

# 見つかる・つながるデータを作るために

このえつこ  
河野江津子

(メディアセンター本部係主任)

さけみ かよ  
酒見 佳世

(日吉メディアセンター)

## 1 はじめに

Google や YAHOO! などの検索エンジンに思いついたキーワードを何語か打ち込めば、関連度順に情報が表示されるのが当たり前となっている昨今、図書館が利用者に提供できるのは、一体どのような検索システムでありデータであろうか。

図書館が所蔵する、読み手の知識となって蓄えられるような思索の成果や研究結果などの学術的な資料の形態には、紙媒体、それを電子化したもの、初めから電子のものがある。アクセスの形態も、実際に手に取れるものから、電子的にアクセスするもの、アクセスに制限があるものなど様々である。現在、KOSMOS (慶應の OPAC 名称) 上では、紙の資料も電子の資料も区別なく、同時に検索できる。検索結果には紙と電子のデータが混在して並ぶが、実際の両者のデータ作成は、全く別の流れを通じて行われている。本稿では、新システム KOSMOS III での主にデータ作成・管理の側面から、紙と電子それぞれの現状と展望を述べることにしたい。

## 2 目録データ

慶應義塾大学の目録データは、KOSMOS II という図書館システムが稼働した 1998 年から、目録規則として英米目録規則第 2 版 (和資料では一部日本目録規則) を採用し、データ規格としては MARC21 フォーマットに準拠している。

KOSMOS III へと業務システムが変わっても MARC フォーマットは同じであるためデータ要素自体が大きく変わるわけではない。それをどうユーザーに見せるのか、検索させるのかという点が目録データにとっては重要な側面である。

例えば、書誌作成単位を見直し、和書は著作責任単位でなく各冊の書誌データを提供するようになったが、このことにより KOSMOS 上では表紙画像や目次情報のような外部提供データとのリンクが可能となり、ユーザーは自分の必要な「1 冊」の特定がし

やすくなった。また、和書に付与してもワード検索ではあまり活かされない LCSH も、検索結果を絞り込むためのファセットナビゲーション機能においてリンク項目として表示され、主題検索の一要素として一役買うことができるようになった。

また、典拠ファイルの参照運用を再開し、個人名や団体名などの表記のゆれによる検索もれが起こらないように調整するための環境が整った。

目録データの外部ソースとしては、主に和書は TRC、洋書は OCLC の書誌データを使用している。MARC21 へのコンバージョンプログラムを工夫し、データの修正率を下げることで、作業の効率化を図っている。外部 MARC を活用することで標準的な目録データの骨組みは簡単に作成できるが、大学図書館の目録として慶應なりの味付けも必要である。コピーカタロギングで全てを済ませるのでなく、ケースによっては本タイトルの取り方を見直す、図書・雑誌のどちらが相応しいか検討するなど、ローカルなレベルでの判断も欠かせない。また、的確な分類や件名、典拠に基づいた標目を付与することで、より品質の高いデータを提供できるよう日々作業を行っている。

一方、外部 MARC が流用できない古典籍や手稿資料など慶應が所蔵する貴重な資料のデータ作成にも力を注ぎ、慶應の目録の独自性を高めていきたい。加えて、過去を振り返れば、長い伝統で培われた紙のコレクションの目録データは完成度も詳細度も均質ではない。未遡及の資料群も存在する。OPAC に慣れた利用者にとって「データが見つからない≡存在しない」ことになってしまう昨今、それらのデータ整備も今後の重要な課題である。

## 3 電子資源のメタデータと管理の状況

電子資源は、基本的に契約に基づく外部アクセスによって提供されるものであるため、従来の図書館のシステムや管理の体制には馴染まない部分があ

る。一度購入すれば、半永久的に保存の対象となる紙の資料のためのデータと、予算の関係でキャンセル/再開したり、サービス自体が停止されたりするような電子のデータとは、おのずと扱いは異なってくる。電子の場合、データそのものは版元側にあるのが通常であり、そのメタデータの作成者は、図書館ではなく版元である方が効率が良い。紙の目録も、版元や書店・取次等のデータを生かすという方向性は同じだが、電子では版元が作成するデータの重要度が更に高いと言える。

電子ジャーナルの特徴として、利用範囲が購入した館に限定されないこと、購読タイトルに加えて非購読タイトルが存在すること、利用可能なタイトルが流動的で版元間の移動が頻繁に行われること、同じタイトルが複数のプラットフォームで提供されることなどが挙げられる。電子ジャーナルを管理・検索するためのメタデータも、そうした特徴を反映しやすいものであることが望ましい。

KOSMOSに登録する電子ジャーナルのデータには、SFX(リンクリゾルバ)のMARCitサービスを利用しているが、このサービスで使用されている大元のデータは、CONSER(Cooperative Online Serials)のMARCデータである。基本的に質の高いデータソースであるが、日本語タイトルに関しては、タイトルがローマ字となっていたり、漢字が文字化けしたり、そもそもデータが存在しないといった問題がある。

SFXによる全文表示機能を支えているのは、版元ごとのタイトル情報、リンク情報を入れたナレッジベースである。このデータの質が、システム自体の使い勝手を決めるといっても過言ではない。特に日本語資料のデータの質については、当初から問題となっている部分であるが、質の改善には、地道なデータ修正作業を継続する一方で、提供元である版元等のデータ自体の精度を上げていくような働きかけが必要である。

一方、電子ブックのメタデータについては、現在、版元から無料で提供されるMARCデータを利用している。今後、更に規模が拡大することが予想される中で、版元毎に異なるMARCデータをそのまま

使うのではなく、一旦、統一的なデータに置き換えるなど、新たな作業や仕組みが必要かもしれない。

データベースについては、MARCデータの入手経路がなく、現在KOSMOSには搭載されていないが、今後検討すべき課題の一つである。

電子資源の管理の状況については、主に契約情報の管理にVerde(電子資源管理システム)、支払作業にKOSMOSⅢの受入システムを使用している。どちらも、ようやく端緒についたばかりであるが、外部へのアクセスが保証されてはじめて成立する電子資源利用の後ろ盾となるのは契約であり、これらをきちんと管理し、共有することは、電子資源の規模が拡大している現況においては更に重要になってくる。契約、支払、アクセスを一つのシステムで結び付けることにより、データの正当性を保証し、確かなアクセスを提供できる環境を維持していきたいと考えている。

#### 4 終わりに

上述の通り、紙と電子それぞれの媒体でデータ作成方法は異なっている。データ作成と言っても、全てを1つの大学図書館で作成することはほぼ不可能であり、紙、電子共に様々なルートから調達しているのが現実である。

電子化された資料であれば、タイトルごとのメタデータに加えて、全文が検索対象となる場合もあるだろうし、タイトル単位だけでなく、論文単位、章単位のメタデータも必要になるかもしれない。KOSMOS=1冊単位の資料を探すものという概念を捨て、研究者データベースや二次情報のデータベースなど、多様なデータソースとの連携を深めることも不可能ではない。様々な素材をいかにうまく一つのプラットフォームに乗せ、なるべく違和感なく、分かりやすく料理して並べるか、それこそが我々図書館員の腕の見せ所となるはずである。

このため、今後も目録規則やMARCに対する国際的な動向、出版界の事情にも目を配り、外部ソースと連携しやすいデータ作りを意識し、より付加価値の高い目録を提供していけるよう努めたい。