

CA1860 オープンなナレッジベースの進展と その背景

うえのともき*
上野友稔*
かがわともこ†
香川朋子†
ふるはしはなえ‡
古橋英枝‡
しおのまゆみ§
塩野真弓§

0. はじめに

近年の学術情報流通においては、論文などの学術情報をインターネットを通じて誰でも閲覧できるようにする Open Access、研究データなどを特別な制限無しに利用・再掲載できる形で公開する Open Data、論文や研究データの公開を通して新たなイノベーションを創出することを目指す新たなサイエンスの進め方を指す Open Science といった動きがある。

加えて、研究成果を利用者にナビゲートするために図書館が作成するデータ、例えば電子ジャーナルのタイトルリストなどをオープンにし、図書館コミュニティで管理する動きが世界的に広まっている。

本稿では、オープンな学術情報流通の国際的動向とその背景について、タイトルレベルの電子リソース流通の鍵となるナレッジベース (KnowledgeBase: KB) (CA1784 参照) に焦点を当て、論じる。

1. KB とは

米国情報標準化機構 (NISO)⁽¹⁾ の定義によれば、KB とは「電子リソースに関するタイトルリストや収録範囲、リンク情報等、リンクリゾルバのベンダーが構築する豊かなデータベース」のこととされている。これに各機関において、特定のコンテンツに関する電子媒体でのアクセス可否や印刷媒体での所蔵の有無を反映したものは、特にローカル KB として区別されている。

KB はリンクリゾルバを支える一つのデータベースとして誕生した。1990 年代後半にソンペル (Herbert Van de Sompel) によって構築された「Resolver モデル」は、従来、電子ジャーナルや電子書籍に代表される電子リソースについて、リンク元とリンク先が 1 対 1 でしか提示できなかったのに対して、リンク元とリンク先の間に中間窓を設定することによって、1 つのリンク元から複数のリンク先へとアクセスすることを可能にした⁽²⁾。この複数のリンク先情報を管理するデータベースとして「SFX base」が構築され⁽³⁾、の

ちの商用 KB の原型となった。その後、このモデルを適用した商用のリンクリゾルバが次々に開発され、現在ではディスカバリサービスや、冊子体資料と電子リソースを共通に管理する Library Services Platform (CA1861 参照) の中核としても活用されている。

活用の場が広がる一方で、商用 KB には次のような課題があることも明らかになった⁽⁴⁾。

第一に、商用 KB ベンダーがベンダーにとって都合のよいアクセス先のみ表示させたり、優先順位を高く設定させたりするリスクが存在しており、図書館側ではコントロールできない課題が生じていた。第二に商用 KB ベンダーが用意した KB に対して、各図書館が個別にメンテナンスを実行することで、同一データに関して複数機関で作業を行う、作業効率の低下が課題となっていた。

これらの課題を解決するために始まったのが Global Open Knowledgebase (以下、「GOKb」とする。)⁽⁵⁾ と Knowledge Base Plus (以下、「KB+」とする。)⁽⁶⁾ に代表されるオープンな KB の構築プロジェクトである。KB をオープンにする理由は、コミュニティで KB を管理することにより、図書館が KB のデータをコントロールするイニシアチブを取り戻すとともに、作業効率の低下を解消するためである。

2. GOKb と KB+

2.1 GOKb

GOKb は、Kuali OLE プロジェクト⁽⁷⁾ (E1003 参照) と英国の Jisc⁽⁸⁾ がアンドリュー・メロン財団⁽⁹⁾ から助成を受け、ノースカロライナ州立大学をリーダー機関として、2012 年から構築されている電子リソースの、グローバルでオープンな KB である。GOKb は、サービス指向型アーキテクチャに基づく図書館システムを設計するプロジェクト Kuali OLE で利用されることを目的としている。GOKb 構築の目的は、Kuali OLE を設計するにあたって、電子リソース管理を支援するオープンかつ図書館コミュニティによって管理される KB がないという課題を解決することにあった。

GOKb のデータは、電子リソースを保有する参加図書館から入手している。2015 年 9 月現在、約 2 万 4,000 のタイトル情報、約 300 のタイトルの全部または一部をまとめて販売する形態であるパッケージの情報、18 の利用者の範囲などの利用条件や契約条件の情報をもつテンプレートライセンスなどを含み、データは CC0 で公開され、誰でもどのような目的でも利用できる。

