

和図書遷及入力作業の終了について

図書部書誌課

はじめに

国立国会図書館所蔵の大正時代に刊行された和図書について、書誌データの遷及入力が終了した。現在、冊子体の蔵書目録が刊行中であり、また、磁気テープ版およびCD-ROM版を順次刊行していく予定である。

今回の大正期和図書データの遷及入力終了により、およそ二〇年にわたる和図書の遷及入力作業が完了し、当館が所蔵する明治時代以後現在に至る和図書の書誌データについて、すべてデータベースでの検索が可能になったわけである。

この機会に、当館の和図書遷及入力について、今回終了した大正期を中心に、その歩みを概観したい。

和図書データ遷及入力計画

当館では、昭和五二年から機械編纂システムによる和図書の入力が始まり、新規収集の和図書については逐次その

書誌情報を入力し、そのデータをもとに『日本全国書誌』を編纂・刊行、併せて、機械可読の磁気テープ版「JAP AN/MARC」を頒布してきた。

一方、入力開始以前の書誌データの入力、すなわち、遷及入力は、冊子体の『国立国会図書館蔵書目録 昭和四四年―五一年』（第三期）の機械編纂を目的として、昭和五四年に入力作業に着手したことに始まる（平成元年三月入力終了）。

昭和六〇年には、あらためて和図書データの遷及入力計画を策定した。冊子目録の機械編纂という目的はもろろんだが、そのことよりも、むしろ、図書館の機械化が進展する中で、オンライン閲覧目録や地域的・全国的総合目録を構築していくための基盤として、遷及入力の必要性が増してきたことをふまえたからである。

この新たな遷及入力計画は、当館が所蔵する和図書のうち、明治期刊行のものから昭和四三年受入までのものを、時代を四つに、すなわち、明治期、大正期、昭和前期（昭

和図書遊及入力作業の成果

(平成11年3月1日現在)

時 期	入力原稿とした目録	NOREN 件数 (オンライン・データベース)	蔵書目録 (遊及入力したデータを機械編纂した冊子体目録)	JAPAN/MARC (磁気テープ)	CD-ROM
明 治 期	『国立国会図書館所蔵明治期刊行図書目録』	109,800	『国立国会図書館蔵書目録明治期』	遊及版明治期	NDL CD-ROM Line 明治期
大 正 期	甲部:『帝国図書館和漢図書書名目録』第四編 乙部:カード目録	79,123	『国立国会図書館蔵書目録大正期』刊行中	平成11年度刊行予定	平成11年度刊行予定
昭和前期	甲部:『帝国図書館和漢図書書名目録』第五、六、七編 『帝国図書館国立図書館和漢書分類目録』 乙部:(原資料) 発禁:『国立国会図書館所蔵発禁図書目録—1945年以前』	188,338	『国立国会図書館蔵書目録昭和元年—24年3月』	遊及版昭和元年—24年3月	NDL CD-ROM Line 昭和元年—24年3月
第一・二期	カード目録	284,244	『国立国会図書館蔵書目録 昭和23年—43年』	遊及版('48—'55) 遊及版('56—'68)	J-BISC 遊及版 ('48—'68)
第 三 期	カード目録	196,880	『国立国会図書館蔵書目録 昭和44年—51年』	遊及版('69—'76) 9編に分割	J-BISC 遊及版 ('69—'83)

和(二・四四年)、第一・二期(昭和二二—四三年)に区分して書誌データを入力するというもので、昭和六三年から実行に移った。

計画全体にかかる入力方針は、次のとおりである。

① 入力に際しては、基本的に、各時期に既に作成されているカード目録や冊子目録を原稿とする(つまり、原資料に個別にあたりデータを作成する方式は採らない)。

② それら既存の目録は時代によって内容・形式が異なるが、その統一は行わない。

既存の目録を利用するというこの方針は、作業の大幅な省力化を目的としたものであった。

第一・二期

― 自動カナフリシステムとアクセスポイントの拡充 ―

第一・二期(昭和二三年—四三年)の入力作業は、それ以前に入力した第三期(昭和四四年—五一年)と同様に、カード目録を入力原稿とした。よって、基本的には、カードに記入されている各書誌事項にフィールドタグをつけて入力すればよく、また、アクセスポイントについても同様である。

しかしながら、そう簡単にはいかなかった。第一・二期の場合、たとえば、原稿としたカードのトレーシングに書名のヨミが欠けているという問題があった。そこで導入し

たのが、日本語の漢字かな混じり文を分かち書きし、フリガナ付けとキーワード抽出を自動的に処理する、自動カナフリ・自動キーワード抽出システムである。このシステムによって書名を読ませ、不正確な部分については、職員が手を加えるという方法を採用した。さらに、この方法を他の書誌事項（副書名、出版者、叢書名）にも活用することで、従来アクセスポイントとしていなかった部分にも、その拡充をはかった。この方法は、以後、他の時期の入力に際しても適用されていくことになる。

