Knowledge Base Profiles

The current knowledge base market includes a wide range of proprietary and open-source products. Proprietary products usually support a wide range of library services, including management and discovery tools. These profiles attempt to capture a snapshot of the current knowledge base landscape and describe the functionality associated with products. Each supplier included responded to a short profile questionnaire provided by the author.

Some knowledge base providers were unavailable to complete the questionnaire, resulting in a few major products being omitted from the following list. These include the SFX Knowledge Base from Ex Libris; the Innovative Central Knowledge Base from Innovative Interfaces, Inc.; and JournalFinder from W. T. Cox.

Commercial Knowledge Bases

EBSCO Integrated Knowledge Base

ORGANIZATION NAME

EBSCO Information Services (https://www.ebsco.com)

ORGANIZATION DESCRIPTION

EBSCO is a leading producer and provider of content and services serving the needs of researchers from libraries of all types and sizes. EBSCO has developed an end-to-end open discovery services platform around EBSCO Discovery Service that supports all content types, features advanced search logic, delivers discovery and holdings management tools, and ensures extensibility with an array of third-party applications. The platform streamlines staff functionality that directly impacts the user, with the EBSCO Integrated Knowledge Base playing a key role in supporting features such as holdings management, publication searching and browsing, OpenURL and direct linking to full-text, e-resource management, consolidation of COUNTER statistics, analysis, reporting, and in-workflow decision support.

NUMBER OF USERS

4,200 customers are using products that rely on the EBSCO Integrated Knowledge Base.

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

- OpenURL link resolver
- publication browse
- discovery service
- Google Scholar/PubMed holdings update services
- KBART-1 and KBART-2 holdings exports
- MARC records service
- usage consolidation
- e-resource management

DISTINCTIVE FEATURES

The EBSCO Integrated Knowledge Base is truly a global knowledge base representing over 10,000 databases and packages from over 1,400 providers. The integrated nature of the knowledge base with its identifier mappings allows EBSCO to automate holdings management for databases, e-journals, e-packages, and e-books ordered through EBSCO. Financial information, license terms, and access and registration information are also automatically updated for e-journals and e-packages ordered through EBSCO.

The same integration allows EBSCO to offer in-workflow decision support by providing access to COUNTER statistics, cost-per-use information, and analytics within the subscription workflow.

The link resolver and discovery service leverage an article-level knowledge base of over 120 million article links to offer a first-of-its-kind direct linking technology (introduced in 2001) that provides confirmed direct links to subscribed content, greatly improving the quality of linking and combatting a common problem of link resolvers where poor quality data on OpenURLs compromise link quality. This same technology allows EBSCOhost and EDS to provide access to more of the library's collection by integrating direct links to subscribed content into search results.

FUTURE DEVELOPMENT PLANS

EBSCO continues to focus development efforts on improving and expanding our knowledge base–related services. Future plans include

- improving librarians' user experience by offering a single interface for managing, reporting, and analyzing holdings, usage, licenses, and e-resource data
- expanding cost-per-use analysis and analytics to cover the entire collection
- supporting more COUNTER reports
- creating open integration with ILS systems to allow EBSCO and ILS partner systems to operate as one

Gold Rush

ORGANIZATION NAME

Colorado Alliance of Research Libraries (www.coalliance.org)

ORGANIZATION DESCRIPTION

The Colorado Alliance of Research Libraries is a nonprofit organization of fourteen research libraries in Colorado and Wyoming (thirteen academic and one public library) established in 1971 and incorporated as a nonprofit 501c3 in 1981. The driving force is cooperation and the sharing of purchasing power, materials, and ideas. Among the services offered by the consortium are the Prospector union catalog, the Gold Rush ERMS, consortial e-resource licensing, a shared print program, and continuing education and training.

NUMBER OF USERS

About fifty libraries in North America use one or more modules of the Gold Rush ERMS.