GOKb の基本精神がオープンかつ図書館コミュニティによる共同管理であることに伴い、KB の情報は電子リソースのメタデータ交換における標準規格であ

*電気通信大学学術情報課

†お茶の水女子大学附属図書館

‡国立情報学研究所

§京都大学附属図書館

る KBART⁽¹⁰⁾形式で公開されている。また、データの再利用を促進するため、GOKb Linked Data Ontology を策定して公開する予定である。同 Ontology は GOKb が定義した独自の語彙に加え、BIBFRAME⁽¹¹⁾、DataCite⁽¹²⁾、Dublin Core⁽¹³⁾等の汎用的な語彙を取り入れた複合的なメタデータスキーマ⁽¹⁴⁾で構成される。また、電子リソースのライセンス契約で定められた利用条件（ライセンス情報）は、電子商取引における標準化推進団体 EDItEUR⁽¹⁵⁾が提供する標準規格 ONIX for Publications Licenses (ONIX-PL)⁽¹⁶⁾形式で公開されている。これは、2013～2014年の ONIX-PL Encoding Project⁽¹⁷⁾において作成されたもので、テンプレートの共有により、各図書館が電子情報資源管理システム（ERMS）に重複して登録していたライセンス情報の作成コストを削減し、かつ正確なライセンス情報をユーザに公開されやすくすることで契約資料の最大限の有効活用につながる効果をもたらしている。

2.2 KB+

KB+は、Jisc と、HEFCE（英国高等教育助成会議）⁽¹⁸⁾から委任された Jisc Collections⁽¹⁹⁾により、2012年から英国で共同構築されている KB である。KB+は、英国内で KB を管理しつつ、英国の各大学・研究機関が持つローカルなデータを含めたデータベースである。KB+ 構築の目的は、英国の大学および高等教育機関向けに正確な KB を提供し、電子リソース情報の管理のために図書館員が費やしている労力と時間を最小化することにあった。

KB+には、NESLi2⁽²⁰⁾、Jisc eCollections⁽²¹⁾、SHEDL⁽²²⁾等のコンソーシアム向けの協定に基づいた KB に加えて、GOKb と異なり、英国の各大学・研究機関の実際の契約に沿った情報、例えばローカルなライセンス情報、JUSP⁽²³⁾の利用統計などのデータを併せ持っている。また、Jisc は GOKb パートナーとなっているので、GOKb のデータが KB+に収録されていることに加え、ローカルな情報を含まない KB+を KBART 形式により CC0 で公開している。Jisc は、2007年にライセンス情報の一部を ONIX-PL 形式にエンコーディングした⁽²⁴⁾ことを皮切りに商用の ERMS への ONIX-PL 形式でのデータ取り込みを開始する⁽²⁵⁾等、英国コミュニティの枠を超え、ライセンスデータ共有のための基盤整備を国際的に牽引している⁽²⁶⁾。

3. 広がりを見せるオープンな KB の連携

この他にもこれまで世界的にさまざまなプロジェクトで、国またはコンソーシアムレベルのオープンな KB が構築されてきた。例としてカナダのサイモン・フレーザー大学図書館の逐次刊行物管理システム

CUFTS の Open Knowledgebase⁽²⁷⁾、ドイツの ZDB (Zeitschriftendatenbank)⁽²⁸⁾、フランス高等教育書誌センター (ABES) の BACON⁽²⁹⁾があげられる。これらは GOKb にデータを搭載する予定である。この他、スウェーデン国立図書館、中国の大学図書館コンソーシアム CALIS、後述する日本の ERDB-JP も、GOKb にデータを提供することに関心を寄せている⁽³⁰⁾。この他、世界規模の広義の KB として、ISSN 国際センターがユネスコの支援を受けて提供するサービス ROAD (Directory of Open Access scholarly Resources) があげられる⁽³¹⁾。

GOKb のような包括的な KB を一つの図書館コミュニティで維持管理していくことは容易ではない。さまざまなコミュニティが作成する、収録対象やデータ内容が異なる KB と連携することで、国際的に作業の重複を避け、より正確で多様なデータを共有することを目指している。

