

階層性を意識した公文書目録の作成

豊見山 和美
Kazumi TOMIYAMA

はじめに

沖縄県公文書館では、5年前の開館当初からコンピューターによる資料管理を導入しており、平成12年度のハードウェアリソース更新にあたって、さらなる性能向上を目的としたシステム開発を実施する機会を得た。資料目録作成・検索機能の面では、より検索精度の高い目録を利用者に提供してサービス向上を実現するためデータベースを再構築する。したがってこの新しいデータベースに、既存のデータを加工して移行する作業が発生するのだが、この作業の核となるのは、階層的な資料記述の導入である。新システムの稼働は2001年2月中旬の予定で、本稿執筆の現時点ではまだ移行作業の最終段階にいるが、編集・出版委員会のご好意により、このようなテーマで紙面をいただくことができたので、沖縄県公文書館が所蔵する琉球政府文書（以下「琉政文書」という。）約16万簿冊の目録データベース再編における基本コンセプトを紹介したい。

1. データベース目録の改良：階層構造のインターフェイス

公文書館が収蔵する資料のうち、圧倒的な比重を占めているのは琉政文書であり、その総簿冊数は約16万冊、書架総延長にして約5kmに及ぶ。この文書群の出处である琉球政府は、1952年（昭和27）4月1日、琉球列島米国民政府布告第13号によって設立され、以後1972年（昭和47）5月15日の

日本復帰による施政権返還に至るまで、立法・司法・行政の三権を備えた沖縄の自治機構として機能した。琉政文書は、1995年（平成7）8月から公文書館で一般に公開されており、米国統治下の沖縄の現代史を研究する上で欠かせない資料群として利用のニーズも大きい。琉政文書はデータベース目録化されており、1998年（平成10）以降はホームページでの目録検索サービスも提供している。

琉政文書は、琉球政府閉庁時の政府機構図をもとにした組織分類によるコードおよび資料タイトルのすべてがコンピューターに入力されている。現在のところ利用者の資料検索は組織名による項目検索や、文字列検索によって行われるが、1簿冊単位で記述された内容情報が資料タイトルのみであるので、検索精度が高いとは言い難い。現状の琉政文書目録は受入台帳というべきレベルのものであり、検索キーを増やして利用者の便宜を図るために、琉政文書目録の記述を充実させることが、公文書館の公文書専門員に与えられた大きな課題であり続けている。データベース目録をさらに使いやすい、つまりより検索精度の高いユーザーフレンドリーなものに改良することが整理計画のゴールである。使いやすい目録とは検索しやすい目録のことであるが、大量の簿冊から成る資料群の中から求める資料を特定するには、どのような検索手法を提供すべきなのだろうか。たとえばインタ

*とみやま かずみ：(財)沖縄県文化振興会公文書管理部公文書専門員（沖縄県公文書館）

ーネット上でのホームページ検索の場合、膨大な数に上るサイトの中から瞬時に必要なものを引っ張ってくるポータル・サイトには、ディレクトリ型検索とロボット型検索の2種類がある。ディレクトリ型は、登録されたサイトがカテゴリー（ジャンル）ごとに分類されており、ユーザーは階層を意識して検索を行うことができる。一方ロボット型は、自動収集プログラムにより機械的にサイトをリストアップするので、より多くの検索結果を出すことができるかもしれないが、検索結果リスト自体はさほど秩序があるものにはならないので、さらにリストを絞り込んでいく必要が生じる。現行システムの検索タイプはほぼ後者で、絞り込みが難しい理由には、まず現在の琉政文書データベース目録に採取されている情報が乏しいこと、また現行データベースシステムの構造上、検索結果自体も無秩序配列で現れることが挙げられる。そもそも、現状では検索の初期画面に、県政文書・琉球政府文書・行政刊行物・USCAR文書・地域資料といったカテゴリーが表示されるだけで、利用者が収蔵資料の全体構造を階層的に把握しながら検索することができない。したがって、新たに構築する目録データベースでは、資料群の全体構造をツリービュー形式で表現したものを提示し、利用者の資料把握の便宜を図ることが重要となる。そのため、資料目録記述の階層化として「資料群構造の階層化」と「資料記述の階層化」を行うことになった。この両者がリンクしたデータベースを構築することによって目録の機能向上を図るため、琉政文書データベース目録作成のコンセプトは「階層構造のインターフェイスの実現」と設定された。