昭和六三年一〇月に入力を開始した第一・二期は、平成五年四月に終了した。これにより、まず、当館設立（昭和一三年）以後に収集・整理した和図書について、すべてデータベース化されたこととなった。

明治期—ヨミの困難とマイクロファイッシュ番号—

次に着手したのが、明治期である。当初の計画は、文字どおり時代を順に遡る仕方で行う予定ではあった。ところが、明治期刊行図書のマイクロ化の事業が実施されることとなり、そこで、資料のマイクロ化に、さらに書誌情報のデータベース化を併せることで、明治文化の宝庫というべきこの時代の図書の広汎な利用を可能にすることを旨とし、明治期の入力作業を始めた。

入力原稿としたのは冊子目録であるが、その記入は、

記述の部分のみで、書名・著者名のヨミや分類記号は新たに付与した。

明治期作業のなかでもっとも困難だったのは、著者名のヨミであった。当館の著者名典拠録はもちろん、各種参考図書や原資料にあたりできるだけ調査したが、古い資料ということもあり、ヨミを推定したのも多々ある。調査の結果は、カードに記録して典拠コントロールを行った。この著者名の典拠コントロールに関しては、後述するように、昭和前期入力作業時に、新たなシステムを導入することになる。

また、明治期データには、マイクロファイッシュに付与された固有の番号も入力した。原資料は保存し利用についてはマイクロファイッシュで、という方向への一助となった。

明治期入力作業は、平成元年五月より開始、平成五年一月に終了した。

昭和前期

多様な資料群と書庫作業と著者名典拠ファイルシステム—

昭和前期は、その対象となった資料群が多様性に富んでいる。入力対象としたのは、甲部図書、乙部図書、さらに発禁図書である。

甲部・乙部、というのは、当時の収集方針による区別である。明治八年の納本制度の制定にはじまる戦前の資料を、

帝国図書館では、甲、乙、丙の三部に分けて収蔵していた。このうち、甲、乙のみを保存（丙については一年間保存のうえ廃棄）、さらに、乙については、一応は保存するが利用の価値については後日の判断を待つものとされ、ほぼ未整理の状態ですべての片隅に置かれていた。

よって、乙部の整理作業は、資料を段ボール箱より出すことから始まり、半世紀分の汚れを除去した後、原資料をもとに入力データ票を作成し（この準備作業は平成二年春から始め平成三年八月に終了）、入力を行った。政治的意見書や私家版詩文集などの小冊子を多く含む、七万冊に近い数のこの乙部図書は整理によって、戦前の出版物の歴史的空白が埋められたといってもよからう。

甲部および発禁図書については、それぞれ既存の冊子目録を入力原稿とした。ところが、この時代の甲部図書の冊子目録は、年代ごとに四つの編に別れており、さらに、古い年代の目録になるほど、書誌としての必須の項目である出版地、出版者、出版年、ページ数、大きさ等の欠落が多くなっている。この欠落を補充するためには、やはり、現物にあたるしかなく、書庫内での現物確認作業を大々的に行った。かなりの労力を割いた部分である。

入力作業上の画期的な試みとしては、機械化された著者名典拠ファイルシステムの導入による著者標目の典拠コントロールがあげられる。

入力データの著者名の漢字形と著者名典拠ファイルとを機械で突き合わせ、一致したもののリストを出す。このリストをもとに同定・識別を行い、ヨミ・形を決定する。また、典拠ファイルと一致しなかった著者、すなわち、初出の著者についても、原資料のルビや人名辞典等で調査してヨミ・形を決定し、順次典拠ファイルへ登録していく。つまり、昭和前期の作業以前にすでに典拠ファイルに登録されていた著者名のみならず、この作業の過程で初めて出くわす著者名についても、効率よく典拠コントロールができるわけである。平成五年六月より開始した昭和前期の入力は、平成九年三月に終了した。

ところで、先に述べたように、入力原稿とする既存の目録は時代によって内容・形式が異なるが、その統一は行わない、という方針ではあった。しかしながら、多様な資料群を、多様な冊子目録に基づいて（乙部については原資料をもとに）入力した昭和前期は、そのデータをまとめるにあたって、データ間の調整を行わざるをえなかった。主な調整作業は、重複データの削除と書誌階層の統一である。入力した書誌データを、単に当館の蔵書目録としてのみならず、この時代の『全国書誌』という観点でもとらえるなら、必要とされる調整であった。

