

目録データについて

1 目録のあり方についての調査研究の経緯

国立公文書館の所蔵史料は、平成16年度末現在各省庁から移管された公文書等約58万冊と、旧内閣文書所蔵の古典籍等約46万冊の合計約104万冊に及ぶ。これらの資料の検索手段については、簿冊を単位に提供情報の整理＝目録の作成が行なわれ、平成10年度以来順次目録データベースに搭載され、一般の利用に供されてきた。

平成13年度からは、一般になじみの少ない公文書について、利用者の検索等の便の向上を図るため、その在り方に関する調査研究を実施してきた。平成13年度においては、既存の目録の調査・分析を行うとともに、受け入れた公文書は一定期間内（平成13年度から16年度までの第一期中期目標期間内では1年以内）に一般の利用に供することが求められているので、受入れた時点での秩序は崩さずまず公開しようという考え方にたった第一次目録、移管年度毎・移管省庁毎の「簿冊一件名目録」を整備すること、その上で所蔵資料をその文脈に即した形で提供するための検索補助手段として第二次目録「所蔵資料要覧」を作成するとの基本的方針を定めた。それを受けて平成14年度は移管年度毎の目録ではなく、第二次目録、記録の内容（作成組織・作成年月日等）に即した検索補助手段を作成していくための具体的方法の検討を開始し、独立行政法人評価委員会からの「国際標準を視野にいたした目録整備の必要性」の指摘もあり、国際的な目録作成様式とどのように関連づけることができるかについての検討を行った。

現在、記録資料についての国際的な目録作成様式としては、「General International Standard Archival Description」（記録資料についての国際標準記述。以下ISAD（G）と略す。）がある。当館では2002年に公表された第2版にもとづいて検討を行った¹。これは、資料全体をトータルで、いわば上から眺めて、概要から詳細へ階層的に検索できるように構成し、その情報を26の記述要素によって記載していくことを定めている。平成14年度に行った各省庁から移管されている公文書等の実態の分析からは、省庁別の実態から概要を把握することは可能（例えば、保存分類規則（内閣総理府・文部省）・文書作成組織（内閣総理府）・文書の内容（自治省）²であるが、公文書全体を対象として統一的に把握する方法は現時点では見出せないことが判明したが、作成方法にばらつきがあったとしても、移管元省庁ごとの概要を作成すること

¹ 『記録史料記述の国際標準』（北海道大学図書刊行会、2001）

² 詳細は「（事例による紹介）資料群の構成について」（P.28以下）を参照

は、多様な検索補助手段の提供として有意であると思われるので、概要の作成を実施することとした。また、記述要素の内容については、既存簿冊目録の項目との比較検討も行った結果、できるものから記述し、調査に時間がかかるもの等については、「任意記述」あるいは「記述せず」と整理してISAD (G) の26記述要素に対応させながら図るという作成方針をまとめた。なお、実際に利用に供するに当たっては、記述要素をすべて提示すると空欄がでてくる等見にくくなるため、いくつかの記述要素をまとめたものを「項目名」として整理し提示することとした。

「概要」作成の具体的作業は、平成13年1月6日の省庁再編、平成13年度の情報公開法の施行に伴う公文書の移管基準変更等もあり、平成13年度が当館における公文書の受入れについて、ひとつの節目となったので、平成12年度まで受け入れた公文書全体（約50万冊）を対象に、平成15年度に実施された。移管元省庁毎に、受け入れた公文書（簿冊）をその内容別各省庁の業務別等でまとめ「資料群」とし、「資料群」同士の関係を階層的に位置づけた。ISAD (G) の重要な要素である「記述レベル」については、上述のとおり受け入れ資料の内容、及び物理的数量の差も各省庁間で大きいため、各移管元省庁を横断して統一的に階層を確定することは難しく、各省庁単位で移管元と簿冊をつなぐ上下関係を示すものに留めざるを得なかった。なお、平成15年度段階では、各省庁単位の受入れ資料の概要：資料案内を階層的に示すにとどまり、閲覧申込の単位である簿冊の特定までにはいたらなかったが、以上の作業により172の資料群が作成され、検索補助手段として平成15年度末に「資料群案内」として目録データベースに搭載された。