2. 資料群構造の階層化：資料ガイドの作成

資料群構造の階層化とは、利用者が資料のカテゴリーを認識し、階層を下降しながら必要な資料に到達するツールとしての資

料ガイドの作成である。出処の階層に、簿冊の内容を階層的に表現した文書シリーズ（次項参照）を従属させることによって、ツリーを完成する。これにより利用者が、資料群およびその配下のレベルである文書シリーズごとの固まりで資料を一覧あるいは検索できるようにする。また、公文書館収蔵資料全体の中での琉政文書群の位置付けを示すために、資料ガイドは琉政文書以外の所蔵資料も含めて構成する必要がある（これはもちろんシステム上の要請でもある）。図示されているのは資料検索画面上で提供される公文書館資料ガイドの一部で、琉政文書はガイド中0112という資料群コード（ガイドコード）を与えられ、これを解読すると沖縄県公文書館資料／沖縄県資料／文書／琉球政府という資料群インデックスとなる。琉球政府に従属する出処の階層はさらに、ほぼ局別に対応するAからWのアルファベットコードで分類されており、それぞれの下位階層には課レベルでの出処分類がリンクされている。画面上ではツリーの継ぎ目をクリックすると配下の枝が順次現れる。資料ガイド中のツリー構造になっている資料群を特定すると資料群解説が読め、配下の資料が検索対象となる（選択しない場合は最上位のレベル、つまり沖縄県公文書館の全資料が指定されたことになる）。文字列検索を行うと検索条件に合致した資料が文書シリーズ／タイトルごとに表示され、シリーズ配下のすべてのタイトルを見ることもできる。検索結果も階層的に表現されるわけだが、これは各階層がリレーショナル型データベースマネジメントシステムの手法でリンクされている結果である。

3. 資料記述の階層化：文書のシリーズ化

階層化された資料群に、それぞれの文書シリーズが従属して公文書館資料の階層構造の全面的な表現が完成する。この文書シリーズの作成とは、一連の簿冊に概念上の

上位レベルを設定することで、これが資料記述の階層化ということになる。たとえば琉球政府法務局の「確認証及び使用料供託」としてタイトルがデータベースに採録されている簿冊を例にとろう。これは軍用地代の支払いに関するもので、この簿冊にアクセスしようとする利用者のたいていは「軍用地」というキーワードを用いるだろうが、このままではそれでヒットすることはない。レファレンスを受けた職員は、軍用地事務所のコードで分類されている簿冊すべてを一覧として抽出する。こうしてべったりした姿で出てきた数千件のヒットのうちから、利用者が必要と思われる簿冊をさらに絞り込んでいかねばならない。

この場合、より検索精度を高めるためには、その簿冊単体レベルの記述に「軍用地」というキーワードを追加する方法があるが、今回の整理ではそれを採用していない。一群の資料が特定のキーワードを共有するとすれば、階層のより下位のレベル（個々の簿冊の詳細情報の記述レベル）において、コントロールされたキーワードなどの手段により記述内容を標準化するのでなく、むしろそれを個々の簿冊が従属する上位概念としてカテゴリライズし検索キーとする方法を使ったのである。もちろんその上位概念自体も詳細記述と等しく検索対象となる。ディレクトリ型検索を利用者に提供するため、現行の目録を素材として資料群の構成要素を階層的にカテゴリライズすると、この例の場合は「軍用地に関する書類」というシリーズを設定し、この「確認証及び使用料供託」や「借賃前払申請書類」や「収容告知・契約・宣言一覧番号」といった一連の簿冊をその配下に置くことになる。この「文書のシリーズ化」作業が、新システムへのデータ移行のための整理の中心となった。実際の作業は、現行のデータベース目録で採取されている資料タイトルを、シリーズ／タイトル／サブタイトルの3階層に分節することである。シリーズ名付与にあ

たっての基本ツールは各組織の事務分掌表で、また琉政文書自体にいくらか散見される文書事務に関する書類も参考となった。

作業用のワークシートは表1のとおりである。現在のデータベース目録の情報をMS Excelのファイルに取り込み、シリーズ、タイトル、サブタイトルのフィールドを追加した。目録作成ソフトとしてはMS Accessが主流だが、こと定型的な公文書資料の整理については、MS Excelの優位性も強調しておきたい。とくにMS Excelのデータ並べ替えやフィルタといったツール機能や関数を活用すると、入力作業がかなり効率化されるだけでなく、さまざまな角度から資料群を見ることによって組織や文書シリーズの階層性自体がよく見えてくる。筆者が得た文書シリーズ化作業にMS Excelを用いるアイデアも、目録整理などのいわゆるアーキビストとしての本体業務からというより、庶務的業務でMS Excelを利用している経験の中から生まれたといつてよい。