こうして、書誌データの調整を経て、昭和前期の作業は平成九年一月に終了した。

大正期 ― やはり書庫作業と入力作業の集大成 ―

このたび終了した大正期は、八万件弱というデータ総件数こそ他の時期に比べて少ないが、その入力作業は困難を来した。

入力対象としたのは、この時期の甲部図書と乙部図書である。乙部については、手書きのカード目録を入力原稿とした。この手書きのカードは、昭和五年に冊子体の「国立国会図書館所蔵 明治期刊行図書目録」刊行後、それまで未整理だった大正期乙部図書についても、整理し閲覧に供すべきであるとの判断によって作成されていたものである。この時点では、書名順冊子目録の作成が計画されていたが、諸般の事情から刊行は実現せず、このたび新たな形で陽の目を見ることになった。

甲部については、昭和前期同様、冊子目録を入力原稿とした。しかしながら、この時期の目録は、他の時期のそれに比べて、必須の書誌事項の欠落が多い。遡及入力作業は、入力のもとになる原稿の完成度が高いほど容易に進むが、その意味では、この大正期入力作業は最難関だったといえるよう。

書誌事項の欠落を補充するために、甲部図書約五万五、〇〇〇件のすべてについて、書庫内での現物確認作業を行った。作業を行うには必ずしも環境が良いとはいえない書庫

内での、根気の要する作業であった。なお、この書庫作業時に発見した、既存の冊子目録未収録分約一、六〇〇件についても、今回の作業で新規に入力した。

また一方で、遡及入力作業の最後にあたる大正期では、これまでに導入してきたシステムや方法をフルに活用することができた。著者名については、昭和前期作業時に導入した前述の著者名典拠ファイルシステムによって典拠コントロールを行い、また、その他のアクセスポイントについては、第一・二期作業時より使ってきた自動カナフリシステムを活用した。さらに、NDC分類記号の付与については、これも昭和前期作業時に導入した、分類記号自動付与システムを活用した。これは、書名中の主題を表すことばを分類記号に変換することで、NDCを機械的に付与するシステムである。

こうした機械処理の後、職員による確認・修正が必要なのはいうまでもないが、しかし、これらのシステムの活用により、正確で効率の良い入力作業を行うことができた。

平成七年六月から開始した大正期の入力は、平成一〇年七月に終了した。入力終了後は、大正期データとしてのまとまりを考慮し、昭和前期と同様、重複データの削除および書誌階層の統一など、入力データの調整を行ったが、多様性を有した昭和前期ほどのデータ間のばらつきはなく、平成一〇年一〇月に作業を終えた。

既入力データの整備と未入力資料の追加入力

大正期入力作業と並行して、すでに終了していた明治期について、それを補充する作業を行った。ひとつは著者名典拠の整備であり、もうひとつは未入力資料の新規入力である。

前述のように、著者名典拠ファイルシステムによる典拠コントロール作業は、明治期入力作業には導入できなかった。今回、このシステムによって、日本人著者については新たに手直しし、また、明治期データにおいて初出の著者は、典拠ファイルへ登録、生没年・別名称等の参照事項も追記入力した。一方、明治期作業時に入力しなかった資料約九〇〇件（主として、地図・版画などの一枚ものや講義録）について、原資料よりデータ票を作成して入力した。

こうした既入力データの整備および未入力資料の追加入力については、確かに和図書週及入力はひとまず終わったとはいえず、今後も積極的に取り組んでいくべき作業である。

既入力データの整備に関しては、たとえば、第三期および第一・二期のデータについて、入力作業時には開発されていなかった著者名典拠ファイルシステムによる、典拠コントロール作業を行う必要がある。この時期のデータの著者名は、基本的には典拠ファイルに登録されているが、この典拠形と当課ですでに入力したデータの著者標目の形

とが一致していないものが多々ある。これを一致させておくことによって、書誌データと典拠ファイルとのリンクが可能になり、データ訂正等の作業に（将来的には利用者の検索にも）、便利なものとなる。

未入力資料に関しては、書庫内で新たに大正期および昭和前期に該当する図書が見つかっている（これらのうち、すでに入力したものと重複がだけだけあるのかは、調査を必要とする）。これらを貴重なものとみるかみないかは異論があろうが、乙部図書などと同様に、当時の出版物の歴史を示す資料には変わりなかるう。あるいはまた、明治期を除いて、国内刊行の洋書は網羅されていない。これまで週及入力してきたデータを、『全国書誌』として、また、全国的な総合目録の基盤としてとらえるなら、こうした未入力分についても追加入力していかなければならない。

なお、特に明治期から昭和前期までの書誌データについては、入力対象が古い時代の図書であり、また、入力終了期限が定められるなど、時間との兼ね合いもあって、書名・著者名のヨミなどを推定せざるを得なかったものがある。あるいはまた、意識せずとも数々の誤りをデータ中に忍び込ませてきているであろう。こうした点について、すでに館内外からいくつもの指摘を受けており、その都度データの訂正を行ってきた。入力データをより正確なものにしていくためにも、今後も各方面からのご教示をお願いする。