2 デジタルアーカイブ・システムでのEADの適用について

2.1 概要

EAD2002 (Encoded Archival Description 2002・符号化記録史料記述) の内容等については、既に詳細な内容が発表されているので、そちらを参照されたい³。ここでは、当館デジタルアーカイブ・システム（以下本システムという）でのその適用について簡単に紹介することとする。

EADは、ヘッダー・前付け・記述の三部構成となっている。そのうち「記述」部分が、記録資料の情報を記載する部分で、最大限12階層までの階層的な記述が可能である。本システムでは、所蔵資料についてのデータを検索システムとして稼働させるこ

³ EAD公式HP www.loc.gov/ead 「日本におけるEAD検索手段のデータ記載形式：EAD実践ガイドラインをもとにして」（国文学研究資料館アーカイブズ研究系助手 五島敏芳、国文学研究資料館アーカイブズ研究系COE非常勤研究員 戸森麻衣子）公開研究会「日本のアーカイブズの電子的検索手段のために」（20050124）

とを主目的としているので、記述部分の構成に重点をおき、ヘッダー及び前付の部分は、系統的に必要最小限の要素のみの構成とし、具体的データの入力は行っていない。階層については、7階層の構成とした。

また本システムでは、所蔵資料を各省庁から移管された「公文書」と「内閣文庫」に分けている（これをデータ種別という）。また、所蔵資料についてのデータは大きく「資料群データ」と「目録データ」に分けられる（これを目録種別という）。前者は、移管元省庁毎の概要をまとめた「資料群」についてのデータであり、文章で記述されている。後者は従来の目録データベース（RDB）に搭載されていた「簿冊目録」及び「件名（内閣文庫の細目も含む。以下同じ）目録」の2階層のデータであり、項目によりコードにより入力されているものもある。

2.2 資料群及び階層構造

データの階層構造は、第1階層fondsとし、その下に最大4階層までのseriesをおき、第6階層にfile、第7階層にitemをおいた。公文書の場合は、第1階層を移管元省庁、第6階層を簿冊、第7階層を件名、内閣文庫では、第1階層を和書・漢籍・洋書の別とし、第6階層を簿冊、第7階層を細目とすることで両者のレベルを統一した。

移管元省庁毎の公文書の概要を示す「資料群」はすべてseriesとし、省庁により第2階層から第5階層までの展開とした。本システム搭載に当たって、従来の資料群の細分化、及び省庁再編後の平成13年度移管計画以降受け入れた公文書の資料群の作成が行われた。上記のとおり資料群作成の方法が、省庁により異なっていることに加え、検索画面で一覧表示された際の見易さといった視点も加わり、その粗密及び細分化の基準は様々にならざるを得なかった。したがって、第2階層から第5階層までの間については、各省庁間及びデータ種別を超えての横断的統一性はなく、特定の階層を指定（例えば第2階層すべてといった指定）しての横断的検索は出来ない。

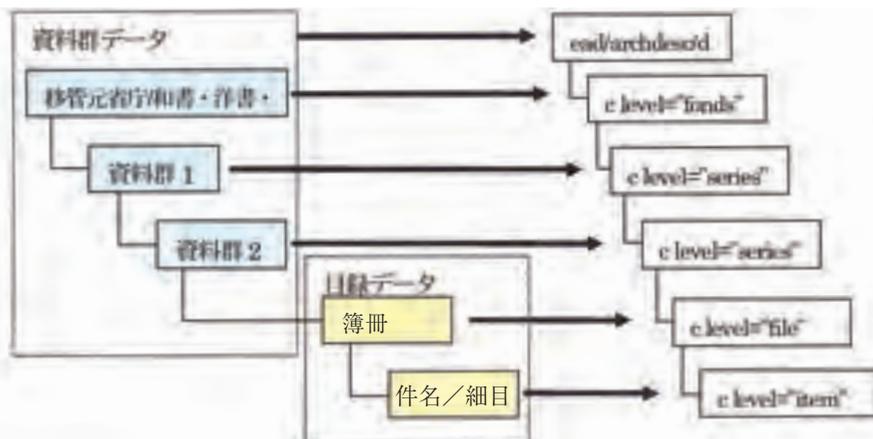


図1 資料群データと目録データの階層構造

2.3 マッピングについて

記述部分のデータの項目とEAD要素のマッピングのうち、資料群データについては、上記のとおり、ISAD (G) の記述要素26項目に準拠した形でのデータの整備が行われていたため、「EAD Tag Library A.1. ISAD (G) ・EAD対照表」を基本に行った⁴ (表1参照)。