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

The Gold Rush service (https://www.coalliance.org/software/gold-rush) includes a link resolver, *A–Z* service for serials, ERMS for managing subscriptions, and Gold Rush Decision Support. The service is centrally managed, and libraries may subscribe to any needed module at a cost far below commercial counterparts. The Gold Rush Decision Support supports a knowledge base of over 1,700 title lists, which include publishers, aggregators, abstracting and indexing services, and specialty lists (e.g., Portico, CLOCKSS, CrossRef, shared print serial sets, open-access lists, etc.).

DISTINCTIVE FEATURES

The Gold Rush Decision Support service allows libraries to do content overlap between electronic resource packages from primary publishers, aggregators, and indexing/abstracting services. Users can compare one-to-one or many-to-many in the same simple interface. Results are displayed in graphical form, and analyses can easily be downloaded as needed. Libraries may also load title lists from other services such as a commercial ERMS, RapidILL, or other sources that may be used for comparative purposes.

FUTURE DEVELOPMENT PLANS

A new area of development that was released in fall 2015 is the Gold Rush Library Content Comparison System (https://www.coalliance.org/faq-library-content-comparison-system), which was developed to allow libraries to load their MARC records and compare them with other libraries in the system. It was initially developed to support the Shared Print program of the Colorado Alliance of Research Libraries but is now available to any library or consortium for a reasonable fee. There are many possible use cases for the system, some of which could include

- shared print programs among a group of libraries so that libraries can make better decisions about what to weed or put in storage
- adding a new program at an institution where the library wants to see how its collection compares to an institution that has a similar program in the same area
- a library loading a special collection of titles that are under consideration for weeding or storage to determine what is unique in that particular set
- performing quick exports of data sets for participation in other cooperative programs
- analyzing a collection for accreditation or membership in another organization

ProQuest Knowledgebase

ORGANIZATION NAME

ProQuest LLC (www.proquest.com)

ORGANIZATION DESCRIPTION

ProQuest is committed to empowering researchers and librarians around the world. The company's portfolio of assets—including content, technologies, and deep expertise—drives better research outcomes for users and greater efficiency for the libraries and organizations that serve them. ProQuest is headquartered in Ann Arbor, Michigan, with offices around the world.

NUMBER OF USERS

2,800+ libraries in more than 150 countries worldwide

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

- 360 Core (*A*-to-*Z* list)
- 360 Link (link resolver)
- 360 MARC Updates (OPAC updating service)
- 360 Resource Manager (electronic resource management)
- Intota (library services platform)
- Intota Assessment (print and electronic collection analysis and assessment)
- Summon (discovery service)

DISTINCTIVE FEATURES

At ProQuest, we have an integrated, centrally managed knowledgebase. From its origins in the year 2000 as "Serials Solutions KnowledgeWorks"—the first dedicated e-resource knowledgebase in the library industry—our knowledgebase has been a repository of high-quality, continuously updated metadata about e-journals, e-books, and other resources that is used across our services. For this reason, we can deliver consistent, synchronized metadata to any and all of the products that use the knowledgebase.

The fact that our knowledgebase is centrally curated and managed means that ProQuest libraries can utilize the same high-quality metadata across its librarian-facing tools (including ERM and assessment), as well as its discovery and access services. With our hosted software-as-a-service (SaaS) model, we make updates to the knowledgebase that are shared across all of our customers' services at once.

For the past fifteen years, ProQuest has developed and used increasingly comprehensive processes for cleaning, verifying, reconciling ("normalizing"), and updating the data we gather from content aggregators, hosts, publishers, and other providers. These processes create a corrected and consistent set of metadata that can be used across our products so that librarians and researchers don't have to worry about the quirks or inconsistencies that are inherent in a surprising percentage of the source data.

FUTURE DEVELOPMENT PLANS

Over the past three years, ProQuest has been hard at work behind the scenes, transforming our knowledgebase and expanding its scope, scale, and capabilities into a new, even more comprehensive knowledgebase. The new knowledgebase includes all of the e-resource metadata ProQuest curates, plus the serials and provider metadata we maintain in our Ulrich's Global Serials Directory, as well as our expansive store of MARC source records and data from new sources. The work we have accomplished enables us to bring together electronic, print, microform, and digital resource metadata in one place—on a new knowledgebase platform—and share it across a wider array of ProQuest services through APIs and web services. The new knowledgebase is also cloud-based, so we are able to innovate and scale the knowledgebase for future growth and expansion easily and effectively.