このワークシートは琉政文書のうち法務局（総簿冊数約30,300）の一部のものである。少ないサンプル数ではあるが、原タイトルがどのように分節されたかを示した。ID6から19は「出入域管理に関する書類」という文書シリーズに従属し、出入域記録送付簿というタイトルをもっている。（ちなみに法務局内での文書シリーズ数はほぼ50前後でおさまっている）。この出入域記録送付簿をさらに特定するためにはサブタイトルレベルの記述が必要になるが、原タイトルの情報だけでは出入域記録簿を差異化できないのであらためて原資料を参照し、情報を追加した。また、検索結果がある程度の秩序を持った配列として得られるようにするため、サブタイトルレベル内の数字は最大に合わせて桁数を揃えた。たとえばシート中ID4と5は「軍用地に関する書類」（このシリーズに含まれる簿冊数は約16,000）というシリーズの配下にある

先述した「確認証及び使用料供託」という一連の約1,100冊の一部であり、文書作成者によってアルファベットと数字を組み合わせたファイル番号が付与されているので、これをキーとして、数字部分が最大の3桁に揃うよう修正入力したのである。ID7の簿冊は原タイトルが「議事録綴」となっているが、これでは内容を推定することもできないので、この場合も原資料を参照し、「出入域管理連絡協議会に関する書類」というタイトル情報を入力し、「議事録綴」はサブタイトルレベルの記述に下ろした。作業に入る前はある程度ワークシート上だけで情報を論理的に配列できらうと考えていたのだが、実際は現物を見る必要に迫られることが非常に多かった。このワークシートの情報採取作業には42人月を費やした（それでも単純計算でひと月1万件以上を処理したことになる）。これらの作業により、新システムで提供される検索結果において、データが整理された帳票形式で現れることが可能になる。

4. 今後の作業

琉政文書資料群の目録データベース作成における今後の作業として、目録メタデータの記述（具体的には文書群解説および文書シリーズ解説）の追加がある。文書群解説は、出処分類された資料群単位に、文書の発生源となった組織および文書との連関、すなわち組織体史（Administrative History）であり、一方、文書シリーズ解説は、行政文書になじみの薄い一般の利用者に提供しうる有力なファインディング・エイズである。琉球政府およびその前身機関の組織変遷については、すでに公文書館で資料集2冊を刊行している。この組織図

に組織体史および組織が作成した文書シリーズの記述をビルトインすることができれば、利用者にとっては適切なガイドとなるだろうし、親機関の活動の証拠資料を残し説明責任を担保しようとする公文書館の存在意義を完遂できていると思っている。

以上、階層性を意識した公文書目録作成のコンセプトと新しい検索システムへのデータ移行作業の内容を簡単に述べた。新システムのデータベース構造の概要については、大城博光「公文書目録情報のデータベースモデル～階層構造を持つ目録情報のリレーショナルデータベースでの実装～」(沖縄県公文書館研究紀要第2号 2000年3月 所収)にまとめられているので併せ読まれない。冒頭でふれたが、新システムはアーカス21 (ARCHAS 21: ARCHives Administration System) という愛称で、2001年2月には稼動予定である。ただし、言い訳が許されるなら、本文で述べたような力まかせの突貫作業を続けてきたため、たとえば資料群・シリーズ解説などのメタ情報の搭載など、稼動時に積み残しがあることは残念ながら確実で、稼動後もしばらくはデータベースへの積み込み作業が続くはずだ。公文書館ホームページ「ねっとOPA」で資料目録検索ができるので、内容更新を楽しみながら(?)ブラウズのうえ、本稿へのコメントと合わせてご意見をお寄せいただければ幸いである。また、筆者の「公文書目録データベースにおける階層構造の表現に関する試み」という論文が沖縄県公文書館研究紀要第3号に掲載予定であるので、本稿との重複部分が多いがそちらも参照していただきたい。