4. ERDB-JP プロジェクト

日本の大学図書館でも 2000 年代半ばから多くの機関で海外製商用 KB の導入が進んできた。その中で、特に海外製商用 KB に含まれる国内刊行の電子リソースのデータが少ないことが問題になっていた。このため、「大学図書館と国立情報学研究所との連携・協力推進会議」⁽³²⁾を母体とする「これからの学術情報システム構築検討委員会」⁽³³⁾の電子リソースデータ共有ワーキンググループが、国内の電子リソースに関する情報を集約・管理するデータベース“ERDB-JP” (Electronic Resources Database-JAPAN)⁽³⁴⁾を構築し、2015年4月に公開した (E1678 参照)。

ERDB-JP のデータ形式は KBART および KBART2 にタイトルのヨミ、NACSIS-CAT 書誌レコード ID などの独自拡張フィールドを追加したものである。収録されたデータは CC0 で公開しており、GOKb との連携等、国内外へのデータ流通に向けた活動を予定している。

ERDB-JP は、収録された国内刊行の OA ジャーナルの情報を GOKb に提供することで、国際的な学術情報基盤整備と流通の潮流に、そのデータを載せる可能性を持つ。GOKb や KB+との国際連携が進めば、ERDB-JP データを含む KB が、各種ディスカバリサービスやリンクリゾルバなどの様々なシステムを通して世界中の利用者から利用され、その結果として日本語の研究成果へのアクセス増加が期待できる。

5. オープンな学術情報流通の国際的な展望

GOKb、KB+、ERDB-JP は、それぞれデータを CC0 としており、それぞれのデータを、誰が、どの

ような目的でも、いかなる負担もなしに幅広く利用できることが期待されている。例えば、複数ある商用KBがこれらのデータを再利用し、より正確でリッチなKBが各種商用ツールをとおして流通することで、より多くの機関で利用されることも考えられる。実際にGOKbは、他国で構築が進むオープンなKBの収録を図りつつ、商用KBのベンダーと協議の場をもち、今後も多くの商用KBベンダーにGOKbデータを提供しようとしている。

世界的なKBができてきた結果、図書館は商用ベンダーと協力しながら、KBの管理に主体的に関与し、KBのデータを正確で最新に保つ役割を担う。図書館がこの世界的なKBをディスカバリサービスなどで活用することで、利用者は電子リソースへの正確なアクセスを享受できる。立場や国を超えて図書館が必要とするデータを構築・管理する世界的なKBの取り組みは、今後、図書館が電子書籍など、より多種類の電子リソースのデータをさらに広く流通させるために、重要なモデルケースになりうる。

