

早慶目録共同運用の実現まで

このえっこ
河野江津子

(メディアセンター本部課長)

1 はじめに

慶應義塾大学メディアセンター（以下「慶應」とする）と早稲田大学図書館（以下「早稲田」とする）の図書館システム共同運用において、その柱となるのが目録データの共同作成とデータベースの共有化である。予算や個人情報など秘匿性の高いデータもあるなかで、目録はOCLC WorldCat（以下「OCLC」とする）やNACSIS-CAT（以下「CAT」とする）など図書館業界でシェアする仕組みや意識も発達しており、共同運用のメリットをもっとも活かせるデータと言える。

本稿では、早慶図書館システム共同運用プロジェクトで目録チーム¹⁾が目指したこと、実現に向けて行ったこと、移行後の課題について紹介する。

2 目録共同運用で目指したこと

第1の目標は、早慶で重複する書誌を共有することによる目録作業の効率化、目録作成の共同運用によるコスト削減である。実際に書誌データの重複率を2015年度の新規作成成分から算出したところ、ISBNでマッチングできたものだけでも慶應で40%、早稲田で55%、和書に限ると慶應で54%、早稲田で67%と高い割合となることが裏付けられた。

第2の目標は目録の標準化と国際流通である。早慶ではデータフォーマットとして国際的に広く利用されているMARC21を採用しているが、日本語資料を扱うための独自仕様を定めている。これを改め、より標準的な形式に変更することで、外部サービスとのデータ交換や登録をしやすくし、データの視認性を高める狙いがある。

第3の目標は目録データ検索の利便性向上である。データベースを統合し、利用者が早慶の目録をシームレスに（もちろん別々にも）検索できるようにすることで、機関を越えたりクエストやシェアード・プリントなど、将来への展望も広がることになる。

3 目録共同運用体制

(1) 早慶目録ユニットの枠組

目録作成を共同で行うにあたり、その実働組織を

どのように編成するかが検討の出発点となった。早慶ともに業務委託を導入しているのは共通だったが、慶應はオリジナルや逐次刊行物の書誌作成などある程度スキルの必要な作業は職員が引き受け、目録知識を深める方針を取っている。一方早稲田は目録業務を業者へ完全委託し、職員はその成果の管理をするのみという体制の違いがあった。

効率化のためには、委託先を早慶で統一し、作業場所を1か所にまとめる必要がある。新システムでは早慶共通のルールやフォーマットを定め、マニュアルの刷新、初期のスタッフトレーニングも行うことになる。最終的に、これらを主導する慶應の職員が関わりやすい組織運営をすることとなり、「早慶目録ユニット」（以下「ユニット」とする）と名付けた組織の慶應への設置が決定した。当初は受入業務も1か所で行う案もあったが、取引書店との関係などから納品は各大学で受け、早稲田分は受入処理が済んでから、ユニットの作業対象のみを早慶便²⁾で慶應へ送付するフローとした。

職員が担う役割についての早慶の方針は維持し、ユニットでのコピーカタログ作業は業者へ委託し、オリジナル目録の作成や逐次刊行物の処理は早稲田分も含めて慶應の職員が担うことになった。早稲田の資料のうち慶應の職員が整理する分は、業者から慶應へ再委託されるという形を取っている。

委託に適さない資料群は、早慶が自館で処理するよう対象から除外した。具体的には書誌をシェアしない貴重書、和古書・漢籍、学位論文、新システム稼働以前に受け入れた遡及資料などである。

委託契約は早慶それぞれが業者と締結し、什器、作業端末、ソフトウェア提供などの必要経費は慶應、早慶便のコンテナの用意および搬送費用は早稲田が負担するという分担になっている。

(2) システム選定

目録共同運用で目指す目標を実現し、ユニット体制で利用するために目録チームで重視した主なシステム要件は、以下のようなものである。

- ・早慶2機関の目録業務を同一のラインで行うことができるが、目録に付随する発注・支払や予算情報など早慶が共有しないデータは別に管理できる
- ・早慶の既存書誌データを移行時に統合できる
- ・EditorがMARC21フォーマットに対応している
- ・OCLCと連携している
- ・日本語のローマ字生成機能の組み込みができる