目録データの整備については、公文書と内閣文庫所蔵の古典籍という大きく資料の性格の異なるものを同一の規則により電子符号化すること、既に約240万件にも及ぶデータが存在していること、当館における資料の保存・利用の単位が簿冊単位であり、受入れから一定期間内に一般の利用に供するためには、上記の第一次目録である簿冊・件名目録をさきに作成する必要があること等々から、従来の簿冊・件名目録のデータ項目(名)を生かす方法をとった。従って、「上位のレベルで記述された内容は繰り返さない」といったISAD (G) の原則等は反映させず、簿冊・件名各レベルでは、資料群とは別の個別の項目(名)(上位の資料群の解説がなくても、該当する資料を特定するに足る情報の項目)の設定及び記述方法(コード入力 of 継続等)をとることとした。

一方、今後のデータ整備についても視野に入れて、「EADマッピング定義」は、ISAD (G) の項目に合わせる形で章立てし、最後に「画像情報ファイルへのリンク情報」「資料固有情報」等ISAD (G) にない物理的資料の管理情報等をまとめる構成とした。さらに各項目のEAD要素とのマッピングについても、項目(名)が異なっても、データが内容的にまとめられるものは、まとめて統一したEAD要素を採用した。しかし、「項目名」については、資料の特性を具現するものであり、利用者の利便性の観点からも、従来の「名称」を表示できるように、表示を規定するlabel属性等を利用して、検索結果の表示画面には従来の項目名を表示できるようマッピング作業を行った。

表示を規定するlabel属性を利用した一例が、基本的なタイトルの表記である。タイトルは、「資料群」では資料群の「名称」であり、公文書の簿冊では「簿冊標題」、公文書の件名では「件名」、内閣文庫の簿冊や細目では「書名」が使われている。これらの「呼び名」は、資料の特性を具現するものでもあるので、EADの要素は、全データ共通で「unittitle」要素を使用した⁴が、従来の名称をlabel属性として採用し、検索結果画面に表示するようにした。

⁴ 「Encoded Archival Description Tag Library version2002」 <http://www.loc.gov/ead/tglib>

(表1) ISADとEAD対応表

ISAD(G)	ISAD(G)訳語	E A D	EAD訳語案	公文書館項目名
		<eadheader>	eadヘッダー	
		<eadid>	eadid	
		<titleproper>	記述タイトル	
		<author>	記述作成者名	
3.1.1 Reference code(s)	レファレンス・コード	<unitid>	レファレンス・コード	
3.1.2 Title	タイトル	<unittitle>	タイトル	名称
3.1.3 Dates	[年月日]	<unitdate>	年月日	年月日
3.1.4 Level of description	記述レベル	<archdesc> and <c> LEVEL attribute	記述レベル	
3.1.5 Extent and medium of the unit	数量	<physdesc> and subelements <extent>, <origination>and subelements	数量	数量
3.2.1 Name of creator	作成者名称	<persname><corpname>	作成者名称	出所(部局)
3.2.2 Administrative/Biographical history	組織歴または履歴	<bioghist>	組織歴・履歴	出所(部局)
3.2.3 Archival history	伝来	<custodhist>	資料履歴(伝来)	出所(部局)
3.2.4 Immediate source of acquisition	入手先	<acqinfo>	入手先	出所(部局)
3.3.1 Scope and content	資料内容	<scopecontent>	資料内容	内容
3.3.2 Appraisal, destruction and scheduling	評価, 廃棄処分, 保存年限	<appraisal>	評価, 廃棄処分, 保存年限	内容
3.3.3 Accruals	追加受入	<accruals>	追加受入	内容
3.3.4 System of arrangement	編成	<arrangement>	編成	内容
3.4.1 Conditions governing access	公開条件	<accessrestrict>	公開条件	利用情報
3.4.2 Conditions governing reproduction	利用または複写条件	<userrestrict>	利用/複写条件	利用情報
3.4.3 Language/scripts of material	使用言語	<langmaterial>and subelements <language>	資料使用言語	利用情報
3.4.4 Physical characteristics and technical requirements	物的特徴	<phystech>	物的特徴	利用情報
3.4.5 Finding aids	検索手段	<otherfindaid>	検索手段	検索手段
3.5.1 Existence and location of originals	オリジナル資料	<originalsloc>	オリジナル資料	関連資料
3.5.2 Existence and location of copies	複製	<altformavail>	複製	関連資料
3.5.3 Related units of description	関連資料	<relatedmaterial>	関連資料	関連資料
3.5.4 Publication note	出版書誌情報	<bibliography>	出版書誌情報	参考文献
		<note>	ノート	注記
3.7.1 Archivist's note	[アーキビストのノート]	<processinfo>	アーキビストのノート	注記
3.7.2 Rules or conventions	[記述規則]	<processinfo>	記述規則	注記
3.7.3 Date(s) of descriptions	[記述作成年月日]	<processinfo><p><date>	記述作成年月日	注記
		<physloc>	排架場所	請求番号