TDNet Discover

ORGANIZATION NAME

TDNet (www.tdnet.io)

ORGANIZATION DESCRIPTION

TDNet is a leading provider of information technology solutions for libraries and knowledge centers. TDNet is dedicated to helping knowledge workers work faster and more efficiently while enhancing user experience. TDNet's highly flexible solutions meet the needs of individual libraries, knowledge centers, and consortia doing much of the work and saving both time and expenses. TDNet's company flagship—TDNet Discover—leverages years of experience and understanding of customer needs, reduces administrative workload, simplifies discovery, and enables library personnel to focus on serving their patrons.

NUMBER OF USERS

Hundreds of customers worldwide

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

- TDNet Discover—discovery web-scale search
- TDNet Discover—Library e-Resources—e-resources discovery and access gateway
- TDNet Discover—OpenURL link resolver

- TDNet Discover—TOC alerts service
- TDNet Core ERM—electronic resource management system
- TDNet Holdings Manager—MARC records and other knowledge base–extracted information service

DISTINCTIVE FEATURES

TDNet Discover uniquely combines technology and content, together with services. At TDNet, we believe that the search process and its results are a significant stage in a much broader and complex organizational process. Based on this approach, TDNet Discover is not a stand-alone platform but part of a collection of organizational research workflow tools and processes. As such, discovery-to-delivery must be adapted to the organization's entire work environment.

These are TDNet Discover's features that enable users to discover and access information in enterprise content repositories, external repositories, licensed and open-access publishers' content, the web, and more:

- full library portal with efficient information deployment
- advanced, comprehensive content and search capabilities
- multisite, consortia, group support
- extensive statistics reporting tools
- built-in SUSHI statistics harvester
- responsive interface for mobile
- compatibility with authentication protocols
- full interoperability with enterprise workflows and infrastructures and full API support

FUTURE DEVELOPMENT PLANS

Aiming to best serve our core customer base—corporate, biomedical, government, and other special libraries and information centers—TDNet's development road map follows the holistic approach of developing all components of our offering. We are pursuing continued development of our comprehensive knowledge base and index, optimization of search and retrieval processes and open-access exposure.

WorldCat Knowledge Base

ORGANIZATION NAME

OCLC (www.oclc.org)

ORGANIZATION DESCRIPTION

OCLC is a global library cooperative that provides shared technology services, original research, and community programs for its membership and the library community at large. We are librarians, technologists, researchers, pioneers, leaders, and learners. With thousands of library members in more than 100 countries, we come together as OCLC to make information more accessible and more useful, because what is known must be shared.

NUMBER OF USERS

More than 4,700 total member libraries use the World-Cat Knowledge Base.

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

As OCLC has built new services and transformed our foundational services for the age of electronic resources, the WorldCat Knowledge Base has been placed alongside WorldCat at the center of everything OCLC does:

- WorldCat Discovery (web-based discovery service)
- A–Z List (public-facing inventory of e-resources)
- WorldShare ILL (resource sharing service)
- WorldShare Acquisitions (ordering and procurement)
- WorldShare License Manager (license management and usage statistics solution)
- WorldShare Analytics (collection analysis tool)
- WorldShare Collection Manager
- MARC record delivery service

DISTINCTIVE FEATURES

The WorldCat Knowledge Base aggregates e-resource data from over 5,900 different vendors and provides link resolution for 3.7 million open-access titles. As a content-neutral knowledge base provider, OCLC is proud to work across the broadest possible range of vendors and content partners.

OCLC was first to implement direct holdings feeds from content providers into the WorldCat Knowledge Base, updating a library's coverage quickly and accurately, and that program continues to expand today. Partners in this program as of November 2015 include EBL Ebook Library, ebrary, Ingram MyiLibrary, Elsevier ScienceDirect (journals and e-books), JSTOR, and Teton Data Systems.