沖縄県公文書館ホームページ

<http://www.archives.pref.okinawa.jp>

公文書館収蔵資料ガイド

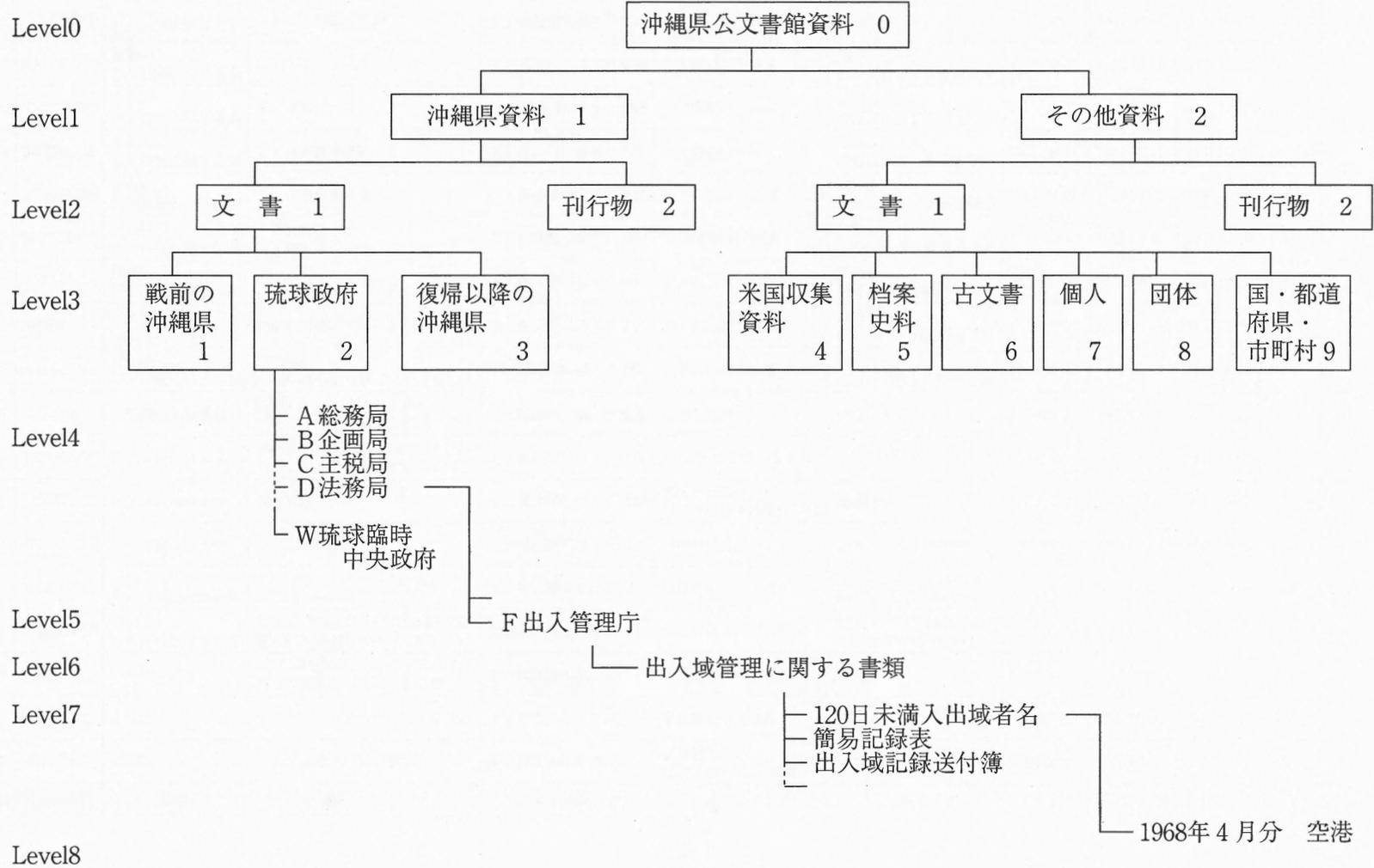


表1

ID	資料コード	文書主管課	原タイトル	シリーズ	タイトル	サブタイトル	作成年月 自	作成年月 至	書架コード	収納箱番号	備考
1	R00021468B	土地課	告知・契約・宣告一覧番号	軍用地に関する書類	収用告知・契約・宣告一覧番号	1960年	1960-06-01	1960-06-01	09-F-05-05	0006	
2	R00024127B	土地課	軍用地の借賃前払に関する書類	軍用地に関する書類	借賃前払申請書類	1963年	1963-06-01	1963-06-01	09-F-07-01	0103	
3	R00024115B	民事部土地課	軍用地借賃前払に関する書類 (決定表送付伺) (66年7月以降)	軍用地に関する書類	借賃前払決定表通知	1965年	1965-10-29	1965-12-16	09-F-07-01	0102	
4	R00050327B	軍用地関係事務所業務課	確認証及び使用料供託 CC99-3 石川市・コザ・読谷 取用委員会裁決	軍用地に関する書類	確認証及び使用料供託	コザ市 石川市 読谷村 CC-099-03 取用委員会裁決	1957-03-08	1957-03-08	09-H-15-05	0409	
5	R00050328B	軍用地関係事務所業務課	確認証及び使用料供託 CC320-1 石川・美里・金武 布告33号 1956年度	軍用地に関する書類	確認証及び使用料供託	石川市 美里村 金武村 CC-320-01 布告33号 1956年度	1956-09-28	1956-09-28	09-H-15-05	0409	
6	R00027175B	出入管理庁出張所	違反事件処理簿	出入域管理に関する書類	違反事件処理簿	1968年 宮古出張所	1968-06-01	1968-06-01	09-H-03-05	0011	
7	R00024957B	出入管理庁審査課	議事録綴	出入域管理に関する書類	出入域管理連絡協議会に関する書類	議事録綴	1965-11-04	1967-10-31	09-F-08-01	0014	
8	R00030552B	出入管理庁審査課	審査に関する書類 120日未満入出域者名簿 (空港) 1966年4月分	出入域管理に関する書類	120日未満入出域者名簿	1966年4月分 空港	1966-04-01	1966-04-30	09-G-14-05	0748	
9	R00029243B	出入管理庁審査課	入域及び出域に関する書類 (簡易記録) 空港 1969年4月	出入域管理に関する書類	簡易記録表	1969年4月	1969-04-01	1969-04-30	09-G-10-03	0447	