他チームの要件とも合わせて選定した結果、2017年11月に、上記を満たすEx Libris社のAlmaを採用することが決定し、稼働開始を2019年9月と定めて準備を進めることになった。

(3) 委託業者選定

ユニットの業務を委託する業者は入札形式で選定した。仕様書・契約書案を作成して募集を行ったのが2018年8月で、応札企業への説明会や技術審査を経て最終的に2018年10月に株式会社キャリアパワーを選定した。

同社は早慶にとって従来の委託先とは異なる業者であり、処理能力は未知数であった。早慶の目録作成ルールは国内では特殊で経験者も少なく、スタッフが総入れ替えとなるリスクも大きい。それでも、過去のやり方を引きずらず、新しいシステムで新しい体制と一緒に築いていくという期待を込め、将来性に賭けて決定した。

システム切り替えおよび共同運用体制への移行は年度途中となるため、慶應が先行して2019年4月に委託先を切り替えた。新スタッフへはまず旧システムで業務トレーニングを行い、7月後半から8月に設けたシステム停止期間に改めて新体制に向けたトレーニングを行った。

Almaは2019年9月に無事稼働し、同時に「早慶目録ユニット」が正式に発足して業務を開始した。

4 目録の標準化

(1) データフォーマット

先述のように、早慶ではもともと国際標準として広く普及しているMARC21をベースに書誌を作成してきた。国内の大学図書館の多くがCAT対応のシステムを使用しているなか独自路線を走ってきたのだが、これが早慶協働を可能にした大きな要因でもある。

慶應では3つ前のシステムまではCATに準拠していたが、週及事業や外部サービスとの書誌調達・提供などでデータ交換をする際、フォーマット変換が困難だったこと、CATの多階層書誌によってデータベースが肥大化し、システムの処理機能を圧迫したことなどを理由にCAT対応をやめる決断をした。処理速度を上げるために多階層書誌の畳み込みを実施し、シンプルなデータ構造へ変換したこともあり、Almaの2つ前のシステム移行時に階層なしで記述できるMARC21への変更へ踏み切った。

国際標準に寄せたことで、海外の各種サービスとのデータ調達、登録、交換が以前より容易になった。また、海外ベンダーから提供される電子資料のknowledge baseのデータもMARC21に準拠しており、冊子体と電子の両方を扱うAlmaではそのフォーマットが揃っていることも有利に働いている。

逆に、国内MARCはそのまま利用することができないため、CATではMARC21へのマッピング・データを使用し、TRC MARCはMARC21へ変換する外付けプログラムを開発してAlmaと連携させる措置を取っている。

ただし、MARC21とはいえ、旧システムでは早慶ともに日本語データを扱うための独自の形を採用していた。本来はレギュラーのタグフィールドとAlternative Graphic Representationを表す880をセットで使用するのが基本のところ、慶應では同じフィールドでレギュラー形（日本語形）、漢字分かち形、カナ読み形、ローマ字読み形という4点セット、早稲田は漢字分かち形のない3点セットのデータを作成していた。移行にあたっては、なるべくAlmaの仕様を活かし、国際標準に準拠することを前提としたため、漢字分かち、カナ読みを捨てることにした。また、早慶でローマ字化のルールが異なる部分があり（長音記号の使用の有無など）、ALA-LC Romanization Tablesや国立国会図書館（以下「NDL」とする）の形も参照し、慶應が早稲田に合わせて長音記号を採用することになった。これは慶應にとって大きな仕様変更であり、既存データの一括修正、日本語からローマ字を自動生成するツールであるHappinessの追加開発も伴うことになった。

典拠データベースは慶應のものを移行し、早慶で共有することになった。書誌と同様に日本人は漢字形とローマ字形のみを残す形とし、旧システムでは

典拠形アクセスポイントにもカナを含めていたが、その部分を削除することになった。

(2) 目録規則

目録規則は、慶應ではシステム移行に先んじて2017年4月から洋書にRDA (Resource Description and Access) を適用していた。英米目録規則第2版 (以下「AACR2」とする) に代わるものとして、既にデータ調達元である米国議会図書館やOCLCが採用していたことによる。

