再整理を行い、文書群を分析し階層構造の復元を試みる。その際に解説の執筆、標題変更や枝番号の付加等も行う。従って最も情報量が多く、情報に変更が生じにくく、かつ入力データの原本が確固として存在する印刷化された目録を対象とした。

③**印刷化された目録に存在する情報は出来る限り盛り込む。**：既存の目録よりも極端に情報量が減ることを避け、また復元を試みた階層構造の提示や解説等の情報も出来る限り盛り込むことを前提に、データ項目や実際の検索機能等を考えることにした。

④**データの更新及びシステムのバージョンアップが容易であること。**：運用費が嵩むことで既存の業務を圧迫することの無いよう、館内でデータの更新・追加が行えること、また将来必ず起こる新システムへの移行に際し支障が発生しないよう、特殊な技術は使わず、単純な構成とすることにした。

2 データ項目

当館では「藩政文書」「行政文書」「行政資料」「諸家文書」「特設文庫」という5つの区分を最上位に設定し、整理を行っている。従ってこの区分毎に（中には区分内においても）記述項目や様式が異なる。今回のデータ入力（パンチング）にあたっては、検索の利便性を高めるため、項目を全て統一した（表1）。

また文書群の階層構造を表現するため、原則①に則って現状の整理区分を生かし、所蔵文書全体と目録1件（1レコード）の間に「分類1」～「同4」を設定した（表2参照。必ずしも組織の階層構造ではないため、あえて「分類」

表1 データ項目 として)。

フィールド名	データ型	
分類1	文字列	
分類2	文字列	
分類3	文字列	
分類4	文字列	
冊子別ID	数値	
整理番号	文字列	
標題	文字列	
表示用年代	文字列	
作成年代1	元号 1-1	文字列
	年 1-1	数値
	西暦 1-1	数値
	閏 1-1	文字列
	月 1-1	数値
	日 1-1	数値
	元号 1-2	文字列
	年 1-2	数値
	西暦 1-2	数値
	閏 1-2	文字列
月 1-2	数値	
日 1-2	数値	
分量 (形態)	文字列	
分量 (数)	数値	
作成・受取	文字列	
記述レベル	文字列	
作成年代2	元号 2	文字列
	年 2	数値
	西暦 2	数値
	閏 2	文字列
	月 2	数値
日 2	数値	
摘要	文字列	
複製	文字列	
内容	文字列	

※ゴシックのデータが画面上に表示される

また原則③のもと、「表示用年代」を用意し、「三月初秋」「未四月」等のイレギュラーな値も全て表示可能にした。年月日により検索を行う場合は「作成年代1」が検索対象データとなるが、画面上には「表示用年代」しか表示されない。

表2 分類1～4の例



さらに「内容」を用意し、『毛利利家文庫目録』における解題部分等を適宜含めた。なお、作業は外注し、入力に関する細

かい質問・回答はFAXで行った。結果として入力件数は約8万件となった。

3 検索機能の概要

検索機能としては「階層検索」「簡易検索」「詳細検索」という3種を用意した(「詳細検索」については紙幅の都合上、説明を割愛する)。

○階層検索(写真1)：「分類1」～「同4」の説明を上位から下位へ選択しつつ読みすすめ、一覧表示する。「分類」名が画面左に、説明が画面右下に、また包含関係が画面右上にそれぞれ表示される。

○簡易検索：検索窓に文字列を打ち込むと、その文字列が「標題」「作成受取」「内容」の中に含まれているレコードを抽出する。文字列間にスペースを

狭むことにより絞り込みが可能となる。また最上位5区分を取捨選択できるようチェックボックスを用意した。

○結果表示：検索結果一覧画面（写真2）では、「標題」「分類」「整理番号」「年代」を計3行に表示する。「分類1」～「同4」は包含関係が分かり易いように">"で繋いだ。また見やすさを意識し、「標題」のフォントを他の3項目より大きく設定した。この「標題」部分をクリックすると目録詳細画面が開く。同

画面では上記4項目の他に「内容」「記述レベル」「分量」「作成・受取」「摘要」「複製」を表示する。

以上、各画面の基本的なデザインや機能については、担当職員が仕様を固め、委託先である コアと協議・調整し決定した。

おわりに ～今後の課題～

今回の事業によって、所蔵文書点数43万点の内、約8万件のデータがホームページ上で検索可能となった。しかし検索結果から上位の「分類」の説明へと遡ることが出来ないこと、さらには本来優先されるべき全収蔵文書群の概要情報についてはリストを掲載しているにすぎない等、課題も多いことをあえて付け加えておきたい。なお所蔵文書目録データベースの提供は当事業の一部である。全体については南方長との共著「文書館館藏品情報発信事業について」（『山口県文書館研究紀要』第30号、2003）をご覧いただきたい。

写真1 階層検索画面



写真2 検索結果一覧画面