10	R00026434B	出入管理庁出入管理事務所	審査に関する書類 (出入域者記録送付簿)	出入域管理に関する書類	出入域記録送付簿	那覇空港出入管理事務所 1968年	1968-06-01	1968-06-01	09-G-16-07	0038	
11	R00026652B	出入管理庁出入管理事務所	出入域記録送付簿 ①	出入域管理に関する書類	出入域記録送付簿	那覇空港出入管理事務所 1970年 1	1970-06-01	1970-06-01	09-H-03-03	0060	
12	R00026653B	出入管理庁出入管理事務所	出入域記録送付簿	出入域管理に関する書類	出入域記録送付簿	那覇空港出入管理事務所 1970年 2	1970-06-01	1970-06-01	09-H-03-03	0060	
13	R00026649B	出入管理庁出入管理事務所	出入域記録送付簿 1月～6月 ①	出入域管理に関する書類	出入域記録送付簿	那覇空港出入管理事務所 1971年 1	1971-06-01	1971-06-01	09-H-03-03	0060	
14	R00026650B	出入管理庁出入管理事務所	出入域記録送付簿 2	出入域管理に関する書類	出入域記録送付簿	那覇空港出入管理事務所 1971年 2	1971-06-01	1971-06-01	09-H-03-03	0060	
15	R00026651B	出入管理庁出入管理事務所	出入域記録送付簿 3	出入域管理に関する書類	出入域記録送付簿	那覇空港出入管理事務所 1971年 3	1971-06-01	1971-06-01	09-H-03-03	0060	
16	R00026646B	出入管理庁出入管理事務所	出入域記録送付簿 1	出入域管理に関する書類	出入域記録送付簿	那覇空港出入管理事務所 1972年 1	1972-06-01	1972-06-01	09-H-03-03	0059	
17	R00026647B	出入管理庁出入管理事務所	出入域記録送付簿 2	出入域管理に関する書類	出入域記録送付簿	那覇空港出入管理事務所 1972年 1	1972-06-01	1972-06-01	09-H-03-03	0059	
18	R00026811B	出入管理庁出張所	出入域者記録送付簿	出入域管理に関する書類	出入域記録送付簿	名護出張所 1971年	1971-06-01	1971-06-01	09-H-03-04	0002	
19	R00030793B	出入管理庁審査課	入域及び出域に関する書類 (非琉球人出域記録) 海港 No7 2月26日～2月28日	出入域管理に関する書類	非琉球人出入域記録	1967年 7	1967-02-26	1967-02-28	09-G-12-07	0630	