一方、和書についてはAACR2形を維持していたが、RDAに準拠した日本目録規則の2018年版 (以下「NCR2018」とする) が2018年12月に刊行され、新規則での記録方法がイメージできるようになったこともあり、委託業者切り替えのタイミングで和書・洋書にRDAを適用することになった。日本独自の出版慣行に影響される箇所、例えば情報源規定や出版関連事項などではNCR2018を採用している部分もあるが、全体としては和書についてもRDA準拠、NCR2018は参照用という扱いである。

(3) 分類

主題を表す分類記号は、早慶ともに日本十進分類法 (以下「NDC」とする) 9版を採用していたが、移行時に10版に変更した。慶應では請求記号に分類記号を使用しない一部の洋書では分類を付与してこなかったが、移行後は早慶すべての和洋図書に付与することにした。分類対象が増えることになったが、逆に分類の要・不要を判定する手間をなくしてフローを簡略化することを目指した。

(4) OCLCへの書誌・所蔵登録

新体制では、OCLCへの書誌・所蔵登録を必須とした。早稲田では以前から行われていたが、慶應では過去のデータが一部登録されているだけで、システムの仕組みと効率化の面から直近では行っていなかった。登録はOCLC参加館の義務でもあり、移行を機に慶應も網羅的に登録するフローを整えることになった。

OCLC参加の主目的は洋書の書誌を調達することであるが、最近ではNDL、TRCからOCLCへの登録もあり、和書の書誌調達も可能になっている。さらに、世界規模の書誌ユーティリティであるOCLC

への登録により国際的に慶應の所蔵の可視化が図れ、視認性のアップが大学自体の広い認知にも役立つという意義でのメリットも考えられる。

また、今回最も重視したのは、データの識別子として書誌にOCLC IDを持たせることだった。書誌の同定に使われる識別子としてはISBNやISSNが考えられるが、年代の古いものや非売品などこれらを持たない資料も数多い。このような資料群の識別子として国際的にはOCLC IDがキーとなることが多く、移行時のデータマッチングにも使われることがわかった。このため、将来的にも早慶で共有する書誌にはこのIDを持たせることが重要と考えている。

なお、OCLCのデータフォーマットでは、ラテン系言語以外の書誌の場合、レギュラーフィールドにローマ字形、880に漢字形などの原語形を入れるのが標準である。早慶データを登録する際、レギュラーと880のどちらを漢字形にするか、また、和書の目録言語を日本語にするか英語にするかというのが問題となった。インターフェースにレギュラーしか表示されない検索システムもあるため、日本語書誌のためにはレギュラーに漢字形を入れることを優先したいが、その形ではOCLC登録時のValidation Checkが通らない。このため、最終的にはAlmaではレギュラーを漢字形とし、OCLC登録時には漢字形が880になるようにフィールドを転換させることに決定した。

5 データ移行・統合

(1) 移行手順

早慶ともに200万件超の目録データから成るデータベースを保持しており、それをどう統合するののかも大きな課題であった。同じMARC21を採用していても、階層がある書誌の本タイトルの採用基準、書誌の作成単位、図書と雑誌の扱いなどの差異もあり、内容が全く同じである訳ではない。どちらの書誌を活かすかについて、最終的には慶應側に統合することに決定した。MARC21のフィールド記述の独自ルールが少なかったことに加え、ユニットの運用を慶應が主導することなどが理由である。

Almaには早慶で共有するNetwork Zone (以下「NZ」とする) と機関別のデータを置くInstitution Zone (以下「IZ」とする) がある。基本的に書誌データはNZ、所蔵データはIZに置く想定である。移行