資料群	<unittitle label="名称">運輸省</unittitle>
公文書簿冊	<unittitle label="簿冊表題">炭鉱労働者雇用安定法（1992） </unittitle>
公文書件名	<unittitle label="件名">炭鉱離職者臨時措置法改正案の概要図 </unittitle>
内閣文庫	<unittitle label="書名">大乘院寺社雑事記</unittitle>

label属性のほかに、採用した要素にさらに下位の要素を加える方法をとった一例に「作成者」がある。「作成者」の「項目名」の表記は、資料群では「出所」、公文書では「作成部局」、内閣文庫では「人名」となっている。そのためまず、EAD要素を「origination」で統一したあと、タイトルと同様に、従来の名称をlabel属性として採用した。ところが内閣文庫の場合は「人名」を、書籍の場合は「著者」「編者」「校訂者」等であったり、古文書の場合は「差出」、「宛名」であったりと、さらに個別の資料の特性により使いわけてきていた。そこで「origination label=人名」の下位要素として、「persname」要素を設定し、「著者」「宛名」といった役割を表示するために「role」属性をつけた。

資料群	<origination label="出所部局">鉄道院，鉄道省，運輸通信省，運輸省鉄道局（国土交通省鉄道局）</origination>
公文書	<origination label="作成部局">労働省職業安定局</origination>
内閣文庫	<origination label="人名"> <persname role="01">尋尊</persname> </origination>

(注：role属性はコード入力となっており、01は著者を示す)

その他、具体的なEADのマッピング内容については「国立公文書館版XML/EAD第1.07版」、EADによるデータの記述例については「国立公文書館のファイル形式・構造目録データ」として本システム上に公開しているのでそれを参照されたい。

なお、文字コードはUTF8を使用しているが、ブラウザ上の表示との兼ね合いから、UTF8のextension A以降については「外字」扱いとし、外字部分は文字図形番号（「今昔文字鏡」）のコードを指定することによって、フォント画像をGIF形式イメージとして配信している。

2.4 「辞書」について

公文書には、利用者が日常使用しない、または馴染みのない用語が多く使用されていることが現在でもよく指摘される。そのため、利用者が想定するキーワードと公文書用語の間の乖離を埋めることを目的とする検索補助手段の必要性がいわれている。

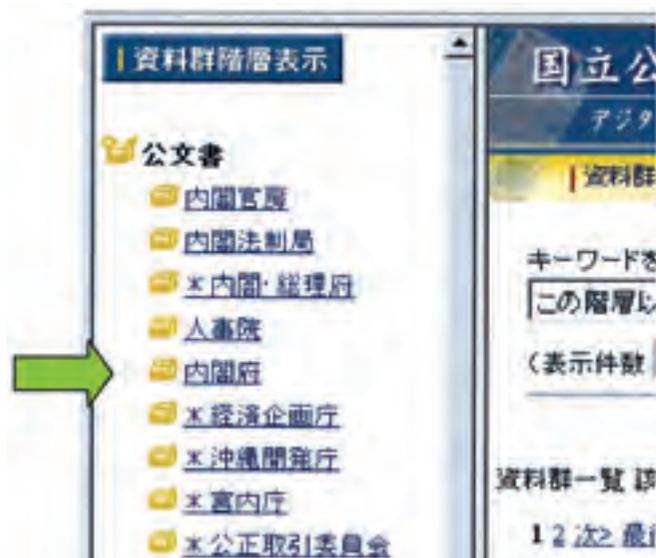
昭和元年から29年までの歴史を跡付ける基本公文書である「御署名原本」「公文類聚（閣議案件）」「枢密院会議録」「公文雑纂（閣議案件以外に内閣の受け付けた公文書）」の約10万件の目録データの用語分析からも、事件・出来事を示す歴史用語はほとんど含まれず、前三者の資料群で法律・条約名や組織名が全体の用語の25%を占めており、「公文雑纂」については人名が60%近くに上っていることが判明した（詳細は本システムに「辞書搭載データ作成方法調査報告書」抜粋を掲載しているのでそちらを参照されたい）。