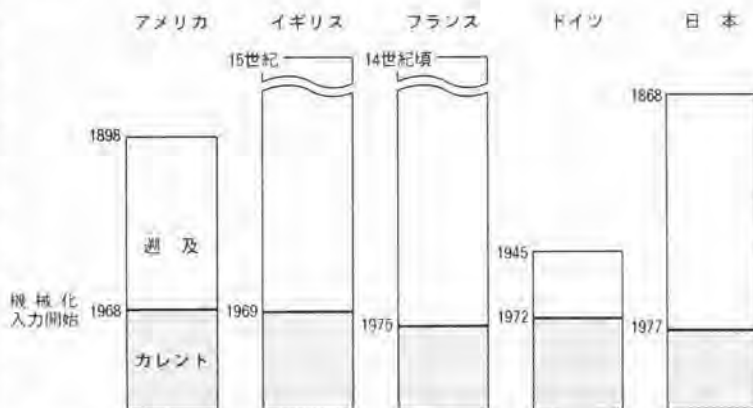
おわりに

今年度より当館で事業を開始した、都道府県立および政令指定都市立図書館を結ぶ「国立国会図書館総合目録ネットワーク」のデータベースに、今年度中に明治期と昭和前期のデータが、来年度早々には大正期のデータが順次投入される予定である。この時期の書誌データを入力している図書館は多くないため、当館への期待が大きい部分である。図書部書誌課は、この総合目録の書誌データの内容にかかる連絡・調整等の業務を担当している。

当館は、国際子ども図書館を平成一二年度に、国立国会図書館関西館（仮称）を平成一四年度に、それぞれ

主要国国立図書館の国内刊行図書遡及入力概要

いち早く図書館の機械化が始まった欧米では、1970年前後に書誌データ作成の機械化が始まった。機械化以前の冊子目録・カード目録などの入力、すなわち遡及入力も相次いで行われ、特に70年代後半より大規模な遡及入力プロジェクトが開始された。以下は、主要国国立図書館の遡及入力の概況である。ここでは、機械化後の書誌データ作成を「カレント」、機械化以前の書誌データの入力を「遡及」として示す。データの入力方法・提供媒体等、国によって様々であるが、簡略化した。



【参考資料】

各国国立図書館のホームページ

国立国会図書館専門資料部参考課，世界各国の全国書誌：主要国を中心に一改訂増補版，1995.

IFLA UBCIM Programme. *International guide to MARC databases and services: national magnetic tape, online and CD-ROM services*. 3rd. edn. ed. 1993.

れ開館を予定している。書誌課では、現在、児童書の選及入力について、データ調整の最後の大詰めを迎えている。入力されたデータは、児童書総合目録データベースに投入される。また、これと並行して、関西館に配置される和図書の組織化作業に取り組んでいる。

今後、東京本館、関西館、および国際子ども図書館の三館に所蔵資料と機能が分散する中で、それらが一体となつて図書館業務を円滑に遂行していくためには、当館は、その所蔵する図書館資料について、書誌データを入力し、データベースでの検索を可能にしておかなければならない。また、当館が構想している、「図書館が通信ネットワークを介して行う一次情報および二次情報の電子的提供とそのための基盤」である「電子図書館」の、特に二次情報の提供に関して、書誌データベースは核となる。

若干の未入力資料が存在するとはいえ、今回の和図書選及入力作業の終了により、当館書誌データベースの主要な柱のひとつである和図書誌データの入力に関しては、大きな山を越えたといえよう。

時代を「選及」して行ってきた入力作業は、当館の「今後」に確実に結びついている。和図書選及入力作業の終了は、新しい始まりへのひとつの区切りである。

△参考文献▽

○和図書データ選及入力計画および第一・二期について

石川史士「和図書データの選及入力計画」(本誌三三四号 一九八九・一)

○明治期について

柿崎幸子「明治期刊行図書選及入力この一年―特に書名・著者名について―」(本誌三五三号 一九九〇・八)

岡本英夫「明治期刊行図書の選及入力を終えて」(本誌三九八号 一九九四・五)

○昭和前期について

瀬川弘悦・菅原篤子「昭和前期和図書の選及入力開始にあたって」(本誌三九一号 一九九三・一〇)

図書部書誌課「昭和前期和図書選及入力の終了と『国立国会図書館蔵書目録 昭和元年―二十四年三月』の刊行について」(本誌四四五号 一九九八・四)

○大正期について

図書部書誌課「和図書選及入力計画最終段階へ―大正期分の入力作業に着手―」(本誌四一四号 一九九五・九)

(図書部書誌課 文責…大柴忠彦)