- (1) "KBART: Knowledge Bases and Related Tools". National Information Standards Organization. <http://www.niso.org/publications/rp/RP-2010-09.pdf>, (accessed 2015-09-25).
- (2) 増田豊. 学術リンク S・F・X と OpenURL—S・F・X と OpenURL—. 情報管理. 2002, vol.45, no.9, p.613-620. <http://doi.org/10.1241/johokanri.45.613>, (accessed 2015-09-25).
- (3) Herbert Van de Sompel, Patrick Hochstenbach. Reference Linking in a Hybrid Library Environment Part2: SFX, a Generic Linking Solution. D-Lib Magazine. 1999, vol.5, no.4. http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt2.html, (accessed 2015-09-25).
- (4) James Culling. "Link Resolvers and the Serials Supply Chain". http://www.uksg.org/sites/uksg.org/files/uksg_link_resolvers_final_report.pdf, (accessed 2015-09-25).
- (5) Global Open Knowledgebase. <http://gokb.org/>, (accessed 2015-09-30).
- (6) Knowledge Base Plus. <https://www.kbplus.ac.uk/kbplus/>, (accessed 2015-09-30).
- (7) Quali Open Library Environment. <https://www.quali.org/ole>, (accessed 2015-09-16).
- (8) The Joint Informatin Systems Committee. <https://www.jisc.ac.uk/>, (accessed 2015-09-16).
- (9) The Andrew W. Mellon Foundation. <https://mellon.org/>, (accessed 2015-09-16).
- (10) "Knowledge Base And Related Tools (KBART)". National Information Standards Organization. <http://www.niso.org/workrooms/kbart>, (accessed 2015-09-16).
- (11) "BIBFRAME.ORG". Bibliographic Framework Initiative. <http://bibframe.org/>, (accessed 2015-09-16).
- (12) DataCite. <https://www.datacite.org/>, (accessed 2015-09-16).
- (13) "DCMI Metadata Terms". Dublin Core Metadata Initiative. <http://dublincore.org/documents/dcmi-terms/>, (accessed 2015-09-16).
- (14) その他, "Bibliographic Ontology", "FOAF", "MODS RDF Ontology", "RDF Schema", "The Service Ontology", "SKOS Simple Knowledge Organization System Namespace", "STAC (Security Toolbox: Attack & Countermeasure) ontology" が採用されている。
- (15) EDItEUR. <http://www.editeur.org/>, (accessed 2015-09-16).
- (16) "ONIX-PL". EDItEUR. <http://www.editeur.org/21/ONIX-PL/>, (accessed 2015-09-16).
- (17) "ONIX-PL Encoding Project". National Information Standards Organization. <http://www.niso.org/workrooms/onixpl-encoding/>, (accessed 2015-09-16).
- (18) Higher Education Founding Council for England. <http://www.hefce.ac.uk/>, (accessed 2015-09-16).
- (19) Jisc Collections. <https://www.jisc-collections.ac.uk/>, (accessed 2015-09-16).
- (20) NESLi2. <https://www.jisc-collections.ac.uk/journals/>, (accessed 2015-09-16).
- (21) Jisc eCollections. <http://www.jisc-collections.ac.uk/Catalogue/Overview/index/1819>, (accessed 2015-09-16).
- (22) SHEDL. <http://scurl.ac.uk/what-we-do/procurement/shedl/>, (accessed 2015-09-16).
- (23) 出版・アグリゲーターのジャーナルの利用実績データを、一元的に閲覧する英国のポータルサイト。Journal Usage Statistics Portal. <http://jusp.mimas.ac.uk/>, (accessed 2015-09-16).
- (24) Brian Green, Liam Earney. The importance of linking electronic resources and their licence terms: a project to implement ONIX for Licensing Terms for UK academic institutions. Serials. 2007, 20(3), p. 235-239. <http://serials.uksg.org/articles/abstract/10.1629/20235/>, (accessed 2015-09-16).
- (25) "JISC Collections licence information to be included in Serials Solutions 360 Resource Manager". Jisc Collections. <http://www.jisc-collections.ac.uk/News/JISC-Collections-licences-in-360-Resource-Manager/>, (accessed 2015-09-16).
- (26) ONIX-PL および編集ツール the ONIX-PL Editor (OPLE) は、Jisc と Publishers Licensing Society (PLS) の共同出資により構築された。Jisc は単独でも ONIX-PL Encoding Project に 1 万ドルを出資し、ウェビナーの実施や e-ラーニング資料の公開等、ONIX-PL の普及活動を行っている。
- (27) "Open Knowledgebase: reSearcher". Simon Fraser University Library. <http://www.lib.sfu.ca/about/initiatives/researcher/open-knowledgebase>, (accessed 2015-09-18).
- (28) Zeitschriftendatenbank. <http://www.zeitschriftendatenbank.de>, (accessed 2015-09-18).
- (29) BACON. <https://bacon.abes.fr>, (accessed 2015-09-18).
- (30) Kristen Antelman, Liam Earney, and Kristen Wilson. "GOKb: The Global Open Knowledgebase". <http://www.slideshare.net/gokb/gokb-the-global-open-knowledgebase>, (accessed 2015-09-18).
- (31) ROAD は ISSN を有するオープンアクセスの学術資料を対象に、書誌情報の他、主要データベースへの採録状況、ジャーナル評価指標などを CC BY-NC 4.0 で公開している。ISSN 日本センターも 2015 年 3 月から ROAD への登録を行っている。ROAD. <http://road.issn.org>, (accessed 2015-09-18).
- (32) 大学図書館と国立情報学研究所との連携、協力推進会議. <http://www.nii.ac.jp/content/cpc/>, (参照 2015-10-30).
- (33) これからの学術情報システム構築検討委員会. <http://www.nii.ac.jp/content/korekara/>, (参照 2015-10-30).
- (34) ERDB-JP は日本語が主な使用言語となっている、または編集・発行の責任主体が日本にある電子リソース (有料のものを含む) のタイトル情報を収録している。ERDB-JP. <https://erdb-jp.nii.ac.jp/>, (参照 2015-09-18).

[受理：2015-11-16]

Ueno Tomoki
 Kagawa Tomoko
 Furuhashi Hanae
 Shiono Mayumi
 Progress and background of open knowledgebase