検索補助手段としての「辞書」については、アジア歴史資料センターでの実績があるので、アジ歴搭載「辞書」の用語を精査し、新たに法律名の略称等を加えた「辞書」機能を本システムに搭載した。

（事例による紹介）資料群の構成について

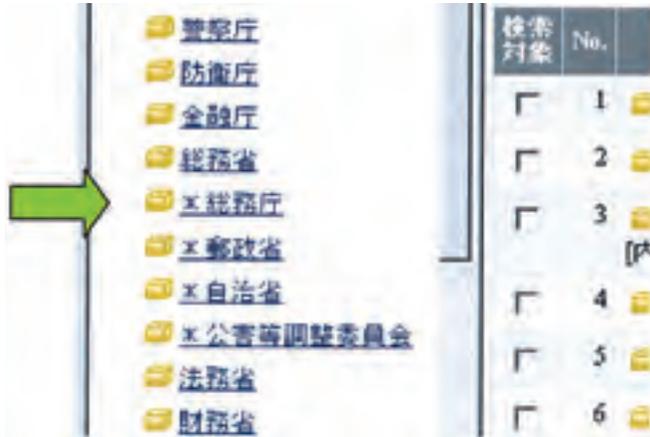
資料群から閲覧申込の単位である簿冊までの階層検索が新たに提供されることになりましたので、ここで資料群の構成の現状（公文書）について紹介します。公文書の階層は、移管府省等別・資料群別に作成しています。まず、公文書の第1階層には、平成13年1月の中央省庁等再編に伴う現府省等名とそれ以前の旧府省等名を混在させて表示していますが、原則として現府省等名順を基本にし、旧府省等名は主な統合先の現府省等名のもとに「※」印を付して並べてあります。また、名称に変更のなかった省庁は新旧を区別していません。

（事例1——新旧府省等名——）



例えば、左の画像の「内閣府」、「総務省」は現在の省名です。それぞれ、その下にある「※」印を付した「経済企画庁」、「沖縄開発庁」、あるいは「総務庁」、「郵政省」、「自治省」が統合された省庁名です。

なお、宮内庁は現在でも存在していますが、現行の移管制度では宮内庁の歴史公文書等は宮内庁書陵部に移管されることとなり、13年度以降国立公文書館へは移管されないため、「※」印としています。

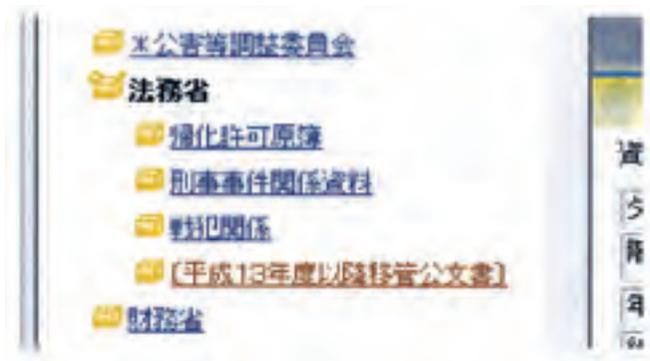


また、公害等調整委員会は、現在も総務省の外局として存在していますが、現行の移管制度では総務省から移管されることになっており、現時点では「~~某~~」を付して表示しています。