WorldCat Knowledge Base has been designed and deployed to be leveraged at any level the library needs and chooses. It can be integrated with OCLC applications like WorldCat Discovery or WMS, easily synchronized with another knowledge base, integrated with third-party applications as a data platform, or used to enrich data for use in external systems.

The WorldCat Knowledge Base is the first cooperatively managed knowledge base. Each institution has the option to deny or approve updates to collection data from vendors before they are loaded to

the knowledge base. Institutions can also contribute brand-new collection data, which the rest of community can then make use of. With the help of members OCLC is building a collaborative and comprehensive global knowledge base.

FUTURE DEVELOPMENT PLANS

With a goal of getting as close as possible to real-time updates, OCLC is continually investing in architecture and in exploration of better, faster methods of getting updates from partners. OCLC is committed to gaining new partnerships with vendors and implementing direct holdings feeds to create a "hands-off" e-resource management system for libraries. OCLC is also experimenting with an option to receive vendor data on demand through APIs instead of depending on file loading.

OCLC's recent focus has been on improving the scalability of the system. OCLC is building a system to handle continual growth as the data ingested from providers and libraries continues to grow. Comprehensiveness is a goal libraries can achieve in cooperatively managing the WorldCat Knowledge Base.

The user experience is the ultimate goal of this work, and near–100 percent Google-style reliability of links is a critical component. Medium-term strategies include a move to direct linking to complement or in some cases supplant OpenURL linking. OCLC is currently testing a direct linking solution using Gale collection data and plans to expand this testing to other vendors.

Open Knowledge Bases

BAse de COnnaissance Nationale (BACON)

RESPONSIBLE ORGANIZATION

Agence Bibliographique de l'Enseignement Supérieur (ABES; http://en.abes.fr)

ORGANIZATION DESCRIPTION

ABES was created in 1994 to implement Sudoc (Système Universitaire de Documentation, or University Documentation System), the union catalog of France's higher education libraries. Sudoc opened in 2001 and has proved a resounding success. It covers the collections of 1,419 "deployed" or member libraries, along with the 1,793 public or private libraries from the Sudoc-PS network, which specializes in referencing serial publications. With over 10 million bibliographic records, 32 million localized documents and 24 million public queries in 2013, it plays a leading role in the French higher education and research information system.

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

BACON provides trusted KBART v2 formatted metadata for e-resources packages available for French higher education institutions. These metadata, put under a CC0 license, can be downloaded via BACON's website (https://bacon.abes.fr) and is accessible via web services. KBART files can be then used by knowledge base vendors and libraries.

DISTINCTIVE FEATURES

BACON focuses on French content. Data that can be fetched from other trusted community knowledge bases (KB+, GOKb) [is] integrated as is. For French content, we spend a lot of time encouraging French academic publishers to enhance their own metadata, and we insist that the KBART files be produced from the metadata used by the publishers' platforms. We have built a semi-automated workflow that analyses the data sent to us by the publishers and converts it to trusted metadata sources (SUDOC, ISSN registry, French National Library catalog). We are then able to produce a detailed report that helps the publishers spot the mistakes or the inconsistencies of their metadata. If the publishers correct their metadata, ABES grants them a "quality label." The major benefit for the publishers—and for everyone in the supply chain—is that the corrected and enhanced metadata can be used in any metadata feed, including ONIX files, MARC records, and data sent to discovery tools vendors.

FUTURE DEVELOPMENT PLANS

Future development plans include full automation of the file analysis workflow and full coverage of French academic publishers.

CUFTS Knowledgebase

RESPONSIBLE ORGANIZATION

Simon Fraser University Library (www.sfu.ca)

ORGANIZATION DESCRIPTION

Simon Fraser University (SFU) is a medium-sized publicly funded institution serving a student population of approximately 19,990 FTE. SFU offers comprehensive undergraduate and graduate programs with three campuses located in the Metro Vancouver region of British Columbia, Canada. The SFU Library employs approximately 113 FTE personnel.