は、先に慶應の書誌データをNZにアップロードし、早稲田の書誌をそこにマッチングさせて、

- ・重複していない早稲田の書誌と所蔵はそのままロードする
- ・重複していた場合は早稲田の所蔵を慶應書誌に付け替え、元の早稲田書誌は移行しない

という方法で行うことになった。

移行前にはシステム担当の協力を得て、既存データの改修作業も段階的に実施した。例えば、慶應では主な案件として固定長フィールドの桁ずれや矛盾のあるコードが入っているものの整理・修正、ローカルフィールドの使用タグの書き換え、漢字分かち・カナ読みフィールドの削除、ローマ字の再生成などを行った。また、所蔵データに対しても配架タイトルの入力フィールドの変更などを行った。移行後の新ルールに合わせつつ、MARC21対応のシステムでエラーが発生しない形式にすること、利用者がディスカバリで検索結果を見たときに、重要な情報が見やすくなることなども考慮した。

書誌データの統合は移行過程で行われるが、その際のマッチキーはOCLC IDである。すなわち、早慶で同じOCLC IDを持つ書誌が機械的に統合されるのだが、慶應の書誌にはOCLC IDは網羅的には入っていない。よって、事前に早慶データをマッチングさせて、同定された場合は慶應の書誌に早稲田側からOCLC IDをコピーして書き込んでおくという処理をすることになった。この事前マッチングは2回行い、既存OCLC IDやLCCN、ISBNと出版年などをキーとして重複チェックした。また、ここでマッチした場合は、早稲田で残したいフィールドを

慶應書誌に事前に書き込んでおくという処理も行った。具体的にはNDC、NDLSH、内容注記、必要なローカルフィールドなどである。

なお、除籍資料、貴重書、Google Library Project書誌、リポジトリ書誌、学位論文などはマッチングから除外した。

事前マッチングは、第1回は2018年9月、第2回はその後作成された書誌を拾うために2019年6月に行った。両方合わせてのマッチ件数は、雑誌が早慶ともに約7万件の書誌のうち21,476件、雑誌以外が早慶ともに210万件強のうち、515,263件となった。割合として早稲田の書誌25%程度が慶應のものにマッチし、所蔵が慶應側に統合されたことになる。

(2) 移行後の書誌データ例

上述のような移行作業を経て、2019年9月にAlmaでの業務が開始されたが、実際にユニットで作成されているデータは、次のような形式となっている。(図1)

AlmaのMetadata Editorで見ると、レギュラーに対応する880のフィールドは、対であることがわかりやすいよう、“880”なしでレギュラーの直下に並んで表示される。アクセスポイントに虫眼鏡マークが付いているのは、典拠データとリンクしていることを表している。040 \$eの“rda”のコードは準拠する目録規則がRDAであることを表し、出版事項に264フィールドを使用していること、336、337、338というコンテンツ、メディア、キャリア種別のフィールドがあることなどからも、RDA対応された書誌であることがわかる。

LDR	01389nam#a2200301##4500	245 1 0	\$S6 880-02 \$Sa 図書館のための和漢古書目録法入門 / \$Sc 伊藤洪二著
001	9912018270404033	1 0	\$S6 245-02 \$Sa Toshokan no tame no Wakan kosho mokurokuhō nyūmon / \$Sc Itō Kōji cho
005	20200210163210.0	264 1	\$S6 880-03 \$Sa 東京 : \$B 樹村房, \$Sc 2019.11
008	191118s2019###ja#####jpn d	1	\$S6 264-03 \$Sa Tōkyō : \$B Jusonbō, \$Sc 2019.11
020	\$Sa 9784883673292	300	\$Sa viii, 262 p ; \$Sc 19 cm
035	\$Sa (JaToTRC)19050735	336	\$Sa text \$S2 rdacontent
035	\$Sa (OCoLC)1130748314	337	\$Sa unmediated \$S2 rdamedia
040	\$Sa JTKU \$Sb jpn \$Se rda \$Sc JTKU	338	\$Sa volume \$S2 rdacarrier
084	\$Sa 014.3 \$S2 njb/10	504	\$Sa 文献 : p 255-257
099	\$Sa AM1	650 7	\$S6 880-04 \$Sa 資料目録法 \$S2 jlabsh
100 1	\$S6 880-01 \$Sa 伊藤, 洪二, \$Sd 1971- \$S9 A \$S0 (sokeiaut)9812030680504033	7	\$S6 650-04 \$Sa Shiryō mokurokuhō \$S2 jlabsh
1	\$S6 100-01 \$Sa Itō, Kōji, \$Sd 1971- \$S9 R \$S0 (sokeiaut)9812030680504033		