これに対して、「人事院」、「警察庁」、「防衛庁」、「法務省」などは、新旧を区別していない例です。

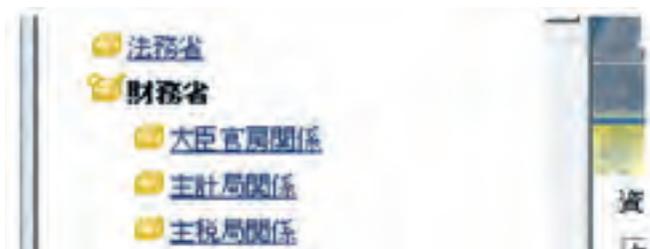
また、平成13年度以降は、国立公文書館法（平成11年法律第79号）第15条に基づく移管基準（「歴史資料として重要な公文書等の適切な保存のために必要な措置について」（平成13年3月30日閣議決定）等）により移管されていることに伴って、移管公文書が多様化しており、〔平成13年度以降移管公文書〕としてまとめる、あるいは、移管元府省等の局別に仮の資料群を作成するなどの方法によって編成しています。

（事例2——平成13年度以降(1)——）

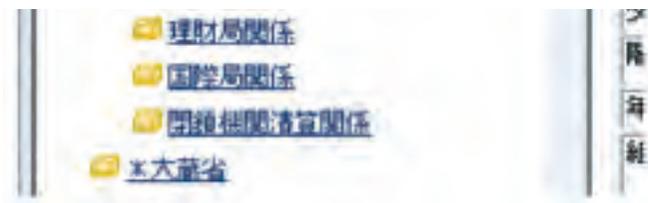


「法務省」は、公文書第1階層で新旧を区別していない例ですが、平成13年度以降移管された公文書は、〔平成13年度以降移管公文書〕としてまとめています。それ以外の「帰化許可原簿」、「刑事事件関係資料」、「戦犯関係」は、平成12年度までに法務省から公文書館に移管された資料群です。

（事例3——平成13年度以降(2)——）



「財務省」は、平成13年1月の中央省庁等再編後の現在の省名であり、その財務省から平成13年度以降に移管された文書は、移管元府省等の局別に仮の資料群を作成しています。例外として、「閉鎖機関清算関係」



は数量が多く、単独の資料群として整理し、局別の資料群とはしていません。

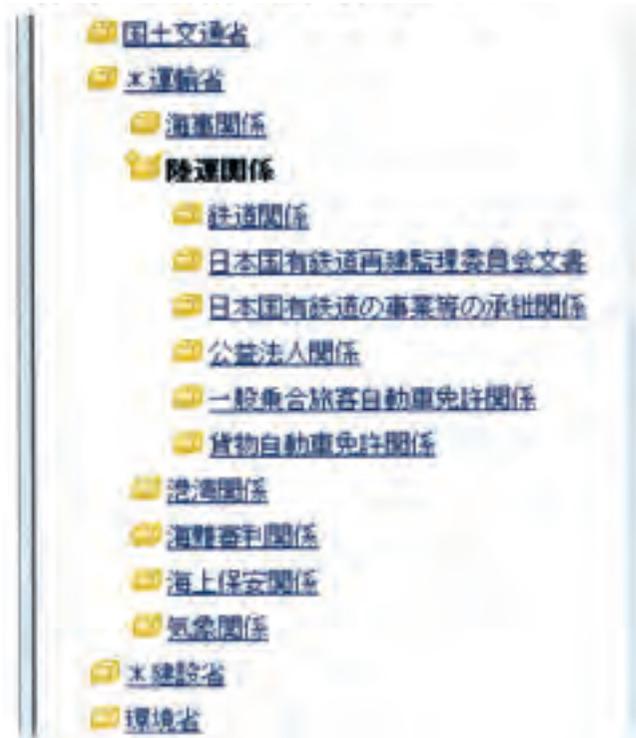
また、平成12年度までに大蔵省から移管された資料群は、「大蔵省」に含まれています。