NUMBER OF USERS

Approximately 66

- GODOT: OpenURL link resolver and interlibrary loan–requesting software
- CJDB: CUFTS Journal Database, a public, web-based *A–Z* electronic journal listing
- integration with CUFTS ERM for public display of license information via the CJDB
- simple MARC record service (title, ISSN, e-ISSN, and holdings by provider on a single record) for import into integrated library systems
- import of print MARC journal holdings for integration into the CJDB *A–Z* public display
- automated monthly export of Google Scholar XML holdings for Google Scholar Library links
- automated monthly export of holdings for use in the BrowZine service
- CUFTS Resource Comparison Tool—compares up to four CUFTS targets in the knowledgebase to find duplicate and unique coverage
- Journal Search—finds out which CUFTS targets in the knowledgebase contain full text for a specific title
- off-campus authentication services (such as EZproxy or Innovative's WAM) supported, and a proxy prefix can be added automatically by selecting proxy for each target

DISTINCTIVE FEATURES

Developed by an academic library for use in academic libraries in a consortia environment, the CUFTS knowledgebase is maintained by staff at the SFU Library. The knowledgebase contains the majority of the popular aggregator databases from EBSCO, Gale, and ProQuest as well as journal collections from large commercial academic publishers, university presses, and scholarly societies. In addition, the CUFTS open knowledgebase includes the Canadian Research Knowledge Network (CRKN) consortia journal packages. With Simon Fraser University Library's commitment to establishing leading-edge scholarly communications support, significant efforts are made to populate the knowledgebase with open-access journal targets and free back issue targets. Open Journal Systems (OJS) targets are also well represented in the knowledgebase

All targets in the CUFTS knowledgebase display a "title list scanned" date, which provides the date the target was last updated. Whenever partially activated targets are updated in the global knowledgebase, the contact listed in CUFTS will receive an e-mail message detailing the number of new titles added, modified, and deleted during the update as well as tab-delimited text files for each of the new, modified, and deleted titles that affect the library's holdings. Library contacts receive a deleted file only if any of their own activated titles were deleted by the global update.

The CJDB can also be integrated with the CUFTS Electronic Resources Management (ERM) module to display relevant license information for end users. License information appears in easy-to-read tabbed format and offers simple icons and plain language for end users and library staff. Some Canadian academic institutions have opted out of the Access Copyright agreement and rely on the Canadian Copyright Act and their own existing license agreements made directly with publishers and providers. So there has been an emphasis among Canadian academic institutions to make their license details publicly accessible.

FUTURE DEVELOPMENT PLANS

CUFTS is currently in a "steady state." There is a committed user community, but it is not growing dramatically. Similarly, ongoing incremental development is always underway, but at present there are no plans for any major development initiatives.

Electronic Resources Database-JAPAN: ERDB-JP

ORGANIZATION NAME

A Working Group for E-Resource Data Sharing (https://erdb-jp.nii.ac.jp/ja)

ORGANIZATION DESCRIPTION

A Working Group for E-Resource Data Sharing was established by the Future Scholarly Information Systems Committee to handle ERDB-JP. The Future Scholarly Information Systems Committee operates under the Cooperation Promotion Council set up by the Inter-University Research Institute Corporation, the Research Organization of Information and Systems, the National Institution of Informatics (NII), and the Japanese Coordinating Committee for University Libraries.

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

Link resolver and web-scale discovery service

DISTINCTIVE FEATURES

ERDB-JP is a one-of-a-kind knowledge base describing electronic journals and books written in Japanese and electronic journals and books edited or published in Japan. ERDB-JP covers more than 11,000 journal titles as of October 2015.

FUTURE DEVELOPMENT PLANS

- *Quality improvement of ERDB-JP data*: We are continuing to evaluate the optimal maintenance

organization needed to provide accurate and current ERDB-JP data.

- *Increasing ERDB-JP partners*: ERDB-JP partners maintain ERDB-JP data along with the working group. We are encouraging electronic resources publishers, commercial knowledge base vendors, and academic conferences to consider ERDB-JP partnership.
- *International collaboration*: We are going to transmit ERDB-JP data to GOKb for the distribution of Japanese research outcomes.
- *Electronic books and licensing*: We are evaluating the possibility of adding collections of electronic books and electronic resources licenses to ERDB-JP.