図1

6 移行後の問題点・課題

(1) データ、システム関連

早慶で機械的に事前マッチングできなかった75%の書誌のうち、その後の目録作業中に重複と確認できたものは、都度統合処理を行っている。単純な一対一の書誌であれば良いが、継続刊行されているものの書誌の作成単位や図書・雑誌の書誌種別、発注種別が異なっていると手間のかかる調整が必要となる。このような判断はユニットでは難しく、スピーディな処理の妨げとなっている。

典拠形アクセスポイントの齟齬も問題である。事前にわかっていたことではあるが、慶應は日本人についてNDLとNII準拠が混在し、早稲田はNDL準拠であったため、同一著者でも典拠形が割れているものが存在する。現時点では過去データを網羅的に修正していく予定はなく、すぐの解消は期待できない状況である。

システムの機能に不具合がある場合は、Ex Libris社へCaseをあげて対応を依頼するが、すぐには修正されずにバージョンアップを待ったり、日本でしか確認されない些細なトラブルでは保留のまま検討の俎上に載らないこともある。これまでのところ致命的な案件は残っておらず、操作の工夫で回避できるものが多いが、改善されれば使い勝手が向上するため、引き続き解決に向けて要望を伝えている。

(2) 滞貨

目録作成業務においては、移行時から蓄積した滞貨が生じており、この解消が喫緊の課題となっている。滞貨の原因としては、移行前に1か月半目録業務を停止したこと、委託スタッフのスキルアップが必要なこと、書誌のRDA対応、洋書のNDC分類付与に時間がかかること等があげられる。まだ移行直後で、早慶の重複統合もれを頻繁に手作業で処理する負荷も高い。

利用者に資料を迅速に提供できていないことが心苦しいが、1日で処理できる量に限りがあるなか、長期間に渡って作業を停止すると、その分の遅れを取り戻すのは一朝一夕には難しい。また、2020年度はCOVID-19感染防止による思いがけない業務停止も強いられ、その影響も甚大である。

新規スタッフのスキルアップと処理のスピードアップ、データ品質の維持、これらを満たしながら

滞貨を減らすべく、職員とユニットの委託業者で改善策を模索しながら作業に尽力している。

7 まとめ

共同目録体制の枠組作り、移行仕様の検討にあたっては、チームメンバーが日常業務もこなしながら準備を進めてきた。また、RDAに対応した新しいマニュアル作り、総入れ替えとなった委託スタッフのトレーニングには慶應の目録担当の職員が全員で取り組み、今もユニットの運用を支えている。この体制における早慶の立場は平等を前提としているが、実際にはユニットの状況を現場で監督し、実作業にも従事している慶應の負担はかなり大きい。これは当初から想定していたことではあるが、目録業務は職員が行うべきものかどうかという思想の違いに依っている。データ作成の大部分を委託業者に任せたととしても、そのためのルールを策定し、品質を判定できる知識を持ち、システムの機能を検証して熟知すべきはやはり職員である。そして、これらは実際の業務経験を通して養われるのではないだろうか。今回のデータ統合やデータ改修の手順・方法、外付けシステムの仕様などを考えられたのも、実作業を通して経験と知識を深めた職員が何人もいてくれたからこそである。共同目録体制が稼働し、滞貨などの課題はあっても大過なく動いているのは彼らの努力に依るところが大きく、ここで深く感謝したい。

また、もちろん早稲田側での作業、調整に尽力してくださった早稲田の目録チームとの協力関係も今回得られた貴重な財産である。今後も意見交換を重ねつつ、緊密に連携しながら体制の安定化を図っていきたい。

注

- 1) 早慶図書館システム共同運用の実現に向け、業務内容ごとにチームに分かれ、各分野で検討を行った。目録チームはその1つ。
- 2) 早慶の図書館間で資料の搬送を行う連絡便。