平成12年度までに移管された公文書の資料群は、各府省から移管された際の送付目録や公文書館で作成した簿冊目録を参考に、原本の一部を確認しながら編成しています。編成の方法は、簿冊の現状が、移管前に各府省等（大臣官房）において集中管理・分類されていたものが移管公文書の大半を占める場合には、その分類を生かすようにしています（例：内閣・総理府、文部省、厚生省、労働省、建設省）。ただし、分類規則そのものが確認できないため、送付目録や原本のラベルから推定しているものもあります（例：厚生省、労働省、建設省）。また、分類規則の第1分類が組織（例えば、局）別になっているものは、組織別で分類されることになり、分類規則のもつ意味はそれほどありませんが、個々の簿冊を資料群に振り分ける際に使った場合があります（例：労働省）。それ以外のものは、文書の内容や作成（取得）組織を参考に、省庁ごとに移管文書の把握がしやすいと考えた方法をとっています。例えば、文書内容にそれほど変化は無いが、組織の機構変遷が激しいものなどは文書の内容を優先し（例：自治省）、特定の組織の特定の文書という場合には、組織を優先して資料群を構成しています（農林水産省も分類規則が存在したようですが、簿冊の現状との関連づけができていないため、現時点では大括りとなっています）。

資料群の名称は、「大内閣・総理府」に含まれる資料群は従来からの名称をほぼ踏襲していますが、その他のものは、内容を把握できるよう留意しながら、新たに与えたものが多くあります。原則として、組織単位のものについては「〇〇文書」とし、内容別なのは「〇〇関係」としています。資料群全体が同種類の文書綴の場合は、「文書」や「関係」という語を付けていません。また、公文書館に移管される前に何らかの資料群と認識されて名称が付されているものは、そのまま踏襲したものがあります（法務省、警察庁、大蔵省（焼残文書））。

（事例4——平成12年度以前(1)——）

平成12年度までに運輸省から移管された文書は、「大運輸省」に含まれます。運輸省の下位階層は、内容別に「海事関係」、「陸運関係」などとし、「陸運関係」の下には、「鉄道関係」、「日本鉄道再建監理委員会文書」などと内容や組織によって資料群を作成しています。内容からみれば、「日本鉄道再建監理委員会文書」は「鉄道関係」



の下位階層に位置づけてもよいものですが、ある程度のまとまりがあるので、鉄道関係と同じ階層に位置づけています。また、名称も組織名を掲げたほうがわかりやすいと考えられるため組織名を付しています。

また、運輸省には「航空関係」という項目があってもよさそうですが、平成12年度までに運輸省から移管された文書には、資料群を構成するような文書が含まれていないため、航空関係という項目はありません。

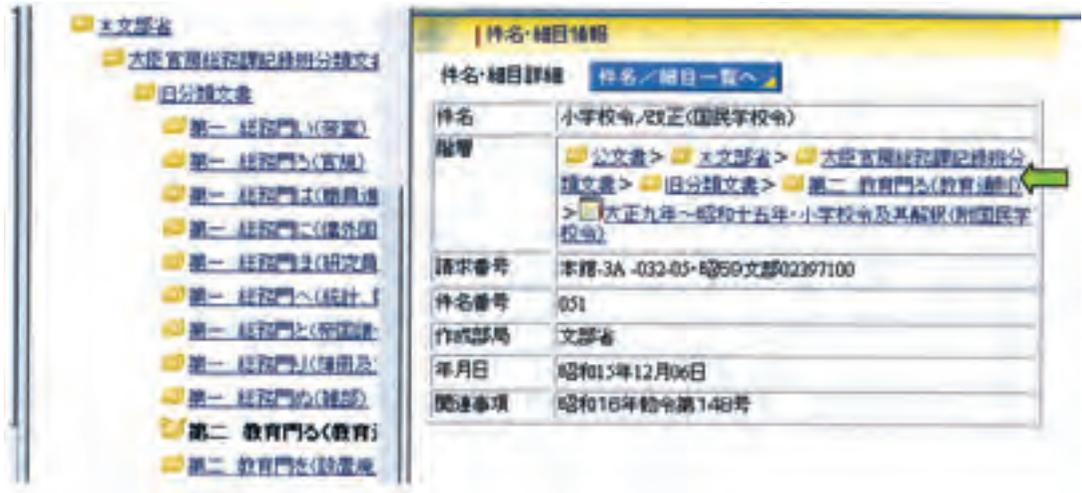
このように階層として表現された資料群を選んでいくことで、閲覧申込の単位である簿冊