Global Open Knowledgebase (GOKb)

RESPONSIBLE ORGANIZATIONS

The Kuali Foundation (https://www.kuali.org) and Jisc (https://www.jisc.ac.uk)

ORGANIZATION DESCRIPTIONS

The Kuali Foundation is a nonprofit organization that develops open-source administration software for higher education. Kuali is also the parent organization to Kuali OLE, a community source library management system and sister project to GOKb. Jisc is a not-for-profit organization that supports digital services and solutions for the UK higher education sector. Jisc Collections supports the Knowledge Base Plus (KB+) project, also a project partner to GOKb. Kuali OLE and Jisc Collections have been working together since 2012 to develop GOKb as an open, community-managed knowledge base to support the broader community as well as their own individual projects.

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

GOKb aims to make knowledge base data freely available to the library community and provide the infrastructure necessary for partners to participate in the data management process. While GOKb does not support typical knowledge base–powered tools such as a discovery platform or ERMS, its open data and APIs are designed to allow external systems to consume the data in support of these functions.

Features include

- a web interface for browsing and searching data
- editor functionality that allows GOKb partners to deposit new data, correct errors, and contribute data enhancements like title history information
- OAI-PMH standards–based APIs designed for easy consumption and integration of data

DISTINCTIVE FEATURES

In addition to traditional knowledge base metadata, GOKb offers an enhanced data model that tracks changes over time, relationships between resources, and an extensible set of external identifiers. A co-referencing service within the knowledge base allows users to submit an identifier and receive a results set of all known identifiers associated with the same resource, through either the web interface or an API. All of the data in GOKb can be accessed through the web interface, API, or export tools.

The data found in GOKb is completely managed by the GOKb partners, which include the Kuali OLE partners, Jisc, and a number of additional library partners with an interest in the service. GOKb's data is openly available under a CC0 license. It can be used by anyone, for any purpose, without attribution. Academic institutions and commercial publishers and vendors are encouraged to collaborate in building and sharing GOKb's data.

FUTURE DEVELOPMENT

The GOKb development team is completing several development initiatives as part of its second round of grant funding from the Andrew W. Mellon Foundation. Features planned for release in 2016 include support for e-book packages, more advanced data-loading and management tools, and exposure of the knowledge base as linked data. GOKb will also continue to engage in community-building activities and is actively seeking new partnerships with libraries, consortia, publishers, and vendors.

Knowledge Base Plus (KB+)

RESPONSIBLE ORGANIZATION

Jisc (https://www.jisc.ac.uk)

ORGANIZATION DESCRIPTION

Jisc is the UK higher, further education, and skills sectors' not-for-profit organization for digital services and solutions.

SERVICES SUPPORTED BY THE KNOWLEDGE BASE

KB+ is a knowledge base that includes electronic resources management tools. All of the KB+ data is made available under an open license and disseminated throughout the library supply chain so that the right organizations have the data they need when they need it. Currently Ex Libris, ProQuest, OCLC, and EBSCO all use KB+ data in their systems. KB+ data is also used by other Jisc services or projects including JUSP and Safenet.

- A centrally maintained and managed knowledge base in which Jisc Collections collates, verifies, and updates knowledge base data to avoid costly and wasteful duplication of effort by libraries all trying to do the same thing by themselves.
- Verified, accurate, and up-to-date publication information for e-journal agreements, including national and regional consortium agreements from across the United Kingdom and a growing number of non-Jisc packages.
- Subscription information and management tools to help institutions track details of entitlements and journal coverage, manage renewals, compare different journal packages, view usage statistics from JUSP, and export files formatted for use with link resolvers.
- License information covering key values such as walk-in users, concurrent access, post cancellation access, and more. Institutions can create their own license information, making use of templates created by Jisc Collections or their own licenses.

FUTURE DEVELOPMENT PLANS

Incorporation of financial data will enable measurement of value (i.e., cost per use) and assessment of the strategic value of a title on a dimension other than raw usage.