にたどり着くようになっていきます。

次に、文書分類規則で資料群を構成し資料群データを作成している「文部省」を事例として取り上げます。平成12年度までに文部省から移管された文書のほとんどは、文部省大臣官房総務課記録班で集中管理されていたものであり、簿冊の現状からおおよそ「文部省文書保存及分類規則」（昭和10年1月7日訓令）によって編綴されたもの（階層名称は「旧分類文書」と「文部省記録文書分類表」（昭和45年2月21日訓令第6号）によって編綴されたもの（「新分類文書」）に分けられます。旧分類は、第一総務門、第二教育門、第三図書門、第四会計門、第五宗教門に分かれ、さらに「い、ろ、は、に、…」の細目に分かれています。ただし、再編綴した形跡があり細かい部分では分類どおりでない場合もあります。

例えば、小学校令改正（昭和16年勅令第148号）に関する文部省公文書は、「第二教育門る（教育通則）」に分類されています。「小学校令」という言葉がわかっている場合には、キーワード検索が可能です。しかし、それがわからない場合には、どの分類項目にありそうか、図書・会計・宗教の各門ではなさそうだから、総務門か教育門かと探してみるうちに、見つかるかもしれません。少なくとも端から順に見るよりは効率的に探すことができます。

(事例5 —— 平成12年度以前(2)分類規則 (件名詳細) ——)



資料群データは、はじめて探す場合よりは、類似の資料を探す際の補助手段としての方が有効かもしれません。小学校令が含まれている「第二 教育門(教育通則)」をクリックすると、この資料群に含まれる簿冊一覧が表示されます。簿冊標題「大正九年～昭和十五年・小学校令及其解釈(附国民学校令)」の前後には、教育通則としてまとめられた国民学校令、青年学校令、実業学校令、高等女学校令、中学校令などに関する簿冊が並んでいます。

(事例6 —— 平成12年度以前(2)分類規則 (簿冊一覧) ——)



以上、資料群の構成の現状について、事例を挙げながら紹介してきました。個々の資料群情報を含め資料群データは未作成のものも多くありますが、徐々に充実していくことを期して、検索手段のひとつとして提供しているものです。また、資料群の構成は、いくつか考えられるまとめ方のうちのひとつを、現時点で採用し、資料群データとして提供しているに過ぎません。そのため、今後検索にヨリ有効な資料群データとするために、資料群の構成・資料群データを変更することもあります（これに対して、簿冊の請求番号は、変更することを想定していません。資料群データが変更されても、一度利用した簿冊の請求番号が変わるということはありませんので御安心ください）。

平成13年度以降移管公文書は、暫定的に資料群を作成していると紹介しましたが、新たな移管公文書の資料群については、①移管公文書の多様化、②既移管公文書、特に平成12年度以前移管公文書（※印をつけた資料群）との関係、などにどのように対応するかということが課題となっています。①に関していえば、今年度の移管基準改正によりさらに多様化が予想されます。②に関しては、現状では、「関連資料」欄に資料群名を記入するという対応をとっていますが、移管年度が異なるだけで第1階層から異なっているというのは、検索上不便であり、新旧どちらかの資料群に統一するなどのルールを決めて、対応していく必要があると考えています（例えば、「天気図」。現状では、キーワード検索の画面で検索対象から「簿冊」と「件名」をはずし、「資料群」のみを検索対象として「天気図」で検索すれば、国土交通省から移管された「天気図」と運輸省から移管された「天気図」があることがわかります）。

最後に、資料群データは、所蔵資料を簿冊の現状から資料群として編成し作成しているため、移管公文書がなければ、当然資料群データはありません。しかし、公文書館に移管されていなくても将来的には移管される可能性がある公文書も含めて資料群データを作成するという方法も考えられます（その際には、資料群データはあるけれども、簿冊情報が存在しない（空の資料群）という場合もでてくることになります）。また、現在の第1階層は、公文書館に移管され所蔵している簿冊全体を例外なく一定の水準でカバーするために、文書の内容や作成（取得）組織ではなく、移管府省等を優先していますが、作成（取得）組織を優先して資料群データとし、移管公文書がある場合にはそこに簿冊情報がぶら下がるような資料群の構成もあるかもしれません。所蔵資料と空の資料群とのバランス、簿冊情報の精度、検索場面での有効性等を勘案して考えられていくものだと思います。